Karlsruhe Institute of Technology

On the NLS Hierarchy with Nonzero Boundary Conditions

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der KIT-Fakultät für Mathematik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

Robert Wegner, M.Sc.

Tag der mündlichen Prüfung:	11. Februar 2026
Referent:	Prof. Dr. Xian Liao
Korreferenten:	Prof. Dr. Guido Schneider
	apl. Prof. Dr. Peer Kunstmann

Acknowledgments

I am grateful to Xian Liao for her supervision over the last years. Her advice and guidance have been invaluable to me and I will always fondly remember our long discussions. I am especially thankful for the freedom she gave me to explore any direction I wanted, and that she always had an open ear for whatever was on my mind.

I am also thankful to Guido Schneider and Peer Kunstmann for their role as co-referees.

I would like to thank Emile, Joannis, Julia, Kevin, Louis, Lukas, Marion, Niklas, Rebekka, Sebastian, and all my other colleagues and former colleagues from the work group, and also the CRC, who have made my time as a PhD a truly enjoyable experience. They were always available for discussions, or to help one another, or just to chat. Many hats were built! I especially thank Sebastian and Niklas (and others!) for letting me discover tabletop gaming, and Julia for always being a great office mate. I would like to also mention here my half-colleagues Hugo and Zihui whom I hope to chance upon in the future, perhaps at a conference.

Next, I would like to thank Herbert Koch, Leonardo Tolomeo and Nikolay Tzvetkov for the opportunities they gave me, the discussions we had, and for inspiring me.

I am thankful to my parents Barbara and Christoph for their support, on which I can always rely, and to my brother Fred, to whom I must explain all this one day.

Lastly, I would like to thank my love Stef. You were always there for me and I cannot imagine the last years without you. They were wonderful and I am looking forward to so many more!

Contents

CHAPTER 1

Introduction

The purpose of this dissertation is threefold:

1. To advance the understanding of the transformation of one-dimensional Schrödinger equations into hydrodynamic equations by establishing new mapping properties of the Madelung transform.
2. To initiate the quantitative study of the NLS hierarchy with nonzero boundary condition at infinity, and to introduce the $\widetilde{\mathsf{GP}}$ hierarchy as a *better* renormalization of the NLS hierarchy.
3. To establish a new connection between the NLS hierarchy and the KdV hierarchy in the form of a quantitative approximation result.

The central objects are therefore the NLS hierarchy with nonzero boundary condition, its renormalized versions, its hydrodynamic formulation, and the KdV hierarchy. Before we define these objects and explain our results, we would like to motivate the study of these hierarchies.

1.1. The path to the NLS and KdV hierarchies

Let us take a step back, start with the abstract, and then move to the specific, taking various detours along the way. We shall attempt to give a rudimentary answer to the following question:

How does one naturally encounter the NLS *and* KdV *hierarchies?*

1.1.1. Models for time-evolving physical systems

Suppose we wish to model a time-evolving physical system. First, we shall denote by $\mathcal{X}$ the set of all possible configurations of the system, or perhaps all those which we consider non-pathological. It is natural for $\mathcal{X}$ to be equipped with a Hausdorff topology.

Second, we need a set of *trajectories*, which represent possible time evolutions of the system. We equip $(0,\infty] = (0,\infty) \cup \{\infty\}$ with the order topology so that the intervals $(T,\infty] = (T,\infty) \cup \{\infty\}$, $T > 0$ are open neighborhoods of ∞. We define the set of trajectories to be

$$\mathcal{T}(\mathcal{X}) = \bigcup_{T\in(0,\infty]} \{T\} \times C((-T,T);\mathcal{X}) .$$

Here we have made some choices: we restrict ourselves to continuous trajectories, we work with both positive and negative times, and we have introduced a *time of existence* T of a trajectory, i.e. trajectories do not need to be defined globally in time. By replacing $C((-T,T);\mathcal{X})$ with $C((-T,T)^N;\mathcal{X})$ for some $N \in \mathbb{N}$ one can also model systems with multiple time variables. We consider only one-dimensional time here.

The set $\mathcal{T}(\mathcal{X})$ needs a topology as well. For $T^* > 0$ we equip $C((-T^*,T^*);\mathcal{X})$ with the compact-open topology, which is Hausdorff. Next, we define for $T^* > 0$ the subsets

$$\mathcal{T}_{\geq T^*}(\mathcal{X}) = \{(T,\varphi) \in \mathcal{T}(\mathcal{X}) : T \geq T^*\}$$

and equip each of them with the initial topology induced by the map

$$\begin{aligned} \mathcal{T}_{\geq T^*}(\mathcal{X}) &\longrightarrow (0,\infty] \times C((-T^*,T^*);\mathcal{X}) \\ (T,\varphi) &\longmapsto (T,\varphi|_{(-T^*,T^*)}) , \end{aligned}$$

i.e. the coarsest topology for which this map is continuous. Note that for $T_2 \geq T_1 > 0$ the inclusion $\mathcal{T}_{\geq T_2}(\mathcal{X}) \hookrightarrow \mathcal{T}_{\geq T_1}(\mathcal{X})$ is continuous. Now we equip $\mathcal{T}(\mathcal{X})$ with the final topology induced by the inclusions

$$\mathcal{T}_{\geq T^*}(\mathcal{X}) \hookrightarrow \mathcal{T}(\mathcal{X})$$

for $T^* > 0$, i.e. the finest topology for which these inclusions are continuous. By working directly with the definition, one can show that $\mathcal{T}(\mathcal{X})$ is Hausdorff.

Lastly, our model should decide which trajectories are admissible, in the sense that we deem them an appropriate description of the physical time evolution, and which are not. We describe this mathematically with a set $\mathcal{A} \subseteq \mathcal{T}(\mathcal{X})$ consisting of pairs (T, φ) where T is a *time of existence* and φ is an *admissible trajectory.* Most crucially, we require $\mathcal{A}$ to be translation invariant in the following sense:

$$\begin{gathered}\forall\,(T_1, \varphi_1) \in \mathcal{A} \quad \forall\, t_1 \in (-T_1, T_1) \quad \exists\,(T_2, \varphi_2) \in \mathcal{A} \quad \forall\, t_2 \in (-T_2, T_2) \\ |t_1 + t_2| < T_1 \Longrightarrow \varphi_1(t_1 + t_2) = \varphi_2(t_2)\,.\end{gathered} \tag{1.1.1}$$

In the language of ODEs, this means the systems we consider are autonomous.

We call the pair $(\mathcal{X}, \mathcal{A})$ a *model.* By an abuse of notation, we identify $(T, \varphi) \in \mathcal{A}$ with φ and write $T(\varphi)$ for the time of existence of φ.

1.1.2. Well-posedness of models and their realization as dynamical systems

The first step in the study of such a model is to prove its well-posedness.

Definition 1.1.1 (Well-posedness in the sense of Hadamard)**.** We say that the model $(\mathcal{X}, \mathcal{A})$ is locally well-posed if the following hold:

- **Existence.** There exists $\Phi : \mathcal{X} \to \mathcal{A}$ such that $\Phi(x)(0) = x$ for every $x \in \mathcal{X}$.
- **Uniqueness.** For all $x \in \mathcal{X}$ and $\varphi \in \mathcal{A}$ with $\varphi(0) = x$ we have $\varphi(t) = \Phi(x)(t)$ as long as $|t| < \min\{T(\varphi), T(\Phi(x))\}$.
- **Continuous dependence on data.** The map $\Phi : \mathcal{X} \to \mathcal{A}$ is continuous. Equivalently, for every $T^* > 0$ the maps

$$\begin{gathered}T \circ \Phi : \mathcal{X} \longrightarrow (0, \infty] \\ \Phi\big|_{(-T^*, T^*)} : \{x \in \mathcal{X} : T(\Phi(x)) \geq T^*\} \longrightarrow C((-T^*, T^*); \mathcal{X})\end{gathered}$$

  are continuous.

In this case we call Φ a *flow map* for $(\mathcal{X}, \mathcal{A})$. We say that the model $(\mathcal{X}, \mathcal{A})$ is globally well-posed if in addition:

- **Global existence.** There exists a choice of Φ with $T(\Phi(x)) = \infty$ for every $x \in \mathcal{X}$.

A proof of local well-posedness usually involves the explicit construction of the flow map Φ, for example via a fixed point iteration. Then global well-posedness is most classically proven in the following steps:

1. Extend the flow map Φ to a *maximal* flow map Φ^*, where maximal means that there exists no flow map which has for any $x \in \mathcal{X}$ a strictly larger time of existence.
2. Find a conserved energy, i.e. a function (or a family of functions) $E : \mathcal{X} \to [0, \infty)$ such that $E(\Phi^*(x)(t)) = E(\Phi^*(x)(0))$ as long as $|t| < T(\Phi^*(x))$.
3. Prove a blow-up alternative: either $\limsup_{|t| \to T(\Phi^*(x))} E(\Phi^*(x)(t)) = \infty$ for some $x \in \mathcal{X}$, or Φ^* is not maximal in the above sense.

If such a conserved energy is not available, then one faces a hard problem and more sophisticated techniques are necessary.

One can understand a proof of well-posedness as the model being *realized as a continuous dynamical system* in the following sense:

Lemma 1.1.2. *Assume that the model $(\mathcal{X}, \mathcal{A})$ is locally well-posed with flow map Φ and define, with an abuse of notation, new maps $\Phi(x, t) = \Phi(x)(t)$ and $T(x) = T(\Phi(x))$. Then $(\mathcal{X}, \Phi)$ forms a continuous dynamical system in the following sense:*

(i) *The flow map $\Phi : \{(x, t) \in \mathcal{X} \times \mathbb{R} : |t| < T(x)\} \longrightarrow \mathcal{X}$ is continuous.*

(ii) *We have $\Phi(x, 0) = x$ for all $x \in \mathcal{X}$.*

(iii) *For all* $x \in \mathcal{X}$ *and* $t_1, t_2 \in \mathbb{R}$ *with* $|t_1|, |t_1+t_2| < T(x)$ *and* $|t_2| < T(\Phi(x,t_1))$ *we have*

$$\Phi(\Phi(x,t_1),t_2) = \Phi(x,t_1+t_2)\,. \tag{1.1.2}$$

Proof. We give the derivation of the flow property (iii) from (1.1.1) because it is nontrivial. The first step is to prove the following short-time flow property: for all $x \in \mathcal{X}$ and $t_1 \in \mathbb{R}$ there exists some $T_2 > 0$ such that for all $t_2 \in \mathbb{R}$ with $|t_1|, |t_1+t_2| < T(x)$ and $|t_2| < \min\{T_2, T(\Phi(x,t_1))\}$ we have

$$\Phi(\Phi(x,t_1),t_2) = \Phi(x,t_1+t_2)\,. \tag{1.1.3}$$

To do so, we set $(T_1, \varphi_1) = (T(x), \Phi(x)) \in \mathcal{A}$ and let $t_1 \in \mathbb{R}$ with $|t_1| < T(x)$. By (1.1.1) there exists $(T_2, \varphi_2) \in \mathcal{A}$ so that $\varphi_1(t_1+t_2) = \varphi_2(t_2)$ as long as $|t_2| < T_2$ and $|t_1+t_2| < T(x)$. Note that $\Phi(x,t_1) = \varphi_2(0)$, so by uniqueness we have $\Phi(\Phi(x,t_1),t_2) = \varphi_2(t_2) = \Phi(x,t_1+t_2)$ as long as $|t_1+t_2| < T(x)$ and $|t_2| < \min\{T_2, T(\Phi(x,t_1))\}$.

Now we extend this flow property from $|t_2| < \min\{T_2, T(\Phi(x,t_1))\}$ to $|t_2| < T(\Phi(x,t_1))$. Define

$$\begin{aligned} I &= \{t_2 \in \mathbb{R} : |t_1+t_2| < T(x) \text{ and } |t_2| < T(\Phi(x,t_1))\} \\ J &= \{t_2 \in \mathbb{R} : |t_1+t_2| < T(x) \text{ and } |t_2| < T(\Phi(x,t_1)) \text{ and } \Phi(\Phi(x,t_1),t_2) = \Phi(x,t_1+t_2)\}\,. \end{aligned}$$

The interval I is a connected space which contains J. Since every trajectory $\Phi(x)$ is continuous in time and $\mathcal{X}$ is Hausdorff, the set J is closed in I. From the short-time flow property we know that J is nonempty. It therefore remains to show that J is open in I. We fix some $t_2 \in J$. Applying the short-time flow property to the left hand side of (1.1.2) yields a time $T_3 > 0$ so that

$$\Phi(\Phi(\Phi(x,t_1),t_2),t_3) = \Phi(\Phi(x,t_1),t_2+t_3)$$

for all $t_3 \in \mathbb{R}$ with $|t_2+t_3| < T(\Phi(x,t_1))$ and $|t_3| < \min\{T_3, T(\Phi(\Phi(x,t_1),t_2))\}$. On the other hand, applying it to the right hand side of (1.1.2) yields a time $\widetilde{T}_3 > 0$ so that

$$\Phi(\Phi(\Phi(x,t_1),t_2),t_3) = \Phi(\Phi(x,t_1+t_2),t_3) = \Phi(x,t_1+t_2+t_3)$$

for all $t_3 \in \mathbb{R}$ with $|t_1+t_2+t_3| < T(x)$ and $|t_3| < \min\{\widetilde{T}_3, T(\Phi(x,t_1+t_2))\}$. Therefore with $T_4 = \min\{T_3, \widetilde{T}_3, T(\Phi(x,t_1+t_2))\}$ we have $(t_2 - T_4, t_2 + T_4) \cap I \subseteq J$, concluding the proof that J is open in I. □

1.1.3. KdV and the pursuit of lower regularity

Many continuous time-evolving models in physics with continuous space are partial differential equations (PDEs), in the sense that $\mathcal{X}$ is a function space and the admissibility criterion is given by a PDE. Models may also be given by integral equations, nonlocal PDEs, difference equations, etc. We are interested in the PDE case. Let us discuss an example: the Korteweg-de Vries equation[1]

$$U_t = 6UU_X - U_{XXX} \tag{KdV}$$

for a function $U = U(t,X) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a model for shallow water waves [125], and one of the most well-studied nonlinear PDEs in existence. The easiest way to represent it as a model in our sense is as follows:

$$\begin{aligned} \mathcal{X}_\infty &= \mathcal{S}(\mathbb{R};\mathbb{R}) \\ \mathcal{A}_\infty &= \{(T,\varphi) : T > 0, \varphi \in C((-T,T), \mathcal{S}(\mathbb{R};\mathbb{R})) \text{ so that } U = \varphi \in C^{1,3}_{t,x}(\mathbb{R};\mathbb{R}) \text{ solves (KdV)}\}\,. \end{aligned}$$

Here $\mathcal{S}(\mathbb{R};\mathbb{R})$ is the Schwartz space, consisting of smooth functions which decay, together with their derivatives, faster than any polynomial. Initial results for the global well-posedness of $(\mathcal{X}_\infty, \mathcal{A}_\infty)$ are [32, 103]. Since it is advantageous to model a broader set of physical trajectories, it is natural to seek a model with the largest possible configuration space $\mathcal{X}$. In the case of PDE this usually means lower regularity, or weaker assumptions on integrability or decay. The pursuit of such a weakening of assumptions is directly tied to the strength of the model. The problem

[1] We use the capitalized variables U and X to be consistent with the notation in Chapter 4.

one encounters here is that (KdV) stops making sense if we use large function spaces such as $L^2(\mathbb{R};\mathbb{R})$. Instead, we need to write (KdV) in its distributional formulation, requiring that for all $\psi \in C_c^\infty([-T,T]\times\mathbb{R};\mathbb{R})$

$$\begin{aligned}&\int_{\mathbb{R}} U(t,X)\psi(t,X) - U(0,X)\psi(0,X)\,\mathrm{d}X - \int_0^t \int_{\mathbb{R}} U(t',X)\psi_t(t',X)\,\mathrm{d}X\,\mathrm{d}t' \\ &= \int_0^t \int_{\mathbb{R}} -3U(t',X)^2\psi_X(t',X) + U(t',X)\psi_{XXX}(t',X)\,\mathrm{d}X\,\mathrm{d}t'\,.\end{aligned} \tag{1.1.4}$$

We are now able to consider the model

$$\begin{aligned}&\mathcal{X}_0 = L^2(\mathbb{R};\mathbb{R})\\ &\mathcal{A}_0 = \{(T,\varphi) : T>0, \varphi\in C((-T,T),L^2(\mathbb{R};\mathbb{R})) \text{ so that } U=\varphi \text{ solves } (1.1.4)\}\,.\end{aligned}$$

The global well-posedness of $(\mathcal{X}_0,\mathcal{A}_0)$ was first obtained by J. Bourgain in his seminal work [35]. Can we consider an even larger model?

Suppose $(\mathcal{X},\mathcal{A})$ is a model and $\mathcal{X}'$ is a larger topological space so that $\mathcal{X}\subset\mathcal{X}'$ is a continuous and dense inclusion. Then

$$\mathcal{A}' = \overline{\mathcal{A}}^{\mathcal{T}(\mathcal{X}')}$$

is the natural way to *extend* the model $(\mathcal{X},\mathcal{A})$ to a model $(\mathcal{X}',\mathcal{A}')$. The well-posedness of $(\mathcal{X}',\mathcal{A}')$ decomposes into the well-posedness of $(\mathcal{X},\mathcal{A})$ and the continuity of the flow map with respect to the weaker topology. These claims are made precise by the following lemma.

Lemma 1.1.3. *The model $(\mathcal{X}',\mathcal{A}')$ is locally/globally well-posed with flow map Φ' if the model $(\mathcal{X},\mathcal{A})$ is locally/globally well-posed with a flow map Φ that satisfies:*

- ***Continuous dependence on data with respect to the weaker topology.*** *The map $\Phi:\mathcal{X}\to\mathcal{A}$ is continuous when $\mathcal{X}\subseteq\mathcal{X}'$ and $\mathcal{A}\subseteq\mathcal{A}'$ are equipped with the weaker topologies, i.e. the subspace topologies induced by the inclusions. Equivalently, for every $T^*>0$ the maps*

$$\begin{aligned}T\circ\Phi : \mathcal{X} &\longrightarrow (0,\infty]\\ \Phi\big|_{(-T^*,T^*)} : \{x\in\mathcal{X} : T(\Phi(x))\geq T^*\} &\longrightarrow C((-T^*,T^*);\mathcal{X})\end{aligned}$$

  are continuous when $\mathcal{X}\subseteq\mathcal{X}'$ is equipped with the subspace topology.

In this case, since $\mathcal{A}'$ is Hausdorff, Φ' is the unique continuous extension of Φ from the dense subset $\mathcal{X}\subseteq\mathcal{X}'$ to the whole space. In particular, if $(\mathcal{X}',\mathcal{A}_0')$ is another locally well-posed model with $\mathcal{A}\subseteq\mathcal{A}_0'$ and a flow map Φ_0' satisfying $T\circ\Phi_0'\big|_{\mathcal{X}} = T\circ\Phi$, then $\Phi_0' : \mathcal{X}'\to\mathcal{A}_0'$ is another continuous extension of Φ, and hence $\Phi_0'=\Phi'$. In the case of global well-posedness this implies $\mathcal{A}_0'=\mathcal{A}'$.

Let us apply this philosophy to our example of (KdV). We embed $\mathcal{X}_\infty = \mathcal{S}(\mathbb{R};\mathbb{R})\subseteq H^{-1}(\mathbb{R};\mathbb{R})$ and define

$$\begin{aligned}&\mathcal{X}_{-1} = H^{-1}(\mathbb{R};\mathbb{R})\\ &\mathcal{A}_{-1} = \overline{\mathcal{A}_\infty}^{\mathcal{T}(H^{-1}(\mathbb{R};\mathbb{R}))}\,.\end{aligned}$$

We cannot make sense of either of the formulations (KdV) or (1.1.4) for such a large class of functions, but the soft formulation purely based on the continuous dependence on initial data nevertheless allows us to give meaning to the question of well-posedness. In fact, the global well-posedness of $(\mathcal{X}_{-1},\mathcal{A}_{-1})$ was achieved in the seminal paper [114]. Furthermore, this is indeed the *largest model* on the scale of Sobolev spaces for which KdV is well-posed, i.e. for $s<-1$ the model with space $\mathcal{X}_s = H^s(\mathbb{R};\mathbb{R})$ is ill-posed [146]. We see that in the analysis of well-posedness of a model, or a family of models, the first step is to prove well-posedness in a high regularity setting, and the second step is to study the continuity of the flow map with respect to weaker topologies. In this dissertation we are concerned with the first step, working in settings of high regularity.

1.1.4. Three kinds of dynamics

After having attained a satisfactory understanding of the well-posedness of the model under consideration, in particular having realized it as a dynamical system, the next step is to study its dynamics with the goal of understanding and being able to predict them, ideally on long timescales. We pnow resent an informal classification of such dynamics into three kinds.

In the best case the dynamics are *predictable*, in the sense that the future configuration can be determined when the initial configuration is known to a sufficient degree of accuracy. We exemplify the 19th century viewpoint that dynamical systems should, in principle, be regarded as predictable with a well-known quote from P.-S. Laplace [130, Chapter II, §3][2]:

> *We ought then to regard the present state of the universe as the effect of its anterior state and as the cause of the one which is to follow. Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present to its eyes.*

Regardless of the degree to which Laplace and his peers were aware of the phenomenon of instability, the discovery of *chaotic* dynamics towards the end of the 19th century (see e.g. [90, 154]) certainly introduced a new perspective, establishing fundamental restrictions to the predictability of many dynamical systems. Chaotic dynamics are highly sensitive to their initial data, meaning that a small error quickly propagates to major, qualitative differences in trajectories. Even though they are mathematically deterministic, they are, for all intents and purposes, *totally unpredictable* on long timescales.

Subsequently in the early 20th century, the theory of computation was developed, from which we highlight the introduction of Turing machines and Turing universality [174], the λ-calculus [51, 52], and the formulation of the so-called *Church–Turing thesis*, which claims that all reasonable models of computation are equivalent to, or can be described by, the aforementioned ones. Towards the end of the 20th century this led to the discovery (or perhaps awareness) of a third kind of *computational* dynamics, which sit somewhere between *predictable* and *chaotic*. This can be summarized by the idea that *computation happens at the edge of chaos* [55, 129, 184]. These kinds of dynamics are predictable in principle, but in practice there is no way to make predictions any more efficiently than simply observing the evolution of the system. Characteristic of computational dynamics are complex self-organizing structures, pattern formation, and the possibility to emulate any algorithm (as in Turing machine) by a careful preparation of the initial data.

1.1.5. Complete integrability

There is no accepted unifying and rigorous definition of complete integrability. Let us vaguely characterize this notion by framing it as a fourth kind of dynamics: those which are *predictable when you know the trick.* The suggestion here is that, although the system in question was already understood to be generally non-chaotic, there had been no easy way to give an explicit description of the dynamics until a *trick* was discovered that simplifies the dynamics to ones which are already well-understood. I was first exposed to this viewpoint in the informal lecture series [61] by P. A. Deift. They give the following definition.

Definition 1.1.4 (Completely integrable dynamical system)**.** A dynamical system $(\mathcal{X}, \Phi)$ is completely integrable if there exists a homeomorphism $\eta : \mathcal{X} \to \mathcal{Y}$ to another space $\mathcal{Y}$ so that both the transformation η and the dynamical system $(\mathcal{Y}, \eta \circ \Phi \circ \eta^{-1})$ are, in a sense, *well-understood.*

We would also like to mention work in the 1990s and early 2000s (see e.g. [3, 59, 98]) investigating the degree to which completely integrable dynamics can be of computational character. For example, the Manakov system (which is a kind of vector NLS system), and more specifically its

[2] The original reference is [131]

soliton interactions, have been found by K. Steiglitz in [165] to be Turing universal, in a certain sense.

In practice, complete integrability describes a number of phenomena and has various definitions, depending on the context. The following is a list of properties which one can consider to be characteristic of completely integrable dynamics/systems/problems.

(i) The existence of a transformation that renders the system explicitly solvable, usually by the use of integrals (hence "integrability").

(ii) The existence of an infinite number of *commuting* symmetries. Here "symmetry" refers to a continuous one-parameter family of transformations $\mathcal{X} \to \mathcal{X}$ that maps admissible trajectories into admissible trajectories, and "commuting" means that the application of multiple symmetries is independent of the ordering.

(iii) The existence of a set of conserved quantities which is, in a certain sense, *maximal* and *orthogonal*. In the case of a Hamiltonian system on a $2n$-dimensional symplectic manifold this is known as Liouville-Arnold integrability. Here "maximal" means n conserved quantities, and "orthogonal" means pairwise Poisson commuting and having linearly independent differentials.

(iv) The existence of an explicitly computable change of variables into *actions* and *angles*, where the time-evolution of the actions is constant, and the evolution of the angles is linear.

(v) The existence of a Lax pair structure, which consists of a mapping from states $x \in \mathcal{X}$ to operators $L(x)$ and $P(x)$ so that a trajectory $\varphi \in C((-T,T);\mathcal{X})$ is admissible, i.e. corresponds to an evolution of the dynamical system, if and only if

$$\frac{\mathrm{d}}{\mathrm{d}t} L(\varphi(t)) = [P(\varphi(t)), L(\varphi(t))] \,. \tag{1.1.5}$$

(vi) The existence of *soliton* solutions. A "soliton" is a solitary (i.e. isolated, lonely, single) wave, kink, or front, that interacts with other solitons in a way that is particle-like. PDEs that admit solitons are also called *soliton equations*.

The study of complete integrability in this sense began with the analysis of various finite-dimensional Hamiltonian systems. The classic examples here are the Euler, Lagrange, and Kowalevskaya spinning tops, which are in a certain sense the only rigid body spinning tops that have completely integrable dynamics. We emphasize the role of examples because complete integrability is generally understood to be a fragile property that is only true for a relatively small number of dynamical systems, and hence the theory hinges, in particular in infinite dimensions, on an ever-growing collection of known examples.

The complete integrability of finite-dimensional Hamiltonian systems has a satisfying characterization in the form of the Liouville-Arnold theorem (proven in its first version in 1855) [15, 136], which states in essence that (iii) $\iff$ (iv). Since the evolution of the action-angle variables is trivial, we can say that (iii) $\implies$ (i). Noether's theorem yields a correspondence (ii) $\iff$ (iii). It took another century for the first genuinely infinite-dimensional and completely integrable Hamiltonian systems to be discovered. We refer to [150, Chapter 3] for an overview. To summarize it briefly, in the famous Fermi-Pasta-Ulam and later Kruskal-Zabusky experiments, unexpected behaviors were observed in numerical simulations of various time-evolving lattice systems. These were later connected to the Korteweg-de Vries (see (KdV) above) and Boussinesq equations, which are now known to be completely integrable.

Indeed, the first PDE which was proven to be completely integrable is KdV, first via the inverse scattering transform (IST) method [75] and later by giving action-angle coordinates [185], making it the first infinite-dimensional completely integrable Hamiltonian system. Here the IST method can be understood as a kind of complete integrability in the sense of (i). The second such PDE is the *focusing* cubic nonlinear Schrödinger equation

$$iq_t = -q_{xx} - 2|q|^2 q \tag{fNLS}$$

for a wavefunction $q = q(t,x) : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$ (see [188] for the IST method and [175] for action-angle variables). We are interested instead in the *defocusing* nonlinear Schrödinger equation

$$iq_t = -q_{xx} + 2|q|^2 q \tag{NLS}$$

with either of the following boundary conditions:

$$\lim_{x\to\pm\infty} q(t,x) = 0 \tag{ZBC}$$

$$\lim_{x\to\pm\infty} q(t,x) = e^{-2it}q_\pm \text{ where } |q_\pm| = 1\,. \tag{NZBC}$$

We refer to the seminal paper [189] for the initial application of the IST method in the setting NLS–NZBC.

The modern approach to the complete integrability of PDEs is via the Lax pair formalism (v) above, introduced in [132] by P. Lax. Here the key objects are the so-called Lax operators, which for the systems KdV and NLS are

$$L^{\mathsf{NLS}} = \begin{pmatrix} i\partial_x & -iq \\ i\overline{q} & -i\partial_x \end{pmatrix} \qquad L^{\mathsf{KdV}} = \begin{pmatrix} -\partial_X^2 + U & 0 \\ -U_X & -\partial_X^2 + U \end{pmatrix}. \tag{1.1.6}$$

There exist corresponding operators P^{KdV}, P^{fNLS} and P^{NLS}, depending on the potentials q and U so that (1.1.5) is equivalent to respectively (KdV), (fNLS), and (NLS). A consequence of (1.1.5) is that eigenvalues of the Lax operators are conserved quantities. A deeper consequence is that by studying the scattering problems associated to the Lax operators, one can construct bijective transformations from the potentials q and U, given regularity, integrability and decay assumptions, to their *scattering data*. The time-evolution of the PDEs then becomes trivial for the scattering data (in a way which can be understood as action-angle coordinates), rendering the PDEs explicitly solvable if one is able to perform the inverse scattering transform, i.e. from the scattering data back to the potentials. We therefore have a connection (v) $\Longrightarrow$ (i)–(iv).

Establishing appropriate mapping properties of the IST is generally difficult, which is why the study of KdV, NLS–ZBC, and NLS–NZBC is not trivialized by their complete integrability. Satisfactory bijective mapping properties for the IST have been established for L^{KdV} [60, 101, 102] and for L^{NLS} [156, 192]. For potentials q satisfying NZBC the understanding of the IST is not as complete (see the survey [155]). In the absence of appropriate mapping properties, one can still use the IST method to construct special solutions, such as the soliton and multi-soliton solutions. These are connected to the spectral information of the Lax operator, and the invariance of this spectral information under time evolution explains the surprising stability of the solitons. In this sense, we can say that *(v)* $\Longrightarrow$ *(vi)*. A natural question that remains is whether one can say that (vi) $\Longrightarrow$ (i)–(v). This seems to be not well-understood.

1.1.6. Transmission coefficients and the Hamiltonians

In this dissertation we are less interested in (i), (iv), and (vi), i.e. the IST method, explicit solvability, solitons, etc., and more in (ii) and (iii), as these are closely connected to the concept of *integrable hierarchies*, of which the NLS and KdV hierarchies are examples. Specifically, the symmetries mentioned in (ii) are flow maps corresponding to those Hamiltonian PDEs, which have as Hamiltonians the conserved quantities mentioned in (iii). We describe now how these conserved quantities can be obtained from the Lax operators (1.1.6). Let $j \in \{1,2\}$ and set $\gamma_1 = 1$ and $\gamma_2 = -1$. The first step is to consider the so-called *Jost solutions* $\Phi_j^{\mathsf{NLS},\pm}$ and $\Phi_j^{\mathsf{KdV},\pm}$, which are defined for $\lambda, k \in \mathbb{C}$ as the solutions to the scattering problems

$$L^{\mathsf{NLS}}\, \Phi_j^{\mathsf{NLS},\pm}(x,\lambda;q) = \lambda\, \Phi_j^{\mathsf{NLS},\pm}(x,\lambda;q)\,, \qquad L^{\mathsf{KdV}}\, \Phi_j^{\mathsf{KdV},\pm}(X,k;U) = k^2\, \Phi_j^{\mathsf{KdV},\pm}(X,k;U)\,,$$

$$\lim_{x\to\pm\infty} \Phi_j^{\mathsf{NLS},\pm}(x,\lambda;q)e^{xi\lambda\gamma_j} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{-,j}, \qquad \lim_{X\to\pm\infty} \Phi_j^{\mathsf{KdV},\pm}(X,k;U)e^{Xik\gamma_j} = \begin{pmatrix} 1 & 1 \\ ik & -ik \end{pmatrix}_{-,j}.$$

If the potentials q and U are Schwartz functions and $\operatorname{Im}\lambda \geq 0$ (resp $\operatorname{Im} k \geq 0$), then the Jost solutions $\Phi_1^{\mathsf{NLS},-}, \Phi_2^{\mathsf{NLS},+}$ (resp. $\Phi_1^{\mathsf{KdV},-}, \Phi_2^{\mathsf{KdV},+}$) exist and are smooth in x (resp. X) and analytic in λ (resp. k), up to a finite set of points. If $\operatorname{Im}\lambda = 0$ (resp. $\operatorname{Im} k = 0$) then this holds true for all the Jost solutions. In the latter case there exist numbers $a^{\mathsf{NLS}}(\lambda;q), b^{\mathsf{NLS}}(\lambda;q), a^{\mathsf{KdV}}(k;U), b^{\mathsf{KdV}}(k;U) \in \mathbb{C}$ so that

$$\Phi_1^{\mathsf{NLS},-} = a^{\mathsf{NLS}}\Phi_1^{\mathsf{NLS},+} + b^{\mathsf{NLS}}\Phi_2^{\mathsf{NLS},+} \qquad \Phi_1^{\mathsf{KdV},-} = a^{\mathsf{KdV}}\Phi_1^{\mathsf{KdV},+} + b^{\mathsf{KdV}}\Phi_2^{\mathsf{KdV},+}\,.$$

Together with the eigenvalues of the Lax operator and certain associated normalizing constants, these form the so-called *scattering data* mentioned above, from which the potentials q and U can be reconstructed. We are interested only in the *transmission coefficients* a^{NLS} and a^{KdV}. They can be extended analytically to the upper half-plane, and if $\operatorname{Im}\lambda$ and $\operatorname{Im}k$ are sufficiently large they have the representation

$$\log a^{\mathsf{NLS}}(\lambda;q) = \int_{\mathbb{R}} \frac{\partial_x \Phi_{1,1}^{\mathsf{NLS},-}(x,\lambda;q)}{\Phi_{1,1}^{\mathsf{NLS},-}(x,\lambda;q)} + i\lambda \,\mathrm{d}x \quad \log a^{\mathsf{KdV}}(k;U) = \int_{\mathbb{R}} \frac{\partial_X \Phi_{1,1}^{\mathsf{KdV},-}(X,k;U)}{\Phi_{1,1}^{\mathsf{KdV},-}(X,k;U)} + ik \,\mathrm{d}X\,.$$

These integrals admit asymptotic expansions as $\lambda, k \to \infty$, and we define the aforementioned conserved quantities to be precisely the following expansion coefficients:

$$\log a^{\mathsf{NLS}}(\lambda;q) \sim i\sum_{n=0}^{\infty} \frac{\mathcal{H}_n^{\mathsf{NLS}}(q)}{(2\lambda)^{n+1}} \qquad\qquad \log a^{\mathsf{KdV}}(k;U) \sim i\sum_{n=-1}^{\infty} \frac{\mathcal{E}_n^{\mathsf{KdV}}(U)}{(2k)^{2n+3}}\,. \tag{1.1.7}$$

Their conservation under the flows of NLS and KdV is a consequence of the conservation of the transmission coefficient itself. In fact, they are pairwise Poisson commuting with respect to the relevant Poisson brackets. By deriving recurrence relations for the densities of these quantities, they can be explicitly computed. For the case of KdV the initial ones are

$$\mathcal{E}_{-1}^{\mathsf{KdV}}(U) = \int_{\mathbb{R}} U \,\mathrm{d}X \tag{Mass}$$

$$\mathcal{E}_{0}^{\mathsf{KdV}}(U) = \int_{\mathbb{R}} U^2 \,\mathrm{d}X \tag{Momentum}$$

$$\mathcal{E}_{1}^{\mathsf{KdV}}(U) = \int_{\mathbb{R}} U_X^2 + 2U^3 \,\mathrm{d}X \tag{Energy}$$

$$\mathcal{E}_{2}^{\mathsf{KdV}}(U) = \int_{\mathbb{R}} U_{XX}^2 + 10U_X^2 U + 5U^4 \,\mathrm{d}X\,.$$

For the case of NLS–ZBC they are

$$\mathcal{H}_{0}^{\mathsf{NLS}}(q) = \int_{\mathbb{R}} q\overline{q} \,\mathrm{d}x \tag{Mass}$$

$$\mathcal{H}_{1}^{\mathsf{NLS}}(q) = -i\int_{\mathbb{R}} q\overline{q}_x \,\mathrm{d}x \tag{Momentum}$$

$$\mathcal{H}_{2}^{\mathsf{NLS}}(q) = \int_{\mathbb{R}} q_x\overline{q}_x + q^2\overline{q}^2 \,\mathrm{d}x \tag{Energy}$$

$$\mathcal{H}_{3}^{\mathsf{NLS}}(q) = i\int_{\mathbb{R}} q\overline{q}_{xxx} - 4q^2\overline{q}\overline{q}_x - q_x q\overline{q}^2 \,\mathrm{d}x\,.$$

In Sections 3.2 and 4.2 the Jost solutions, the transmission coefficients, and their asymptotic expansions are presented in detail.

1.1.7. The NLS and KdV hierarchies

We define the KdV hierarchy as the infinite hierarchy of Hamiltonian PDEs

$$\partial_{T_n} U = \partial_X \frac{\delta \mathcal{E}_n^{\mathsf{KdV}}(U)}{\delta U}$$

for $n \geq 0$ and $U = U(T_n, X) : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$. Similarly, we define the NLS hierarchy as the infinite hierarchy of Hamiltonian PDEs

$$i\partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\mathsf{NLS}}(q)}{\delta \overline{q}}$$

for $n \geq 0$ and $q = q(t_n, x) : \mathbb{R}\times\mathbb{R} \to \mathbb{C}$. Since the Hamiltonians are pairwise Poisson commuting with respect to the relevant Poisson bracket, and hence their flow maps commute, an alternative viewpoint is to consider instead for a fixed $N \in \mathbb{N}$ and a single function $U = U(T_0, \dots, T_N, X) : \mathbb{R}^{N+1} \times \mathbb{R} \to \mathbb{R}$ the system of equations

$$\partial_{T_n} U = \partial_X \frac{\delta \mathcal{E}_n^{\mathsf{KdV}}(U)}{\delta U} \qquad \forall n \in \{0, \dots, N\}, \tag{1.1.8}$$

and analogously for a single function $q = q(t_0, \dots, t_N, x) : \mathbb{R}^{N+1} \times \mathbb{R} \to \mathbb{C}$ the system of equations

$$i\partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\mathsf{NLS}}(q)}{\delta \overline{q}} \qquad \forall n \in \{0, \dots, N\}. \tag{NLS_n}$$

We sometimes refer to these as the N*-truncated* NLS and KdV hierarchies.

The central object of this dissertation is the NLS hierarchy together with the nonzero boundary condition

$$\lim_{x \to \pm\infty} q(t_0, \dots, t_N, x) = q_\pm e^{i\varphi(t_0, \dots, t_N)} \text{ where } |q_\pm| = 1 \text{ and } \varphi(t_0, \dots, t_N) \in \mathbb{R}. \tag{NZBC}$$

Note that the Hamiltonians $\mathcal{H}_n^{\mathsf{NLS}}(q)$ are not well-defined in this setting. Instead, we work with two alternative families of Hamiltonians $\mathcal{H}_n^{\mathsf{GP}}(q)$ and $\mathcal{H}_n^{\widetilde{\mathsf{GP}}}(q)$, which are well-defined for q satisfying NZBC. The initial ones are

$$\mathcal{H}_{-1}^{\mathsf{GP}}(q) = -i \log\left(\frac{q_-}{q_+}\right) \tag{Asymptotic phase change}$$

$$\mathcal{H}_0^{\mathsf{GP}}(q) = \int_{\mathbb{R}} q\overline{q} - 1 \,\mathrm{d}x \tag{Mass}$$

$$\mathcal{H}_1^{\mathsf{GP}}(q) = -i \int_{\mathbb{R}} q\overline{q}_x \,\mathrm{d}x \tag{Momentum}$$

$$\mathcal{H}_2^{\mathsf{GP}}(q) = \int_{\mathbb{R}} q_x\overline{q}_x + (q\overline{q} - 1)^2 \,\mathrm{d}x \tag{Energy}$$

$$\mathcal{H}_3^{\mathsf{GP}}(q) = i \int_{\mathbb{R}} q\overline{q}_{xxx} - 4q^2\overline{q}\overline{q}_x - q_x q\overline{q}^2 + 4q\overline{q}_x \,\mathrm{d}x$$

and

$$\begin{aligned}
\mathcal{H}_{-1}^{\widetilde{\mathsf{GP}}}(q) &= \frac{1}{2}\mathcal{H}_{-1}^{\mathsf{GP}}(q) \\
\mathcal{H}_0^{\widetilde{\mathsf{GP}}}(q) &= \mathcal{H}_0^{\mathsf{GP}}(q) \\
\mathcal{H}_1^{\widetilde{\mathsf{GP}}}(q) &= -i \int_{\mathbb{R}} q\overline{q}_x \,\mathrm{d}x - \mathcal{H}_{-1}^{\mathsf{GP}}(q) \\
\mathcal{H}_2^{\widetilde{\mathsf{GP}}}(q) &= \mathcal{H}_2^{\mathsf{GP}}(q) \\
\mathcal{H}_3^{\widetilde{\mathsf{GP}}}(q) &= i \int_{\mathbb{R}} q\overline{q}_{xxx} - 4q^2\overline{q}\overline{q}_x - q_x q\overline{q}^2 + 6q\overline{q}_x \,\mathrm{d}x + 3\mathcal{H}_{-1}^{\mathsf{GP}}(q).
\end{aligned}$$

Here log denotes the principal logarithm. They can be understood as renormalized versions of the Hamiltonians $\mathcal{H}^{\mathsf{NLS}}$. The superscript GP stands for the Gross-Pitaevskii equation

$$iq_t + q_{xx} = 2q(|q|^2 - 1), \tag{GP}$$

which is formally equivalent to NLS under the transformation $q \mapsto e^{2it}q$. These alternative Hamiltonians can also be defined from the transmission coefficient associated to the Lax operator $L^{\mathsf{GP}} = L^{\mathsf{NLS}}$. To do this we let q be a smooth potential satisfying NZBC, and more specifically $q - q_\pm \in \mathcal{S}(\mathbb{R}; \mathbb{C})$ for certain $q_\pm \in \mathbb{C}$ with $|q_\pm| = 1$. Let $\lambda, z \in \mathbb{C}$ with $\lambda^2 = z^2 - 1$, $\operatorname{Im}\lambda, \operatorname{Im} z \geq 0$, and consider the scattering problem

$$\begin{aligned}
L^{\mathsf{GP}}\ \Phi_j^{\mathsf{GP},\pm}(x, \lambda; q) &= \lambda\ \Phi_j^{\mathsf{GP},\pm}(x, \lambda; q) \\
\lim_{x \to \pm\infty} \Phi_j^{\mathsf{GP},\pm}(x, \lambda; q) e^{xiz\gamma_j} &= \begin{pmatrix} q_\pm & i(z - \lambda) \\ i(\lambda - z) & \overline{q}_\pm \end{pmatrix}_{-,j}.
\end{aligned}$$

The definition of the transmission coefficient $a^{\mathsf{GP}}(\lambda; q)$ and its asymptotic expansion is analogous to the above, but more involved due to the nonzero boundary condition. Once accomplished, one obtains the asymptotic expansion

$$\log a^{\mathsf{GP}}(\lambda) \sim i \sum_{n=-1}^{\infty} \frac{\mathcal{H}_n^{\mathsf{GP}}(q)}{(2z)^{n+1}} \sim i \sum_{n=0}^{\infty} \frac{\mathcal{H}_{2n}^{\widetilde{\mathsf{GP}}}(q)}{(2z)^{2n+1}} + i\frac{\lambda}{z} \sum_{n=0}^{\infty} \frac{2^{\mathbb{1}_{\{n=0\}}} \mathcal{H}_{2n-1}^{\widetilde{\mathsf{GP}}}(q)}{(2z)^{2n}} . \tag{1.1.9}$$

From this one can calculate that $\mathcal{H}_{2n}^{\mathsf{GP}} = \mathcal{H}_{2n}^{\widetilde{\mathsf{GP}}}$ and

$$\mathcal{H}_{2n-1}^{\widetilde{\mathsf{GP}}} = \sum_{m=0}^{n} \binom{-\frac{1}{2}}{n-m} 4^{n-m} 2^{-\mathbb{1}_{\{m=0\}}} \mathcal{H}_{2m-1}^{\mathsf{GP}} . \tag{1.1.10}$$

Now we define the GP and $\widetilde{\mathsf{GP}}$ hierarchies from the corresponding Hamiltonians, in analogy with the definition of the NLS hierarchy, i.e. as the systems

$$i\partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\mathsf{GP}}(q)}{\delta \bar{q}} \qquad \forall\, n \in \{0, \dots, N\} \tag{GP_n}$$

$$i\partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\widetilde{\mathsf{GP}}}(q)}{\delta \bar{q}} \qquad \forall\, n \in \{0, \dots, N\} \tag{$\widetilde{\mathsf{GP}}_n$}$$

for a function $q = q(t_0, \dots, t_N, x) : \mathbb{R}^{N+1} \times \mathbb{R} \to \mathbb{C}$. They are equivalent to the NLS hierarchy in a sense which needs to be extracted from the relation

$$\frac{\delta \mathcal{H}_{2n}^{\mathsf{NLS}}(q)}{\delta \bar{q}} = \sum_{m=0}^{n} \binom{n-\frac{1}{2}}{n-m} 4^{n-m} \frac{\delta \mathcal{H}_{2m}^{\mathsf{GP}}(q)}{\delta \bar{q}} \tag{1.1.11}$$

$$\frac{\delta \mathcal{H}_{2n-1}^{\mathsf{NLS}}(q)}{\delta \bar{q}} = \sum_{m=0}^{n} \binom{n-1}{n-m} 4^{n-m} \frac{\delta \mathcal{H}_{2m-1}^{\mathsf{GP}}(q)}{\delta \bar{q}} . \tag{1.1.12}$$

They are more suitable for working with NZBC, and the $\widetilde{\mathsf{GP}}$ hierarchy in particular plays a crucial role in Chapters 3 and 4. We consider the recognition that in various circumstances the $\widetilde{\mathsf{GP}}$ hierarchy is the "better" hierarchy to work with to be one of the main contributions of this dissertation.

For more discussion and references on the relevance of these hierarchies in mathematical physics, the theory of long-wave modulation equations, and algebraic geometry, we refer to the introductions of Chapters 3 and 4.

1.2. MAIN RESULTS

Let us now revisit the three points that we gave at the very beginning and expand on them.

1. Chapter 2 consists of the published paper [180]. We prove that in the absence of vacuum and for $s > \frac{1}{2}$, the Madelung transform $q \mapsto (|q|^2, \operatorname{Im}[q^{-1}\partial_x q])$ is a locally bilipschitz map between the complete metric spaces (X^s, d^s), introduced in [119], and the complete metric space $(\mathcal{Y}^s, \theta^s) = (1 + H^s(\mathbb{R})) \times H^{s-1}(\mathbb{R})$. As a result, the global well-posedness result for NLS–NZBC from [119, 120] can be transferred to the hydrodynamic formulation.
2. Chapter 3 consists of the preprint [135], which is joint work with X. Liao. We prove the local well-posedness of the NLS hierarchy with nonzero boundary data in weighted Sobolev spaces at high regularity, by determining the structure of the higher PDEs sufficiently in order to apply methods for general dispersive nonlinear systems of PDEs. This involves generalized local smoothing and maximal function estimates, which are of independent interest. Subsequently, we use the conserved energies constructed in [119, 120] to conclude global well-posedness.
3. Chapter 4 consists of a preprint by the author, which is in preparation. We prove that long-wave solutions to the NLS hierarchy, of the kind that were constructed in Chapter 3, when transformed to the hydrodynamic variables using the mapping studied in Chapter 1,

behave approximately like solutions to the KdV hierarchy. This is based on an approximation of the Hamiltonians $\mathcal{H}^{\widetilde{\mathsf{GP}}}$ in hydrodynamic variables by the energies $\mathcal{E}^{\mathsf{KdV}}$, which is related to an approximation result for the transmission coefficients of the associated Lax operators. This can also be viewed as a result on the splitting of the one-dimensional Dirac operator in the nonrelativistic limit.

Each chapter is self-contained, so they may be read in any order.

We conclude the introduction with a presentation of the main results from each chapter in a simplified manner. Since the literature is reviewed in full detail in the respective introductions, we give only the most directly relevant references in the following.

1.2.1. MAIN RESULTS OF CHAPTER 2

In this chapter we study the one-dimensional defocusing Gross-Pitaevskii equation

$$i\partial_t q + \partial_{xx} q - 2(|q|^2 - 1)q = 0 \tag{GP}$$

for a wavefunction $q = q(t, x) : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$, subject to the nonzero boundary condition

$$\lim_{|x|\to\infty} |q(t,x)|^2 = 1\,.$$

We are interested in solutions with

$$E(q) = \frac{1}{2}\int_{\mathbb{R}} |\partial_x q|^2 + (|q|^2 - 1)^2 \, \mathrm{d}x < \frac{4}{3}\,. \tag{1.2.1}$$

The energy[3] $E(q)$ is a conserved quantity and controls the distance from vacuum, meaning that if $E(q) < \frac{4}{3}$, then there exists some $\delta = \delta(E(q)) > 0$ such that $|q| > \delta$ (see e.g. Lemma 2.1.2 below). In this setting, we define the Madelung transform

$$\mathcal{M}(q) = (\rho, v) \text{ where } \rho = |q|^2 \text{ and } v = \operatorname{Im}\left[\frac{q_x}{q}\right]$$

with corresponding inverse

$$\mathcal{M}^{-1}(\rho, v) = \sqrt{\rho}e^{i\varphi} \text{ where } \partial_x \varphi = v\,.$$

Note that here φ is only defined up to a choice of constant, as this is information which the Madelung transform drops. Furthermore, there is little difference between the use of the variable $\rho = |q|^2$ as opposed to $a = |q|$: all results can be rewritten to use (a, v) instead. Under the Madelung transform, the Gross-Pitaevskii equation is transformed into the hydrodynamic system of PDEs

$$\begin{cases} \partial_t \rho + 2\partial_x(\rho v) = 0\,, \\ \partial_t v + \partial_x(v^2) + 2\partial_x \rho = \partial_x\left(\partial_x\left(\frac{1}{2}\frac{\partial_x \rho}{\rho}\right) + \left(\frac{1}{2}\frac{\partial_x \rho}{\rho}\right)^2\right)\,, \end{cases} \tag{hGP}$$

which is commonly referred to as the hydrodynamic Gross-Pitaevskii equations. The corresponding energy[4] is $\mathcal{E}(\rho, v) = E(q)$. Our main result is the global well-posedness of (hGP) in the metric space $(\mathcal{Y}^s, \theta^s)$, defined by

$$\mathcal{Y}^s = (1 + H^s(\mathbb{R})) \times H^{s-1}(\mathbb{R})$$
$$\theta^s((\rho, v), (\eta, w)) = \|\rho - \eta\|_{H^s} + \|v - w\|_{H^{s-1}}\,.$$

Theorem 2.1.9 (simplified)**.** Let $s \geq 1$. The system (hGP) is globally well-posed in the metric space $(\{(\rho_0, v_0) \in \mathcal{Y}^s : \mathcal{E}(\rho_0, v_0) < \frac{4}{3}\}, \theta^s)$.

The integer-regularity cases of this theorem were previously obtained in [145]. The global well-posedness we obtain follows from the well-posedness theory for (GP) that was previously established

[3] In Chapters 3 and 4 we use $\mathcal{H}_2^{\mathsf{GP}}(q) = 2E(q)$

[4] In Chapter 4 we use $\mathcal{H}_2^{\mathsf{hGP}}(w) = 2\mathcal{E}(\rho, v)$

in [119, 120]. Here global well-posedness is proven for $s \geq 0$ in the complete metric space (X^s, d^s), where

$$X^s = \{q \in H^s_{\mathrm{loc}}(\mathbb{R};\mathbb{C}) : E^s(q) < \infty\} \,/\, \mathbb{S}^1$$
$$d^s(q,p) = \left(\int_{\mathbb{R}} \inf_{\lambda \in \mathbb{C}, |\lambda|=1} \|\operatorname{sech}(y-\cdot)(\lambda q - p)\|^2_{H^s} \,\mathrm{d}y\right)^{\frac{1}{2}} .$$

The core of our result is correspondingly the following corollary.

Theorem 2.1.6 (simplified)**.** Let $s > \frac{1}{2}$ and $\delta > 0$. The map

$$\left(\{q \in X^s : |q| > \delta\}, d^s\right) \xrightarrow{\mathcal{M}} \left(\{(\rho, v) \in \mathcal{Y}^s : \rho > \delta^2\}, \theta^s\right)$$

is a bilipschitz equivalence, i.e. a map which is bijective, Lipschitz continuous, and has a Lipschitz continuous inverse.

Remark 1.2.1. *In Theorem 2.1.9 we only restrict ourselves to $s \geq 1$ because the product of distributions $\rho \cdot v$ is not defined anymore if $s < 1$, as then it is possible that $v \notin L^2(\mathbb{R})$. Nevertheless, we can still transfer the flow map for (GP) to the hydrodynamic setting for $s \in (\frac{1}{2}, 1)$, meaning the global well-posedness can be stated in the sense described in Section 1.1.3. In order to obtain a satisfying well-posedness result for $s \in (\frac{1}{2}, 1)$, one should prove that due to some smoothing effect the trajectories have more regularity than expected and indeed solve (hGP) in the sense of distributions.*

The proof uses an intermediate metric: we prove that for a covering $(B_k)_{k\in\mathbb{Z}}$ of $\mathbb{R}$ consisting of evenly spaced balls of identical radius, we have

$$d^s \sim \sum_{k\in\mathbb{Z}} d^s_*\big|_{B_k} \sim \theta^s ,$$

where for a ball B we define

$$d^s_*\big|_B(q,p) = \inf_{\lambda \in \mathbb{S}^1} \|\lambda q - p\|_{W^{s,2}(B)} .$$

Since Theorem 2.1.6 establishes an equivalence between natural function spaces for the variables q and (ρ, v), it can be of use in settings much more general than Theorem 2.1.9. For example, we use it in Chapter 4 to transfer the well-posedness result from Chapter 3 for the NLS hierarchy to the alternative hydrodynamic variables

$$w_{\pm} = \frac{v}{2} \pm (\sqrt{\rho} - 1) = \frac{v}{2} \pm (a-1) . \tag{1.2.2}$$

1.2.2. Main results of Chapter 3

In this chapter we study the aforementioned NLS, GP, and $\widetilde{\mathsf{GP}}$ hierarchies in the setting of NZBC with the goal of proving their well-posedness. Conveniently, H. Koch and X. Liao have already constructed in [119, 120] a family of conserved energies for GP that control every Sobolev norm $H^s(\mathbb{R})$ for $s \geq 0$. These energies are conserved for every flow in the NLS, GP, and $\widetilde{\mathsf{GP}}$ hierarchies because they are defined from the transmission coefficient $a^{\mathsf{GP}}(\lambda; q)$, which is associated with the scattering problem for the Lax operator $L^{\mathsf{GP}} = L^{\mathsf{NLS}}$ for a complex parameter $\lambda \in \mathbb{C}$. This is a central object that is studied in detail in Sections 3.2 and 4.2. As a result, the main difficulty is to prove local well-posedness. The equations (NLS_n) are dispersive nonlinear PDEs of high order, so once we understand their linear part we can apply known techniques for general dispersive nonlinear PDEs. This approach restricts us to settings of high regularity. By working with the definition of the Hamiltonians and the transmission coefficient, we are able to prove the following theorem on the structure of the equations.

Theorem 3.1.10 (simplified)**.** The right hand sides of (NLS_n), (GP_n), and ($\widetilde{\mathsf{GP}}_n$) consist of sums of monomials involving q, $\overline{q}$, and up to n derivatives thereof. We give explicit formulas for these right hand sides up to monomials which have two or more factors that are derivatives of q or $\overline{q}$.

For further efforts to calculate explicit coefficients in integrable hierarchies, we refer to [4], where the coefficients of all cubic terms in the dNLS hierarchy are calculated, and also [19] for the KdV hierarchy. We now choose a fixed front $q_* : \mathbb{R} \to \mathbb{C}$ with $\lim_{x\to\pm\infty} q_*(x) = q_\pm$ so that q_* satisfies NZBC. Then we rewrite the well-posedness problem for a solution q of the higher (i.e. $n \geq 2$, as the cases $n \in \{0, 1\}$ are trivial) flows of the NLS, GP and $\widetilde{\text{GP}}$ hierarchies in terms of the variable problem for the perturbation $p = q - q_*$. This yields dispersive nonlinear PDEs with coefficients that depend on q_*. From our structure result we derive the following formulation.

Proposition 3.1.21 (simplified)**.** Let $n \geq 2$. The n-th equation in the $\widetilde{\text{GP}}$ hierarchy ($\widetilde{\text{GP}}_n$) can be written for the vector $\mathbf{p} = (p, q_*^2\overline{p}, \overline{p}, \overline{q}_*^2 p)$ in the form

$$i\partial_{t_n}\mathbf{p} = \mathcal{L}^n[\mathbf{p}] + \mathcal{N}^n[\mathbf{p}]\,, \tag{1.2.3}$$

where with $D_x = -i\partial_x$ we have

$$\mathcal{L}^{2m} = D_x^{2m-2}\begin{pmatrix} D_x^2+2 & 2 & 0 & 0 \\ -2 & -D_x^2-2 & 0 & 0 \\ 0 & 0 & -D_x^2-2 & -2 \\ 0 & 0 & 2 & D_x^2+2 \end{pmatrix} \tag{1.2.4}$$

$$\mathcal{L}^{2m+1} = -D_x^{2m+1}\,, \tag{1.2.5}$$

and for each $j \in \{1, 2, 3, 4\}$ the terms $\mathcal{N}^n[\mathbf{p}]$ are sums of monomials in $q_*, \overline{q}_*, p, \overline{p}$ and their derivatives. They involve at most 2 fewer derivatives of $\mathbf{p}$ than the linear part, and always have at least one factor which has "good integrability" like p, $(q_*)_x$, $|q_*|^2 - 1$, or a derivative thereof, and at least two such factors if the monomial depends on p or $\overline{p}$.

The well-posedness theory developed in the 1990s by C. E. Kenig, G. Ponce and L. Vega [106, 108, 109, 110, 111] is in principle able to tackle such systems. After proving generalized local smoothing and maximal function estimates, we obtain a well-posedness theory for general systems of dispersive nonlinear PDEs of the form (1.2.3). This uses the (weighted) Sobolev spaces

$$\begin{aligned}\big(H^s(\mathbb{R}), \|\cdot\|_{H^s}\big) &= \big(\{u \in L^2(\mathbb{R}) : (1+|\xi|^2)^{\frac{s}{2}}\widehat{u}(\xi) \in L^2(\mathbb{R})\}, \|(1+|\xi|^2)^{\frac{s}{2}}\widehat{\cdot}(\xi)\|_{L^2(\mathbb{R})}\big)\\ \big(H^{s',1}(\mathbb{R}), \|\cdot\|_{H^{s',1}(\mathbb{R})}\big) &= \big(\{u \in H^{s'}(\mathbb{R}) : xu \in H^{s'}(\mathbb{R})\}, \|x\cdot\|_{H^{s'}(\mathbb{R})} + \|\cdot\|_{H^{s'}(\mathbb{R})}\big)\,.\end{aligned} \tag{1.2.6}$$

We obtain the following results for the hierarchies.

Theorem 3.1.2 (simplified)**.** Let $n \in \mathbb{N}$ with $n \geq 2$ and let $s, s' \in \mathbb{N}$ such that $s \geq \frac{5n-1}{2}$ and $\frac{s}{2} + \frac{n-1}{4} \leq s' \leq s - n$. Let $q_* : \mathbb{R} \to \mathbb{C}$ be a front for which $(q_*)_x$ and $|q_*|^2 - 1$ are Schwartz functions. The perturbative formulation of the n-th flow in the $\widetilde{\text{GP}}$ hierarchy ($\widetilde{\text{GP}}_n$), using the variable $p = q - q_*$, is globally well-posed in $H^s(\mathbb{R}) \cap H^{s',1}(\mathbb{R})$.

By stitching together the flow maps, we obtain as a corollary a well-posedness result that uses $N + 1$ time variables, i.e. for the N-truncated hierarchy in a direct sense.

Corollary 3.1.7 (simplified)**.** Let $N \in \mathbb{N}$ and let $s, s' \in \mathbb{N}$ such that $s \geq \frac{5N-1}{2}$ and $\frac{s}{2} + \frac{N-1}{4} \leq s' \leq s - N$. Let q_* be given as in the previous theorem and let $p_0 \in H^s(\mathbb{R}) \cap H^{s',1}(\mathbb{R})$. There exists a function $q = q(t_0, \ldots, t_N, x) : \mathbb{R}^{N+1} \times \mathbb{R} \to \mathbb{C}$ with initial data $q(0, \ldots, 0, x) = q_*(x) + p_0(x)$ which solves (NLS_n) in the sense of distributions for every $n \in \{0, \ldots, N\}$. It is unique in the sense of the previous theorem. The same result holds for the GP and $\widetilde{\text{GP}}$ hierarchies.

We mention as relevant literature the work [118] on the global well-posedness of the KdV hierarchy, the works [88, 5] on the well-posedness of the NLS (and also KdV) hierarchy with zero boundary data, as well as [4] for the dNLS hierarchy.

1.2.3. Main results of Chapter 4

Let $N \in \mathbb{N}$ and define $\mathbf{t} = (t_0, \ldots, t_N)$. In this chapter, we are interested in solutions q of the NLS hierarchy with the nonzero boundary condition

$$\lim_{x\to\pm\infty} q(\mathbf{t}, x) = q_\pm e^{i\varphi(\mathbf{t})} \text{ where } |q_\pm| = 1 \text{ and } \varphi(\mathbf{t}) \in \mathbb{R} \tag{NZBC}$$

that arise via the Ansatz

$$q(\mathbf{t},x) = \left(1 + \frac{1}{2}(w_+ - w_-)(\mathbf{t},x)\right) \exp\left(i\int_0^x (w_+ + w_-)(\mathbf{t},y)\,\mathrm{d}y\right) e^{i\varphi(\mathbf{t})}\,,$$

where the functions $w_\pm$ were defined in (1.2.2). We consider for a small parameter $\varepsilon > 0$ the rescaled variables

$$X = \sqrt{2}\varepsilon x \qquad\qquad \mathbf{T} = (T_n)_{0\le n\le N} = \left(\frac{\mathrm{d}T_n}{\mathrm{d}t_n} t_n\right)_{0\le n\le N}$$

where

$$\frac{\mathrm{d}T_{2n}}{\mathrm{d}t_{2n}} = 2(\sqrt{2}\varepsilon)^{2n-1} \qquad\qquad \frac{\mathrm{d}T_{2n-1}}{\mathrm{d}t_{2n-1}} = (\sqrt{2}\varepsilon)^{2n-1}\,, \tag{1.2.7}$$

and obtain with

$$W_\pm(\mathbf{T},X) = \frac{1}{\varepsilon^2} w_\pm(\mathbf{t},x)$$

the long-wave Ansatz

$$q(\mathbf{t},x) = \left(1 + \frac{\varepsilon^2}{2}(W_+ - W_-)(\mathbf{T},X)\right) \exp\left(i\frac{\varepsilon}{\sqrt{2}}\int_0^X (W_+ + W_-)(\mathbf{T},Y)\,\mathrm{d}Y\right) e^{i\varphi(\mathbf{t})}\,. \tag{1.2.8}$$

Here the functions $W_\pm$ are real-valued and decay at infinity. We write $W = (W_+, W_-)$ for the pair and define

$$\mathcal{H}^{\varepsilon\mathsf{GP}}(W) = \mathcal{H}^{\mathsf{GP}}(q) \qquad\qquad \mathcal{H}^{\varepsilon\widetilde{\mathsf{GP}}}(W) = \mathcal{H}^{\widetilde{\mathsf{GP}}}(q)\,,$$

which allows us to define the $\varepsilon\mathsf{GP}$ and $\varepsilon\widetilde{\mathsf{GP}}$ hierarchies as the hierarchies of Hamiltonian flows

$$i\partial_{T_n} W = \frac{\mathrm{d}t_n}{\mathrm{d}T_n}\frac{1}{2\varepsilon^2}\begin{pmatrix} -\partial_X A^{-1} - A^{-1}\partial_X & \partial_X A^{-1} - A^{-1}\partial_X \\ -\partial_X A^{-1} + A^{-1}\partial_X & \partial_X A^{-1} + A^{-1}\partial_X \end{pmatrix}\begin{pmatrix} \frac{\delta}{\delta W_+}\mathcal{H}_n^{\varepsilon\mathsf{GP}}(W) \\ \frac{\delta}{\delta W_-}\mathcal{H}_n^{\varepsilon\mathsf{GP}}(W)\end{pmatrix} \tag{$\varepsilon\mathsf{GP}_n$}$$

$$i\partial_{T_n} W = \frac{\mathrm{d}t_n}{\mathrm{d}T_n}\frac{1}{2\varepsilon^2}\begin{pmatrix} -\partial_X A^{-1} - A^{-1}\partial_X & \partial_X A^{-1} - A^{-1}\partial_X \\ -\partial_X A^{-1} + A^{-1}\partial_X & \partial_X A^{-1} + A^{-1}\partial_X \end{pmatrix}\begin{pmatrix} \frac{\delta}{\delta W_+}\mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}(W) \\ \frac{\delta}{\delta W_-}\mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}(W)\end{pmatrix} \tag{$\varepsilon\widetilde{\mathsf{GP}}_n$}$$

for $n \in \{0, \dots, N\}$. Here

$$A(\mathbf{T},X) = 1 + \frac{\varepsilon^2}{2}(W_+ - W_-)(\mathbf{T},X)$$

is the rescaled amplitude. Formally, it is indeed the case that if q solves (GP_n) (resp. ($\widetilde{\mathsf{GP}}_n$)), then W solves ($\varepsilon\mathsf{GP}_n$) (resp. ($\varepsilon\widetilde{\mathsf{GP}}_n$)).

In [27, 28] an approximation of the Hamiltonians $\mathcal{H}_0^{\widetilde{\mathsf{GP}}}, \dots, \mathcal{H}_8^{\widetilde{\mathsf{GP}}}$ by $\mathcal{E}_{-1}^{\mathsf{KdV}}, \dots, \mathcal{E}_3^{\mathsf{KdV}}$ was discovered, and an analogous result was speculated to hold for all $n \in \mathbb{N}$. We prove the following theorem.

Theorem 4.1.1 (simplified)**.** For all $n \in \mathbb{N}$ there exist polynomials $\mathbf{P}_n$, depending on ε, W_+, W_-, and derivatives thereof, such that

$$\mathcal{H}_{2n}^{\varepsilon\widetilde{\mathsf{GP}}}(W) \mp 2\mathcal{H}_{2n-1}^{\varepsilon\widetilde{\mathsf{GP}}}(W) = (\sqrt{2}\varepsilon)^{2n+1}\left(\mathcal{E}_{n-1}^{\mathsf{KdV}}(\pm W_\pm) + \frac{\varepsilon^2}{2}\int_{\mathbb{R}} \mathbf{P}_{2n} \mp \mathbf{P}_{2n-1}\,\mathrm{d}X\right).$$

In order to prove statements about these infinite families of Hamiltonians, we define them via the asymptotic expansions of the logarithms of the transmission coefficients (see (1.1.7) and (1.1.9) above). The above theorem is then a consequence of the following proposition. Here we define $a^{\varepsilon\mathsf{GP}}(\lambda,k;W) = a^{\mathsf{GP}}(\lambda,z;q)$ where $\lambda^2 = z^2 - 1 = 2\varepsilon^2k^2 - 1$. We define also the projections onto the even and odd parts Ev_λ and Od_λ by

$$\mathrm{Ev}_\lambda f(\lambda) = \frac{f(\lambda) + f(-\lambda)}{2} \qquad\qquad \mathrm{Od}_\lambda f(\lambda) = \frac{f(\lambda) - f(-\lambda)}{2}\,. \tag{1.2.9}$$

Proposition 4.1.3 (simplified)**.** Let $W = (W_+, W_-) \in \mathcal{S}(\mathbb{R}; \mathbb{R}^2)$, $\varepsilon \in (0,1)$ and let the parameters $\lambda, k \in \{b \in \mathbb{C} : \operatorname{Re} b, \operatorname{Im} b \geq 0\}$ be large so that the asymptotic expansions below exist. Then

$$\mathrm{Ev}_\lambda\big[\log a^{\varepsilon\mathsf{GP}}(\lambda, k; W)\big] \mp \frac{1}{\lambda}\,\mathrm{Od}_\lambda\big[\log a^{\varepsilon\mathsf{GP}}(\lambda, k; W)\big] \sim \log a^{\mathsf{KdV}}(k; (\pm W_\pm)) + \sum_{n=-1}^{\infty} \frac{O(\varepsilon^2)}{(2k)^{n+1}}$$

in the sense of asymptotic expansions as $2ik \to \infty$. Here $O(\varepsilon^2)$ must be understood symbolically in ε, W_+, W_-, and derivatives thereof.

Now assume in addition that $k \in i\,\mathbb{R}_+$ and $\kappa = -ik < \frac{1}{\sqrt{2}\varepsilon}$ satisfies

$$\max\left\{\frac{1}{\kappa}, \frac{1}{\kappa^3}\right\}\|W\|_{L^2}^2 \leq \frac{1}{3}\,.$$

Then

$$\mathrm{Ev}_\lambda\big[\log a^{\varepsilon\mathsf{GP}}(\lambda, k; W)\big] \mp \frac{1}{\lambda}\,\mathrm{Od}_\lambda\big[\log a^{\varepsilon\mathsf{GP}}(\lambda, k; W)\big] = \log a^{\mathsf{KdV}}(k; \pm W_\pm) + \mathbf{E}(\kappa; W)$$

with

$$|\mathbf{E}(\kappa; W)| \leq 8\varepsilon^2\kappa^{-1}\|W\|_{L^2}^2\,.$$

This can be understood as a result on the splitting of the Dirac operator in the nonrelativistic limit. We explain this in Remark 4.1.4 below. By exploiting an additional approximation result for the relevant symplectic/Poisson/Hamiltonian structures, we obtain the following formal approximation of the hierarchies. This uses the upper triangular coefficient matrix $\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}} \in \mathbb{R}^{(N+1)\times(N+1)}$, defined by

$$\mathbf{c}^{2n,2m}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}} = \binom{n-\frac{1}{2}}{n-m}\left(\frac{2}{\varepsilon^2}\right)^{n-m} \qquad \mathbf{c}^{2n,2m-1}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}} = 0$$

$$\mathbf{c}^{2n-1,2m}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}} = 0 \qquad \mathbf{c}^{2n-1,2m-1}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}} = \binom{n-\frac{1}{2}}{n-m}\left(\frac{2}{\varepsilon^2}\right)^{n-m}.$$

It represents a combination of the relations (1.1.10), (1.1.11), (1.1.12), and (1.2.7).

Theorem 4.1.8 (simplified)**.** Suppose that q is formally a solution of the N-truncated NLS hierarchy (NLS_n). If $W(\mathbf{T})$ is defined by (1.2.8), then $W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T})$ formally solves the N-truncated $\varepsilon\widetilde{\mathsf{GP}}$ hierarchy ($\varepsilon\widetilde{\mathsf{GP}}_n$) and

$$\frac{\mathrm{d}}{\mathrm{d}T_n} W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T}) = \begin{pmatrix} -(-1)^n\frac{1}{2}\partial_X \frac{\delta}{\delta W_+}\mathcal{E}^{\mathsf{KdV}}_{\lfloor\frac{n-1}{2}\rfloor}(W_+(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T})) \\ -\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)}\mathcal{E}^{\mathsf{KdV}}_{\lfloor\frac{n-1}{2}\rfloor}(-W_-(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T})) \end{pmatrix} + O(\varepsilon^2)\,.$$

We want to strengthen this to a quantitative approximation result. Here it is of great convenience that the global well-posedness of both the KdV hierarchy and the NLS hierarchy with nonzero boundary data has already been established. For the KdV hierarchy, this is a combination of the local well-posedness in high regularity Sobolev spaces from [104] and the conserved energies from [122]. For the NLS hierarchy with nonzero boundary we use our Corollary 3.1.7. Furthermore, we can use Theorem 2.1.6 to transfer estimates for $q \in q_* + H^s(\mathbb{R}) \subset X^s$ first to $(\rho, v) \in (1 + H^s(\mathbb{R})) \times H^{s-1}(\mathbb{R})$, and then to $W \in H^{s-1}(\mathbb{R})$.

Define now $M = \lfloor\frac{N-1}{2}\rfloor$ and $\mathbf{b}^+, \mathbf{b}^- \in \mathbb{R}^{(M+1)\times(N+1)}$ by

$$\mathbf{b}^+_{m,n} = -(-1)^n\mathbb{1}_{\{m=\lfloor\frac{n-1}{2}\rfloor\}} \qquad \mathbf{b}^-_{m,n} = \mathbb{1}_{\{m=\lfloor\frac{n-1}{2}\rfloor\}}\,.$$

The purpose of this matrix is to match the correct flows in the NLS hierarchy with the corresponding flows in the KdV hierarchy, as described by Theorem 4.1.8. Define $\mathbf{0} = (0, \ldots, 0)$.

Theorem 4.1.12 (simplified)**.** Let $N, s \in \mathbb{N}$ and $q_0 : \mathbb{R} \to \mathbb{C}$ satisfy the assumptions of Corollary 3.1.7 for some $s' \in \mathbb{N}$ and $q_*, p_0 : \mathbb{R} \to \mathbb{C}$. Assume in addition that $\mathcal{H}_2^{\mathsf{GP}}(q_0) < \frac{8}{3}$ so that $|q(\mathbf{t}, x)| > \delta > 0$, and hence we can define for some $\varepsilon > 0$ the function W according to the Ansatz (1.2.8). We have $W \in C(\mathbb{R}^{N+1}; H^{s-1}(\mathbb{R}))$, and if ε is sufficiently small then

$$\|W\|_{L^\infty_{\mathbf{T}} H^{h-2}_X} \leq C(h, \|W(\mathbf{0})\|_{H^h}) \|W(\mathbf{0})\|_{H^h} \qquad \forall\, h \in \{2, \dots, s-1\}\,.$$

Furthermore, $W(\mathbf{c}^{\mathsf{NLS} \mapsto \varepsilon \widetilde{\mathsf{GP}}} \mathbf{T}) \in C(\mathbb{R}^{N+1}; H^{s-1}(\mathbb{R}))$ solves ($\varepsilon\widetilde{\mathsf{GP}}_n$) in the sense of distributions.
Let $\pm U_\pm \in C(\mathbb{R}^{N+1} \times \mathbb{R}; H^{s-1}(\mathbb{R}))$ be the solutions to the N-truncated KdV hierarchy with initial data $U_\pm(\mathbf{0}) = W_\pm(\mathbf{0})$ which are known to exist and admit the bound

$$\|U(\mathbf{T})\|_{L^\infty_{\mathbf{T}} H^h_X} \leq C(h, \|U(\mathbf{0})\|_{H^{-1}}) \|U(\mathbf{0})\|_{H^h} \qquad \forall\, h \in \{0, \dots, s-1\}\,. \tag{1.2.10}$$

Define $[\mathbf{T}] = |T_2| + \cdots + |T_N|$. For all $h \in \{0, \dots, \lfloor \frac{s-N}{2} \rfloor\}$ there exists a constant

$$C = C(N, h, \delta, \|W\|_{L^\infty_{\mathbf{T}} H^{N+1+2h}_X}) > 0$$

such that

$$\left\| \begin{pmatrix} U_+(\mathbf{b}^+ \mathbf{T}) \\ U_-(\mathbf{b}^- \mathbf{T}) \end{pmatrix} - W(\mathbf{c}^{\mathsf{NLS} \mapsto \varepsilon \widetilde{\mathsf{GP}}} \mathbf{T}) \right\|_{H^h} \leq \varepsilon^2 C [\mathbf{T}]^{\frac{1}{2}} e^{C[\mathbf{T}]}\,. \tag{1.2.11}$$

This result represents a collection of approximations, and the timescales involved have several edge cases. In the Examples 4.1.13–4.1.20 in Chapter 4 this is discussed in detail. We present here an informal summary of the implications of this theorem for the flows of (NLS_n), (GP_n), and ($\widetilde{\mathsf{GP}}_n$) with $n \in \{0, \dots, 7\}$.

Example 4.1.13 ((NLS_0) = (GP_0) = ($\widetilde{\mathsf{GP}}_0$) = “phase rotation”)**.** The flows (NLS_0), (GP_0), and ($\widetilde{\mathsf{GP}}_0$) for the variable W are all constant, which matches exactly (KdV_{-1}).

Example 4.1.14 ((NLS_1) = (GP_1) = ($\widetilde{\mathsf{GP}}_1$) = “transport”)**.** The flows (NLS_1), (GP_1), and ($\widetilde{\mathsf{GP}}_1$) for the variable W are all transport flows, matching exactly the transport equation (KdV_0).

Example 4.1.15 ((NLS_2) = NLS and (GP_2) = ($\widetilde{\mathsf{GP}}_2$) = GP)**.** The flows (NLS_2), (GP_2), and ($\widetilde{\mathsf{GP}}_2$) are identical for the variable W and approximated by (KdV_0) on the timescale $|t_2| \lesssim \varepsilon^{-1}$, which is natural here. This result corresponds to [25, Theorem 2], i.e. it is an approximation of GP by two transport equations. In the introduction of Chapter 4 we elaborate on the following fact: it is known from [27, 28] that in the subsequent order KdV = (KdV_1) shows up as the approximating equation. It is natural to conjecture that this may be generalized, i.e. to characterize the effect of the terms $O(\varepsilon^2)$ up to $O(\varepsilon^4)$, even for the higher flows.

Example 4.1.16 ((NLS_3) = mKdV, (GP_3), and ($\widetilde{\mathsf{GP}}_3$))**.** The flows ($\mathsf{NLS}_3$), ($\mathsf{GP}_3$), and ($\widetilde{\mathsf{GP}}_3$) for the variable W are approximated by a combination of (KdV_0) and the KdV equation (KdV_1) on the timescale $|t_3| \lesssim \varepsilon^{-3}$, which is natural here.

Example 4.1.17 ((NLS_4) = “4-th order NLS”, (GP_4) = ($\widetilde{\mathsf{GP}}_4$) = “4-th order GP”)**.** The flows (NLS_4), (GP_4) for the variable W are approximated by a combination of (KdV_0) and (KdV_1) on the timescale $|t_4| \lesssim \varepsilon^{-1}$, which is only long enough to see (KdV_0). The flow ($\widetilde{\mathsf{GP}}_4$) for the variable W is approximated by a combination of (KdV_0) and (KdV_1) on the timescale $|t_4| \lesssim \varepsilon^{-3}$, which is natural here.

Example 4.1.18 ((NLS_5) = “5-th order mKdV”, (GP_5) = ($\widetilde{\mathsf{GP}}_5$))**.** The flows ($\mathsf{NLS}_5$), ($\mathsf{GP}_5$) for the variable W are approximated by a combination of (KdV_0), (KdV_1), and the fifth order KdV equation (KdV_2) on the timescale $|t_5| \lesssim \varepsilon^{-3}$, which is only long enough to see (KdV_0) and (KdV_1). The flow ($\widetilde{\mathsf{GP}}_5$) for the variable W is approximated by a combination of (KdV_0), (KdV_1), and (KdV_2) on the timescale $|t_5| \lesssim \varepsilon^{-5}$, which is natural here.

Example 4.1.19 ((NLS$_6$) = "6-th order NLS", (GP$_6$) = ($\widetilde{\mathsf{GP}}_6$) = "6-th order GP")**.** The flows (NLS$_6$), (GP$_6$) for the variable W are approximated by a combination of (KdV$_0$), (KdV$_1$), and (KdV$_2$) on the timescale $|t_6| \lesssim \varepsilon^{-1}$, which is only long enough to see (KdV$_0$). The flow ($\widetilde{\mathsf{GP}}_6$) for the variable W is approximated by a combination of (KdV$_0$), (KdV$_1$), and (KdV$_2$) on the timescale $|t_6| \lesssim \varepsilon^{-5}$, which is natural here.

Example 4.1.20 ((NLS$_7$) = "7-th order mKdV", (GP$_7$) = ($\widetilde{\mathsf{GP}}_7$))**.** The flows (NLS$_7$), (GP$_7$) for the variable W are approximated by a combination of (KdV$_0$), (KdV$_1$), (KdV$_2$), and the seventh order KdV equation (KdV$_3$) on the timescale $|t_7| \lesssim \varepsilon^{-3}$, which is only long enough to see (KdV$_0$) and (KdV$_1$). The flow ($\widetilde{\mathsf{GP}}_7$) for the variable W is approximated by a combination of (KdV$_0$), (KdV$_1$), (KdV$_2$), and (KdV$_3$) on the timescale $|t_7| \lesssim \varepsilon^{-7}$, which is natural here.

At this point we stop because all edge cases have been exhibited and the pattern continues as expected.

1.3. Outlook

We give for each chapter a brief outlook on possible future directions of research.

1. Regarding Chapter 2, there is the obvious question of lower regularity. Since the well-posedness results for (GP) from [119, 120] hold for $s \geq 0$, it would be useful to understand the Madelung transform with domain X^s for $s \in [0, \frac{1}{2}]$. Here the Sobolev embedding $H^s(\mathbb{R}) \longrightarrow L^\infty(\mathbb{R})$ is no longer available, so one needs to tackle the problem *with vacuum.*
2. Regarding Chapter 3, we would like to mention that much more recent techniques than the ones we have used are available, such as the Fourier restriction norm method (see e.g. [5, 4, 86, 87, 88]), as well as the method of commuting flows [38, 91, 114, 118]. This is, of course, much more technically involved. It would be of particular interest to find a natural adaptation of the method of commuting flows to the setting of NZBC. More generally, we hope that our initial well-posedness result inspires future work on the NLS hierarchy with nonzero boundary data. Lastly, we believe that with a lot of work Theorem 3.1.10 can be extended to describe the terms with two factors which have derivatives, and with even more work those with three factors with derivatives. If then one is able to make a guess for a formula for the complete structure of the NLS hierarchy, it may be possible to give an elegant proof using generating functions.
3. Regarding Chapter 4, we note that the main result Theorem 4.1.12 does not generalize the approximation result [27, Theorem 1], which states that (GP) is approximated by the transport equation (KdV$_0$) in the leading order, and (KdV) = (KdV$_1$) in the next-to-leading order. We reproduce only the leading order transport equation approximation. It is therefore plausible that every result in Chapter 4 might have a refined version, where the terms which we treat as $O(\varepsilon^2)$ are described up to $O(\varepsilon^4)$. In another direction, it would be of interest to explore further the implications of our result for the splitting of the one-dimensional Dirac operator in the nonrelativistic limit.

CHAPTER 2

Global-in-time Well-posedness of the One-dimensional Hydrodynamic Gross-Pitaevskii Equations without Vacuum

Abstract

We establish global-in-time well-posedness of the one-dimensional hydrodynamic Gross-Pitaevskii equations in the absence of vacuum in $(1 + H^s) \times H^{s-1}$ with $s \geq 1$. We achieve this by a reduction via the Madelung transform to the previous global-in-time well-posedness result for the Gross-Pitaevskii equation in [119, 120]. Our core result is a local bilipschitz equivalence of the relevant function spaces, which enables the transfer of results between the two equations.

2.1. Introduction

We consider in one dimension the Gross-Pitaevskii equation

$$i\partial_t q + \partial_{xx} q - 2(|q|^2 - 1)q = 0\,, \tag{GP}$$

where $q(t, x) : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{C}$ represents an unknown wave function, subject to the boundary condition at infinity $\lim_{|x|\to\infty} |q(t, x)| = 1$. In the absence of vacuum, meaning that

$$|q| > 0\,,$$

the Gross-Pitaevskii equation has a hydrodynamic formulation

$$\begin{cases} \partial_t \rho + 2\partial_x(\rho v) = 0\,, \\ \partial_t v + \partial_x(v^2) + 2\partial_x \rho = \partial_x\left(\partial_x\left(\frac{1}{2}\frac{\partial_x \rho}{\rho}\right) + \left(\frac{1}{2}\frac{\partial_x \rho}{\rho}\right)^2\right), \end{cases} \tag{hGP}$$

which we call the hydrodynamic Gross-Pitaevskii equations. Here the functions $\rho(t, x) : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}_+$ and $v(t, x) : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ may be understood as the unknown density and velocity of a quantum fluid. The system (hGP) belongs to the class of quantum hydrodynamical models, which may be used to model various physical phenomena such as Bose-Einstein condensation [73, 83], superfluidity [72, 127, 138] and quantum semiconductors [76]. We refer to [8, 36, 93] for more information on quantum hydrodynamical models and their relation to nonlinear Schrödinger equations.

The relation between (GP) and (hGP) is given by the Madelung transform

$$\mathcal{M}(q) = \left(|q|^2, \operatorname{Im}\left[\frac{\partial_x q}{q}\right]\right), \tag{2.1.1}$$

which formally transforms a solution q of (GP) into a solution $(\rho, v) = \mathcal{M}(q)$ of (hGP). Note that ρ and v are real-valued. One immediately sees that the Madelung transform $\mathcal{M}$ only makes sense when $|q| > 0$. In this case, we may recover q from its Madelung transform by the formula

$$q = \sqrt{\rho}e^{i\varphi}\,,$$

where φ is some spatial primitive of v, i.e.

$$\partial_x \varphi = v\,.$$

One furthermore sees that the inverse Madelung transform $(\rho, v) \longmapsto q$ is only defined up to multiplication with $\mathbb{S}^1$, i.e. a constant rotation in phase (see (2.1.13) below for more details). We refer the reader to [43] for a survey of the Madelung transform and the hydrodynamic Gross-Pitaevskii equations.

2.1.1. Related results for the Gross-Pitaevskii equation

E.P. Gross [85] and L.P. Pitaevskii [153] introduced the Gross-Pitaevskii equation as a model for a Bose-Einstein Condensate, a type of Boson gas at very low density and temperature. For rigorous justification of the model, we refer to the mean-field approximation established by L. Erdős, B. Schlein, and H. Yau [70], as well as references therein. As the Gross-Pitaevskii equation is a kind of defocusing cubic nonlinear Schrödinger equation, its well-posedness has been extensively studied. Due to the non-zero boundary condition, finite-energy solutions to (GP) can clearly not be in traditional function spaces that require global integrability, such as $L^p(\mathbb{R})$. For integers $k \geq 1$ and in any dimension $n \geq 1$, P.E. Zhidkov [191] established local-in-time well-posedness in the so-called Zhidkov space $Z^k(\mathbb{R}^n)$, which is the closure of $\{u \in C_b^k(\mathbb{R}^n) : \partial_x u \in H^{k-1}(\mathbb{R}^n)\}$ under the norm

$$\|u\|_{Z^k(\mathbb{R}^n)} = \|u\|_{L^\infty(\mathbb{R}^n)} + \sum_{1 \leq |\alpha| \leq k} \|\partial_x^\alpha u\|_{L^2(\mathbb{R}^n)}. \tag{2.1.2}$$

This led to a first global-in-time well-posedness result in one dimension in $Z^1(\mathbb{R})$, as the Ginzburg-Landau energy

$$E(q) = \frac{1}{2} \int_{\mathbb{R}^n} |\partial_x q|^2 + (|q|^2 - 1)^2 \, \mathrm{d}x \tag{2.1.3}$$

is conserved. The Gross-Pitaevskii equation (GP) can be interpreted as the Hamiltonian evolutionary equation associated with this energy. The well-posedness result in Zhidkov spaces was expanded to the cases $n = 2, 3$ by C. Gallo [74]. Global-in-time well-posedness in the energy space $\{q \in H^1_{\mathrm{loc}}(\mathbb{R}^n) : E(q) < \infty\}$ equipped with the metric

$$d_E(q,p) = \|q - p\|_{Z^1(\mathbb{R}^n) + H^1(\mathbb{R}^n)} + \||q|^2 - |p|^2\|_{L^2(\mathbb{R}^n)}$$

was obtained by P. Gérard [77, 78] for $n = 1, 2, 3$, and for $n = 4$ under smallness assumptions. Later R. Killip, T. Oh, O. Pocovnicu, and M. Vișan [113] established global-in-time well-posedness in the energy space for $n = 4$. More recently, the problem has been studied by P. Antonelli, L.E. Hientzsch, and P. Marcati [7] in $n = 2, 3$ for general nonlinearities f satisfying a Kato-type assumption. Local-in-time well-posedness is obtained in the energy space and extended globally in the defocusing case under some further assumptions. Regarding the case of non-finite energy, H. Pecher [151] established global-in-time well-posedness in three dimensions in $1 + H^s(\mathbb{R}^3)$ for $s \in (\frac{5}{6}, 1)$.

We are concerned with the case $n = 1$. For $s \in \mathbb{R}$, we associate with solutions of (GP) the energy functionals

$$E^s(q) = \frac{1}{2}\|\partial_x q\|^2_{H^{s-1}(\mathbb{R})} + \frac{1}{2}\||q|^2 - 1\|^2_{H^{s-1}(\mathbb{R})}. \tag{2.1.4}$$

Note that indeed $E^1 = E$. Our results are consequences of a pair of papers [119, 120] by H. Koch and X. Liao, where for $s \geq 0$ they proved the global-in-time well-posedness of (GP) in the complete metric space

$$X^s = \{q \in H^s_{\mathrm{loc}}(\mathbb{R}) : E^s(q) < \infty\} / \mathbb{S}^1, \tag{2.1.5}$$

equipped with the distance function

$$d^s(q,p) = \left(\int_{\mathbb{R}} \inf_{\lambda \in \mathbb{S}^1} \|\operatorname{sech}(y - \cdot)(\lambda q - p)\|^2_{H^s} \, \mathrm{d}y \right)^{\frac{1}{2}}. \tag{2.1.6}$$

We summarize several of their results, taken from [119, Theorem 1.2, 1.3, Lemma 6.1] and [120, Theorem 1.5], in the following theorem.

Theorem 2.1.1 (Global-in-time well-posedness of (GP) [119, 120])**.** *Let $s \geq 0$. The pair (X^s, d^s) is a complete metric space, and the energy functional $E^s : X^s \longrightarrow \mathbb{R}$ is continuous. There exists a constant $C_0 > 0$ such that $d^s(1,q) \leq C_0\sqrt{E^s(q)}$ for all $q \in X^s$.*

The Gross-Pitaevskii equation (GP) is globally-in-time well-posed in the metric space (X^s, d^s) in the following sense: For any initial data $q_0 \in X^s$ there exists a unique global-in-time solution $q \in C(\mathbb{R}; X^s)$ of (GP) (see Definition 2.3.2 below). For any $t \geq 0$ the Gross-Pitaevskii flow map $X^s \ni q_0 \mapsto q \in C([-t,t]; X^s)$ is continuous. There exists a constant $\tilde{C}_0(s, E^s(q_0))$ such that

$$\sup_{t \in \mathbb{R}} E^s(q(t)) \leq \tilde{C}_0(s, E^s(q_0))\, E^s(q_0), \tag{2.1.7}$$

and in the case $s \geq 1$ the energy $E(q(t))$, defined in (2.1.3), is conserved.

2.1.2. Related results for the hydrodynamic Gross-Pitaevskii equations

The question of equivalence between Schrödinger equations and quantum hydrodynamical equations is relevant for the validity of classical approaches to quantum mechanics such as de Broglie-Bohm Theory [31] and stochastic mechanics [147]. It became a topic of controversy when Wallstrom raised some objections [177, 178, 179]. We recommend [157] for a review of these issues. The difficulty arises from possible vacuum regions, which complicate the definition of the inverse Madelung transform. This can be resolved via an additional "Takabayasi's quantization condition" [29, 168], which requires the winding numbers of the velocity field on closed loops to be quantized. Note that this condition trivially holds in one dimension. Wallstrom also raised objections regarding uniqueness [179]. There are indeed non-uniqueness results for weak solutions to quantum hydrodynamical systems, which are related to a change in the number of non-vacuum connected components [139].

Nevertheless, P. Antonelli and P. Marcati [9, 10] constructed weak solutions vanishing at infinity with finite but arbitrarily large energy, meaning that vacuum may appear, in $n = 2, 3$ for a general quantum hydrodynamical system. They use a polar decomposition technique in order to define the velocity field in the vacuum regions. In collaboration with H. Zheng, they extended this to $n = 1$ via a purely hydrodynamical approach [12, 13]. An alternative approach to finite energy weak solutions is explored in [6, 11].

The well-posedness of the Euler-Korteweg system, a generalization of the compressible Euler equations which includes capillarity effects and contains (hGP) as a special case, was studied in higher dimensions by C. Audiard and B. Haspot [17, 18]. Similar to the approach we take is a paper by C. Audiard [16], in which global-in-time well-posedness of (hGP) under smallness assumptions is shown in certain spaces for $n \geq 2$ by applying the Madelung transform to solutions to (GP). While they used scattering results to bound the solution away from 0, we use a rather elementary argument that leads us to the aforementioned energy bound $E < \frac{4}{3}$.

A closely related paper is a work by H. Mohamad [145], which in [145, Proposition 1.1] states similar relations to our Theorem 2.1.6, and then uses a well-posedness result for (GP) to obtain a well-posedness result for (hGP) in $(1 + H^{k+1}) \times H^k$ with $k \in \mathbb{N}_{\geq 0}$ up to the appearance of vacuum. We discuss the similarities and differences in Remark 2.1.10.

2.1.3. Functional analytic framework

Our goal is to show a novel global-in-time well-posedness result for (hGP) with $(\rho, v) \in (1 + H^s) \times H^{s-1}$ where $s \geq 1$ and $n = 1$. We achieve this under the assumptions $s \geq 1$ and $E < \frac{4}{3}$ by passing the well-posedness result for (GP) in Theorem 2.1.1 through the Madelung transform (2.1.1). The first assumption $s \geq 1$ ensures sufficient regularity for the energy E to be defined, and for (hGP) to be interpretable in the sense of distributions. As an example, consider that $s \geq 1$ implies $v \in L^2(\mathbb{R})$, and so the problematic square of a distribution v^2 appearing in $(\text{hGP})_2$ does indeed exist. The second assumption $E < \frac{4}{3}$ can also be understood as a "regularity" assumption: solutions below the critical energy of $\frac{4}{3}$ can not have vacuum, that is points or intervals where $|q| = \sqrt{\rho} = 0$ (see Corollary 2.1.3 below). As a result, singularities are avoided in the hydrodynamic formulation. Due to conservation of energy, the absence of vacuum is guaranteed for all times. Note that this energy assumption is sharp in the sense that the black soliton solution $q(t, x) = \tanh(x)$ to (GP) has a zero $\tanh(0) = 0$, while also having energy $E(\tanh) = \frac{4}{3}$.

We collect now some results that ensure the absence of vacuum, given certain energy bounds. We start with the following lemma.

Lemma 2.1.2. *Consider the function* $\tilde{b} : [0,1] \longrightarrow [0, \frac{4}{3}]$ *defined by*

$$\tilde{b}(\delta) = \frac{4}{3} - 2\delta + \frac{2}{3}\delta^3.$$

This is a strictly decreasing bijection (see Fig. 1) whose inverse we denote by $\tilde{\delta}(b) : [0, \frac{4}{3}] \longrightarrow [0,1]$*. We have*

$$\tilde{b}(\delta) = \min\left\{E(q) : q \in H^1_{\mathrm{loc}}(\mathbb{R}),\ \inf_{x\in\mathbb{R}} |q(x)| \leq \delta\right\},$$

$$\tilde{\delta}(b) = \min\left\{\inf_{x\in\mathbb{R}} |q(x)| : q \in H^1_{\mathrm{loc}}(\mathbb{R}), E(q) \leq b\right\}.$$

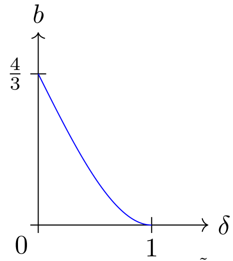


Fig. 1 Graph of $\tilde{b}$

This lemma is a stronger version of [26, Lemma 1]. The proof of a slightly more general Lemma 2.A.1 is given in the appendix. As a consequence of Lemma 2.1.2, the "energy gap" $\frac{4}{3} - E(q)$ yields an explicit lower bound for the distance of $|q|$ to zero. Due to conservation of the energy $E(q)$, we obtain the following corollary.

Corollary 2.1.3. *For any solution* $q \in C(\mathbb{R}; X^1)$ *of (GP) (see Definition 2.3.2), we have*

$$E(q_0) < b < \frac{4}{3} \Longrightarrow \inf_{(t,x)\in\mathbb{R}^2} |q(t,x)| > \tilde{\delta}(b) > 0\,. \tag{2.1.8}$$

We thus consider solutions q of (GP) in X^s, $s \geq 1$ with energy

$$E(q) < \frac{4}{3}\,, \tag{2.1.9}$$

recalling the definitions (2.1.3) - (2.1.6) of X^s and E. We look for solutions (ρ, v) of (hGP) in the function space

$$\mathcal{Y}^s = (1 + H^s(\mathbb{R};\mathbb{R})) \times H^{s-1}(\mathbb{R};\mathbb{R})\,, \tag{2.1.10}$$

equipped with the metric

$$\theta^s((\rho, v),(\eta, w)) = \|\rho - \eta\|_{H^s} + \|v - w\|_{H^{s-1}}\,. \tag{2.1.11}$$

We define the analogous energy

$$\mathcal{E}(\rho, v) = E(\mathcal{M}^{-1}(\rho, v)) = \frac{1}{2}\int_{\mathbb{R}} \frac{(\partial_x\rho)^2}{4\rho} + \rho v^2 + (\rho - 1)^2\,\mathrm{d}x. \tag{2.1.12}$$

Here the inverse Madelung transform is defined as

$$\mathcal{M}^{-1}(\rho, v)(x) = \left(\sqrt{\rho(x)}e^{i\varphi(x)}\right)\mathbb{S}^1 = \left\{\lambda\sqrt{\rho(x)}e^{i\varphi(x)} : \lambda \in \mathbb{S}^1\right\}, \tag{2.1.13}$$

where φ is any spatial primitive of v, i.e. $\partial_x\varphi = v$. Note that the energy E is indeed well-defined on equivalence classes under multiplication by $\mathbb{S}^1$, and furthermore that the space X^s consists of such equivalence classes, and is hence a suitable domain for the Madelung transform $\mathcal{M}$, given in (2.1.1).

In order to transform solutions of (GP) into solutions of (hGP) via the Madelung transform, we establish an equivalence between the relevant function spaces (X^s, d^s) and $(\mathcal{Y}^s, \theta^s)$. Specifically, we prove a local bilipschitz equivalence between the distance functions d^s and θ^s for all $s > \frac{1}{2}$. While our main result only holds for $s \geq 1$, our approach has the potential to be extended to the case $\frac{1}{2} < s < 1$ if one finds a way to make sense of $(\mathrm{hGP})_2$ in such a low regularity setting. For example, this may be possible via a local smoothing result, as in [114] (see Remark 2.1.12 below). When $\frac{1}{2} < s < 1$, the absence of vacuum can still be ensured by a smallness assumption of the form

$$E^\mu(q) < \varepsilon_0(\mu) << 1\,,$$

where $\mu > \frac{1}{2}$ (see (2.1.16) below). This smallness condition can also replace $E < \frac{4}{3}$ in the case $s \geq 1$, $\mu \leq s$. Specifically, we have the following Lemma 2.1.4 as a replacement for Lemma 2.1.2.

Lemma 2.1.4. *For $\delta \in [0,1]$ and $\mu > \frac{1}{2}$ define*

$$E^\mu_\delta = \inf\left\{E^\mu(q) : q \in H^\mu_{\mathrm{loc}}, \inf_{x\in\mathbb{R}} |q(x)| \le \delta\right\}.$$

Then $E^\mu_1 = 0$, the function $\delta \mapsto E^\mu_\delta$ is decreasing, and there exists a constant $\tilde{C}(\mu) > 0$ so that

$$E^\mu_\delta \ge \frac{(1-\delta)^2}{\tilde{C}(\mu)}. \tag{2.1.14}$$

This Lemma is also a special case of Lemma 2.A.1. By (2.1.7) there exists for any $\mu > \frac{1}{2}$ a constant $c(\mu) > 0$ such that

$$E^\mu(q_0) < \varepsilon \Longrightarrow \sup_{t\in\mathbb{R}} E^\mu(q(t)) < c(\mu)\,\varepsilon \tag{2.1.15}$$

for all $\varepsilon \in (0,1)$ and any solution $q \in C(\mathbb{R}; X^\mu)$ of (GP). Not attempting to obtain a sharp bound, we state the analogous of Corollary 2.1.3.

Corollary 2.1.5. *Let $\mu > \frac{1}{2}$ and define*

$$\varepsilon_0(\mu) = \max\left\{\frac{1}{2}, \frac{1}{4c(\mu)\tilde{C}(\mu)}\right\}. \tag{2.1.16}$$

For any solution $q \in C(\mathbb{R}; X^\mu)$ of (GP) (see Definition 2.3.2), we have

$$E^\mu(q_0) < \varepsilon < \varepsilon_0(\mu) \Longrightarrow \inf_{(t,x)\in\mathbb{R}^2} |q(t,x)| > 1 - \sqrt{\varepsilon}\sqrt{c(\mu)\tilde{C}(\mu)} > \frac{1}{2}. \tag{2.1.17}$$

Proof. We prove the contrapositive. Suppose $\inf_{(t,x)\in\mathbb{R}^2} |q(t,x)| \le \delta := 1 - \sqrt{\varepsilon}\sqrt{c(\mu)\tilde{C}(\mu)}$ and note that $\delta \in (0,1)$. Using the definition of $E^\mu_{\tilde{\delta}}$ and (2.1.14), this implies that for any $\tilde{\delta} > \delta$ there exists $t \in \mathbb{R}$ with

$$E^\mu(q(t)) \ge E^\mu_{\tilde{\delta}} \ge \frac{(1-\tilde{\delta})^2}{\tilde{C}(\mu)}.$$

In particular

$$\sup_{t\in\mathbb{R}} E^\mu(q(t)) \ge \frac{(1-\delta)^2}{\tilde{C}(\mu)} = c(\mu)\varepsilon,$$

so (2.1.15) implies $E^\mu(q_0) \ge \varepsilon$. □

As the energies E^μ still provide a lower bound for the distance of $|q|$ to zero, we can use the smallness assumption $E^\mu < \varepsilon_0(\mu)$ as a substitute for $E < \frac{4}{3}$. We define for $\mu > \frac{1}{2}$ the energies

$$\mathcal{E}^\mu(\rho, v) = E^\mu(\mathcal{M}^{-1}(\rho, v)). \tag{2.1.18}$$

2.1.4. Main results

For both the Gross-Pitaevskii equation (GP) and its hydrodynamic formulation (hGP), there are three key objects in our function framework: the energy, the space and the metric. We summarize the definitions given in §1.2 in the following diagram:

$$\boxed{\begin{array}{c}
E^s(q) = \frac{1}{2}\|\partial_x q\|^2_{H^{s-1}} + \frac{1}{2}\||q|^2 - 1\|^2_{H^{s-1}} \\
X^s = \{q \in H^s_{\mathrm{loc}}(\mathbb{R};\mathbb{C}) : E^s(q) < \infty\} / \mathbb{S}^1 \\
d^s(q,p) = \left(\int_{\mathbb{R}} \inf_{\lambda\in\mathbb{S}^1} \|\operatorname{sech}(y-\cdot)(\lambda q - p)\|^2_{H^s}\, dy\right)^{\frac{1}{2}} \\
\\
\mathcal{M}\Big\downarrow \quad \begin{array}{l} q = \sqrt{\rho}e^{i\varphi},\ \rho = |q|^2,\ v = \partial_x\varphi \\ p = \sqrt{\eta}e^{i\psi},\ \eta = |p|^2,\ w = \partial_x\psi \end{array} \quad \Big\uparrow\mathcal{M}^{-1} \\
\\
\mathcal{E}^s(\rho,v) = E^s(\mathcal{M}^{-1}(\rho,v)) \\
\mathcal{Y}^s = \{(\rho,v) \in (1 + H^s(\mathbb{R};\mathbb{R})) \times H^{s-1}(\mathbb{R};\mathbb{R})\} \\
\theta^s((\rho,v),(\eta,w)) = \|\rho - \eta\|_{H^s} + \|v - w\|_{H^{s-1}}
\end{array}} \tag{2.1.19}$$

Here the Madelung transform and its inverse

$$\mathcal{M}(q) = \left(|q|^2, \operatorname{Im}\left[\frac{\partial_x q}{q}\right]\right) \qquad\qquad \mathcal{M}^{-1}(\rho, v) = \left(\sqrt{\rho}e^{i\varphi}\right)\mathbb{S}^1, \quad \partial_x\varphi = v$$

are given in (2.1.1) and (2.1.13) respectively. Recall also the explicit forms of the energies E and $\mathcal{E}$ in the most important case $s = 1$:

$$E(q) = \frac{1}{2}\int_{\mathbb{R}} |\partial_x q|^2 + (|q|^2 - 1)^2 \,\mathrm{d}x \qquad\qquad \mathcal{E}(\rho, v) = \frac{1}{2}\int_{\mathbb{R}} \frac{(\partial_x\rho)^2}{4\rho} + \rho v^2 + (\rho - 1)^2 \,\mathrm{d}x\,.$$

Our first main result is the following theorem, which is central to our strategy as it establishes a local bilipschitz equivalence between the metrics d^s and θ^s. We require $s > \frac{1}{2}$ to use L^∞ embeddings and certain product estimates.

Theorem 2.1.6 (Local bilipschitz equivalence of d^s and θ^s)**.** *Let $s > \frac{1}{2}$ and $r, \delta > 0$. Consider measurable functions $\rho, \eta, \varphi, \psi : \mathbb{R} \longrightarrow \mathbb{R}$ so that $q, p \in \mathcal{S}'(\mathbb{R}) \cap H^s_{\mathrm{loc}}(\mathbb{R})$ and $|q|, |p| > \delta$, where $q = \sqrt{\rho}e^{i\varphi}$ and $p = \sqrt{\eta}e^{i\psi}$. There exist constants $C_1(s, \delta, r), C_2(s, r) > 0$ so that the following hold:*

(i) *If $d^s(1, q), d^s(1, p) < r$, then*

$$\theta^s((\rho, \partial_x\varphi), (\eta, \partial_x\psi)) \leq C_1(s, \delta, r)\, d^s(q, p)\,.$$

(ii) *If $\theta^s((1, 0), (\rho, \partial_x\varphi)), \theta^s((1, 0), (\eta, \partial_x\psi)) < r$, then*

$$d^s(q, p) \leq C_2(s, r)\, \theta^s((\rho, \partial_x\varphi), (\eta, \partial_x\psi))\,.$$

Corollary 2.1.7. *Let $s \geq 1$, $\frac{1}{2} < \mu < 1$. For all $b < \frac{4}{3}$ and $\varepsilon < \varepsilon_0(\mu)$ the maps*

$$\left(\left\{q \in X^s : E(q) < b\right\}, d^s\right) \xrightarrow{\mathcal{M}} \left(\left\{(\rho, v) \in \mathcal{Y}^s : \mathcal{E}(\rho, v) < b\right\}, \theta^s\right)$$

and

$$\left(\left\{q \in X^s : E^\mu(q) < \varepsilon\right\}, d^s\right) \xrightarrow{\mathcal{M}} \left(\left\{(\rho, v) \in \mathcal{Y}^s : \mathcal{E}^\mu(\rho, v) < \varepsilon\right\}, \theta^s\right)$$

are bilipschitz equivalences. Recall that a bilipschitz equivalence is a map which is bijective, Lipschitz continuous, and has a Lipschitz continuous inverse. Here $\varepsilon_0(\mu)$ is a constant defined in (2.1.16).

Our second main result is the global-in-time well-posedness of the hydrodynamic Gross-Pitaevskii equations.

Definition 2.1.8 (Solution to (hGP))**.** Let $s \geq 1$ and $0 \in I \subset \mathbb{R}$ be an open time interval or the real line, and let $(\rho_0, v_0) \in \mathcal{Y}^s$ with $\rho_0 > 0$. A solution to (hGP) with initial data (ρ_0, v_0) is a pair $(\rho, v) \in C(I; \mathcal{Y}^s)$ with $\rho > 0$ which solves (hGP) in the sense of distributions and fulfills $(\rho, v)(0) = (\rho_0, v_0)$.

Theorem 2.1.9 (Global-in-time well-posedness of (hGP) for $s \geq 1$)**.** *Let $s \geq 1$. The hydrodynamic Gross-Pitaevskii equations (hGP) are globally-in-time well-posed in the metric space $(\mathcal{Y}^s, \theta^s)$ for initial data $(\rho_0, v_0) \in \mathcal{Y}^s$ with $\mathcal{E}(\rho_0, v_0) < \frac{4}{3}$ in the following sense:*

There exists a solution $(\rho, v) \in C_b(\mathbb{R}; \mathcal{Y}^s)$ to (hGP) (see Definition 2.1.8). It is the unique solution that fulfills

$$\mathcal{E}(\rho(t), v(t)) = \mathcal{E}(\rho_0, v_0) < \frac{4}{3} \tag{2.1.20}$$

for all $t \in \mathbb{R}$. For any $T \geq 0$ the solution map

$$\begin{aligned} \left\{(\rho_0, v_0) \in \mathcal{Y}^s : \mathcal{E}(\rho_0, v_0) < \frac{4}{3}\right\} &\longrightarrow C_b([-T, T]; \mathcal{Y}^s) \\ (\rho_0, v_0) &\longmapsto (\rho, v) \end{aligned}$$

is continuous.

For all $\frac{1}{2} < \mu < 1$ there exist constants $c(\mu), \varepsilon_0(\mu) > 0$, defined in (2.1.15) and (2.1.16), so that if we replace the assumption $\mathcal{E}(\rho_0, v_0) < \frac{4}{3}$ by $\mathcal{E}^\mu(\rho_0, v_0) < \varepsilon < \varepsilon_0(\mu)$, then the above statement holds with (2.1.20) replaced by $\mathcal{E}^\mu(\rho(t), v(t)) < c(\mu)\,\varepsilon$.

Remark 2.1.10. *In [145, Theorem 1.2] the well-posedness of (hGP) until the appearance of vacuum is shown in* $(1 + H^k) \times H^{k-1}$ *for* $k \in \mathbb{N}_{\geq 1}$. *Furthermore, continuity properties for the Madelung transform between* $(\mathcal{Y}^k, \theta^k)$ *and the space*

$$\tilde{E}^k = \{q \in L^\infty(\mathbb{R}; \mathbb{C}) : 1 - |q|^2 \in L^2(\mathbb{R}), \partial_x q \in H^{k-1}(\mathbb{R})\}$$

are established. Specifically, a strong metric

$$\tilde{d}^k(q,p) = \|q-p\|_{L^\infty(\mathbb{R})} + \||q|^2 - |p|^2\|_{L^2(\mathbb{R})} + \|\partial_x q - \partial_x p\|_{H^{k-1}(\mathbb{R})}$$

and a weak metric

$$\tilde{d}^k_{\mathrm{loc}}(q,p) = \|q-p\|_{L^\infty([-1,1])} + \||q|^2 - |p|^2\|_{L^2(\mathbb{R})} + \|\partial_x q - \partial_x p\|_{H^{k-1}(\mathbb{R})}$$

are considered, and the following is shown [145, Proposition 1.1]:

$$\begin{tikzcd}
(\tilde{E}^k \cap \{|q|>0\}, \tilde{d}^k) \arrow[dr, "\textit{loc. Lipschitz}"] & & (\tilde{E}^k \cap \{|q|>0\}, \tilde{d}^k) \\
& (\mathcal{Y}^k \cap \{\rho>0\}, \theta^k) \arrow[ur, "\textit{not cont.}"] \arrow[dr, "\textit{cont.}"'] & \\
(\tilde{E}^k \cap \{|q|>0\}, \tilde{d}^k_{\mathrm{loc}}) \arrow[ur, "\textit{not loc. Lipschitz}"'] & & (\tilde{E}^k \cap \{|q|>0\}, \tilde{d}^k_{\mathrm{loc}})
\end{tikzcd}$$

There are several ways in which Theorem 2.1.6 improves upon [145, Theorem 1.2]. The first is that our result covers the fractional cases as well. The second is that we obtain a full local bilipschitz equivalence

$$(X^s \cap \{|q|>0\}, d^s) \xrightarrow{\textit{loc. Lipschitz}} (\mathcal{Y}^s \cap \{\rho>0\}, \theta^s) \xrightarrow{\textit{loc. Lipschitz}} (X^s \cap \{|q|>0\}, d^s)$$

and do so down to $s > \frac{1}{2}$, *while they have to work with two different metrics on the side of (GP) for the two directions of estimates. Furthermore, they state well-posedness up to the appearance of vacuum, while we use energy bounds to ensure this absence of vacuum for all times. Lastly, we find solutions which are uniformly bounded.*

Remark 2.1.11. *Previously, P.E. Zhidkov [190, Theorem III.3.1] studied the stability of solutions in the Zhidkov space* $Z^1(\mathbb{R})$ *(see (2.1.2)) near space-homogeneous solutions* Φ, *such as the constant solution* $\Phi = 1$, *with respect to the distance* θ^1. *For the case* $s = 1$ *he derived similar estimates as above under smallness assumptions, although he did not formulate a well-posedness result. Curiously, in [190, Cor. III.3.5] he proved furthermore that for any ball* $B \subset \mathbb{R}$, *if the initial* θ^1*-distance between the perturbed and the space-homogeneous solution is small, then for all times also the distance*

$$\inf_{\lambda \in \mathbb{S}^1} \|\lambda q - \Phi\|_{W^{1,2}(B)}$$

is small. This can be interpreted as a weaker form of the estimate $d^1 \lesssim \theta^1$ *we derive (see Lemma 2.2.6 and Remark 2.2.11 below).*

Remark 2.1.12. *As both Theorem 2.1.1 and Theorem 2.1.6 work for all* $s > \frac{1}{2}$, *it may be possible to extend Theorem 2.1.9 to the case* $\frac{1}{2} < s < 1$. *The problem is that for* $v \in H^{s-1} \not\subseteq L^2$ *the product of distributions* $v^2 = v \cdot v$ *is not necessarily defined. Nevertheless, it may be possible to find global distributional solutions. For example, in the paper [114] by R. Killip and M. Vișan global-in-time well-posedness of the KdV equation in* H^{-1} *is first shown in the sense that the solution map* $\mathbb{R} \times \mathcal{S} \longrightarrow \mathcal{S}$ *extends to a continuous mapping* $\mathbb{R} \times H^{-1} \longrightarrow H^{-1}$, *and some other conditions are fulfilled. In our case, it is similarly true that for any* $T \geq 0$ *the solution map*

$$\Big\{(\rho_0, v_0) \in \mathcal{Y}^1 : \mathcal{E}^s(\rho_0, v_0) < \varepsilon_0(s)\Big\} \longrightarrow C_b([-T,T]; \mathcal{Y}^1)$$

has a unique continuous extension to a map

$$\Big\{(\rho_0, v_0) \in \mathcal{Y}^s : \mathcal{E}^s(\rho_0, v_0) < \varepsilon_0(s)\Big\} \longrightarrow C_b([-T,T]; \mathcal{Y}^s)\,.$$

This extension is given by the conjugation of the corresponding solution map for (GP) at regularity s with the Madelung transform. R. Killip and M. Vișan then furthermore show a local smoothing result, which implies that the solution map produces functions in $L^2_{\mathrm{loc},t,x}$. As a result, the equation is indeed solved in the sense of distributions. We do not know if such a local smoothing result holds in our case.

Organization of the paper. In §2 we prove Theorem 2.1.6, the local bilipschitz equivalence of (X^s, d^s) and $(\mathcal{Y}^s, \theta^s)$. In §3 we prove Theorem 2.1.9, the global-in-time well-posedness of the hydrodynamic Gross-Pitaevskii equations.

Acknowledgements. I would like to thank Sarah Hofbauer for her help with reviewing the literature, and Phillipe Gravejat for directing my attention to the highly relevant paper [145] by H. Mohamad. I would also like to thank the anonymous reviewer for his detailed and valuable comments. I am especially thankful to my supervisor Xian Liao for proposing this problem and strategy, and for her patience during many hours of discussion.

Statements and Declarations.
Funding: This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173.
Version: This version of the article has been accepted for publication, after peer review but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1007/s00033-023-02089-4

2.2. Local bilipschitz equivalence of (X^s, d^s) and $(\mathcal{Y}^s, \theta^s)$

The goal of this section is to prove Theorem 2.1.6. In §2.1, we introduce the necessary notations, definitions, and basic results required for the rest of the paper. We split the proof of the two statements (i) and (ii) of Theorem 2.1.6 into §2.2 and §2.3.

2.2.1. Notations and preliminaries

We use the notations $\mathbb{R}_{\geq} = \{r \in \mathbb{R} : r \geq 0\}$ and $\mathbb{R}_+ = \{r \in \mathbb{R} : r > 0\}$. We write C or $C(...)$ for various constants with possible dependence on other quantities. These may change from one line to the next. We denote by $\mathcal{D}' = \mathcal{D}'(\mathbb{R}) = \mathcal{D}'(\mathbb{R};\mathbb{C})$ the space of distributions and by $\mathcal{S}' = \mathcal{S}'(\mathbb{R}) = \mathcal{S}'(\mathbb{R};\mathbb{C})$ the space of tempered distributions. In general, if for a family of function spaces, such as the L^p-spaces, we write just "L^p", then we mean $L^p(\mathbb{R};\mathbb{C})$.

We write $\widehat{f}$ for the Fourier transform

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ix\xi} f(x)\,\mathrm{d}x$$

for a Schwartz function $f \in \mathcal{S}$ and extend the definition as usual to the tempered distributions $f \in \mathcal{S}'$. Let $s \in \mathbb{R}$. We define the Sobolev space

$$H^s = H^s(\mathbb{R}) = H^s(\mathbb{R};\mathbb{C}) = \{f \in \mathcal{S}'(\mathbb{R};\mathbb{C}) : \|f\|_{H^s} < \infty\}$$

with norm

$$\|f\|_{H^s} = \|f\|_{H^s(\mathbb{R})} = \left\|\langle\xi\rangle^s \widehat{f}\right\|_{L^2(\mathbb{R})}.$$

Here $\langle\xi\rangle = \sqrt{1+|\xi|^2}$. We also define the quasinorm of the homogeneous Sobolev space

$$\|f\|_{\dot{H}^s} = \|f\|_{\dot{H}^s(\mathbb{R})} = \left\||\xi|^s \widehat{f}\right\|_{L^2(\mathbb{R})}.$$

Let $p \in [1,\infty)$. For $s \geq 0$, let $\alpha \in [0,1)$ and $m \in \mathbb{Z}$ so that $s = m + \alpha$. Let $B \subset \mathbb{R}$ be a non-empty open interval. We define the Sobolev-Slobodeckij space

$$W^{s,p}(B) = \{f \in \mathcal{D}'(B) : \|f\|_{W^{s,p}(B)} < \infty\}$$

with norm

$$\|f\|_{W^{s,p}(B)} = \left(\sum_{k=0}^{m} \|\partial^k f\|^p_{L^p(B)} + \int_B \int_B \frac{|\partial^m f(x) - \partial^m f(y)|^p}{|x-y|^{1+\alpha p}} \,\mathrm{d}x \,\mathrm{d}y\right)^{\frac{1}{p}}.$$

We define $W_0^{s,p}(B) = \overline{\mathcal{D}(B)}^{W^{s,p}(B)}$. For $s < 0$ we define $W^{s,p'}(B) = (W_0^{-s,p}(B))^*$, where $\frac{1}{p} + \frac{1}{p'} = 1$. We refer the reader to the book [140] by W. McLean for a comprehensive exposition. For the convenience of the reader, let us recall some well-known results on Sobolev spaces that may be used without mention.

2.2.1.1. FRACTIONAL SOBOLEV SPACES ON B AND $\mathbb{R}$

The only bounded domains we use are balls B, and on those we use the Sobolev-Slobodeckij spaces $W^{s,2}(B)$. On the whole real line $\mathbb{R}$ we use $H^s = H^s(\mathbb{R})$.

Lemma 2.2.1 ($W^{s,2}(B)$ and $H^s(\mathbb{R})$)**.** *Let $s \in \mathbb{R}$ and $R > 0$. Let $\tilde{B} \subset B \subseteq \mathbb{R}$ be concentric balls of radius $\frac{R}{2}$ and R. Set $B_k = B + kR$ and $\tilde{B}_k = \tilde{B} + kR$ for $k \in \mathbb{Z}$.*

(i) *There exists a natural isomorphism $H^{-s} \cong (H^s)^*$ (see [140, p. 76]).*

(ii) *$H^s = W^{s,2}(\mathbb{R})$ and*

$$\|f\|_{W^{s,2}(B)} \le C_1 \min\{\|F\|_{H^s} : F\big|_B = f\} \le C_2 \|f\|_{W^{s,2}(B)} .$$

(see [140, p. 77, (3.23) + Theorem 3.18, 3.19]).

(iii) *For $s \ge 0$, there exists a bounded linear extension operator $E : W^{s,2}(B) \longrightarrow H^s$ with $Ef\big|_B = f$ (see [140, Theorem A.4]).*

(iv) *For $s \ge 0$, there exists a constant $C(s, R)$ so that*

$$\sum_{k\in\mathbb{Z}} \|f\|^2_{W^{s,2}(\tilde{B}_k)} \le \|f\|^2_{H^s} \le C(s,R) \sum_{k\in\mathbb{Z}} \|f\|^2_{W^{s,2}(B_k)} \le 4C(s,R)\|f\|^2_{H^s}$$

and

$$\|f\|^2_{H^{-s}} \le C(s,R) \sum_{k\in\mathbb{Z}} \|f\|^2_{W^{-s,2}(B_k)}.$$

(v) *If $s > \frac{1}{2}$, then $\|fg\|_{H^s} \le C(s)\|f\|_{H^s}\|g\|_{H^s}$ (see [20, Cor. 2.87]). By use of the extension operator, we also have $\|fg\|_{W^{s,2}(B)} \le C(s)\|f\|_{W^{s,2}(B)}\|g\|_{W^{s,2}(B)}$.*

Proof. We only have to prove (iv). We start with the first inequality in the sequence. The case $s = 0$ is trivial, so we assume $s > 0$. Here the statement is trivial for the terms with integer regularity, and for the fractional terms we estimate

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} \int_{\tilde{B}_k} \int_{\tilde{B}_k} \frac{|\partial^m f(x) - \partial^m f(y)|^2}{|x-y|^{1+2\alpha}} \,\mathrm{d}x \,\mathrm{d}y &\le \sum_{j,k\in\mathbb{Z}} \int_{\tilde{B}_j} \int_{\tilde{B}_k} \frac{|\partial^m f(x) - \partial^m f(y)|^2}{|x-y|^{1+2\alpha}} \,\mathrm{d}x \,\mathrm{d}y \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|\partial^m f(x) - \partial^m f(y)|^2}{|x-y|^{1+2\alpha}} \,\mathrm{d}x \,\mathrm{d}y\,.
\end{aligned}$$

The third inequality in the sequence follows trivially. We show the second inequality first for $s > 0$. For the terms with integer regularity the statement is again trivial, so we focus on the fractional part[1]. Here

$$\begin{aligned}
\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|\partial^m f(x) - \partial^m f(y)|^2}{|x-y|^{1+2\alpha}} \,\mathrm{d}x \,\mathrm{d}y &= \sum_{j,k\in\mathbb{Z}} \int_{\tilde{B}_k} \int_{\tilde{B}_j} \frac{|\partial^m f(x) - \partial^m f(y)|^2}{|x-y|^{1+2\alpha}} \,\mathrm{d}x \,\mathrm{d}y \\
&\le \sum_{k\in\mathbb{Z}} \int_{B_{k-1}\cup B_{k+1}} \int_{B_{k-1}\cup B_{k+1}} \frac{|\partial^m f(x) - \partial^m f(y)|^2}{|x-y|^{1+2\alpha}} \,\mathrm{d}x \,\mathrm{d}y \\
&\quad + \sum_{k\in\mathbb{Z}} \int_{\tilde{B}_k} 2(|\partial^m f(x)|^2 + |\partial^m f(x)|^2) \sum_{\substack{j\in\mathbb{Z} \\ |k-j|\ge 2}} \int_{\tilde{B}_j} \frac{1}{|x-y|^{1+2\alpha}} \,\mathrm{d}x \,\mathrm{d}y
\end{aligned}$$

[1]The proof in the published version of this paper contains a mistake, which is corrected in this version.

$$= (I) + (II)\,.$$

Clearly

$$(II) \leq C(s,R) \sum_{k\in\mathbb{Z}} \|\partial^m f\|^2_{L^2(B_k)}$$

and

$$(I) \leq C(s,R) \sum_{k\in\mathbb{Z}} \|\partial^m f\|^2_{W^{\alpha,2}(B_{k-1}\cup B_{k+1})} \leq C(s,R) \sum_{k\in\mathbb{Z}} \|f\|^2_{W^{s,2}(B_k)}\,.$$

The final estimate for (I) uses a partition of unity with the covering $\bigcup_{j=-2}^{2} B_{k+j}$ of $B_{k-1}\cup B_{k+1}$. Now we show the inequality for negative regularities. Let $f \in H^{-s}(\mathbb{R})$, $g \in H^s(\mathbb{R})$ and denote by $\langle f, g\rangle$ the dual pairing. We decompose $g = \sum_{k\in\mathbb{Z}} \eta_k g$, where η_k is a smooth partition of unity with $\operatorname{supp}\eta_k \subset B_k$, $\sum_{k\in\mathbb{Z}} \eta_k = 1$ and $\eta_k(x) = \eta_k(x+kR)$. Since $\eta_k g \in W^{s,2}_0(B_k)$ with $\|\eta_k g\|_{W^{s,2}(B_k)} \leq C(s,R,\eta_0)\|g\|_{W^{s,2}(B_k)}$, we can estimate

$$\begin{aligned}|\langle f,g\rangle| &\leq \sum_{k\in\mathbb{Z}} |\langle f, \eta_k g\rangle| \leq \sum_{k\in\mathbb{Z}} \|f\|_{W^{-s,2}(B_k)}\|\eta_k g\|_{W^{s,2}(B_k)} \leq \left(\sum_{k\in\mathbb{Z}} \|f\|^2_{W^{-s,2}(B_k)}\right)^{\frac12}\left(\sum_{k\in\mathbb{Z}} \|\eta_k g\|^2_{W^{s,2}(B_k)}\right)^{\frac12}\\ &\leq C(s,R,\eta_0)\left(\sum_{k\in\mathbb{Z}} \|f\|^2_{W^{-s,2}(B_k)}\right)^{\frac12} \|g\|_{H^s(\mathbb{R})}\,.\end{aligned}$$

□

2.2.1.2. Estimates in H^s

The following lemma states two crucial estimates. Such kinds of product estimates are well-known in the literature, see for example [58, Proposition 2.7].

Lemma 2.2.2. *Let $s > \frac12$ and $f, g \in \mathcal{S}'$. There exists a constant $C(s)$ so that*

$$\|fg\|_{H^s} \leq C(s)\|g\|_{H^s}\big(\|f\|_{L^\infty} + \|f'\|_{H^{s-1}}\big) \tag{2.2.1}$$

and

$$\|fg\|_{H^{s-1}} \leq C(s)\|g\|_{H^{s-1}}\big(\|f\|_{L^\infty} + \|f'\|_{H^{s-1}}\big)\,. \tag{2.2.2}$$

Proof. See Appendix B. □

In this section we often write f' for the spatial derivative $\partial_x f$. Recall the definitions (2.1.19). For notational convenience, we sometimes prefer to use the variables

$$A = \sqrt{\rho} \quad \text{and} \quad B = \sqrt{\eta}\,.$$

These variables are equivalent for the sake of our estimates, by which we mean specifically Lemma 2.2.4. In order to prove this, we state two estimates regarding the action of a smooth function on Sobolev spaces. They are a direct consequence of some results in [20].

Lemma 2.2.3 ([20, Theorem 2.87, Corollary 2.91])**.** *Let $s > \frac12$ and $F \in C^\infty(\mathbb{R};\mathbb{R})$ with $F'(0) = F(0) = 0$. Let $u, v \in H^s(\mathbb{R};\mathbb{R}) \cap L^\infty(\mathbb{R};\mathbb{R})$. We have the estimates*

$$\|F\circ u\|_{H^s} \leq C(s, F', \|u\|_{L^\infty})\|u\|_{H^s} \tag{2.2.3}$$

and

$$\|F\circ u - F\circ v\|_{H^s} \leq C(s, F'', \|u\|_{H^s}, \|v\|_{H^s})\|u - v\|_{H^s}\,. \tag{2.2.4}$$

An analysis of the proof in [20] reveals that, more precisely, the constants depend on $\|F'\|_{C^{\lceil s\rceil+1}(B_{\|u\|_{L^\infty}})}$ and $\|F''\|_{C^{\lceil s\rceil+1}(B_{\|u\|_{L^\infty}})}$ respectively.

Lemma 2.2.4. *Let $s > \frac{1}{2}$ and $\rho, \eta \in \mathcal{S}'(\mathbb{R};\mathbb{R}) \cap H^s_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ with $\rho, \eta > 0$. Define $A = \sqrt{\rho}$ and $B = \sqrt{\eta}$. We have the estimates*

$$\|\rho - \eta\|_{H^s} \le C_1\big(s, \|A-1\|_{H^s}, \|B-1\|_{H^s}\big)\, \|A-B\|_{H^s}$$

and

$$\|A - B\|_{H^s} \le C_2\big(s, \|\rho-1\|_{H^s}, \|\eta-1\|_{H^s}\big)\, \|\rho-\eta\|_{H^s}\,.$$

Proof. We apply Lemma 2.2.3 with $F(u) = u^2$ and obtain

$$\begin{aligned}\|A^2 - B^2\|_{H^s} &\le \|(A-1)^2 - (B-1)^2\|_{H^s} + 2\|A-B\|_{H^s}\\ &\le C\big(s, \|A-1\|_{H^s}, \|B-1\|_{H^s}\big)\, \|A-B\|_{H^s}\,.\end{aligned}$$

Similarly with any function $F \in C^\infty(\mathbb{R};\mathbb{R})$ that fulfills $F(u) = \sqrt{u+1} - \frac{1}{2}u - 1$ for $x \ge 0$, we obtain

$$\begin{aligned}\|\sqrt{\rho} - \sqrt{\eta}\|_{H^s} \le{}& \left\| \left(\sqrt{(\rho-1)+1} - \frac{1}{2}(\rho-1)\right) - \left(\sqrt{(\eta-1)+1} - \frac{1}{2}(\eta-1)\right)\right\|_{H^s}\\ &+ \frac{1}{2}\|\rho-\eta\|_{H^s}\\ \le{}& C\big(s, \|\rho-1\|_{H^s}, \|\eta-1\|_{H^s}\big)\, \|\rho-\eta\|_{H^s}\,.\end{aligned}$$

□

The following lemma is also a consequence of Lemma 2.2.3 and will be used frequently in the subsequent section.

Lemma 2.2.5. *Let $s, \delta, R > 0$. There exists $C(s,\delta,R) > 0$ such that for any ball $B_0 \subset \mathbb{R}$ of radius R and all $u \in W^{s,2}(B_0)$ with $|u| > \delta > 0$ we have*

$$\|u^{-1}\|_{W^{s,2}(B_0)} \le C(s,\delta)\|u\|_{W^{s,2}(B_0)}\,.$$

Proof. This follows by applying Lemma 2.2.3 with any function $F \in C^\infty(\mathbb{R};\mathbb{R})$ so that $F(0) = F'(0) = 0$ and $F(x) = x^{-1}$ for $|x| > \frac{\delta}{2}$, and using the existence of an extension operator from Lemma 2.2.1 (*iii*). Note that Lemma 2.2.3 requires real-valued functions, so we apply it to the real and imaginary parts of u^{-1} separately. □

2.2.2. PROOF OF THEOREM 2.1.6 (I)

Recall the definitions (2.1.19), in particular $q = \sqrt{\rho}e^{i\varphi}$ and $p = \sqrt{\eta}e^{i\psi}$, as well as $A = \sqrt{\rho}$ and $B = \sqrt{\eta}$. We assume $s > \frac{1}{2}$, $d^s(1,q), d^s(1,p) < r$ and $|q|, |p| > \delta > 0$. We have to prove that

$$\theta^s((\rho, \partial_x\varphi), (\eta, \partial_x\psi)) \le C(s,\delta,r)\, d^s(q,p)\,.$$

We do this by showing an estimate of the form

$$\theta^s \lesssim \sum_{k\in\mathbb{Z}} d^s_*\big|_{B_k} \lesssim d^s\,. \tag{2.2.5}$$

Let us elaborate on the quantity in the middle before we start the proof. Given a ball $B \subset \mathbb{R}$, we define for convenience the following notations:

$$d^s_*\big|_B(q,p) = \inf_{\lambda\in\mathbb{S}^1} \|\lambda q - p\|_{W^{s,2}(B)}\,, \tag{2.2.6}$$

$$d^s\big|_B(q,p) = \left(\int_{\mathbb{R}} \inf_{\lambda\in\mathbb{S}^1} \|\operatorname{sech}(y-\cdot)(\lambda q - p)\|^2_{W^{s,2}(B)}\,\mathrm{d}y\right)^{\frac{1}{2}}. \tag{2.2.7}$$

Lemma 2.2.6. *Let $s > \frac{1}{2}$ and let $B_0 = \{x \in \mathbb{R} : |x| < R\}$ be an open ball of radius $R > 0$ with center 0. There exists $C(s,R) > 0$ so that*

$$d^s_*\big|_{B_0}(q,p) \le C(s,R)\, d^s\big|_{B_0}(q,p) \tag{2.2.8}$$

for all $q, p \in \mathcal{S}' \cap H^s_{\mathrm{loc}}$. As a consequence, for families of balls $B_k = B_0 + kR$ with $k \in \mathbb{Z}$ we have

$$\sum_{k\in\mathbb{Z}} d^s_*\big|_{B_k}(q,p)^2 \le C(s,R)\, d^s(q,p)^2\,. \tag{2.2.9}$$

Proof. As $\{y-x : x,y \in B_0\} \subseteq \{x \in \mathbb{R} : |x| < 2R\}$, there exists a finite constant $C(s,R) > 0$ such that $\sup_{y\in B_0} \|\operatorname{sech}(y-\cdot)^{-1}\|^2_{W^{s,2}(B_0)} \leq C(s,R)$. The first estimate follows:

$$\begin{aligned}\inf_{\lambda\in\mathbb{S}^1} \|\lambda q - p\|^2_{W^{s,2}(B_0)} &\leq C(s,R) \inf_{y\in B_0} \inf_{\lambda\in\mathbb{S}^1} \|\operatorname{sech}(y-\cdot)(\lambda q - p)\|^2_{W^{s,2}(B_0)} \\ &\leq C(s,R) \int_{\mathbb{R}} \inf_{\lambda\in\mathbb{S}^1} \|\operatorname{sech}(y-\cdot)(\lambda q - p)\|^2_{W^{s,2}(B_0)} \,\mathrm{d}y\,.\end{aligned}$$

Using this and Lemma 2.2.1 (iv), we obtain the second estimate:

$$\begin{aligned}\sum_{k\in\mathbb{Z}} d^s_*\big|_{B_k}(q,p)^2 &\leq C(s,R) \sum_{k\in\mathbb{Z}} d^s\big|_{B_k}(q,p)^2 \\ &\leq C(s,R) \int_{\mathbb{R}} \inf_{\lambda\in\mathbb{S}^1} \sum_{k\in\mathbb{Z}} \|\operatorname{sech}(y-\cdot)(\lambda q - p)\|^2_{W^{s,2}(B_k)} \,\mathrm{d}y \\ &\leq C(s,R)\, d^s(q,p)^2\,.\end{aligned}$$

□

Proof of Theorem 2.1.6 (i). Let $B_0 = \{x \in \mathbb{R} : |x| < 1\}$ and observe that

$$\begin{aligned}&\||q|^2 - |p|^2\|_{W^{s,2}(B_0)} \\ &= \inf_{\lambda,\nu\in\mathbb{S}^1} \||\lambda q|^2 - |\nu p|^2\|_{W^{s,2}(B_0)} \\ &= \inf_{\lambda,\mu,\nu\in\mathbb{S}^1} \||\lambda q - \mu|^2 - |\nu p - \mu|^2 + 2(\operatorname{Re}(\lambda\overline{\mu}q) - \operatorname{Re}(\overline{\mu}\nu p))\|_{W^{s,2}(B_0)}\,.\end{aligned}$$

We can estimate

$$\begin{aligned}&\||q|^2 - |p|^2\|_{W^{s,2}(B_0)} \qquad (2.2.10)\\ &\leq \inf_{\lambda,\mu,\nu\in\mathbb{S}^1} \Big\| \operatorname{Re}\Big(\big((\lambda q - \mu) - (\nu p - \mu)\big)\big(\overline{(\lambda q-\mu) + (\nu p - \mu)}\big)\Big)\Big\|_{W^{s,2}(B_0)} \\ &\qquad + 2\|\operatorname{Re}(\lambda\overline{\mu}q - \overline{\mu}\nu p)\|_{W^{s,2}(B_0)} \\ &\leq C(s) \inf_{\lambda,\nu\in\mathbb{S}^1} \|\lambda q - \nu p\|_{W^{s,2}(B_0)} \inf_{\mu\in\mathbb{S}^1} \big(\|\lambda q - \mu\|_{W^{s,2}(B_0)} + \|\nu p - \mu\|_{W^{s,2}(B_0)} + 2\big) \\ &= C(s) \inf_{\lambda\in\mathbb{S}^1} \|\lambda q - p\|_{W^{s,2}(B_0)} \inf_{\mu\in\mathbb{S}^1} \inf_{\nu\in\mathbb{S}^1} \big(\|\lambda q - \mu\|_{W^{s,2}(B_0)} + \|\nu p - \mu\|_{W^{s,2}(B_0)} + 2\big) \\ &\leq C(s)\, d^s_*\big|_{B_0}(q,p)\big(2 + d^s_*\big|_{B_0}(1,q) + d^s_*\big|_{B_0}(1,p)\big) \\ &\leq C(s,r)\, d^s_*\big|_{B_0}(q,p),\end{aligned}$$

where in the last line we used (2.2.8). Now we set $B_k = B_0 + k$ and see with Lemma 2.2.1 (iv) and (2.2.9) that

$$\begin{aligned}\|\rho - \eta\|^2_{H^s} &\leq C(s) \sum_{k\in\mathbb{Z}} \||q|^2 - |p|^2\|^2_{W^{s,2}(B_k)} \\ &\leq C(s,r) \sum_{k\in\mathbb{Z}} d^s_*\big|_{B_k}(q,p)^2 \\ &\leq C(s,r)\, d^s(q,p)^2.\end{aligned}$$

It remains to estimate $\|\varphi' - \psi'\|_{H^{s-1}}$. Applying Lemma 2.2.2 yields

$$\begin{aligned}\|(\varphi-\psi)'\|^2_{H^{s-1}} &= \|(e^{i(\varphi-\psi)})' e^{-i(\varphi-\psi)}\|^2_{H^{s-1}} \\ &\leq C(s)\|(e^{i(\varphi-\psi)})'\|^2_{H^{s-1}} \big(\|e^{-i(\varphi-\psi)}\|^2_{L^\infty} + \|(e^{-i(\varphi-\psi)})'\|^2_{H^{s-1}}\big) \\ &\leq C(s)\|(e^{i(\varphi-\psi)})'\|^2_{H^{s-1}} \big(1 + \|(e^{i(\varphi-\psi)})'\|^2_{H^{s-1}}\big)\,.\end{aligned}$$

It therefore suffices to derive the estimate for the quantity $\|(e^{i(\varphi-\psi)})'\|^2_{H^{s-1}}$. Observe with Lemma 2.2.1 (iv) that

$$\|(e^{i(\varphi-\psi)})'\|^2_{H^{s-1}} \leq C(s) \sum_{k\in\mathbb{Z}} \inf_{\lambda\in\mathbb{S}^1} \|(e^{i(\varphi-\psi)} - \lambda)'\|^2_{W^{s-1,2}(B_k)}$$

$$
\begin{aligned}
&= C(s) \sum_{k\in\mathbb{Z}} \inf_{\theta\in\mathbb{R}} \|e^{i(\varphi-\psi-\theta)} - 1\|^2_{W^{s,2}(B_k)} \\
&= C(s) \sum_{k\in\mathbb{Z}} \inf_{\theta\in\mathbb{R}} \|e^{-i\psi}(e^{i(\varphi+\theta)} - e^{i\psi})\|^2_{W^{s,2}(B_k)} \,.
\end{aligned}
$$

We now carefully introduce the amplitudes:

$$
\begin{aligned}
&\inf_{\theta\in\mathbb{R}} \|e^{-i\psi}(e^{i(\varphi+\theta)} - e^{i\psi})\|^2_{W^{s,2}(B_k)} \\
&= \inf_{\theta\in\mathbb{R}} \left\| Be^{-i\psi(x)}B^{-1}\left(\frac{Ae^{i(\varphi(x)+\theta)} - Be^{i\psi(x)}}{A} + Be^{i\psi(x)}\left(\frac{1}{A} - \frac{1}{B}\right)\right)\right\|^2_{W^{2,s}(B_k)} \\
&\leq C(s)\|Be^{-i\psi}\|^2_{W^{s,2}(B_k)}\|B^{-1}\|^2_{W^{s,2}(B_k)} \\
&\times \left(\inf_{\theta\in\mathbb{R}} \left\| \frac{Ae^{i(\varphi(x)+\theta)} - Be^{i\psi(x)}}{A} \right\|^2_{W^{s,2}(B_k)} + \left\| Be^{i\psi(x)}\left(\frac{1}{A} - \frac{1}{B}\right)\right\|^2_{W^{s,2}(B_k)} \right) .
\end{aligned}
$$

As $A, B > \delta > 0$, we can apply Lemma 2.2.5. Together with (2.2.10) we obtain

$$
\|A^{-1}\|^2_{W^{s,2}(B_k)} \leq C(s,\delta)\|A\|^2_{W^{s,2}(B_k)} \leq C(s,\delta,r)\big(\inf_{\lambda\in\mathbb{S}^1} \|A-\lambda\|^2_{W^{s,2}(B_k)} + |B_k|\big) \leq C(s,\delta,r) \,,
$$

and similarly $\|B^{\pm 1}\|_{W^{s,2}(B_k)}, \|q^{\pm 1}\|_{W^{s,2}(B_k)}, \|p^{\pm 1}\|_{W^{s,2}(B_k)} \leq C(s,\delta,r)$. We conclude again by reducing the situation to an application of Lemma 2.2.6 and the previously shown estimate (2.2.10):

$$
\begin{aligned}
\|(e^{i(\varphi-\psi)})'\|^2_{H^{s-1}} &\leq C(s) \sum_{k\in\mathbb{Z}} \|p\|^2_{W^{s,2}(B_k)}\|B^{-1}\|^2_{W^{s,2}(B_k)} \\
&\quad\times \Big(\inf_{\lambda\in\mathbb{S}^1} \|\lambda q - p\|^2_{W^{s,2}(B_k)}\|A^{-1}\|^2_{W^{s,2}(B_k)} \\
&\qquad + \|A-B\|^2_{W^{s,2}(B_k)}\|p\|^2_{W^{s,2}(B_k)}\|A^{-1}\|^2_{W^{s,2}(B_k)}\|B^{-1}\|^2_{W^{s,2}(B_k)}\Big) \\
&\leq C(s,\delta,r) \sum_{k\in\mathbb{Z}} d^s_*\big|_{B_k}(q,p)^2 \\
&\leq C(s,\delta,r)\, d^s(q,p)^2 \,.
\end{aligned}
$$

□

2.2.3. PROOF OF THEOREM 2.1.6 (II)

We assume

$$
\theta^s((1,0),(\rho,\partial_x\varphi)), \theta^s((1,0),(\eta,\partial_x\psi)) < r \,,
$$

and $\sqrt{\rho}, \sqrt{\eta} > \delta > 0$. We have to prove that

$$
d^s(q,p) \leq C(s,r)\, \theta^s((\rho,\partial_x\varphi),(\eta,\partial_x\psi)) \,.
$$

As mentioned above, due to Lemma 2.2.4 it suffices to prove this with ρ, η replaced by $A = \sqrt{\rho}$ and $B = \sqrt{\eta}$. Recall the definitions (2.1.19). We define

$$
\tilde{d}^s(q,p) = \left(\int_{\mathbb{R}} \inf_{\lambda\in\mathbb{S}^1} \|\sqrt{\operatorname{sech}(y-\cdot)}(\lambda q - p)\|^2_{H^s}\, \mathrm{d}y \right)^{\frac{1}{2}} ,
$$

where we have replaced the sech in the definition of d^s with $\sqrt{\operatorname{sech}}$. Some of the hard work for this direction has already been done in the proof of the following Lemma 2.2.7. This was proven for d^s in [119, Lemma 6.1], but the proof is identical for $\tilde{d}^s$ as $\sqrt{\operatorname{sech}}$ is positive and still has sufficiently fast decay.

Lemma 2.2.7 ([119, Lemma 6.1])**.** *For all $s \geq 0$ the energy $E^s : X^s \longrightarrow \mathbb{R}_{\geq}$ is continuous with respect to d^s, and there exists $C(s) > 0$ so that*

$$
d^s(1,q) \leq C(s)\sqrt{E^s(q)} \quad \text{and} \quad \tilde{d}^s(1,q) \leq C(s)\sqrt{E^s(q)}
$$

for all $q \in X^s$.

Remark 2.2.8. *The appearance of the square root is explained by a clash of notation: the energies E^s as defined in [119] correspond to $\sqrt{2E^s}$ in our notation.*

We first prove two Lemmas.

Lemma 2.2.9. *Let $s > \frac{1}{2}$. There exists a constant $C(s) > 0$ so that for all $\varphi \in \mathcal{S}' \cap H^s_{\mathrm{loc}}$ we have*

$$\|(e^{i\varphi})'\|_{H^{s-1}} \leq C(s)(1 + \|\varphi'\|_{H^{s-1}})^\gamma \|\varphi'\|_{H^{s-1}} \,,$$

where $\gamma = 2s - 2$ if $s \geq 1$ and $\gamma = \frac{1-s}{s-\frac{1}{2}}$ if $s < 1$.

Proof. We assume $\|\varphi'\|_{H^{s-1}} \neq 0$. By Lemma 2.2.2 there exists a constant $C(s)$ so that

$$\|\varphi' e^{i\varphi}\|_{H^{s-1}} \leq C(s)\|\varphi'\|_{H^{s-1}}\big(\|e^{i\varphi}\|_{L^\infty} + \|(e^{i\varphi})'\|_{H^{s-1}}\big)\,. \tag{2.2.11}$$

For $\varepsilon \in (0, 1)$ and $f \in H^s_{\mathrm{loc}}$ define $f_\varepsilon(x) = f(\varepsilon x)$. This has the scaling estimates

$$\min\{\varepsilon^{s-\frac{1}{2}}, \varepsilon^{\frac{1}{2}}\}\|f'\|_{H^{s-1}} \leq \|(f_\varepsilon)'\|_{H^{s-1}} \leq \max\{\varepsilon^{s-\frac{1}{2}}, \varepsilon^{\frac{1}{2}}\}\|f'\|_{H^{s-1}}\,. \tag{2.2.12}$$

Define $s_{\min} \leq s_{\max}$ so that $\{s_{\min}, s_{\max}\} = \{s - \frac{1}{2}, \frac{1}{2}\}$. Then we can rewrite the above as

$$\varepsilon^{s_{\max}}\|f'\|_{H^{s-1}} \leq \|(f_\varepsilon)'\|_{H^{s-1}} \leq \varepsilon^{s_{\min}}\|f'\|_{H^{s-1}}\,. \tag{2.2.13}$$

We choose $\varepsilon = (1 + 2C(s)\|\varphi'\|_{H^{s-1}})^{-\frac{1}{s_{\min}}}$ so that

$$\|(\varphi_\varepsilon)'\|_{H^{s-1}} \leq \varepsilon^{s_{\min}}\|\varphi'\|_{H^{s-1}} = \frac{\|\varphi'\|_{H^{s-1}}}{1 + 2C(s)\|\varphi'\|_{H^{s-1}}} \leq \frac{1}{2C(s)}\,.$$

Combining this with (2.2.11) yields

$$\|(e^{i\varphi_\varepsilon})'\|_{H^{s-1}} \leq C(s)\|(\varphi_\varepsilon)'\|_{H^{s-1}} + \frac{1}{2}\|(e^{i\varphi_\varepsilon})'\|_{H^{s-1}}\,,$$

so we obtain

$$\|(e^{i\varphi_\varepsilon})'\|_{H^{s-1}} \leq 2C(s)\|(\varphi_\varepsilon)'\|_{H^{s-1}}\,.$$

We conclude with the scaling estimates (2.2.12) that

$$\|(e^{i\varphi})'\|_{H^{s-1}} \leq \varepsilon^{-s_{\max}}\|(e^{i\varphi_\varepsilon})'\|_{H^{s-1}} \leq 2C(s)\varepsilon^{-s_{\max}}\|(\varphi_\varepsilon)'\|_{H^{s-1}} \leq 2C(s)\varepsilon^{s_{\min}-s_{\max}}\|\varphi'\|_{H^{s-1}}\,.$$

Lastly, note that

$$\varepsilon^{s_{\min}-s_{\max}} = \big(1 + 2C(s)\|\varphi'\|_{H^{s-1}}\big)^{\frac{|s-1|}{s_{\min}}} = \big(1 + 2C(s)\|\varphi'\|_{H^{s-1}}\big)^\gamma\,.$$

□

Lemma 2.2.10. *Let $s > \frac{1}{2}$ and $r > 0$. There exists $C(s, r) > 0$ so that for all $q = Ae^{i\varphi} \in \mathcal{S}' \cap H^s_{\mathrm{loc}}$ with $\theta^s((1, 0), (A, \varphi')) < r$ we have*

$$E^s(q) \leq C(s, r)\, \theta^s((1, 0), (A, \varphi'))^2\,.$$

Proof. For the amplitudinal part of the energy, we know from Lemma 2.2.4 that

$$\||q|^2 - 1\|_{H^{s-1}} \leq \|A^2 - 1\|_{H^s} \leq C(s, r)\|A - 1\|_{H^s}\,.$$

For the remainder, we use Lemma 2.2.2:

$$\begin{aligned}
\|q'\|_{H^{s-1}} &\leq \|A' e^{i\varphi}\|_{H^{s-1}} + \|A(e^{i\varphi})'\|_{H^{s-1}} \\
&\leq C(s)\|A'\|_{H^{s-1}}\big(\|e^{i\varphi}\|_{L^\infty} + \|(e^{i\varphi})'\|_{H^{s-1}}\big) \\
&\quad + \|(e^{i\varphi})'\|_{H^{s-1}}\big(\|A - 1\|_{L^\infty} + 1 + \|A'\|_{H^{s-1}}\big)\,.
\end{aligned}$$

We now conclude by estimating both appearances of $\|(e^{i\varphi})'\|_{H^{s-1}}$ with Lemma 2.2.9. □

Proof of Theorem 2.1.6 (ii). We split the distance $d^s(q,p)$ into two parts:

$$\begin{aligned} d^s(q,p)^2 &\le 2\int_{\mathbb{R}} \inf_{\theta\in\mathbb{R}} \|\operatorname{sech}(y-\cdot)A(e^{i(\varphi+\theta)} - e^{i\psi})\|_{H^s}^2 \,\mathrm{d}y \\ &\quad + 2\int_{\mathbb{R}} \|\operatorname{sech}(y-\cdot)(B-A)e^{i\psi}\|_{H^s}^2 \,\mathrm{d}y \\ &= (I) + (II)\,. \end{aligned}$$

We use the algebra property of H^s and Lemma 2.2.7 to estimate

$$\begin{aligned} (I) &\le C(s)\sup_{y\in\mathbb{R}} \left\|\sqrt{\operatorname{sech}(y-\cdot)}Ae^{i\psi}\right\|_{H^s}^2 \tilde{d}^s(1, e^{i(\varphi-\psi)})^2 \\ &\le C(s,r)\big(\|A-1\|_{H^s}^2 + 1\big)\sup_{y\in\mathbb{R}} \left\|\sqrt{\operatorname{sech}(y-\cdot)}e^{i\psi}\right\|_{H^s} E^s(e^{i(\varphi-\psi)})\,. \end{aligned}$$

With Lemma 2.2.10 we can estimate $E^s(e^{i(\varphi-\psi)})$ by $\|\varphi'-\psi'\|_{H^{s-1}}^2$, and Lemma 2.2.2 yields

$$\sup_{y\in\mathbb{R}} \left\|\sqrt{\operatorname{sech}(y-\cdot)}e^{i\psi}\right\|_{H^s} \le C(s)\sup_{y\in\mathbb{R}} \left\|\sqrt{\operatorname{sech}(y-\cdot)}\right\|_{H^s} (\|e^{i\psi}\|_{L^\infty} + \|(e^{i\psi})'\|_{H^{s-1}}) \le C(s,r)\,.$$

It follows that

$$(I) \le C(s,r)\,\theta^s((A,\varphi'),(B,\psi'))^2\,.$$

Note that

$$\begin{aligned} (II) &\le C(s)\int_{\mathbb{R}} \left\|\sqrt{\operatorname{sech}(y-\cdot)}(A-B)\right\|_{H^s}^2 \Big(1 + \inf_{\lambda\in\mathbb{S}^1} \left\|\sqrt{\operatorname{sech}(y-\cdot)}(e^{i\psi}-\lambda)\right\|_{H^s}^2\Big)\,\mathrm{d}y \\ &\le C(s)\left(\int_{\mathbb{R}} \left\|\sqrt{\operatorname{sech}(y-\cdot)}(A-B)\right\|_{H^s}^2 \mathrm{d}y + \|\sqrt{\operatorname{sech}}\|_{H^s}^2 \|A-B\|_{H^s}^2\,\tilde{d}^s(1,e^{i\psi})^2\right). \end{aligned}$$

We can deal with the second term as before. For the first one, we use Lemma 2.2.1 (iv) and Young's convolution inequality:

$$\begin{aligned} \int_{\mathbb{R}} \left\|\sqrt{\operatorname{sech}(y-\cdot)}(A-B)\right\|_{H^s}^2 \mathrm{d}y &\le \sum_{k\in\mathbb{Z}} \sup_{y\in[k,k+1]} \sum_{j\in\mathbb{Z}} \left\|\sqrt{\operatorname{sech}(y-\cdot)}(A-B)\right\|_{W^{s,2}([j,j+3])}^2 \\ &\le \sum_{j,k\in\mathbb{Z}} \|\sqrt{\operatorname{sech}}\|_{C^{\lceil s\rceil+1}([k-j-3,k-j+1]))}^2 \|A-B\|_{W^{s,2}([j,j+3])}^2 \\ &\le \sum_{k\in\mathbb{Z}} \|\sqrt{\operatorname{sech}}\|_{C^{\lceil s\rceil+1}([k-3,k+1]))}^2 \sum_{j\in\mathbb{Z}} \|A-B\|_{W^{s,2}([j,j+3])}^2 \\ &\le C(s)\|A-B\|_{H^s}^2\,. \end{aligned}$$

Therefore

$$(II) \le C(s,r)\,\theta^s((A,\varphi),(B,\psi))^2\,.$$

To conclude, we have shown that

$$d^s(q,p)^2 \le (I) + (II) \le C(s,r)\,\theta^s((A,\varphi),(B,\psi))^2\,.$$

□

Remark 2.2.11. *Recall the definition of $d^s_*\big|_B$ (see (2.2.6)). We have shown in particular that there exist constants such that*

$$\left(\sum_{k\in\mathbb{Z}} d^s_*\big|_{B_k}(q,p)^2\right)^{\frac{1}{2}} \le C(s,r)\,d^s(q,p) \le C(s,\delta,r)\left(\sum_{k\in\mathbb{Z}} d^s_*\big|_{B_k}(q,p)^2\right)^{\frac{1}{2}}$$

for all $q,p\in X^s$ with $|q|,|p|>\delta>0$ and $d^s(1,q), d^s(1,p) < r$. Here the first estimate is Lemma 2.2.6, while the second estimate actually follows from (ii) together with the fact that we showed (i) by proving (2.2.5).

Let us say a few words on how Corollary 2.1.7 follows from Theorem 2.1.6.

Proof of Corollary 2.1.7. The Madelung transform is well-defined on equivalence classes under multiplication by $\mathbb{S}^1$, as $v = \varphi'$ ignores changes by a constant in the phase φ. Note also that $s \geq 1$, and so for any $(\rho, v) \in \mathcal{Y}^s$ we have $v \in L^2 \subset L^1_{\mathrm{loc}}$. Therefore we can define

$$\varphi(x) = \int_0^x v(y)\,\mathrm{d}y\,.$$

Recall that $b < \frac{4}{3}$ and $\varepsilon < \varepsilon_0(\mu)$ (see (2.1.16)). Due to (2.1.17) and (2.1.8), there exists $\delta > 0$ such that $|q| > \delta$ for all $q \in X^s$ with $E(q) < b$ or $E^\mu(q) < \varepsilon$. With Lemma 2.2.7 we find some $r = r(s, \varepsilon, b) > 0$ such that $d^s(1, q) < r$. Then Theorem 2.1.6 establishes the bilipschitz estimates. □

2.3. Proof of Theorem 2.1.9

Given that Theorem 2.1.6 establishes an equivalence between the relevant function spaces (X^s, d^s) and $(\mathcal{Y}^s, \theta^s)$, the proof of Theorem 2.1.9 is now primarily a matter of carefully carrying over the results of Theorem 2.1.1. This is straightforward for the existence and continuity results. Uniqueness requires a further Lemma.

Lemma 2.3.1. *Let $I \ni 0$ be an open time interval and $q_0 \in L^\infty \cap \dot{H}^1$. Suppose*

$$q_1, q_2 \in C(I; L^2_{\mathrm{loc}}) \cap L^\infty(I; L^\infty \cap \dot{H}^1)$$

are two distributional solutions to (GP) with $q_1(0) = q_2(0) = q_0$. Then $q_1 = q_2$.

Proof. See Appendix C. □

This result is necessary because Theorem 2.1.1, in the way it is stated in [119], only yields uniqueness for the following class of solutions, which for the case $s \geq 1$ is a priori smaller.

Definition 2.3.2 (Solutions to (GP) [119])**.** Let $s \geq 0$. We say that $q \in C(I; X^s)$ is a solution of the Gross-Pitaevskii equation (GP) with initial data $q_0 \in X^s$ on the open time interval $I \ni 0$ if there exists $\tilde{q} : I \longrightarrow H^s_{\mathrm{loc}}$ such that the following hold:

(i) $\tilde{q}$ solves (GP) in the sense of distributions on $I \times \mathbb{R}$.

(ii) $\tilde{q}$ projects to q, which means that $\tilde{q}\mathbb{S}^1 = q$.

(iii) We have
$$\big[t \mapsto \tilde{q}(t) - \tilde{q}(0)\big] \in C(I; L^2(\mathbb{R}))\,.$$

(iv) For all compact intervals $[a, b] \subset I$ and for some (and hence for all) regularized initial data $\tilde{q}_0^*$ of $\tilde{q}(0)$ we have
$$\big[t \mapsto \tilde{q}(t) - \tilde{q}_0^*\big] \in L^4([a, b] \times \mathbb{R})\,.$$

The uniqueness result in Theorem 2.1.1 for $s \geq 1$ is therefore weaker than the one in Lemma 2.3.1. The proofs, however, are almost identical: in [119] uniqueness is shown by a classical argument with an energy estimate and Grönwall's inequality. We extend this argument for $s \geq 1$ to gain Lemma 2.3.1.

Remark 2.3.3. *If $\tilde{p} \in C(I; L^2_{\mathrm{loc}}) \cap L^\infty(I; L^\infty \cap \dot{H}^1)$ is a distributional solution to (GP), as in Lemma 2.3.1, with initial data $\tilde{p}(0)\mathbb{S}^1 \in X^1$, then $\tilde{p}\mathbb{S}^1$ is also a solution in the sense of Definition 2.3.2. The reason is that by Theorem 2.1.1 there exists a solution $q \in C(I; X^1)$ in the sense of Definition 2.3.2 with initial data $q(0) = \tilde{p}(0)\mathbb{S}^1$. One can see that this has a representative $\tilde{q} \in C(I; L^2_{\mathrm{loc}}) \cap L^\infty(I; L^\infty \cap \dot{H}^1)$ which solves (GP) in distribution, so Lemma 2.3.1 implies $\tilde{q} = \tilde{p}$.*

Theorem 2.1.9 states that (hGP) is globally-in-time well-posed, meaning that there exist solutions, they are unique, and the flow map is continuous. The structure of the proof is to transfer the existence and continuity result for (GP) from Theorem 2.1.1 via the Madelung transform over to (hGP). This requires the absence of vacuum, which we obtain by the energy assumptions $E < \frac{4}{3}$

or $E^\mu < \varepsilon_0(\mu)$ (see (2.1.17) and (2.1.8)). Uniqueness for (hGP) is similarly inferred from the uniqueness result for (GP) in Lemma 2.3.1.

Recall that by Lemma 2.2.7 the energy functionals $E^s : X^s \longrightarrow \mathbb{R}_{\geq}$ are continuous. Recall furthermore the definitions (2.1.19).

Proof of Theorem 2.1.9.

Existence. We are given an initial data $(\rho_0, v_0) \in \mathcal{Y}^s$ which fulfills one of the bounds $\mathcal{E}(\rho_0, v_0) < \frac{4}{3}$ or $\mathcal{E}^\mu(\rho_0, v_0) < \varepsilon_0(\mu)$. We define $q_0 = \mathcal{M}^{-1}(\rho_0, v_0)$ and obtain via Theorem 2.1.1 a solution $q \in C_b(\mathbb{R}; X^s)$ of (GP) in the sense of Definition 2.3.2. Our solution q has a special representative $\tilde{q} \in \mathcal{S}'(\mathbb{R} \times \mathbb{R})$. In both cases $E(q_0) < \frac{4}{3}$ and $E^\mu(q_0) < \varepsilon_0(\mu)$, we obtain either (2.1.8) or (2.1.17), so there exists some $\delta > 0$ depending on the initial data such that $|\tilde{q}| > \delta > 0$. Now Corollary 2.1.7 implies $(\rho, v) = \mathcal{M}(\tilde{q}) \in C_b(\mathbb{R}; \mathcal{Y}^s)$.

$$\begin{tikzcd}
(\rho_0, v_0) \arrow[r, squiggly, "\text{(hGP)}"] \arrow[d, "\mathcal{M}^{-1}"'] & (\rho, v)(t) \\
q_0 \arrow[r, squiggly, "\text{(GP)}"'] & q(t) \arrow[u, "\mathcal{M}"']
\end{tikzcd}$$

We show that (ρ, v) is a distributional solution of (hGP) in the sense of Definition 2.1.8. We fix a ball $B_0 \subset \mathbb{R}$ and a time interval $J = (a, b) \subset \mathbb{R}$ with $0 \in J$. It suffices to verify that (hGP) holds in distribution, i.e. when tested against any test function $f \in \mathcal{D}(J \times B_0)$.

On regularity. Due to Lemmas 2.2.5 and 2.2.6 for $s \geq 1$, we know that $\tilde{q}, \tilde{q}^{-1} \in L^\infty(J; W^{1,2}(B_0))$. From these considerations $\partial_{xx}\tilde{q} \in L^\infty(J; W^{-1,2}(B_0))$ and $(|\tilde{q}|^2 - 1)\tilde{q} \in L^\infty(J; W^{1,2}(B_0))$ directly follow. Then $\partial_t \tilde{q} \in L^\infty(J; W^{-1,2}(B_0))$ holds because $\tilde{q}$ solves (GP) in the sense of distributions.

As a consequence of duality and the algebra property of H^1, one obtains the product estimate $\|fg\|_{H^{-1}} \leq C\|f\|_{H^1}\|g\|_{H^{-1}}$. From this we obtain some regularity for some of the more difficult terms appearing in the subsequent calculations, for example $\partial_t \tilde{q}\overline{\tilde{q}}$, $\partial_{xx}\tilde{q}\overline{\tilde{q}} \in L^\infty(J; W^{-1,2}(B_0))$. We now present approximation arguments that derive $(\text{hGP})_1$ and $(\text{hGP})_2$ from (GP).

Obtaining (hGP)$_1$ from (GP). Set $\tilde{q}_\varepsilon = \eta_\varepsilon * \tilde{q}$ for a standard mollifier $(\eta_\varepsilon)_{\varepsilon > 0}$, i.e. some $\eta_\varepsilon(x) = \eta(\varepsilon^{-1}(\varepsilon^{-1}x))$ where $\eta \in C_c^\infty(\mathbb{R}; \mathbb{R}_{\geq})$ with $\int \eta \, \mathrm{d}x = 1$. Note that $|\tilde{q}| > \delta$ implies $|\tilde{q}_\varepsilon| > \frac{\delta}{2}$ for sufficiently small $\varepsilon > 0$ as we have sufficient regularity. We define $\rho_\varepsilon = |\tilde{q}_\varepsilon|^2$ and $v_\varepsilon = \operatorname{Im}\big[\frac{\partial_x \tilde{q}_\varepsilon}{\tilde{q}_\varepsilon}\big]$. Note furthermore the identity $\frac{\partial_x \tilde{q}_\varepsilon}{\tilde{q}_\varepsilon} = \frac{1}{2}\frac{\partial_x \rho_\varepsilon}{\rho_\varepsilon} + iv_\varepsilon$, which we use below. Equation $(\text{hGP})_1$ can be obtained by multiplying (GP) for $\tilde{q}_\varepsilon$ with $\overline{\tilde{q}}_\varepsilon$, taking the imaginary part, and then the limit:

$$\begin{aligned}
0 = \operatorname{Im}\big(\overline{\tilde{q}}\,(\text{GP})\big) &\xleftarrow{\varepsilon \to 0} \operatorname{Im}\big[i\partial_t \tilde{q}_\varepsilon \overline{\tilde{q}_\varepsilon} + \partial_{xx}\tilde{q}_\varepsilon \overline{\tilde{q}_\varepsilon} - 2\tilde{q}_\varepsilon(|\tilde{q}_\varepsilon|^2 - 1)\overline{\tilde{q}_\varepsilon}\big] \\
&= \operatorname{Re}\big[\partial_t \tilde{q}_\varepsilon \overline{\tilde{q}_\varepsilon}\big] + \partial_x \operatorname{Im}\left[\frac{\partial_x \tilde{q}_\varepsilon}{\tilde{q}_\varepsilon}\tilde{q}_\varepsilon \overline{\tilde{q}_\varepsilon}\right] - \operatorname{Im}\big[\partial_x \tilde{q}_\varepsilon \partial_x \overline{\tilde{q}_\varepsilon}\big] - \operatorname{Im}\big[2|\tilde{q}_\varepsilon|^2(|\tilde{q}_\varepsilon|^2 - 1)\big] \\
&= \frac{1}{2}\partial_t(|\tilde{q}_\varepsilon|^2) + \partial_x\left(|\tilde{q}_\varepsilon|^2 \operatorname{Im}\left[\frac{\partial_x \tilde{q}_\varepsilon}{\tilde{q}_\varepsilon}\right]\right) \\
&= \frac{1}{2}\partial_t \rho_\varepsilon + \partial_x(\rho_\varepsilon v_\varepsilon) \\
&\xrightarrow{\varepsilon \to 0} \frac{1}{2}\partial_t \rho + \partial_x(\rho v)\,.
\end{aligned}$$

We have to justify the limits in distribution on both sides. Observe that

$$\begin{aligned}
\left|\int_J \int_{B_0} (\partial_t \tilde{q}_\varepsilon \overline{\tilde{q}_\varepsilon} - \partial_t \tilde{q}\overline{\tilde{q}})\overline{f}\right| &\lesssim \|\partial_t \tilde{q}_\varepsilon - \partial_t \tilde{q}\|_{L^\infty(J; W^{-1,2}(B_0))}\|\tilde{q}_\varepsilon\|_{L^\infty(J; W^{1,2}(B_0))}\|f\|_{L^\infty(J; W^{1,2}(B_0))} \\
&\quad + \|\partial_t \tilde{q}\|_{L^\infty(J; W^{-1,2}(B_0))}\|\tilde{q}_\varepsilon - \tilde{q}\|_{L^\infty(J; W^{1,2}(B_0))}\|f\|_{L^\infty(J; W^{1,2}(B_0))} \\
&\xrightarrow{\varepsilon \to 0} 0\,.
\end{aligned}$$

With the same estimates, we can take the limit of the distribution $\partial_{xx}\tilde{q}_\varepsilon\overline{\tilde{q}_\varepsilon}$. The convergence of the nonlinear term follows similarly. We have shown that

$$i\partial_t\tilde{q}_\varepsilon\overline{\tilde{q}_\varepsilon} + \partial_{xx}\tilde{q}_\varepsilon\overline{\tilde{q}_\varepsilon} - 2\tilde{q}_\varepsilon(|\tilde{q}_\varepsilon|^2-1)\overline{\tilde{q}_\varepsilon} \xrightarrow{\varepsilon\to 0} \overline{\tilde{q}}\big(i\partial_t\tilde{q} + \partial_{xx}\tilde{q} - 2\tilde{q}(|\tilde{q}|^2-1)\big) = 0$$

in distribution on $J\times B_0$. As

$$\begin{aligned}
&\left|\int_J\int_{B_0}(\rho_\varepsilon v_\varepsilon - \rho v)\overline{\partial_x f}\,\mathrm{d}x\,\mathrm{d}t\right| \\
&\lesssim \big(\|\tilde{q}\|_{L^\infty(J\times B_0)} + \|\tilde{q}_\varepsilon\|_{L^\infty(J\times B_0)}\big)\|\tilde{q}_\varepsilon - \tilde{q}\|_{L^\infty(J\times B_0)}\|v\|_{L^\infty(J;L^2(B_0))}\|\partial_x f\|_{L^2(J\times B_0)} \\
&+ \big(\|\tilde{q}\|_{L^\infty(J\times B_0)} + \|\tilde{q}_\varepsilon\|_{L^\infty(J\times B_0)}\big)\|\tilde{q}_\varepsilon\|_{L^\infty(J\times B_0)}\|v_\varepsilon - v\|_{L^\infty(J;L^2(B_0))}\|\partial_x f\|_{L^2(J\times B_0)} \xrightarrow{\varepsilon\to 0} 0\,,
\end{aligned}$$

and we can similarly show $\partial_t\rho_\varepsilon \xrightarrow{\varepsilon\to 0} \partial_t\rho$, we have

$$\frac{1}{2}\partial_t\rho_\varepsilon + \partial_x(\rho_\varepsilon v_\varepsilon) \xrightarrow{\varepsilon\to 0} \frac{1}{2}\partial_t\rho + \partial_x(\rho v)$$

in distribution on $J\times B_0$.

Obtaining $(\mathrm{hGP})_2$ from (GP). We now repeat these arguments for the second equation $(\mathrm{hGP})_2$. Here we divide (GP) for $\tilde{q}_\varepsilon$ by $\tilde{q}_\varepsilon$, take a further derivative, the real part, and then the limit:

$$\begin{aligned}
0 = \operatorname{Re}\partial_x\left(\frac{(\mathrm{GP})}{\tilde{q}}\right) \xleftarrow{\varepsilon\to 0} &\operatorname{Re}\left[\partial_x\left(\frac{i\partial_t\tilde{q}_\varepsilon}{\tilde{q}_\varepsilon} + \frac{\partial_{xx}\tilde{q}_\varepsilon}{\tilde{q}_\varepsilon} - 2(|\tilde{q}_\varepsilon|^2-1)\right)\right] \\
&= -\operatorname{Im}[\partial_x\partial_t(\ln\tilde{q}_\varepsilon)] + \operatorname{Re}\left[\partial_x\left(\partial_x\left(\frac{\partial_x\tilde{q}_\varepsilon}{\tilde{q}_\varepsilon}\right) + \left(\frac{\partial_x\tilde{q}_\varepsilon}{\tilde{q}_\varepsilon}\right)^2\right)\right] - 2\partial_x(|\tilde{q}_\varepsilon|^2) \\
&= -\partial_t v_\varepsilon - \partial_x(v_\varepsilon^2) - 2\partial_x\rho_\varepsilon + \partial_x\left(\partial_x\left(\frac{1}{2}\frac{\partial_x\rho_\varepsilon}{\rho_\varepsilon}\right) + \left(\frac{1}{2}\frac{\partial_x\rho_\varepsilon}{\rho_\varepsilon}\right)^2\right) \\
&\xrightarrow{\varepsilon\to 0} -\partial_t v - \partial_x(v^2) - 2\partial_x\rho + \partial_x\left(\partial_x\left(\frac{1}{2}\frac{\partial_x\rho}{\rho}\right) + \left(\frac{1}{2}\frac{\partial_x\rho}{\rho}\right)^2\right).
\end{aligned}$$

Of course, the limits have to be justified again. For the left-hand side, we can proceed just as before since $\tilde{q}^{-1}\in L^\infty(J;W^{1,2}(B_0))$. On the right hand side the difficult terms are v^2 and $\left(\frac{1}{2}\frac{\partial_x\rho}{\rho}\right)^2$, as here the square of a distribution in H^{s-1} is taken. The situation would be much more difficult if we did not assume $s\geq 1$. In our case we indeed have $v, \frac{\partial_x\rho}{\rho}\in C(J;L^2(B_0))$, which implies that the squares are trivially defined. Furthermore

$$\begin{aligned}
&\left|\int_J\int_{B_0}(v_\varepsilon^2 - v^2)\partial_x f\,\mathrm{d}x\right| \\
&\leq \|v_\varepsilon - v\|_{L^\infty(J;L^2(B_0))}\big(\|v_\varepsilon\|_{L^\infty(J;L^2(B_0))} + \|v\|_{L^\infty(J;L^2(B_0))}\big)\|\partial_x f\|_{L^\infty(J\times B_0)} \xrightarrow{\varepsilon\to 0} 0\,,
\end{aligned}$$

and the same estimate works for $\left(\frac{1}{2}\frac{\partial_x\rho}{\rho}\right)^2$. The remaining terms are strictly easier to deal with.

Uniqueness. Let $0\in I\subset\mathbb{R}$ be a bounded open interval and let $(\rho_1,v_1),(\rho_2,v_2)\in C(I;\mathcal{Y}^s)$ be two solutions to (hGP) in the sense of the theorem, both with initial data $(\rho_0,v_0)\in\mathcal{Y}^s$. In particular, they satisfy one of the energy bounds $\mathcal{E}<b<\frac{4}{3}$ or $\mathcal{E}^\mu < c(\mu)\varepsilon < c(\mu)\varepsilon_0(\mu)$ (see (2.1.15) and (2.1.16)). As before, this implies that there exists a $\delta>0$ so that $\sqrt{\rho_k}>\delta$, where $k\in\{0,1,2\}$. Since $v_k\in L^2\subset L^1_{\mathrm{loc}}$, we can define $\varphi_k(x) = \int_0^x v_k(y)\,\mathrm{d}y$ and $\tilde{q}_k = \sqrt{\rho_k}e^{i\varphi_k}$. Note that $\tilde{q}_k$ having uniformly bounded energy E^1 implies $\tilde{q}_k\in L^\infty(I;L^\infty\cap\dot{H}^1)$. We now fix $j\in\{1,2\}$. Writing $q_j = \tilde{q}_j\mathbb{S}^1$ for the equivalence class, we know from Corollary 2.1.7 that $q_j\in C(I;X^s)$.

Just as in the existence part of the proof, one can show that (ρ_j,v_j) solving (hGP) implies that for the quantity

$$Q_j = i\partial_t\tilde{q}_j + \partial_{xx}\tilde{q}_j - 2\tilde{q}_j(|\tilde{q}_j|^2-1) \tag{2.3.1}$$

we have

$$\operatorname{Im}\big[Q_j\overline{\tilde{q}_j}\,\big] = 0 \qquad \text{and} \qquad \partial_x\operatorname{Re}\left[\frac{Q_j}{\tilde{q}_j}\right] = 0 \tag{2.3.2}$$

in the sense of distributions. We sketch the argument that follows with a diagram.

$$\begin{tikzcd}[column sep=large, row sep=large]
(\rho_1, v_1)(t) \arrow[d, mapsto, "\mathcal{M}^{-1}"'] & (\rho_0, v_0) \arrow[l, squiggly, "\text{(hGP)}"'] \arrow[r, squiggly, "\text{(hGP)}"] \arrow[d, mapsto, "\mathcal{M}^{-1}"'] & (\rho_2, v_2)(t) \arrow[d, mapsto, "\mathcal{M}^{-1}"] \\
\tilde{q}_1(t) \arrow[d, mapsto, "\cdot e^{iG_1(t)}"'] & \tilde{q}_0 \arrow[l, dashed, "(2.3.1)\text{ - }(2.3.2)"] \arrow[r, dashed, "(2.3.1)\text{ - }(2.3.2)"'] \arrow[d, equal] & \tilde{q}_2(t) \arrow[d, mapsto, "\cdot e^{iG_2(t)}"] \\
p_1(t) \arrow[rr, equal, bend right, "\text{Lemma 2.3.1}"'] & p_0 \arrow[l, squiggly, "\text{(GP)}"] \arrow[r, squiggly, "\text{(GP)}"'] & p_2(t)
\end{tikzcd}$$

Due to (2.3.2) we have in particular

$$\operatorname{Im}\left[\frac{Q_j}{\tilde{q}_j}\right] = \operatorname{Im}\left[\frac{Q_j\overline{\tilde{q}_j}}{|\tilde{q}_j|^2}\right] = \frac{\operatorname{Im}[Q_j\overline{\tilde{q}_j}]}{|\tilde{q}_j|^2} = 0\,,$$

and hence for every $t \in I$ there exists a $g_j(t) \in \mathbb{R}$ so that

$$g_j(t) = \frac{Q_j}{\tilde{q}_j}\,.$$

We see that, in fact, $\tilde{q}_j$ does not necessarily solve (GP). The reason is that for each time $t \in I$ we had to make an arbitrary choice of a constant-in-space phase rotation, as this information is lost in the Madelung transform. This choice was the arbitrary lower limit 0 in the integral $\varphi(t) = \int_0^t v(s)\,\mathrm{d}s$. In order to find solutions to (GP), we would now like to define

$$p_j(t) = e^{iG_j(t)}\tilde{q}_j(t) \qquad \text{where} \qquad G_j(t) = \int_0^t g_j(s)\,\mathrm{d}s\,.$$

Then

$$i\partial_t p_j + \partial_{xx} p_j - 2p_j(|p_j|^2 - 1) = Q_j - G'(t)\tilde{q}_j = 0\,.$$

This argument requires $g_j : I \longrightarrow \mathbb{R}$ to be locally integrable. We show that $g_j \in C(I;\mathbb{R})$ by verifying that $Q_j \in C(I;W^{-1,2}(B_0))$ for any ball $B_0 \subset \mathbb{R}$. With the same reasoning as in the existence part of the proof, $\rho_j \in C(I;W^{1,2}(B_0))$ and $v_j \in C(I;L^2(B_0))$ solving (hGP) in distribution implies $\rho_j \in C^1(I;W^{-1,2}(B_0))$ and $v_j \in C^1(I;W^{-1,1}(B_0))$. In particular we have $\partial_t\varphi_j \in C(I;L^1(B_0))$. Observe that

$$Q_j = i\frac{\partial_t\rho_j}{\rho_j}\tilde{q}_j + i(\partial_t\varphi_j)\tilde{q}_j + \partial_{xx}\tilde{q}_j - 2\tilde{q}_j(|\tilde{q}_j|^2 - 1)\,.$$

Verifying the products of distributions, each term can now be seen to be in $C(I;W^{-1,2}(B_0))$. We have shown that for any bounded interval $I \ni 0$, both p_1 and p_2 are distributional solutions to (GP) with initial data $p_1(0) = p_2(0) = \tilde{q}_0$. At the same time $p_j \in C(I;L^2_{\mathrm{loc}}) \cap L^\infty(I;L^\infty_{t,x} \cap \dot{H}^1)$. Therefore Lemma 2.3.1 implies $p_1 = p_2$, from which $q_1 = q_2$ in $C(I;X^s)$ and $(\rho_1, v_1) = (\rho_2, v_2)$ follow.

Continuity. This is a direct consequence of the continuity result for (GP) from Theorem 2.1.1, the continuity of the energy functionals from Lemma 2.2.7, and the local bilipschitz equivalence from Theorem 2.1.6. ☐

2.A. Appendix: Absence of vacuum for small energies

Lemma 2.A.1. *For $\delta \in [0,1]$ and $s \in \left(\frac{1}{2}, 1\right]$ define*

$$E^s_\delta = \inf\left\{E^s(q) : q \in H^s_{\mathrm{loc}}, \inf_{x\in\mathbb{R}} |q(x)| \le \delta\right\}.$$

Then $E_1^s = 0$, the function $\delta \mapsto E_\delta^s$ is decreasing, and there exists a constant $\tilde{C}(s) > 0$ so that

$$E_\delta^s \geq \frac{(1-\delta)^2}{\tilde{C}(s)}\,. \tag{2.A.1}$$

Assume $s = 1$ and write $E_\delta = E_\delta^1$. Set $q_0 = \tanh$, $q_1 = 1$, and for $\delta \in (0,1)$ define

$$q_\delta(x) = \tanh\big(|x| + \tanh^{-1}(\delta)\big)\,. \tag{2.A.2}$$

We have

$$E_\delta = E(q_\delta) = \frac{4}{3} - 2\delta + \frac{2}{3}\delta^3\,.$$

There exists a strictly decreasing inverse function $\tilde{\delta} : [0, \frac{4}{3}] \longrightarrow [0,1]$ with $\tilde{\delta}(0) = 1$, $\tilde{\delta}(\frac{4}{3}) = 0$ and

$$\tilde{\delta}(b) = \inf\big\{ \inf_{x\in\mathbb{R}} |q(x)| : q \in H^1_{\mathrm{loc}}, E(q) \leq b\big\}.$$

Proof. We see that $E_1^s = 0$ by choosing $q = 1$. Clearly the set over which the infimum is taken increases with δ, and hence the infimum is decreasing. Recall that Lemma 2.2.6 implies

$$\inf_{\lambda\in\mathbb{S}^1} \|q-\lambda\|_{W^{s,2}(B_k)} \leq C(s)\, d^s(1,q)\,,$$

where $B_k = B_0 + k, k \in \mathbb{Z}$ are balls of radius 1. Estimating with Lemma 2.2.7 on the right and the Sobolev embedding $W^{1,2}(B_k) \hookrightarrow L^\infty$ on the left, we obtain

$$1-\delta \leq \sup_{k\in\mathbb{Z}} \inf_{\lambda\in\mathbb{S}^1} \|q-\lambda\|_{L^\infty(B_k)} \leq C(s)\sqrt{E^s(q)}$$

for every $q \in H^s_{\mathrm{loc}}$ with $\inf_{x\in\mathbb{R}} |q(x)| \leq \delta$. This proves (2.A.1).

Now we assume $s = 1$. We first rewrite the problem as $E_\delta = \inf_{\nu\in[0,\delta]} \tilde{E}_\nu$ with

$$\tilde{E}_\nu = \inf\{E(q) : q \in H^1_{\mathrm{loc}}, \inf_{x\in\mathbb{R}} |q(x)| = \nu\}\,, \qquad \nu \in [0,\delta]\,.$$

Of course we expect that $\tilde{E}_\nu$ is decreasing in ν and hence $E_\delta = \tilde{E}_\delta$. This will be verified once we have calculated $\tilde{E}_\nu$. Using invariance under translations, phase rotations, and mirror symmetry, we can equivalently consider the minimization problem

$$\tilde{E}_\nu = 2\inf\bigg\{\frac{1}{2}\int_0^\infty |\partial_x q|^2 + (|q|^2-1)^2 \,\mathrm{d}x : q \in H^1_{\mathrm{loc}}(\mathbb{R}_\geq), q(0) = \nu\bigg\}\,.$$

We now follow the same arguments as in [26, Lemma 1] to find a minimizer. Consider a minimizing sequence $(q_n)_{n\in\mathbb{N}}$. As $E^s(q_n)$ is uniformly bounded, so is $\|\partial_x q_n\|_{L^2(\mathbb{R}_\geq)}$. The Banach-Alaoglu theorem then implies, up to a subsequence, that $\partial_x q_n \longrightarrow p'_\nu$ for some $p'_\nu \in L^2(\mathbb{R}_\geq)$. Furthermore as $q_n(0) = \nu$ is fixed, we have a Poincare inequality $\|q_n\|_{W^{1,2}(B_0)} \leq C(s,B)\,\|\partial_x q_n\|_{L^2(B_0)}$ on any finite interval $B_0 \subset \mathbb{R}_\geq$. Then we can use compactness of the Sobolev embedding $H^1 \hookrightarrow L^\infty$ to find, up to a subsequence, that $q_n \longrightarrow p_\nu$ in $L^\infty_{\mathrm{loc}}(\mathbb{R}_\geq)$ for some $p_\nu \in H^1_{\mathrm{loc}}(\mathbb{R}_\geq)$, with p'_ν indeed being its distributional derivative. Now we can conclude with Fatou's lemma that p_ν is a minimizer for $\tilde{E}_\nu$:

$$\begin{aligned}\int_0^\infty (|p_0|^2-1)^2 + |p'_0|^2 \,\mathrm{d}x &= \int_0^\infty \liminf_n (|q_n|^2-1)^2 + \liminf_n |q'_n|^2 \,\mathrm{d}x \\ &\leq \liminf_n \int_0^\infty (|q_n|^2-1)^2 + \liminf_n |q'_n|^2 \,\mathrm{d}x \\ &= E_\nu\,.\end{aligned}$$

For the case $\nu = 0$, we obtain the Euler-Lagrange equation

$$p''_0 - 2p_0(1-|p_0|^2) = 0\,.$$

Then as $p_0(0) = 0$ and $E(p_0) < \infty$, [26, Theorem 1] implies that $p_0 = \tanh$ is the unique solution. Consequently, it must be the case that for $a > 0$ the function $p_\nu(x) = \tanh(x+a)$ is a minimizer for the problem with $\nu = \tanh(a)$, as otherwise one could modify p_0 on $[r, \infty)$ to find an admissible function with strictly smaller energy for the minimization problem of $\tilde{E}_0$. This implies that the q_δ defined in (2.A.2) are minimizers for $\tilde{E}_\nu$.

With $a = \tanh^{-1}(\nu)$, and noting the identities

$$\cosh(a) = \frac{1}{\sqrt{1-\nu^2}}, \qquad \cosh(2a) = -\frac{1+\nu^2}{1-\nu^2}, \qquad \operatorname{sech}^2(a) = 1-\nu^2,$$

we compute

$$\begin{aligned} E(q_\nu) &= 2 \cdot \frac{1}{2} \int_a^\infty (\tanh(x)')^2 + (\tanh(x)^2 - 1)^2 \,\mathrm{d}x \\ &= \int_a^\infty \operatorname{sech}^4(x) + \operatorname{sech}^4(x) \,\mathrm{d}x \,. \end{aligned}$$

Evaluating the integral yields

$$\begin{aligned} E(q_\nu) &= \left[\frac{2}{3}(\cosh(2x) + 2)\tanh(x)\operatorname{sech}^2(x)\right]_r^\infty \\ &= \frac{2}{3}\Big(2 - \nu(1-\nu^2)\Big(2 + \frac{1+\nu^2}{1-\nu^2}\Big)\Big) \\ &= \frac{4}{3} - 2\nu + \frac{2}{3}\nu^3 \,. \end{aligned}$$

□

2.B. APPENDIX: LITTLEWOOD-PALEY THEORY AND PROOF OF LEMMA 2.2.2

The proof uses the Bony decomposition

$$fg = T_f g + R(f,g) + T_g f\,,$$

which J.M. Bony introduced in his 1981 paper [33]. It relies on the Littlewood-Paley theory, for which we refer the reader to [20, Chp. 2]. We give a brief introduction below, always only considering the one-dimensional case.

Let $\varphi \in C_c^\infty(\{\xi : \frac{3}{4} < |\xi| < \frac{8}{3}\}$ and $\chi \in C_c^\infty(\{\xi : |\xi| < \frac{4}{3}\})$ be non-negative functions on $\mathbb{R}$ so that

$$\chi(\xi) + \sum_{j=0}^{\infty} \varphi(2^{-j}) = 1\,.$$

This is called a dyadic partition of unity. We define for $j \in \mathbb{Z}$ the operators

$$\begin{aligned} \Delta_j : \mathcal{S}' &\longrightarrow \mathcal{S}' \\ f &\longmapsto \Delta_j f = \begin{cases} \varphi(2^{-j}\cdot)\widehat{f} & , j \geq 0 \\ \chi\widehat{f} & , j = -1 \\ 0 & , j \leq -2 \end{cases} \end{aligned}$$

and $S_j = \sum_{j'<j} \Delta_{j'}$. These operators have nice properties such as $\|S_j f\|_{L^p} \leq C(p)\|f\|_{L^p}$, $p \in [1, \infty]$. At least formally, we have the decomposition

$$\mathrm{Id} = \lim_{j\to\infty} S_j = \sum_j \Delta_j\,.$$

The Bony decomposition is given by

$$fg = \sum_{j,k} \Delta_j f \Delta_k g = T_f g + R(f,g) + T_g f\,,$$

where we define

$$T_f g = \sum_j S_{j-1} f \Delta_j g \qquad\qquad R(f,g) = \sum_j \sum_{|\nu|\le 1} \Delta_{j+\nu} f \Delta_j g\,.$$

In the Littlewood-Paley setting it is easy to define the Besov spaces $B^s_{p,q}$ for $1 \le p, r \le \infty$, $s \in \mathbb{R}$ by

$$B^s_{p,q} = \left\{ f \in \mathcal{S}'(\mathbb{R};\mathbb{C}) : \|f\|_{B^s_{p,q}} < \infty \right\},$$

where

$$\|f\|_{B^s_{p,q}} = \left\| \left(2^{js}\|\Delta_j f\|_{L^p}\right)_{j\in\mathbb{Z}} \right\|_{\ell^r}.$$

It is evident that

$$B^s_{2,2} = \{ f \in \mathcal{S}'(\mathbb{R};\mathbb{C}) : \|\langle\xi\rangle^s \widehat{f}\|_{L^2(\mathbb{R})} < \infty \} = H^s(\mathbb{R};\mathbb{C})\,.$$

Proof of Lemma 2.2.2. We only prove (2.2.2) as the proof of (2.2.1) is analogous and strictly simpler. Consider the decomposition $fg = gS_0 f + g(1-S_0)f$. Since $S_0 f$ is spectrally supported in a fixed ball there exists a constant $N \in \mathbb{N}$ so that $\Delta_k(S_0 f \Delta_j g) = 0$ unless $|k-j| \le N$. Consequently,

$$\begin{aligned}
\|gS_0 f\|^2_{H^{s-1}} &= \sum_{j\in\mathbb{Z}} 2^{2j(s-1)} \left\| \sum_{|\nu|\le N} \Delta_j (S_0 f \Delta_{j+\nu} g) \right\|^2_{L^2} \\
&\le C(N) \|S_0 f\|^2_{L^\infty} \sum_{j\in\mathbb{Z}} 2^{2j(s-1)} \|\Delta_k g\|^2_{L^2} \\
&= C(N) \|S_0 f\|^2_{L^\infty} \|g\|^2_{H^{s-1}}\,.
\end{aligned}$$

It remains to estimate $\|g(1-S_0)f\|_{H^{s-1}}$. To simplify notation, we now write f for $(1-S_0)f$ and derive an estimate by $\|f\|_{H^s}$. Note that $S_{j-1} f \Delta_j g$ is only non-zero if $j \ge 1$, and in that case it is a convolution of a ball with an annulus of much larger radius. As a result, there exists an annulus $\mathcal{C}$ so that $\mathcal{F}[S_{j-1} f \Delta_j g]$ is supported in $2^j \mathcal{C}$, and so [20, Lemma 2.69] implies

$$\|T_f g\|_{H^{s-1}} \lesssim \left\| 2^{j(s-1)} \|S_{j-1} f \Delta_j g\|_{L^2} \right\|_{\ell^2(\mathbb{Z})}.$$

Since

$$\|S_{j-1} f \Delta_j g\|_{L^2} \le \|S_{j-1} f\|_{L^\infty} \|\Delta_j g\|_{L^2} \le \|f\|_{L^\infty} \|\Delta_j g\|_{L^2}\,,$$

this implies

$$\|T_f g\|_{H^{s-1}} \lesssim \|g\|_{H^{s-1}} \|f\|_{L^\infty}\,.$$

For the same reason as before, we have

$$\|T_g f\|_{H^{s-1}} \lesssim \left\| 2^{j(s-1)} \|S_{j-1} g \Delta_j f\|_{L^2} \right\|_{\ell^2}.$$

Here we consider two cases. If $s \le 1$ then we use the Bernstein inequality [20, Lemma 2.1]. It states that

$$\operatorname{supp} \widehat{u} \subset \lambda B \Longrightarrow \|u\|_{L^\infty} \le C(B) \lambda^{\frac{1}{2}} \|u\|_{L^2}$$

for any fixed ball B. This yields

$$\begin{aligned}
2^{j(s-1)} \|S_{j-1} g \Delta_j f\|_{L^2} &\le 2^{j(s-1)} \|S_{j-1} g\|_{L^2} \|\Delta_j f\|_{L^\infty} \\
&\lesssim \|S_{j-1} g\|_{H^{s-1}} 2^{\frac{j}{2}} \|\Delta_j f\|_{L^2} \\
&\lesssim \|g\|_{H^{s-1}} 2^{\frac{j}{2}} \|\Delta_j f\|_{L^2}\,.
\end{aligned}$$

Here we have used

$$2^{2j(s-1)} \|S_{j-1} g\|^2_{L^2} = \sum_{j'<j-1} \underbrace{2^{2(j-j')(s-1)}}_{\le 1} 2^{2j'(s-1)} \|\Delta_{j'} g\|^2_{L^2} \le \|S_{j-1} g\|^2_{H^{s-1}} \le \|g\|^2_{H^{s-1}}\,.$$

We see that

$$\|T_g f\|_{H^{s-1}} \lesssim \|g\|_{H^{s-1}} \|f\|_{H^{\frac{1}{2}}} .$$

For the case $s > 1$ we estimate

$$\begin{aligned} 2^{j(s-1)} \|S_{j-1} g \Delta_j f\|_{L^2} &\le 2^{j(s-1)} \|S_{j-1} g\|_{L^\infty} \|\Delta_j f\|_{L^2} \\ &\lesssim \|2^{-j} S_{j-1} g\|_{H^1} 2^{js} \|\Delta_j f\|_{L^2} \\ &\lesssim \|g\|_{L^2} 2^{js} \|\Delta_j f\|_{L^2} \end{aligned}$$

and obtain

$$\|T_g f\|_{H^{s-1}} \lesssim \|g\|_{H^{s-1}} \|f\|_{H^s} .$$

It remains to estimate the remainder terms $R(f, g)$. Here let it be noted that there exists an integer $N > 0$, independent of j, so that $\sum_{|\nu| \le 1} \Delta_{j-\nu} f \Delta_j g$ is spectrally supported in a ball of radius 2^{j+N-1}. In this case we know by [20, Lemma 2.84] that

$$\|R(f, g)\|_{B^{\tilde{s}}_{p,r}} \le C(p, r, \tilde{s}) \Big\| 2^{j\tilde{s}} \| \sum_{|\nu|=1} \Delta_{j-\nu} g \Delta_j f \|_{L^p} \Big\|_{\ell^r(\mathbb{Z})} \tag{2.B.1}$$

for $\tilde{s} > 0$. This does not work in general if $\tilde{s} < 0$. Therefore, we use the embedding

$$\|R(f, g)\|_{H^{s-1}} \le C(s) \|R(f, g)\|_{B^{s-\frac{1}{2}}_{1,1}}$$

in order to apply (2.B.1) with $\tilde{s} = s - \frac{1}{2} > 0$ and $p, r = 1$. Now we can conclude via Hölder's and Young's inequalities for sequences:

$$\begin{aligned} \Big\| 2^{j(s-\frac{1}{2})} \Big\| \sum_{|\nu|=1} \Delta_{j-\nu} f \Delta_j g \Big\|_{L^1} \Big\|_{\ell^1(\mathbb{Z})} &\lesssim \sum_{|\nu| \le 1} 2^{\frac{\nu}{2}} \big\| 2^{\frac{j-\nu}{2}} \|\Delta_{j-\nu} f\|_{L^2} \big\|_{\ell^2(\mathbb{Z})} \big\| 2^{j(s-1)} \|\Delta_j g\|_{L^2} \big\|_{\ell^2(\mathbb{Z})} \\ &\lesssim \|g\|_{H^{s-1}} \|f\|_{H^{\frac{1}{2}}} . \end{aligned}$$

□

2.C. Appendix: Uniqueness for the Gross-Pitaevskii equation

Proof of Lemma 2.3.1. Recall that $q_1, q_2 \in C(I; L^2_{\mathrm{loc}}) \cap L^\infty(I; L^\infty \cap \dot{H}^1)$ are two distributional solutions of (GP) on an open interval $I \ni 0$ with the same initial data q_0. It follows that $q_1, q_2 \in L^\infty(I; W^{1,2}(B))$ for any arbitrary ball $B \subset \mathbb{R}$. We define $b = q_1 - q_2$ and compute that it solves in distribution the following equation:

$$\begin{aligned} i\partial_t b + \partial_{xx} b &= 2q_1(|q_1|^2 - 1) - 2q_2(|q_2|^2 - 1) \\ &= 2b(|q_1|^2 - 1) + 2b(|q_2|^2 - 1) + 2q_2(|q_1|^2 - 1) - 2q_1(|q_2|^2 - 1) \\ &= 2b(|q_1|^2 + |q_2|^2 - 2 + 1) + 2(q_2|q_1|^2 - q_1|q_2|^2) \\ &= 2b((b + q_2)\overline{(b + q_2)} + |q_2|^2 - 1) + 2(q_2|b + q_2|^2 - (b + q_2)|q_2|^2) \\ &= 2b(|b|^2 + b\overline{q_2} + \overline{b}q_2 + 2|q_2|^2 - 1) \\ &+ 2(q_2|b|^2 + b|q_2|^2 + \overline{b}q_2^2 + q_2|q_2|^2 - b|q_2|^2 - q_2|q_2|^2) \\ &= 2b(|b|^2 + b\overline{q_2} + \overline{b}q_2 + 2|q_2|^2 - 1) + 2(q_2|b|^2 + \overline{b}q_2^2) \\ &= 2|b|^2 b + 4|b|^2 q_2 + 2b^2\overline{q_2} + 2b(2|q_2|^2 - 1) + 2\overline{b}q_2^2 . \end{aligned}$$

We know that $\partial_{xx} b \in L^\infty(I; W^{-1,2}(B))$. Then b solving the equation implies $\partial_t b \in L^\infty(I; W^{-1,2}(B))$. Using duality and the algebra property of $W^{1,2}(B)$, we find that also $\overline{b}\,\partial_t b, \partial_t(|b|^2) \in L^\infty(I; W^{-1,2}(B))$.

Let $\varphi_n(x) = \varphi(\frac{x}{n})$ where $\varphi \in C_c^\infty([-2,2];[0,1])$ and $\varphi\big|_{[-1,1]} = 1$. One may choose φ in such a way that there exists $K > 0$ with $|\partial_x \varphi| \leq K\sqrt{\varphi}$ and in particular $|\partial_x \varphi_n| \leq K n^{-1}\sqrt{\varphi_n}$. We test the above with $\overline{b}\varphi_n$ and take the imaginary part. On the left-hand side, we have

$$\int_{\mathbb{R}} \mathrm{Im}[i(\partial_t b)\overline{b}\varphi_n] + \mathrm{Im}[(\partial_{xx} b)\overline{b}\varphi_n]\,\mathrm{d}x = \int_{\mathbb{R}} \frac{1}{2}\varphi_n \partial_t(|b|^2) - \mathrm{Im}[(\partial_x b)\overline{b}\partial_x\varphi_n]\,\mathrm{d}x\,.$$

Therefore for a fixed time $t \in I$,

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}} \varphi_n |b|^2\,\mathrm{d}x &= \int_{\mathbb{R}} \mathrm{Im}[(\partial_x b)\overline{b}\partial_x\varphi_n]\,\mathrm{d}x \\ &\quad + \int_{\mathbb{R}} \big(2|b|^4 + 4|b|^2\overline{b}q_2 + 2|b|^2 b\overline{q_2} + 2|b|^2(2|q_2|^2 - 1) + 2\overline{b}^2 q_2^2\big)\varphi_n\,\mathrm{d}x \\ &= (I) + (II)\,.\end{aligned}$$

We estimate

$$(I) \leq \|\partial_x b\|_{L^2_x}\|b\partial_x\varphi_n\|_{L^2_x} \leq \big(\|q_1\|_{L^\infty_t \dot{H}^1_x} + \|q_2\|_{L^\infty_t \dot{H}^1_x}\big)Kn^{-1}\big\|b\sqrt{\varphi_n}\big\|_{L^2_x}$$

and

$$\begin{aligned}(II) &\leq C\big\|b\sqrt{\varphi_n}\big\|^2_{L^2_x}\big\||b|^2 + |b||q_2| + |q_2|^2 + 1\big\|_{L^\infty_{t,x}} \\ &\leq C\big\|b\sqrt{\varphi_n}\big\|^2_{L^2_x}\big(1 + \|q_1\|^2_{L^\infty_{t,x}} + \|q_2\|^2_{L^\infty_{t,x}}\big)\,.\end{aligned}$$

We have shown that there exists some $C > 0$, depending on q_1, q_2 but independent of time, such that

$$\frac{1}{2}\frac{d}{dt}\big(\|b\sqrt{\varphi_n}\|^2_{L^2_x}\big) \leq C\left(\frac{1}{\sqrt{n}}\|b\sqrt{\varphi_n}\|_{L^2_x} + \|b\sqrt{\varphi_n}\|^2_{L^2_x}\right).$$

In particular

$$\frac{d}{dt}\|b\sqrt{\varphi_n}\|_{L^2_x} \leq C\left(\frac{1}{\sqrt{n}} + \|b\sqrt{\varphi_n}\|_{L^2_x}\right).$$

Now Grönwall's inequality implies for any fixed $t > 0$ that

$$\|b(t)\|_{L^2_x} \xleftarrow{n\to\infty} \|b(t)\sqrt{\varphi_n}\|_{L^2_x} \leq \left(\underbrace{\|b(0)\sqrt{\varphi_n}\|_{L^2_x}}_{=0} + C\frac{t}{\sqrt{n}}\right)e^{Ct} \xrightarrow{n\to\infty} 0\,,$$

hence $q_1 = q_2$ for positive times. The argument for negative times is analogous. $\square$

CHAPTER 3

Global Well-posedness of the NLS Hierarchy with Nonzero Boundary Conditions

Abstract

We consider the NLS hierarchy with the nonzero boundary conditions $q(t,x) \to q_\pm \in \mathbb{S}^1$ as $x \to \pm\infty$ and prove that it is globally well-posed for initial data of high regularity. Specifically, we prove well-posedness of the problem for the perturbation $p = q - q_*$ from a time-independent front q_* connecting q_- to q_+.

We present a new proof for the asymptotic expansion of the logarithmic derivative of the Jost solutions associated with the Lax operator, from which the equations in the NLS hierarchy are defined. Using the corresponding recurrence relation, we derive *new explicit formulas* for all terms in the NLS hierarchy with at most one factor that is q_x, $\overline{q}_x$, or a derivative thereof.

This allows us to view the equations for the perturbation p as dispersive nonlinear PDEs with explicit linear parts and sufficiently localized nonlinearities. They are part of a large class of systems for which we develop a local well-posedness theory in weighted Sobolev spaces. We establish local smoothing and maximal function estimates for general dispersion relations with finitely many critical points, which are of independent interest.

Finally, we globalize the solutions using the conserved energies constructed in [119, 120].

3.1. Introduction

Consider the defocusing Nonlinear Schrödinger (NLS) equation

$$iq_t + q_{xx} = 2q|q|^2 \tag{NLS}$$

for a wavefunction $q = q(t,x) : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{C}$. While NLS is most naturally associated with the zero boundary condition

$$\lim_{|x|\to\infty} q(t,x) = 0\,, \tag{ZBC}$$

we are interested in the nonzero boundary conditions

$$\lim_{x\to\pm\infty} e^{2it} q(t,x) = q_\pm \text{ where } |q_\pm| = 1\,. \tag{3.1.1}$$

Because solutions of NLS–(3.1.1) are not stationary at infinity, it is preferable to work instead with the so-called Gross-Pitaevskii (GP) equation

$$iq_t + q_{xx} = 2q(|q|^2 - 1)\,, \tag{GP}$$

which is formally equivalent to NLS under the transformation $q \mapsto e^{2it}q$. The boundary conditions (3.1.1) are replaced by

$$\lim_{x\to\pm\infty} q(t,x) = q_\pm \text{ where } |q_\pm| = 1\,. \tag{NZBC}$$

From now on, we write NLS for the system NLS–ZBC and GP for the system GP–NZBC, respectively. Before stating our main result, we discuss various aspects of NLS, GP, and the associated hierarchies.

On applications of NLS and GP. The systems NLS and GP, also in their focusing variants, have been intensely studied due to their ubiquity in the analysis of wave phenomena in various physical systems. We refer to the review papers [92] for applications in nonlinear optics and specifically high-speed communications, [23] for applications in Bose-Einstein condensation and plasma physics, and [24] for applications in the study of wave collapse.

We remark that GP is equivalent (under the Madelung transform, see [145] and Chapter 2) to a quantum hydrodynamical system, which is relevant in the study of Bose-Einstein condensation [73, 83], superfluidity [72, 127, 138], and quantum semiconductors [76].

Well-posedness results and function spaces in view of NZBC. We start with a brief review of some well-posedness results for NLS, focusing only on global results on $\mathbb{R}$ due to the breadth of the literature. For global well-posedness results in Sobolev spaces $H^s(\mathbb{R})$, we note, in descending regularity, [21] for $s=2$, [82] for $s=1$, [173] for $s=0$, and [49] for a priori bounds and existence of weak solutions with $s>-\frac{1}{12}$. In [121, 122, 123] a priori bounds are given down to $s>-\frac{1}{2}$. Besides the Sobolev scale, there are global well-posedness results in Fourier-Lebesgue spaces $\widehat{H}^s_r(\mathbb{R})$, with $s\geq 0$ [87], and modulation spaces $M^s_{p,q}(\mathbb{R})$, also with $s\geq 0$ [117, 148]. We also note [176], where global well-posedness is shown for a certain class of initial data with an infinite $L^2(\mathbb{R})$-norm. For the reader interested in other settings, we refer to the books [44, 170] as well as the recent program started in [96] by M. Ifrim and D. Tataru.

Because NZBC is not preserved under addition and scalar multiplication of functions, it is not compatible with traditional scales of function spaces, such as the Sobolev scale. As a result, the well-posedness of GP is more delicate. Spaces compatible with ZBC contain the trivial background solution $q=0$ to NLS, while spaces compatible with NZBC contain, for example, the trivial solution $q=1$ of GP, the stationary "black soliton" solution $q=\tanh$, or any of the time-dependent "dark soliton" solutions $q(t,x)=\mathrm{Re}[\zeta]+i\,\mathrm{Im}[\zeta]\tanh(\mathrm{Im}[\zeta](x+2\,\mathrm{Re}[\zeta]t))$, $\zeta\in\mathbb{S}^1$ to GP. Generally, for any $q_\pm\in\mathbb{S}^1$, there exists a dark soliton profile connecting q_- to q_+ (see (3.2.1) below for an explicit formula).

P. E. Zhidkov gave a first result in 1987 [191], establishing local well-posedness of GP in the Zhidkov space $Z^k(\mathbb{R}^n)$ with $n,k\in\mathbb{N}_{\geq 1}$, which is the closure of $C_c^\infty(\mathbb{R}^n)$ under the norm $\|\cdot\|_{Z^k(\mathbb{R}^n)}=\|u\|_{L^\infty(\mathbb{R}^n)}+\|\nabla_x u\|_{H^{k-1}(\mathbb{R}^n)}$. This implies global well-posedness in $Z^1(\mathbb{R})$ due to conservation of energy (see $\mathcal{H}_2^{\mathsf{GP}}$ below) in one space dimension. In [74] this is extended to $n=2,3$. A global result in the energy space was obtained by P. Gérard [77, 78] for $n=1,2,3$, and for $n=4$ under smallness assumptions. The smallness assumptions for $n=4$ are lifted in [113]. In [89] and preceding works, global stability and scattering results were obtained for small initial data in higher dimensions. We note [151] for a global well-posedness result permitting infinite energy in $1+H^s(\mathbb{R}^3)$, $s\in\left(\frac{5}{6},1\right)$, and [7] for more general nonlinearities in the cases $n=2,3$. Recently in [119, 120], when $n=1$, H. Koch and X. Liao proved the global well-posedness of GP in a complete metric space (X^s,d^s) with $s\geq 0$, which we define below in (3.1.15). They use the complete integrability to construct conserved energies that control the solution at every regularity $s\geq 0$ (see (3.4.35) below). We make use of these energies in our globalization argument in Section 3.4.2.

Complete integrability and the conserved Hamiltonians of NLS and GP. The systems NLS and GP are completely integrable in the sense that they admit Lax pairs. A Lax pair for a time-evolving PDE is a pair of operators L and P, depending on a time-dependent potential function, such that the Lax equation $\partial_t L=[P,L]=PL-LP$ holds true if and only if the potential is a solution to the PDE. For NLS and GP, it is shown in [189] that the operators

$$L=L^{\mathsf{NLS}}=L^{\mathsf{GP}}=i\begin{pmatrix}\partial_x & -q\\ \overline{q} & -\partial_x\end{pmatrix} \tag{3.1.2}$$

and

$$P^{\mathsf{NLS}}=i\begin{pmatrix}2\partial_x^2-q\overline{q} & -2q\partial_x-q_x\\ 2\overline{q}\partial_x+\overline{q}_x & -2\partial_x^2+q\overline{q}\end{pmatrix}\qquad P^{\mathsf{GP}}=i\begin{pmatrix}2\partial_x^2-q\overline{q}+1 & -2q\partial_x-q_x\\ 2\overline{q}\partial_x+\overline{q}_x & -2\partial_x^2+q\overline{q}-1\end{pmatrix} \tag{3.1.3}$$

yield Lax pairs. A consequence of the Lax pair formulation is that eigenvalues of the Lax operator are stationary, and furthermore that the time evolution of the scattering data (including the reflection coefficient) is characterized by a *linear* equation. One may try to recover the potential from the evolved scattering data. This is indeed possible for NLS and GP, where the so-called inverse scattering transform (IST) method can be applied. For NLS, the IST method was already well-developed in the seminal paper [2] (see also [34] and references therein for more recent work). Recently, attention toward the nonzero boundary data case of GP, where the IST method is harder to deploy, has grown, and significant progress has been made. We refer to the pioneering work [189] and the recent review [155]. An application of the IST method in our setting can be found in [149], where the authors prove orbital stability of the dark soliton solutions to the real mKdV hierarchy, i.e. the flows (NLS_{2n+1}) with real-valued q.

Another consequence of complete integrability is the existence of an infinite number of conserved quantities. For NLS, we denote these by $\mathcal{H}_n^{\mathsf{NLS}}$ and list the initial ones:

$$\mathcal{H}_0^{\mathsf{NLS}}(q) = \int_{\mathbb{R}} q\overline{q}\,\mathrm{d}x \qquad \text{(Mass)}$$

$$\mathcal{H}_1^{\mathsf{NLS}}(q) = -i\int_{\mathbb{R}} q\overline{q}_x\,\mathrm{d}x \qquad \text{(Momentum)}$$

$$\mathcal{H}_2^{\mathsf{NLS}}(q) = \int_{\mathbb{R}} q_x\overline{q}_x + q^2\overline{q}^2\,\mathrm{d}x \qquad \text{(Energy)}$$

$$\mathcal{H}_3^{\mathsf{NLS}}(q) = i\int_{\mathbb{R}} q\overline{q}_{xxx} - 4q^2\overline{q}\overline{q}_x - q_xq\overline{q}^2\,\mathrm{d}x$$

$$\mathcal{H}_4^{\mathsf{NLS}}(q) = \int_{\mathbb{R}} q_{xx}\overline{q}_{xx} - 6q^2\overline{q}\overline{q}_{xx} - 5q^2\overline{q}_x^2 - 6qq_x\overline{q}\overline{q}_x - qq_{xx}\overline{q}^2 + 2q^3\overline{q}^3\,\mathrm{d}x\,.$$

Using the energy, NLS can be written in Hamiltonian form

$$iq_t = \frac{\delta\mathcal{H}_2^{\mathsf{NLS}}(q)}{\delta\overline{q}}\,.$$

The Hamiltonian structure is commonly considered for a pair of variables (q,r), where the identification $r=\overline{q}$ is made later. For convenience, we use $\overline{q}$ from the beginning, writing $q\overline{q}$ instead of $|q|^2$ when we intend to highlight the functional dependencies.

The Hamiltonians $\mathcal{H}_n^{\mathsf{NLS}}$ are formally still conserved under the flow of GP, but the even ones $\mathcal{H}_{2n}^{\mathsf{NLS}}$ are ill-defined in the setting of NZBC. We use instead a renormalized sequence of Hamiltonians $\mathcal{H}_n^{\mathsf{GP}}$, which are conserved quantities for GP and compatible with NZBC. The initial ones are:

$$\mathcal{H}_0^{\mathsf{GP}}(q) = \int_{\mathbb{R}} q\overline{q} - 1\,\mathrm{d}x \qquad \text{(Mass)}$$

$$\mathcal{H}_1^{\mathsf{GP}}(q) = -i\int_{\mathbb{R}} q\overline{q}_x\,\mathrm{d}x \qquad \text{(Momentum)}$$

$$\mathcal{H}_2^{\mathsf{GP}}(q) = \int_{\mathbb{R}} q_x\overline{q}_x + (q\overline{q}-1)^2\,\mathrm{d}x \qquad \text{(Energy)}$$

$$\mathcal{H}_3^{\mathsf{GP}}(q) = i\int_{\mathbb{R}} q\overline{q}_{xxx} - 4q^2\overline{q}\overline{q}_x - q_xq\overline{q}^2 + 4q\overline{q}_x\,\mathrm{d}x$$

$$\begin{aligned}\mathcal{H}_4^{\mathsf{GP}}(q) = \int_{\mathbb{R}} & q_{xx}\overline{q}_{xx} - 6q^2\overline{q}\overline{q}_{xx} - 5q^2\overline{q}_x^2 - 6qq_x\overline{q}\overline{q}_x - qq_{xx}\overline{q}^2 \\ & - 6q_x\overline{q}_x + 2q^3\overline{q}^3 - 6q^2\overline{q}^2 + 6q\overline{q} - 2\,\mathrm{d}x\,.\end{aligned}$$

Using the energy, GP can be written in Hamiltonian form

$$iq_t = \frac{\delta\mathcal{H}_2^{\mathsf{GP}}(q)}{\delta\overline{q}}\,.$$

The Hamiltonians $\mathcal{H}_n^{\mathsf{NLS}}$ and $\mathcal{H}_n^{\mathsf{GP}}$ satisfy

$$\frac{\delta\mathcal{H}_{2m}^{\mathsf{GP}}(q)}{\delta\overline{q}} = \sum_{k=0}^{m}\binom{m-\frac{1}{2}}{m-k}(-4)^{m-k}\frac{\delta\mathcal{H}_{2k}^{\mathsf{NLS}}(q)}{\delta\overline{q}} \tag{3.1.4}$$

$$\frac{\delta\mathcal{H}_{2m+1}^{\mathsf{GP}}(q)}{\delta\overline{q}} = \sum_{k=0}^{m}\binom{m}{m-k}(-4)^{m-k}\frac{\delta\mathcal{H}_{2k+1}^{\mathsf{NLS}}(q)}{\delta\overline{q}}\,. \tag{3.1.5}$$

More specifically, the densities of $\mathcal{H}_n^{\mathsf{GP}}$ are affine linear combinations of the densities of $\mathcal{H}_n^{\mathsf{NLS}}$, and vice versa (see Remark 3.2.10 below, and also [71, Chapter 1, (10.25)]). In the setting of NZBC there is an additional nontrivial conserved quantity, which plays no significant role in this work:

$$\mathcal{H}_{-1}^{\mathsf{GP}}(q) = -i\log_{\mathrm{pr}}\left(\frac{q_-}{q_+}\right). \qquad \text{(Phase change)}$$

Here $\log_{\mathrm{pr}}$ denotes the principal logarithm.

We also consider another sequence of Hamiltonians $\mathcal{H}_n^{\widetilde{\mathsf{GP}}}$, defined by

$$\mathcal{H}_{2m}^{\widetilde{\mathsf{GP}}}(q) = \mathcal{H}_{2m}^{\mathsf{GP}}(q) \qquad \mathcal{H}_{2m+1}^{\widetilde{\mathsf{GP}}}(q) = \sum_{k=-1}^{m} \binom{-\frac{1}{2}}{m-k} 4^{m-k} 2^{-\mathbb{1}_{\{k=-1\}}} \mathcal{H}_{2k+1}^{\mathsf{GP}}(q)\,, \tag{3.1.6}$$

or equivalently

$$\mathcal{H}_{2m}^{\mathsf{GP}}(q) = \mathcal{H}_{2m}^{\widetilde{\mathsf{GP}}}(q) \qquad \mathcal{H}_{2m+1}^{\mathsf{GP}}(q) = 2^{\mathbb{1}_{\{m=-1\}}} \sum_{k=-1}^{m} \binom{\frac{1}{2}}{m-k} 4^{m-k} \mathcal{H}_{2k+1}^{\widetilde{\mathsf{GP}}}(q)\,.$$

Here the initial odd Hamiltonians $\mathcal{H}_{2n-1}^{\widetilde{\mathsf{GP}}}$ are

$$\begin{aligned}
\mathcal{H}_{-1}^{\widetilde{\mathsf{GP}}}(q) &= \frac{1}{2}\mathcal{H}_{-1}^{\mathsf{GP}}(q) \\
\mathcal{H}_{1}^{\widetilde{\mathsf{GP}}}(q) &= -i\int_{\mathbb{R}} q\overline{q}_x \,\mathrm{d}x - \mathcal{H}_{-1}^{\mathsf{GP}}(q) \\
\mathcal{H}_{3}^{\widetilde{\mathsf{GP}}}(q) &= i\int_{\mathbb{R}} q\overline{q}_{xxx} - 4q^2\overline{q}\overline{q}_x - q_x q\overline{q}^2 + 6q\overline{q}_x \,\mathrm{d}x + 3\mathcal{H}_{-1}^{\mathsf{GP}}(q)\,.
\end{aligned}$$

A relation similar to (3.1.4)–(3.1.5) holds between $\mathcal{H}_n^{\mathsf{NLS}}$ and $\mathcal{H}_n^{\widetilde{\mathsf{GP}}}$:

$$\frac{\delta \mathcal{H}_{2m}^{\widetilde{\mathsf{GP}}}(q)}{\delta \overline{q}} = \sum_{k=0}^{m} \binom{m-\frac{1}{2}}{m-k} (-4)^{m-k} \frac{\delta \mathcal{H}_{2k}^{\mathsf{NLS}}(q)}{\delta \overline{q}} \tag{3.1.7}$$

$$\frac{\delta \mathcal{H}_{2m+1}^{\widetilde{\mathsf{GP}}}(q)}{\delta \overline{q}} = \sum_{k=0}^{m} \binom{m+\frac{1}{2}}{m-k} (-4)^{m-k} \frac{\delta \mathcal{H}_{2k+1}^{\mathsf{NLS}}(q)}{\delta \overline{q}}\,. \tag{3.1.8}$$

The Hamiltonians $\mathcal{H}_n^{\mathsf{NLS}}$, $\mathcal{H}_n^{\mathsf{GP}}$ and $\mathcal{H}_n^{\widetilde{\mathsf{GP}}}$ appear naturally in asymptotic expansions of the logarithm of the transmission coefficient (see (3.2.50) and (3.2.51) below).

Definitions of the NLS, GP, and $\widetilde{\mathsf{GP}}$ hierarchies. It is natural to consider for $n \geq 0$ the infinite sequence of Hamiltonian PDEs

$$iq_t = \frac{\delta \mathcal{H}_n^{\mathsf{NLS}}(q)}{\delta \overline{q}}\,. \tag{3.1.9}$$

This is referred to as the NLS hierarchy. Some caution is needed, as sometimes only the flows where n is even are called the NLS hierarchy, while the flows where n is odd are called the complex mKdV hierarchy. The Hamiltonians $\mathcal{H}_n^{\mathsf{NLS}}$ are conserved quantities for every one of these flows. Equivalently, for any $n, m \geq 0$ we have

$$\left\{\mathcal{H}_n^{\mathsf{NLS}}, \mathcal{H}_m^{\mathsf{NLS}}\right\} := \frac{\delta \mathcal{H}_n^{\mathsf{NLS}}}{\delta q}\frac{\delta \mathcal{H}_m^{\mathsf{NLS}}}{\delta \overline{q}} - \frac{\delta \mathcal{H}_n^{\mathsf{NLS}}}{\delta \overline{q}}\frac{\delta \mathcal{H}_m^{\mathsf{NLS}}}{\delta q} = 0\,, \tag{3.1.10}$$

i. e. they pairwise Poisson commute (see [71, III.§2] or [118, Theorem B.7]). As a result, their Hamiltonian flows commute and we can consider instead for a single function $q = q(t_0, t_1, t_2, \ldots, x) : \mathbb{R}^{\mathbb{N}} \times \mathbb{R} \longrightarrow \mathbb{C}$ the infinite hierarchy of PDEs

$$i\partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\mathsf{NLS}}(q)}{\delta \overline{q}}\,. \tag{NLS_n}$$

In this paper, we denote the set of all nonnegative integers by $\mathbb{N}$. By an abuse of notation, we exclude 0 from $\mathbb{N}$ when the case $n = 0$ would obviously not make sense.

As is the case for NLS, solutions to (NLS_n) with nonzero boundary data are not stationary at infinity, and so it is preferable to work instead with the infinite sequence of Hamiltonian PDEs

$$iq_t = \frac{\delta \mathcal{H}_n^{\mathsf{GP}}(q)}{\delta \overline{q}}\,, \tag{3.1.11}$$

which we call the GP hierarchy.[1] Here we take $n \in \mathbb{N}$, excluding the case $n = -1$ as the flow generated by $\mathcal{H}_{-1}^{\mathsf{GP}}$ is the constant flow $q(t,x) = q(0,x)$. In the case $n = 0$ equation (3.1.11) is $iq_t = q$, whose solution is given by the rotation in phase $q(t,x) = e^{-it}q(0,x)$. All flows with $n \geq 1$ are compatible with NZBC in the sense that they are stationary at infinity (see also Corollary 3.1.7, (iii) below).

Of course, the same Poisson commutation relations $\{\mathcal{H}_n^{\mathsf{GP}}, \mathcal{H}_m^{\mathsf{GP}}\} = 0$ for $n, m \geq 0$ hold, so we can again consider for a single function $q = q(t_0, t_1, \dots, x)$ the infinite hierarchy of PDEs

$$i\partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\mathsf{GP}}(q)}{\delta \overline{q}}\,. \tag{GP_n}$$

Lastly, we define the $\widetilde{\mathsf{GP}}$ hierarchy as the infinite hierarchy of PDEs

$$i\partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\widetilde{\mathsf{GP}}}(q)}{\delta \overline{q}}\,. \tag{$\widetilde{\mathsf{GP}}_n$}$$

Our motivation to introduce the Hamiltonians $\mathcal{H}_n^{\widetilde{\mathsf{GP}}}$ is that later, the linear part of ($\widetilde{\mathsf{GP}}_{2m+1}$) is easier to work with than the linear part of (GP_{2m+1}).

On applications of the NLS, GP, and other integrable hierarchies. First and foremost, the study of any particular higher equation of an integrable hierarchy is also the study of every other equation in the hierarchy due to Noether's theorem: the flows generated by the higher Hamiltonians are symmetries for every other flow, in particular the usually eponymous initial PDE (NLS, KdV, etc.) of the hierarchy under consideration. As such, the well-posedness result for the NLS hierarchy that we present here represents the construction of an infinite number of symmetries for GP. Integrable equations like NLS and KdV appear frequently as amplitude equations for the long-wave regime of other equations (see e. g. [24, 163, 172, 183, 187]). In such cases, they may be modulated using the symmetries in their hierarchies (see [22] for an example involving higher symmetries). On another note, integrable hierarchies are also relevant in mathematical physics, in connection with algebraic geometry [67, 68].

Renormalization of hierarchies. Let us state clearly that taking affine linear combinations of the (densities of the) Hamiltonians in an integrable hierarchy yields flows which are essentially equivalent to those already in the hierarchy. We make this precise with a proposition.

Proposition 3.1.1 (Equivalence between hierarchies)**.** *Let $\mathbf{c} = (\mathbf{c}_{n,k}) \in \mathbb{C}^{\mathbb{N}\times\mathbb{N}}$ be an invertible infinite upper triangular matrix and set $\mathbf{t} = (t_0, t_1, \dots)$. Formally, $q(\mathbf{t}, x)$ solves (NLS_n) for all $n \in \mathbb{N}$ if and only if $u(\mathbf{t}, x) = q(\mathbf{ct}, x)$ solves the renormalized NLS hierarchy*

$$i\partial_{t_n} u = \sum_{k=0}^{n} \mathbf{c}_{k,n} \frac{\delta \mathcal{H}_k^{\mathsf{NLS}}(u)}{\delta \overline{u}}$$

for all $n \in \mathbb{N}$.

Proof. The Hamiltonians only depend on q, $\overline{q}$ and their spatial derivatives, but not on $\mathbf{t}$, so the proof follows from the chain rule. □

The relations (3.1.4)–(3.1.5) represent precisely such an equivalence between the NLS and GP hierarchies. Here $\mathbf{c} = \mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}$ where

$$\begin{aligned}
\mathbf{c}_{2k,2m}^{\mathsf{NLS}\mapsto\mathsf{GP}} &= \mathbb{1}_{\{m\geq k\}} \binom{m-\frac{1}{2}}{m-k} (-4)^{m-k} & \qquad \mathbf{c}_{2k+1,2m}^{\mathsf{NLS}\mapsto\mathsf{GP}} &= 0 \\
\mathbf{c}_{2k,2m+1}^{\mathsf{NLS}\mapsto\mathsf{GP}} &= 0 & \qquad \mathbf{c}_{2k+1,2m+1}^{\mathsf{NLS}\mapsto\mathsf{GP}} &= \mathbb{1}_{\{m\geq k\}} \binom{m}{m-k} (-4)^{m-k}\,,
\end{aligned} \tag{$\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}$}$$

and this matrix is invertible with its diagonal entries being 1 (with $(-4)^{m-k}$ replaced by 4^{m-k}). As an example, we have

$$(\mathbf{c}_{n,k}^{\mathsf{NLS}\mapsto\mathsf{GP}})_{0\leq n,k\leq 2} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

[1] Not to be confused with the "Gross–Pitaevskii hierarchy" from kinetic theory (see e. g. [70]).

so in the case $n = 2$ the energy of NLS is renormalized with the mass to yield the energy of GP. Similarly, the relations (3.1.7)–(3.1.8) correspond to the choice $\mathbf{c} = \mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}$ where

$$\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}_{2k,2m} = \mathbb{1}_{\{m\geq k\}}\binom{m-\frac{1}{2}}{m-k}(-4)^{m-k} \qquad \mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}_{2k+1,2m} = 0$$
$$\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}_{2k,2m+1} = 0 \qquad \mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}_{2k+1,2m+1} = \mathbb{1}_{\{m\geq k\}}\binom{m+\frac{1}{2}}{m-k}(-4)^{m-k}\,. \tag{$\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}$}$$

We can also derive $\mathbf{c} = \mathbf{c}^{\mathsf{GP}\mapsto\widetilde{\mathsf{GP}}}$ from the relation (3.1.6) to be

$$\mathbf{c}^{\mathsf{GP}\mapsto\widetilde{\mathsf{GP}}}_{2k,2m} = \mathbb{1}_{\{m=k\}} \qquad \mathbf{c}^{\mathsf{GP}\mapsto\widetilde{\mathsf{GP}}}_{2k+1,2m} = 0$$
$$\mathbf{c}^{\mathsf{GP}\mapsto\widetilde{\mathsf{GP}}}_{2k,2m+1} = 0 \qquad \mathbf{c}^{\mathsf{GP}\mapsto\widetilde{\mathsf{GP}}}_{2k+1,2m+1} = \mathbb{1}_{\{m\geq k\}}\binom{-\frac{1}{2}}{m-k}4^{m-k}2^{-\mathbb{1}_{\{k=-1\}}}\,, \tag{$\mathbf{c}^{\mathsf{GP}\mapsto\widetilde{\mathsf{GP}}}$}$$

with its inverse $\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}$ similarly defined, except $\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}_{2k+1,2m+1} = 2^{\mathbb{1}_{\{m=-1\}}}\binom{\frac{1}{2}}{m-k}4^{m-k}$. Then $q(\mathbf{t},x)$ solves the GP hierarchy (GP_n) if and only if $u(\mathbf{t},x) = q(\mathbf{c}^{\mathsf{GP}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t},x)$ solves the $\widetilde{\mathsf{GP}}$ hierarchy ($\widetilde{\mathsf{GP}}_n$), and correspondingly $q(\mathbf{t},x)$ solves the $\widetilde{\mathsf{GP}}$ hierarchy if and only if $u(\mathbf{t},x) = q(\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}\mathbf{t},x)$ solves the GP hierarchy.

These equivalences allow us to transfer well-posedness results among the NLS, GP, and $\widetilde{\mathsf{GP}}$ hierarchies.

3.1.1. Main result

We denote by $\mathcal{D}(\mathbb{R};\mathbb{C}) = C_c^\infty(\mathbb{R};\mathbb{C})$ the space of test functions. For a function $f \in \mathcal{D}(\mathbb{R};\mathbb{C})$, we define the Fourier transform

$$\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-ix\cdot\xi}f(x)\,\mathrm{d}x$$

and extend it to tempered distributions in the usual manner. Before stating our main result, we define for $s, s' \in \mathbb{N}$ the (weighted) Sobolev spaces

$$\begin{aligned}\big(H^s(\mathbb{R}),\|\cdot\|_{H^s(\mathbb{R})}\big) &= \big(\{u \in L^2(\mathbb{R}) : (1+|\xi|^2)^{\frac{s}{2}}\widehat{u}(\xi) \in L^2(\mathbb{R})\}, \|(1+|\xi|^2)^{\frac{s}{2}}\widehat{\cdot}(\xi)\|_{L^2(\mathbb{R})}\big)\\ \big(H^{s',1}(\mathbb{R}),\|\cdot\|_{H^{s',1}(\mathbb{R})}\big) &= \big(\{u \in H^{s'}(\mathbb{R}) : xu \in H^{s'}(\mathbb{R})\}, \|x\cdot\|_{H^{s'}(\mathbb{R})} + \|\cdot\|_{H^{s'}(\mathbb{R})}\big)\end{aligned} \tag{3.1.12}$$

and the energy functionals

$$E^s(q) = \||q|^2-1\|_{H^{s-1}(\mathbb{R})} + \|q_x\|_{H^{s-1}(\mathbb{R})} \qquad E^{s',1}(q) = \||q|^2-1\|_{H^{s'-1,1}(\mathbb{R})} + \|q_x\|_{H^{s'-1,1}(\mathbb{R})}\,. \tag{3.1.13}$$

In this paper, we do not search for optimal regularity assumptions, and shall always take the Sobolev regularity indices s, s' to be sufficiently large integers. For $\omega \in \mathbb{R}$ we use the notations $\lceil\omega\rceil = \min\{k \in \mathbb{Z} : k \geq \omega\}$ and $\lfloor\omega\rfloor = \max\{k \in \mathbb{Z} : k \leq \omega\}$.

Theorem 3.1.2 (Global well-posedness of each flow (GP_n) in the GP hierarchy)**.** *Let $n \geq 2$ and define $\omega = \frac{n-1}{2}$. Let $s, s' \in \mathbb{N}$ such that $s \geq 3n-1+\omega$ and $\frac{s+3\omega}{2} \leq s' \leq s-n$. Then ($\mathsf{GP}_n$) is globally well-posed in the following sense: For every $q_* = q_*(x) \in H^{s+1-\lceil\omega\rceil+n}_{\mathrm{loc}}$ with $E^{s+1-\lceil\omega\rceil+n,1}(q_*) < \infty$ and every $p_0 = p_0(x) \in H^s \cap H^{s',1}$, there exists a unique $p = p(t,x) \in C(\mathbb{R}; H^s \cap H^{s',1})$ with $p(0) = p_0$ such that $q_* + p(t)$ is a solution of (GP_n) in the sense of distributions. The solution is strong in the sense that p solves the corresponding perturbative formulation (see (3.1.21) below) strongly in $H^{s-n} \cap H^{s'-n,1}$. For every $T > 0$ the map*

$$\begin{aligned} H^s \cap H^{s',1} &\longrightarrow C_b([-T,T]; H^s \cap H^{s',1})\\ p_0 &\longmapsto p\end{aligned}$$

is Lipschitz continuous. Lastly, there exists $C = C(n, s, E^{s+1-\lceil\omega\rceil+n,1}(q_*), \|p_0\|_{H^s\cap H^{s',1}}) > 0$ such that for all $t \geq 0$ we have

$$\|p(t)\|_{L^\infty_x} + \|p_x(t)\|_{H^{s-1}_x} \leq C \qquad \|p(t)\|_{H^{s',1}_x} \leq Ce^{Ct}\,. \tag{3.1.14}$$

Remark 3.1.3 (On the $\widetilde{\mathsf{GP}}$ hierarchy)**.** *The same result holds for the $\widetilde{\mathsf{GP}}$ hierarchy. In fact, the proof is given for the $\widetilde{\mathsf{GP}}$ hierarchy because its perturbative formulation has a better linear part (see (3.1.25)).*

Remark 3.1.4 (On the proof)**.** *The difficulty lies in the nonzero boundary conditions, which is why we solve for a localized perturbation p of a time-independent front q_*. The main step of the proof is a study of the coefficients appearing in (GP_n) in order to show that the equation for p is part of a broad class of nonlinear dispersive PDEs (Proposition 3.1.14) for which we can show local well-posedness (Theorem 3.4.6) in $H^s \cap H^{s',1}$. In [119, 120] the authors construct conserved energies for GP that control the solution at every level of regularity in certain metric spaces. In particular, the energies $E^s(q)$ remain bounded. These energies are strong enough to prevent finite-time blow-up of the local solution p in $H^s \cap H^{s',1}$ (Lemma 3.4.8), allowing us to conclude global well-posedness. The reason that the $H^{s',1}$-norm of p is not globally bounded is that we cannot directly control it with the energies, instead having to use Grönwall's inequality in an energy estimate. An explicit example of the L^2-norm blow-up at infinity for the first flow in the GP hierarchy is given in (3.1.16).*

Remark 3.1.5 (On the energy assumption and NZBC)**.** *The assumption $E^{s+1-\lceil\omega\rceil+n,1}(q_*) < \infty$ implies that $(q_*)_x \in L^1$ and hence $\lim_{x\to\pm\infty} q_*(x) = q_\pm$ for some $q_\pm \in \mathbb{S}^1$, i. e. NZBC holds. Conversely, we can connect any end state pair (q_-, q_+) with a dark soliton profile (see (3.2.1)) that satisfies $E^{s+1-\lceil\omega\rceil+n,1}(q_*) < \infty$.*

One may be tempted to consider a more general finite-energy setting with $|q_*|^2 - 1 \in L^2$ and $(q_*)_x \in L^2$. This permits $(q_*)_x \notin L^1$, and hence solutions which do not have limits $q_\pm$ at $\pm\infty$ could be studied. However, it turns out that the well-posedness theory in Section 3.4 requires $|q_*|^2 - 1$ and $(q_*)_x$ in $H^{s-\lceil\omega\rceil+n,1}$ for some large s. This implies $(q_*)_x \in L^1$, so our method cannot consider boundary conditions more general than NZBC.

Remark 3.1.6 (On the time-independence of q_* and stability of solitons)**.** *When $q_\pm = \pm 1$, we can take q_* to be the black soliton solution $\tanh(x)$ of the Gross-Pitaevskii equation (GP) (i. e. (GP_n) for $n = 2$). Note that our result does not show the stability of the black soliton solution. We refer to [56] for the asymptotic behavior of the perturbed solution under further decay assumptions on p_0.*

It would be natural to choose q_* as a moving front, for example a dark soliton, and attempt to study its stability using the perturbative formulation for p. Currently we choose q_* to be time-independent, which causes terms depending only on q_* to appear on the right-hand side in (3.1.24) (this also causes $g \neq 0$ in (3.4.2)). Removing these terms would lead to minor improvements of certain estimates and regularity assumptions. Since our well-posedness theory can handle these inhomogeneous terms, and for the sake of simplicity, we choose q_* to be time-independent.

Recall the invertible matrices in ($\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}$) and ($\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}$), and define $\mathbf{c}^{\mathsf{GP}\mapsto\mathsf{NLS}} = (\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}})^{-1}$ and $\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} = (\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}})^{-1}$. For the reader interested in a formulation of well-posedness explicitly for the NLS *hierarchy*, i. e. for a function $q = q(t_0, \dots, t_N, x)$ with several time variables that solves (NLS_n) for all $0 \le n \le N$, we draw a corollary. Note that we use the notation $\mathbf{t} = (t_0, \dots, t_N), \mathbf{0} = (0, \dots, 0) \in \mathbb{R}^{N+1}$.

Corollary 3.1.7 (Global well-posedness of the N-truncated NLS, GP, and $\widetilde{\mathsf{GP}}$ hierarchies)**.** *Let $N \ge 2$ and define $\omega = \frac{N-1}{2}$. Let $s, s' \in \mathbb{N}$ such that $s \ge 3N - 1 + \omega$ and $\frac{s+3\omega}{2} \le s' \le s - N$. For every $q_0 : \mathbb{R} \to \mathbb{C}$ of the form $q_0 = q_* + p_0$, where $q_* \in H^{s+1-\lceil\omega\rceil+N}_{\mathrm{loc}}$ with $E^{s+1-\lceil\omega\rceil+N,1}(q_*) < \infty$ and $p_0 \in H^s \cap H^{s',1}$, there exists a unique function $q = q(\mathbf{t}, x) : \mathbb{R}^{N+1} \times \mathbb{R} \to \mathbb{C}$ with $q(\mathbf{0}) = q_0$ which satisfies the following:*

- (i) *For every $n \in \{0, \dots, N\}$ the functions $q(\mathbf{t}, x)$, $q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t}, x)$, and $q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t}, x)$ respectively solve (NLS_n), (GP_n), and ($\widetilde{\mathsf{GP}}_n$) in the sense of distributions.*
- (ii) *We have $q \in C_b(\mathbb{R}^{N+1}; X^s)$ and*

$$\sup_{\mathbf{t}\in\mathbb{R}^{N+1}} E^s(q(\mathbf{t})) \le C(N, s, E^s(q_0)) \,.$$

Here X^s stands for the complete metric space (X^s, d^s), $s \geq 0$ defined in [119, 120]

$$X^s = \{q \in H^s_{loc}(\mathbb{R}) : E^s(q) < \infty\}/\,\mathbb{S}^1 \tag{3.1.15}$$
$$d^s(q,p) = \left(\int_{\mathbb{R}} \inf_{\lambda\in\mathbb{S}^1} \|\operatorname{sech}(\cdot - y)(\lambda p - q)\|^2_{H^s}\,\mathrm{d}y\right)^{\frac{1}{2}}.$$

(iii) *The function $p = q - e^{-i(\mathbf{c}^{\mathsf{GP}\mapsto\mathsf{NLS}}\mathbf{t})_0} q_*$ satisfies $p \in C(\mathbb{R}^{N+1}; H^s \cap H^{s',1})$. Furthermore, $p(\mathbf{t})$, $p(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})$, and $p(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t})$ solve the respective perturbative formulations of (NLS_n), (GP_n), and ($\widetilde{\mathsf{GP}}_n$) (see Proposition 3.1.14 below), for the fronts $e^{-i(\mathbf{c}^{\mathsf{GP}\mapsto\mathsf{NLS}}\mathbf{t})_0} q_*$, $e^{-it_0}q_*$, and $e^{-it_0}q_*$, strongly in $H^{s-n} \cap H^{s'-n,1}$.*

Remark 3.1.8. *Notice that the zeroth* GP *flow*

$$i\partial_{t_0} q = \frac{\delta\mathcal{H}^{\mathsf{GP}}_0(q)}{\delta\overline{q}} = q$$

is simply a phase rotation, which changes a front q_* into $e^{-it_0}q_*$:

$$q_*(x) + p_0(x) \overset{(\mathsf{GP}_0)}{\longmapsto} e^{-it_0}q_*(x) + e^{-it_0}p_0(x).$$

This results in the change of the profile in (iii) above. The first GP *flow*

$$i\partial_{t_1} q = \frac{\delta\mathcal{H}^{\mathsf{GP}}_1(q)}{\delta\overline{q}} = i\partial_x q$$

is a transport equation, which transports a front q_* to the left and preserves its limits at infinity:

$$q_*(x) + p_0(x) \overset{(\mathsf{GP}_1)}{\longmapsto} q_*(x+t_1) + p_0(x+t_1).$$

The nth GP *flow with $n \geq 2$, by Theorem 3.1.2, preserves also the limits of a front for any finite time:*

$$q_*(x) + p_0(x) \overset{(\mathsf{GP}_n)}{\longmapsto} q_*(x) + p(t_n, x), \quad p(t_n,x) \in C(\mathbb{R}; H^s \cap H^{s',1}).$$

Note, however, that we have not shown uniform-in-time control of the norm $\|p(t_n)\|_{H^s\cap H^{s',1}}$ in Theorem 3.1.2. Observe that for the first GP *flow the L^2_x-norm of the t_1-dependent perturbation of the front $q_* = \tanh$ blows up at infinity:*

$$\|\tanh(x+t_1) - \tanh(x)\|_{L^2_x} \to \infty \text{ as } t_1 \to \infty\,. \tag{3.1.16}$$

The reason is that $\tanh(x+t_1) - \tanh(x) \geq \frac{1}{2}$ on the interval $[-t_1, 0]$ for sufficiently large t_1, while the perturbation derivative is uniformly bounded in time due to $\tanh' = \operatorname{sech}^2 \in \mathcal{S}$.

Remark 3.1.9. *In principle, solutions with infinitely many nonzero time variables can also be constructed, given q_0 satisfies the assumptions of Corollary 3.1.7 for every $N \in \mathbb{N}$, but this requires a non-trivial limit argument which we do not wish to study here.*

3.1.2. Organization of the paper and further results

All sections and appendices are largely self-contained and can in principle be read in any order.

3.1.2.1. Overview

In Subsection 3.1.2.2 we motivate Section 3.2, which is not directly relevant for the proof of the main result and instead is concerned with the rigorous definition and asymptotic expansion of the logarithm of the transmission coefficient associated with the Lax operator L (see (3.1.2)). In addition, we review some of the literature on the Inverse Scattering Transform (IST) method for NLS/GP with a nonzero background and elaborate on our contribution. Section 3.2 concludes with the derivation of the well-known recurrence relation (3.2.47), which is fundamental for the computations in Section 3.3.

In Subsection 3.1.2.3 we discuss Section 3.3, which contains an essential result on the structure of the NLS and GP hierarchies. Using generating functions, we are able to give explicit formulas for all terms in the hierarchies containing at most one factor that is q_x, $\overline{q}_x$, or a derivative thereof. Section 3.3 concludes with the proof of Proposition 3.1.14, which is a perturbative formulation of the GP hierarchy that fits into the well-posedness theory developed in Section 3.4.

In Subsection 3.1.2.4 we discuss Section 3.4. We begin with a review of different approaches to the well-posedness of the NLS and KdV hierarchies. Subsequently, we motivate our choice to use techniques developed by C. E. Kenig, G. Ponce and L. Vega in the 1990s. The key result is the local well-posedness of a large class of dispersive nonlinear systems, of which the perturbative formulation of the GP hierarchy is a special case. Section 3.4 concludes with the proof of Theorem 3.1.2, our main result.

3.1.2.2. Section 2: Direct scattering theory revisited and further developed

In Section 3.2 we recall some aspects of the Lax operator L, associated with GP, in the setting of NZBC, with the goal of rigorously defining the transmission coefficient $a(\lambda)$ and proving that its logarithm has asymptotic expansions in powers of the spectral parameters λ and $z = \sqrt{\lambda^2 - 1}$ at infinity. For λ and z in the first quadrant of the complex plane, we define Hamiltonians $\mathcal{H}_n^{\mathsf{NLS}}$, $\mathcal{H}_n^{\mathsf{GP}}$, and $\mathcal{H}_n^{\widetilde{\mathsf{GP}}}$ as the following expansion coefficients (see (3.2.50)–(3.2.51)):

$$\log a(\lambda) \sim i\sum_{n=0}^{\infty} \frac{\mathcal{H}_n^{\mathsf{NLS}}}{(2\lambda)^{n+1}} \qquad \log a(\lambda) \sim i\sum_{n=-1}^{\infty} \frac{\mathcal{H}_n^{\mathsf{GP}}}{(2z)^{n+1}} \sim i\sum_{n=0}^{\infty} \frac{\mathcal{H}_{2n}^{\widetilde{\mathsf{GP}}}}{(2z)^{2n+1}} + i\frac{\lambda}{z}\sum_{n=0}^{\infty} \frac{2^{\mathbb{1}_{\{n=0\}}}\mathcal{H}_{2n-1}^{\widetilde{\mathsf{GP}}}}{(2z)^{2n}}$$

in the different functional settings (ZBC) and (NZBC). Their densities satisfy certain recurrence relations (see Lemma 3.2.8 below), which are the basis of our calculations in Section 3.3. The recurrence relation for the densities of $\mathcal{H}_n^{\mathsf{NLS}}$ is well-known and also appears in the seminal paper [189] by V. E. Zakharov and A. B. Shabat, which initiated the study of the IST method for GP with NZBC. We refer to [30, 56, 63, 64, 65, 142] and the review [155] for further development of the IST method for NLS-type integrable equations with nonzero boundary data.

Note that most of these references are concerned with the formal and rigorous establishment of the inverse scattering transform, but not the rigorous asymptotic expansion of the logarithm of the transmission coefficient. The standard reference for this is the monograph [71, I.§3-§8], where first the periodic case on an interval of length L is considered, and then the limit $L \to \infty$ is taken. More recently in [119, 120], the renormalized transmission coefficient is defined and asymptotically expanded in the general finite-energy setting $E^s(q) < \infty$ for $s \geq 0$ (see (3.1.13)), by finding a change of variables for the Zakharov-Shabat scattering problem that essentially causes only the quantities $|q|^2 - 1$ and q_x to appear.

We believe there is expository value in a direct path from the Zakharov-Shabat scattering problem to the rigorous expansion of the transmission coefficient, using the classical approach described in [71], but not working with the periodic case first. Specifically, we use the integral representation ansatz (3.2.12), which to the best of our knowledge is not present in the literature. In order to explain, we consider the two Jost solutions $\Phi_1^{\pm}(x,\lambda)$, with $x \in \mathbb{R}$ and $\lambda \in \mathbb{C}$, to the Zakharov-Shabat scattering problem associated with the Lax operator (3.1.2)

$$L\Phi_1^{\pm}(x,\lambda) = \lambda\Phi_1^{\pm}(x,\lambda)\,, \quad \text{or equivalently} \quad \partial_x\Phi_1^{\pm}(x,\lambda) = \begin{pmatrix} -i\lambda & q(x) \\ \overline{q}(x) & i\lambda \end{pmatrix}\Phi_1^{\pm}(x,\lambda)\,, \tag{ZS}$$

for a potential of interest q which connects two limit states $q_{\pm} \in \mathbb{S}^1$. Then we search for an appropriate reference potential $q_* \in C^{\infty}(\mathbb{R};\mathbb{R})$ with the following properties

- $q_* \to q_{\pm}$ as $x \to \pm\infty$
- The Jost solutions $\Phi_{*,1}^{\pm}$ are known.

Fortunately such q_* exists (see the dark soliton profile (3.2.1)), and we now restrict ourselves to the functional setting $q \in q_* + \mathcal{S}(\mathbb{R};\mathbb{R})$.

We make the triangular representation ansatz

$$\Phi_1^{\pm}(x,\lambda) = \Phi_{*,1}^{\pm}(x,\lambda) + \int_{\pm\infty}^{x} \Gamma^{\pm}(x,y)\Phi_{*,1}^{\pm}(y,\lambda)\,\mathrm{d}y\,, \tag{3.1.17}$$

where $\Gamma^{\pm}(x,y)$ is a kernel that is independent of λ. Such triangular representations are well-known in the constant-potential cases $q_* = q_-$ or $q_* = q_+$ (see e. g. [65, 71, 141, 189, 149]). Since the kernel $\Gamma^{\pm}$ is independent of λ, these integral representations are suitable for the asymptotic analysis in the parameter λ of the Jost solutions via integration by parts. In fact, the density of the difference of the logarithm of the transmission coefficient $\log a(\lambda)$ and $\log(q_-\overline{q}_+)$ is $\partial_x \log \Psi_{1,1}^-$, i. e. the logarithmic derivative of the first component of the modified Jost solution $\Psi_1^- = \Phi_1^- e^{ixz}$ (see (3.2.30) below). Therefore, we can use (3.1.17) to derive the desired asymptotic expansion for $\log a(\lambda)$, as long as the error terms can be controlled in L_x^1. An essential source of this desired integrability when working with (3.1.17) is the difference $q - q_* \in \mathcal{S}$.

We comment on the existence and decay properties of the kernels $\Gamma^{\pm}$. It is important for $\Phi_{*,1}^{\pm}$ to be a solution of the Zakharov-Shabat scattering problem (ZS) in order for $\Gamma^{\pm}(\cdot,\cdot)$ to solve a *boundary value problem* which does not involve λ (see the system (3.2.15)–(3.2.17) below, which generalizes [71, Chapter 1, (8.18)–(8.19)]). We subsequently state and prove a well-posedness result for $\Gamma^{\pm}$ in Lemma 3.2.2, which we believe to be of independent interest. The decay estimates in Lemma 3.2.2 allow us to deduce the desired expansion of the transmission coefficient in Lemma 3.2.5.

Note that we work in the setting where q and q_* are smooth, and $q - q_*$ is Schwartz. The reader interested in weaker assumptions may adapt the proof for this purpose. We see potential for the triangular representation (3.1.17) to be useful in other aspects of the IST method, such as WKB expansions of various scattering data. Furthermore, it may be of use in the setting with asymmetric boundary conditions (i. e. $|q_-| \neq |q_+|$), when a Jost solution $\Phi_{*,1}^-$ for a reference potential q_* is known.

3.1.2.3. Section 3: Explicit formulas for parts of the hierarchies and the perturbative setting

In Section 3.3 we determine the coefficients of all terms in the NLS hierarchy which have at most one factor that is q_x, $\overline{q}_x$, or a derivative thereof. When using a perturbative ansatz $q(t,x) = q_*(x) + p(t,x)$, these are precisely the terms which we need to know in order to determine the linear part of the PDEs for p.

To describe the nonlinear terms in the PDEs in a quantitative way, we introduce the notation $\mathcal{O}_A^{M,L}(B)$, where $M, L \in \mathbb{N}$ and A, B are sets of formal complex-valued functions. We define the sets of formal functions

$$A \cup \overline{A} = \{f : f \in A \text{ or } \overline{f} \in A\} \qquad [\![B \cup \overline{B}]\!] = \{\partial_x^n f : f \in B \text{ or } \overline{f} \in B, n \in \mathbb{N}\}$$

and the set $\mathcal{O}_A^{M,L}(B)$ consists of sums of products with at least M factors from the set $[\![B \cup \overline{B}]\!]$, and at most L derivatives in total:

$$\mathcal{O}_A^{M,L}(B) = \left\{ \sum_{k=M}^{K} \sum_{\substack{l=(l_1,\dots,l_k)\in\mathbb{N}^k \\ |l|\le L}} \sum_{\substack{h_1,\dots,h_k \in A\cup\overline{A} \\ \partial_x^{l_1}h_1,\dots,\partial_x^{l_M}h_M \in [\![B\cup\overline{B}]\!]}} c_{h_1,\dots,h_k}^{k,l} \prod_{j=1}^{k} \partial_x^{l_j} h_j : K \in \mathbb{N}, c_{h_1,\dots,h_k}^{k,l} \in \mathbb{C} \right\}. \tag{3.1.18}$$

We write $f = \mathcal{O}_A^{M,L}(B)$ if $f \in \mathcal{O}_A^{M,L}(B)$. If any of the parameters of $\mathcal{O}$ are missing, we set by default M, L to zero and A, B to the empty set.

Theorem 3.1.10 (Structure of the NLS, GP, and $\widetilde{\mathsf{GP}}$ hierarchies)**.** *The following hold true.*

(i) *Let $m = \lfloor \frac{n}{2} \rfloor \geq 1$. The equations ($\mathsf{NLS}_n$) of the NLS hierarchy have the form*

$$i\partial_{t_{2m}} q = \sum_{j=0}^{m-2} J_{2m+1,j}\, q^{j+2}\overline{q}^j (i\partial_x)^{2m-2-2j}\overline{q} - \sum_{j=0}^{m-1} K_{2m+1,j}\, q^j \overline{q}^j (i\partial_x)^{2m-2j} q \tag{NLS_{2m}}$$

$$+ (m+1)C_m\overline{q}^m q^{m+1} + \mathcal{O}_q^{2,n-2}(q_x)$$

$$i\partial_{t_{2m+1}} q = \sum_{j=0}^{m} K_{2m+2,j} q^j \overline{q}^j (i\partial_x)^{2m+1-2j} q + \mathcal{O}_q^{2,n-2}(q_x)\,. \tag{NLS_{2m+1}}$$

In the above, the real-valued coefficients $J_{k,j}$, $K_{k,j}$ are defined as follows. Let $[u^j]$ denote the functional that extracts the coefficient in front of u^j from a formal power series in the symbol u, e. g. $[u^2](1+u)^3 = 3$. With this notation, we define

$$J_{k,j} = \begin{cases} 0 & k \text{ even} \\ [u^j]2(1+4u)^{\frac{k-2}{2}} & k \text{ odd} \end{cases} \qquad K_{k,j} = \begin{cases} [u^j](1+4u)^{\frac{k-1}{2}} & k \text{ even} \\ [u^j](-1-2u)(1+4u)^{\frac{k-2}{2}} & k \text{ odd} \end{cases}. \tag{3.1.19}$$

The coefficients C_m are the Catalan numbers, defined by the recurrence relation: $C_0 = 1$, $C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}$ (see also (3.2.45) below).

The nonlinear term $\mathcal{O}_q^{2,n-2}(q_x)$, as defined in (3.1.18), refers to the class of polynomial expressions in q, $\overline{q}$, and their derivatives, for which each monomial has at most $n-2$ derivatives in total, and at least two factors are q_x, $\overline{q}_x$, or derivatives thereof.

(ii) *For $n \geq 3$ with $m = \lfloor \frac{n}{2} \rfloor \geq 1$ the equations of the* GP *hierarchy* (GP_n) *have the form*

$$i\partial_{t_{2m}} q = \big((i\partial_x)^{2m} + 2(i\partial_x)^{2m-2}\big)q + 2q^2(i\partial_x)^{2m-2}\overline{q} + \mathcal{O}^{2,n-2}_{q,|q|^2-1}(q_x, |q|^2-1) \tag{GP_{2m}}$$

$$i\partial_{t_{2m+1}} q = \sum_{j=0}^{m} \binom{\frac{1}{2}}{m-j} 4^{m-j}(i\partial_x)^{2j+1} q + \mathcal{O}^{2,n-2}_{q,|q|^2-1}(q_x, |q|^2-1)\,, \tag{GP_{2m+1}}$$

while equations of the $\widetilde{\mathsf{GP}}$ *hierarchy* ($\widetilde{\mathsf{GP}}_n$) *have the form*

$$i\partial_{t_{2m}} q = \big((i\partial_x)^{2m} + 2(i\partial_x)^{2m-2}\big)q + 2q^2(i\partial_x)^{2m-2}\overline{q} + \mathcal{O}^{2,n-2}_{q,|q|^2-1}(q_x, |q|^2-1) \tag{$\widetilde{\mathsf{GP}}_{2m}$}$$

$$i\partial_{t_{2m+1}} q = (i\partial_x)^{2m+1} q + \mathcal{O}^{2,n-2}_{q,|q|^2-1}(q_x, |q|^2-1)\,. \tag{$\widetilde{\mathsf{GP}}_{2m+1}$}$$

Recall that for $n = 2m$ the equations ($\widetilde{\mathsf{GP}}_{2m}$) = ($\mathsf{GP}_{2m}$) are identical. Here $\mathcal{O}^{2,n-2}_{q,|q|^2-1}(q_x, |q|^2-1)$, defined in (3.1.18), refers to the class of polynomial expressions in q, $\overline{q}$, $|q|^2-1$, and their derivatives, for which each monomial has at most $n-2$ derivatives in total, and at least two factors are q_x, $\overline{q}_x$, $|q|^2-1$, or derivatives thereof.

Remark 3.1.11. *To the best of our knowledge these coefficients are not known in the literature. For further efforts to calculate explicit coefficients in integrable hierarchies, we refer to [4], where the coefficients of all cubic terms in the* dNLS *hierarchy are calculated, and also [19] for the* KdV *hierarchy. The general theory in [66] may also be of use.*

Remark 3.1.12. *The equation* (NLS)=(NLS_2) *corresponds to* (NLS_{2m}) *with $m = 1$, $n = 2$ and with $\mathcal{O}_q^{2,n-2}(q_x)$ replaced by 0. However,* (GP)=(GP_2) *is excluded from the general form* (GP_{2m}) *because there is a mismatch in the coefficient before the term q. Nevertheless,* (GP) *can be formulated in the perturbative form (3.1.21) below, with an additional term $2(|q_*|^2-1)q_*$ when $n = 2$, which is included in $\mathcal{O}^{1,0}_{q_*,|q_*|^2-1}(|q_*|^2-1)$ in (3.1.24).*

Remark 3.1.13. *We present an elegant reformulation of (i), whose proof is found at the end of Section 3.3. We have as $|\lambda| \to \infty$ the asymptotic expansion*

$$\frac{\delta \log a(\lambda)}{\delta \overline{q}} \sim \sum_{n=0}^{\infty} \frac{\mathcal{O}_q^{2,n}(q_x)}{(2i\lambda)^{n+1}} \tag{3.1.20}$$

$$+ \frac{i}{2(\lambda^2 - q\overline{q})^{\frac{1}{2}}} \left(q + \frac{2\lambda + i\partial_x}{4(\lambda^2 - q\overline{q}) + \partial_x^2}[iq_x] + \frac{1}{2(\lambda^2 - q\overline{q})(4(\lambda^2 - q\overline{q}) + \partial_x^2)}[q\partial_x^2(\lambda^2 - q\overline{q})] \right).$$

This is a formal statement, which is to be understood in the sense of expansion in λ by the geometric series and subsequent comparison of orders. Note that ∂_x^2 and the multiplication operator $q\overline{q}$ commute up to terms in $\mathcal{O}_q^{2,1}(q_x)$. The square brackets denote operator application.

We would like to point out that (3.1.20) might "by chance" describe more terms in the hierarchy than (NLS_{2m})–(NLS_{2m+1}) *do. To explain, observe that different arrangements of the order of* ∂_x^2 *and multiplication by* $q\overline{q}$ *in the geometric series expansion of the resolvent yield different residual terms in* $(2i\lambda)^{-n-1}\mathcal{O}_q^{2,n}(q_x)$*. One may hope that by choosing the correct ordering the residual terms can vanish, such that the formula above would become exact. We have concluded, after testing various natural arrangements, that this does not seem to be the case.*

Given a choice of q_*, we can now write down a dispersive PDE for $p = q - q_*$.

Proposition 3.1.14 (Perturbative formulation of the GP and $\widetilde{\mathsf{GP}}$ hierarchies)**.** *Let* $q(t,x) = q_*(x) + p(t,x)$*, where* $\lim_{x\to\pm\infty} q_*(x) = q_\pm \in \mathbb{S}^1$*. Let* $n \geq 2$ *and* $m = \lfloor \frac{n}{2} \rfloor$*. We consider the extended system for* $\mathbf{p} = (\mathbf{p}_j)_{1\leq j\leq 2} = (p, \overline{p})$*. Then the following hold.*

(i) *q solves* (GP_n) *if and only if* $\mathbf{p}$ *solves*

$$i\partial_{t_n}\mathbf{p} = \mathcal{L}^n[\mathbf{p}] + \mathcal{N}^n(\mathbf{p})\,, \tag{3.1.21}$$

where with $D_x = -i\partial_x$ *we have the linear operator*

$$\mathcal{L}^{2m} = D_x^{2m-2}\begin{pmatrix} D_x^2+2 & 2q_*^2 \\ -2\overline{q}_*^2 & -D_x^2-2 \end{pmatrix} \tag{3.1.22}$$

$$\mathcal{L}^{2m+1} = -\sum_{k=0}^{m}\binom{\frac12}{m-k}4^{m-k}D_x^{2k+1}\,, \tag{3.1.23}$$

and for $j \in \{1,2\}$ *we have*

$$(\mathcal{N}^n[\mathbf{p}])_j = \mathcal{O}^{1,0}_{q_*,|q_*|^2-1}(|q_*|^2-1) + \mathcal{O}^{1,n}_{q_*}((q_*)_x) + \mathcal{O}^{2,n-2}_{p,\overline{p},q_*,|q_*|^2-1}(p,(q_*)_x,|q_*|^2-1) \tag{3.1.24}$$

where the notation $\mathcal{O}$ *is defined in* (3.1.18).

(ii) *If q solves* ($\widetilde{\mathsf{GP}}_n$) *instead, then the above still holds with the only difference being an improved linear part when* $n = 2m+1$ *is odd:*

$$\mathcal{L}^{2m+1} = -D_x^{2m+1}\,. \tag{3.1.25}$$

(iii) *Assume* $n = 2m$*. Let* $R > 0$ *and* $\eta \in C_c^\infty(\mathbb{R};[0,1])$ *with* $\eta|_{[-R,R]} = 1$ *and* $\operatorname{supp}\eta \subseteq [-2R, 2R]$*. We change to a new variable*

$$\mathbf{u} = \mathcal{T}^{-1}\mathbf{p} \quad \text{where} \quad \mathcal{T} = 1 + B\psi(D_x) \quad \text{and} \quad \psi(\xi) = \frac{1-\eta(\xi)}{\xi^2+2} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1-q_*^2 \\ 1-\overline{q}_*^2 & 0 \end{pmatrix}.$$

Then (3.1.21) *can be rewritten, using modified nonlinear parts* $\mathcal{N}^n$ *satisfying* (3.1.24)*, as*

$$\begin{aligned} i\partial_{t_{2m}}\mathbf{u} = {} & D_x^{2m-2}\begin{pmatrix} D_x^2+2 & 2 \\ -2 & -D_x^2-2 \end{pmatrix}\mathbf{u} \\ & + \mathcal{T}^{-1}\Big(\mathcal{N}^n(\mathcal{T}\mathbf{u}) + \sum_{\gamma=-1}^{0}\big(\mathcal{N}^n(D_x^\gamma\psi(D_x)\mathbf{u}) + \mathcal{N}^n(D_x^\gamma(1-\eta(D_x))\mathbf{u})\big) - [\sigma_3, B]D_x^{2m-2}\eta(D_x)\mathbf{u}\Big)\,. \end{aligned} \tag{3.1.26}$$

Remark 3.1.15. *As an example, in the case* $n = 2$*,* $m = 1$ *and* $q_* = 1$ *the system* (3.1.22) *reads*

$$\begin{aligned} i\partial_t p &= (D_x^2+2)p + 2\overline{p} + 2p^2\overline{p} + 4p\overline{p} + 2p^2 \\ i\partial_t \overline{p} &= -2p + (-D_x^2-2)\overline{p} - 2\overline{p}^2 p - 4\overline{p}p - 2\overline{p}^2\,. \end{aligned}$$

We refer to [89] for a specific study of this system, also in higher dimensions.

Remark 3.1.16. *For the even flows we switch to the variable* $\mathbf{u}$ *in order to eliminate the non-constant coefficient* q^2 *that can be seen in* (GP_{2m})*, and would otherwise appear in the linear part of the equation for* p*. Then the key point is that the linear part is benign, and that each term in the "nonlinear" part has sufficient integrability, in the sense that it either has two factors which are* p *or* $\overline{p}$ *or derivatives thereof, or a coefficient which provides integrability, e. g.* $|q_*|^2 - 1$*,* $(q_*)_x$*. The extra term* $[\sigma_3, B]D_x^{2m-2}\eta(D_x)$ *is of order zero, which is the key improvement over* (3.1.21)*.*

3.1.2.4. SECTION 4: LOCAL WELL-POSEDNESS OF A LARGE CLASS OF NONLINEAR DISPERSIVE SYSTEMS INCLUDING THE $\widetilde{\mathsf{GP}}$ HIERARCHY

In Section 3.4 we construct a local well-posedness theory for a large class of dispersive nonlinear systems of PDEs that includes the perturbative formulation of the $\widetilde{\mathsf{GP}}$ hierarchy in Proposition 3.1.14. We then deduce global well-posedness in Theorem 3.1.2 by using the conserved energies constructed in [119, 120].

Before we explain our argument in detail, let us review existing well-posedness results for integrable hierarchies and general dispersive nonlinear equations. Due to their nature, when working with integrable hierarchies, one expects to have available a plethora of conserved quantities. Nevertheless, it is not a trivial matter to construct *useful* conserved quantities, especially at low regularity (see for example [119, 120, 122] and [91, 115, 158]). We work with high, integer regularity, and for us *local* well-posedness is the essential problem. Here it is not easy to derive a benefit from the integrability of the equation under consideration. As such, the local well-posedness of large, not necessarily integrable, classes of dispersive nonlinear equations may be studied, and in turn integrable hierarchies serve as natural applications for such theories.

We start with the KdV hierarchy, which we define for a function $u = u(t_0, t_1, \dots, x) : \mathbb{R}^{\mathbb{N}} \times \mathbb{R} \to \mathbb{R}$ as the infinite sequence of PDEs

$$\partial_{t_n} u = \frac{1}{2}\partial_x \frac{\delta \mathcal{E}_n^{\mathsf{KdV}}(u)}{\delta u} \qquad \mathcal{E}_n^{\mathsf{KdV}}(u) = (-1)^n \int_{\mathbb{R}} \sigma_{2n+3}^{\mathsf{KdV}}(u)\,\mathrm{d}x \tag{KdV_n}$$

$$\sigma_0^{\mathsf{KdV}}(u) = 0 \qquad \sigma_1^{\mathsf{KdV}}(u) = -u \qquad \sigma_{n+1}^{\mathsf{KdV}}(u) = \partial_x \sigma_n^{\mathsf{KdV}}(u) + \sum_{k=0}^{n} \sigma_k^{\mathsf{KdV}}(u)\sigma_{n-k}^{\mathsf{KdV}}(u) \quad \forall n \geq 1\,.$$

The derivation of this recurrence relation, up to a different choice of signs, can be found in [186]. J.-C. Saut first proved in [159] the existence of global distributional solutions to (KdV_n) for initial data in Sobolev spaces. Subsequently, M. Schwarz in [99] showed global existence and uniqueness of solutions in Sobolev spaces on the torus, again at high regularity.

We are most interested in a theory developed in the 1990s by C. E. Kenig, G. Ponce and L. Vega [106, 108, 109, 110, 111]. Here *local smoothing* and *maximal function estimates* for the linear evolution are proven and used to derive local and global well-posedness results for dispersive PDEs for large classes of linear and nonlinear parts. We focus in particular on the papers [105, 107], where local (and in some cases global) well-posedness results in weighted Sobolev spaces are proven for a class of equations that includes the KdV hierarchy.

We use their methods to set up a local well-posedness theory for a large class of dispersive nonlinear systems of PDEs that covers the case of Proposition 3.1.14. Before we provide the details, let us discuss more recent methods and results, and why we do not use them here.

D. Pilod [152] studied a family of equations similar to the KdV hierarchy, but with only quadratic nonlinearities, and proved local well-posedness for small initial data in weighted L^2-based Sobolev spaces. However, they also proved ill-posedness in $H^s(\mathbb{R})$ for any $s \in \mathbb{R}$ in the sense that the solution map cannot be C^2 at zero. Subsequently in [104], D. Pilod and C. E. Kenig established local well-posedness in $H^s(\mathbb{R})$ for initial data of arbitrary size for a class of equations that includes the KdV hierarchy, using subtle energy estimates and parabolic regularization. Technically, the smallness of the initial data is used in the proofs in both of these papers, although this smallness assumption is later lifted in the second paper by using the scaling symmetry. We want to avoid smallness assumptions in the proof because there is no scaling symmetry available in the perturbative setting of Proposition 3.1.14.

Based on their work on the Fourier restriction norm method and the use of Fourier-Lebesgue spaces for dispersive nonlinear equations (see e. g. [86]), A. Grünrock proved in [88] well- and ill-posedness results in the aforementioned spaces for the mKdV and KdV hierarchies at low regularity. Recently, J. Adams continued work in this direction [4, 5], proving well- and ill-posedness results for the NLS- and dNLS hierarchies in Fourier-Lebesgue and modulation spaces at low regularity.

We certainly expect the Fourier restriction norm method to be applicable to our setting, but not without considerable work. We face the following complications for our system (3.1.21) in Proposition 3.1.14 (see Lemma 3.4.1 below for details):

- We are dealing with a genuine *system*, which is only diagonalizable with a singular change of variables (see (3.4.13));
- The dispersion relation is *nonpolynomial*: for $n \in 2\,\mathbb{N}$ it is of the form $\varphi(\xi) = \pm\xi^{n-1}\sqrt{\xi^2+4}$.
- The "nonlinearity" contains *quadratic* terms, and also *linear* terms with non-constant coefficients.

For these reasons, we focus on the classical theory developed by C. E. Kenig, G. Ponce and L. Vega, working with high regularity and the weighted Sobolev spaces $(H^s \cap H^{s',1})(\mathbb{R})$. We attempt to push their arguments for local smoothing and maximal function estimates to the limit in terms of the generality of the dispersion relation. Although the linear estimates we obtain are, to the best of our knowledge, not in the literature, we have relegated them to Appendix 3.A due to their independence from the rest of the work. Our general linear estimates admit dispersion relations φ that are truly "not like a derivative" in the sense that critical points $\varphi'(\xi) = 0$ with $\xi \neq 0$ are possible. We use these linear estimates in an adaptation of the method in [107] to obtain a local well-posedness result for a large class of dispersive nonlinear systems of PDEs. This is the content of Theorem 3.4.6.

In this local well-posedness theory we need to restrict ourselves to dispersion relations that are rather "like a derivative", in the sense that $\varphi'(\xi) = 0 \implies \xi = 0$ and $\varphi(\xi) \sim \langle\xi\rangle^n$ for large frequencies. The reason is that by the local smoothing effect we have a gain of regularity of order $|\varphi'(D_x)|^{\frac{1}{2}}$. In practice, this assumption is encoded in the requirement of *boundedness* of the auxiliary symbol $\widetilde{a}(\xi)$, which plays a subtle role that we would like to explain further. For example, consider the case $n = 2$, where the perturbative formulation of ($\widetilde{\mathsf{GP}}_2$) from Proposition 3.1.14 has the matrix dispersion relation

$$\Phi(\xi) = \begin{pmatrix} \xi^2+2 & 2 \\ -2 & -\xi^2-2 \end{pmatrix}$$

with eigenvalues $\varphi(\xi) = \pm\xi\sqrt{\xi^2+4}$. Here we choose

$$\widetilde{a}(\xi) = \underbrace{\frac{|\xi|^{\frac{1}{2}}}{|\varphi'(\xi)|^{\frac{1}{2}}}}_{(I)}\underbrace{\frac{\xi}{\langle\xi\rangle}}_{(II)} \cdot \underbrace{\left(\text{linear combinations of } \frac{\xi}{\langle\xi\rangle} \text{ and } \frac{\langle\xi\rangle}{\xi}\right)}_{(III)}$$

This symbol appears in the local smoothing estimate in the following way:

$$\|(\widetilde{a}|\varphi'|^{\frac{1}{2}})(D_x)e^{it\varphi(D_x)}u_0\|_{L^\infty_x L^2_t} \lesssim \|u_0\|_{L^2}\,.$$

The purpose of the term (I) is to obtain an estimate which is convenient for use with Sobolev spaces. A complication which was also encountered in [89] is that the diagonalization of $\Phi(\xi)$ is singular at $\xi = 0$, which causes the terms (III) to appear. We relax the estimate that we aim to show by introducing the factor (II), which removes this singularity. Conveniently, the resulting symbol $\widetilde{a}(\xi)$ is still admissible for our linear estimates. For details we refer to Lemma 3.4.1 and Remark 3.4.2.

After establishing the local well-posedness results, we prove a blow-up alternative and use Grönwall's inequality to prevent the $H^{s',1}$-norm from blowing up, provided the H^s-norm does not (see Lemma 3.4.8). This allows us to prove Theorem 3.1.2.

Finally, let us also mention the method of commuting flows, which was developed and used by R. Killip and M. Vișan to prove global well-posedness of KdV in $H^{-1}(\mathbb{R})$ in [114], and subsequently applied to other equations such as Benjamin-Ono [112] and NLS [91]. We would like to mention that in [114] the flow of KdV is approximated by a flow whose Hamiltonian involves the logarithm of the transmission coefficient. One may understand this as a modulation of the solution along the higher flows, giving an example of higher symmetries aiding in the understanding of an equation in an integrable hierarchy. In the paper [38] by B. Bringmann and the aforementioned authors, this method was applied to the fifth-order equation in the KdV hierarchy, and finally to the whole hierarchy in the work [118] by F. Klaus, H. Koch and B. Liu. We again believe that it is in principle possible to apply this method to the NLS hierarchy with nonzero boundary data, but recognize that it would take considerable work.

3.1.3. Notations and definitions

When we write $A \lesssim_{\text{things}} B$ we mean $A \leq C(\text{things})B$, where the constant is a function of the objects in the parentheses, and must be continuous in the real or complex parameters.

Whenever we consider a function space without explicit domain, the domain is implicitly assumed to be $\mathbb{R}$. When functions in the variables (t, x) are considered, we use subscripts to denote with respect to which variable the function space should be considered.

For Banach spaces X and Y, we denote by $L(X, Y)$ the space of bounded linear operators from X to Y.

Set $D_x = -i\partial_x$ and note that $\widehat{D_x f} = \xi \widehat{f}$ and $\widehat{xf} = -D_\xi \widehat{f}$ for $f \in \mathcal{S}(\mathbb{R}; \mathbb{C})$.

For matrices $\mathbf{c} \in \mathbb{C}^{N\times N}$ we denote the entry in the n-th row and m-th column by $\mathbf{c}_{n,m}$. The only exception is the matrix of Jost solutions $\Phi^\pm = (\Phi_1^\pm | \Phi_2^\pm)$. We denote by $|\mathbf{c}|$ the operator norm and remark that the choice of underlying vector norm does not matter up to universal constants.

If $u(t)$ is a formal power series in t, then $[t^n]u(t)$ denotes the application of the linear functional that extracts the coefficient in front of t^n.

For real numbers $x, y \in \mathbb{R}$, we define $x \wedge y = \min\{x, y\}$ and $x \vee y = \max\{x, y\}$, as well as the Japanese bracket $\langle x \rangle = \sqrt{x^2 + 1}$ and a modified version $\langle x \rangle_\sim = \sqrt{x^2 + 4}$. We use the notations $\lceil x \rceil = \min\{k \in \mathbb{Z} : k \geq x\}$ and $\lfloor x \rfloor = \max\{k \in \mathbb{Z} : k \leq x\}$.

Lastly, we write $\mathcal{Q}_j$, $j \in \{1, 2, 3, 4\}$ for the four open quadrants of the complex plane.

Acknowledgements.
We thank our colleagues Julia Henninger and Sebastian Ohrem for many insightful discussions. We thank Nikolay Tzvetkov for pointing out a reference that led to Remark 3.A.2.

Statements and Declarations.
This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173.

No datasets were generated or analysed during the current study.

The authors have no competing interests to declare that are relevant to the content of this article.

SymPy [144] was used to carry out symbolic computations for the purposes of exploration and verification of various identities involving recurrence relations. Various versions of the OpenAI services ChatGPT and Codex were used for mathematical and grammatical proofreading, literature search, and exploratory discussion. Furthermore, the normal form transformation in Proposition 3.1.14 was proposed by GPT-5.5.

3.2. Asymptotic expansions of direct scattering data and definitions of $\mathcal{H}_n^{\mathsf{NLS}}$ and $\mathcal{H}_n^{\mathsf{GP}}$

In this section we shall study the spectral problem $L\Phi = \lambda\Phi$ of the Lax operator (3.1.2) for a potential $q \in q_* + \mathcal{S}(\mathbb{R}; \mathbb{C})$, where $q_* = q_*(x)$ is the reference potential with limits $q_\pm \in \mathbb{S}^1$ at infinity. We choose q_* to be the dark soliton profile

$$\begin{aligned} q_*(x) &= q_+\overline{\zeta_+}(\mathrm{Re}[\zeta_+] + i\,\mathrm{Im}[\zeta_+]\tanh(\mathrm{Im}[\zeta_+]x)) \\ &= q_-\overline{\zeta_-}(\mathrm{Re}[\zeta_-] + i\,\mathrm{Im}[\zeta_-]\tanh(\mathrm{Im}[\zeta_-]x)). \end{aligned} \tag{3.2.1}$$

Here $\zeta_+ \in e^{i[0,\pi)}$ and $\zeta_- = \overline{\zeta_+} \in e^{i(\pi,2\pi]}$ are defined by $\zeta_+^2 = \frac{q_+}{q_-}$ and $\zeta_-^2 = \frac{q_-}{q_+}$.

The time dependence is largely ignored. We aim to establish the triangular representations for the Jost solution matrices $\Phi^\pm(x, \lambda)$ in terms of λ-independent kernels $\Gamma^\pm(x, y)$, which will be shown to have sufficiently strong decay properties (see Lemma 3.2.2 below).

Based on Lemma 3.2.2, in Section 3.2.4 we rigorously show the asymptotic expansion of the logarithm of the transmission coefficient $\log a(\lambda)$ as $|z| \to \infty$ (see Lemma 3.2.5 below). The relation between the two spectral parameters λ, z is explained in Subsection 3.2.1 below. The coefficients in this expansion are the GP-Hamiltonians $\mathcal{H}_n^{\mathsf{GP}}$ (see Lemma 3.2.9 below). The NLS-Hamiltonians $\mathcal{H}_n^{\mathsf{NLS}}$ appear as coefficients in the asymptotic expansion in the ZBC setting $q \in \mathcal{S}(\mathbb{R};\mathbb{C})$ instead (see (3.2.50)). The sequence of densities σ_n of these Hamiltonians plays an important role in the analysis, and its recurrence relation is crucial for the structure analysis in Section 3.3 below.

3.2.1. A Riemann surface

Consider the Riemann surface

$$\mathbb{K} = \{(\lambda, z) \in \mathbb{C}^2 : \lambda^2 - z^2 = 1\}$$

and decompose $\mathbb{K} = \overline{\mathbb{K}_+} \cup \overline{\mathbb{K}_-}$ using the two sheets

$$\mathbb{K}_\pm = \{(\lambda, z) \in \mathbb{K} : \lambda \in \mathbb{C} \setminus ((-\infty, -1] \cup [1, \infty)), \pm \operatorname{Im} z > 0\}.$$

We define two branches of the square root:

$$\sqrt{re^{i\theta}} = \sqrt{r}e^{i\frac{\theta}{2}} \qquad \forall\, \theta \in (-\pi, \pi] \quad \forall\, r > 0 \qquad \tilde{\sqrt{re^{i\theta}}} = \sqrt{r}e^{i\frac{\theta}{2}} \qquad \forall\, \theta \in (0, 2\pi] \quad \forall\, r > 0\,.$$

In Figure 3.1 we depict pictographically the mapping properties between λ and z on the four open quadrants $\mathcal{Q}_j$, $j \in \{1, 2, 3, 4\}$ of the complex plane.

Figure 3.1.: Mapping properties of $\lambda \leftrightarrow z$. An arrow is bold if and only if the corresponding map is continuous from closed quadrant to closed quadrant.

With the help of this pictogram, we can find on each closed quadrant an explicit homeomorphism that maps $\lambda \leftrightarrow z$. There is no single choice which works for all quadrants. It is convenient to introduce the complex variable ζ, which is defined by

$$\zeta = \lambda + z \qquad \zeta^{-1} = \lambda - z \qquad \lambda = \frac{\zeta + \zeta^{-1}}{2} \qquad z = \frac{\zeta - \zeta^{-1}}{2}\,. \tag{3.2.2}$$

When considering $(\lambda, z) \in \mathbb{K}$ and $\zeta \in \mathbb{C}$, we shall freely use the maps depicted in Figure 3.1, as well as the relations (3.2.2), to consider functions of one parameter also as functions of the others. In particular, we may write $\lambda \in \mathbb{K}_\pm$ or $z \in \mathbb{K}_\pm$ to refer to a pair $(\lambda, z) \in \mathbb{K}_\pm$. In situations where the details of the mapping matter, we split into cases depending on the quadrant under consideration.

Abusing notation, we define

$$\gamma_1 = 1 = \text{“}+\text{”} \qquad\qquad \gamma_2 = -1 = \text{“}-\text{”}$$

and write

$$\mathbb{K}_{+\gamma_1} = \mathbb{K}_{-\gamma_2} = \mathbb{K}_+ \qquad\qquad \mathbb{K}_{-\gamma_1} = \mathbb{K}_{+\gamma_2} = \mathbb{K}_-\,.$$

We shall define below Jost solutions $\Phi_j^\pm = \Phi_j^\pm(x, \lambda)$ for $\lambda \in \overline{\mathbb{K}_{\mp\gamma_j}}$, that is, Φ_1^- and Φ_2^+ on $\overline{\mathbb{K}_+}$, and Φ_2^- and Φ_1^+ on $\overline{\mathbb{K}_-}$. We recognize $\Phi_j^\pm$, $j = 1, 2$ as the first and second columns in the Jost solution matrices $\Phi^\pm = (\Phi_1^\pm | \Phi_2^\pm)$.

3.2.2. Jost solutions

Let $q \in C^\infty(\mathbb{R};\mathbb{C})$ have boundary values $\lim_{x\to\pm\infty} q(x) = q_\pm \in \mathbb{S}^1$ such that $q - q_\pm \in \mathcal{S}(\mathbb{R}_\pm;\mathbb{C})$. Consider for $j \in \{1, 2\}$ the solutions to the Zakharov-Shabat scattering problem

$$\Phi_j^\pm : \mathbb{R} \times \overline{\mathbb{K}_{\mp\gamma_j}} \longrightarrow \mathbb{C}^2$$

$$\partial_x \Phi_j^\pm(x,\lambda) = \begin{pmatrix} -i\lambda & q(x) \\ \overline{q}(x) & i\lambda \end{pmatrix} \Phi_j^\pm(x,\lambda) \tag{3.2.3}$$

$$\lim_{x\to\pm\infty} \Phi_j^\pm(x,\lambda) e^{xiz\gamma_j} = E_j^\pm(\lambda) \quad \text{where} \quad E^\pm(\lambda) = \begin{pmatrix} q_\pm & i(z-\lambda) \\ i(\lambda-z) & \overline{q}_\pm \end{pmatrix}. \tag{3.2.4}$$

Here $E_1^\pm$ is the first and $E_2^\pm$ the second column of $E^\pm$, and the functions $E_j^\pm e^{-xiz\gamma_j}$ are solutions to (3.2.3) if q is replaced by $q_\pm$. We call $\Phi_j^\pm$ the Jost solutions for the potential q.

For notational simplicity, we can rewrite the system (3.2.3) as

$$\partial_x \Phi_j^\pm(x,\lambda) = U(x,\lambda)\Phi_j^\pm(x,\lambda) \text{ with } U(x,\lambda) = Q(x) - i\lambda\sigma_3, \quad Q(x) = \begin{pmatrix} 0 & q(x) \\ \overline{q}(x) & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

where the limit matrices $U_\pm = Q_\pm - i\lambda\sigma_3$ with $Q_\pm = \begin{pmatrix} 0 & q_\pm \\ \overline{q}_\pm & 0 \end{pmatrix}$ have eigenvalues $-iz\gamma_j$ with corresponding eigenvectors $E_j^\pm$, $j=1,2$, that is, $e^{xU_\pm(\lambda)} = E^\pm(\lambda)e^{-xiz\sigma_3}E^\pm(\lambda)^{-1}$.

We define the modified Jost solutions $\Psi_j^\pm$ by

$$\Psi_j^\pm(x,\lambda) = \Phi_j^\pm(x,\lambda)e^{xiz\gamma_j}.$$

They are the (unique) solutions to the system

$$\partial_x \Psi_j^\pm(x,\lambda) = \begin{pmatrix} -i\lambda + iz\gamma_j & q(x) \\ \overline{q}(x) & i\lambda + iz\gamma_j \end{pmatrix} \Psi_j^\pm(x,\lambda) = \big(U_\pm(\lambda) + (Q-Q_\pm)(x) + iz\gamma_j\big)\Psi_j^\pm(x,\lambda) \tag{3.2.5}$$

$$\lim_{x\to\pm\infty} \Psi_j^\pm(x,\lambda) = E_j^\pm(\lambda). \tag{3.2.6}$$

Consider the following equivalent integral representations for the modified Jost solutions, each of which yields a corresponding Neumann series ansatz:

$$\Psi_j^\pm(x,\lambda) = E_j^\pm(\lambda) + \int_{\pm\infty}^x e^{(x-y)U_\pm(\lambda)}(Q-Q_\pm)(y)\Psi_j^\pm(y,\lambda)e^{(x-y)iz\gamma_j}\,\mathrm{d}y \tag{3.2.7}$$

$$\Psi_j^\pm(x,\lambda) = E_j^\pm(\lambda) + \int_{\pm\infty}^x e^{x\widetilde{U}(x,\lambda)}e^{-y\widetilde{U}(y,\lambda)}(Q-\widetilde{Q})(y)\Psi_j^\pm(y,\lambda)e^{(x-y)iz\gamma_j}\,\mathrm{d}y. \tag{3.2.8}$$

Here $\widetilde{Q}(x) = \mathbb{1}_{\{x\le 0\}}Q_- + \mathbb{1}_{\{x>0\}}Q_+$ and $\widetilde{U}(x,\lambda) = \widetilde{Q}(x) - i\lambda\sigma_3$. By solving the above integral equations, one proves the following well-known existence result for the Jost solutions. Note that we are interested in the behavior for large $|\lambda|$ and hence do not put emphasis on points of nonanalyticity.

Lemma 3.2.1 (Existence of Jost solutions, [64, Proposition 3])**.** *If $q - q_\pm \in L^1(\mathbb{R}_\pm)$ then the Jost solutions $\Phi_j^\pm$ and their modified versions $\Psi_j^\pm$ exist. They are analytic in λ except for a finite set of points. They are smooth in x if q is smooth and all derivatives are in $L^1(\mathbb{R})$.*

The representation (3.2.8) has the advantage that the integrand contains $Q - \widetilde{Q}$, which decays at *both* infinities. As explained in the introduction, we are interested in integral representations of modified Jost solutions because we wish to study their asymptotic expansion in powers of λ and z at infinity. Since we aim to control these asymptotic expansions at *both* limits $x \to \pm\infty$, it is of importance to use an appropriate integral representation. To obtain such an expansion, we may attempt to repeatedly integrate the factor $e^{(x-y)iz\gamma_j}$ in (3.2.8) using integration by parts. Since the remaining integrand depends nontrivially on λ, this is inconvenient. We therefore need an integral representation that has an integrand with decay at both infinities, and is also suitable for asymptotic expansion via integration by parts.

3.2.3. Triangular representations

Another integral representation of the modified Jost solutions is given by the triangular representation (see [71, Chapter 1, §8] and [189]):

$$\Psi_j^\pm(x,\lambda) = E_j^\pm(\lambda) + \int_{\pm\infty}^x \Gamma^\pm(x,y)E_j^\pm(\lambda)e^{(x-y)iz\gamma_j}\,\mathrm{d}y. \tag{3.2.9}$$

Here the kernel $\Gamma^{\pm}$ is independent of λ and $E_j^{\pm}(\lambda)$ is bounded and analytic (up to a finite set of points) in λ. Therefore, this representation is suitable for the asymptotic expansion of $\Psi_j^{\pm}(x,\lambda)$ in powers of z at infinity. In [71] it is shown that $\Gamma^{\pm} \in C^{\infty}(\{(x,y) \in \mathbb{R}^2 : \pm(x-y) < 0\}; \mathbb{C}^{2\times 2})$ exists and satisfies $\Gamma^{\pm}(x,\cdot) \in \mathcal{S}(\{y \in \mathbb{R} : \pm(x-y) < 0\}; \mathbb{C}^{2\times 2})$. Subsequently, the authors use this integral representation to derive various asymptotic expansions of the Jost solutions, but not the crucial expansion (3.2.32) in Lemma 3.2.5 for the transmission coefficient that we are interested in. Nevertheless, (3.2.32) is proven in [71], but the proof is heavily abridged and works by referring to previous chapters, where the Zakharov-Shabat problem is considered on a torus of length L, and then transferring expansions obtained in this setting through the limit $L \to \infty$. We present here a direct approach to the expansion of the Jost solutions and the transmission coefficient that uses the triangular representation (3.2.12) below.

Instead of using $q - \widetilde{q}$, with $\widetilde{q}(x) = \mathbb{1}_{\{x\le 0\}} q_- + \mathbb{1}_{\{x>0\}} q_+$, as a source of decay at both infinities, we compare q to the reference profile q_* in (3.2.1) that assumes our boundary data $q_{\pm}$ at infinity. Most conveniently, explicit formulas are given in [46] for the modified Jost solutions of the dark soliton. They are

$$\Psi_{*,1}^{-}(x,\lambda) = \begin{pmatrix} q_- \frac{\zeta-\zeta_+}{\zeta-\zeta_-} + iq_- \frac{1}{\zeta-\zeta_-} \frac{2\,\mathrm{Im}[\zeta_+]}{e^{2\,\mathrm{Im}[\zeta_+]x}+1} \\ i\zeta^{-1} \frac{\zeta-\zeta_+}{\zeta-\zeta_-} \zeta_-^2 - \frac{1}{\zeta-\zeta_-} \zeta_- \frac{2\,\mathrm{Im}[\zeta_+]}{e^{2\,\mathrm{Im}[\zeta_+]x}+1} \end{pmatrix} \quad \Psi_{*,2}^{-}(x,\lambda) = \begin{pmatrix} -i\zeta^{-1} - \frac{1}{\zeta-\zeta_+} \zeta_+ \frac{2\,\mathrm{Im}[\zeta_-]}{e^{2\,\mathrm{Im}[\zeta_-]x}+1} \\ \overline{q}_- - i\overline{q}_- \frac{1}{\zeta-\zeta_+} \frac{2\,\mathrm{Im}[\zeta_-]}{e^{2\,\mathrm{Im}[\zeta_-]x}+1} \end{pmatrix} \tag{3.2.10}$$

$$\Psi_{*,1}^{+}(x,\lambda) = \begin{pmatrix} q_+ \frac{\zeta-\zeta_-}{\zeta-\zeta_+} + iq_+ \frac{1}{\zeta-\zeta_+} \frac{2\,\mathrm{Im}[\zeta_-]}{e^{2\,\mathrm{Im}[\zeta_-]x}+1} \\ i\zeta^{-1} \frac{\zeta-\zeta_-}{\zeta-\zeta_+} \zeta_+^2 - \frac{1}{\zeta-\zeta_+} \zeta_+ \frac{2\,\mathrm{Im}[\zeta_-]}{e^{2\,\mathrm{Im}[\zeta_-]x}+1} \end{pmatrix} \quad \Psi_{*,2}^{+}(x,\lambda) = \begin{pmatrix} -i\zeta^{-1} - \frac{1}{\zeta-\zeta_-} \zeta_- \frac{2\,\mathrm{Im}[\zeta_+]}{e^{2\,\mathrm{Im}[\zeta_+]x}+1} \\ \overline{q}_+ - i\overline{q}_+ \frac{1}{\zeta-\zeta_-} \frac{2\,\mathrm{Im}[\zeta_+]}{e^{2\,\mathrm{Im}[\zeta_+]x}+1} \end{pmatrix} \tag{3.2.11}$$

in our notation. The subscript $*$ always denotes that the potential in the Zakharov-Shabat problem is q_* instead of q. We may now study $\Psi_j^{\pm}$ as perturbations of $\Psi_{*,j}^{\pm}$ by use of the integral representation formula

$$\Psi_j^{\pm}(x,\lambda) = \Psi_{*,j}^{\pm}(x,\lambda) + \int_{\pm\infty}^{x} \Gamma^{\pm}(x,y) \Psi_{*,j}^{\pm}(y,\lambda) e^{(x-y)iz\gamma_j}\, \mathrm{d}y\,, \tag{3.2.12}$$

where crucially the kernel $\Gamma^{\pm}(x,y)$ is independent of λ, and $\Psi_{*,j}^{\pm}(y,\lambda)$ is bounded and analytic (up to a finite set of points) in λ, i. e. a benign factor.

We have to show the existence of the kernel $\Gamma^{\pm}(x,y)$. To this end, we are going to derive boundary value problems for $\Gamma^{\pm}(x,y)$ from the system (3.2.5)–(3.2.6). Define

$$Q_*(x) = \begin{pmatrix} 0 & q_*(x) \\ \overline{q}_*(x) & 0 \end{pmatrix} \qquad\qquad U_*(x,\lambda) = Q_*(x) - i\lambda\sigma_3\,,$$

and recall that $\Psi_{*,j}^{\pm}$ solves

$$\partial_x \Psi_{*,j}^{\pm}(x,\lambda) = (U_*(x,\lambda) + iz\gamma_j)\Psi_{*,j}^{\pm}(x,\lambda) \tag{3.2.13}$$

$$\lim_{x\to\pm\infty} \Psi_{*,j}^{\pm}(x,\lambda) = E_j^{\pm}(\lambda)\,. \tag{3.2.14}$$

We write the differential equation (3.2.5) as $0 = (\partial_x - U(x,\lambda) - iz\gamma_j)\Psi_j^{\pm}(x,\lambda)$ and substitute into this the ansatz (3.2.12), yielding

$$\begin{aligned} 0 &= -(U(x,\lambda) + iz\gamma_j - U_*(x,\lambda) - iz\gamma_j)\Psi_{*,j}^{\pm}(x,\lambda) + \Gamma^{\pm}(x,x)\Psi_{*,j}^{\pm}(x,\lambda) \\ &\quad + \int_{\pm\infty}^{x} \big((\partial_x - U(x,\lambda))\Gamma^{\pm}(x,y)\big)\Psi_{*,j}^{\pm}(y,\lambda)e^{(x-y)iz\gamma_j}\,\mathrm{d}y \\ &= \big(\Gamma^{\pm}(x,x) - (U - U_*)(x,\lambda)\big)\Psi_{*,j}^{\pm}(x,\lambda) + \int_{\pm\infty}^{x} \big((\partial_x - U(x,\lambda))\Gamma^{\pm}(x,y)\big)\Psi_{*,j}^{\pm}(y,\lambda)e^{(x-y)iz\gamma_j}\,\mathrm{d}y\,. \end{aligned}$$

We denote the diagonal and off-diagonal parts of the 2×2 matrices $\Gamma^{\pm}$ by

$$\Gamma_{\mathrm{d}}^{\pm} = \frac{1}{2}(\Gamma^{\pm} + \sigma_3\Gamma^{\pm}\sigma_3) \qquad\qquad \Gamma_{\mathrm{od}}^{\pm} = \frac{1}{2}(\Gamma^{\pm} - \sigma_3\Gamma^{\pm}\sigma_3)\,.$$

They satisfy

$$\Gamma^{\pm} = \Gamma^{\pm}_{\mathrm{d}} + \Gamma^{\pm}_{\mathrm{od}} = \sigma_3\Gamma^{\pm}\sigma_3 + 2\Gamma^{\pm}_{\mathrm{od}}.$$

We note that $(U - U_*)_{\mathrm{d}} = (Q - Q_*)_{\mathrm{d}} = 0$ and prescribe the value of $\Gamma^{\pm}_{\mathrm{od}}$ on the diagonal $\{(x, x) : x \in \mathbb{R}\}$ to be

$$\Gamma^{\pm}_{\mathrm{od}}(x, x) = \frac{1}{2}(U - U_*)_{\mathrm{od}}(x) = \frac{1}{2}(Q - Q_*)(x)\,.$$

Thus, the differential equation (3.2.5) reduces to

$$0 = \sigma_3\Gamma^{\pm}(x, x)\sigma_3\Psi^{\pm}_{*,j}(x, \lambda) + \int_{\pm\infty}^{x} \big((\partial_x - U(x, \lambda))\Gamma^{\pm}(x, y)\big)\Psi^{\pm}_{*,j}(y, \lambda)e^{(x-y)iz\gamma_j}\,\mathrm{d}y\,.$$

Define $\Phi^{\pm}_{*,j}(x, \lambda) = e^{-xiz\gamma_j}\Psi^{\pm}_{*,j}(x, \lambda)$ and let

$$F(x, y) = \sigma_3\Gamma^{\pm}(x, y)\sigma_3\Phi^{\pm}_{*,j}(y, \lambda) + \int_{\pm\infty}^{y} (\partial_x - U(x, \lambda))\Gamma^{\pm}(x, s)\Phi^{\pm}_{*,j}(s, \lambda)\,\mathrm{d}s\,.$$

Then the above equation is a special case of $F(x, y) = 0$. Observe that

$$\begin{aligned}\partial_y F(x, y) &= \big(\sigma_3\partial_y\Gamma^{\pm}(x, y)\sigma_3 + \sigma_3\Gamma^{\pm}(x, y)\sigma_3U_*(y, \lambda) + (\partial_x - U(x, \lambda))\Gamma^{\pm}(x, y)\big)\Phi^{\pm}_{*,j}(y, \lambda)\\ \lim_{y\to\pm\infty} F(x, y) &= \lim_{y\to\pm\infty} \sigma_3\Gamma^{\pm}(x, y)\sigma_3\Phi^{\pm}_{*,j}(y, \lambda)\,.\end{aligned}$$

We see that $F(x, y) = 0$ is satisfied if $\Gamma^{\pm}$ solves the boundary value problem

$$\begin{gathered}\sigma_3\partial_y\Gamma^{\pm}(x, y)\sigma_3 = -\sigma_3\Gamma^{\pm}(x, y)\sigma_3U_*(y, \lambda) - (\partial_x - U(x, \lambda))\Gamma^{\pm}(x, y)\\ \lim_{y\to\pm\infty} \sigma_3\Gamma^{\pm}(x, y)\sigma_3\Phi^{\pm}_{*,j}(y, \lambda) = 0 \qquad \Gamma^{\pm}_{\mathrm{od}}(x, x) = \frac{1}{2}(Q - Q_*)(x)\,.\end{gathered}$$

It is necessary at this point that $\Phi^{\pm}_{*,j}(y, \lambda)$ remains bounded as $y \to \pm\infty$, which is valid on the domain of definition $\mathbb{R} \times \overline{\mathbb{K}_{\mp\gamma_j}}$. Then a solution to the boundary value problem

$$\begin{gathered}\sigma_3\partial_y\Gamma^{\pm}(x, y)\sigma_3 = -\sigma_3\Gamma^{\pm}(x, y)\sigma_3Q_*(y) - (\partial_x - Q(x))\Gamma^{\pm}(x, y)\\ \lim_{y\to\pm\infty} \Gamma^{\pm}(x, y) = 0 \qquad \Gamma^{\pm}_{\mathrm{od}}(x, x) = \frac{1}{2}(Q - Q_*)(x)\end{gathered}$$

indeed yields a solution to (3.2.5)–(3.2.6), as long as (3.2.12) is well-defined.

We state now our well-posedness result for this boundary value problem, which is equivalent to (3.2.15)–(3.2.17) below. It provides all the estimates necessary for the aforementioned asymptotic expansion of the Jost solutions and the transmission coefficient in powers of λ and z. In the following, for a matrix $M \in \mathbb{C}^{2\times 2}$ we write $\overrightarrow{M}$ to denote the operator of multiplication from the right by M, and treat it as if it were a matrix in our use of language and notation.

Lemma 3.2.2 (Existence and decay properties of kernels $\Gamma^{\pm}$ in the triangular representations)**.** *Let $q \in q_* + \mathcal{S}(\mathbb{R}; \mathbb{C})$ with the dark soliton profile q_* being defined in (3.2.1). There exist smooth solutions $\Gamma^{\pm} : \{(x, y) : \pm(x - y) \leq 0\} \to \mathbb{C}^{2\times 2}$ to the boundary value problem*

$$(\partial_x + \partial_y)\Gamma^{\pm}_{\mathrm{d}}(x, y) = \Big(Q(x) + \overrightarrow{Q_*(y)}\Big)\Gamma^{\pm}_{\mathrm{od}}(x, y) \tag{3.2.15}$$

$$(\partial_x - \partial_y)\Gamma^{\pm}_{\mathrm{od}}(x, y) = \Big(Q(x) - \overrightarrow{Q_*(y)}\Big)\Gamma^{\pm}_{\mathrm{d}}(x, y) \tag{3.2.16}$$

$$\lim_{y\to\pm\infty} \Gamma^{\pm}(x, y) = 0 \qquad \Gamma^{\pm}_{\mathrm{od}}(x, x) = \frac{1}{2}(Q - Q_*)(x) \tag{3.2.17}$$

with the following properties:

(i) *For every $k, m \in \mathbb{N}$ there exist bounded, monotonic control functions $c_\pm \in C_b(\mathbb{R}; \mathbb{R}_+)$, decreasing faster than any power of $\frac{1}{x}$ as $x \to \pm\infty$, such that*

$$\|((\partial_x + \partial_y)\partial_x^m \partial_y^k \Gamma^-)(s+y, s-y)\|_{(L^1 \cap L^\infty)_s((-\infty,x])} \le c_-(x) e^{\int_{x-y}^{x+y} c_-(s)\,\mathrm{d}s} \tag{3.2.18}$$

$$\|((\partial_x + \partial_y)\partial_x^m \partial_y^k \Gamma^+)(s-y, s+y)\|_{(L^1 \cap L^\infty)_s([x,\infty))} \le c_+(x) e^{\int_{x-y}^{x+y} c_+(s)\,\mathrm{d}s} \tag{3.2.19}$$

for all $x \in \mathbb{R}$ and $y \ge 0$. Integrating along the direction $\partial_x + \partial_y$ yields

$$|\partial_x^m \partial_y^k \Gamma^-(x,y)| \lesssim c_-\left(\frac{x+y}{2}\right) e^{\int_y^x c_-(s)\,\mathrm{d}s} \tag{3.2.20}$$

$$|\partial_x^m \partial_y^k \Gamma^+(x,y)| \lesssim c_+\left(\frac{x+y}{2}\right) e^{\int_x^y c_+(s)\,\mathrm{d}s}\,. \tag{3.2.21}$$

(ii) *On the diagonal $y = x$, we have*

$$(\partial_y^k \Gamma^\pm)|_{y=x} \in C_b(\mathbb{R}; \mathbb{C}^{2\times 2}) \cap \mathcal{S}(\mathbb{R}_\pm; \mathbb{C}^{2\times 2}) \quad ((\partial_x + \partial_y)\partial_y^k \Gamma^\pm)|_{y=x} \in \mathcal{S}(\mathbb{R}; \mathbb{C}^{2\times 2})\,. \tag{3.2.22}$$

Consequently, the formula (3.2.12) gives the solutions

$$\Psi_j^\pm : \mathbb{R} \times \overline{\mathbb{K}_{\mp\gamma_j}} \longrightarrow \mathbb{C}^2$$

to the Zakharov-Shabat scattering problem (3.2.5)–(3.2.6).

Proof. Appendix 3.B contains the proof of (3.2.18) and (3.2.19), as well as the statements on the diagonal (see Claim 3.B.2). □

Remark 3.2.3. *If the reference profile q_* is chosen as $q_* = q_-$, then the system (3.2.15)–(3.2.17) matches [71, Chapter 1, (8.18)–(8.19)].*

3.2.4. Transmission coefficient, asymptotic expansions, recurrence relations, and definitions of $\mathcal{H}_n^{\mathsf{NLS}}$ and $\mathcal{H}_n^{\mathsf{GP}}$

We focus only on the sheet $\overline{\mathbb{K}_+}$ here for simplicity, but every result in this section has an analogous statement and proof on $\overline{\mathbb{K}_-}$.

3.2.4.1. Definitions of the transmission coefficient and its logarithm

If $\operatorname{Im} z = 0$ then $\Phi^- = (\Phi_1^- | \Phi_2^-)$ and $\Phi^+ = (\Phi_1^+ | \Phi_2^+)$ are two fundamental solution matrices to (3.2.3), hence there exist $a(\lambda), b(\lambda) \in \mathbb{C}$ such that

$$\Phi_1^-(x,\lambda) = a(\lambda)\Phi_1^+(x,\lambda) + b(\lambda)\Phi_2^+(x,\lambda)\,, \tag{3.2.23}$$

$$\text{or equivalently, } \Psi_1^-(x,\lambda) = a(\lambda)\Psi_1^+(x,\lambda) + b(\lambda)e^{2xiz}\Psi_2^+(x,\lambda)\,. \tag{3.2.24}$$

We call $a(\lambda)$ the transmission coefficient of the Lax operator (3.1.2) with the potential q. It fulfills

$$a(\lambda) = \frac{\det(\Psi_1^-, \Psi_2^+)}{\det(\Psi_1^+, \Psi_2^+)} = \frac{\Psi_{1,1}^- \Psi_{2,2}^+ - \Psi_{1,2}^- \Psi_{2,1}^+}{\det(\Phi_1^+, \Phi_2^+)} = \frac{\Psi_{1,1}^- \Psi_{2,2}^+ - \Psi_{1,2}^- \Psi_{2,1}^+}{2z(\lambda - z)}\,. \tag{3.2.25}$$

For the last equality we used the fact that $\det(\Phi_1^+, \Phi_2^+)$ is independent of $x \in \mathbb{R}$ (due to the trace-free matrix in (3.2.3)), and hence equal to $\det E^+ = 2z(\lambda - z)$. Using this formula, we extend $a(\lambda)$ analytically to $\lambda \in \mathbb{K}_+$, where Ψ_1^-, Ψ_2^+ are well-defined, up to a finite set of points. For the dark soliton profile q_* given in (3.2.1), we can explicitly determine $a_*(\lambda) = \frac{q_-}{q_+}\frac{\zeta - \zeta_+}{\zeta - \zeta_-}$ and $b_*(\lambda) = 0$.

Suppose for now that

$$\text{the limits } \lim_{x\to\infty} \Psi_1^-(x,\lambda),\ \lim_{x\to-\infty} \Psi_2^+(x,\lambda) \text{ and } \lim_{x\to\infty} \partial_x \Psi_1^-(x,\lambda),\ \lim_{x\to-\infty} \partial_x \Psi_2^+(x,\lambda) \text{ exist.} \tag{3.2.26}$$

In particular, the latter two limits must be zero. Suppose further that for every $q \in q_* + \mathcal{S}(\mathbb{R};\mathbb{C})$ there exist $c = c(q) > 0$ and $C = C(q) > 0$ such that $\operatorname{Im} z > c$ implies

$$\sup_{x\in\mathbb{R}} |\Psi^-_{1,1}(x,\lambda) - q_-| + \sup_{x\in\mathbb{R}} |\Psi^+_{2,2}(x,\lambda) - \overline{q}_+| < \frac{C}{|z|} < \frac{1}{2}\,. \tag{3.2.27}$$

Using (3.2.5) to substitute $\Psi^-_{1,2}$ in (3.2.25) and taking the limit $x \to \infty$, or respectively substituting $\Psi^+_{2,1}$ and taking the limit $x \to -\infty$, we find from (3.2.6) that

$$\lim_{x\to\infty} \Psi^-_{1,1}(x,\lambda) = a(\lambda) q_+ \qquad\qquad \lim_{x\to-\infty} \Psi^+_{2,2}(x,\lambda) = a(\lambda)\overline{q}_-\,. \tag{3.2.28}$$

Consequently, if $\operatorname{Im} z > c$, the functions $\Psi^-_{1,1}(\cdot,\lambda)\overline{q}_+$ and $\Psi^+_{2,2}(\cdot,\lambda) q_-$ have similar limits at infinity in the following sense:

$$\begin{aligned} \lim_{x\to-\infty} \Psi^-_{1,1}(x,\lambda)\overline{q}_+ &= q_-\overline{q}_+ = \lim_{x\to+\infty} \Psi^+_{2,2}(x,\lambda) q_- \\ \lim_{x\to+\infty} \Psi^-_{1,1}(x,\lambda)\overline{q}_+ &= a(\lambda) = \lim_{x\to-\infty} \Psi^+_{2,2}(x,\lambda) q_-. \end{aligned} \tag{3.2.29}$$

By (3.2.27) there exists a choice of branch log of the logarithm, depending only on $q_-\overline{q}_+$, for which $\log(\Psi^-_{1,1}(\cdot,\lambda)\overline{q}_+)$ and $\log(\Psi^+_{2,2}(\cdot,\lambda)q_-)$ are continuously differentiable in x when $\operatorname{Im} z > c$, and

$$\log a(\lambda) - \log(q_-\overline{q}_+) = \int_{\mathbb{R}} \sigma^{\mathsf{GP}}(x,\lambda)\,\mathrm{d}x = \int_{\mathbb{R}} \widetilde{\sigma}^{\mathsf{GP}}(x,\lambda)\,\mathrm{d}x\,, \tag{3.2.30}$$

where

$$\sigma^{\mathsf{GP}}(x,\lambda) = \frac{\partial_x \Psi^-_{1,1}(x,\lambda)}{\Psi^-_{1,1}(x,\lambda)} \qquad\qquad \widetilde{\sigma}^{\mathsf{GP}}(x,\lambda) = -\frac{\partial_x \Psi^+_{2,2}(x,\lambda)}{\Psi^+_{2,2}(x,\lambda)}\,. \tag{3.2.31}$$

In Lemma 3.2.5 below, we establish that these densities are indeed integrable and (3.2.26)–(3.2.27) hold.

3.2.4.2. Asymptotic expansions for Ψ_1^-, Ψ_2^+, σ^{GP}, $\widetilde{\sigma}^{\mathsf{GP}}$, and $\log a(\lambda)$

Definition 3.2.4 (Asymptotic expansion)**.** Let $D \subset \mathbb{C}$ be an unbounded set, $d \in \mathbb{N}$, and $1 \le p \le \infty$.

(i) We say that a function $f = f(z) : D \to \mathbb{C}^d$ has an **asymptotic expansion in powers of** $2iz$ **at infinity on** D if there exist $(f_n)_{n\in\mathbb{N}} \subset \mathbb{C}^d$ such that

$$\forall\, N \in \mathbb{N} \qquad \lim_{\substack{|z|\to\infty \\ z\in D}} (2iz)^N \left(f(z) - \sum_{n=0}^{N} \frac{f_n}{(2iz)^n} \right) = 0\,.$$

We call f_n the **expansion coefficients** of f and note that they are unique. If f and g have such an expansion then $2izf$ and fg do as well. If $g(z) \ne 0$ for $|z|$ sufficiently large, then fg^{-1} also has such an expansion.

(ii) We say that a function $f = f(x,z) : D \times \mathbb{R} \to \mathbb{C}^d$ has an L^p**-smooth asymptotic expansion in powers of** $2iz$ **at infinity on** D if $f(\cdot,z)$ is smooth for every $z \in D$, and there exist $(f_n)_{n\in\mathbb{N}} \subset (C^\infty \cap L^p)(\mathbb{R};\mathbb{C}^d)$ such that

$$\forall\, k, N \in \mathbb{N} \qquad \lim_{\substack{|z|\to\infty \\ z\in D}} \left\| (2iz)^N \left(\partial_x^k f(x,z) - \sum_{n=0}^{N} \frac{\partial_x^k f_n(x)}{(2iz)^n} \right) \right\|_{L^p_x(\mathbb{R})} = 0\,.$$

In this case $\partial_x f$ has such an expansion as well. If g has such an L^∞-smooth expansion and $\inf_{x\in\mathbb{R}} |g(x,z)| > 0$ for $|z|$ sufficiently large, then fg^{-1} also has an L^p-smooth asymptotic expansion. If $p = 1$ then $\int_{-\infty}^{\infty} f(x,z)\,\mathrm{d}x$ has an asymptotic expansion in powers of $2iz$ at infinity on D with expansion coefficients $\int_{-\infty}^{\infty} f_n(x)\,\mathrm{d}x$.

When f has an asymptotic expansion of some kind with expansion coefficients $(f_n)_{n\in\mathbb{N}}$, we write

$$f \sim \sum_{n=0}^{\infty} \frac{f_n}{(2iz)^n}.$$

We now make use of the reference Jost solution (3.2.10) and the decay properties of kernels $\Gamma^{\pm}$ in Lemma 3.2.2 to show the asymptotic expansions for $\log a(\lambda)$.

Lemma 3.2.5 (Asymptotic expansions of $\sigma^{\mathsf{GP}}, \widetilde{\sigma}^{\mathsf{GP}}, \log a(\lambda)$)**.** *Assume $q \in C^\infty(\mathbb{R};\mathbb{C})$ with $q - q_\pm \in \mathcal{S}(\mathbb{R}_\pm;\mathbb{C})$. Then Ψ_1^-, Ψ_2^+, $\partial_x\Psi_1^-$, $\partial_x\Psi_2^+$, σ^{GP}, and $\widetilde{\sigma}^{\mathsf{GP}}$ have L^∞-smooth asymptotic expansions in powers of $2iz$ at infinity on the unbounded domain $D = \mathbb{K}_+ \cap \{\operatorname{Im} z > c, \operatorname{Im}\lambda > 0\}$ for some $c = c(q) > 0$, and our assumptions (3.2.26)–(3.2.27) are true there.*

Furthermore, the asymptotic expansions for $\partial_x\Psi_1^-$, $\partial_x\Psi_2^+$, σ^{GP}, and $\widetilde{\sigma}^{\mathsf{GP}}$ are L^1-smooth, and as a result $\log a(\lambda)$ has an asymptotic expansion of the form

$$\log a(\lambda) - \log(q_-\overline{q}_+) \sim \sum_{n=1}^{\infty} \frac{\int \sigma_n^{\mathsf{GP}}(x)\,\mathrm{d}x}{(2iz)^n} = \sum_{n=1}^{\infty} \frac{\int \widetilde{\sigma}_n^{\mathsf{GP}}(x)\,\mathrm{d}x}{(2iz)^n}, \tag{3.2.32}$$

where σ_n^{GP} and $\widetilde{\sigma}_n^{\mathsf{GP}}$ (with $\sigma_0^{\mathsf{GP}} = \widetilde{\sigma}_0^{\mathsf{GP}} = 0$) denote the corresponding expansion coefficients of σ^{GP} and $\widetilde{\sigma}^{\mathsf{GP}}$.

Remark 3.2.6. *We can analogously consider the same asymptotic expansions on the unbounded set $D_- = \mathbb{K}_+ \cap \{\operatorname{Im} z > c,\ \operatorname{Im}\lambda < 0\}$. Notice the difference between the expansions of λ in powers of $2z$ on D (with $\operatorname{Im}\lambda > 0$) and on D_- (with $\operatorname{Im}\lambda < 0$):*

$$\zeta = \lambda + z \sim 2z\left(1 + \sum_{m=0}^{\infty} \frac{(-1)^m}{2m+1}\binom{2m+1}{m}\frac{1}{(2z)^{2m+2}}\right), \quad \text{if } \operatorname{Im}\lambda > 0, \tag{3.2.33}$$
$$\frac{1}{\zeta} = \lambda - z \sim -2z\left(1 + \sum_{m=0}^{\infty} \frac{(-1)^m}{2m+1}\binom{2m+1}{m}\frac{1}{(2z)^{2m+2}}\right), \quad \text{if } \operatorname{Im}\lambda < 0.$$

In particular σ_n^{GP} and $\widetilde{\sigma}_n^{\mathsf{GP}}$ in (3.2.32) differ between D and D_-.

Remark 3.2.7. *Note that from (3.2.12) and the same dominated convergence argument, we obtain on D the integral representation formula*

$$a(\lambda) = a_*(\lambda)\left(1 - \int_0^\infty \lim_{x\to\infty} \Gamma^-(x, x-y)\begin{pmatrix} 1 \\ \frac{1}{\lambda - z}\end{pmatrix} e^{yiz}\,\mathrm{d}y\right).$$

Proof. We give the proof only for the case of Ψ_1^- and σ^{GP}, as it is analogous for Ψ_2^+ and $\widetilde{\sigma}^{\mathsf{GP}}$.

Recall the triangular representation formula (3.2.12). For $k \in \mathbb{N}$ it implies

$$\partial_x^k \Psi_1^-(x,\lambda) = \partial_x^k \Psi_{*,1}^-(x,\lambda) - \int_0^\infty \partial_x^k\big(\Gamma^-(x,x-y)\Psi_{*,1}^-(x-y,\lambda)\big)e^{yiz}\,\mathrm{d}y. \tag{3.2.34}$$

Integrating by parts N times, we have the following expression for $\partial_x^k\Psi_1^-$ in terms of the reference Jost solution $\Psi_{*,1}^-$ and the kernel $\Gamma^-(x,y)$:

$$\begin{aligned}\partial_x^k \Psi_1^-(x,\lambda) = \partial_x^k \Psi_{*,1}^-(x,\lambda) &- \sum_{n=1}^{N} \partial_x^k\Big(\partial_y^{n-1}\big|_{y=x}\big(\Gamma^-(x,y)\Psi_{*,1}^-(y,\lambda)\big)\Big)(iz)^{-n} \\ &- \int_0^\infty \partial_x^k(-\partial_y)^N\big(\Gamma^-(x,x-y)\Psi_{*,1}^-(x-y,\lambda)\big)e^{yiz}(iz)^{-N}\,\mathrm{d}y.\end{aligned}$$

We summarize the known facts concerning $\Psi_{*,1}^-$ and $\Gamma^-(x,y)$.

- From the explicit formula (3.2.10), we know that

$$\Psi^-_{*,1} = \begin{pmatrix} q_- \frac{\zeta-\zeta_+}{\zeta-\zeta_-} + iq_- \frac{1}{\zeta-\zeta_-} \frac{2\,\mathrm{Im}[\zeta_+]}{e^{2\,\mathrm{Im}[\zeta_+]x}+1} \\ i\zeta^{-1}\frac{\zeta-\zeta_+}{\zeta-\zeta_-}\zeta_-^2 - \frac{1}{\zeta-\zeta_-}\zeta_- \frac{2\,\mathrm{Im}[\zeta_+]}{e^{2\,\mathrm{Im}[\zeta_+]x}+1}\end{pmatrix} \tag{3.2.35}$$

and we recall

$$\zeta = \lambda + z, \quad \zeta_+^2 = q_+\overline{q}_-, \quad \mathrm{Im}[\zeta_+] \geq 0, \quad \zeta_- = \overline{\zeta_+}\,. \tag{3.2.36}$$

Note that $\Psi^-_{*,1,1}\overline{q}_+$ has the limits $q_-\overline{q}_+$ at $-\infty$ and $a_*(\lambda)$ at $+\infty$, which matches (3.2.29) for $\Psi^-_{1,1}\overline{q}_+$. By virtue of (3.2.33), $\Psi^-_{*,1}$ has an L^∞-smooth asymptotic expansion in powers of $2iz$ at infinity on the unbounded set D:

$$\forall\, N\in\mathbb{N} \qquad \lim_{\substack{|z|\to\infty\\ z\in D}} (2iz)^N \left\| \Psi^-_{*,1}(x,\lambda) - \sum_{n=0}^N \frac{\Psi^{-,n}_{*,1}(x)}{(2iz)^n}\right\|_{L^\infty_x(\mathbb{R})} = 0\,. \tag{3.2.37}$$

Here $\Psi^{-,0}_{*,1} = \begin{pmatrix} q_- \\ 0\end{pmatrix}$.

- The derivative

$$\partial_x\Psi^-_{*,1} = \begin{pmatrix} -iq_- \frac{1}{\zeta-\zeta_-}\frac{(2\,\mathrm{Im}[\zeta_+])^2 e^{2\,\mathrm{Im}[\zeta_+]x}}{(e^{2\,\mathrm{Im}[\zeta_+]x}+1)^2} \\ \frac{1}{\zeta-\zeta_-}\zeta_-\frac{(2\,\mathrm{Im}[\zeta_+])^2 e^{2\,\mathrm{Im}[\zeta_+]x}}{(e^{2\,\mathrm{Im}[\zeta_+]x}+1)^2}\end{pmatrix} \tag{3.2.38}$$

has an L^1- and L^∞-smooth asymptotic expansion in powers of $2iz$ at infinity on the unbounded set D as follows:

$$\forall\, k\geq 1, N\in\mathbb{N} \qquad \lim_{\substack{|z|\to\infty\\ z\in D}} \left\| (2iz)^N \left(\partial_x^k\Psi^-_{*,1}(x,\lambda) - \sum_{n=1}^N \frac{\partial_x^k\Psi^{-,n}_{*,1}(x)}{(2iz)^n}\right)\right\|_{L^1_x\cap L^\infty_x(\mathbb{R})} = 0\,. \tag{3.2.39}$$

Here $\partial_x\Psi^{-,0}_{*,1} = 0$.

- Using the estimates (3.2.18)–(3.2.20), we know that for every $k,N\in\mathbb{N}$ with $k\geq 1$ there exists some control function $c^-\in L^\infty(\mathbb{R};\mathbb{R})$, decreasing faster than any power of $\frac{1}{x}$ at $-\infty$, such that

$$\left\|(-\partial_y)^N\big(\Gamma^-(x,x-y)\Psi^-_{*,1}(x-y,\lambda)\big)e^{yiz}\right\|_{L^\infty_x((-\infty,x_0])} \leq c^-\Big(x_0-\frac{y}{2}\Big)e^{yc^-(x_0)-y\,\mathrm{Im}\, z}$$

$$\left\|\partial_x^k(-\partial_y)^N\big(\Gamma^-(x,x-y)\Psi^-_{*,1}(x-y,\lambda)\big)e^{yiz}\right\|_{(L^1\cap L^\infty)_x((-\infty,x_0])} \leq c^-\Big(x_0-\frac{y}{2}\Big)e^{yc^-(x_0)-y\,\mathrm{Im}\, z}$$

for all $y\geq 0$ and λ with $\mathrm{Im}\, z > \|c^-\|_{L^\infty}$.

We now rewrite carefully the above expression for $\partial_x^k\Psi^-_1$, $k\geq 0$, by use of the asymptotic expansion of $\Psi^-_{*,1}$:

$$\begin{aligned}
\partial_x^k\Psi^-_1(x,\lambda) &= \sum_{n=0}^N (2iz)^{-n}\partial_x^k\left(\Psi^{-,n}_{*,1}(x) - \left(\sum_{m=0}^{n-1} 2^{n-m}\partial_y^{n-1-m}\big|_{y=x}\big(\Gamma^-(x,y)\Psi^{-,m}_{*,1}(y)\big)\right)\right)\\
&\quad + \left(\partial_x^k\Psi^-_{*,1}(x,\lambda) - \sum_{n=0}^N\frac{\partial_x^k\Psi^{-,n}_{*,1}(x)}{(2iz)^n}\right)\\
&\quad - \sum_{n=1}^N (2iz)^{-n}2^n\partial_x^k\left(\partial_y^{n-1}\Big|_{y=x}\left(\Gamma^-(x,y)\left(\Psi^-_{*,1}(y,\lambda) - \sum_{m=0}^{N-n}\frac{\Psi^{-,m}_{*,1}(y)}{(2iz)^m}\right)\right)\right)\\
&\quad - (iz)^{-N}\int_0^\infty \partial_x^k(-\partial_y)^N\big(\Gamma^-(x,x-y)\Psi^-_{*,1}(x-y,\lambda)\big)e^{yiz}\,\mathrm{d}y\,.
\end{aligned}$$

By combining the above facts concerning $\Psi^-_{*,1}$ and Γ^-, we derive that $\Psi^-_1(x,\lambda)$ has an L^∞-smooth, and $\partial_x\Psi^-_1(x,\lambda)$ an L^∞- and L^1-smooth asymptotic expansion in powers of $2iz$ at infinity on D. The expansion coefficients are

$$\Psi^{-,n}_1(x) = \Psi^{-,n}_{*,1}(x) - \sum_{m=0}^{n-1} 2^{n-m}\partial_y^{n-1-m}\big|_{y=x}\big(\Gamma^-(x,y)\Psi^{-,m}_{*,1}(y)\big)\,.$$

In particular

$$\Psi_1^{-,0} = \begin{pmatrix} q_- \\ 0 \end{pmatrix} \qquad \partial_x \Psi_1^{-,0} = 0\,,$$

which, together with the knowledge of $\Psi^-_{*,1}$ and Γ^-, implies the assumptions (3.2.26)–(3.2.27). Thus $\inf_{x\in\mathbb{R}} |\Psi^-_{1,1}(x,\lambda)| > 0$ when $\operatorname{Im} z$ is sufficiently large, and hence $\sigma^{\mathsf{GP}}(x,\lambda) = \frac{\partial_x \Psi^-_{1,1}}{\Psi^-_{1,1}}$ has the same type of asymptotic expansion as $\partial_x \Psi^-_{1,1}$. That is, an L^∞- and L^1-smooth asymptotic expansion in powers of $2iz$ at infinity on D, with expansion coefficients σ_n^{GP} (note that $\sigma_0^{\mathsf{GP}} = 0$). Finally, the expansion (3.2.32) for $\log a(\lambda)$ follows from (3.2.30). □

3.2.4.3. The expansion coefficients σ_n^{GP} and σ_n^{NLS}

In this subsection we would like to establish the recurrence relations for the expansion coefficients σ_n^{GP} from Lemma 3.2.5. The first idea would be to apply (3.2.5) to (3.2.31), which yields the Riccati equation

$$(2iz)\sigma^{\mathsf{GP}} - \sigma_x^{\mathsf{GP}} - (\sigma^{\mathsf{GP}})^2 + q\overline{q} - 1 + q_x \frac{\sigma^{\mathsf{GP}} + i(\lambda - z)}{q} = 0\,. \tag{3.2.40}$$

We can then obtain a recurrence relation for the expansion coefficients σ_n^{GP} by substituting σ^{GP} in (3.2.40) with the power series $\sum_{n=0}^\infty \frac{\sigma_n^{\mathsf{GP}}}{(2iz)^n}$ and comparing coefficients.

However, due to the presence of λ on the right-hand side, this yields an awkward recurrence relation, so we consider instead the quantity [2]

$$\sigma^{\mathsf{NLS}}(x,\lambda) = \sigma^{\mathsf{GP}}(x,\lambda) + i(\lambda - z)\,, \tag{3.2.41}$$

for which we show the following.

Lemma 3.2.8. *Assume the setting of Lemma 3.2.5, i. e. $q \in C^\infty(\mathbb{R};\mathbb{C})$ with $q - q_\pm \in \mathcal{S}(\mathbb{R}_\pm;\mathbb{C})$. With σ^{NLS} defined by (3.2.41), the following hold:*

- *σ^{NLS} has an L^∞- (while not L^1-) smooth asymptotic expansion in powers of $2i\lambda$ at infinity on D:*

$$\sigma^{\mathsf{NLS}}(x,\lambda) \sim \sum_{n=0}^\infty \frac{\sigma_n^{\mathsf{NLS}}(x)}{(2i\lambda)^n} \quad \text{with } \sigma_0^{\mathsf{NLS}} = 0\,. \tag{3.2.42}$$

- *The following relations between expansion coefficients σ_n^{NLS} and σ_n^{GP} hold*

$$\sigma_{2m}^{\mathsf{NLS}} = \sum_{k=0}^m \binom{m-1}{m-k} (-4)^{m-k} \sigma_{2k}^{\mathsf{GP}} \qquad \sigma_{2m+1}^{\mathsf{NLS}} = \sum_{k=0}^m \binom{m-\frac{1}{2}}{m-k} (-4)^{m-k} \sigma_{2k+1}^{\mathsf{GP}} + (-1)^{m+1} C_m \tag{3.2.43}$$

$$\sigma_{2m}^{\mathsf{GP}} = \sum_{k=0}^m \binom{m-1}{m-k} 4^{m-k} \sigma_{2k}^{\mathsf{NLS}} \qquad \sigma_{2m+1}^{\mathsf{GP}} = \sum_{k=0}^m \binom{m-\frac{1}{2}}{m-k} 4^{m-k} \sigma_{2k+1}^{\mathsf{NLS}} + C_m\,. \tag{3.2.44}$$

  Here $(C_n)_{n\geq 0}$ are the Catalan numbers. They can be defined either by the recurrence relation they solve, or by their generating function:

$$C_0 = 1 \qquad C_{n+1} = \sum_{k=0}^n C_k C_{n-k} \qquad \sum_{n=0}^\infty X^n C_n = \frac{1 - \sqrt{1-4X}}{2X}\,. \tag{3.2.45}$$

- *σ^{NLS} satisfies the Riccati equation*

$$(2i\lambda)\sigma^{\mathsf{NLS}} = \partial_x \sigma^{\mathsf{NLS}} - \frac{q_x}{q}\sigma^{\mathsf{NLS}} + (\sigma^{\mathsf{NLS}})^2 - q\overline{q}\,, \tag{3.2.46}$$

[2] In the (NLS)–(ZBC) setting $\sigma^{\mathsf{NLS}}(x,\lambda)$ is indeed the density of the transmission coefficient, i. e. it fulfills (3.2.46) and also (3.2.30) in the absence of the term $\log(q_- \overline{q}_+)$, and it is σ^{GP} which is not integrable.

which yields for σ_n^{NLS} the recurrence relation

$$\sigma_0^{\mathsf{NLS}} = 0 \qquad \sigma_1^{\mathsf{NLS}} = -q\bar{q} \qquad \sigma_{n+1}^{\mathsf{NLS}} = \partial_x \sigma_n^{\mathsf{NLS}} - \frac{q_x}{q}\sigma_n^{\mathsf{NLS}} + \sum_{k=0}^{n} \sigma_k^{\mathsf{NLS}} \sigma_{n-k}^{\mathsf{NLS}} . \tag{3.2.47}$$

The recurrence relation for σ_n^{GP} follows from (3.2.44) and (3.2.47).

Proof.

Recall from Lemma 3.2.5 that σ^{GP} has an L^∞- and L^1-smooth asymptotic expansion in powers of $2iz$:

$$\sigma^{\mathsf{GP}}(x,\lambda) \sim \sum_{n=0}^{\infty} \frac{\sigma_n^{\mathsf{GP}}(x)}{(2iz)^n}\,, \text{ with } \sigma_0^{\mathsf{GP}} = 0 \tag{3.2.48}$$

at infinity on $D = \mathbb{K}_+ \cap \{\operatorname{Im} z > c, \operatorname{Im}\lambda > 0\}$ for some sufficiently large $c = c(q) > 0$. Together with the equivalence between $\lambda = \sqrt{z^2+1}$ and z as $|z| \to \infty$ in the case $\operatorname{Im} z > 0, \operatorname{Im}\lambda > 0$ (see also (3.2.33) for more detail), this implies the validity of the expansion (3.2.42). The fact that $\sigma_0^{\mathsf{NLS}} = 0$ follows from $\sigma_0^{\mathsf{GP}} = 0$ and $\lambda - z = O(\frac{1}{|z|})$ as $|z| \to \infty$.

We can then derive the relations (3.2.43)–(3.2.44) from (3.2.41), (3.2.42) and (3.2.48) by direct computation.

The Riccati equation (3.2.46) is equivalent to (3.2.40), and the recurrence relation (3.2.47) follows from (3.2.46) and the L^∞-smooth asymptotic expansion (3.2.42). □

3.2.4.4. THE HAMILTONIANS $\mathcal{H}_n^{\mathsf{NLS}}$ AND $\mathcal{H}_n^{\mathsf{GP}}$

In the classical setting $q \in \mathcal{S}(\mathbb{R};\mathbb{C})$ one can similarly define the Jost solutions, the transmission coefficient $a^{\mathsf{NLS}}(\lambda)$, and $\sigma^{\mathsf{NLS}}(x,\lambda)$ (see e. g. [71]). Then σ^{NLS} has an L^∞- and L^1-smooth asymptotic expansion in powers of $2i\lambda$ on $\{\lambda \in \mathbb{C} : \operatorname{Im}\lambda > c(q)\}$, with expansion coefficients σ_n^{NLS}. We define

$$\mathcal{H}_n^{\mathsf{NLS}} = -(-i)^n \int_{\mathbb{R}} \sigma_{n+1}^{\mathsf{NLS}}(x)\,\mathrm{d}x, \quad n \geq 0\,. \tag{3.2.49}$$

The Hamiltonians $\mathcal{H}_n^{\mathsf{NLS}}$ are real-valued functionals and they Poisson commute: $\{\mathcal{H}_n^{\mathsf{NLS}}, \mathcal{H}_m^{\mathsf{NLS}}\} = 0$. We have a corresponding asymptotic expansion of $\log a^{\mathsf{NLS}}(\lambda)$ in powers of 2λ of the form

$$\log a^{\mathsf{NLS}}(\lambda) \sim i \sum_{n=0}^{\infty} \frac{\mathcal{H}_n^{\mathsf{NLS}}}{(2\lambda)^{n+1}}\,. \tag{3.2.50}$$

We return now to the setting with the nonzero boundary condition (NZBC).

Lemma 3.2.9. *Assume the setting of Lemma 3.2.5, i. e. $q \in C^\infty(\mathbb{R};\mathbb{C})$ with $q - q_\pm \in \mathcal{S}(\mathbb{R}_\pm;\mathbb{C})$. We define*

$$\mathcal{H}_{-1}^{\mathsf{GP}} = -i \log_{\mathrm{pr}}(q_- \bar{q}_+) \qquad \mathcal{H}_n^{\mathsf{GP}} = -(-i)^n \int_{\mathbb{R}} \sigma_{n+1}^{\mathsf{GP}}(x)\,\mathrm{d}x \quad n \geq 0.$$

Recall also that the Hamiltonians $\mathcal{H}_n^{\widetilde{\mathsf{GP}}}$ are defined from $\mathcal{H}_n^{\mathsf{GP}}$ via (3.1.6).

(i) *The Hamiltonians $\mathcal{H}_n^{\mathsf{GP}}$ are real-valued functionals.*

(ii) *For $n, m \in \mathbb{N}$ we have*

$$\{\mathcal{H}_n^{\mathsf{GP}}, \mathcal{H}_m^{\mathsf{GP}}\} = 0\,.$$

More generally, for all $\lambda_1, \lambda_2 \in \mathbb{K}_+$ we have $\{a(\lambda_1), a(\lambda_2)\} = 0$.

(iii) *There exists some $c = c(q) > 0$ for which the functional $\log a(\lambda)$ has an asymptotic expansion in powers of $2z$ at infinity on $D = \mathbb{K}_+ \cap \{\operatorname{Im} z > c, \operatorname{Im}\lambda > 0\}$ of the form*

$$\log a(\lambda) \sim i \sum_{n=-1}^{\infty} \frac{\mathcal{H}_n^{\mathsf{GP}}}{(2z)^{n+1}} \sim i \sum_{n=0}^{\infty} \frac{\mathcal{H}_{2n}^{\widetilde{\mathsf{GP}}}}{(2z)^{2n+1}} + i\frac{\lambda}{z} \sum_{n=0}^{\infty} \frac{2^{\mathbb{1}_{\{n=0\}}} \mathcal{H}_{2n-1}^{\widetilde{\mathsf{GP}}}}{(2z)^{2n}}\,. \tag{3.2.51}$$

Proof.

(i) This follows from symmetry: one can derive the Riccati equation for $\widetilde{\sigma}^{\mathsf{GP}}$ by applying (3.2.5) to (3.2.31), which yields $(2iz)\widetilde{\sigma}^{\mathsf{GP}} + \widetilde{\sigma}^{\mathsf{GP}}_x - (\widetilde{\sigma}^{\mathsf{GP}})^2 + q\overline{q} - 1 - \overline{q}_x \frac{\widetilde{\sigma}^{\mathsf{GP}} + i(\lambda - z)}{\overline{q}} = 0$. We notice that $\widetilde{\sigma}^{\mathsf{GP}}(-x, \lambda; \overline{q}^\sharp)$ with $q^\sharp(x) = q(-x)$ also solves (3.2.40).

A consequence of the recurrence relations (3.2.44) and (3.2.47) is that σ_n^{GP} consists of monomials with a total number of derivatives that has the opposite parity of n. Then $\overline{(\partial_x^k q^\sharp)(x)} = (-1)^k (\partial_x^k q)(-x)$ together with (3.2.30) yields $\int \sigma_n^{\mathsf{GP}}(x,\lambda)\,\mathrm{d}x = (-1)^{n+1} \int \overline{\sigma_n^{\mathsf{GP}}(x,\lambda)}\,\mathrm{d}x$. Hence $\mathcal{H}_n^{\mathsf{GP}} = -(-i)^n \int_{\mathbb{R}} \sigma_{n+1}^{\mathsf{GP}}(x)\,\mathrm{d}x$ is real-valued.

(ii) We refer to [71, III.§2] and [118, Theorem B.7].

(iii) This follows from Lemma 3.2.5 and a calculation with the definitions of the Hamiltonians, in particular (3.1.6).

□

Remark 3.2.10 (Relations between $\frac{\delta}{\delta \overline{q}}\mathcal{H}_n^{\mathsf{NLS}}$ and $\frac{\delta}{\delta \overline{q}}\mathcal{H}_n^{\mathsf{GP}}$)**.** *As explained above in Subsection 3.2.4.3, in the setting of* NZBC*, σ^{NLS}, defined in* (3.2.41)*, is not integrable with respect to $x \in \mathbb{R}$, and thus $\mathcal{H}_n^{\mathsf{NLS}}$ cannot be well-defined in* (3.2.49)*. Nevertheless, the L^∞-smooth asymptotic expansion for σ^{NLS} in D implies the well-definedness of $\frac{\delta \mathcal{H}_n^{\mathsf{NLS}}}{\delta \overline{q}}$, and in particular the relations* (3.1.4)*–*(3.1.5) *between $\frac{\delta \mathcal{H}_n^{\mathsf{NLS}}}{\delta \overline{q}}$ and $\frac{\delta \mathcal{H}_n^{\mathsf{GP}}}{\delta \overline{q}}$ follow from* (3.2.44)*.*

3.3. Analysis of the structure of the NLS and GP hierarchies

The goal of this section is to prove Theorem 3.1.10 and Proposition 3.1.14. We do this by studying the recurrence relation (3.2.47) from Lemma 3.2.8. The arguments are computational in nature, involving lengthy calculations with generating functions and recurrence relations. Since these calculations are simple, in the sense that they can be carried out mechanically, we do not present the details. In order to help the reader follow the argument, and indeed convince themselves of the correctness of the claimed formulas, we provide a "Python companion"[3] to go along with this section. It contains 39 Python functions, named `test1` to `test39`, that test specific identities (for a finite number of cases) symbolically, using the "SymPy" library [144]. Note that these Python functions are not part of the proof; they are only supplementary material. We indicate with footnotes when an identity has a corresponding Python function.

For a function $F = F(q, \overline{q})$, we use the shorthand notation

$$\delta F = \delta F(q, \overline{q}) = \frac{\delta}{\delta \overline{q}} \int_{\mathbb{R}} F(q(x), \overline{q}(x))\,\mathrm{d}x$$

for the functional derivative with respect to $\overline{q}$. With this notation, we write (NLS_n) as

$$i\partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\mathsf{NLS}}}{\delta \overline{q}} = -(-i)^n \delta \sigma_{n+1}^{\mathsf{NLS}}$$

where $\mathcal{H}_n^{\mathsf{NLS}}$ is given in terms of $\sigma_{n+1}^{\mathsf{NLS}}$ in (3.2.49), and σ_n^{NLS} are the expansion coefficients of σ^{NLS} given in Lemma 3.2.8. In this section we study these coefficients in a purely symbolic sense, so the question of boundary conditions ((ZBC) or (NZBC)) is not relevant. For the rest of this section we write $\sigma = \sigma^{\mathsf{NLS}}$.

Given a polynomial Q of $q, \overline{q}$ and their derivatives, we write $\pi_k Q$ for the sum of all monomials in Q which have exactly k factors with derivatives. We want to obtain an explicit formula for $\pi_0 \delta \sigma_{n+1}$ and $\pi_1 \delta \sigma_{n+1}$. The reason is that in our well-posedness theory we make a perturbative ansatz $q(x,t) = q_*(x) + p(x,t)$, and the non-trivial linear part of the equation for p depends only on this part of $\delta \sigma_{n+1}$. Using the notation $\mathcal{O}_A^{M,L}(B)$ in (3.1.18), we decompose $\delta \sigma_{n+1}$ as follows:

$$\delta \sigma_{n+1} = \pi_0 \delta \sigma_{n+1} + \pi_1 \delta \sigma_{n+1} + \mathcal{O}_q^{2,n-2}(q_x)\,. \tag{3.3.1}$$

[3] https://gist.github.com/Robert-Wegner/05451952fee3019bb784bc659a85409c

Recall the recurrence relation (3.2.47) for $\sigma_n = \sigma_n^{\mathsf{NLS}}$:

$$\sigma_0 = 0 \qquad \sigma_1 = -q\overline{q} \qquad \sigma_{n+1} = \partial_x \sigma_n - \frac{q_x}{q}\sigma_n + \sum_{k=0}^{n} \sigma_k \sigma_{n-k} \,. \tag{3.3.2}$$

We make a number of observations:

- Observation for $\pi_0\sigma_n$. (3.3.2) implies

$$\begin{aligned} \pi_0\sigma_1 &= -|q|^2 \\ \pi_0\sigma_{n+1} &= \sum_{k=1}^{n} \pi_0\sigma_k \pi_0\sigma_{n-k} \,. \end{aligned} \tag{3.3.3}$$

  This allows us to find an explicit formula for $\pi_0\sigma_n$, which we state in Lemma 3.3.1 below.
- Observation for $\pi_1\sigma_n$. (3.3.2) implies the following recurrence relation for $\pi_1\sigma_n$, involving $\pi_0\sigma_n$:

$$\begin{aligned} \pi_1\sigma_1 &= 0 \\ \pi_1\sigma_{n+1} &= \partial_x \pi_0\sigma_n - \frac{q_x}{q}\pi_0\sigma_n + \pi_1\partial_x\pi_1\sigma_n + \sum_{k=1}^{n} 2\pi_1\sigma_k\pi_0\sigma_{n-k} \,. \end{aligned} \tag{3.3.4}$$

  This allows us to find an explicit formula for $\pi_1\sigma_n$, which we state in Lemma 3.3.2 below.
- Observation for $\pi_1\delta\sigma_n$. Unfortunately, $\pi_1\delta\sigma_n$ depends not just on $\delta\pi_1\sigma_n$, but also on certain terms from $\pi_2\sigma_n$. Let us carefully perform this analysis, starting with the trivial observation

$$\pi_1\delta\sigma_n = \sum_{k=0}^{\infty} \pi_1\delta\pi_k\sigma_n \,.$$

  Clearly, $\delta\pi_0\sigma_n$ has no derivatives and is therefore in the kernel of π_1. Similarly, $\delta\pi_k\sigma_n$ for $k \geq 3$ always has at least 2 factors in each monomial which have derivatives, so it is also in the kernel of π_1. We suppose here that we already have an explicit formula for $\pi_1\delta\pi_1\sigma_n$, so it remains to study $\pi_1\delta\pi_2\sigma_n$. In fact, we have

$$\pi_1\delta\pi_2\sigma_n = \pi_1\delta\widetilde{\pi}_2\sigma_n,$$

  where $\widetilde{\pi}_2$ projects onto sums of monomials which have exactly two factors with derivatives, and one of those factors is of the form $\partial_x^k\overline{q}$. Similarly, we define $\widetilde{\pi}_1$ to project onto sums of monomials which have only one factor with derivatives, and this factor is $\partial_x^k\overline{q}$ for some $k \geq 1$. We obtain the decomposition

$$\pi_1\delta\sigma_n = \pi_1\delta\pi_1\sigma_n + \pi_1\delta\widetilde{\pi}_2\sigma_n \,. \tag{3.3.5}$$

  With our explicit formula for $\pi_1\sigma_n$, we can compute the first term, so it remains to find an explicit formula for $\widetilde{\pi}_2\sigma_n$.
- Observation for $\widetilde{\pi}_2\sigma_n$. (3.3.2) implies

$$\begin{aligned} \widetilde{\pi}_2\sigma_1 &= 0 \\ \widetilde{\pi}_2\sigma_{n+1} &= \widetilde{\pi}_2\partial_x\widetilde{\pi}_2\sigma_n - \frac{q_x}{q}\widetilde{\pi}_1\sigma_n + \widetilde{\pi}_2\partial_x\pi_1\sigma_n + 2\sum_{k=1}^{n-1}(\widetilde{\pi}_2\sigma_k\pi_0\sigma_{n-k} + \widetilde{\pi}_1\sigma_k\pi_1\sigma_{n-k}) \,. \end{aligned} \tag{3.3.6}$$

  This helps us derive the formula for $\pi_1\delta\widetilde{\pi}_2\sigma_n$ in Lemma 3.3.3 below.

We finally state and prove in Lemma 3.3.4 our formulas for $\pi_0\delta\sigma_n$ and $\pi_1\delta\sigma_n$, which correspond to part (i) of Theorem 3.1.10. In Subsection 3.3.4 we use this to deduce part (ii) of Theorem 3.1.10, which concerns the GP hierarchy. Lastly, we prove Proposition 3.1.14, which describes the perturbative formulation of the GP and $\widetilde{\mathsf{GP}}$ hierarchies.

3.3.1. Formulas for $\pi_0\sigma_n$ and $\pi_1\sigma_n$

Lemma 3.3.1. *We have*

$$\pi_0\sigma_n = \mathbb{1}_{\{n\ odd\}} C_{\frac{n-1}{2}} \overline{q}^{\frac{n+1}{2}} (-q)^{\frac{n+1}{2}} , \tag{3.3.7}$$

where C_n are the Catalan numbers defined in (3.2.45).

Proof. We prove by induction. The base case is trivial, so we assume the formula holds for n. Then

$$\begin{aligned}
\pi_0\sigma_{n+1} &= \sum_{k=1}^{n} \pi_0\sigma_k \pi_0\sigma_{n-k} \\
&= \sum_{k=1}^{n} \mathbb{1}_{\{k\text{ odd}\}} C_{\frac{k-1}{2}} \overline{q}^{\frac{k+1}{2}} (-q)^{\frac{k+1}{2}} \mathbb{1}_{\{n-k\text{ odd}\}} C_{\frac{n-k-1}{2}} \overline{q}^{\frac{n-k+1}{2}} (-q)^{\frac{n-k+1}{2}} \\
&= \mathbb{1}_{\{n+1\text{ odd}\}} \overline{q}^{\frac{n+2}{2}} (-q)^{\frac{n+2}{2}} \sum_{k=1}^{n} \mathbb{1}_{\{k\text{ odd}\}} C_{\frac{k-1}{2}} C_{\frac{n-k-1}{2}} ,
\end{aligned}$$

so by the definition of the Catalan numbers (3.2.45), we are done. □

Lemma 3.3.2. *We have the formula*

$$\pi_1\sigma_n = \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor -1} D_{n,j} (-q)^{j+1} \overline{q}^j \partial_x^{n-1-2j} \overline{q} + \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor -2} E_{n,j} (-q)^{j+1} \overline{q}^{j+2} \partial_x^{n-3-2j} q , \tag{3.3.8}$$

where

$$D_{n,j} = 4^j \binom{\frac{n}{2}-1}{j} \tag{3.3.9}$$

$$E_{n,j} = \sum_{l=0}^{j} (-1)^l (C_{l+1} - 2C_l) 4^{j-l} \binom{\frac{n-1}{2}-1}{j-l} . \tag{3.3.10}$$

Proof. Using (3.3.7), the iteration (3.3.4) simplifies to

$$\begin{aligned}
\pi_1\sigma_{n+1} = \partial_x (\mathbb{1}_{\{n\text{ odd}\}} C_{\frac{n-1}{2}} \overline{q}^{\frac{n+1}{2}} (-q)^{\frac{n+1}{2}}) + \pi_1 \partial_x \pi_1 \sigma_n + q_x \mathbb{1}_{\{n\text{ odd}\}} C_{\frac{n-1}{2}} \overline{q}^{\frac{n+1}{2}} (-q)^{\frac{n-1}{2}} \\
+ \sum_{k=1}^{n} 2\pi_1\sigma_k \mathbb{1}_{\{n-k\text{ odd}\}} C_{\frac{n-k-1}{2}} \overline{q}^{\frac{n-k+1}{2}} (-q)^{\frac{n-k+1}{2}} .
\end{aligned}$$

We plug (3.3.8) as an ansatz into this recurrence relation. After a lengthy calculation, which we do not present, we obtain the following recurrence relations for the coefficients:[4]

$$D_{2,0} = 1 \qquad D_{n+1,j} = 2\sum_{k=0}^{j-1} C_k D_{n-2k-1,j-1-k} + \begin{cases} C_{\frac{n-1}{2}} \frac{n+1}{2} & , j = \frac{n-1}{2} \\ D_{n,j} & , \text{ else} \end{cases} \qquad \forall 0 \le j \le \left\lfloor \frac{n+1}{2} \right\rfloor - 1$$

$$E_{4,0} = -1 \qquad E_{n+1,j} = 2\sum_{k=0}^{j-1} C_k E_{n-2k-1,j-1-k} + \begin{cases} -C_{\frac{n-1}{2}} \frac{n-1}{2} & , j = \frac{n-3}{2} \\ E_{n,j} & , \text{ else} \end{cases} \qquad \forall 0 \le j \le \left\lfloor \frac{n+1}{2} \right\rfloor - 2 .$$

When writing down such recurrence relations, we always set the coefficients that are not explicitly defined to zero (e. g. $D_{1,0} = E_{3,0} = 0$).

We now verify that (3.3.9)–(3.3.10) solve this recurrence relation. Note that

$$C_m(m+1) = \binom{2m}{m} = (-4)^m \binom{-\frac{1}{2}}{m} = 4^m \binom{\frac{1}{2}+m-1}{m} .$$

[4] See `test1`, `test2`, `test3`, and `test4` for a "sanity check" that for $n \in \{0, \ldots, 12\}$ our coefficients indeed satisfy these recurrence relations.

Since solutions to the recurrence are unique, to verify (3.3.9) it suffices to show that

$$\begin{aligned}
&4^j\binom{\frac{n+1}{2}-1}{j} = 2\sum_{k=0}^{j-1} C_k 4^{j-1-k}\binom{\frac{n-1}{2}-1-k}{j-1-k} + 4^j\binom{\frac{n}{2}-1}{j} \qquad 0\le j\le \left\lfloor\frac{n+1}{2}\right\rfloor -1\\
\iff &\binom{\frac{n+1}{2}-1}{j+1} - \binom{\frac{n}{2}-1}{j+1} = \frac{1}{2}\sum_{k=0}^{j} 4^{-k}C_k\binom{\frac{n-1}{2}-1-k}{j-k} \qquad \forall -1\le j\le \left\lfloor\frac{n+1}{2}\right\rfloor\\
\iff &\binom{\frac{s}{2}+1}{j+1} - \binom{\frac{s}{2}+\frac{1}{2}}{j+1} = \frac{1}{2}\sum_{k=0}^{j} 4^{-k}C_k\binom{\frac{s}{2}-k}{j-k} \qquad \text{where } s=n-3\,.
\end{aligned}$$

This identity is verified by tedious calculation and subsequent comparison of the generating functions[5]

$$\sum_{s,j\ge 0} X^sY^j\binom{\frac{s}{2}+1}{j+1} - \sum_{s,j\ge 0} X^sY^j\binom{\frac{s}{2}+\frac{1}{2}}{j+1} = \frac{1+Y-\sqrt{1+Y}}{Y}\frac{1}{1-X\sqrt{1+Y}}$$

and[6]

$$\sum_{s,j\ge 0} X^sY^j\frac{1}{2}\sum_{k=0}^{j}4^{-k}C_k\binom{\frac{s}{2}-k}{j-k} = \frac{1+Y-\sqrt{1+Y}}{Y}\frac{1}{1-X\sqrt{1+Y}}\,.$$

We proceed similarly for the coefficients $E_{n,j}$. Here it suffices to show that for $0\le j\le \left\lfloor\frac{n+1}{2}\right\rfloor -2$ we have

$$\begin{aligned}
\sum_{l=0}^{j}(-1)^l(C_{l+1}-2C_l)4^{-l}\binom{\frac{n}{2}-1}{j-l} &= 2\sum_{k=0}^{j-1}C_k\sum_{l=0}^{j-1-k}(-1)^l(C_{l+1}-2C_l)4^{-1-k-l}\binom{\frac{n}{2}-2-k}{j-1-k-l}\\
&\quad+\begin{cases} -4^{-j}C_{\frac{n-1}{2}}\frac{n-1}{2} & ,j=\frac{n-3}{2}\\ \sum_{l=0}^{j}(-1)^l(C_{l+1}-2C_l)4^{-l}\binom{\frac{n-1}{2}-1}{j-l} & ,0\le j\le \left\lfloor\frac{n}{2}\right\rfloor -2\end{cases}.
\end{aligned}$$

If $j=\frac{n-3}{2}$ the two expressions at the end agree, which we verify by calculating [7]

$$\sum_{m=0}^{\infty}Y^m\sum_{l=0}^{m}(-1)^l(C_{l+1}-2C_l)4^{-l}\binom{m}{m-l} = \frac{8}{Y^2}-\frac{4\sqrt{1-Y}}{Y^2}-\frac{4}{Y^2\sqrt{1-Y}}$$

and[8]

$$-\sum_{m=0}^{\infty}4^{-m}Y^mC_{m+1}(m+1) = \frac{8}{Y^2}-4\frac{\sqrt{1-Y}}{Y^2}-\frac{4}{Y^2\sqrt{1-Y}}\,.$$

After removing the case distinction, it remains to show that

$$\begin{aligned}
&\sum_{l=0}^{j}(-1)^l(C_{l+1}-2C_l)4^{-l}\left(\binom{\frac{n}{2}-1}{j-l}-\binom{\frac{n-1}{2}-1}{j-l}\right)\\
&=2\sum_{k=0}^{j-1}C_k\sum_{l=0}^{j-1-k}(-1)^l(C_{l+1}-2C_l)4^{-1-k-l}\binom{\frac{n}{2}-2-k}{j-1-k-l}\,.
\end{aligned}$$

This follows by yet another lengthy calculation of the generating functions[9]

$$\begin{aligned}
&\sum_{n,j\ge 0}X^nY^j\sum_{l=0}^{j}(C_{l+1}-2C_l)4^{-l}\left(\binom{\frac{n}{2}-1}{j-l}-\binom{\frac{n-1}{2}-1}{j-l}\right)\\
&=\left(2\left(\frac{4}{Y}-2\right)\frac{1-\sqrt{1-Y}}{Y}-\frac{4}{Y}\right)\frac{1}{1+Y}\frac{1}{1-X\sqrt{1+Y}}\left(1-\frac{1}{\sqrt{1+Y}}\right)
\end{aligned}$$

[5] See `test5`
[6] See `test6`
[7] See `test7`
[8] See `test8`
[9] See `test9`

and[10]

$$\sum_{n,j\geq 0} X^n Y^j 2\sum_{k=0}^{j-1} C_k \sum_{l=0}^{j-1-k} (C_{l+1}-2C_l)4^{-1-k-l}\binom{\frac{n}{2}-2-k}{j-1-k-l}$$
$$=\left(2\Big(\frac{4}{Y}-2\Big)\frac{1-\sqrt{1-Y}}{Y}-\frac{4}{Y}\right)\frac{1}{1+Y}\frac{1}{1-X\sqrt{1+Y}}\left(1-\frac{1}{\sqrt{1+Y}}\right).$$

□

3.3.2. FORMULAS FOR $\pi_1\delta\widetilde{\pi}_2\sigma_n$

Lemma 3.3.3. *We have the formula*

$$\pi_1\delta\widetilde{\pi}_2\sigma_n = \sum_{j=0}^{\lfloor\frac{n-1}{2}\rfloor-2}(-q)^{j+2}\overline{q}^j\partial_x^{n-3-2j}\overline{q}\widetilde{F}_{n,j} + \sum_{j=0}^{\lfloor\frac{n-1}{2}\rfloor-1}(-q)^j\overline{q}^j\partial_x^{n-1-2j}q\widetilde{G}_{n,j}\,, \tag{3.3.11}$$

where

$$\widetilde{F}_{2m,j} = 0 \tag{3.3.12}$$
$$\widetilde{F}_{2m+1,j} = -(8m-8j-6)4^j\binom{m-\frac{1}{2}}{j} \tag{3.3.13}$$
$$\widetilde{G}_{2m,j} = -\sum_{k=0}^{j-1}C_k4^{j-k}(m-1-j)\binom{m-\frac{3}{2}-k}{j-1-k} \tag{3.3.14}$$
$$\widetilde{G}_{2m+1,j} = 4^{j-1}2\binom{m-\frac{1}{2}}{j-1}+\sum_{k=0}^{j-1}C_k4^{j-k}(m-j)\binom{m-\frac{1}{2}-k}{j-1-k}\,. \tag{3.3.15}$$

Proof. We first consider $\widetilde{\pi}_2\sigma_n$ and then $\pi_1\delta\widetilde{\pi}_2\sigma_n$.

Recurrence relation for the coefficients $G_{n,j,t}, F_{n,j,t}$ in $\widetilde{\pi}_2\sigma_n$. Recall the recurrence relation (3.3.6) for $\widetilde{\pi}_2\sigma_n$:

$$\widetilde{\pi}_2\sigma_1 = 0$$
$$\widetilde{\pi}_2\sigma_{n+1} = \widetilde{\pi}_2\partial_x\widetilde{\pi}_2\sigma_n - \frac{q_x}{q}\widetilde{\pi}_1\sigma_n + \widetilde{\pi}_2\partial_x\pi_1\sigma_n + 2\sum_{k=1}^{n-1}(\widetilde{\pi}_2\sigma_k\pi_0\sigma_{n-k}+\widetilde{\pi}_1\sigma_k\pi_1\sigma_{n-k})\,.$$

We make the ansatz

$$\begin{aligned}\widetilde{\pi}_2\sigma_n = &\sum_{j=1}^{\lfloor\frac{n-1}{2}\rfloor-1}(-q)^j\overline{q}^j\sum_{t=1}^{n-2-2j}G_{n,j,t}\partial_x^{n-1-2j-t}q\partial_x^t\overline{q}\\ &+\sum_{j=0}^{\lfloor\frac{n-1}{2}\rfloor-2}(-q)^{j+2}\overline{q}^j\sum_{t=1}^{\lfloor\frac{n-1}{2}\rfloor-1-j}F_{n,j,t}\partial_x^{n-3-2j-t}\overline{q}\partial_x^t\overline{q}\,.\end{aligned} \tag{3.3.16}$$

Unfortunately, we must now plug this ansatz, as well as the formulas for $\pi_0\sigma_n$ and $\pi_1\sigma_n$ that we have obtained, into this recurrence relation. A lengthy calculation, which we spare the reader, yields the following recurrence relations for $G_{n,j,t}$ and $F_{n,j,t}$. We set $G_{5,1,1}=-6$ and for $n\geq 0$, $1\leq j\leq \lfloor\frac{n}{2}\rfloor-1$, $1\leq t\leq n-1-2j$ we have the following recurrence relation for $G_{n,j,t}$:

$$\begin{aligned}G_{n+1,j,t} = &\,2\sum_{k=0}^{j-2}G_{n-2k-1,j-k-1,t}C_k + 2\sum_{k=0}^{j-2}D_{t+1+2k,k}E_{n-1-2k-t,j-2-k}\\ &+\mathbb{1}_{\{j\neq\frac{n}{2}-1\}}\mathbb{1}_{\{t\neq n-1-2j\}}G_{n,j,t}+\mathbb{1}_{\{j\neq\frac{n}{2}-1\}}\mathbb{1}_{\{t\neq 1\}}G_{n,j,t-1}\\ &+\mathbb{1}_{\{t=1\}}(j+1)E_{n,j-1}-\mathbb{1}_{\{t=n-1-2j\}}jD_{n,j}\,.\end{aligned}$$

[10] See `test10`

Similarly, we set $F_{5,0,1} = 5$ and for $n \geq 5$, $0 \leq j \leq \lfloor \frac{n}{2} \rfloor - 2$, $1 \leq t \leq \lfloor \frac{n}{2} \rfloor - 1 - j$ we have

$$\begin{aligned} F_{n+1,j,t} &= 2\sum_{k=0}^{j-1} F_{n-2k-1,j-k-1,t}C_k + (1 + \mathbb{1}_{\{t\neq\frac{n}{2}-1-j\}})\sum_{k=0}^{j} D_{t+1+2k,k}D_{n-1-2k-t,j-k} \\ &\quad + \mathbb{1}_{\{t=1\}}(j+1)D_{n,j+1} + \mathbb{1}_{\{j\neq\frac{n}{2}-2\}}\Big(\mathbb{1}_{\{t\neq\frac{n}{2}-1-j\}}F_{n,j,t} + \mathbb{1}_{\{t\neq1\}}F_{n,j,t-1} + \mathbb{1}_{\{t=\frac{n-1}{2}-1-j\}}F_{n,j,t}\Big). \end{aligned}$$

Recurrence relations for the coefficients $\widetilde{G}_{n,j}, \widetilde{F}_{n,j}$ in $\pi_1\delta\widetilde{\pi}_2\sigma_n$. The ansatz (3.3.16) implies the formula for $\pi_1\delta\widetilde{\pi}_2\sigma_n$ of the desired form (3.3.11), with the coefficients given by

$$\widetilde{G}_{n,j} = \sum_{t=1}^{n-2-2j} (-1)^t G_{n,j,t} \qquad\qquad \widetilde{F}_{n,j} = (1-(-1)^n)\sum_{t=1}^{\lfloor\frac{n-1}{2}\rfloor-1-j} (-1)^t F_{n,j,t}\,.$$

Then[11]

$$\begin{aligned} \widetilde{G}_{n+1,j} &= 2\sum_{k=0}^{j-2}\sum_{t=1}^{n-1-2j} (-1)^t G_{n-2k-1,j-k-1,t}C_k + 2\sum_{k=0}^{j-2}\sum_{t=1}^{n-1-2j} (-1)^t D_{t+1+2k,k}E_{n-1-2k-t,j-2-k} \\ &\quad + \mathbb{1}_{\{j\neq\frac{n}{2}-1\}}\sum_{t=1}^{n-1-2j} (-1)^t(\mathbb{1}_{\{t\neq n-1-2j\}}G_{n,j,t} + \mathbb{1}_{\{t\neq1\}}G_{n,j,t-1}) \\ &\quad + \sum_{t=1}^{n-1-2j} (-1)^t(\mathbb{1}_{\{t=1\}}(j+1)E_{n,j-1} - \mathbb{1}_{\{t=n-1-2j\}}jD_{n,j}) \\ &= 2\sum_{k=0}^{j-1}\widetilde{G}_{n-2k-1,j-k-1}C_k + 2\sum_{k=0}^{j-2}\sum_{t=1}^{n-1-2j} (-1)^t D_{t+1+2k,k}E_{n-1-2k-t,j-2-k} \\ &\quad - (j+1)E_{n,j-1} + (-1)^n jD_{n,j} \end{aligned}$$

and[12]

$$\begin{aligned} \widetilde{F}_{n+1,j} &= 2\mathbb{1}_{\{n \text{ even}\}}\Bigg(2\sum_{k=0}^{j-1}\sum_{t=1}^{\lfloor\frac{n}{2}\rfloor-1-j} (-1)^t F_{n-2k-1,j-k-1,t}C_k \\ &\quad + \sum_{k=0}^{j}\sum_{t=1}^{\lfloor\frac{n}{2}\rfloor-1-j} (-1)^t(1+\mathbb{1}_{\{t\neq\frac{n}{2}-1-j\}})D_{t+1+2k,k}D_{n-1-2k-t,j-k} + \sum_{t=1}^{\lfloor\frac{n}{2}\rfloor-1-j}(-1)^t\Big(\mathbb{1}_{\{t=1\}}(j+1)D_{n,j+1} \\ &\quad + \mathbb{1}_{\{j\neq\frac{n}{2}-2\}}\Big(\mathbb{1}_{\{t\neq\frac{n}{2}-1-j\}}F_{n,j,t} + \mathbb{1}_{\{t\neq1\}}F_{n,j,t-1} + \mathbb{1}_{\{t=\frac{n-1}{2}-1-j\}}F_{n,j,t}\Big)\Big)\Bigg) \\ &= 2\mathbb{1}_{\{n \text{ even}\}}\Bigg(\sum_{k=0}^{j-1}\widetilde{F}_{n-2k-1,j-k-1}C_k + \sum_{k=0}^{j}\sum_{t=1}^{\lfloor\frac{n}{2}\rfloor-1-j}(-1)^t(1+\mathbb{1}_{\{t\neq\frac{n}{2}-1-j\}})D_{t+1+2k,k}D_{n-1-2k-t,j-k} \\ &\quad - (j+1)D_{n,j+1} + \mathbb{1}_{\{n \text{ odd}\}}2(-1)^{\frac{n-1}{2}-1-j}F_{n,j,\frac{n-1}{2}-1-j}\Bigg) \\ &= \mathbb{1}_{\{n \text{ even}\}}\Bigg(2\sum_{k=0}^{j-1}\widetilde{F}_{n-2k-1,j-k-1}C_k + 2\sum_{k=0}^{j}\sum_{t=1}^{\lfloor\frac{n}{2}\rfloor-1-j}(-1)^t(1+\mathbb{1}_{\{t\neq\frac{n}{2}-1-j\}})D_{t+1+2k,k}D_{n-1-2k-t,j-k} \\ &\quad - 2(j+1)D_{n,j+1}\Bigg). \end{aligned}$$

Verification of the formula (3.3.14)–(3.3.15) for $\widetilde{G}_{n,j}$. Recall the formulas (3.3.9) and (3.3.10) for $D_{n,j}$ and $E_{n,j}$. A calculation with binomial identities (or alternatively verified manually using

[11] See `test11`
[12] See `test12`

generating functions) yields[13]

$$\sum_{k=0}^{j-2}\sum_{t=1}^{n-1-2j}(-1)^t D_{t+1+2k,k}E_{n-1-2k-t,j-2-k} = -\mathbb{1}_{\{n \text{ even}\}}E_{n,j-2}\,.$$

Therefore the recurrence relation for $\widetilde{G}_{n,j}$ becomes[14]

$$\begin{aligned}\widetilde{G}_{n+1,j} &= 2\sum_{k=0}^{j-2}\widetilde{G}_{n-2k-1,j-k-1}C_k - 2\mathbb{1}_{\{n \text{ even}\}}\sum_{l=0}^{j-2}(C_{l+1}-2C_l)(-1)^l 4^{j-2-l}\binom{\frac{n-1}{2}-1}{j-2-l}\\ &\quad-(j+1)\sum_{l=0}^{j-1}(C_{l+1}-2C_l)(-1)^l 4^{j-1-l}\binom{\frac{n-1}{2}-1}{j-1-l}+(-1)^n j4^j\binom{\frac{n}{2}-1}{j}\\ &= 2\sum_{k=0}^{j-2}\widetilde{G}_{n-2k-1,j-k-1}C_k+(-1)^n j4^j\binom{\frac{n}{2}-1}{j}\\ &\quad-\sum_{l=0}^{j-1}(C_{l+1}-2C_l)(-1)^l\left((j+1)4^{j-1-l}\binom{\frac{n-1}{2}-1}{j-1-l}+2\mathbb{1}_{\{n \text{ even}\}}4^{j-2-l}\binom{\frac{n-1}{2}-1}{j-2-l}\right).\end{aligned}$$

A set of tedious generating function calculations needs to be performed to verify that the desired formulas (3.3.14)–(3.3.15) for $\widetilde{G}_{n,j}$:

$$\widetilde{G}_{n,j} = (-1)^{n+1}4\left(\left\lfloor\frac{n-1}{2}\right\rfloor - j\right)\sum_{k=0}^{j-1}C_k 4^{j-1-k}\binom{\lfloor\frac{n-1}{2}\rfloor-\frac{1}{2}-k}{j-1-k}+2\mathbb{1}_{\{n \text{ odd}\}}4^{j-1}\binom{\frac{n}{2}-1}{j-1}$$

satisfy the above recurrence relation. We calculate from this formula that[15]

$$\begin{aligned}\sum_{n=0,j=0}^{\infty}X^nY^j\widetilde{G}_{n+1,j} &= \frac{-4Y}{\sqrt{1+4Y}}\frac{-X}{1-X^2(1+4Y)}+\frac{-2Y}{\sqrt{1+4Y}}\frac{1}{1-X^2(1+4Y)}\\ &\quad+\left(1-\frac{1}{\sqrt{1+4Y}}\right)2\sqrt{1+4Y}\frac{X^2(1-X)}{(1-X^2(1+4Y))^2}\,.\end{aligned}$$

Next, we calculate[16]

$$\begin{aligned}\sum_{n=0,j=0}^{\infty}X^nY^j 2\sum_{k=0}^{j-2}C_k\widetilde{G}_{n-2k-1,j-k-1} &= \left(1-\frac{1}{\sqrt{1+4Y}}\right)^2 2\sqrt{1+4Y}\frac{X^2(1-X)}{(1-X^2(1+4Y))^2}\\ &\quad+\left(1-\frac{1}{\sqrt{1+4Y}}\right)\frac{2Y}{\sqrt{1+4Y}}\left(-2-\frac{2}{\sqrt{1+4Y}}\right)\frac{-X}{1-X^2(1+4Y)}\\ &\quad+\left(1-\frac{1}{\sqrt{1+4Y}}\right)\frac{2Y}{\sqrt{1+4Y}}\left(-1-\frac{2}{\sqrt{1+4Y}}\right)\frac{1}{1-X^2(1+4Y)}\,.\end{aligned}$$

Lastly, we calculate[17]

$$\begin{aligned}&\sum_{n=0,j=0}^{\infty}X^nY^j(-1)^n j4^j\binom{\frac{n}{2}-1}{j}\\ &=\left(\frac{1}{\sqrt{1+4Y}}-X\right)\frac{4Y}{\sqrt{1+4Y}}\frac{X^2}{(1-X^2(1+4Y))^2}+\left(\frac{1}{2}X-\frac{1}{\sqrt{1+4Y}}\right)\frac{4Y}{\sqrt{1+4Y}^3}\frac{1}{1-X^2(1+4Y)}\end{aligned}$$

and[18]

$$-\sum_{n=0,j=0}^{\infty}X^nY^j\sum_{l=0}^{j-1}(C_{l+1}-2C_l)(-1)^l\left((j+1)4^{j-1-l}\binom{\frac{n-1}{2}-1}{j-1-l}+2\mathbb{1}_{\{n \text{ even}\}}4^{j-2-l}\binom{\frac{n-1}{2}-1}{j-2-l}\right)$$

---

[13] See `test13`
[14] See `test14`
[15] See `test15`
[16] See `test16`
[17] See `test17`
[18] See `test18`

$$= \left(\sqrt{1+4Y} + \frac{1}{\sqrt{1+4Y}} - 2\right)\left(X + \frac{1}{\sqrt{1+4Y}}\right)\frac{X^2}{(1-X^2(1+4Y))^2}$$
$$+ \frac{-2Y}{\sqrt{1+4Y}^3}\frac{-X}{1-X^2(1+4Y)} + \left(\frac{1}{2} - \frac{1}{\sqrt{1+4Y}} + \frac{1}{2}\frac{1}{1+4Y} + \frac{1}{\sqrt{1+4Y}^3} - \frac{1}{(1+4Y)^2}\right)\frac{1}{1-X^2(1+4Y)}.$$

Summing up all the contributions[19] verifies that the claimed formula (3.3.14)–(3.3.15) for $\widetilde{G}_{n,j}$ satisfies the given recurrence relation.

Verification of the formula (3.3.12)–(3.3.13) **for $\widetilde{F}_{n,j}$.** For even $n = 2m$ we calculate[20]

$$\sum_{k=0}^{j}\sum_{t=1}^{\lfloor\frac{n}{2}\rfloor-1-j}(-1)^t(1+\mathbb{1}_{\{t\neq\frac{n}{2}-1-j\}})D_{t+1+2k,k}D_{n-1-2k-t,j-k} = -\mathbb{1}_{\{j\leq m-2\}}(-4)^j\binom{j-m}{j}.$$

We obtain the following recurrence relation

$$\widetilde{F}_{2m+1,j} = 2\sum_{k=0}^{j-1}\widetilde{F}_{2m-2k-1,j-k-1}C_k - 2(-4)^j\binom{j-m}{j} - 2(j+1)4^{j+1}\binom{m-1}{j+1}.$$

We then aim to verify that the formula given in (3.3.12)–(3.3.13)

$$\widetilde{F}_{n,j} = -\mathbb{1}_{\{n\text{ odd}\}}\left(8\left\lfloor\frac{n-1}{2}\right\rfloor - 8j - 6\right)4^j\binom{\frac{n}{2}-1}{j} = \begin{cases} -(8m-8j-6)4^j\binom{m-\frac{1}{2}}{j} & , n = 2m+1 \\ 0 & , n = 2m \end{cases}$$

solves this recurrence relation.

Plugging everything into the recurrence relation yields the trivial identity for odd n, and for even $n = 2m$ the binomial identity[21]

$$(8m-8j-6)\binom{m-\frac{1}{2}}{j} = 2\sum_{k=0}^{j-1}(8m-8j-6)4^{-k-1}\binom{m-k-\frac{3}{2}}{j-k-1}C_k + 2(-1)^j\binom{j-m}{j} + 8(j+1)\binom{m-1}{j+1}.$$

It remains to prove this for all $(m,j)\in\mathbb{Z}^2$. Once more, we calculate and compare the generating functions[22]

$$\sum_{m=0}^{\infty}\sum_{j=0}^{\infty}X^mY^j2\sum_{k=0}^{j-1}(8(m-j)-6)4^{-k-1}\binom{m-k-1-\frac{1}{2}}{j-k-1}C_k$$
$$= \left(-8\frac{Y}{\sqrt{1+Y}^3} + 8\frac{1}{\sqrt{1+Y}}\right)\frac{X(1+Y)}{(1-X(1+Y))^2} + \left(4\frac{Y}{\sqrt{1+Y}^3} - 6\frac{1}{\sqrt{1+Y}}\right)\frac{1}{1-X(1+Y)}$$
$$- \frac{8}{(1+Y)^2}\frac{X(1+Y)}{(1-X(1+Y))^2} + \frac{6-2Y}{(1+Y)^2}\frac{1}{1-X(1+Y)}$$

and[23]

$$\sum_{m=0}^{\infty}\sum_{j=0}^{\infty}X^mY^j2(-1)^j\binom{-m+j}{j} + \sum_{m=0}^{\infty}\sum_{j=0}^{\infty}X^mY^j8(j+1)\binom{m-1}{j+1}$$
$$= \frac{2Y-6}{(1+Y)^2}\frac{1}{1-X(1+Y)} + \frac{8}{(1+Y)^2}\frac{X(1+Y)}{(1-X(1+Y))^2}$$

with[24]

$$\sum_{m=0}^{\infty}\sum_{j=0}^{\infty}X^mY^j(8(m-j)-6)\binom{m-\frac{1}{2}}{j}$$
$$= \left(-8\frac{Y}{\sqrt{1+Y}^3} + 8\frac{1}{\sqrt{1+Y}}\right)\frac{X(1+Y)}{(1-X(1+Y))^2} + \left(4\frac{Y}{\sqrt{1+Y}^3} - 6\frac{1}{\sqrt{1+Y}}\right)\frac{1}{1-X(1+Y)}.$$

□

[19] See `test19`
[20] See `test20`
[21] See `test21`
[22] See `test22`
[23] See `test23`
[24] See `test24`

3.3.3. Formulas for $\pi_0\delta\sigma_n$ and $\pi_1\delta\sigma_n$

Lemma 3.3.4. *We have*

$$\pi_0\delta\sigma_n = \mathbb{1}_{\{n \ odd\}}\frac{n+1}{2}C_{\frac{n-1}{2}}\overline{q}^{\frac{n-1}{2}}(-q)^{\frac{n+1}{2}} \tag{3.3.17}$$

$$\pi_1\delta\sigma_n = \sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-2} J_{n,j}(-q)^{j+2}\overline{q}^j\partial_x^{n-3-2j}\overline{q} + \sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-1} K_{n,j}(-q)^j\overline{q}^j\partial_x^{n-1-2j}q\,, \tag{3.3.18}$$

where $J_{n,j}$ and $K_{n,j}$ are defined in (3.1.19).

Proof. Recall Lemmas 3.3.1, 3.3.2 and 3.3.3. We easily obtain the formula (3.3.17) for $\pi_0\delta\sigma_n = \pi_0\delta\pi_0\sigma_n$ from Lemma 3.3.1. From Lemma 3.3.2 we have

$$\begin{aligned}
\pi_1\delta\pi_1\sigma_n &= \sum_{j=1}^{\lfloor\frac{n}{2}\rfloor-1} jD_{n,j}(-q)^{j+1}\overline{q}^{j-1}\partial_x^{n-1-2j}\overline{q} + \sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-2}(j+2)E_{n,j}(-q)^{j+1}\overline{q}^{j+1}\partial_x^{n-3-2j}q\\
&+\pi_1\sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-1}(-1)^{n-1-2j}D_{n,j}\sum_{k=0}^{n-1-2j}\binom{n-1-2j}{k}\partial_x^k((-q)^{j+1})\partial_x^{n-1-2j-k}(\overline{q}^j)\\
&=\sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-2}(j+1)D_{n,j+1}(-q)^{j+2}\overline{q}^j\partial_x^{n-3-2j}\overline{q} + \sum_{j=1}^{\lfloor\frac{n}{2}\rfloor-1}(j+1)E_{n,j-1}(-q)^j\overline{q}^j\partial_x^{n-1-2j}q\\
&+\sum_{j=1}^{\lfloor\frac{n}{2}\rfloor-1} j(-1)^{n-1}D_{n,j}(-q)^{j+1}\overline{q}^{j-1}\partial_x^{n-1-2j}\overline{q} + \sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-1}(j+1)(-1)^nD_{n,j}(-q)^j\overline{q}^j\partial_x^{n-1-2j}q\,,
\end{aligned}$$

from which we obtain

$$\begin{aligned}
\pi_1\delta\pi_1\sigma_n = (1-(-1)^n)&\sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-2}(j+1)D_{n,j+1}(-q)^{j+2}\overline{q}^j\partial_x^{n-3-2j}\overline{q}\\
&+\sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-1}(j+1)(\mathbb{1}_{\{j\geq1\}}E_{n,j-1}+(-1)^nD_{n,j})(-q)^j\overline{q}^j\partial_x^{n-1-2j}q\,.
\end{aligned}$$

This, together with Lemma 3.3.3, implies

$$\begin{aligned}
\pi_1\delta\sigma_n &= \sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-2}\big(\mathbb{1}_{\{0\leq j\leq\lfloor\frac{n-1}{2}\rfloor-2\}}\widetilde{F}_{n,j}+\mathbb{1}_{\{n\text{ odd}\}}2(j+1)D_{n,j+1}\big)(-q)^{j+2}\overline{q}^j\partial_x^{n-3-2j}\overline{q}\\
&+\sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-1}\big(\mathbb{1}_{\{1\leq j\leq\lfloor\frac{n-1}{2}\rfloor-1\}}\widetilde{G}_{n,j}+(j+1)(\mathbb{1}_{\{j\geq1\}}E_{n,j-1}+(-1)^nD_{n,j})\big)(-q)^j\overline{q}^j\partial_x^{n-1-2j}q\\
&=\sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-2}J_{n,j}(-q)^{j+2}\overline{q}^j\partial_x^{n-3-2j}\overline{q} + \sum_{j=0}^{\lfloor\frac{n}{2}\rfloor-1}K_{n,j}(-q)^j\overline{q}^j\partial_x^{n-1-2j}q
\end{aligned}$$

for certain coefficients $J_{n,j}$ and $K_{n,j}$, whose claimed formulas remain to be verified by simplifying the expressions above.

Verification of the formula (3.1.19) **for $J_{n,j}$.** Recall the definitions of $D_{n,j}$, $E_{n,j}$, $\widetilde{F}_{n,j}$, and $\widetilde{G}_{n,j}$. Let $0\leq j\leq\lfloor\frac{n}{2}\rfloor-2$ and $n=2m$ or $n=2m+1$. We have[25]

$$\begin{aligned}
J_{2m,j} &= \mathbb{1}_{\{j\leq\lfloor\frac{2m-1}{2}\rfloor-2\}}\widetilde{F}_{2m,j}+\mathbb{1}_{\{2m\text{ odd}\}}2(j+1)D_{2m,j+1}=0\\
J_{2m+1,j} &= \mathbb{1}_{\{j\leq\lfloor\frac{2m+1-1}{2}\rfloor-2\}}\widetilde{F}_{2m+1,j}+\mathbb{1}_{\{2m+1\text{ odd}\}}2(j+1)D_{2m+1,j+1}\\
&=-\mathbb{1}_{\{j\leq m-2\}}(8m-8j-6)4^j\binom{m-\frac{1}{2}}{j}+2(j+1)4^{j+1}\binom{m-\frac{1}{2}}{j+1}
\end{aligned}$$

[25]See `test25`

$$= (8j+8)4^j \binom{m+\frac{1}{2}}{j+1} - (8m+2)4^j \binom{m-\frac{1}{2}}{j}.$$

A computation shows that[26]

$$\sum_{m=0}^{\infty}\sum_{j=0}^{\infty} X^m Y^j (8j+8)4^j \binom{m+\frac{1}{2}}{j+1} - \sum_{m=0}^{\infty}\sum_{j=0}^{\infty} X^m Y^j (8m+2)4^j \binom{m-\frac{1}{2}}{j}$$
$$= \frac{2}{\sqrt{1+4Y}} \frac{1}{1-X(1+4Y)} = \sum_{m=0}^{\infty}\sum_{j=0}^{\infty} X^m Y^j 4^j 2 \binom{m-\frac{1}{2}}{j}.$$

Hence

$$J_{2m+1,j} = 4^j 2 \binom{m-\frac{1}{2}}{j} \qquad \text{and} \qquad J_{n,j} = 2\mathbb{1}_{\{n \text{ odd}\}} 4^j \binom{\frac{n}{2}-1}{j}. \tag{3.3.19}$$

Using the notation $[u^j]$ for the extraction of the coefficient in front of u^j from a formal power series in the symbol u, we can also write this as in (3.1.19):

$$J_{n,j} = \mathbb{1}_{\{n \text{ odd}\}} [u^j] 2(1+4u)^{\frac{n-2}{2}}. \tag{3.3.20}$$

Verification of the formula (3.1.19) for $K_{n,j}$. Now let $0 \le j \le \lfloor \frac{n}{2} \rfloor - 1$. We have[27]

$$\begin{aligned}
K_{n,j} &= \mathbb{1}_{\{1 \le j \le \lfloor \frac{n-1}{2} \rfloor - 1\}} \widetilde{G}_{n,j} + (j+1)(\mathbb{1}_{\{j \ge 1\}} E_{n,j-1} + (-1)^n D_{n,j}) \\
&= \widetilde{G}_{n,j} + (j+1)(E_{n,j-1} + (-1)^n D_{n,j}) \\
&= (-1)^{n+1} 4 \left(\left\lfloor \frac{n-1}{2} \right\rfloor - j \right) \sum_{k=0}^{j-1} C_k 4^{j-1-k} \binom{\lfloor \frac{n-1}{2} \rfloor - \frac{1}{2} - k}{j-1-k} + 2\mathbb{1}_{\{n \text{ odd}\}} 4^{j-1} \binom{\frac{n}{2}-1}{j-1} \\
&\quad + (j+1) \left(\mathbb{1}_{\{j \ge 1\}} \sum_{l=0}^{j-1} (-1)^l (C_{l+1} - 2C_l) 4^{j-1-l} \binom{\frac{n-1}{2}-1}{j-1-l} + (-1)^n 4^j \binom{\frac{n}{2}-1}{j} \right).
\end{aligned}$$

We claim that[28]

$$K_{n,j} = (-1)^n D_{n+1,j} + 2\mathbb{1}_{\{n \text{ odd}\}} E_{n,j-2} \tag{3.3.21}$$

$$= (-1)^n 4^j \binom{\frac{n+1}{2}-1}{j} + 2\mathbb{1}_{\{n \text{ odd}\}} \sum_{l=0}^{j-2} (-1)^l (C_{l+1} - 2C_l) 4^{j-2-l} \binom{\frac{n-1}{2}-1}{j-2-l}. \tag{3.3.22}$$

$$= \begin{cases} [u^j](1+4u)^{\frac{n-1}{2}} & , n \text{ even} \\ [u^j](-1-2u)(1+4u)^{\frac{n-2}{2}} & , n \text{ odd} \end{cases}. \tag{3.3.23}$$

Note that[29]

$$\sum_{n=0,j=0}^{\infty} X^n Y^j (-1)^n 4^j \binom{\frac{n+1}{2}-1}{j} + \sum_{n=0,j=0}^{\infty} X^n Y^j 2\mathbb{1}_{\{n \text{ odd}\}} \sum_{l=0}^{j-2} (-1)^l (C_{l+1} - 2C_l) 4^{j-2-l} \binom{\frac{n-1}{2}-1}{j-2-l}$$
$$= \frac{1}{\sqrt{1+4Y}} \frac{1}{1-X^2(1+4Y)} + \left(-\frac{1}{2} \frac{1}{\sqrt{1+4Y}} - \frac{1}{2}\sqrt{1+4Y} \right) \frac{X}{1-X^2(1+4Y)}.$$
$$= \frac{1}{\sqrt{1+4Y}} \frac{1-X(1+2Y)}{1-X^2(1+4Y)}$$
$$= \sum_{m=0,j=0}^{\infty} X^{2m} Y^j [u^j](1+4u)^{\frac{2m-1}{2}} + \sum_{m=0,j=0}^{\infty} X^{2m+1} Y^j [u^j](-1-2u)(1+4u)^{\frac{2m-1}{2}}.$$

This shows that (3.3.22) and (3.3.23) agree.

---

[26] See `test26`
[27] See `test27`
[28] See `test28`
[29] See `test29`

To show that these are indeed formulas for $K_{n,j}$, we have to compute the following generating function, which mostly reduces to previous calculations:

$$\sum_{n=0,j=0}^{\infty} X^n Y^j K_{n,j}$$
$$= X \sum_{n=0,j=0}^{\infty} X^n Y^j \widetilde{G}_{n+1,j} + \sum_{j=0}^{\infty} Y^j \widetilde{G}_{0,j} + \sum_{n=0,j=0}^{\infty} X^n Y^j (j+1) E_{n,j-1} + \sum_{n=0,j=0}^{\infty} X^n Y^j (j+1)(-1)^n D_{n,j}\,.$$

Recall that

$$X \sum_{n=0,j=0}^{\infty} X^n Y^j \widetilde{G}_{n+1,j} = X\Bigg(\frac{-4Y}{\sqrt{1+4Y}} \frac{-X}{1-X^2(1+4Y)} + \frac{-2Y}{\sqrt{1+4Y}} \frac{1}{1-X^2(1+4Y)}$$
$$+ \left(1 - \frac{1}{\sqrt{1+4Y}}\right) 2\sqrt{1+4Y} \frac{X^2(1-X)}{(1-X^2(1+4Y))^2} \Bigg)\,.$$

For the new term involving $\widetilde{G}_{0,j}$, we have[30]

$$\sum_{j=0}^{\infty} Y^j \widetilde{G}_{0,j} = \frac{1}{2}\left(1 - \frac{1}{\sqrt{1+4Y}}\right)\left(-\frac{1}{2} \frac{16Y}{\sqrt{1+4Y}^3} + \frac{4}{\sqrt{1+4Y}}\right) + \frac{Y}{\sqrt{1+4Y}}\left(\frac{-16Y}{\sqrt{1+4Y}^3} + \frac{4}{\sqrt{1+4Y}}\right)$$
$$= \frac{1}{\sqrt{1+4Y}} + \frac{1}{\sqrt{1+4Y}^3} - \frac{2}{(1+4Y)^2}\,.$$

Lastly, we calculate[31]

$$\sum_{n=0,j=0}^{\infty} X^n Y^j (j+1)(-1)^n D_{n,j} = \left(\frac{1}{\sqrt{1+4Y}} - \sqrt{1+4Y}\right) \frac{X^3}{(1-X^2(1+4Y))^2} + \frac{4Y}{1+4Y} \frac{X^2}{(1-X^2(1+4Y))^2}$$
$$- \frac{1}{2} \frac{1}{\sqrt{1+4Y}} \left(1 + \frac{1}{1+4Y}\right) \frac{X}{1-X^2(1+4Y)} + \frac{1}{(1+4Y)^2} \frac{1}{1-X^2(1+4Y)}$$

and[32]

$$\sum_{n=0,j=0}^{\infty} X^n Y^j (j+1) E_{n,j-1}$$
$$= \left(-\sqrt{1+4Y} - \frac{1}{\sqrt{1+4Y}} + 2\right) \frac{X^3}{(1-X^2(1+4Y))^2} + \left(-1 - \frac{1}{1+4Y} + \frac{2}{\sqrt{1+4Y}}\right) \frac{X^2}{(1-X^2(1+4Y))^2}$$
$$- \frac{2Y}{\sqrt{1+4Y}^3} \frac{X}{1-X^2(1+4Y)} + \left(\frac{1}{(1+4Y)^2} - \frac{1}{\sqrt{1+4Y}^3}\right) \frac{1}{1-X^2(1+4Y)}\,.$$

Adding up all the contributions[33], we find that the generating functions of both of our formulas for $K_{n,j}$ agree. □

3.3.4. Proofs of Theorem 3.1.10 and Proposition 3.1.14

Proof of Theorem 3.1.10. We can write (NLS_n) as

$$i\partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\mathsf{NLS}}}{\delta \overline{q}} = -(-i)^n \delta\sigma_{n+1}\,,$$

so (3.3.1) and Lemma 3.3.4 directly imply (NLS_{2m}) and (NLS_{2m+1}).

Combining (NLS_{2m})–(NLS_{2m+1}) with (3.1.4)–(3.1.5), we can write the equations (GP_n) of the GP hierarchy as

$$i\partial_{t_{2m}} q = \sum_{k=0}^{m} \binom{m-\frac{1}{2}}{m-k} (-4)^{m-k} \left(\sum_{j=0}^{k-2} J_{2k+1,j} q^2 (i\partial_x)^{2k-2-2j} \overline{q} - \sum_{j=0}^{k-1} K_{2k+1,j} (i\partial_x)^{2k-2j} q \right) \quad (3.3.24)$$

---
[30] See `test30`
[31] See `test31`
[32] See `test32`
[33] See `test33`

$$
\begin{aligned}
&+\sum_{k=0}^{m}\binom{m-\frac{1}{2}}{m-k}(-4)^{m-k}\Bigg(\sum_{j=0}^{k-2}J_{2k+1,j}(|q|^{2j}-1)q^2(i\partial_x)^{2k-2-2j}\overline{q}\\
&-\sum_{j=0}^{k-1}K_{2k+1,j}(|q|^{2j}-1)(i\partial_x)^{2k-2j}q+(k+1)C_k|q|^{2k}q\Bigg)+\mathcal{O}_q^{2,n-2}(q_x)\\
i\partial_{t_{2m+1}}q=&\sum_{k=0}^{m}\binom{m}{m-k}(-4)^{m-k}\Bigg(\sum_{j=0}^{k}K_{2k+2,j}(i\partial_x)^{2k+1-2j}q\Bigg) \qquad (3.3.25)\\
&+\sum_{k=0}^{m}\binom{m}{m-k}(-4)^{m-k}\Bigg(\sum_{j=0}^{k}K_{2k+2,j}(|q|^{2j}-1)(i\partial_x)^{2k+1-2j}q\Bigg)+\mathcal{O}_q^{2,n-2}(q_x)\,.
\end{aligned}
$$

Note that all terms appearing above which contain both a derivative of q or $\overline{q}$ and the factor $|q|^{2j}-1$ are in $\mathcal{O}^{2,n-2}_{q,|q|^2-1}(q_x,|q|^2-1)$. We calculate[34]

$$
\begin{aligned}
&\sum_{m=0}^{\infty}\sum_{k=0}^{m}X^m\binom{m-\frac{1}{2}}{m-k}(-4)^{m-k}(k+1)C_k|q|^{2k}q=\big(1-4X(|q|^2-1)\big)^{-\frac{1}{2}}q\\
&=\sum_{m=0}^{\infty}X^m\binom{2m}{m}(|q|^2-1)^m q=\sum_{m=0}^{\infty}X^m\Big(\mathbb{1}_{\{m=0\}}q+2\mathbb{1}_{\{m=1\}}(q^2\overline{q}-q)+\mathcal{O}^{2,0}_{|q|^2-1}(|q|^2-1)\Big)\,.
\end{aligned}
$$

It remains to simplify the sums in the first lines of (3.3.24) and (3.3.25). Swapping the order of summation, the goal is to evaluate

$$
\sum_{j=1}^{m-1}q^2(i\partial_x)^{2j}\overline{q}\sum_{k=j+1}^{m}\binom{m-\frac{1}{2}}{m-k}(-4)^{m-k}J_{2k+1,k-1-j}-\sum_{j=1}^{m}(i\partial_x)^{2j}q\sum_{k=j}^{m}\binom{m-\frac{1}{2}}{m-k}(-4)^{m-k}K_{2k+1,k-j}
$$

and $\displaystyle\sum_{j=0}^{m}(i\partial_x)^{2j+1}q\sum_{k=j}^{m}\binom{m}{m-k}(-4)^{m-k}K_{2k+2,k-j}\,.$

We calculate[35]

$$
\sum_{k=j+1}^{m}\binom{m-\frac{1}{2}}{m-k}(-4)^{m-k}J_{2k+1,k-1-j}=2\mathbb{1}_{\{j=m-1\}}
$$

and[36]

$$
-\sum_{k=j}^{m}\binom{m-\frac{1}{2}}{m-k}(-4)^{m-k}K_{2k+1,k-j}=4^{m-j}\binom{\frac{1}{2}}{m-j}-2\sum_{l=0}^{m-2-j}(-1)^l(C_{l+1}-2C_l)4^{m-2-j-l}\binom{-\frac{1}{2}}{m-2-j-l}\,.
$$

Studying the generating function[37]

$$
\sum_{k=-2}^{\infty}Y^k4^{k+2}\binom{\frac{1}{2}}{k+2}-2\sum_{k=-2}^{\infty}Y^k\sum_{l=0}^{k}(-1)^l(C_{l+1}-2C_l)4^{k-l}\binom{-\frac{1}{2}}{k-l}=\frac{2}{Y}+\frac{1}{Y^2}\,,
$$

we find that[38]

$$
-\sum_{k=j}^{m}\binom{m-\frac{1}{2}}{m-k}(-4)^{m-k}K_{2k+1,k-j}=\mathbb{1}_{\{j=m\}}+2\mathbb{1}_{\{j=m-1\}}\,.
$$

Lastly, we calculate[39]

$$
\sum_{k=j}^{m}\binom{m}{m-k}(-4)^{m-k}K_{2k+2,k-j}=4^{m-j}\binom{\frac{1}{2}}{m-j}\,.
$$

The results of our calculations now imply (GP_{2m}) and (GP_{2m+1}). Combining this with (3.1.6) yields the equations ($\widetilde{\mathsf{GP}}_{2m}$) and ($\widetilde{\mathsf{GP}}_{2m+1}$). □

[34] See `test34`
[35] See `test35`
[36] See `test36`
[37] See `test37`
[38] See `test38`
[39] See `test39`

Proof of Proposition 3.1.14. We plug the ansatz $q(t,x) = q_*(x) + p(t,x)$ into (GP_{2m})–(GP_{2m+1}) for $n \geq 3$. First, observe that

$$\mathcal{O}^{2,n-2}_{q,|q|^2-1}(q_x, |q|^2 - 1) \subseteq \mathcal{O}^{2,n-2}_{p,q_*,|q_*|^2-1}(p, (q_*)_x, |q_*|^2 - 1)\,.$$

From (GP_{2m+1}) we can already deduce (3.1.23), from ($\widetilde{\mathsf{GP}}_{2m+1}$) similarly (3.1.25).

We write $\mathbf{p} = (p, \overline{p})$ and define the shorthand notation

$$\mathbf{O}_n(\mathbf{p}) = \mathcal{O}^{1,0}_{q_*,|q_*|^2-1}(|q_*|^2 - 1) + \mathcal{O}^{1,n}_{q_*}((q_*)_x) + \mathcal{O}^{2,n-2}_{p,q_*,|q_*|^2-1}(p, (q_*)_x, |q_*|^2 - 1)\,,$$

describing those terms that can be viewed as "nonlinear" in (3.1.24). Recall (GP) for $n = 2$ and observe that for all $n = 2m \geq 2$ we can rewrite (GP_{2m}) as

$$(i\partial_x)^2 q + 2(|q|^2 - 1)q = (D_x^2 + 2)p + 2q_*^2\overline{p} + \mathbf{O}_2(\mathbf{p})$$

$$((i\partial_x)^{2m} + 2(i\partial_x)^{2m-2})q + 2q^2(i\partial_x)^{2m-2}\overline{q} + \mathbf{O}_n(\mathbf{p}) = \big(D_x^{2m} + 2D_x^{2m-2}\big)p + 2q_*^2 D_x^{2m-2}\overline{p} + \mathbf{O}_n(\mathbf{p}) \text{ for } m \geq 2.$$

To obtain (3.1.22), we calculate for all $n = 2m \geq 2$

$$i\partial_{t_{2m}} \begin{pmatrix} p \\ \overline{p} \end{pmatrix} = \begin{pmatrix} \big(D_x^{2m} + 2D_x^{2m-2}\big)p + 2q_*^2 D_x^{2m-2}\overline{p} \\ -\big(D_x^{2m} + 2D_x^{2m-2}\big)\overline{p} - 2\overline{q}_*^2 D_x^{2m-2} p \end{pmatrix} + \mathbf{O}_n(\mathbf{p})\,.$$

Here the term $2q_*^2 D_x^{2m-2}\overline{p}$ is problematic as it has a non-constant (if $q_+^2 \neq q_-^2$) and non-decaying coefficient, which means it cannot be treated either as a constant-coefficient linear term or as a nonlinear term in the well-posedness theory of Section 3.4. We eliminate this term via a change of variables $\mathbf{p} \mapsto \mathbf{u}$.

Normal form transformation. Define

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 - q_*^2 \\ 1 - \overline{q}_*^2 & 0 \end{pmatrix} \qquad E = \begin{pmatrix} 2 & 2q_*^2 \\ -2\overline{q}_*^2 & -2 \end{pmatrix} \qquad E_0 = \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix}.$$

Then we can write

$$i\partial_{t_{2m}} \mathbf{p} = (\sigma_3 D_x^{2m} + E D_x^{2m-2})\mathbf{p} + \mathbf{O}_n(\mathbf{p})\,.$$

Let $R > 0$ be large and $\eta \in C_c^\infty(\mathbb{R}; [0,1])$ with $\eta|_{[-R,R]} = 1$ and $\operatorname{supp}\eta \subseteq [-2R, 2R]$. We define the operators

$$\mathcal{T} = 1 + B\psi(D_x) \quad \text{where} \quad \psi(\xi) = \frac{1 - \eta(\xi)}{\xi^2 + 2}\,.$$

The new variable we shall use is

$$\mathbf{u} = \mathcal{T}^{-1}\mathbf{p}\,.$$

Then the equation for $\mathbf{u}$ reads

$$i\partial_{t_{2m}} \mathbf{u} = \mathcal{T}^{-1}(\sigma_3 D_x^{2m} + E D_x^{2m-2})\mathcal{T}\mathbf{u} + \mathcal{T}^{-1}\mathbf{O}_n(\mathcal{T}\mathbf{u})\,.$$

We calculate

$$\begin{aligned}
(\sigma_3 D_x^{2m} + E D_x^{2m-2})\mathcal{T} &= \sigma_3 D_x^{2m} + E D_x^{2m-2} + \sigma_3 D_x^{2m} B\psi(D_x) + E D_x^{2m-2} B\psi(D_x) \\
&= \sigma_3 D_x^{2m} + E_0 D_x^{2m-2} + B(\sigma_3 D_x^{2m} + E_0 D_x^{2m-2})\psi(D_x) \\
&\quad + (E - E_0 + (B(E - E_0) + [E, B] + [\sigma_3, B]D_x^2)\psi(D_x))D_x^{2m-2} \\
&\quad + \sigma_3[D_x^{2m}, B]\psi(D_x) + E[D_x^{2m-2}, B]\psi(D_x)\,.
\end{aligned}$$

Since $\psi(D_x)$ commutes with the constant matrices σ_3 and E_0, the first line simplifies to

$$\sigma_3 D_x^{2m} + E_0 D_x^{2m-2} + B(\sigma_3 D_x^{2m} + E_0 D_x^{2m-2})\psi(D_x) = \mathcal{T}(\sigma_3 D_x^{2m} + E_0 D_x^{2m-2})\,.$$

We calculate the second line to be

$$\begin{aligned}&(E - E_0 + (B(E - E_0) + [E, B] + [\sigma_3, B]D_x^2)\psi(D_x))D_x^{2m-2}\\&= (-[\sigma_3, B] + [\sigma_3, B](D_x^2 + 2)\psi(D_x) + 2(1 - |q_*|^4)\sigma_3\psi(D_x))D_x^{2m-2}\\&= 2(1 - |q_*|^4)\sigma_3 D_x^{2m-2}\psi(D_x) - [\sigma_3, B]D_x^{2m-2}\eta(D_x)\,.\end{aligned}$$

Similarly for the third line, we use the algebraic identity

$$\begin{aligned}&\sigma_3[D_x^{2m}, B]\psi(D_x) + E[D_x^{2m-2}, B]\psi(D_x)\\&= \big(\sigma_3[D_x^2, [D_x^{2m-2}, B]] + (E - 2\sigma_3)[D_x^{2m-2}, B] + 4i\sigma_3 B_x D_x^{2m-3} - \sigma_3 B_{xx} D_x^{2m-2}\big)\psi(D_x)\\&+ \big(\sigma_3[D_x^{2m-2}, B] - 2i\sigma_3 B_x D_x^{2m-3}\big)(1 - \eta(D_x))\,.\end{aligned}$$

Therefore

$$\begin{aligned}i\partial_{t_{2m}}\mathbf{u} &= (\sigma_3 D_x^{2m} + E_0 D_x^{2m-2})\mathbf{u}\\&\quad + \mathcal{T}^{-1}\big(\mathbf{O}_n(\mathcal{T}\mathbf{u}) + \mathbf{O}_n(\psi(D_x)\mathbf{u}) + \mathbf{O}_n((1 - \eta(D_x))\mathbf{u}) - [\sigma_3, B]D_x^{2m-2}\eta(D_x)\mathbf{u}\big)\,.\end{aligned}$$

□

Proof of Remark 3.1.13. In order to proceed, we must use the Lagrange inversion formula. We refer the reader to [143] for a detailed exposition and adopt the notations from this reference. If $u(t)$ is a formal power series in t, then $[t^n]u(t)$ denotes the operation of extracting the coefficient in front of t^n (e. g. $[t^2](1+t)^3 = 3$). With this notation, a version of the Lagrange inversion formula reads as follows. Let $F(t)$, $u(t)$, and $\varphi(t)$ be formal power series. Assume that $[t^0]\varphi(t) \neq 0$, which implies that there exists a formal power series solution $w(t)$ to the implicit equation

$$w(t) = t\varphi(w(t))\,.$$

Then

$$[t^n]F(w(t)) = [t^n]F(t)\varphi(t)^{n-1}(\varphi(t) - t\varphi'(t))\,.$$

In the subsequent calculations, the relation $\approx$ shall refer to equality up to terms in $\mathcal{O}_q^{2,n-2}(q_x)$. Set $\rho = q\overline{q} = |q|^2$. Square brackets after a differential operator denote operator application. We use the Lagrange inversion formula to calculate $\frac{\delta\mathcal{H}_{2m+1}^{\mathsf{NLS}}}{\delta\overline{q}}$, using the right-hand side of (NLS_{2m+1}), to be

$$\begin{aligned}\frac{\delta\mathcal{H}_{2m+1}^{\mathsf{NLS}}}{\delta\overline{q}} &\approx \sum_{j=0}^{m}[t^j]\rho^j(1+4t)^{\frac{2m+1}{2}}(-\partial_x^2)^{m-j}[i\partial_x q] \approx [t^m]\frac{(1+4\rho t)^{\frac{3}{2}}}{1+t\partial_x^2}(1+4\rho t)^{m-1}[i\partial_x q]\\&\approx [t^m]\frac{(1+4\rho\frac{t}{1-4\rho t})^{\frac{3}{2}}}{1+\frac{t}{1-4\rho t}\partial_x^2}[i\partial_x q] \approx [t^m]\frac{1}{(1-4\rho t)^{\frac{1}{2}}(1-4\rho t+t\partial_x^2)}[i\partial_x q]\,,\end{aligned}$$

and similarly $\delta\mathcal{H}_{2m}^{\mathsf{NLS}}$, which is the right-hand side of (NLS_{2m}), to be

$$\begin{aligned}\frac{\delta\mathcal{H}_{2m}^{\mathsf{NLS}}}{\delta\overline{q}} &\approx [t^m](1-4t)^{-\frac{1}{2}}\rho^m q + \sum_{j=0}^{m-2}[t^j]\rho^j 2q^2(1+4t)^{\frac{2m-1}{2}}(-\partial_x^2)^{m-2-j}[-\partial_x^2\overline{q}]\\&+ \sum_{j=0}^{m-1}[t^j]\rho^j(1+2t)(1+4t)^{\frac{2m-1}{2}}(-\partial_x^2)^{m-1-j}[-\partial_x^2 q]\\&\approx [t^m](1-4\rho t)^{-\frac{1}{2}}q + [t^m]\frac{2q^2t^2(1+4\rho t)^{\frac{1}{2}}}{1+t\partial_x^2}(1+4\rho t)^{m-1}[-\partial_x^2\overline{q}]\\&+ [t^m]\frac{t(1+2\rho t)(1+4\rho t)^{\frac{1}{2}}}{1+t\partial_x^2}(1+4\rho t)^{m-1}[-\partial_x^2 q]\\&\approx [t^m](1-4\rho t)^{-\frac{1}{2}}q + [t^m]\frac{2q^2\frac{t^2}{(1-4\rho t)^2}(1+4\rho\frac{t}{1-4\rho t})^{\frac{1}{2}}}{1+\frac{t}{1-4\rho t}\partial_x^2}[-\partial_x^2\overline{q}]\end{aligned}$$

$$
\begin{aligned}
&+[t^m]\frac{\frac{t}{1-4\rho t}(1+2\rho\frac{t}{1-4\rho t})(1+4\rho\frac{t}{1-4\rho t})^{\frac12}}{1+\frac{t}{1-4\rho t}\partial_x^2}[-\partial_x^2 q] \\
&\approx [t^m](1-4\rho t)^{-\frac12}q + [t^m]\frac{1}{(1-4\rho t)^{\frac32}(1-4\rho t+t\partial_x^2)}[-2t^2\partial_x^2(\rho q)+(6t^2\rho-t)\partial_x^2 q]\,.
\end{aligned}
$$

Here we have used

$$
2t^2q^2\partial_x^2\overline{q}+t(1-2\rho t)\partial_x^2 q \approx t\partial_x^2 q + 2t^2(q^2\partial_x^2\overline{q}-\rho\partial_x^2 q) \approx t\partial_x^2 q + 2t^2(\partial_x^2(\rho q)-3\rho\partial_x^2 q)\,.
$$

We have from (3.2.50) the asymptotic expansion

$$
-2i\lambda\frac{\delta\log a(\lambda)}{\delta\overline{q}} = \sum_{n=0}^{\infty}(2\lambda)^{-n}\frac{\delta\mathcal{H}_n^{\mathsf{NLS}}}{\delta\overline{q}} \sim \sum_{m=0}^{\infty}(4\lambda^2)^{-m}\left(\frac{\delta\mathcal{H}_{2m}^{\mathsf{NLS}}}{\delta\overline{q}}+(2\lambda)^{-1}\frac{\delta\mathcal{H}_{2m+1}^{\mathsf{NLS}}}{\delta\overline{q}}\right)
$$

for $\operatorname{Im}\lambda > c(q)$. Substituting the above formulas replaces each t with $(4\lambda^2)^{-1}$. After further simplification (up to $\approx$), we obtain the claimed formula (3.1.20). $\square$

3.4. Well-posedness results

3.4.1. Local well-posedness for a large class of dispersive nonlinear systems

As mentioned in the introduction, the well-posedness theory presented here builds on works by C. E. Kenig, G. Ponce, and L. Vega from the 1990s (see [106, 108, 109, 110, 111]), specifically [107].

3.4.1.1. Notations and the general dispersive system

Let $n, J, K \in \mathbb{N}$ with $n \geq 2$ and define $\omega = \frac{n-1}{2} \in \frac12\mathbb{N}$. For $\mathbf{u} = \mathbf{u}(t,x) : \mathbb{R}^2 \to \mathbb{C}^J$, we consider the system

$$
\partial_t\mathbf{u} = i\Phi(D_x)\mathbf{u} + \mathcal{N}(\mathbf{u})\,. \tag{3.4.1}
$$

Here the nonlinear part consists of terms with a total number of derivatives not exceeding $n-2$. More precisely

$$
\begin{aligned}
(\mathcal{N}(\mathbf{u}))_j &= \mathbf{g}_j + \sum_{k=1}^{K}\sum_{b\in\{1,\dots,J\}^k}\sum_{l\in\mathbb{N}^k,|l|\leq n-2}\mathcal{J}_{j,b}^{k,l}\mathbf{f}_{j,b}^{k,l}\mathcal{N}_b^{k,l}(\mathbf{u}) \\
\mathcal{N}_b^{k,l}(\mathbf{u}) &= \prod_{h=1}^{k}\partial_x^{l_h}\mathcal{K}_{j,b,h}^{k,l}\mathbf{u}_{b_h}
\end{aligned} \tag{3.4.2}
$$

for $1 \leq j \leq J$, with $\mathbf{g}_j, \mathbf{f}_{j,b}^{k,l} : \mathbb{R} \to \mathbb{C}$ being space-dependent coefficients, and $\mathcal{J}_{j,b}^{k,l}, \mathcal{K}_{j,b,h}^{k,l}$ being bounded linear operators on L_x^2 that behave similarly to the identity. The precise mapping properties we need are given in (3.4.23)–(3.4.25) below.

Note that the theory described in this section also applies for nonlinear parts which are finite sums of expressions of the form (3.4.2).

As for the linear dispersive part, we assume first of all local Lipschitz continuity of the continuous symbol $\Phi(\xi) : \mathbb{R} \to \mathbb{R}^{J\times J}$ and the following growth of $\Phi'(\xi)$ at infinity:

$$
|\Phi'(\xi)| \leq C|\xi|\langle\xi\rangle^{n-2}\,. \tag{3.4.3}
$$

Furthermore, we assume that Φ' is almost symmetric in the sense that

$$
\text{there exists } S = S(\xi) \in L^\infty(\mathbb{R};\mathbb{R}^{J\times J}) \text{ with } \langle Sv, v\rangle_{\mathbb{R}^J} \geq c > 0 \text{ and } S\Phi - \Phi^T S \in L^\infty(\mathbb{R};\mathbb{R}^{J\times J})\,. \tag{3.4.4}
$$

Since we want to apply linear estimates for scalar symbols, as opposed to the matrix dispersion symbol $\Phi(\xi)$, we assume as given measurable functions $\varphi_j, \rho^1_{j,p,q}, \rho^2_{j,p,q} : \mathbb{R} \to \mathbb{R}$ for $j,p,q \in \{1,\dots,J\}$ such that

$$\Big(e^{it\Phi(\xi)}\Big)_{p,q} = \sum_{j=1}^{J} \big(\rho^1_{j,p,q}(\xi) + \operatorname{sign}(\xi)\rho^2_{j,p,q}(\xi)\big)e^{it\varphi_j(\xi)} \tag{3.4.5}$$

for almost all $\xi \in \mathbb{R}$. One should think of $\varphi_j(\xi)$ as the eigenvalues of $\Phi(\xi)$ and understand this decomposition as a convenient alternative to implementing a change of variables that diagonalizes $\Phi(\xi)$ (as is done, for example, in [89]). Informally, we want to force $\rho^1_{j,p,q}$ and $\rho^2_{j,p,q}$ to have for large ξ the same parity as $\varphi_j'(\xi)$. This is made possible by the presence of $\operatorname{sign}(\xi)$. Note that $\frac{1}{i}\operatorname{sign}(D_x) = H$ is the Hilbert transform, which is a benign operator in our situation.

Of course, we require a number of assumptions to be fulfilled by $(\varphi_j, \rho^1_{j,p,q}, \rho^2_{j,p,q})_{1\le j,p,q\le J}$. Notably, we do not necessarily require the boundedness of $\rho^d_{j,p,q}$ near the origin. In order to remove such possible singularities (see the example in Lemma 3.4.1 below), we fix some $\mu \in \frac{1}{2}\mathbb{N}$ with $\mu \le 5\omega$ so that $\mu \in \mathbb{N}$ or $\omega \in \mathbb{N}$ or $\mu \ge \omega$. In addition, we require $\mu \notin \mathbb{N}$ if $\omega = \frac{1}{2}$. Define for $j,p,q \in \{1,\dots,J\}$ and $d \in \{1,2\}$ the auxiliary symbols

$$\widetilde{a}^d_{j,p,q}(\xi) = \frac{|\xi|^\omega}{|\varphi_j'(\xi)|^{\frac{1}{2}}} \cdot \Big(\frac{|\xi|^\mu}{\langle\xi\rangle^\mu}\rho^d_{j,p,q}(\xi)\Big) \tag{3.4.6}$$

$$a^d_{j,p,q}(\xi) = \frac{|\xi|^{2\omega}}{\varphi_j'(\xi)} \cdot \Big(\frac{|\xi|^\mu}{\langle\xi\rangle^\mu}\rho^d_{j,p,q}(\xi)\Big). \tag{3.4.7}$$

We now formulate $e^{it\Phi(\xi)}$ (recalling (3.4.5)) in terms of the symbols $\widetilde{a}$ and a as

$$\Big(e^{it\Phi(\xi)}\Big)_{p,q} = \sum_{j=1}^{J} \frac{1}{|\xi|^{\omega+\mu}}(\widetilde{a}^1_{j,p,q}(\xi) + \operatorname{sign}(\xi)\widetilde{a}^2_{j,p,q}(\xi))|\varphi_j'(\xi)|^{\frac{1}{2}}e^{it\varphi_j(\xi)}\langle\xi\rangle^\mu \tag{3.4.8}$$

$$= \sum_{j=1}^{J} \frac{1}{|\xi|^{2\omega+\mu}}(a^1_{j,p,q}(\xi) + \operatorname{sign}(\xi)a^2_{j,p,q}(\xi))\varphi_j'(\xi)e^{it\varphi_j(\xi)}\langle\xi\rangle^\mu, \tag{3.4.9}$$

and assume the following conditions on the linear matrix symbol $\Phi(\xi)$ to be satisfied. For every $(\varphi, \widetilde{a}, a) \in \{(\varphi_j, \widetilde{a}^d_{j,p,q}, a^d_{j,p,q}) : j,q,p \in \{1,\dots,J\}, d \in \{1,2\}\}$ we have:

(P1) The symbol $\varphi \in C^3(\mathbb{R};\mathbb{R})$ has at most a single critical point at $\xi_0 = 0$, and in that case there exist $\alpha \in (0,1)$, $r > 0$ and functions $h_- \in C^1([\xi_0 - r, \xi_0];\mathbb{R})$, $h_+ \in C^1([\xi_0, \xi_0 + r];\mathbb{R})$ such that $h_\pm(\xi_0) \ne 0$ and

$$\varphi(\xi) = \varphi(\xi_0) + |\xi - \xi_0|^{\frac{1}{\alpha}}h_\pm(\xi) \qquad \forall\, \xi \in \mathbb{R} \text{ s.t. } \pm(\xi - \xi_0) \in (0,r). \tag{3.4.10}$$

(P2) There exist $R, C, c > 0$ such that for all $k \in \{0,1,2,3\}$ and $\xi \in \mathbb{R}$ with $|\xi| > R$ we have

$$c|\xi|^{n-k} \le |\partial_\xi^k \varphi(\xi)| \le C|\xi|^{n-k}.$$

(P3) We have $a \in \operatorname{Lip}(\mathbb{R};\mathbb{R})$, and there exist $C, \delta > 0$ such that for all $\xi, \zeta \in \mathbb{R}$ with $|\varphi(\xi)| = |\varphi(\zeta)|$

$$|\widetilde{a}(\xi)| + |a(\xi)| + \langle\xi\rangle|a'(\xi)| + \langle\varphi(\xi)\rangle^\delta|a(\xi) - a(\zeta)| \le C.$$

(P4) There exists $M > 0$ such that for any $\tau \in \mathbb{R}$ we can decompose $\mathbb{R}$ into M intervals on whose interiors the functions $\frac{a(\xi)}{\varphi(\xi)-\tau}$ and φ' are monotone.

Note that the boundedness of a in (P3) forces us, recalling (3.4.7), to restrict to the case with at most one single critical point of φ in (P1). Note that the linear estimates from Appendix 3.A cover a more general setting, where several critical points are admissible. The role of $\frac{|\xi|^\mu}{\langle\xi\rangle^\mu}$ with appropriately chosen μ is to remove possible singularities of $\rho^d_{j,p,q}$ at the origin so that a is bounded.

Before moving on with the local well-posedness theory, we verify that the linear part of our system (3.1.21) satisfies the above assumptions.

Lemma 3.4.1 (Dispersion of our perturbative flows with $n \geq 2$)**.** *Recall our linear matrix symbol $\Phi(\xi)$ for which*

$$\text{Case } n \in 2\,\mathbb{N}: \text{ the system (3.1.26) has symbol } \Phi(\xi) = -\xi^{n-2}\begin{pmatrix} \xi^2+2 & 2 \\ -2 & -\xi^2-2 \end{pmatrix} \tag{3.4.11}$$

$$\text{Case } n \in 2\,\mathbb{N}+1: \text{ the system (3.1.25) has the symbol } \Phi(\xi) = \xi^n \in \mathbb{R}^{2\times 2}\,. \tag{3.4.12}$$

For each case both (3.4.3) and the almost symmetry condition (3.4.4) are satisfied, and there exist φ_j, $\widetilde{a}^1_{j,p,q}, \widetilde{a}^2_{j,p,q}$, and $a^1_{j,p,q}, a^2_{j,p,q}$ such that (3.4.8) and (3.4.9) hold, and (P1)–(P4) are satisfied.

Proof. We first consider the easy case $n \in 2\,\mathbb{N}+1$ where the matrix symbol is $\Phi(\xi) = \xi^{2\omega+1} \in \mathbb{R}^{2\times 2}$, $\omega = \frac{n-1}{2} \in \mathbb{N}$. This trivially satisfies (3.4.3), the almost symmetry condition (3.4.4), and (P1)–(P4) with $\mu = 0$, $\varphi_j(\xi) = \xi^{2\omega+1}$, $\rho^1_{j,p,q} = \mathbb{1}_{\{j=p=q\}}$, $\rho^2_{j,p,q} = 0$, and

$$\widetilde{a}^1_{j,p,q}(\xi) = \frac{1}{\sqrt{2\omega+1}}\mathbb{1}_{\{p=q\}}, \quad a^1_{j,p,q}(\xi) = \frac{1}{2\omega+1}\mathbb{1}_{\{p=q\}}, \quad \widetilde{a}^2_{j,p,q}(\xi) = a^2_{j,p,q}(\xi) = 0.$$

When verifying (P3), note that $|\varphi_j(\xi)| = |\varphi_j(\zeta)|$ implies $|\xi| = |\zeta|$ and hence $a^d_{j,p,q}(\xi) - a^d_{j,p,q}(\zeta) = 0$.

In the case $n \in 2\,\mathbb{N}$ the symbol $\Phi(\xi) = -\xi^{2\omega-1}\begin{pmatrix} \xi^2+2 & 2 \\ -2 & -\xi^2-2 \end{pmatrix} : \mathbb{R} \to \mathbb{R}^{2\times 2}$ with $\omega = \frac{n-1}{2} \in \frac{1}{2}\,\mathbb{N}$ satisfies (3.4.3) and the almost symmetry condition with $S(\xi) = \begin{pmatrix} 1 & \frac{2}{2+\xi^2} \\ \frac{2}{2+\xi^2} & 1 \end{pmatrix}$ for $|\xi| \geq 1$, and some uniformly positive definite smooth extension to $|\xi| < 1$. We define for $\xi \neq 0$ the modified Japanese bracket $\langle \xi \rangle_\sim = \sqrt{\xi^2+4}$ and the matrices

$$\mathcal{D} = \xi^{2\omega}\langle \xi \rangle_\sim \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.4.13}$$

$$\mathcal{V}(\xi) = \frac{1}{2}\begin{pmatrix} \frac{\xi}{\langle \xi \rangle_\sim}+1 & \frac{\xi}{\langle \xi \rangle_\sim}-1 \\ \frac{\xi}{\langle \xi \rangle_\sim}-1 & \frac{\xi}{\langle \xi \rangle_\sim}+1 \end{pmatrix} \quad \mathcal{W}(\xi) = \frac{1}{2}\begin{pmatrix} \frac{\langle \xi \rangle_\sim}{\xi}+1 & \frac{\langle \xi \rangle_\sim}{\xi}-1 \\ \frac{\langle \xi \rangle_\sim}{\xi}-1 & \frac{\langle \xi \rangle_\sim}{\xi}+1 \end{pmatrix} \tag{3.4.14}$$

so that we can diagonalize $\Phi = \mathcal{V}\mathcal{D}\mathcal{W}$ with $\mathcal{W} = \mathcal{V}^{-1}$. We set $\varphi_j(\xi) = \xi^{2\omega}\langle \xi \rangle_\sim(-1)^j$, and for $\xi \neq 0$ write

$$\left(e^{it\Phi(\xi)}\right)_{p,q} = \sum_{j=1}^{2} \mathcal{V}_{p,j}(\xi)\mathcal{W}_{j,q}(\xi)e^{it\varphi_j(\xi)}.$$

Then (3.4.5) holds for the following choice of two even functions

$$\rho^1_{j,p,q}(\xi) = \frac{(\mathcal{V}_{p,j}\mathcal{W}_{j,q})(\xi) + (\mathcal{V}_{p,j}\mathcal{W}_{j,q})(-\xi)}{2} \quad \rho^2_{j,p,q}(\xi) = \operatorname{sign}(\xi)\frac{(\mathcal{V}_{p,j}\mathcal{W}_{j,q})(\xi) - (\mathcal{V}_{p,j}\mathcal{W}_{j,q})(-\xi)}{2}\,.$$

We now define $\widetilde{a}^d_{j,p,q}$ and $a^d_{j,p,q}$ as in (3.4.6)–(3.4.7) with $\mu = 1$ if $\omega \geq 1$, and $\mu = \frac{3}{2}$ if $\omega = \frac{1}{2}$. Note that (P1), (P2), and (P4) hold trivially. Regarding (P3), observe that, up to a change of sign, the product $(\mathcal{V}_{p,j}\mathcal{W}_{j,q})(\pm\xi)$ is of the form $\left(\frac{\xi}{\langle \xi \rangle_\sim} \pm 1\right)\left(\frac{\langle \xi \rangle_\sim}{\xi} \pm 1\right)$, and furthermore $(\mathcal{V}_{p,j}\mathcal{W}_{j,q})(\xi) + (\mathcal{V}_{p,j}\mathcal{W}_{j,q})(-\xi)$ is constant in ξ, hence $\rho^1_{j,p,q} \in \mathbb{R}$ and $\operatorname{sign}(\xi)\rho^2_{j,p,q} \in \operatorname{span}_{\mathbb{R}}\left\{\frac{\xi}{\langle \xi \rangle_\sim}, \frac{\langle \xi \rangle_\sim}{\xi}\right\}$. We compute

$$(-1)^j\varphi_j'(\xi) = \frac{\xi^{2\omega-1}(8\omega + (2\omega+1)\xi^2)}{\langle \xi \rangle_\sim} \geq 0,$$

such that (3.4.6)–(3.4.7) with $\mu = 1$ read as

$$\widetilde{a}^1_{j,p,q}(\xi) = \frac{\sqrt{|\xi|\langle \xi \rangle_\sim}}{\sqrt{8\omega + (2\omega+1)\xi^2}} \cdot \frac{|\xi|}{\langle \xi \rangle}\Big(\text{some real constant}\Big)$$

$$\widetilde{a}^2_{j,p,q}(\xi) = \frac{\sqrt{|\xi|\langle \xi \rangle_\sim}}{\sqrt{8\omega + (2\omega+1)\xi^2}} \cdot \frac{\xi}{\langle \xi \rangle}\left(\text{linear combinations of } \frac{\xi}{\langle \xi \rangle_\sim} \text{ and } \frac{\langle \xi \rangle_\sim}{\xi}\right),$$

and similarly for $a^d_{j,p,q}$ with the first factor $\frac{\sqrt{|\xi|\langle\xi\rangle_\sim}}{\sqrt{8\omega+(2\omega+1)\xi^2}}$ replaced by $(-1)^j\frac{|\xi|\langle\xi\rangle_\sim}{8\omega+(2\omega+1)\xi^2}$. Thus $\widetilde{a}^d_{j,p,q}, a^d_{j,p,q}$ are all bounded and $(a^d_{j,p,q})'$ decays as $\langle\xi\rangle^{-1}$ at infinity. The case $\omega=\frac{1}{2}$ and $\mu=\frac{3}{2}$ is analogous. To show the "cancellation" property of $|a(\xi)-a(\zeta)|$ for $|\varphi(\xi)|=|\varphi(\zeta)|$, we notice the monotonicity of $\varphi_j(\xi)$ and the evenness of $a^d_{j,p,q}$ (recalling (3.4.6)–(3.4.7)). This completes the verification of the assumption (P3). □

Remark 3.4.2. *The introduction of the regularization factor $\frac{|\xi|^\mu}{\langle\xi\rangle^\mu}$ with $\mu=1$ (resp.$\mu=\frac{3}{2}$) implies the boundedness of the symbol $\widetilde{a}^2_{j,p,q}$ around the origin above. For the symbol $a^2_{j,p,q}$ it would suffice to set $\mu=0$, but we use non-zero values of μ in both cases for the sake of simplicity.*

Remark 3.4.3. *Thanks to Proposition 3.1.1 and Remark 3.1.3, the well-posedness problems for the* GP *and* $\widetilde{\mathsf{GP}}$ *hierarchies are equivalent. In particular it suffices to consider the well-posedness of the system (3.1.26) with the linear part (3.1.25) in the odd case $n\in 2\,\mathbb{N}+1$ and the linear part (3.1.22) in the even case $n\in 2\,\mathbb{N}$.*

3.4.1.2. Preliminary estimates

Recall that $n\geq 2$, $\omega=\frac{n-1}{2}\geq\frac{1}{2}$ and $\mu\in\frac{1}{2}\,\mathbb{N}$ with $\mu\leq 5\omega$ so that $\mu\in\mathbb{N}$ or $\omega\in\mathbb{N}$ or $\mu\geq\omega$, and in addition $\mu\notin\mathbb{N}$ if $\omega=\frac{1}{2}$. We derive the following preliminary estimates from Theorem 3.A.1 and Theorem 3.A.7 in this subsection.

Lemma 3.4.4. *If (P1)–(P4) are fulfilled, then the following linear estimates are available.*

(i) ***Local smoothing estimate*** *(compare to [108, Theorem 4.1], [105, Theorem 2.1, Corollary 2.2], [110, Theorem 3.5]). It holds that*
- *Smoothing the initial data:*

$$\left\||D_x|^{\omega+\mu}e^{it\Phi(D_x)}\mathbf{u}_0\right\|_{L^\infty_x L^2_{t\in[-T,T]}}\lesssim\|\langle D_x\rangle^\mu\mathbf{u}_0\|_{L^2_x} \tag{3.4.15}$$

- *For $h\in\mathbb{N}$ with $0\leq h\leq\omega+\mu$,*

$$\left\|\int_0^t D^h_x e^{i(t-t')\Phi(D_x)}\mathbf{u}(t',x)\,\mathrm{d}t'\right\|_{L^\infty_{t\in[-T,T]}L^2_x}\lesssim T^{\alpha_h}\|\langle D_x\rangle^{\theta_h\mu}\mathbf{u}\|_{L^{p_h}_x L^2_{t\in[-T,T]}} \tag{3.4.16}$$

where $\theta_h=\frac{h}{\omega+\mu}\in[0,1]$, $\alpha_h=\frac{1-\theta_h}{2}\in[0,\frac{1}{2}]$, and $\frac{1}{p_h}=\frac{\theta_h}{1}+\frac{1-\theta_h}{2}=\frac{1+\theta_h}{2}\in[\frac{1}{2},1]$.
- *For $h\in\mathbb{N}$ with $\omega+\mu\leq h\leq 2\omega+\mu$,*

$$\left\|\int_0^t D^h_x e^{i(t-t')\Phi(D_x)}\mathbf{u}(t',x)\,\mathrm{d}t'\right\|_{L^\infty_x L^2_{t\in[-T,T]}}\lesssim T^{\alpha_h}\|\langle D_x\rangle^{\mu}\mathbf{u}\|_{L^{p_h}_x L^2_{t\in[-T,T]}} \tag{3.4.17}$$

where $\theta_h=\frac{h-\omega-\mu}{\omega}\in[0,1]$, $\alpha_h=\frac{1-\theta_h}{2}\in[0,\frac{1}{2}]$, and $\frac{1}{p_h}=\frac{1+\theta_h}{2}\in[\frac{1}{2},1]$.

(ii) ***Maximal function estimate*** *(compare to [108, Theorem 2.5], [105, Theorem 2.3], [111, Corollary 2.9]). Set $r=\frac{1}{2}\vee\frac{\omega}{2}=\frac{1}{2}\mathbb{1}_{\{\omega=\frac{1}{2}\}}+\frac{\omega}{2}\mathbb{1}_{\{\omega\geq 1\}}$. We have*

$$\left\|e^{it\Phi(D_x)}\mathbf{u}_0\right\|_{L^2_x L^\infty_{t\in[-T,T]}}\lesssim_T\|\mathbf{u}_0\|_{H^r_x}\,. \tag{3.4.18}$$

Proof.

(i) We want to apply Theorem 3.A.1. This requires $\widetilde{a}\in L^\infty$, which we assumed in (P3), and a and φ to fulfill (H1)–(H5). Observe that:
- $(P1)\implies(H1)$ by definition.
- $(P1)\implies(H2)$. Without loss of generality, it suffices to show this for ψ_j in the case $\xi_j=0$, $\eta_j=0$, $\sigma_j=1$. Let α, r and $h_\pm$ be given by the definition of positive steepness at $\xi_j=0$. It suffices to consider $\eta\in(0,\varphi(r))$. We write (3.4.10) as

$$\eta=h_+(\psi_j(\eta))|\psi_j(\eta)|^{\frac{1}{\alpha}}$$

and note that its derivative reads

$$1 = \Big(h_+'(\psi_j(\eta))|\psi_j(\eta)|^{\frac{1}{\alpha}} + h_+(\psi_j(\eta))\frac{1}{\alpha}|\psi_j(\eta)|^{\frac{1}{\alpha}-1}\operatorname{sign}(\psi_j(\eta))\Big)\psi_j'(\eta)\,.$$

This implies

$$\frac{|\eta|^\alpha}{|\psi_j(\eta)|} = |h_+(\psi_j(\eta))|^\alpha$$
$$\frac{|\eta|^{\alpha-1}}{|\psi_j'(\eta)|} = \Big(\frac{|\eta|^\alpha}{|\psi_j(\eta)|}\Big)^{\frac{\alpha-1}{\alpha}}\Big|h_+'(\psi_j(\eta))\psi_j(\eta) + h_+(\psi_j(\eta))\frac{1}{\alpha}\Big|\,.$$

Since $h_+\circ\psi_j$ is non-zero and continuous to the right of $\eta=0$ and $\psi_j(0)=0$, the right-hand sides above are bounded from above and below. Thus (3.A.1) and then also (3.A.2) follow.

- $(P2)\implies(H3)$. We apply the estimates from (P2) to $\xi=\psi_j(\eta)$, where $j\in\{0,N\}$, which yields (H3) with $\beta=\frac{1}{n}$.
- $(P3)$ and $(P4)\implies(H4)$ and $(H5)$.

We obtain (3.4.15) by decomposing the exponential with (3.4.8) as

$$\Big(e^{it\Phi(\xi)}\Big)_{p,q}\mathbf{u}_0 = \frac{1}{|\xi|^{\omega+\mu}}\sum_{j=1}^J\Big((\widetilde{a}^1_{j,p,q}(\xi)+\operatorname{sign}(\xi)\widetilde{a}^2_{j,p,q}(\xi))|\varphi_j'(\xi)|^{\frac12}\Big)e^{it\varphi_j(\xi)}\big(\langle\xi\rangle^\mu\mathbf{u}_0\big)\,,$$

and applying (3.A.6) with $\varphi=\varphi_j$ to $\langle D_x\rangle^\mu\mathbf{u}_0$ and $H\langle D_x\rangle^\mu\mathbf{u}_0$. Here $H=\frac{1}{i}\operatorname{sign}(D_x)$ is the Hilbert transform, which is bounded on L^2.

By [40] the vector-valued Hilbert transform is also bounded on $L^{p_h}_xL^2_{[-T,T]}$. As a result, it suffices to prove (3.4.16) and (3.4.17) with D_x^h replaced by $|D_x|^h=D_x^h(iH)^h$. To derive (3.4.16) we use interpolation[40]. The estimate

$$\Big\|\int_0^t e^{i(t-t')\Phi(D_x)}\mathbf{u}(t',x)\,\mathrm{d}t'\Big\|_{L^\infty_{t\in[-T,T]}L^2_x}\lesssim T^{\frac12}\|\mathbf{u}\|_{L^2_xL^2_{t\in[-T,T]}}\,,$$

which corresponds to the case $h=0$, $\theta_h=0$, follows from Hölder's inequality in t'. The estimate for the case $h=\omega+\mu$, $\theta_h=1$ is the dual of (3.4.15):

$$\Big\|\int_0^t |D_x|^{\omega+\mu}e^{i(t-t')\Phi(D_x)}\mathbf{u}(t',x)\,\mathrm{d}t'\Big\|_{L^\infty_{t\in[-T,T]}L^2_x}\lesssim\|\langle D_x\rangle^\mu\mathbf{u}\|_{L^1_xL^2_{t\in[-T,T]}}\,.$$

Thus (3.4.16) follows from interpolation.

We prove (3.4.17) by interpolation as well. Here the case $h=\omega+\mu$ follows from Minkowski's inequality and (3.4.15):

$$\Big\|\int_0^t |D_x|^{\omega+\mu}e^{i(t-t')\Phi(D_x)}\mathbf{u}(t',x)\,\mathrm{d}t'\Big\|_{L^\infty_xL^2_{t\in[-T,T]}}\lesssim T^{\frac12}\|\langle D_x\rangle^\mu\mathbf{u}\|_{L^2_xL^2_{t\in[-T,T]}}.$$

Recall the decomposition (3.4.9):

$$\Big(e^{i(t-t')\Phi(\xi)}\Big)_{p,q}\mathbf{u}(t',x) = \frac{1}{|\xi|^{2\omega+\mu}}\sum_{j=1}^J\Big((a^1_{j,p,q}(\xi)+\operatorname{sign}(\xi)a^2_{j,p,q}(\xi))\varphi_j'(\xi)\Big)e^{i(t-t')\varphi_j(\xi)}\big(\langle\xi\rangle^\mu\mathbf{u}(t',x)\big)\,.$$

We apply (3.A.8) to obtain the estimate for the case $h=2\omega+\mu$:

$$\Big\|\int_0^t |D_x|^{2\omega+\mu}e^{i(t-t')\Phi(D_x)}\mathbf{u}(t',x)\,\mathrm{d}t'\Big\|_{L^\infty_xL^2_{t\in[-T,T]}}\lesssim\|\langle D_x\rangle^\mu\mathbf{u}\|_{L^1_xL^2_{t\in[-T,T]}}.$$

[40] The necessary interpolation inequality follows from the argument in [166], using a version of the "three-lines lemma".

(ii) Let $\chi \in C_c^\infty(\mathbb{R};\mathbb{R})$. We estimate the low frequency part $\chi(D_x)\mathbf{u}_0$ by applying Theorem 3.A.7 with $a = 0$ and $b(t, D_x) = e^{it\Phi(D_x)}\chi(D_x)$. For the high frequency part $(1-\chi(D_x))\mathbf{u}_0$ we use (3.4.5) to reduce to the scalar case with $\varphi = \varphi_j$, and apply Theorem 3.A.7 with $a(D_x) = \rho^1_{j,p,q}(D_x) + iH\rho^2_{j,p,q}(D_x)$ and $b = 0$. Note that (P1)–(P3) imply that for large frequencies $\rho^1_{j,p,q}$ and $\rho^2_{j,p,q}$ are essentially bounded. The requirements (J1)–(J6) for φ are all direct consequences of (P2).

□

3.4.1.3. Local well-posedness for the general dispersive system

For the rest of the section, we always consider $h \in \mathbb{N}$ so that ∂_x^h is not a fractional derivative. We take the initial data $\mathbf{u}_0$ in the (weighted) Sobolev spaces $H^s \cap H^{s',1}$ with $s, s' \in \mathbb{N}$ defined in (3.1.12). Let $T > 0$ be a positive time and consider the (pseudo-)norms

$$\|\mathbf{u}\|_{Y_{1,T}} = \max_{0\le h\le s} \|\partial_x^h \mathbf{u}\|_{L^\infty_{t\in[-T,T]}L^2_x} \qquad \|\mathbf{u}\|_{Y_{2,T}} = \max_{s\le h\le s+\lfloor\omega\rfloor} \|\partial_x^h \mathbf{u}\|_{L^\infty_x L^2_{t\in[-T,T]}}$$

$$\|\mathbf{u}\|_{Y_{3,T}} = \max_{0\le h\le s-\lceil\omega\rceil-\lceil r\rceil} \|\partial_x^h \mathbf{u}\|_{L^2_x L^\infty_{t\in[-T,T]}} \qquad \|\mathbf{u}\|_{Y_{4,T}} = \max_{0\le h\le s'-\lceil\omega+r\rceil} \|x\partial_x^h \mathbf{u}\|_{L^2_x L^\infty_{t\in[-T,T]}}$$

$$\|\mathbf{u}\|_{Y_{5,T}} = \max_{0\le h\le s'-\lceil\omega\rceil} \|x\partial_x^h \mathbf{u}\|_{L^\infty_{t\in[-T,T]}L^2_x} \qquad \|\mathbf{u}\|_{Y_T} = \sum_{j=1}^5 \|\mathbf{u}\|_{Y_{j,T}}\,.$$

Here $\omega = \frac{n-1}{2} \ge \frac{1}{2}$ and $r = \frac{1}{2} \vee \frac{\omega}{2}$ are the same as in the previous section 3.4.1.2. Our fixed-point argument later is applied in the Banach space

$$(Y_T, \|\cdot\|_{Y_T}) = \left(\left\{\mathbf{u} \in C_{t\in[-T,T]}(H^s \cap H^{s',1})_x : \|\mathbf{u}\|_{Y_T} < \infty\right\}, \|\cdot\|_{Y_T}\right)$$

for a sufficiently small time $T > 0$.

Definition 3.4.5 (Mild solution)**.** For a given $\mathbf{u}_0 \in H^s \cap H^{s',1}$ and a time $T > 0$ we say that $\mathbf{u} \in Y_T$ **is a mild solution with initial data** $\mathbf{u}_0$ **to** (3.4.1) if

$$\Lambda_{\mathbf{u}_0}(\mathbf{u}) = \mathbf{u} \qquad \text{where} \qquad \Lambda_{\mathbf{u}_0}(\mathbf{u})(t) = e^{it\Phi(D_x)}\mathbf{u}_0 + \int_0^t e^{i(t-t')\Phi(D_x)}\mathcal{N}(\mathbf{u}(t'))\,\mathrm{d}t'\,. \tag{3.4.19}$$

Theorem 3.4.6 (Local well-posedness)**.** *Let $s, s' \in \mathbb{N}$ satisfy $s \ge 7\omega$ and $\frac{s+3\omega}{2} \le s' \le s - 2\omega$. Consider the general dispersive equation (3.4.1)–(3.4.2), where the linear part satisfies (3.4.3), (3.4.8)–(3.4.9) and (P1)-(P4) (see Subsection 3.4.1.1), and the coefficients in the nonlinear part satisfy*

$$\mathbf{g}_j \in H^{s+1-\lceil\omega\rceil,1} \tag{3.4.20}$$

$$\mathbf{f}^{k,l}_{j,b} \in W^{s+1-\lceil\omega\rceil,\infty} \tag{3.4.21}$$

$$\mathbf{f}^{1,l}_{j,b} \in H^{s+1-\lceil\omega\rceil,1} \text{ or } \mathcal{K}^{1,l}_{j,b,1} \in \bigcap_{T>0} L(H_x^{s'-\lceil\omega+r\rceil,1}L^\infty_{t\in[-T,T]}; H_x^{|l|,1}L^\infty_{t\in[-T,T]})\,. \tag{3.4.22}$$

Assume in addition that

$$\mathcal{J}^{k,l}_{j,b}, \mathcal{K}^{k,l}_{j,b,h} \in \bigcap_{0\le h\le s} L(H^h) \cap \bigcap_{0\le h\le s'} L(H^{h,1}) \tag{3.4.23}$$

$$\mathcal{J}^{k,l}_{j,b} \in \bigcap_{T>0}\ \bigcap_{0\le h\le s+\lfloor\omega\rfloor}\ \bigcap_{p\in(1,2]} L(W_x^{h,p}L^2_{t\in[-T,T]})\,. \tag{3.4.24}$$

$$\mathcal{K}^{k,l}_{j,b,h} \in \bigcap_{T>0} L(Y_T)\,. \tag{3.4.25}$$

All the constants below depend on s, s' and the norms of the above quantities, which we suppress from the notation.

For any $\mathbf{u}_0 \in H^s \cap H^{s',1}$ there exists a time $T = T(\|\mathbf{u}_0\|_{H^s\cap H^{s',1}}) > 0$, such that

- ***Existence.*** *The system* (3.4.1) *has a mild solution* $\mathbf{u}(t,x) \in Y_T$ *with*

$$\|\mathbf{u}\|_{Y_T} \leq C(\|\mathbf{u}_0\|_{H^s \cap H^{s',1}})\|\mathbf{u}_0\|_{H^s \cap H^{s',1}}. \tag{3.4.26}$$

- ***Uniqueness.*** *This mild solution is unique in the sense that it agrees with any other mild solution in* Y_T.
- ***Lipschitz continuity.*** *The data-to-solution map is locally Lipschitz continuous in the sense that for two initial data* $\mathbf{u}_0, \widetilde{\mathbf{u}}_0 \in H^s \cap H^{s',1}$, *with times of existence* T *and* $\widetilde{T}$ *respectively, we have*

$$\|\mathbf{u} - \widetilde{\mathbf{u}}\|_{Y_{\min\{T,\widetilde{T}\}}} \leq C(\|\mathbf{u}_0\|_{H^s \cap H^{s',1}}, \|\widetilde{\mathbf{u}}_0\|_{H^s \cap H^{s',1}})\|\mathbf{u}_0 - \widetilde{\mathbf{u}}_0\|_{H^s \cap H^{s',1}}.$$

- ***Smooth flow.*** *For any* $k \in \mathbb{N}$, *there exist a neighborhood* $U_0 \subseteq H^s \cap H^{s',1}$ *of* $\mathbf{u}_0$ *and a small time* $T_0 > 0$, *for which the data-to-solution flow map*

$$\begin{aligned} U_0 &\longrightarrow Y_{T_0} \\ \mathbf{u}_0 &\longmapsto \mathbf{u} \end{aligned}$$

is a C^k *mapping between Banach spaces*[41].
- ***Blow-up criterion.*** *This mild solution either exists globally in time or blows up at a finite maximal existence time* $T^* > 0$ *in the sense that*

$$\sup_{t \in (-T^*, T^*)} \|\mathbf{u}(t)\|_{H^s \cap H^{s',1}} = \infty. \tag{3.4.27}$$

Recall the linear estimates related to the matrix symbol $\Phi(\xi)$ in Lemma 3.4.4 and the uniform form in the nonlinearities $(\mathcal{N}(\mathbf{u}))_j$ given in (3.4.2) for all $j = 1, \cdots, J$.

In order to simplify the proof below and the rest of the section, we assume $\mathcal{J}^{k,l}_{j,b} = \mathcal{K}^{k,l}_{j,b,h} = 1$ to be the identity operators. The exceptional case is that (3.4.22) may require $\mathcal{K}^{k,l}_{j,b,h}$ to have smoothing properties, which we remark on when necessary. We nevertheless do not show this operator in our notation.

We also assume $J = 1$, and hence drop the (multi-)indices j and b from the notation, which makes no difference for the majority of the arguments below. Specifically, $\mathbf{u}_j$, $\mathcal{N}^{k,l}_b(\mathbf{u})$, $\mathbf{g}_j$ and $\mathbf{f}^{k,l}_{j,b}$ defined in Subsection 3.4.1.1 are replaced by $\mathbf{u}$, $\mathcal{N}^{k,l}(\mathbf{u})$, $\mathbf{g}$ and $\mathbf{f}^{k,l}$. At certain points where the fact that Φ is a matrix matters, we switch to the general setting $J \geq 1$.

Let us summarize these simplifications by writing

$$\mathcal{N}(\mathbf{u}) = \mathbf{g} + \sum_{k=1}^{K} \sum_{l \in \mathbb{N}^k, |l| \leq n-2} \mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}) \qquad \mathcal{N}^{k,l}(\mathbf{u}) = \prod_{d=1}^{k} \partial_x^{l_d}\mathbf{u} \quad l = (l_1, \cdots, l_k) \in \mathbb{N}^k.$$

Proof. **Step 1. A priori estimates.** We decompose

$$\Lambda_{\mathbf{u}_0}(\mathbf{u}) = e^{it\Phi(D_x)}\mathbf{u}_0 + \int_0^t e^{i(t-t')\Phi(D_x)}\mathcal{N}(\mathbf{u}(t'))\,\mathrm{d}t' = I + II$$

where $\mathcal{N}(\mathbf{u})$ is given in (3.4.2).

Estimate for I. We fix some $T > 0$ to be chosen later and start with the estimates for I:

$$\|e^{it\Phi(D_x)}\mathbf{u}_0\|_{Y_{1,T}} = \max_{0 \leq h \leq s} \|\partial_x^h e^{it\Phi(D_x)}\mathbf{u}_0\|_{L^\infty_{t\in[-T,T]}L^2_x} \leq \max_{0 \leq h \leq s} \|\partial_x^h \mathbf{u}_0\|_{L^2_x} \lesssim \|\mathbf{u}_0\|_{H^s_x}$$

$$\|e^{it\Phi(D_x)}\mathbf{u}_0\|_{Y_{2,T}} = \max_{s \leq h \leq s+\lfloor\omega\rfloor} \|\partial_x^h e^{it\Phi(D_x)}\mathbf{u}_0\|_{L^\infty_x L^2_{t\in[-T,T]}} \overset{(3.4.15)}{\lesssim} \max_{0 \leq h \leq s} \||\partial_x|^h \mathbf{u}_0\|_{L^2_x} \lesssim \|\mathbf{u}_0\|_{H^s_x}$$

$$\|e^{it\Phi(D_x)}\mathbf{u}_0\|_{Y_{3,T}} = \max_{0 \leq h \leq s-\lceil\omega\rceil-\lceil r\rceil} \|\partial_x^h e^{it\Phi(D_x)}\mathbf{u}_0\|_{L^2_x L^\infty_{t\in[-T,T]}} \overset{(3.4.18)}{\lesssim_T} \max_{0 \leq h \leq s-\lceil\omega\rceil-\lceil r\rceil} \||\partial_x|^h \mathbf{u}_0\|_{H^r_x} \lesssim \|\mathbf{u}_0\|_{H^s_x}.$$

[41] In the sense of Fréchet differentiability.

For $Y_{4,T}$ we note that

$$
\begin{aligned}
xe^{it\Phi(D_x)}\mathbf{u} &= \mathcal{F}^{-1}\Big[i\partial_\xi\big(e^{it\Phi(\xi)}\widehat{\mathbf{u}_0}\big)\Big] \\
&= \mathcal{F}^{-1}\Big[e^{it\Phi(\xi)}i\partial_\xi - \int_0^t e^{i(t-\tau)\Phi(\xi)}\Phi'(\xi)e^{i\tau\Phi(\xi)}\,\mathrm{d}\tau\widehat{\mathbf{u}_0}\Big] \\
&= e^{it\Phi(D_x)}x\mathbf{u} - \int_0^t e^{i(t-\tau)\Phi(D_x)}\Phi'(D_x)e^{i\tau\Phi(D_x)}\mathbf{u}_0\,\mathrm{d}\tau\,, \qquad (3.4.28)
\end{aligned}
$$

and hence

$$
\begin{aligned}
\|e^{it\Phi(D_x)}\mathbf{u}_0\|_{Y_{4,T}} &= \max_{0\le h\le s'-\lceil\omega+r\rceil}\|x\partial_x^h e^{it\Phi(D_x)}\mathbf{u}_0\|_{L^2_xL^\infty_{t\in[-T,T]}} \\
&\lesssim_T \max_{0\le h\le s'-\lceil\omega+r\rceil}\|\partial_x^h e^{it\Phi(D_x)}(x\mathbf{u}_0)\|_{L^2_xL^\infty_{t\in[-T,T]}} \\
&\quad+\int_0^T\|\partial_x^h e^{i(t-\tau)\Phi(D_x)}\Phi'(D_x)e^{i\tau\Phi(D_x)}\mathbf{u}_0\|_{L^2_xL^\infty_{t\in[-T,T]}}\,\mathrm{d}\tau+\|e^{it\Phi(D_x)}\mathbf{u}_0\|_{Y_{3,T}}\,,
\end{aligned}
$$

where for the last term we applied the derivative ∂_x to x and used $s'-\lceil\omega+r\rceil-1\le s-\lceil\omega\rceil-\lceil r\rceil$. We estimate

$$
\max_{0\le h\le s'-\lceil\omega+r\rceil}\|\partial_x^h e^{it\Phi(D_x)}(x\mathbf{u}_0)\|_{L^2_xL^\infty_{t\in[-T,T]}} \overset{(3.4.18)}{\lesssim_T} \|x\mathbf{u}_0\|_{H^{s'}_x}
$$

$$
\begin{aligned}
\max_{0\le h\le s'-\lceil\omega+r\rceil}\int_0^T\|\partial_x^h e^{i(t-\tau)\Phi(D_x)}\Phi'(D_x)e^{i\tau\Phi(D_x)}\mathbf{u}_0\|_{L^2_xL^\infty_{t\in[-T,T]}}\,\mathrm{d}\tau &\overset{(3.4.18)}{\lesssim_T}\int_0^T\|\partial_x^h\Phi'(D_x)e^{i\tau\Phi(D_x)}\mathbf{u}_0\|_{H^r_x}\,\mathrm{d}\tau \\
&\overset{(3.4.3)}{\lesssim_T}\|\mathbf{u}_0\|_{H_x^{r+s'-\lceil\omega+r\rceil+n-1}}\lesssim\|\mathbf{u}_0\|_{H^s}\,.
\end{aligned}
$$

For $Y_{5,T}$ we use the same strategy and obtain

$$
\max_{0\le h\le s'-\lceil\omega\rceil}\|\partial_x^h e^{it\Phi(D_x)}(x\mathbf{u}_0)\|_{L^\infty_{t\in[-T,T]}L^2_x}\le\|x\mathbf{u}_0\|_{H^{s'}}
$$

$$
\max_{0\le h\le s'-\lceil\omega\rceil}\int_0^T\|\partial_x^h e^{i(t-\tau)\Phi(D_x)}\Phi'(D_x)e^{i\tau\Phi(D_x)}\mathbf{u}_0\|_{L^\infty_{t\in[-T,T]}L^2_x}\,\mathrm{d}\tau\overset{(3.4.3)}{\lesssim_T}\|\mathbf{u}_0\|_{H^{s'+n-1}}\lesssim\|\mathbf{u}_0\|_{H^s}\,.
$$

In summary,

$$
\|e^{it\Phi(D_x)}\mathbf{u}_0\|_{Y_T}\lesssim_T\|\mathbf{u}_0\|_{H^s}+\|\mathbf{u}_0\|_{H^{s',1}}\,.
$$

Estimate for II. Here we always assume that l is a multiindex with $l_1\ge\cdots\ge l_k$ in the product $\mathcal{N}^{k,l}(\mathbf{u})$. In the proof below, we write everything as if $k\ge2$. The term involving $\partial_x^{l_2}\mathbf{u}$ is the only one which we control with the L^2_x-norm, i. e. the only one that requires integrability.

We estimate first the terms involving coefficients $\mathbf{f}^{k,l}$ with $k\ge2$. At the end, we explain how the terms involving $\mathbf{f}^{1,l}$ and $\mathbf{g}$ can be dealt with analogously.

Let us first state for all $h\le s-1$ the following inequalities, due to Hölder's inequality in t and Sobolev embedding in x:

$$
\|\partial_x^h\mathbf{u}\|_{L^2_{t\in[-T,T]}L^2_x}\le T^{\frac12}\|\partial_x^h\mathbf{u}\|_{L^\infty_{t\in[-T,T]}L^2_x}\le T^{\frac12}\|\mathbf{u}\|_{Y_{1,T}} \qquad (3.4.29)
$$

$$
\|\partial_x^h\mathbf{u}\|_{L^\infty_{t\in[-T,T]}L^\infty_x}\lesssim\|\partial_x^h\mathbf{u}\|_{L^\infty_{t\in[-T,T]}L^2_x}+\|\partial_x^{h+1}\mathbf{u}\|_{L^\infty_{t\in[-T,T]}L^2_x}\lesssim\|\mathbf{u}\|_{Y_{1,T}}\,. \qquad (3.4.30)
$$

In the following we estimate II in the norms $Y_{1,T},\cdots,Y_{5,T}$. We show the $Y_{1,T}$-case in detail, and compute the total number of derivatives for other cases, so that the argument for $Y_{1,T}$-case works the same way.

(1) Estimates for $Y_{1,T}$.

For any $h\in[0,s]$, we apply (3.4.16) with the choice "$h_{(3.4.16)}$" $=\lceil\omega+\mu-1\rceil\wedge h\in[0,\omega+\mu)$ to estimate

$$
\left\|\int_0^t e^{i(t-t')\Phi(D_x)}\partial_x^h\Big(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u})\Big)\,\mathrm{d}t'\right\|_{L^\infty_{t\in[-T,T]}L^2_x}
$$

$$\overset{(3.4.16)}{\lesssim} T^{\alpha}\left\|\partial_x^{h-(\lceil\omega+\mu-1\rceil\wedge h)}\langle D_x\rangle^{\theta\mu}\Big(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u})\Big)\right\|_{L_x^pL^2_{t\in[-T,T]}}.$$

Here $\theta = \frac{\lceil\omega+\mu-1\rceil\wedge h}{\omega+\mu} \in [0,1)$, and $\alpha = \frac{1-\theta}{2} \in (0,\frac{1}{2}]$ and $p = \frac{2}{1+\theta} \in (1,2]$ depend on $\lceil\omega+\mu-1\rceil\wedge h$, which we suppress from our notation. We have chosen "$h_{(3.4.16)}$" $= \lceil\omega+\mu-1\rceil\wedge h < \omega+\mu$ to make sure that $\alpha > 0$, which is the source of smallness when T is chosen sufficiently small later.

Using Lemma 3.A.10 we can replace $\langle D_x\rangle^{\theta\mu}$ with integer derivatives up to order $\lceil\theta\mu\rceil$, and subsequently compute an upper bound for the total number of derivatives on the right-hand side above (recalling $|l| \le 2\omega-1$ in (3.4.2)) to be

$$\begin{aligned}
&\sup_{0\le h\le s} h - (\lceil\omega+\mu-1\rceil\wedge h) + \left\lceil\frac{\lceil\omega+\mu-1\rceil\wedge h}{\omega+\mu}\mu\right\rceil + |l| \\
&\le s - (\lceil\omega+\mu-1\rceil\wedge s) + \left\lceil\frac{\lceil\omega+\mu-1\rceil\wedge s}{\omega+\mu}\mu\right\rceil + 2\omega - 1 \le s + \lfloor\omega\rfloor,
\end{aligned}$$

where we have used for the last inequality that $\mu \le 5\omega$ and hence $\lceil\omega+\mu-1\rceil \le \lceil 7\omega\rceil \le \lceil s\rceil = s$, and furthermore our assumption $\omega,\mu \in \frac{1}{2}\mathbb{N}$ with $\mu\in\mathbb{N}$ or $\omega\in\mathbb{N}$ or $\mu\ge\omega$. The right-hand side above is then bounded by

$$T^{\alpha}\sum_{h\le s+1-\lceil\omega\rceil}\ \sum_{|l'|\le 2\omega-1}\ \sum_{|l|\le s+\omega}\left\|(\partial_x^h\mathbf{f}^{k,l'})(\mathcal{N}^{k,l}(\mathbf{u}))\right\|_{L_x^pL^2_{t\in[-T,T]}}.$$

We separate $\mathbb{R}$ into two integration domains $x\in[-1,1]$ and $x\in(-\infty,-1)\cup(1,\infty)$. Noticing that $\frac{1}{x} \in L^{\frac{2p}{2-p}}((-\infty,-1)\cup(1,\infty))$, we use Hölder's inequality in these two different domains to bound the above by

$$\begin{aligned}
&T^{\alpha}\sum_{h\le s+1-\lceil\omega\rceil}\ \sum_{|l'|\le 2\omega-1}\ \sum_{|l|\le s+\omega}\left\|(\partial_x^h\mathbf{f}^{k,l'})(\mathcal{N}^{k,l}(\mathbf{u}))\right\|_{L_x^2L^2_{t\in[-T,T]}} \\
&+T^{\alpha}\sum_{h\le s+1-\lceil\omega\rceil}\ \sum_{|l'|\le 2\omega-1}\ \sum_{|l|\le s+\omega}\left\|x(\partial_x^h\mathbf{f}^{k,l'})(\mathcal{N}^{k,l}(\mathbf{u}))\right\|_{L_x^2L^2_{t\in[-T,T]}} = II_1 + II_2.
\end{aligned}$$

We first focus on II_1. If $l_1 \le s-1$ then (3.4.29) and (3.4.30) imply

$$\sum_{h\le s+1-\lceil\omega\rceil}\ \sum_{|l'|\le 2\omega-1}\ \sum_{|l|\le s+\omega,\, l_k\le\cdots\le l_1\le s-1}\left\|(\partial_x^h\mathbf{f}^{k,l'})(\mathcal{N}^{k,l}(\mathbf{u}))\right\|_{L_x^2L^2_{t\in[-T,T]}} \lesssim_T \|\mathbf{u}\|^k_{Y_{1,T}}.$$

If $l_1 \ge s$ then $l_2 \le \omega \le s - \lceil\omega\rceil - \lceil r\rceil$, and hence

$$\begin{aligned}
&\sum_{h\le s+1-\lceil\omega\rceil}\ \sum_{|l'|\le 2\omega-1}\ \sum_{|l|\le s+\omega,\, s\le l_1\le s+\omega,\, l_k\le\cdots\le l_2\le s-\omega-r}\left\|(\partial_x^h\mathbf{f}^{k,l'})(\mathcal{N}^{k,l}(\mathbf{u}))\right\|_{L_x^2L^2_{t\in[-T,T]}} \\
&\le \sum_{h\le s+1-\lceil\omega\rceil}\ \sum_{|l'|\le 2\omega-1}\ \sum_{s\le l_1\le s+\omega,\, l_k\le\cdots\le l_2\le s-\omega-r}\|\partial_x^{l_1}\mathbf{u}\|_{L_x^\infty L^2_{t\in[-T,T]}}\left\|(\partial_x^h\mathbf{f}^{k,l'})(\partial_x^{l_2}\mathbf{u})\right\|_{L_x^2L^\infty_{t\in[-T,T]}} \\
&\quad\times\prod_{j=3}^{k}\|\partial_x^{l_j}\mathbf{u}\|_{L_x^\infty L^\infty_{t\in[-T,T]}} \lesssim_T \|\mathbf{u}\|_{Y_{2,T}}\|\mathbf{u}\|_{Y_{3,T}}\|\mathbf{u}\|^{k-2}_{Y_{1,T}}.
\end{aligned}$$

For II_2, we group the weight x with $\partial_x^{l_2}\mathbf{u}$ and replace the usage of $Y_{1,T}$ and $Y_{3,T}$ above by $Y_{5,T}$ and $Y_{4,T}$. This is possible thanks to a simple comparison:

$$l_2 \le \left\lfloor\frac{s+\omega}{2}\right\rfloor \le s' - \lceil\omega\rceil, \qquad \text{and if } l_1\ge s, \text{ then } l_2 \le \lfloor\omega\rfloor \le s' - \lceil\omega\rceil - \lceil r\rceil.$$

(2) Estimates for $Y_{2,T}$.

For $s \le h \le s+\lfloor\omega\rfloor$, we apply (3.4.17) with the choice "$h_{(3.4.17)}$" $= \lceil 2\omega+\mu-1\rceil \in [\omega+\mu, 2\omega+\mu)$, which requires our assumption $\mu\notin\mathbb{N}$ if $\omega=\frac{1}{2}$. This yields

$$\left\|\int_0^t e^{i(t-t')\Phi(D_x)}\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))\,\mathrm{d}t'\right\|_{L_x^\infty L^2_{t\in[-T,T]}}$$

$$\overset{(3.4.17)}{\lesssim} T^{\alpha}\left\|\partial_x^{h-\lceil 2\omega+\mu-1\rceil}\langle D_x\rangle^{\mu}(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))\right\|_{L_x^p L^2_{t\in[-T,T]}}.$$

Here $\theta = \frac{\lceil 2\omega+\mu-1\rceil-\omega-\mu}{\omega} \in [0,1)$, $\alpha = \frac{1-\theta}{2} > 0$ and $p = \frac{2}{1+\theta} \in (1,2]$. Note that $\lceil 2\omega+\mu-1\rceil \le 7\omega \le s \le h$, which ensures that the above is well-defined.

We shall distribute the derivatives and proceed as in (1). This is possible because the total number of derivatives we obtain is bounded by

$$s+\lfloor\omega\rfloor-\lceil 2\omega+\mu-1\rceil+\lceil\mu\rceil+2\omega-1 = s+\lfloor\omega\rfloor,$$

and the total number of derivatives falling on $\mathbf{f}^{k,l}$ is bounded by $s+\lfloor\omega\rfloor-2\omega+1 \le \lfloor s+1-\omega\rfloor = s+1-\lceil\omega\rceil$.

(3) Estimates for $Y_{3,T}$.

For $h \le s-\lceil\omega\rceil-\lceil r\rceil$, we estimate with (3.4.18):

$$\begin{aligned}
&\left\|\int_0^t e^{i(t-t')\Phi(D_x)}\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))\,\mathrm{d}t'\right\|_{L_x^2 L^\infty_{t\in[-T,T]}}\\
&\le \left\|e^{i(t-t')\Phi(D_x)}\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))(t',x)\right\|_{L^1_{t'\in[-T,T]}L_x^2L^\infty_{t\in[-T,T]}}\\
&\overset{(3.4.18)}{\lesssim_T} \left\|\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))\right\|_{L^1_{t\in[-T,T]}H^r_x}\\
&\lesssim T^{\frac12}\left\|\partial_x^h\langle D_x\rangle^r(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))\right\|_{L^2_{t\in[-T,T]}L^2_x}.
\end{aligned}$$

We again use Lemma 3.A.10 to replace $\langle D_x\rangle^r$ with integer derivatives up to order $\lceil r\rceil$, and compute for the total number of derivatives the upper bound $s-\lceil\omega\rceil-\lceil r\rceil+\lceil r\rceil+2\omega-1 \le s+\lfloor\omega\rfloor$, as well as for the derivatives falling on $\mathbf{f}^{k,l}$ the bound $s-\lceil\omega\rceil-\lceil r\rceil+\lceil r\rceil \le s+1-\lceil\omega\rceil$. We then proceed as in (1).

(4) Estimates for $Y_{4,T}$.

For $h \le s'-\lceil\omega+r\rceil$, we estimate

$$\begin{aligned}
&\left\|x\int_0^t e^{i(t-t')\Phi(D_x)}\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))\,\mathrm{d}t'\right\|_{L_x^2 L^\infty_{t\in[-T,T]}}\\
&\overset{(3.4.28)}{=} \left\|\int_0^t\left(e^{i(t-t')\Phi(D_x)}x-\int_0^{t-t'}e^{i(t-t'-\tau)\Phi(D_x)}\Phi'(D_x)e^{i\tau\Phi(D_x)}\,\mathrm{d}\tau\right)\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))\,\mathrm{d}t'\right\|_{L_x^2 L^\infty_{t\in[-T,T]}}\\
&\le \left\|e^{i(t-t')\Phi(D_x)}x\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))(t',x)\right\|_{L^1_{t'\in[-T,T]}L_x^2L^\infty_{t\in[-T,T]}}\\
&+\left\|e^{i(t-t'-\tau)\Phi(D_x)}\Phi'(D_x)e^{i\tau\Phi(D_x)}\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u})(t',x))\right\|_{L^1_{t',\tau\in[-T,T]^2}L_x^2L^\infty_{t\in[-T,T]}}\\
&\overset{(3.4.3),(3.4.18)}{\lesssim_T} T^{\frac12}\left\|x\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))\right\|_{L^2_{t\in[-T,T]}H^r_x}+T^{\frac32}\left\|\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))\right\|_{L^2_{t\in[-T,T]}H^{r+n-1}_x}.
\end{aligned}$$

It is easy to check that the total number of derivatives is bounded by $s+\lfloor\omega\rfloor$, and the number of derivatives falling on $\mathbf{f}^{k,l}$ is bounded by $s+1-\lceil\omega\rceil$. We proceed as in (1).

(5) Estimates for $Y_{5,T}$.

For $h \le s'-\lceil\omega\rceil$, we estimate

$$\begin{aligned}
&\left\|x\int_0^t e^{i(t-t')\Phi(D_x)}\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))\,\mathrm{d}t'\right\|_{L^\infty_{t\in[-T,T]}L_x^2}\\
&\overset{(3.4.28)}{=} \left\|\int_0^t\left(e^{i(t-t')\Phi(D_x)}x-\int_0^{t-t'}e^{i(t-t'-\tau)\Phi(D_x)}\Phi'(D_x)e^{i\tau\Phi(D_x)}\,\mathrm{d}\tau\right)\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}))\,\mathrm{d}t'\right\|_{L^\infty_{t\in[-T,T]}L_x^2}
\end{aligned}$$

$$\overset{(3.4.3)}{\lesssim} T^{\frac{1}{2}} \left\| x\partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u})) \right\|_{L^2_{t\in[-T,T]}L^2_x} + T^{\frac{3}{2}} \left\| \partial_x^h(\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u})) \right\|_{L^2_{t\in[-T,T]}H^{n-1}_x}.$$

We again distribute all derivatives and obtain a number of derivatives less than $s+\lfloor\omega\rfloor$, and number of derivatives falling on $\mathbf{f}^{k,l}$ is bounded by $s+1-\lceil\omega\rceil$, so we can proceed as in (1).

For the case $k=1$ the argument is identical, but each $\partial_x^{l_j}\mathbf{u}$ with $j>1$ must be replaced by 1. The assumption (3.4.22) has two cases. If $\mathbf{f}^{1,l}\in H^{s+1-\lceil\omega\rceil,1}$, then we group the weight x and the coefficient $\partial_x^h\mathbf{f}^{1,l}$ with $\partial_x^{l_2}\mathbf{u}=1$, so that their product is in L^2_x. If instead $\mathbf{f}^{1,l}\in W^{s+1-\lceil\omega\rceil,\infty}$, then we group also $\partial_x^{l_1}$ into the $L^\infty_x L^2_{t\in[-T,T]}$-norm and use the smoothing of the operator $\mathcal{K}^{1,l}$ on the space $H_x^{s'-\lceil\omega+r\rceil,1}L^\infty_{t\in[-T,T]}$, assumed in (3.4.22), to gain $\max\{0,|l|-(s'-\lceil\omega+r\rceil)\}$ derivatives.

Similarly, for the term involving $\mathbf{g}$ (referred to also as the case $k=0$) we replace every $\partial_x^{l_j}\mathbf{u}$ by 1, and we make use of L^2_x-integrability for the product $\partial_x^h\mathbf{g}\cdot\partial_x^{l_2}\mathbf{u}=\partial_x^h\mathbf{g}$ assumed in (3.4.20). When $k\geq 2$ we shall place the coefficients $\mathbf{f}^{k,l}$ in an L^∞_x-norm and keep $\partial_x^{l_2}\mathbf{u}$ in the L^2_x-norm.

Step 2. Fixed-point argument. We have shown in the estimates for I, II above that there exists some $\alpha>0$ (depending on s,ω,μ) such that

$$\begin{aligned}
\|e^{it\Phi(D_x)}\mathbf{u}_0\|_{Y_T} &\lesssim_T \|\mathbf{u}_0\|_{H^s\cap H^{s',1}} \\
\|\Lambda_{\mathbf{u}_0}(\mathbf{u})-e^{it\Phi(D_x)}\mathbf{u}_0\|_{Y_T} &\lesssim_T T^\alpha(1+\|\mathbf{u}\|_{Y_T})^K \\
&\lesssim_T T^\alpha\Big(1+\|\mathbf{u}-e^{it\Phi(D_x)}\mathbf{u}_0\|_{Y_T}+\|\mathbf{u}_0\|_{H^s\cap H^{s',1}}\Big)^K.
\end{aligned}$$

For some fixed $R>0$ to be determined later, we shall study the nonlinear map

$$\Lambda_{\mathbf{u}_0}:\mathbf{u}\mapsto\Lambda_{\mathbf{u}_0}(\mathbf{u})=e^{it\Phi(D_x)}\mathbf{u}_0+\int_0^t e^{i(t-t')\Phi(D_x)}\left(\mathbf{g}+\sum_{k=1}^K\sum_{|l|\leq n-2}\mathbf{f}^{k,l}\mathcal{N}^{k,l}(\mathbf{u}(t'))\right)\mathrm{d}t'$$

in the complete metric space

$$Y_T^R=\Big\{\mathbf{u}\in C_{t\in[-T,T]}(H^s\cap H^{s',1})_x:\|\mathbf{u}-e^{it\Phi(D_x)}\mathbf{u}_0\|_{Y_T}\leq R\Big\}$$

for some small enough time $T>0$. Indeed, notice first that by the a priori estimates in Step 1, for any $R>0$, for $\mathbf{u}\in e^{it\Phi(D_x)}\mathbf{u}_0+\overline{B_R^{Y_T}}$, there exists a small $T=T(\|\mathbf{u}_0\|_{H^s\cap H^{s',1}},R)>0$ such that

$$\|\Lambda_{\mathbf{u}_0}(\mathbf{u})-e^{it\Phi(D_x)}\mathbf{u}_0\|_{Y_T}<R.$$

This nonlinear map is well-defined in Y_T^R, as long as we can show $\Lambda_{\mathbf{u}_0}(\mathbf{u})\in C_{t\in[-T,T]}(H^s\cap H^{s',1})_x$. It follows from dominated convergence and the bounds we have achieved in Step 1. Indeed, the assumption on the initial data $\mathbf{u}_0\in H^s\cap H^{s',1}$ directly implies $e^{it\Phi(D_x)}\mathbf{u}_0\in C_{t\in[-T,T]}(H^s\cap H^{s',1})$. As for the "nonlinear" part, for $0<t'<t$, we decompose

$$\begin{aligned}
&\int_0^t e^{i(t-t'')\Phi(D_x)}\mathcal{N}(\mathbf{u}(t''))\,\mathrm{d}t''-\int_0^{t'}e^{i(t'-t'')\Phi(D_x)}\mathcal{N}(\mathbf{u}(t''))\,\mathrm{d}t'' \\
&=\int_{t'}^t e^{i(t-t'')\Phi(D_x)}\mathcal{N}(\mathbf{u}(t''))\,\mathrm{d}t''-\int_0^{t'}e^{i(t'-t'')\Phi(D_x)}\Big(1-e^{i(t-t')\Phi(D_x)}\Big)\mathcal{N}(\mathbf{u}(t''))\,\mathrm{d}t''.
\end{aligned}$$

It converges to 0 in $H^s\cap H^{s',1}$ as $t'\to t$, by (the proof ideas of) the estimate for II in Step 1.

Contraction mapping with small $T>0$. We study the difference

$$\Lambda_{\mathbf{u}_0}(\mathbf{u})-\Lambda_{\mathbf{u}_0}[\widetilde{\mathbf{u}}]=\sum_{k=1}^K\sum_{|l|\leq n-2}\int_0^t e^{i(t-t')\Phi(D_x)}\mathbf{f}^{k,l}\big(\mathcal{N}^{k,l}(\mathbf{u})-\mathcal{N}^{k,l}[\widetilde{\mathbf{u}}]\big)\,\mathrm{d}t'.$$

We expand the difference of products into a sum of products where exactly one factor is a difference between $\mathbf{u}$ and $\widetilde{\mathbf{u}}$. The a priori estimates for II previously described then directly show that there exists some $\alpha>0$ such that

$$\|\Lambda_{\mathbf{u}_0}(\mathbf{u})-\Lambda_{\mathbf{u}_0}[\widetilde{\mathbf{u}}]\|_{Y_T}\lesssim_T T^\alpha\|\mathbf{u}-\widetilde{\mathbf{u}}\|_{Y_T}(1+\|\mathbf{u}\|_{Y_T}+\|\widetilde{\mathbf{u}}\|_{Y_T})^{K-1} \tag{3.4.31}$$

$$\leq T^{\alpha}\|\mathbf{u}-\widetilde{\mathbf{u}}\|_{Y_T}(1+2R+2\|\mathbf{u}_0\|_{H^s\cap H^{s',1}})^{K-1}.$$

Hence, if we choose $R=\|\mathbf{u}_0\|_{H^s\cap H^{s',1}}$, there exists a sufficiently small $T>0$ depending on $\|\mathbf{u}_0\|_{H^s\cap H^{s',1}}$, such that the map $\Lambda_{\mathbf{u}_0}$ is a contraction map and has a unique fixed point in Y_T^R, which is a mild solution.

Step 3. Final check: Uniqueness, Lipschitz continuity, smooth flow and blow-up criterion. So far we have obtained uniqueness only in Y_T^R, and we now show the uniqueness of the mild solution in general. Let $\mathbf{u}\in Y_{T_1}$ and $\widetilde{\mathbf{u}}\in Y_{T_2}$ be two mild solutions of (3.4.1) with the same initial data $\mathbf{u}_0$. Then their difference $\mathbf{u}-\tilde{\mathbf{u}}=\Lambda_{\mathbf{u}_0}(\mathbf{u})-\Lambda_{\mathbf{u}_0}[\widetilde{\mathbf{u}}]$ satisfies the estimate (3.4.31). We choose $T\leq\min\{T_1,T_2\}$ sufficiently small (depending on $\|\mathbf{u}\|_{Y_{T_1}},\|\widetilde{\mathbf{u}}\|_{Y_{T_2}}$), such that $\|\mathbf{u}-\widetilde{\mathbf{u}}\|_{Y_T}=0$ and hence $\mathbf{u}=\widetilde{\mathbf{u}}$ on the time interval T. We can now start from the initial time T and proceed further up to $2T,3T,\cdots$, until the finite time $\min\{T_1,T_2\}$. The uniqueness follows.

The proof of local Lipschitz continuity of the data-to-solution map is analogous to the proof of the contraction mapping property, requiring again T to be sufficiently small. We omit the details here.

The smoothness of the data-to-solution map follows from the implicit mapping theorem for Banach spaces [128, Theorem 5.9]. Given $k\in\mathbb{N}$, we need to verify two assumptions. The first is that the mapping $\big[(\mathbf{u}_0,\mathbf{u})\mapsto\mathbf{u}-\Lambda_{\mathbf{u}_0}(\mathbf{u})\big]\in C^k((H^s\cap H^{s',1})\times Y_{T_0};Y_{T_0})$. Due to the fact that the nonlinearity is polynomial, this follows readily from the estimates we have demonstrated in the previous steps. The second is that the bounded linear operator $D_{\mathbf{u}}\big[(\mathbf{u}_0,\mathbf{u})\mapsto\mathbf{u}-\Lambda_{\mathbf{u}_0}(\mathbf{u})\big]=1-D_{\mathbf{u}}\Lambda_{\mathbf{u}_0}(\mathbf{u})\in L(Y_{T_0};Y_{T_0})$ has a bounded inverse. This can be ensured by choosing $T_0>0$ sufficiently small. We therefore obtain C^k regularity of the data-to-solution map.

Finally, as the existence time T of the mild solution we obtained from the fixed-point argument depends only on $\|\mathbf{u}_0\|_{H^s\cap H^{s',1}}$, the blow-up criterion (3.4.27) follows from the standard unique continuation argument. □

3.4.1.4. Improved blow-up criterion

In the following we first give some further properties of mild solutions, which help us to perform the energy estimates. These energy estimates subsequently imply an improved blow-up criterion, which finally yields the global well-posedness due to the conserved energies.

Lemma 3.4.7 (Further properties of mild solutions.). *Assume the setting of Theorem 3.4.6. The mild solution $\mathbf{u}\in Y_T$ of the general dispersive equation (3.4.1)–(3.4.2) satisfies the following:*

- $\Phi(D_x)\mathbf{u}\in C_{t\in[-T,T]}(H^{s-n}\cap H^{s'-n,1})_x$ *and* $\mathcal{N}(\mathbf{u})\in C_{t\in[-T,T]}(H^{s-n+2}\cap H^{s'-n+2,1})_x$.
- $\mathbf{u}\in C^1_{t\in[-T,T]}(H^{s-n}\cap H^{s'-n,1})_x$ *with derivative* $\partial_t\mathbf{u}=i\Phi(D_x)\mathbf{u}+\mathcal{N}(\mathbf{u})$*. In particular, for all* $h\leq s-n$ *the* L^2_x*-Bochner integral identity*

$$\partial_x^h\mathbf{u}(t)=\partial_x^h\mathbf{u}(0)+\int_0^t\partial_x^h(i\Phi(D_x)\mathbf{u}(t')+\mathcal{N}(\mathbf{u}(t')))\,\mathrm{d}t' \tag{3.4.32}$$

holds pointwise for almost all $x\in\mathbb{R}$.

Proof. Since $\mathbf{u}\in C_{t\in[-T,T]}(H^s\cap H^{s',1})_x$ we have $\Phi(D_x)\mathbf{u}\in C_{t\in[-T,T]}(H^{s-n}\cap H^{s'-n,1})_x$ due to (3.4.3). We want to show $\mathcal{N}(\mathbf{u})\in C_{t\in[-T,T]}(H^{s-n+2}\cap H^{s'-n+2,1})_x$. By the assumptions (3.4.21), in order to control the H^{s-n+2}-norm, it suffices to consider the limit $t'\to t$ of

$$\Big\|\partial_x^h\big(\mathbf{f}^{k,l}(\mathcal{N}^{k,l}(\mathbf{u})(t)-\mathcal{N}^{k,l}(\mathbf{u})(t'))\big)\Big\|_{L^2_x},\quad h\leq s-n+2,\quad 1\leq k\leq K,\quad |l|\leq n-2.$$

This is bounded by

$$\sum_{1\leq k\leq K}\sum_{h\leq s-n+2}\sum_{|l'|\leq n-2}\sum_{|l|\leq s}\|\partial_x^h\mathbf{f}^{k,l'}\|_{L^\infty}\sum_{j=1}^k\big\|\partial_x^{l_j}\mathbf{u}(t)-\partial_x^{l_j}\mathbf{u}(t')\big\|_{L^{q_j}}\prod_{\substack{i=1\\i\neq j}}^k\big(\big\|\partial_x^{l_i}\mathbf{u}_{b_i}(t)\big\|_{L^{q_i}}+\big\|\partial_x^{l_i}\mathbf{u}_{b_i}(t')\big\|_{L^{q_i}}\big)$$

$$\lesssim\|\mathbf{u}(t)-\mathbf{u}(t')\|_{H^s}(1+\|\mathbf{u}\|_{L^\infty_{t\in[-T,T]}H^s_x})^{K-1}$$

where we chose $q_j = 2$ for the j which maximizes $|l_j|$, and $q_j = \infty$ for all other j, so that the embedding $H^{\frac{s}{2}}(\mathbb{R}) \subset H^1(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$ applied for the inequality. Thus the time continuity of $\mathbf{u}$ in H^s_x implies the time continuity of $\mathcal{N}(\mathbf{u})$ in H^{s-n+2}_x. The time continuity in $H^{s'-n+2,1}$ of $\mathcal{N}(\mathbf{u})$ follows in the same way, grouping the weight x with the j-th factor.

Next, we show that $\mathbf{u} \in C_{t\in[-T,T]}(H^{s-n} \cap H^{s'-n,1})_x$ is continuously differentiable at $t = 0$. Differentiability at any $t_0 \in [-T, T]$ can then be obtained by considering $\mathbf{u}(t + t_0)$ as a mild solution in Y_{T-t_0} with initial data $\mathbf{u}(t_0)$. Applying (3.4.19) yields

$$\begin{aligned}\frac{\mathbf{u}(t)-\mathbf{u}_0}{t} - i\Phi(D_x)\mathbf{u}_0 - \mathcal{N}(\mathbf{u}_0) &= \left(\frac{e^{it\Phi(D_x)}-1}{t} - i\Phi(D_x)\right)\mathbf{u}_0\\ &\quad + \frac{1}{t}\int_0^t (e^{i(t-t')\Phi(D_x)}-1)\mathcal{N}(\mathbf{u}(t')) + \mathcal{N}(\mathbf{u}(t')) - \mathcal{N}(\mathbf{u}_0)\,\mathrm{d}t'\,.\end{aligned}$$

We use (3.4.3) to obtain the vanishing limit of H^{s-n}-norm of the first term as $t \to 0$, while for all $h \le s-n$, the second term can be controlled by dominated convergence, after a change of variables $t' \mapsto tt'$ and then using estimates as above:

$$\left\| \int_0^1 \xi^h \Big((e^{it(1-t')\Phi(\xi)} - 1)\widehat{\mathcal{N}(\mathbf{u})}(tt') + \widehat{\mathcal{N}(\mathbf{u})}(tt') - \widehat{\mathcal{N}(\mathbf{u}_0)}\Big)\,\mathrm{d}t' \right\|_{L^2_\xi} \xrightarrow{t\to 0} 0.$$

The case of $H^{s'-n,1}$-norm follows similarly, with some terms containing an additional $\Phi'(D_x)$ that can be treated since $s' + n - 1 \le s$. ☐

Lemma 3.4.8 (Estimate of $H^{s',1}$-norm and improved blow-up criterion)**.** *Assume the setting of Theorem 3.4.6 and assume in addition $s' \le s - n$, which requires $s \ge 7\omega + 2$. There exists a constant $C > 0$, depending on s, s', n, and the norms assumed to be finite in Theorem 3.4.6, such that for any initial data $\mathbf{u}_0 \in H^s \cap H^{s',1}$ the mild solution $\mathbf{u} \in Y_T$ to the general dispersive equation* (3.4.1)–(3.4.2) *satisfies*

$$\|\mathbf{u}\|_{L^\infty_{t\in[-T,T]}H^{s',1}_x} \le \exp\Big(C(1+T)\big(C + \|\mathbf{u}_0\|_{H^{s',1}} + \|\mathbf{u}\|_{L^\infty_{t\in[-T,T]}L^\infty_x} + \|\mathbf{u}_x\|_{L^\infty_{t\in[-T,T]}H^{s-1}_x}\big)^{K+1}\Big)\,. \tag{3.4.33}$$

In particular, the blow-up criterion (3.4.27) *can be improved into*

$$\sup_{t\in(-T^*,T^*)} (\|\mathbf{u}(t)\|_{L^\infty} + \|\mathbf{u}_x(t)\|_{H^{s-1}}) = \infty. \tag{3.4.34}$$

Proof. Here we do not assume $J = 1$, i. e. $\Phi \in \mathbb{R}^{J\times J}$ may be a matrix. Recall that we have assumed the existence of a real symmetric uniformly positive definite matrix symbol $S = S(\xi) \in L^\infty(\mathbb{R}; \mathbb{R}^{J\times J})$, i. e. we have the equivalence of quadratic forms $\langle S(D_x)\mathbf{u}, \mathbf{u}\rangle_{L^2} \approx \|\mathbf{u}\|_{L^2}$, so that in addition $S\Phi - \Phi^T S \in L^\infty(\mathbb{R}; \mathbb{R}^{J\times J})$.

We shall estimate $\|x\partial_x^h \mathbf{u}\|_{L^\infty_{t\in[-T,T]}L^2_x}$ for $h \le s' \le s - n$ with Grönwall's inequality, and note that the norms $\|\partial_x^h \mathbf{u}\|_{L^\infty_{t\in[-T,T]}L^2_x}$ can be treated analogously.

In the classical continuous differentiable setting, one can easily derive from the differential equation $v' = F(v)$ that $(v\bar{v})' = \bar{v}F(v) + v\overline{F(v)}$. We can formulate it in the integral form as follows: If $v(t) = v(0) + \int_0^t F(v(t'))\,\mathrm{d}t'$, then $v(t)\bar{v}(t) = v(0)\overline{v(0)} + \int_0^t \overline{v(t')}F(v(t')) + v(t')\overline{F(v(t'))}\,\mathrm{d}t'$. We apply this fact to the integral formulation (3.4.32) of our equation in Lemma 3.4.7 to obtain for all $h \le s - n$

$$\begin{aligned}\int_{\mathbb{R}} S(D_x)(x\partial_x^h)\mathbf{u}\cdot\overline{x\partial_x^h\mathbf{u}}\,\mathrm{d}x &= \int_{\mathbb{R}} x\partial_x^h\mathbf{u}_0\cdot\overline{S(D_x)x\partial_x^h\mathbf{u}_0}\,\mathrm{d}x\\ &+ \int_0^t\int_{\mathbb{R}} xi\Phi(D_x)\partial_x^h\mathbf{u}(t')\cdot\overline{S(D_x)x\partial_x^h\mathbf{u}(t')} + x\partial_x^h\mathbf{u}(t')\cdot\overline{S(D_x)xi\Phi(D_x)\partial_x^h\mathbf{u}(t')}\,\mathrm{d}x\,\mathrm{d}t'\\ &+ \int_0^t\int_{\mathbb{R}} x\partial_x^h\mathcal{N}(\mathbf{u}(t'))\cdot S(D_x)\overline{x\partial_x^h\mathbf{u}(t')} + x\partial_x^h\mathbf{u}(t')\cdot\overline{S(D_x)x\partial_x^h\mathcal{N}(\mathbf{u}(t'))}\,\mathrm{d}x\,\mathrm{d}t'\end{aligned}$$

$$= III_1 + III_2 + III_3\,.$$

We first note the trivial estimate $|III_1| \le \|\mathbf{u}_0\|^2_{H_x^{s',1}}$. Next, we estimate the nonlinear part III_3. Here

$$\left|\int_0^t \int_{\mathbb{R}} x\partial_x^h \mathbf{g} \cdot \overline{S(D_x)x\partial_x^h \mathbf{u}(t')}\,\mathrm{d}x\,\mathrm{d}t'\right| \le \int_0^t \|\mathbf{g}\|_{H^{s',1}}\|\mathbf{u}(t')\|_{H^{s',1}}\,\mathrm{d}t' \lesssim \int_0^t \|\mathbf{u}(t')\|^2_{H^{s',1}} + \|\mathbf{g}\|^2_{H^{s',1}}\,\mathrm{d}t'\,.$$

For other terms involving $\mathbf{f}^{k,l}$, it suffices to assume $l_1 \ge l_2 \ge \cdots \ge l_k$ and $|l| \le s-n+n-2 = s-2$, $h \le s'$, $|l'| \le n-2$, and estimate $x^2(\partial_x^h \mathbf{f}^{k,l'})(\mathcal{N}^{k,l}(\mathbf{u}))(\overline{\partial_x^h \mathbf{u}})$. If $k \ge 2$ then $l_k \le \frac{s-1}{2} \le s'$, and we have

$$\begin{aligned}
&\left|\int_0^t \int_{\mathbb{R}} x(\partial_x^h \mathbf{f}^{k,l'})(\mathcal{N}^{k,l}(\mathbf{u})(t')) \cdot \overline{S(D_x)x\partial_x^h \mathbf{u}(t')}\,\mathrm{d}x\,\mathrm{d}t'\right| \\
&\le \int_0^t \|(\partial_x^h \mathbf{f}^{k,l'})(x\partial_x^{l_k}\mathbf{u}(t'))\|_{L^2}\|x\partial_x^h \mathbf{u}(t')\|_{L^2}\prod_{j=1}^{k-1}\|\partial_x^{l_j}\mathbf{u}(t')\|_{L^\infty}\,\mathrm{d}t' \\
&\lesssim \int_0^t \|\mathbf{u}(t')\|^2_{H^{s',1}}(1 + \|\mathbf{f}^{k,l'}\|_{W^{s',\infty}} + \|\mathbf{u}(t')\|_{L^\infty} + \|\mathbf{u}_x(t')\|_{H^{s-1}})^{K-1}\,\mathrm{d}t'\,.
\end{aligned}$$

If $k=1$ then we face the problem that $l_1 = l_k$ may be as large as $s-2$. According to (3.4.22) there are two cases. When $\mathbf{f}^{1,l'} \in H^{s+1-\lceil\omega\rceil,1}$ we place $\partial_x^{l_1}\mathbf{u}$ into the L^∞-norm and group the weight x with $\partial_x^h \mathbf{f}^{k,l'}$, subsequently applying the Sobolev embedding $H^{s'+1} \hookrightarrow W^{s',\infty}$. Alternatively, we use the smoothing provided by $\mathcal{K}^{1,l'} \in L(H^{s',1}; H^{|l'|,1})$.

It remains to estimate the linear part III_2. With the Plancherel theorem, we can write

$$\begin{aligned}
-\frac{1}{2}III_2 &= \mathrm{Im}\left[\int_0^t \int_{\mathbb{R}} x\Phi(D_x)D_x^h \mathbf{u}(t') \cdot \overline{S(D_x)xD_x^h \mathbf{u}(t')}\,\mathrm{d}x\,\mathrm{d}t'\right] \\
&= \mathrm{Im}\left[\int_0^t \int_{\mathbb{R}} \partial_\xi(\Phi(\xi)\xi^h \widehat{\mathbf{u}}(t')) \cdot \overline{S(\xi)\partial_\xi(\xi^h \widehat{\mathbf{u}}(t'))}\,\mathrm{d}\xi\,\mathrm{d}t'\right] \\
&= \frac{i}{2}\int_0^t \int_{\mathbb{R}} (S\Phi - \Phi^T S)(\xi)\partial_\xi(\xi^h \widehat{\mathbf{u}}(t')) \cdot \overline{\partial_\xi(\xi^h \widehat{\mathbf{u}}(t'))}\,\mathrm{d}\xi\,\mathrm{d}t' \\
&\quad + \mathrm{Im}\left[\int_0^t \int_{\mathbb{R}} \Phi'(\xi)\xi^h \widehat{\mathbf{u}}(t') \cdot \overline{S(\xi)\partial_\xi(\xi^h \widehat{\mathbf{u}}(t'))}\,\mathrm{d}\xi\,\mathrm{d}t'\right] \\
&\lesssim \int_0^t \|\mathbf{u}(t')\|^2_{H^{s',1}} + \|\mathbf{u}(t')\|_{H^{s',1}}\|\mathbf{u}_x(t')\|_{H^{s-1}}\,\mathrm{d}t' \lesssim \int_0^t \|\mathbf{u}(t')\|^2_{H^{s',1}} + \|\mathbf{u}_x(t')\|^2_{H^{s-1}}\,\mathrm{d}t'\,.
\end{aligned}$$

Here we have used Cauchy–Schwarz with $h \le s'$ and (3.4.3). As mentioned above, the norms $\|\partial_x^h \mathbf{u}\|^2_{L_x^2}$ can be estimated analogously. To conclude, we have proven

$$\begin{aligned}
\|\mathbf{u}(t)\|^2_{H^{s',1}} \lesssim \|\mathbf{u}_0\|^2_{H_x^{s',1}} &+ \int_0^t \|\mathbf{u}(t')\|^2_{H^{s',1}}(1 + \|\mathbf{f}^{k,l'}\|_{W^{s',\infty}} + \|\mathbf{u}(t')\|_{L^\infty} + \|\mathbf{u}_x(t')\|_{H^{s-1}})^{K-1}\,\mathrm{d}t' \\
&+ \int_0^t \left(\|\mathbf{g}\|^2_{H^{s',1}} + \|\mathbf{u}_x(t')\|^2_{H^{s-1}}\right)\mathrm{d}t'\,,
\end{aligned}$$

and Grönwall's inequality then yields (3.4.33). Since $\|\mathbf{u}\|_{L_x^2} \lesssim \|\mathbf{u}\|_{H_x^{s',1}}$, we have

$$\|\mathbf{u}(t)\|_{H_x^s} \lesssim \|\mathbf{u}(t)\|_{H_x^{s',1}} + \|\mathbf{u}_x(t)\|_{H_x^{s-1}}.$$

Hence the same bound as (3.4.33) holds for $\|\mathbf{u}\|_{L^\infty_{t\in[-T,T]}H_x^s}$. Thus, for any $T \in (0,\infty)$, $\|\mathbf{u}\|_{L^\infty_{t\in[-T,T]}L_x^\infty} + \|\mathbf{u}_x\|_{L^\infty_{t\in[-T,T]}H_x^{s-1}} < \infty$ yields $\|\mathbf{u}\|_{L^\infty_{t\in[-T,T]}(H_x^s \cap H_x^{s',1})} < \infty$. Therefore the blow-up criterion (3.4.27) can be improved into (3.4.34). $\square$

3.4.2. Global well-posedness of the NLS, GP, and $\widetilde{\mathsf{GP}}$ hierarchies

Proof of Theorem 3.1.2 and Corollary 3.1.7. We first show the global well-posedness of ($\widetilde{\mathsf{GP}}_n$), and then of the $\widetilde{\mathsf{GP}}$, NLS, and GP hierarchies.

Step 1. Global well-posedness of $(\widetilde{\mathsf{GP}}_n)$**.** Note that $(\widetilde{\mathsf{GP}}_1)$ is a transport equation with constant speed -1, for which global well-posedness trivially holds. We therefore assume $n \geq 2$ and use the formulation of $(\widetilde{\mathsf{GP}}_n)$ given in Proposition 3.1.14. We show now the global well-posedness of $(\widetilde{\mathsf{GP}}_n)$ in the sense of Theorem 3.1.2. For the odd flows $(\widetilde{\mathsf{GP}}_{2m+1})$ we can use the variable $\mathbf{p}$ and the formulation (3.1.21), while for the even flows $(\widetilde{\mathsf{GP}}_{2m})$ we need to use $\mathbf{u} = \mathcal{T}^{-1}\mathbf{p}$ and the corresponding formulation (3.1.26). Here we recall that $R > 0$ is chosen large, $\eta \in C_c^\infty(\mathbb{R}; [0, 1])$ with $\eta|_{[-R,R]} = 1$ and $\operatorname{supp}\eta \subseteq [-2R, 2R]$, and

$$\mathcal{T} = 1 + B\psi(D_x) \quad \text{where} \quad \psi(\xi) = \frac{1-\eta(\xi)}{\xi^2+2} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1-q_*^2 \\ 1-\overline{q}_*^2 & 0 \end{pmatrix}.$$

We have to verify that (3.1.21) and (3.1.26) satisfy the hypotheses of Theorem 3.4.6, and correspondingly also Lemma 3.4.7 and Lemma 3.4.8. For the verification that the linear parts are admissible, we refer to Lemma 3.4.1. The conditions on the nonlinear part of (3.1.21) are weaker than those required on (3.1.26), so we focus on the latter case.

We now represent everything except the intended linear part in (3.1.26) as a sum of expressions of the form (3.4.2). Below, we use the keyword "or" to denote alternatives that may appear in separate versions of (3.4.2). For the linear operators, we choose

$$\begin{aligned}
\mathcal{J}_{j,b}^{k,l} &= \mathcal{T}^{-1} = (1 + B\psi(D_x))^{-1} \\
\mathcal{K}_{j,b,h}^{k,l} &= \mathcal{T} \quad \text{or} \quad \psi(D_x) \quad \text{or} \quad D_x^{-1}\psi(D_x) \quad \text{or} \quad (1-\eta(D_x)) \quad \text{or} \quad D_x^{-1}(1-\eta(D_x)) \quad \text{or} \quad \mathbb{1}_{\{k=1\}}\eta(D_x)\,.
\end{aligned}$$

The coefficients $\mathbf{g}_j$ and $\mathbf{f}_{j,b}^{k,l}$ have the form

$$\begin{aligned}
\mathbf{g}_j &= \mathbb{1}_{\{n=2\}}\mathcal{O}^{1,0}_{q_*,|q_*|^2-1}(|q_*|^2-1) + \mathcal{O}^{1,n}_{q_*}((q_*)_x) \\
\mathbf{f}_{j,b}^{k,l} &= \mathcal{O}^{0,n-2}_{q_*,|q_*|^2-1}((q_*)_x, |q_*|^2-1) \quad \text{or} \quad -\mathbb{1}_{\{k=1\}}[\sigma_3, B]\,,
\end{aligned}$$

Note that the operator $\eta(D_x)$ and the coefficient $-[\sigma_3, B]$ appear only in the term $-[\sigma_3, B]D_x^{2m-2}\eta(D_x)\mathbf{u}$ of (3.1.26). If the background function q_* has finite energy $E^{s+1-\lceil\omega\rceil+n,1}(q_*) < \infty$, then $\mathbf{g}_j, \mathbf{f}_{j,b}^{1,l} \in H^{s+1-\lceil\omega\rceil,1}$ and $\mathbf{f}_{j,b}^{k,l} \in W^{s+1-\lceil\omega\rceil,\infty}$ for $k \geq 1$, except for the special coefficient $-[\sigma_3, B]$ which may not be in $H^{s+1-\lceil\omega\rceil,1}$. These coefficients satisfy (3.4.20)–(3.4.22) for Theorem 3.4.6. We need to prove that $\mathcal{J}_{j,b}^{k,l}$ and $\mathcal{K}_{j,b,h}^{k,l}$ satisfy (3.4.23)–(3.4.25). This follows from $B \in W^{s+\lfloor\omega\rfloor,\infty}$, the fact that the symbols $\psi(\xi)$, $\partial_\xi\psi(\xi)$, $\eta(\xi)$, and $\partial_\xi\eta(\xi)$ have inverse Fourier transforms in L^1, and the possibility of choosing $R > 0$ sufficiently large to ensure the invertibility of $\mathcal{T}$. An exception is the symbol $\xi(1-\eta(D_x))$, whose inverse Fourier transform is not in L^1. Nevertheless, one can check manually that this choice of $\mathcal{K}_{j,b,h}^{k,l}$ still admits (3.4.23)–(3.4.25). Finally, we note that the operator $\eta(D_x)$ has the smoothing property required by (3.4.22).

Applying Theorem 3.4.6, we obtain a local mild solution $\mathbf{u}$ to (3.1.26) with a time of existence $T > 0$ and initial data $\mathbf{u}_0$. Furthermore, Lemma 3.4.7 and Lemma 3.4.8 apply to this solution. Using $\mathcal{T} \in L(H^s) \cap L(H^{s',1})$ we find that $\mathbf{p}$ solves (3.1.21) in the same sense, i. e. analogues of Theorem 3.4.6, Lemma 3.4.7, and Lemma 3.4.8 hold. Then Lemma 3.4.7 implies that $q = q_* + p$, with initial data $q_* + p_0$, solves $(\widetilde{\mathsf{GP}}_n)$ in $H^{s-n} \cap H^{s'-n,1}$. By the blow-up criterion in Lemma 3.4.8, it suffices to prove

$$\|p\|_{L^\infty_{t\in[-T,T]}L^\infty_x} + \|p_x\|_{L^\infty_{t\in[-T,T]}H^{s-1}_x} \lesssim_{T,s,q_*,p_0} 1$$

in order to obtain the global-in-time well-posedness.

By [119, Lemma 2.1] the map $[p \mapsto q_* + p] : H^s \to X^s$ (see (3.1.15) for the definition of X^s) is continuous. Consequently, we know that $q \in C_{t\in[-T,T]}X^s$. Furthermore, in [119, 120] certain energy functionals $\mathcal{E}^s : X^s \to \mathbb{R}$ (see [119, Theorem 1.3]) are defined, which are conserved under every flow (GP_n) of the hierarchy. They are conserved because they are constructed from the renormalized transmission coefficient (defined in [119, 120]), which by an approximation argument and Lemma 3.2.9 Poisson commutes with the Hamiltonians $\mathcal{H}_n^{\mathsf{GP}}$, and hence by virtue of the relations (3.1.6) with the Hamiltonians $\mathcal{H}_n^{\widetilde{\mathsf{GP}}}$. We therefore obtain, as in [119, Theorem 1.3], the uniform bound

$$\sup_{t\in[-T,T]} E^s(q(t)) \lesssim_{s,q_*,p_0} 1\,. \tag{3.4.35}$$

We conclude the global well-posedness of ($\widetilde{\mathsf{GP}}_n$) with the observation that $p(t) = q(t) - q_*$ implies

$$\sup_{t\in[-T,T]} \left(\|p(t)\|_{L_x^\infty} + \|p_x(t)\|_{H_x^{s-1}}\right) \lesssim E^s(q_*) + \sup_{t\in[-T,T]} E^s(q(t)) \lesssim_{s,q_*,p_0} 1\,.$$

Thus the uniform bound in (3.1.14) follows, and the exponential-in-time bound for $\|p(t)\|_{H_x^{s',1}} \leq Ce^{Ct}$ follows from (3.4.33).

Step 2. Global well-posedness of hierarchies. We first notice that the lower bounds $3n-1+\omega$, $\frac{s+3\omega}{2}$ for the indices s, s' are increasing in n while the upper bound $s-n$ for s' is decreasing in n. Thus if s, s' satisfy these bounds for N as in Corollary 3.1.7, then they satisfy the corresponding bounds for all $n \leq N$.

We call the solution of the perturbative formulation of ($\widetilde{\mathsf{GP}}_n$) constructed just now $p^{(1)}$, which depends only on a single time variable t_n for $n \geq 1$. We now define a perturbation $p^{(2)} \in C(\mathbb{R}^N; H^s \cap H^{s',1})$ which depends on $(t_1, \dots, t_N)$ by moving the initial data for a time t_1 along the flow of ($\widetilde{\mathsf{GP}}_1$), for a time t_2 along the flow of ($\widetilde{\mathsf{GP}}_2$), and so on. As a consequence of the commutativity of the flows, which is shown in Lemma 3.4.9 below, this is independent of the order in which we move along the flows. Then $p^{(2)}$ solves for $n \in \{1, \cdots, N\}$ the perturbative formulation of ($\widetilde{\mathsf{GP}}_n$) for the stationary front q_* strongly in $H^{s-n} \cap H^{s'-n,1}$. We obtain the solution $q_* + p(t)$ of (GP_n) from

$$p(t) = p^{(2)}(\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}\mathbf{t}, x) \text{ with } \mathbf{t} = (0, \cdots, 0, t, 0, \cdots)\,.$$

This, together with the bounds (3.1.14) for $p^{(2)}$, concludes Theorem 3.1.2.

We have excluded so far the case $n = 0$. The flow of ($\widetilde{\mathsf{GP}}_0$) consists simply of multiplication with e^{-it_0}, so we define $p^{(3)} \in C(\mathbb{R}^{N+1}; H^s \cap H^{s',1})$ by

$$p^{(3)}(t_0, t_1, \dots, t_N) = e^{-it_0} p^{(2)}(t_1, \dots, t_N)\,.$$

Then $p^{(3)}$ solves for all $n \in \{0, \cdots, N\}$ the perturbative formulation of ($\widetilde{\mathsf{GP}}_n$) for the rotating front $e^{-it_0}q_*$ strongly in $H^{s-n} \cap H^{s'-n,1}$. The proofs of (i) and (iii) in Corollary 3.1.7 are finished by defining $p(\mathbf{t}) = p^{(3)}(\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{t})$. Finally, the bounds in (ii) follow from [119, Lemma 2.1, Theorem 1.3], as in (3.4.35). $\square$

Lemma 3.4.9. *Assume the setting of Theorem 3.1.2 and define for $n \in \mathbb{N}$ the flow map*

$$\begin{aligned}\Omega_n : \mathbb{R} \times (H^s \cap H^{s',1}) &\longrightarrow (H^s \cap H^{s',1})\\ (t, p_0) &\longmapsto p(t)\end{aligned}$$

where p is the solution to ($\widetilde{\mathsf{GP}}_n$). For all $p_0 \in H^s \cap H^{s',1}$ and $n, m \in \mathbb{N}$, $t, r \in \mathbb{R}$ we have

$$\Omega_n(t, \Omega_m(r, p_0)) = \Omega_m(r, \Omega_n(t, p_0))\,.$$

Proof. Let $p_0 \in H^s \cap H^{s',1}$ be fixed initial data. We know from Lemma 3.4.7 that $\Omega_n(\cdot, p_0) \in C^1(\mathbb{R}; H^{s-n} \cap H^{s'-n,1})$, and from Theorem 3.4.6 that there exists a neighborhood $U_0 \subseteq H^s \cap H^{s',1}$ of p_0, as well as a small time $T_0 > 0$, such that $\Omega_n(t, \cdot) \in C^2(U_0; H^s \cap H^{s',1})$ for all $t \in [-T_0, T_0]$. We denote by

$$\begin{aligned} D\Omega_n(t,p)[u] &= \frac{\mathrm{d}}{\mathrm{d}\mu}\Big|_{\mu=0} \Omega_n(t, p + \mu u)\\ D^2\Omega_n(t,p)[u,v] &= \frac{\mathrm{d}}{\mathrm{d}\mu}\Big|_{\mu=0} \frac{\mathrm{d}}{\mathrm{d}\nu}\Big|_{\nu=0} \Omega_n(t, p + \mu u + \nu v)\end{aligned}$$

the first and second Fréchet derivatives at $p \in U_0$ for directions $u, v \in H^s \cap H^{s',1}$. The argument below requires some more regularity, specifically that $D\Omega_n \in C^1([-T_0, T_0] \times U_0; L(H^{s-n} \cap H^{s'-n,1}))$. Indeed, since continuous partial differentiability implies continuous differentiability, it suffices to prove that

$$\begin{aligned}
D\Omega_n(t, \cdot) &\in C^1(U_0; L(H^{s-n} \cap H^{s'-n,1})) \\
D\Omega_n(\cdot, p) &\in C^1([-T_0, T_0]; L(H^{s-n} \cap H^{s'-n,1})) \\
\partial_t D\Omega_n &\in C^0([-T_0, T_0] \times U_0; L(H^{s-n} \cap H^{s'-n,1}))\,. \\
D^2\Omega_n &\in C^0([-T_0, T_0] \times U_0; L((H^{s-n} \cap H^{s'-n,1})^{\otimes 2}; (H^{s-n} \cap H^{s'-n,1})))\,.
\end{aligned}$$

We already know the first statement to hold true. The second statement follows by substituting $D\Omega_n(\cdot, p) = D\Lambda_{p_0}(p)$ and arguing in a similar fashion to Lemma 3.4.7. The third and fourth statements can be easily seen to hold true once the derivatives are written down.

Define

$$\begin{aligned}
X_n : H^s \cap H^{s',1} &\longrightarrow H^{s-n} \cap H^{s'-n,1} \\
p &\longmapsto \frac{1}{i} \frac{\delta \mathcal{H}_n^{\widetilde{\mathsf{GP}}}(q)}{\delta \overline{q}}\Big|_{q=q_*+p}
\end{aligned}$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Omega_n(t, p) = X_n(\Omega_n(t, p))\,.$$

We can alternatively write

$$\frac{\mathrm{d}}{\mathrm{d}t}\Omega_n(t, p) = \frac{\mathrm{d}}{\mathrm{d}\mu}\Big|_{\mu=0}\Omega_n(t, \Omega_n(\mu, p)) = D\Omega_n(t, p)[X_n(p)]\,.$$

Let $n, m \in \mathbb{N}$. By the continuity of the flow map, we may assume without loss of generality that $p \in U_0 \cap H^{s+(n\vee m)} \cap H^{s'+(n\vee m),1}$. The argument below involves taking two derivatives, which leads to a loss of $n + m$ derivatives. By assuming additional regularity for p, we are able to perform the following calculations in $H^{s-n} \cap H^{s'-n,1}$. Let $t, r \in [-T_0, T_0]$ and define

$$a = \Omega_n(t, p) \qquad b = \Omega_m(r, a) \qquad c = \Omega_n(-t, b) \qquad d = \Omega_m(-r, c)\,.$$

Our goal is to prove $d = p$, for which it suffices to prove $\frac{\mathrm{d}}{\mathrm{d}r}d = 0$. We write

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}r}d &= \frac{\mathrm{d}}{\mathrm{d}r}\Omega_m(-r, \Omega_n(-t, \Omega_m(r, \Omega_n(t, p)))) \\
&= D\Omega_m(-r, c)[D\Omega_n(-t, b)[X_m(b)] - X_m(c)]\,.
\end{aligned}$$

Due to the polynomial structure of the vector fields X_n it is easy to see that $X_n \in C^1(H^s \cap H^{s',1}; H^{s-n} \cap H^{s'-n,1})$. We define the Lie bracket $[X_n, X_m] \in C(H^s \cap H^{s',1}; H^{s-n} \cap H^{s'-n,1})$ by

$$[X_n, X_m](p) = DX_n(p)[X_m(p)] - DX_m(p)[X_n(p)]\,.$$

It is a direct consequence of Lemma 3.2.9 that $[X_n, X_m](p) = 0$ for every $p \in H^s \cap H^{s',1}$.

Since $b = \Omega_n(t, c)$, to show $\frac{\mathrm{d}}{\mathrm{d}r}d = 0$ it suffices to prove for any fixed $c \in U_0 \cap H^{s+(n\vee m)} \cap H^{s'+(n\vee m),1}$ and t in a small neighborhood of 0 that

$$D\Omega_n(-t, \Omega_n(t, c))[X_m(\Omega_n(t, c))] - X_m(c) = 0\,.$$

We set $b = \Omega_n(t, c)$ and calculate

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}D\Omega_n(-t, \Omega_n(t, c))[X_m(\Omega_n(t, c))] = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}\mu}\Big|_{\mu=0}\Omega_n\big(-t, \Omega_n(t, c) + \mu X_m(\Omega_n(t, c))\big) \\
&= \frac{\mathrm{d}}{\mathrm{d}\nu}\Big|_{\nu=0}\frac{\mathrm{d}}{\mathrm{d}\mu}\Big|_{\mu=0}\Big(-\Omega_n\big(-t, \Omega_n(\nu, b + \mu X_m(b))\big) + \Omega_n\big(-t, \Omega_n(t + \nu, c) + \mu X_m(b)\big)
\end{aligned}$$

$$
\begin{aligned}
&\qquad\quad + \Omega_n\big(- t, b + \mu X_m(\Omega_n(t+\nu, c))\big)\Big) \\
&= \frac{\mathrm{d}}{\mathrm{d}\mu}\Big|_{\mu=0}\Big(- D\Omega_n(-t, b + \mu X_m(b))[X_n(b + \mu X_m(b))] + D\Omega_n(-t, b + \mu X_m(b))[X_n(b)] \\
&\qquad + D\Omega_n(-t, b + \mu X_m(b))[\mu D X_m(b)[X_n(b)]]\Big) \\
&= -D\Omega_n(-t, b)[D X_n(b)[X_m(b)]] + \frac{\mathrm{d}}{\mathrm{d}\mu}\Big|_{\mu=0}\Big(- D\Omega_n(-t, b + \mu X_m(b))[X_n(b)] \\
&\qquad + D\Omega_n(-t, b + \mu X_m(b))[X_n(b)]\Big) + D\Omega_n(-t, b)[D X_m(b)[X_n(b)]] \\
&= D\Omega_n(-t, b)[[X_m, X_n](b)] = 0\,.
\end{aligned}
$$

□

3.A. Appendix: Linear Estimates

For $f = f(t,x)$ we write $\widehat{f}^{(t)}(\tau, x)$ and $\widehat{f}^{(x)}(t, \xi)$ for the Fourier transforms in only one variable. We define

$$
\mathcal{D}_{\otimes}(\mathbb{R}^2) = \mathcal{D}(\mathbb{R}) \otimes \mathcal{D}(\mathbb{R}) = \left\{ \sum_{n=1}^{N} f_n \otimes g_n : N \in \mathbb{N}, f = (f_1, \dots, f_N), g = (g_1, \dots, g_N) \in \mathcal{D}(\mathbb{R})^N \right\}.
$$

3.A.1. The local smoothing estimate

We assume the following properties for the symbol $\varphi \in C^1(\mathbb{R};\mathbb{R})$:

(H1) **Finitely many distinct critical points.** φ has finitely many critical points $\xi_1 < \cdots < \xi_N$, i. e. $\varphi'(\xi) = 0$ if and only if $\xi \in \{\xi_1, \dots, \xi_N\}$. We define in addition $\xi_0 = -\infty$ and $\xi_{N+1} = \infty$. Set $\eta_j = \varphi(\xi_j)$. We consider the bijective restrictions of φ to the intervals (ξ_j, ξ_{j+1}), on which φ is strictly monotone, and denote them by

$$
\begin{aligned}
\varphi_j : (\xi_j, \xi_{j+1}) &\longrightarrow (\widehat{\eta}_j, \check{\eta}_j) = \varphi((\xi_j, \xi_{j+1})) \\
\psi_j = \varphi_j^{-1} : (\widehat{\eta}_j, \check{\eta}_j) &\longrightarrow (\xi_j, \xi_{j+1})
\end{aligned}
$$

for $j \in \{0, \dots, N\}$. Here $\widehat{\eta}_j = \eta_j \wedge \eta_{j+1}$ and $\check{\eta}_j = \eta_j \vee \eta_{j+1}$. We write $\sigma_j = \operatorname{sign}(\varphi_j')$ for the sign.

(H2) **Hölder continuity at critical points.** There exists a small $r > 0$ such that for every $j \in \{1, \dots, N\}$ there exist $\alpha_j \in (0,1)$ and $q, Q > 0$ such that for all $\eta \in (0, r)$

$$
q\eta^{\alpha_j} \le |\psi_{j-1}(\eta_j - \sigma_{j-1}\eta) - \xi_j|, |\psi_j(\eta_j + \sigma_j\eta) - \xi_j| \le Q\eta^{\alpha_j} \tag{3.A.1}
$$

$$
q\eta^{\alpha_j - 1} \le |\psi_{j-1}'(\eta_j - \sigma_{j-1}\eta)|, |\psi_j'(\eta_j + \sigma_j\eta)| \le Q\eta^{\alpha_j - 1}\,. \tag{3.A.2}
$$

(H3) **Sublinear growth at infinity.** There exists a large $R > 0$ and some $\beta \in (0,1)$ and $p, P > 0$ such that for all $\eta > R$ we have

$$
p\eta^{\beta} \le |\psi_0(-\sigma_0\eta)|, |\psi_N(\sigma_N\eta)| \le P\eta^{\beta} \tag{3.A.3}
$$

$$
p\eta^{\beta-1} \le |\psi_0'(-\sigma_0\eta)|, |\psi_N'(\sigma_N\eta)| \le P\eta^{\beta-1}\,. \tag{3.A.4}
$$

Because of (H3) we have $\widehat{\eta}_0 = -\sigma_0\infty$ and $\check{\eta}_N = \sigma_N\infty$. Given any function $f : \mathbb{R} \to \mathbb{C}$ for which the integrals below exist, we can perform the change of variables $\eta = \varphi(\xi)$ by splitting the integral:

$$
\int_{\mathbb{R}} f(\xi)\varphi'(\xi)\,\mathrm{d}\xi = \sum_{j=0}^{N} \sigma_j \int_{\widehat{\eta}_j < \eta < \check{\eta}_j} f(\psi_j(\eta))\,\mathrm{d}\eta\,. \tag{3.A.5}
$$

Sometimes we assume without loss of generality that $\sigma_0 = -1$.

We further assume as given a symbol $a \in \operatorname{Lip}(\mathbb{R};\mathbb{R})$ with the following properties:

(H4) **Boundedness and decay at infinity.** There exist $A,\delta>0$ such that for almost all $\xi\in\mathbb{R}$ and $\eta>R$

$$|a(\xi)|+\langle\xi\rangle|a'(\xi)|+\langle\eta\rangle^{\delta}|a(\psi_0(-\sigma_0\eta))-a(\psi_N(\sigma_N\eta))|\le A\,.$$

(H5) **Monotonicity.** There exists $M>0$ such that for any $\tau\in\mathbb{R}$ we can decompose $\mathbb{R}$ into M intervals on whose interiors $\frac{a(\xi)}{\varphi(\xi)-\tau}$ and φ' are monotone. In particular, for any $j\in\{0,\dots,N\}$ we can decompose $(\tau\vee\widehat{\eta}_j,\check{\eta}_j)$ and $(\widehat{\eta}_j,\check{\eta}_j\wedge\tau)$ into M intervals on whose interiors the functions $\eta\mapsto\frac{a(\psi_j(\eta))}{\eta-\tau}$ and $\eta\mapsto\psi_j'$ are monotone.

We can freely choose R to be an arbitrarily large constant and r an arbitrarily small constant, both depending only on φ and a. The goal of this section is to prove the following theorem.

Theorem 3.A.1 (Generalization of [110, Theorem 3.5])**.** *Let $\varphi\in C^1(\mathbb{R};\mathbb{R})$ and $a\in\operatorname{Lip}(\mathbb{R};\mathbb{R})$ fulfill (H1)–(H5). Let $\widetilde{a}\in L^\infty(\mathbb{R};\mathbb{R})$. Then*

$$\left\|\big(\widetilde{a}|\varphi'|^{\frac12}\big)(D_x)e^{it\varphi(D_x)}u_0\right\|_{L^\infty_xL^2_t}\lesssim_{\varphi,\widetilde{a}}\|u_0\|_{L^2_x}\tag{3.A.6}$$

$$\left\|(a\varphi')(D_x)\int_0^\infty e^{i(t-t')\varphi(D_x)}f(t',x)\,\mathrm{d}t'\right\|_{L^\infty_xL^2_t}\lesssim_{\varphi,a}\|f\|_{L^1_xL^2_t}\tag{3.A.7}$$

$$\left\|(a\varphi')(D_x)\int_0^t e^{i(t-t')\varphi(D_x)}f(t',x)\,\mathrm{d}t'\right\|_{L^\infty_xL^2_t}\lesssim_{\varphi,a}\|f\|_{L^1_xL^2_t}\,.\tag{3.A.8}$$

Proof. We directly prove (3.A.6):

$$\begin{aligned}
&\int_{\mathbb{R}}\left|\big(\widetilde{a}|\varphi'|^{\frac12}\big)(D_x)e^{it\varphi(D_x)}u_0(x)\right|^2\mathrm{d}t=\frac{1}{2\pi}\int_{\mathbb{R}}\left|\int_{\mathbb{R}}e^{i(x\xi+t\varphi(\xi))}\big(\widetilde{a}|\varphi'|^{\frac12}\big)(\xi)\widehat{u_0}(\xi)\,\mathrm{d}\xi\right|^2\mathrm{d}t\\
&=\frac{1}{2\pi}\int_{\mathbb{R}}\left|\int_{\mathbb{R}}e^{it\eta}\left(\sum_{j=0}^{N}\mathbb{1}_{(\widehat{\eta}_j,\check{\eta}_j)}(\eta)e^{ix\psi_j(\eta)}\big(\widetilde{a}|\varphi'|^{\frac12}\big)(\psi_j(\eta))\widehat{u_0}(\psi_j(\eta))\psi_j'(\eta)\right)\mathrm{d}\eta\right|^2\mathrm{d}t\\
&\lesssim\int_{\mathbb{R}}\left|\sum_{j=0}^{N}\mathbb{1}_{(\widehat{\eta}_j,\check{\eta}_j)}(\eta)e^{ix\psi_j(\eta)}(\widetilde{a}\circ\psi_j)(\eta)(\widehat{u_0}\circ\psi_j)(\eta)|\varphi_j'\circ\psi_j|^{-\frac12}(\eta)\right|^2\mathrm{d}\eta\\
&\lesssim_{N,\|\widetilde{a}\|_{L^\infty}}\sum_{j=0}^{N}\int_{\xi_j}^{\xi_{j+1}}|\widehat{u_0}(\xi)|^2|\varphi_j'(\xi)|^{-1}|\varphi_j'(\xi)|\,\mathrm{d}\xi\\
&\lesssim_N\|u_0\|_{L^2_x}^2\,.
\end{aligned}$$

From (3.A.6) we know that $\big(\widetilde{a}|\varphi'|^{\frac12}\big)(D_x)e^{it\varphi(D_x)}:L^2_x(\mathbb{R})\to L^\infty_xL^2_t(\mathbb{R}\times\mathbb{R}_+)$ is bounded. The boundedness of the corresponding adjoint yields the estimate

$$\left\|\big(\widetilde{a}|\varphi'|^{\frac12}\big)(D_x)\int_0^\infty e^{i(t-t')\varphi(D_x)}f(t',x)\,\mathrm{d}t'\right\|_{L^\infty_xL^2_t(\mathbb{R}\times\mathbb{R}_+)}\lesssim_{N,A}\|f\|_{L^1_xL^2_t(\mathbb{R}\times\mathbb{R}_+)}\,.$$

Consider now a test function $g\in\mathcal{D}_\otimes(\mathbb{R}^2)$. We apply the above with $\widetilde{a}=|a|^{\frac12}$ and $\widetilde{a}=\operatorname{sign}(a\varphi')|a|^{\frac12}$ to obtain

$$\begin{aligned}
&\int_0^\infty\int_{\mathbb{R}}(a\varphi')(D_x)\int_0^\infty e^{i(t-t')\varphi(D_x)}f(t',x)\,\mathrm{d}t'g(t,x)\,\mathrm{d}x\,\mathrm{d}t\\
&=\int_{\mathbb{R}}\big(|a\varphi'|^{\frac12}\big)(D_x)\int_0^\infty e^{-it'\varphi(D_x)}f(t',x)\,\mathrm{d}t'\,\big(\operatorname{sign}(a\varphi')|a\varphi'|^{\frac12}\big)(D_x)\int_0^\infty e^{it\varphi(D_x)}g(t,x)\,\mathrm{d}t\,\mathrm{d}x\\
&\lesssim_{N,A}\|f\|_{L^1_xL^2_t(\mathbb{R}\times\mathbb{R}_+)}\|g\|_{L^1_xL^2_t(\mathbb{R}\times\mathbb{R}_+)}\,,
\end{aligned}$$

By duality, this implies (3.A.7).

For (3.A.8) we follow [110, §3]. This method is also outlined in [109, Theorem 2.3] and [106]. Consider the solution

$$u(t,x)=\int_0^t e^{i(t-t')\varphi(D_x)}f(t',x)\,\mathrm{d}t'\quad\text{to}\quad D_tu=\varphi(D_x)u-if\quad\text{with initial data}\quad u(0)=0\,.$$

Then

$$\widehat{u}(\tau,\xi) = \frac{-i\widehat{f}(\tau,\xi)}{\tau - \varphi(\xi)}\,.$$

Define

$$v(t,x) = \mathrm{P.V.}\,\frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} e^{it\tau}e^{ix\xi}\frac{-i\widehat{f}(\tau,\xi)}{\tau-\varphi(\xi)}\,\mathrm{d}\xi\,\mathrm{d}\tau\,,$$

which is formally another solution, although it may or may not fulfill $v(0)=0$. Using Proposition 3.A.3 below, which states that one can change the order between P.V. and the integration w. r. t. ξ, and $\widehat{\mathrm{sign}}(\tau) = -i\sqrt{\frac{2}{\pi}}\,\mathrm{P.V.}\big(\frac{1}{\tau}\big)$, we write

$$\begin{aligned}(a\varphi')(D_x)v(t,x) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{ix\xi}\,\mathrm{P.V.}\,\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{it\tau}\frac{-i(a\varphi')(\xi)\widehat{f}(\tau,\xi)}{\tau-\varphi(\xi)}\,\mathrm{d}\tau\,\mathrm{d}\xi \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{ix\xi}\frac{1}{2}\int_{\mathbb{R}} e^{i(t-t')\varphi(\xi)}\,\mathrm{sign}(t-t')(a\varphi')(\xi)\widehat{f}^{(x)}(t',\xi)\,\mathrm{d}t'\,\mathrm{d}\xi \\ &= \frac{1}{2}\int_{\mathbb{R}}(a\varphi')(D_x)e^{i(t-t')\varphi(D_x)}\,\mathrm{sign}(t-t')f(t',x)\,\mathrm{d}t'\,.\end{aligned}$$

We obtain the decomposition

$$\begin{aligned}&(a\varphi')(D_x)u(t,x) \\ &= \frac{1}{2}\int_{\mathbb{R}}(a\varphi')(D_x)e^{i(t-t')\varphi(D_x)}(\mathrm{sign}(t-t')-\mathrm{sign}(-t'))f(t',x)\,\mathrm{d}t' \\ &= (a\varphi')(D_x)v(t,x) + \frac{1}{2}\int_0^\infty(a\varphi')(D_x)e^{i(t-t')\varphi(D_x)}f(t',x)\,\mathrm{d}t' - \frac{1}{2}\int_{-\infty}^0(a\varphi')(D_x)e^{i(t-t')\varphi(D_x)}f(t',x)\,\mathrm{d}t'\,.\end{aligned}$$

These calculations are identical to [110, Proposition 3.1, Lemma 3.4] for $\varphi(\xi)=\xi^3$. We can estimate the integrals at the end with (3.A.7) and a corresponding version of (3.A.7) on $\mathbb{R}_-$, so it remains to show that $\|(a\varphi')(D_x)v\|_{L^\infty_x L^2_t} \lesssim_{a,\varphi} \|f\|_{L^1_x L^2_t}$. By Proposition 3.A.4 below, which states that one may change the order between P.V. and the integration w. r. t. τ, we have

$$(a\varphi')(D_x)v(t,x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{it\tau}\,\mathrm{P.V.}\int_{\mathbb{R}} e^{ix\xi}\frac{-i(a\varphi')(\xi)\widehat{f}(\tau,\xi)}{\tau-\varphi(\xi)}\,\mathrm{d}\xi\,\mathrm{d}\tau\,,$$

and by Proposition 3.A.5 there exists some kernel $K\in L^\infty(\mathbb{R}^2)$ such that

$$(a\varphi')(D_x)v(t,x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{it\tau}\frac{i}{\sqrt{2\pi}}\int_{\mathbb{R}} K(x-y,\tau)\widehat{f}^{(t)}(\tau,y)\,\mathrm{d}y\,\mathrm{d}\tau\,.$$

These propositions correspond to [110, Proposition 3.2, 3.3]. We use Plancherel's theorem and Minkowski's inequality to obtain

$$\begin{aligned}\|(a\varphi')(D_x)v\|_{L^\infty_x L^2_t} &= \frac{1}{(2\pi)^{\frac{3}{2}}}\left\|\left(\int_{\mathbb{R}}\left|\int_{\mathbb{R}} K(x-y,\tau)\widehat{f}^{(t)}(\tau,y)\,\mathrm{d}y\right|^2\mathrm{d}\tau\right)^{\frac{1}{2}}\right\|_{L^\infty_x} \\ &\le \frac{1}{(2\pi)^{\frac{3}{2}}}\|K\|_{L^\infty}\int_{\mathbb{R}}\left(\int_{\mathbb{R}}|\widehat{f}^{(t)}(\tau,y)|^2\,\mathrm{d}\tau\right)^{\frac{1}{2}}\mathrm{d}y \\ &\lesssim_{\varphi,a} \|f\|_{L^1_x L^2_t}\,.\end{aligned}$$

□

Remark 3.A.2. *One may wonder whether it is possible to directly obtain (3.A.8) from (3.A.7) by a Christ-Kiselev type lemma. Applying the original Christ-Kiselev lemma from [50] would require a bounded operator $L^2_t L^1_x \to L^2_t L^\infty_x$. A reversed Christ-Kiselev lemma is given in [41, Theorem B] for bounded operators $L^{p_1}_x L^{q_1}_t \to L^{p_2}_x L^{q_2}_t$ with $p_1 \vee q_1 < p_2 \wedge q_2$. In our situation $p_1 = 1$, $p_2 = \infty$, and $q_1 = q_2 = 2$, so we are in an endpoint case.*

Proposition 3.A.3 (Generalization of [110, Proposition 3.1])**.** *Let $f \in \mathcal{D}_\otimes(\mathbb{R}^2)$. Then for all $(t,x) \in \mathbb{R}^2$ we have*

$$\lim_{\varepsilon\searrow 0} \iint\limits_{\varepsilon<|\varphi(\xi)-\tau|<\frac{1}{\varepsilon}} e^{i(x\xi+t\tau)} \frac{(a\varphi')(\xi)}{\varphi(\xi)-\tau} \widehat{f}(\tau,\xi)\,\mathrm{d}\xi\,\mathrm{d}\tau = \int_{\mathbb{R}} e^{ix\xi} \lim_{\varepsilon\searrow 0} \int\limits_{\varepsilon<|\varphi(\xi)-\tau|<\frac{1}{\varepsilon}} e^{it\tau} \frac{(a\varphi')(\xi)}{\varphi(\xi)-\tau} \widehat{f}(\tau,\xi)\,\mathrm{d}\tau\,\mathrm{d}\xi\,.$$

Moreover, both terms are finite and both limits exist.

Proof. The proof is essentially identical to that of [110, Proposition 3.1]. It suffices to assume that $f(t,x) = h(x)g(t)$. Then

$$\lim_{\varepsilon\searrow 0} \int\limits_{\varepsilon<|\varphi(\xi)-\tau|<\frac{1}{\varepsilon}} e^{it\tau} \frac{(a\varphi')(\xi)}{\varphi(\xi)-\tau} \widehat{f}(\tau,\xi)\,\mathrm{d}\tau = (a\varphi'\widehat{h})(\xi) H[e^{it\tau}\widehat{g}(\tau)](\varphi(\xi))\,,$$

where H denotes the Hilbert transform. Since $H : \mathcal{D}(\mathbb{R}) \to L^\infty(\mathbb{R})$, the integral

$$\int_{\mathbb{R}} e^{ix\xi} (a\varphi'\widehat{h})(\xi) H[e^{it\tau}\widehat{g}(\tau)](\varphi(\xi))\,\mathrm{d}\xi$$

is absolutely convergent. Define

$$H^*[w](x) = \sup_{\varepsilon>0} H_\varepsilon[w](x) = \sup_{\varepsilon>0} \left| \int\limits_{\varepsilon<|x-y|<\frac{1}{\varepsilon}} \frac{w(y)}{x-y}\,\mathrm{d}y \right|$$

and recall that $\lim_{\varepsilon\searrow 0} H_\varepsilon[w](x) = H[w](x)$ for almost all $x \in \mathbb{R}$, and furthermore that $H^* : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ is bounded for $1 < p < \infty$. Now consider

$$\begin{aligned}\iint\limits_{\varepsilon<|\varphi(\xi)-\tau|<\frac{1}{\varepsilon}} e^{i(x\xi+t\tau)} \frac{(a\varphi')(\xi)}{\varphi(\xi)-\tau} \widehat{f}(\tau,\xi)\,\mathrm{d}\xi\,\mathrm{d}\tau &= \int_{\mathbb{R}} e^{ix\xi} (a\varphi'\widehat{h})(\xi) H[e^{it\tau}\widehat{g}(\tau)](\varphi(\xi))\,\mathrm{d}\xi \\ &\quad - \int_{\mathbb{R}} e^{ix\xi} (a\varphi'\widehat{h})(\xi) (H-H_\varepsilon)[e^{it\tau}\widehat{g}(\tau)](\varphi(\xi))\,\mathrm{d}\xi \\ &=: (I) - (II_\varepsilon)\,.\end{aligned}$$

It remains to prove that $\lim_{\varepsilon\searrow 0} |(II_\varepsilon)| = 0$. We have

$$(II_\varepsilon) = \sum_{j=0}^{N} \sigma_j \int_{\widehat{\eta}_j}^{\check{\eta}_j} e^{ix\psi_j(\eta)} (a\widehat{h})(\psi_j(\eta)) (H-H_\varepsilon)[e^{it\tau}\widehat{g}(\tau)](\eta)\,\mathrm{d}\eta\,.$$

Since $a, \widehat{h} \in L^\infty(\mathbb{R})$ and $|(H_\varepsilon - H)[g]| \le 2H^*[g]$ is bounded on $L^p(\mathbb{R})$, the claim follows by dominated convergence. ☐

Proposition 3.A.4 (Generalization of [110, Proposition 3.2])**.** *Let $f \in \mathcal{D}_\otimes(\mathbb{R}^2)$. Then for all $(t,x) \in \mathbb{R}^2$ we have*

$$\lim_{\varepsilon\searrow 0} \iint\limits_{\varepsilon<|\varphi(\xi)-\tau|<\frac{1}{\varepsilon}} e^{i(x\xi+t\tau)} \frac{(a\varphi')(\xi)}{\varphi(\xi)-\tau} \widehat{f}(\tau,\xi)\,\mathrm{d}\xi\,\mathrm{d}\tau = \int_{\mathbb{R}} e^{it\tau} \lim_{\varepsilon\searrow 0} \int\limits_{\varepsilon<|\varphi(\xi)-\tau|<\frac{1}{\varepsilon}} e^{ix\xi} \frac{(a\varphi')(\xi)}{\varphi(\xi)-\tau} \widehat{f}(\tau,\xi)\,\mathrm{d}\xi\,\mathrm{d}\tau\,.$$

Moreover, both integrals are absolutely convergent and both limits exist.

Proof. The proof is again almost identical to that of [110, Proposition 3.2]. In Proposition 3.A.3 the first limit is shown to exist, using that for $g \in \mathcal{D}(\mathbb{R})$ the limit $H[g](x) = \lim_{\varepsilon\searrow 0} H_\varepsilon[g](x)$ exists for almost all $x \in \mathbb{R}$ and that $H^*[g] \in L^p(\mathbb{R})$. Both the statement and the proof of this proposition are analogous to those of the previous one, except H, H_ε and H^* are replaced by $\mathcal{K}, \mathcal{K}_\varepsilon$ and $\mathcal{K}^*$, defined by

$$\mathcal{K}_\varepsilon[g](\tau) = \int\limits_{\varepsilon<|\varphi(\xi)-\tau|<\frac{1}{\varepsilon}} \frac{(a\varphi')(\xi)}{\varphi(\xi)-\tau} g(\xi)\,\mathrm{d}\xi \quad \mathcal{K}[g](\tau) = \lim_{\varepsilon\searrow 0} \mathcal{K}_\varepsilon[g](\tau) \quad \mathcal{K}^*[g](\tau) = \sup_{\varepsilon>0} |\mathcal{K}_\varepsilon[g](\tau)|\,.$$

The key steps in the proof are therefore the existence of the limit $\mathcal{K}[g]$, and $\mathcal{K}^*[g] \in L^p(\mathbb{R})$ for $1 < p < \infty$. We write

$$\mathcal{K}_\varepsilon[g](\tau) = \sum_{j=0}^{N} \sigma_j \int\limits_{\varepsilon<|\eta-\tau|<\frac{1}{\varepsilon}} \frac{\mathbb{1}_{(\widehat{\eta}_j,\check{\eta}_j)}(\eta)(ag)(\psi_j(\eta))}{\eta-\tau}\,\mathrm{d}\eta\,.$$

Since $a \in L^\infty(\mathbb{R})$, it suffices to prove that $g \circ \psi_j \in L^p((\widehat{\eta}_j, \check{\eta}_j))$ for $1 \le p \le \infty$ and use known properties of the Hilbert transform. We have

$$\sigma_j \int_{\widehat{\eta}_j}^{\check{\eta}_j} |g(\psi_j(\eta))|^p\,\mathrm{d}\eta = \int_{\xi_j}^{\xi_{j+1}} |g(\xi)|^p \varphi_j'(\xi)\,\mathrm{d}\xi\,,$$

so our assumptions $\varphi \in C^1(\mathbb{R},\mathbb{R})$ and $g \in \mathcal{D}(\mathbb{R})$ are sufficient. $\square$

Proposition 3.A.5 (Generalization of [110, Proposition 3.3])**.** *For functions φ, a satisfying (H1)–(H5), we define*

$$K_\varepsilon(z,\tau) = \int\limits_{\varepsilon<|\varphi(\xi)-\tau|<\frac{1}{\varepsilon}} e^{iz\xi}\frac{(a\varphi')(\xi)}{\varphi(\xi)-\tau}\,\mathrm{d}\xi \qquad\qquad K(z,\tau) = \lim_{\varepsilon\searrow 0} K_\varepsilon(z,\tau)\,.$$

Then there exists some $\varepsilon_0 > 0$ for which the family $\{K_\varepsilon\}_{0<\varepsilon<\varepsilon_0} \subset L^\infty(\mathbb{R}^2)$ is bounded. Furthermore, the limit $\lim_{\varepsilon\to 0} K_\varepsilon(z,\tau) = K(z,\tau)$ exists for all $(z,\tau) \in \mathbb{R}^2$ and $K \in L^\infty(\mathbb{R}^2)$.

Subsequently, for any $f \in \mathcal{D}_\otimes(\mathbb{R}^2)$ and $(t,x) \in \mathbb{R}^2$ we have

$$\lim_{\varepsilon\searrow 0} \frac{1}{\sqrt{2\pi}} \int\limits_{\varepsilon<|\varphi(\xi)-\tau|<\frac{1}{\varepsilon}} e^{ix\xi}\frac{(a\varphi')(\xi)}{\varphi(\xi)-\tau}\widehat{f}(\tau,\xi)\,\mathrm{d}\xi = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} K(x-y,\tau)\widehat{f}^{(t)}(\tau,y)\,\mathrm{d}y\,. \tag{3.A.9}$$

Here the proof requires some notable modifications from that of [110, Proposition 3.3]. This stems from the fact that the change of variables $\eta = \varphi(\xi)$ is much more involved. In fact, after stating the following Van der Corput type lemma, we shall devote the rest of this section to the proof of Proposition 3.A.5.

Lemma 3.A.6 (Van der Corput type lemma)**.** *Let $f \in C^1([a,b];\mathbb{R})$ such that f' is nonzero and monotone. Then*

$$\left|\int_a^b e^{izf(x)}\,\mathrm{d}x\right| \le \frac{4}{|z|}\sup_{x\in(a,b)}\frac{1}{|f'(x)|}\,, \qquad \forall\, z \in \mathbb{R}\setminus\{0\}\,.$$

If some function $g \in \mathrm{Lip}([a,b];\mathbb{R})$ has the property that $[a,b]$ can be decomposed into M intervals on whose interiors g is monotone, then

$$\left|\int_a^b e^{izf(x)}g(x)\,\mathrm{d}x\right| \le \frac{12M}{|z|}\sup_{x\in[a,b]}\frac{1}{|f'(x)|}\sup_{x\in[a,b]}|g(x)|\,, \qquad \forall\, z \in \mathbb{R}\setminus\{0\}\,.$$

Proof. The first classical estimate follows from integration by parts and the monotonicity of the function $\frac{1}{f'(x)}$. The second estimate follows similarly from integration by parts and the same monotonicity trick to write $\|g'\|_{L^1([a,b])} = |g(b)-g(a)|$. $\square$

Proof of Proposition 3.A.5. We first notice, given the uniform boundedness and convergence results on K_ε, that we can obtain from Plancherel's theorem the representation

$$\frac{1}{\sqrt{2\pi}} \int\limits_{\varepsilon<|\varphi(\xi)-\tau|<\frac{1}{\varepsilon}} e^{ix\xi}\frac{(a\varphi')(\xi)}{\varphi(\xi)-\tau}\widehat{f}(\tau,\xi)\,\mathrm{d}\xi = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} K_\varepsilon(x-y,\tau)\widehat{f}^{(t)}(\tau,y)\,\mathrm{d}y\,, \tag{3.A.10}$$

which implies (3.A.9) thanks to the dominated convergence theorem.

Thus it remains to show the uniform boundedness of $(K_\varepsilon)_{0<\varepsilon<\varepsilon_0} \subseteq L^\infty(\mathbb{R}^2)$ as well as the pointwise existence of the limit as $\varepsilon \to 0$. Recall that

$$K_\varepsilon(z,\tau) = \int\limits_{\varepsilon<|\varphi(\xi)-\tau|<\frac{1}{\varepsilon}} e^{iz\xi}\,\frac{(a\varphi')(\xi)}{\varphi(\xi)-\tau}\,\mathrm{d}\xi = \sum_{j=0}^{N} \sigma_j \int\limits_{\substack{\varepsilon<|\eta-\tau|<\frac{1}{\varepsilon}\\ \widehat{\eta}_j<\eta<\check{\eta}_j}} e^{iz\psi_j(\eta)}\,\frac{a(\psi_j(\eta))}{\eta-\tau}\,\mathrm{d}\eta\,.$$

Once the uniform boundedness is established, an inspection of the proof reveals that the convergence of all integrals is either trivial, in cases where the integral has no singularity for sufficiently small ε, or follows from dominated convergence. All integrals corresponding to the latter case are indeed estimated by exhibiting an integrable majorant, as opposed to using oscillation, i. e. Lemma 3.A.6.

We recall that $r > 0$ is an arbitrarily small constant and $R > 0$ is an arbitrarily large constant. We consider the strip $(\tau - r, \tau + r)$. For every $0 \le j \le N$ exactly one of three cases below holds true, and we aim to show in each case the integral

$$\int\limits_{\substack{\varepsilon<|\eta-\tau|<\frac{1}{\varepsilon}\\ \widehat{\eta}_j<\eta<\check{\eta}_j}} e^{iz\psi_j(\eta)}\,\frac{a(\psi_j(\eta))}{\eta-\tau}\,\mathrm{d}\eta \tag{3.A.11}$$

can be estimated uniformly in τ, z, and ε, by use of the hypotheses (H1)–(H5). After the following case distinction we shall discuss each case, except the trivial case (a), in detail.

(i) **The continuous case** $(\tau - r, \tau + r) \cap (\widehat{\eta}_j, \check{\eta}_j) = \varnothing$.

  (a) **The continuous bounded case** $j \notin \{0, N\}$. Here we integrate a continuous function over a bounded domain. This case is trivial, since for any finite interval I with $|\eta - \tau|\big|_I > r$ we have

$$\left|\int_I e^{iz\psi_j(\eta)}\,\frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}}a(\psi_j(\eta))}{\eta-\tau}\,\mathrm{d}\eta\right| \le \frac{A}{r}|I|\,, \tag{3.A.12}$$

  where the right-hand side is uniform in z, τ, ε.

  (b) **The continuous oscillatory tail case** $j \in \{0, N\}$. Here we have to combine the oscillation with the $\frac{1}{\eta}$ decay to bound the integral. By using (a), we can furthermore assume without loss of generality that the integral ranges over the intervals (R, ∞) or $(-\infty, -R)$, where $R = R(\varphi) > 0$ is arbitrarily large.

(ii) **The non-critical point case** $(\tau - r, \tau + r) \subset (\widehat{\eta}_j, \check{\eta}_j)$.

  (c) **The non-critical point singularity bounded case** $j \notin \{0, N\}$. Here we have to use oscillation as well as cancellation to bound the integral.

  (d) **The non-critical point singularity oscillatory tail case** $j \in \{0, N\}$. Here the singularity is possibly far away in the oscillatory tail. This can be reduced to a combination of (a), (b) and (c), but it requires an improvement of (b) as the oscillation in the integral may become weak far away from the origin.

(iii) **The critical point case** $(\tau - r, \tau + r) \cap (\widehat{\eta}_j, \check{\eta}_j) \neq \varnothing$ and either $\eta_j \in (\tau - r, \tau + r)$ or $\eta_{j+1} \in (\tau - r, \tau + r)$. This case always comes in pairs, meaning that if $\eta_j \in (\tau - r, \tau + r)$, then $j-1$ is also of the critical point case: $(\tau - r, \tau + r) \cap (\widehat{\eta}_{j-1}, \check{\eta}_{j-1}) \neq \varnothing$, and if $\eta_{j+1} \in (\tau - r, \tau + r)$ then $j + 1$ is also of the critical point case: $(\tau - r, \tau + r) \cap (\widehat{\eta}_{j+1}, \check{\eta}_{j+1}) \neq \varnothing$.

  We assume without loss of generality that $\eta_j \in (\tau - r, \tau + r)$ and furthermore that $\sigma_{j-1} = 1$. Then $\sigma_j = -1$ if and only if ξ_j is a strict local maximum. If $\sigma_j = -1$ we decompose

$$\begin{aligned}
&\int\limits_{\substack{\varepsilon<|\eta-\tau|<\frac{1}{\varepsilon}\\ \widehat{\eta}_{j-1}<\eta<\check{\eta}_{j-1}}} e^{iz\psi_{j-1}(\eta)}\,\frac{a(\psi_{j-1}(\eta))}{\eta-\tau}\,\mathrm{d}\eta - \int\limits_{\substack{\varepsilon<|\eta-\tau|<\frac{1}{\varepsilon}\\ \widehat{\eta}_j<\eta<\check{\eta}_j}} e^{iz\psi_j(\eta)}\,\frac{a(\psi_j(\eta))}{\eta-\tau}\,\mathrm{d}\eta \\
&= \int_{\eta_{j-1}}^{\tau-r} e^{iz\psi_{j-1}(\eta)}\,\frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}}a(\psi_{j-1}(\eta))}{\eta-\tau}\,\mathrm{d}\eta - \int_{\eta_{j+1}}^{\tau-r} e^{iz\psi_j(\eta)}\,\frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}}a(\psi_j(\eta))}{\eta-\tau}\,\mathrm{d}\eta
\end{aligned}$$

$$+\int_{\tau-r}^{\eta_j} e^{iz\psi_{j-1}(\eta)}\frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}}a(\psi_{j-1}(\eta))}{\eta-\tau}\,\mathrm{d}\eta-\int_{\tau-r}^{\eta_j} e^{iz\psi_j(\eta)}\frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}}a(\psi_j(\eta))}{\eta-\tau}\,\mathrm{d}\eta\,,$$

and if $\sigma_j=1$ we decompose

$$\begin{aligned}&\int\limits_{\substack{\varepsilon<|\eta-\tau|<\frac{1}{\varepsilon}\\ \widehat{\eta}_{j-1}<\eta<\check{\eta}_{j-1}}} e^{iz\psi_{j-1}(\eta)}\frac{a(\psi_{j-1}(\eta))}{\eta-\tau}\,\mathrm{d}\eta+\int\limits_{\substack{\varepsilon<|\eta-\tau|<\frac{1}{\varepsilon}\\ \widehat{\eta}_{j}<\eta<\check{\eta}_{j}}} e^{iz\psi_{j}(\eta)}\frac{a(\psi_{j}(\eta))}{\eta-\tau}\,\mathrm{d}\eta\\ &=\int_{\eta_{j-1}}^{\tau-r} e^{iz\psi_{j-1}(\eta)}\frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}}a(\psi_{j-1}(\eta))}{\eta-\tau}\,\mathrm{d}\eta+\int_{\tau+r}^{\eta_{j+1}} e^{iz\psi_j(\eta)}\frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}}a(\psi_j(\eta))}{\eta-\tau}\,\mathrm{d}\eta\\ &+\int_{\tau-r}^{\eta_j} e^{iz\psi_{j-1}(\eta)}\frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}}a(\psi_{j-1}(\eta))}{\eta-\tau}\,\mathrm{d}\eta+\int_{\eta_j}^{\tau+r} e^{iz\psi_j(\eta)}\frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}}a(\psi_j(\eta))}{\eta-\tau}\,\mathrm{d}\eta\,.\end{aligned}$$

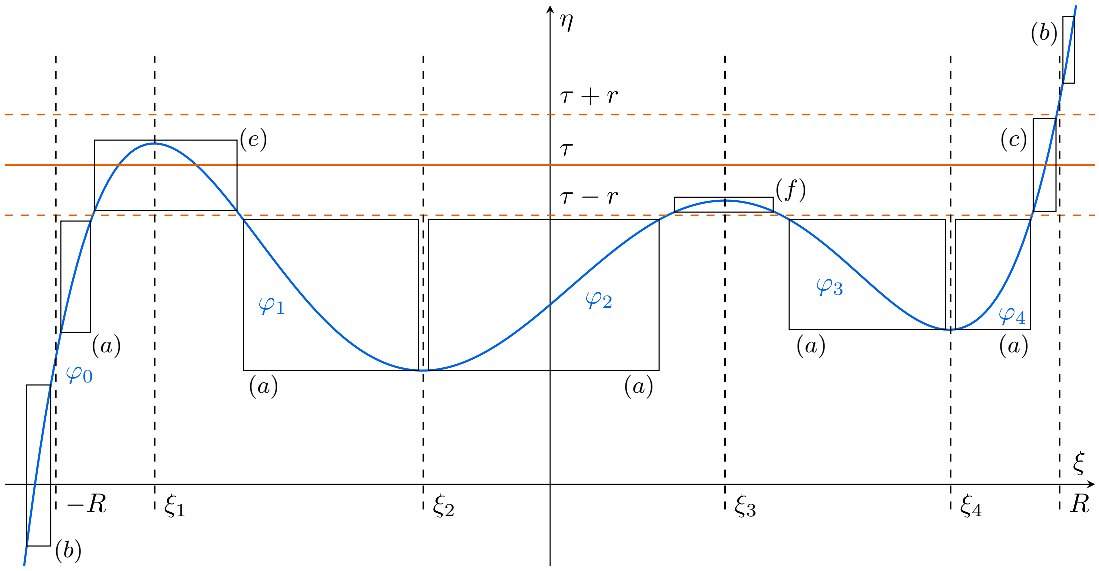
In both cases the integrals in the first line have $|\eta-\tau|\geq r$ and can hence be treated with either (a) or (b), depending on whether the integral is over a bounded or unbounded domain. The remaining integrals need further case distinction.

e) **The singularity near critical point case** $\sigma_j=-1$ and $\tau<\eta_j$, or $\sigma_j=1$. Here the integral has one or two singularities in close proximity to the critical point.

f) **The almost singularity near critical point case** $\sigma_j=-1$ and $\tau>\eta_j$. Here the integral has no singularity, but can be arbitrarily close to being singular. This case is easier than the previous one.

g) **The singularity at critical point case** $\tau=\eta_j$. This case is strictly more difficult to deal with than certain sections of the integrals to estimate in the previous cases, so it serves as the prototypical case.

Figure 3.2.: An example for φ and some of the aforementioned cases.



We now enumerate all nontrivial cases and perform the necessary estimates, assuming here without loss of generality that $z\geq 0$.

(b) The continuous oscillatory tail case. For high frequencies it suffices to estimate each tail individually, but for low frequencies there is little oscillation, and instead cancellation between the two tails has to be exploited.

We assume without loss of generality that $\sigma_0=-1$ and distinguish the cases $\sigma_N=1$ and $\sigma_N=-1$. If $\sigma_N=1$, then we have to bound the quantity

$$-\int_R^\infty e^{iz\psi_0(\eta)}\frac{a(\psi_0(\eta))}{\eta-\tau}\,\mathrm{d}\eta+\int_R^\infty e^{iz\psi_N(\eta)}\frac{a(\psi_N(\eta))}{\eta-\tau}\,\mathrm{d}\eta\,.$$

On the other hand, if $\sigma_N = -1$ we have to bound the quantity

$$-\int_R^\infty e^{iz\psi_0(\eta)} \frac{a(\psi_0(\eta))}{\eta - \tau} \,\mathrm{d}\eta - \int_{-\infty}^{-R} e^{iz\psi_N(\eta)} \frac{a(\psi_N(\eta))}{\eta - \tau} \,\mathrm{d}\eta \,.$$

That is, we aim to estimate the following quantity

$$\int_R^\infty e^{iz\psi_0(\eta)} \frac{a(\psi_0(\eta))}{\eta - \tau} \,\mathrm{d}\eta - \int_R^\infty e^{iz\psi_N(\sigma_N\eta)} \frac{a(\psi_N(\sigma_N\eta))}{\eta - \sigma_N\tau} \,\mathrm{d}\eta \,.$$

Since we are in the case where there is no singularity in the oscillatory tail, we have $\tau < R - r$ and $\sigma_N\tau < R - r$. Then for all $\eta > R$ we have

$$\left|\frac{1}{\eta - \tau}\right| \le \left(1 + \mathbb{1}_{\{\tau \ge 0\}}\left|\frac{1}{\frac{\eta}{\tau} - 1}\right|\right)\frac{1}{\eta} \le \left(1 + \frac{1}{\frac{R}{R-r} - 1}\right)\frac{1}{\eta} \lesssim_{R,r} \frac{1}{\eta} \tag{3.A.13}$$

$$\left|\frac{1}{\eta - \sigma_N\tau}\right| \le \left(1 + \mathbb{1}_{\{\sigma_N\tau \ge 0\}}\left|\frac{1}{\frac{\eta}{\sigma_N\tau} - 1}\right|\right)\frac{1}{\eta} \le \left(1 + \frac{1}{\frac{R}{R-r} - 1}\right)\frac{1}{\eta} \lesssim_{R,r} \frac{1}{\eta} \,. \tag{3.A.14}$$

We decompose

$$\begin{aligned}
&\int_R^\infty e^{iz\psi_0(\eta)} \frac{a(\psi_0(\eta))}{\eta - \tau} \,\mathrm{d}\eta - \int_R^\infty e^{iz\psi_N(\sigma_N\eta)} \frac{a(\psi_N(\sigma_N\eta))}{\eta - \sigma_N\tau} \,\mathrm{d}\eta \\
&= \int_R^\infty e^{iz\psi_N(\sigma_N\eta)} \left(\frac{a(\psi_0(\eta))}{\eta - \tau} - \frac{a(\psi_N(\sigma_N\eta))}{\eta - \sigma_N\tau}\right) \mathrm{d}\eta + \int_R^\infty \left(e^{iz\psi_0(\eta)} - e^{iz\psi_N(\sigma_N\eta)}\right) \frac{a(\psi_0(\eta))}{\eta - \tau} \,\mathrm{d}\eta \\
&= (I) + (II)
\end{aligned}$$

and directly estimate

$$\begin{aligned}
|(I)| &\lesssim_A \int_R^\infty \left|\frac{\sigma_N\tau - \tau}{(\eta - \tau)(\eta - \sigma_N\tau)}\right| + \left|\frac{a(\psi_0(\eta)) - a(\psi_N(\sigma_N\eta))}{(\eta - \sigma_N\tau)}\right| \mathrm{d}\eta \\
&\lesssim_{A,R,P,r} \int_R^\infty \frac{1}{\eta^2} + \frac{1}{\eta^{1+\delta}} \,\mathrm{d}\eta \lesssim_{R,\beta,\delta} 1 \,.
\end{aligned}$$

If $z = 0$ then $(II) = 0$, so we assume $z > 0$ (we have assumed without loss of generality $z \ge 0$). We split the interval (R, ∞) into $(R, R \vee z^{-\frac{1}{\beta}})$ and $(R \vee z^{-\frac{1}{\beta}}, \infty)$. For the finite interval, we estimate

$$\begin{aligned}
\left|\int_R^{R\vee z^{-\frac{1}{\beta}}} \left(e^{iz\psi_0(\eta)} - e^{iz\psi_N(\sigma_N\eta)}\right) \frac{a(\psi_0(\eta))}{\eta - \tau} \,\mathrm{d}\eta\right| &\lesssim_A z \int_R^{R\vee z^{-\frac{1}{\beta}}} \frac{|\psi_0(\eta)| + |\psi_N(\sigma_N\eta)|}{\eta - \tau} \,\mathrm{d}\eta \\
&\lesssim_{R,P,r} z \int_R^{R\vee z^{-\frac{1}{\beta}}} \frac{\eta^\beta}{\eta} \,\mathrm{d}\eta \\
&= z\frac{1}{\beta}\left((R \vee z^{-\frac{1}{\beta}})^\beta - R^\beta\right) \lesssim_\beta 1 \,.
\end{aligned}$$

We decompose the other interval into dyadic blocks

$$[R \vee z^{-\frac{1}{\beta}}, \infty) = \bigcup_{l=L}^\infty [2^l, 2^{l+1}) \cup [R \vee z^{-\frac{1}{\beta}}, 2^L) \,,$$

where $L \in \mathbb{N}$ is the unique integer such that $2^{L-1} \le R \vee z^{-\frac{1}{\beta}} < 2^L$. Now we use the oscillation with Lemma 3.A.6 to estimate

$$\begin{aligned}
&\sum_{l=L}^\infty \left|\int_{2^l}^{2^{l+1}} \left(e^{iz\psi_0(\eta)} - e^{iz\psi_N(\sigma_N\eta)}\right) \frac{a(\psi_0(\eta))}{\eta - \tau} \,\mathrm{d}\eta\right| \\
&+ \left|\int_{R\vee z^{-\frac{1}{\beta}}}^{2^L} \left(e^{iz\psi_0(\eta)} - e^{iz\psi_N(\sigma_N\eta)}\right) \frac{a(\psi_0(\eta))}{\eta - \tau} \,\mathrm{d}\eta\right|
\end{aligned}$$

$$
\begin{aligned}
&\lesssim_{M,R,A,r,p} \frac{1}{z}\left(\sum_{l=L}^{\infty} 2^{(l+1)(1-\beta)}2^{-l} + 2^{L(1-\beta)}(R\vee z^{-\frac{1}{\beta}})^{-1}\right)\\
&\lesssim \left(\frac{1}{z}\sum_{l=L}^{\infty} 2^{-\beta l} + z^{-1}2^{(1-\beta)L}(R\vee z^{-\frac{1}{\beta}})^{-1}\right)\\
&\lesssim \left(\left(2^L(R\vee z^{-\frac{1}{\beta}})^{-1}\right)^{-\beta} + \left(2^L(R\vee z^{-\frac{1}{\beta}})^{-1}\right)^{1-\beta}\right) \lesssim 1\,.
\end{aligned}
$$

Here (H5) implies the monotonicity condition that the lemma requires.

(c) The non-critical point singularity bounded case. Due to (3.A.12) it suffices to estimate the integral (3.A.11) on $(\tau - r, \tau + r)$ instead of the whole interval $(\widehat{\eta}_j, \check{\eta}_j)$. In fact, the interval $(\tau - r', \tau + r')$ where $r' = \frac{r}{2}$ suffices, since the case (iii) takes care of the integral near a critical point. We have ensured that we are not close to a critical point and hence $\sup_{\eta\in(\tau-r',\tau+r')} |\psi_j'(\eta)|$ is bounded uniformly for small τ, i. e. $|\tau| \leq R$. On the other hand, (H3) ensures boundedness for $|\tau| > R$. Another control quantity is $\sup_{\eta\in(\tau-r',\tau+r')} \frac{1}{|\psi_j'(\eta)|}$, which is bounded uniformly in τ for $|\tau| \leq R$ only.

We have

$$
\int_{\tau-r'}^{\tau+r'} e^{iz\psi_j(\eta)} \frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}} a(\psi_j(\eta))}{\eta-\tau}\,\mathrm{d}\eta = \int_{\varepsilon}^{r'} \left(e^{iz\psi_j(\tau+\eta)}a(\psi_j(\tau+\eta)) - e^{iz\psi_j(\tau-\eta)}a(\psi_j(\tau-\eta))\right)\frac{1}{\eta}\,\mathrm{d}\eta\,.
$$

The Lipschitz property yields

$$
\left|\int_{\varepsilon}^{r'\wedge\frac{1}{z}} \left(e^{iz\psi_j(\tau+\eta)}a(\psi_j(\tau+\eta)) - e^{iz\psi_j(\tau-\eta)}a(\psi_j(\tau-\eta))\right)\frac{1}{\eta}\,\mathrm{d}\eta\right| \lesssim_{A,R,\beta} \sup_{\eta\in(\tau-r',\tau+r')} |\psi_j'(\eta)|\,.
$$

If $zr' \leq 1$ then we are done, so we assume $zr' > 1$ and aim to estimate

$$
\int_{\frac{1}{z}}^{r'} \left(e^{iz\psi_j(\tau+\eta)}a(\psi_j(\tau+\eta)) - e^{iz\psi_j(\tau-\eta)}a(\psi_j(\tau-\eta))\right)\frac{1}{\eta}\,\mathrm{d}\eta\,.
$$

Here we apply Lemma 3.A.6 to each summand, using that ψ_j is strictly monotone and that $\frac{a(\psi_j(\tau\pm\eta))}{\eta}$ only changes monotonicity less than M times by (H5). This yields

$$
\left|\int_{\frac{1}{z}}^{r'} e^{iz\psi_j(\tau\pm\eta)}a(\psi_j(\tau\pm\eta))\frac{1}{\eta}\,\mathrm{d}\eta\right| \lesssim_{M,A} \sup_{\eta\in(\tau-r',\tau+r')} \frac{1}{|\psi_j'(\eta)|}\,.
$$

Note that the right-hand side may grow with τ. This is acceptable as we are in the case $j \notin \{0, N\}$, i. e. there exists some fixed large R such that $|\tau| < R$, but it needs to be improved for the case (d) below.

(d) The non-critical point singularity oscillatory tail case. In this case the singularity is far from the origin in the oscillatory tail, i. e. $|\tau| > R$. Note that now $j \in \{0, N\}$. We assume without loss of generality that $\tau > \frac{R}{1-r} > R$, and correspondingly that $\sigma_j = -1$ if $j = 0$, and $\sigma_j = 1$ if $j = N$, as otherwise there is no singularity in the oscillatory tail. We would like to separate the singularity from the oscillatory tail. This requires us to revisit the proofs of (b) and (c).

Note that in (c) we have treated the singularity on a smaller interval $(\tau-r', \tau+r')$. Correspondingly, we now estimate the integral (3.A.11) on a larger interval $(\tau-\tau r, \tau+\tau r) \subset (R, \infty)$. We decompose

$$
\begin{aligned}
&\int_{\varepsilon}^{\tau r} \left(e^{iz\psi_j(\tau+\eta)}a(\psi_j(\tau+\eta)) - e^{iz\psi_j(\tau-\eta)}a(\psi_j(\tau-\eta))\right)\frac{1}{\eta}\,\mathrm{d}\eta\\
&= \int_{\varepsilon}^{\tau r} e^{iz\psi_j(\tau+\eta)}(a(\psi_j(\tau+\eta)) - a(\psi_j(\tau-\eta)))\frac{1}{\eta}\,\mathrm{d}\eta + \int_{\varepsilon}^{\tau r} \left(e^{iz\psi_j(\tau+\eta)} - e^{iz\psi_j(\tau-\eta)}\right)a(\psi_j(\tau-\eta))\frac{1}{\eta}\,\mathrm{d}\eta\,.
\end{aligned}
$$

The Lipschitz property yields

$$
\left|\int_{\varepsilon}^{\tau r} e^{iz\psi_j(\tau+\eta)}(a(\psi_j(\tau+\eta)) - a(\psi_j(\tau-\eta)))\frac{1}{\eta}\,\mathrm{d}\eta\right| \lesssim_{A,P} \tau r \sup_{\eta\in(\tau-\tau r,\tau+\tau r)} (\eta^\beta)^{-1}\eta^{\beta-1} \lesssim_R 1
$$

and

$$\left| \int_{\varepsilon}^{\tau r \wedge \frac{\tau^{1-\beta}}{z}} \left(e^{iz\psi_j(\tau+\eta)} - e^{iz\psi_j(\tau-\eta)} \right) a(\psi_j(\tau-\eta)) \frac{1}{\eta} \, \mathrm{d}\eta \right| \lesssim_{A,P} z \frac{\tau^{1-\beta}}{z} \sup_{\eta \in (\tau - \tau r, \tau + \tau r)} \eta^{\beta - 1} \lesssim 1 \,.$$

If $\frac{\tau^{1-\beta}}{z} < \tau r$ then there is a remainder, which we estimate using the oscillation, i. e. Lemma 3.A.6:

$$\begin{aligned} \left| \int_{\frac{\tau^{1-\beta}}{z}}^{\tau r} \left(e^{iz\psi_j(\tau+\eta)} - e^{iz\psi_j(\tau-\eta)} \right) a(\psi_j(\tau-\eta)) \frac{1}{\eta} \, \mathrm{d}\eta \right| &\lesssim_{M,A,P} \frac{1}{z} \sup_{\eta \in (\tau - \tau r, \tau + \tau r)} \eta^{1-\beta} \frac{z}{\tau^{1-\beta}} \\ &\lesssim 1 \,. \end{aligned}$$

As before, (H5) supplies the required monotonicity assumption.

We now modify the proof of (b) to deal with the remaining oscillatory tail in the integral (3.A.11) with $j \in \{0, N\}$, where the interval $(\tau - \tau r, \tau + \tau r)$ around the singularity has been removed by the above consideration. We again assume that $\sigma_0 = -1$ and distinguish two cases. If $\sigma_N = 1$ we have to bound the quantity

$$\begin{aligned} &\int_R^{\tau - \tau r} e^{iz\psi_0(\eta)} \frac{a(\psi_0(\eta))}{\eta - \tau} \, \mathrm{d}\eta - \int_R^{\tau - \tau r} e^{iz\psi_N(\eta)} \frac{a(\psi_N(\eta))}{\eta - \tau} \, \mathrm{d}\eta \\ &+ \int_{\tau + \tau r}^{\infty} e^{iz\psi_0(\eta)} \frac{a(\psi_0(\eta))}{\eta - \tau} \, \mathrm{d}\eta - \int_{\tau + \tau r}^{\infty} e^{iz\psi_N(\eta)} \frac{a(\psi_N(\eta))}{\eta - \tau} \, \mathrm{d}\eta \,, \end{aligned}$$

while if $\sigma_N = -1$ we have to bound the quantity

$$\begin{aligned} &\int_R^{\tau - \tau r} e^{iz\psi_0(\eta)} \frac{a(\psi_0(\eta))}{\eta - \tau} \, \mathrm{d}\eta + \int_{-\tau + \tau r}^{-R} e^{iz\psi_N(\eta)} \frac{a(\psi_N(\eta))}{\eta - \tau} \, \mathrm{d}\eta \\ &+ \int_{\tau + \tau r}^{\infty} e^{iz\psi_0(\eta)} \frac{a(\psi_0(\eta))}{\eta - \tau} \, \mathrm{d}\eta + \int_{-\infty}^{-\tau - \tau r} e^{iz\psi_N(\eta)} \frac{a(\psi_N(\eta))}{\eta - \tau} \, \mathrm{d}\eta \\ &+ \int_{-\tau - \tau r}^{-\tau + \tau r} e^{iz\psi_N(\eta)} \frac{a(\psi_N(\eta))}{\eta - \tau} \, \mathrm{d}\eta \,. \end{aligned}$$

We can directly estimate all the above integrals which have finite domains as in (3.A.12), as the size of the domain is always bounded by $\tau + \tau r$, while the integrand is bounded by $\frac{A}{\tau r}$. Unifying the remaining integrals for $\sigma_N \in \{-1, 1\}$ as in the treatment of (b) before, it remains to bound the quantity

$$\begin{aligned} &\int_{\tau + \tau r}^{\infty} e^{iz\psi_0(\eta)} \frac{a(\psi_0(\eta))}{\eta - \tau} \, \mathrm{d}\eta - \int_{\tau + \tau r}^{\infty} e^{iz\psi_N(\sigma_N \eta)} \frac{a(\psi_N(\sigma_N \eta))}{\eta - \sigma_N \tau} \, \mathrm{d}\eta \\ &= \int_{\tau + \tau r}^{\infty} e^{iz\psi_N(\sigma_N \eta)} \left(\frac{a(\psi_0(\eta))}{\eta - \tau} - \frac{a(\psi_N(\sigma_N \eta))}{\eta - \sigma_N \tau} \right) \mathrm{d}\eta + \int_{\tau + \tau r}^{\infty} \left(e^{iz\psi_0(\eta)} - e^{iz\psi_N(\sigma_N \eta)} \right) \frac{a(\psi_0(\eta))}{\eta - \tau} \, \mathrm{d}\eta \\ &= (I) + (II) \,. \end{aligned}$$

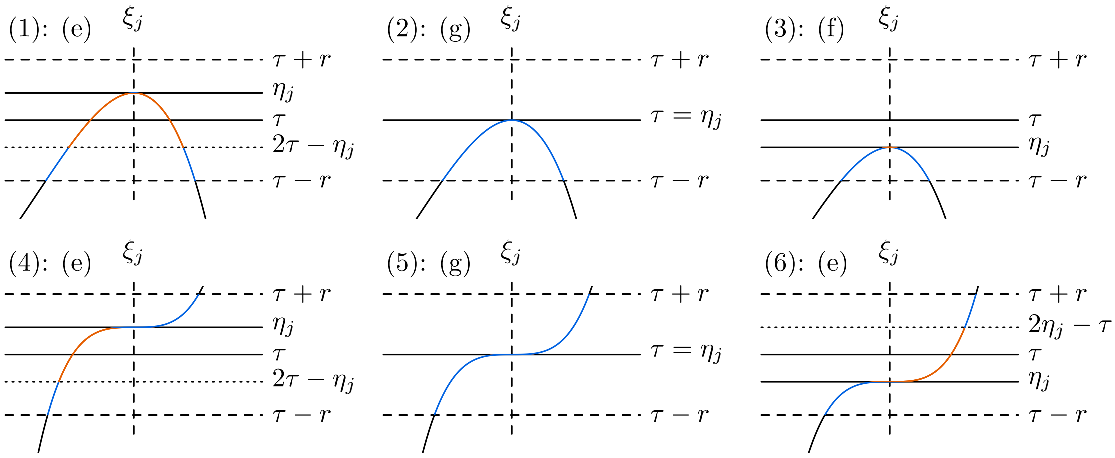
This requires adapting (3.A.13)–(3.A.14). For all $\eta \in (\tau + \tau r, \infty)$ we have

$$\left| \frac{1}{\eta - \tau} \right| \leq \left(1 + \left| \frac{1}{\frac{\eta}{\tau} - 1} \right| \right) \frac{1}{\eta} \leq \left(1 + \left| \frac{1}{\frac{\tau + \tau r}{\tau} - 1} \right| \right) \frac{1}{\eta} \leq \left(1 + \frac{1}{r} \right) \frac{1}{\eta} \qquad \left| \frac{1}{\eta + \tau} \right| \leq \frac{1}{\eta} \,.$$

Both terms (I) and (II) can now be estimated with the same method as in (b).

(iii) The critical point case Outside the interval $(\tau - r, \tau + r)$ we have $|\eta - \tau| \geq r$, so the integrals over these sections can be estimated as in (3.A.12) if the interval is finite and (b) if not. It therefore suffices to consider the integral over $(\tau - r, \tau + r)$. Here we shall describe our approach to the cases (e), (f) and (g) with the help of Figure 3.3. We assume without loss of generality that $\sigma_{j-1} = 1$.

Figure 3.3.: Six possibilities for the cases (e), (f) and (g) when $\sigma_{j-1} = 1$.



We refer to different parts of the integral to estimate as "sections". The orange sections contain the singularity, so cancellation of the sections before and after the singularity is crucial. Note that in case (1) cancellation between the orange section before the point (ξ_j, η_j) and the one after (ξ_j, η_j) is not necessary. We perform the proof only for the section to the right of (ξ_j, η_j) in (1), as it is representative of all other orange sections. All blue sections need to be considered in pairs consisting of a section to the left of (ξ_j, η_j) and one to the right, in order to exploit a cancellation of the form $|e^{ix} - e^{iy}| \le |x| + |y|$. Here the cases (2) and (5) are strictly more difficult than the other ones, because the sections we integrate over go all the way up to the singularity. We treat them as representative of all the other blue sections.

Define $b = \eta_j - \tau$. In the case (1) we have $\tau - r < \tau - b < \tau < \tau + b = \eta_j < \tau + r$, and the integral to estimate is

$$\int_{\tau-b}^{\tau+b} e^{iz\psi_{j-1}(\eta)} \frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}} a(\psi_{j-1}(\eta))}{\eta - \tau} \,\mathrm{d}\eta - \int_{\tau-b}^{\tau+b} e^{iz\psi_j(\eta)} \frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}} a(\psi_j(\eta))}{\eta - \tau} \,\mathrm{d}\eta$$

As mentioned before, we estimate these two terms individually and with the same technique, so it suffices to consider the case of ψ_j. Here

$$\begin{aligned}
&\int_{\tau-b}^{\tau+b} e^{iz\psi_j(\eta)} \frac{\mathbb{1}_{\{|\eta-\tau|>\varepsilon\}} a(\psi_j(\eta))}{\eta - \tau} \,\mathrm{d}\eta \\
&= \int_{\varepsilon}^{b} \Big(e^{iz\psi_j(\tau+\eta)} a(\psi_j(\tau+\eta)) - e^{iz\psi_j(\tau-\eta)} a(\psi_j(\tau-\eta)) \Big) \frac{1}{\eta} \,\mathrm{d}\eta \\
&= \int_{\varepsilon}^{b} \Big(e^{iz\psi_j(\eta_j-(b-\eta))} a(\psi_j(\eta_j-(b-\eta))) - e^{iz\psi_j(\eta_j-(b+\eta))} a(\psi_j(\eta_j-(b+\eta))) \Big) \frac{1}{\eta} \,\mathrm{d}\eta \,.
\end{aligned}$$

To simplify notation, we assume without loss of generality that $\eta_j = 0$ and replace $\psi_j(\eta)$ by $\psi_j(-\eta)$. This means that $\psi_j(\eta_j - (b \pm \eta))$ is replaced by $\psi_j(b \pm \eta)$. We decompose

$$\begin{aligned}
&\int_{\varepsilon}^{b} \frac{e^{iz\psi_j(b+\eta)} a(\psi_j(b+\eta)) - e^{iz\psi_j(b-\eta)} a(\psi_j(b-\eta))}{\eta} \,\mathrm{d}\eta \\
&= \int_{\varepsilon}^{b} e^{iz\psi_j(b-\eta)} \frac{a(\psi_j(b+\eta)) - a(\psi_j(b-\eta))}{\eta} \,\mathrm{d}\eta + \int_{\varepsilon}^{b} \frac{e^{iz\psi_j(b+\eta)} - e^{iz\psi_j(b-\eta)}}{\eta} a(\psi_j(b+\eta)) \,\mathrm{d}\eta \,.
\end{aligned}$$

The first term can be directly estimated:

$$\begin{aligned}
&\int_{\varepsilon}^{b} \frac{|a(\psi_j(b+\eta)) - a(\psi_j(b-\eta))|}{\eta} \,\mathrm{d}\eta \le \int_0^b \frac{1}{\eta} \int_{-\eta}^{\eta} |a'(\psi_j(b+y))\psi_j'(b+y)| \,\mathrm{d}y \,\mathrm{d}\eta \\
&= \int_{-b}^{b} |a'(\psi_j(b+y))\psi_j'(b+y)| \int_{|y|}^{b} \frac{1}{\eta} \,\mathrm{d}\eta \,\mathrm{d}y \lesssim_{A,Q} \int_{-b}^{b} (b+y)^{\alpha-1} \ln\Big(\frac{b}{|y|}\Big) \,\mathrm{d}y
\end{aligned}$$

$$\leq b^\alpha \int_0^2 \frac{1}{y^{1-\alpha}} \ln\left(\frac{1}{|y-1|}\right) \mathrm{d}y \lesssim_{R,\alpha} 1\,.$$

To estimate the second term, we first suppose that $\frac{b^{1-\alpha}}{z} < b$. Then (H5) permits application of Lemma 3.A.6, which yields

$$\begin{aligned}\left|\int_{\frac{b^{1-\alpha}}{z}}^{b} e^{iz\psi_j(b\pm\eta)} \frac{a(\psi_j(b\pm\eta))}{\eta} \mathrm{d}\eta\right| &\lesssim_{M,A,q} \frac{1}{z} \sup_{\eta\in\left(\frac{b^{1-\alpha}}{z},b\right)} (b\pm\eta)^{1-\alpha} \sup_{\eta\in\left(\frac{b^{1-\alpha}}{z},b\right)} \frac{1}{\eta}\\ &\leq \frac{1}{z}(2b)^{1-\alpha}\frac{z}{b^{1-\alpha}} \lesssim 1\,.\end{aligned}$$

The remaining interval is $\left(\varepsilon, \varepsilon \vee \frac{b^{1-\alpha}}{z}\right)$, so we assume $\varepsilon < \frac{b^{1-\alpha}}{z}$ and estimate

$$\begin{aligned}\left|\int_\varepsilon^{\frac{b^{1-\alpha}}{z}} \frac{e^{iz\psi_j(b+\eta)} - e^{iz\psi_j(b-\eta)}}{\eta} a(\psi_j(b+\eta))\,\mathrm{d}\eta\right| &= \left|\int_\varepsilon^{\frac{b^{1-\alpha}}{z}} \frac{a(\psi_j(b+\eta))}{\eta}\int_{-\eta}^{\eta} e^{iz\psi_j(b+y)} iz\psi_j'(b+y)\,\mathrm{d}y\,\mathrm{d}\eta\right|\\ &= \left|\int_{-\frac{b^{1-\alpha}}{z}}^{\frac{b^{1-\alpha}}{z}} e^{iz\psi_j(b+y)} iz\psi_j'(b+y) \int_{\varepsilon\vee|y|}^{\frac{b^{1-\alpha}}{z}} \frac{a(\psi_j(b+\eta))}{\eta}\,\mathrm{d}\eta\,\mathrm{d}y\right|\\ &\lesssim_{A,Q} \frac{z}{b^{1-\alpha}}\int_{-\frac{b^{1-\alpha}}{z}}^{\frac{b^{1-\alpha}}{z}} \frac{b^{1-\alpha}}{(b+y)^{1-\alpha}}\left|\log\left(\frac{b^{1-\alpha}}{z}\frac{1}{|y|}\right)\right|\mathrm{d}y\\ &= \int_{-1}^{1}\frac{1}{\left(1+\frac{t}{b^\alpha z}\right)^{1-\alpha}}\left|\log\left(\frac{1}{|t|}\right)\right|\mathrm{d}t \lesssim_\alpha 1\,.\end{aligned}$$

Note that at the end we have used $b^\alpha z > 1$. It remains to deal with the case $b^\alpha z \leq 1$:

$$\begin{aligned}\left|\int_\varepsilon^{b} \frac{e^{iz\psi_j(b+\eta)} - e^{iz\psi_j(b-\eta)}}{\eta} a(\psi_j(b+\eta))\,\mathrm{d}\eta\right| &= \left|\int_{-b}^{b} e^{iz\psi_j(b+y)} iz\psi_j'(b+y)\int_{\varepsilon\vee|y|}^{b}\frac{a(\psi_j(b+\eta))}{\eta}\,\mathrm{d}\eta\,\mathrm{d}y\right|\\ &\lesssim_{A,Q} \frac{1}{b}\int_{-b}^{b}\frac{b^{1-\alpha}}{(b+y)^{1-\alpha}}\left|\log\left(\frac{b}{|y|}\right)\right|\mathrm{d}y \lesssim_\alpha 1\,.\end{aligned}$$

In the cases (2) and (5) we have $\tau - r < \tau = \eta_j < \tau + r$ and $\sigma_{j-1} = 1$ while $\sigma_j \in \{-1, 1\}$. The integral to estimate can be written for both cases as

$$\begin{aligned}&\int_{\tau-r}^{\tau-\varepsilon} e^{iz\psi_{j-1}(\eta)}\frac{a(\psi_{j-1}(\eta))}{\eta-\tau}\,\mathrm{d}\eta + \int_{\tau+\sigma_j\varepsilon}^{\tau+\sigma_j r} e^{iz\psi_j(\eta)}\frac{a(\psi_j(\eta))}{\eta-\tau}\,\mathrm{d}\eta\\ &= \int_\varepsilon^r\left(e^{iz\psi_j(\eta_j+\sigma_j\eta)} - e^{iz\psi_{j-1}(\eta_j-\eta)}\right)\frac{a(\psi_j(\eta_j+\sigma_j\eta))}{\eta}\,\mathrm{d}\eta\\ &+ \int_\varepsilon^r e^{iz\psi_{j-1}(\eta_j-\eta)}\frac{a(\psi_j(\eta_j+\sigma_j\eta)) - a(\psi_{j-1}(\eta_j-\eta))}{\eta}\,\mathrm{d}\eta\\ &= (III) + (IV)\,.\end{aligned}$$

We have

$$|(IV)| \lesssim_A \int_\varepsilon^r \frac{|\psi_j(\eta_j+\sigma_j\eta) - \xi_j + \xi_j - \psi_{j-1}(\eta_j-\eta)|}{\eta}\,\mathrm{d}\eta \lesssim_Q \int_0^r \frac{\eta^\alpha}{\eta}\,\mathrm{d}\eta \lesssim_\alpha 1\,.$$

It remains to estimate (III). Assuming without loss of generality that $\varepsilon < z^{-\frac{1}{\alpha}}$, we can immediately estimate

$$\begin{aligned}\left|\int_\varepsilon^{z^{-\frac{1}{\alpha}}}\left(e^{iz\psi_j(\eta_j+\sigma_j\eta)} - e^{iz\psi_{j-1}(\eta_j-\eta)}\right)\frac{a(\psi_j(\eta_j+\sigma_j\eta))}{\eta}\,\mathrm{d}\eta\right| &\lesssim_{A,Q} z\int_0^{z^{-\frac{1}{\alpha}}} \eta^\alpha\frac{1}{\eta}\,\mathrm{d}\eta\\ &\lesssim_\alpha z(z^{-\frac{1}{\alpha}})^\alpha = 1\,.\end{aligned}$$

If $z^{-\frac{1}{\alpha}} \geq r$ then this is the whole integral, so we assume $z^{-\frac{1}{\alpha}} < r$ and decompose

$$[z^{-\frac{1}{\alpha}}, r) = \bigcup_{l=0}^{L-1} [r2^{-l-1}, r2^{-l}) \cup [z^{-\frac{1}{\alpha}}, r2^{-L}),$$

where $L \in \mathbb{N}$ is the unique integer such that $2^{-L-1} \leq \frac{z^{-\frac{1}{\alpha}}}{r} < 2^{-L}$. With (H5) we can use the oscillation through Lemma 3.A.6 to estimate

$$\begin{aligned}
&\sum_{l=0}^{L-1} \left| \int_{r2^{-l-1}}^{r2^{-l}} \left(e^{iz\psi_j(\eta_j+\sigma_j\eta)} - e^{iz\psi_{j-1}(\eta_j-\eta)} \right) \frac{a(\psi_j(\eta_j+\sigma_j\eta))}{\eta} \, \mathrm{d}\eta \right| \\
&+ \left| \int_{z^{-\frac{1}{\alpha}}}^{r2^{-L}} \left(e^{iz\psi_j(\eta_j+\sigma_j\eta)} - e^{iz\psi_{j-1}(\eta_j-\eta)} \right) \frac{a(\psi_j(\eta_j+\sigma_j\eta))}{\eta} \, \mathrm{d}\eta \right| \\
&\lesssim_{M,A,q} \frac{1}{z} \left(\sum_{l=0}^{L-1} (r2^{-l})^{1-\alpha} (r2^{-l-1})^{-1} + (r2^{-L})^{1-\alpha} (z^{-\frac{1}{\alpha}})^{-1} \right) \\
&\lesssim \left(\frac{1}{z} \sum_{l=0}^{L-1} 2^{\alpha l} + z^{\frac{1}{\alpha}-1} 2^{(\alpha-1)L} \right) \lesssim_\alpha \left(\left(2^{-L} z^{\frac{1}{\alpha}} \right)^{-\alpha} + \left(2^{-L} z^{\frac{1}{\alpha}} \right)^{1-\alpha} \right) \lesssim_{r,\alpha} 1 .
\end{aligned}$$

□

3.A.2. The maximal function estimate

Let $\varphi \in C^3(\mathbb{R}; \mathbb{R})$. We assume that there exists a large $R > 0$ for which the following hold.

(J1) There exists some $A > 0$ and $n > 1$ such that

$$|\varphi'(\xi)| \geq R \Longrightarrow |\varphi'(\xi)| \leq A|\xi|^{n-1} .$$

(J2) There exists some $B > 0$ such that

$$|\varphi'(\xi)| \geq R \Longrightarrow |\varphi'(\xi)| \leq \frac{1}{B} |\xi \varphi''(\xi)| .$$

(J3) We have

$$\left\| \frac{\varphi'''}{\varphi'} \right\|_{L^1(\{|\varphi'(\xi)| \geq R\})} + \left\| \frac{\varphi''}{\varphi'} \right\|_{(L^2 \cap L^\infty)(\{|\varphi'(\xi)| \geq R\})} < \infty .$$

(J4) There exists $N \in \mathbb{N}$ such that $|(\varphi')^{-1}(\{\eta\}) \cap \{|\varphi'(\xi)| \geq R\}| \leq N$ for all $\eta \in \mathbb{R}$.

(J5) There exist $E, r > 0$ such that for all $R' > R$ we have

$$\frac{1}{2} R' \leq |\varphi'(\xi)| \leq \frac{3}{2} R' \Longrightarrow \sup_{|\zeta| < r} |\varphi''(\xi + \zeta)| \leq E R' .$$

(J6) We have $\lim_{|\xi| \to \infty} |\varphi'(\xi)| = \infty$.

Theorem 3.A.7 (Generalization of [111, Theorem 2.7, Corollary 2.8, Corollary 2.9])**.** *Let $\varphi \in C^3(\mathbb{R}; \mathbb{R})$ satisfy (J1)–(J6) and let $a \in L^\infty(\mathbb{R}; \mathbb{R})$. Let $T > 0$. Suppose $b = b(t, \xi) \in C^1([-T, T]; L^\infty(\mathbb{R}))$ has bounded support in ξ, uniformly in t. For any $u_0 \in \mathcal{D}(\mathbb{R})$ we have*

$$\left(\sum_{j \in \mathbb{Z}} \sup_{|t| \leq T} \sup_{j \leq x \leq j+1} |(a(D_x) e^{it\varphi(D_x)} + b(t, D_x)) u_0(x)|^2 \right)^{\frac{1}{2}} \lesssim_{T,s,\varphi,a,b} \|u_0\|_{H^{\frac{1}{2} \vee \frac{n-1}{4}}} , \tag{3.A.15}$$

and consequently for any $s \geq \frac{1}{2} \vee \frac{n-1}{4}$ also

$$\|(a(D_x) e^{it\varphi(D_x)} + b(t, D_x)) u_0\|_{L^2_{x \in \mathbb{R}} L^\infty_{t \in [-T,T]}} \lesssim_{T,s,\varphi,a,b} \|u_0\|_{H^s} . \tag{3.A.16}$$

We first recall another version of the Van der Corput lemma, whose proof can be found in [167, pp. 309-311].

Lemma 3.A.8 (Van der Corput lemma). *Let $\psi \in C_0^\infty(\mathbb{R};\mathbb{R})$ and assume that $\Phi \in C^2(\mathbb{R};\mathbb{R})$ with $\Phi''\big|_{\operatorname{supp}\psi} > \lambda > 0$. Then*

$$\left| \int_{\mathbb{R}} e^{i\Phi(\xi)} \psi(\xi)\,\mathrm{d}\xi \right| \leq 10\lambda^{-\frac{1}{2}} \left(\|\psi\|_{L^\infty} + \|\psi'\|_{L^1} \right). \tag{3.A.17}$$

The proof of Theorem 3.A.7 relies on the following crucial lemma.

Lemma 3.A.9 (Generalization of [111, Proposition 2.6]). *Let $\varphi \in C^3(\mathbb{R};\mathbb{R})$ fulfill (J1)–(J6) and let $\psi \in C^\infty(\mathbb{R};\mathbb{R})$ with $\operatorname{supp}\psi \subseteq [2^{k-1}, 2^{k+1}]$ for some $k \in \mathbb{N}$. Let $T > 0$. There exists a constant $c = c(\varphi) > 0$ such that the following pointwise estimate holds:*

$$\sup_{|t|\leq T} \left| \int_{\mathbb{R}} e^{i(t\varphi(\xi)+x\xi)} \psi(\xi)\,\mathrm{d}\xi \right| \lesssim_{T,\varphi} H_k^{n-1}(x) \left(\|\psi''\|_{L^1} + \|\psi'\|_{L^1\cap L^\infty} + \|\psi\|_{L^\infty} \right),$$

where

$$H_k^{n-1}(x) = \begin{cases} 2^k & , |x| \leq 2TR \\ 2^{\frac{k}{2}} |x|^{-\frac{1}{2}} & , 2TR < |x| \leq cT2^{(n-1)k} \\ (1+|x|^2)^{-1} & , |x| > cT2^{(n-1)k}. \end{cases} \tag{3.A.18}$$

Proof. We may assume $t \in (0, T]$, $x \in \mathbb{R}$. If $|x| \leq 2TR$ we estimate

$$\left| \int_{\mathbb{R}} e^{i(t\varphi(\xi)+x\xi)} \psi(\xi)\,\mathrm{d}\xi \right| \leq 2(2^{k+1} - 2^{k-1}) \|\psi\|_{L^\infty} = 2^k 3 \|\psi\|_{L^\infty}.$$

We therefore assume $|x| > 2TR > 1$ and define

$$\Omega = \left\{ \xi \in [2^{k-1}, 2^{k+1}] : |t\varphi'(\xi) + x| \leq \frac{|x|}{2} \right\}$$
$$\widetilde{\Omega} = \left\{ \xi \in [2^{k-1}, 2^{k+1}] : |t\varphi'(\xi) + x| \leq \frac{|x|}{3} \right\}.$$

Note that $\xi \in \Omega$ implies

$$\frac{1}{2}\frac{|x|}{t} \leq |\varphi'(\xi)| \leq \frac{3}{2}\frac{|x|}{t},$$

and furthermore that $|x| > 2TR$ and $t \leq T$ ensure $\frac{1}{2}\frac{|x|}{t} \geq R$. Since $R > 1$ can be chosen arbitrarily large, we are in a regime where $|\varphi'(\xi)|$ is large and hence (J1)–(J6) may be applied.

Let $\eta \in C^\infty(\mathbb{R};\mathbb{R})$ with $\operatorname{supp}\eta \subseteq \Omega$ and $\eta\big|_{\widetilde{\Omega}} = 1$. We know that $\|\eta\|_{L^\infty} \leq 1$ and now prove furthermore that it is possible to choose η so that $\|\eta'\|_{L^\infty} \lesssim_\varphi 1$. We do this by showing $\widetilde{\Omega} + B_r(0) \subseteq \Omega$ for some small radius $r > 0$ independent of t. Let $0 < |\zeta| < r$ and $\xi \in \widetilde{\Omega}$. Then

$$|t\varphi'(\xi+\zeta) + x| \leq \frac{|x|}{3} + tr \sup_{|a|<r} |\varphi''(\xi+a)|.$$

With (J5) we obtain

$$tr \sup_{|a|<r} |\varphi''(\xi+a)| \leq trE\frac{|x|}{t} = Er|x|,$$

so a choice of r with $Er < \frac{1}{6}$ suffices.

Since (J4) implies that Ω and $\widetilde{\Omega}$ can each be written as unions of $2N+1$ or fewer closed intervals, there exists a choice of η for which also $\|\eta'\|_{L^1}, \|\eta''\|_{L^1} \lesssim_\varphi 1$. If $\xi \in \operatorname{supp}\eta \subseteq \Omega$, then (J2) implies

$$|(t\varphi(\xi) + x\xi)''| = t|\varphi''(\xi)| \geq tB\left|\frac{\varphi'(\xi)}{\xi}\right| \geq \frac{B}{2^{k+1}}\frac{|x|}{2} = \frac{B}{4}2^{-k}|x|.$$

We apply (3.A.17) to obtain

$$\left|\int_{\mathbb{R}} e^{i(t\varphi(\xi)+x\xi)}(\eta\psi)(\xi)\,\mathrm{d}\xi\right| \lesssim_\varphi 2^{\frac{k}{2}}|x|^{-\frac{1}{2}}(\|(\eta\psi)'\|_{L^1}+\|\eta\psi\|_{L^\infty}) \lesssim_\varphi 2^{\frac{k}{2}}|x|^{-\frac{1}{2}}(\|\psi'\|_{L^1}+\|\psi\|_{L^\infty})\,.$$

On the other hand if $\xi \in \operatorname{supp}(1-\eta)$, then

$$|(t\varphi(\xi)+x\xi)'| = |t\varphi'(\xi)+x| \geq \frac{|x|}{3}\,.$$

Due to ψ having compact support we can perform the integration by parts below without any boundary terms appearing. We estimate

$$\begin{aligned}
&\left|\int_{\mathbb{R}} e^{i(t\varphi(\xi)+x\xi)}((1-\eta)\psi)(\xi)\,\mathrm{d}\xi\right| \\
&= \left|\int_{\mathbb{R}} e^{i(t\varphi(\xi)+x\xi)}\partial_\xi\left(\frac{1}{t\varphi'(\xi)+x}\partial_\xi\left(\frac{1}{t\varphi'(\xi)+x}((1-\eta)\psi)(\xi)\right)\right)\mathrm{d}\xi\right| \\
&\leq \int_{\operatorname{supp}(1-\eta)}\left|\frac{((1-\eta)\psi)''(t\varphi'+x)-t\varphi'''(1-\eta)\psi}{(t\varphi'+x)^3}\right|\mathrm{d}\xi \\
&\qquad + \int_{\operatorname{supp}(1-\eta)}\left|\frac{3t\varphi''\big(((1-\eta)\psi)'(t\varphi'+x)-t\varphi''(1-\eta)\psi\big)}{(t\varphi'+x)^4}\right|\mathrm{d}\xi \\
&\lesssim \frac{1}{|x|^2}\left(1+\left\|\frac{t\varphi'''}{t\varphi'+x}\right\|_{L^1(\mathbb{R}\setminus\widetilde{\Omega})}+\left\|\frac{t\varphi''}{t\varphi'+x}\right\|_{L^\infty(\mathbb{R}\setminus\widetilde{\Omega})}+\left\|\frac{(t\varphi'')^2}{(t\varphi'+x)^2}\right\|_{L^1(\mathbb{R}\setminus\widetilde{\Omega})}\right) \\
&\qquad (\|((1-\eta)\psi)''\|_{L^1}+\|((1-\eta)\psi)'\|_{L^1}+\|(1-\eta)\psi\|_{L^\infty}) \\
&\lesssim_\varphi \frac{1}{|x|^2}\left(\left(1+\left\|\frac{\varphi'''}{\varphi'}\right\|_{L^1(\{|\varphi'(\xi)|\geq R\})}+\left\|\frac{\varphi''}{\varphi'}\right\|^2_{(L^2\cap L^\infty)(\{|\varphi'(\xi)|\geq R\})}\right)\left\|\left\langle 1-\frac{x}{t\varphi'+x}\right\rangle^2\right\|_{L^\infty(\mathbb{R}\setminus\widetilde{\Omega})}\right. \\
&\qquad \left. +\|\varphi'''\|_{L^1(\{|\varphi'(\xi)|\leq R\})}+\|\varphi''\|^2_{(L^2\cap L^\infty)(\{|\varphi'(\xi)|\leq R\})}\right)(\|\psi''\|_{L^1}+\|\psi'\|_{L^1\cap L^\infty}+\|\psi\|_{L^\infty})\,.
\end{aligned}$$

Here we intersect $\mathbb{R}\setminus\widetilde{\Omega}$ with the sets $\{|\varphi'(\xi)<R\}$ and $\{|\varphi'(\xi)\geq R\}$. We apply (J3) to estimate the norms with $|\varphi'(\xi)|\geq R$, and (J6) together with continuity of φ'' and φ''' to estimate the norms with $|\varphi'(\xi)|\leq R$. We have shown that

$$\left|\int_{\mathbb{R}} e^{i(t\varphi(\xi)+x\xi)}((1-\eta)\psi)(\xi)\,\mathrm{d}\xi\right| \lesssim_\varphi |x|^{-2}(\|\psi''\|_{L^1}+\|\psi'\|_{L^1\cap L^\infty}+\|\psi\|_{L^\infty})\,.$$

In summary,

$$\begin{aligned}
\left|\int_{\mathbb{R}} e^{i(t\varphi(\xi)+x\xi)}\psi(\xi)\,\mathrm{d}\xi\right| &\lesssim_\varphi 2^{\frac{k}{2}}|x|^{-\frac{1}{2}}(\|\psi'\|_{L^1}+\|\psi\|_{L^\infty})+|x|^{-2}(\|\psi''\|_{L^1}+\|\psi'\|_{L^1}+\|\psi\|_{L^\infty}) \\
&\lesssim_{T,\varphi} 2^{\frac{k}{2}}|x|^{-\frac{1}{2}}(\|\psi''\|_{L^1}+\|\psi'\|_{L^1\cap L^\infty}+\|\psi\|_{L^\infty})\,.
\end{aligned}$$

Note that for all $\xi \in \operatorname{supp}\psi \subseteq [2^{k-1},2^{k+1}]$ we have

$$|t\varphi'(\xi)+x| \leq \frac{1}{2}|x| \Longrightarrow \frac{t|\varphi'(\xi)|}{|x|} \geq \frac{1}{2} \overset{\text{(J1)}}{\Longrightarrow} |x| \leq 2TA2^{(n-1)(k+1)}\,,$$

so we can choose a constant $c(\varphi)$ so that in the case $|x| > c(\varphi)T2^{(n-1)k}$ only the term with decay $|x|^{-2}$ appears. Since in this case $|x| > cT$, we have $|x|^{-2} \lesssim_{T,\varphi} (1+|x|^2)^{-1}$. □

Proof of Theorem 3.A.7. The proof of (3.A.15) is identical to that of [111, Theorem 2.7], except [111, Proposition 2.6] is replaced by our Lemma 3.A.9, and the time interval $|t|\leq 1$ is replaced with $|t|\leq T$. The extension of the time interval is performed separately in [111, Corollary 2.8] using a scaling argument, which yields an explicit algebraic growth bound $(1+T)^\rho, \rho > \frac{3}{4}$ for the constant in the time T. Since $\varphi(\xi)$ is not necessarily homogeneous in our situation, this scaling argument cannot be trivially generalized. As we do not need to know a growth bound for the constant in T, we can simply skip this argument and work directly with $|t| \leq T$.

The additional operators $a(D_x)$ and $b(t, D_x)$ are not present in the reference. In the proof the claim is reduced to estimating an $L^2_x(\mathbb{R})$-norm of a linear function of u, and at that point the weight $a(\xi)$ disappears. The term involving b can be controlled by $\|u_0\|_{L^2}$ using the Sobolev inequality. In this case the oscillation from the semigroup is not necessary, i. e. Lemma 3.A.9 is not used.

The estimate (3.A.16) was shown for $\varphi(\xi) = \xi|\xi|^{n-1}$, $n \geq 2$ in [111, Corollary 2.9] with explicit time growth bound $C(\varphi, s, T) = (1+T)^{\rho}\widetilde{C}(s, n-1)$, $\rho > \frac{3}{4}$. The corollary is a consequence of [111, Corollary 2.8], which itself is a consequence of [111, Theorem 2.7]. This theorem crucially relies on [111, Proposition 2.6]. We have stated and proven a generalization of this proposition above. The remaining proofs need no significant changes. Then (3.A.16) is a direct consequence of (3.A.15).

□

3.A.3. Further estimates

Here we use the standard Japanese bracket $\langle \xi \rangle = \sqrt{1+\xi^2}$.

Lemma 3.A.10. *Let $s \geq 0$ and $p, q \in [1, \infty) \cup \{\infty\}$. For all $f \in L^p_x L^q_t$ with $|D_x|^s f \in L^p_x L^q_t$ we have*

$$\| |D_x|^s f\|_{L^p_x L^q_t} \lesssim \|\langle D_x \rangle^s f\|_{L^p_x L^q_t} \lesssim \|f\|_{L^p_x L^q_t} + \||D_x|^s f\|_{L^p_x L^q_t} . \tag{3.A.19}$$

Furthermore,

$$\|\langle D_x \rangle^{-s} f\|_{L^p_x L^q_t} \lesssim \|f\|_{L^p_x L^q_t} . \tag{3.A.20}$$

Similarly if $s \in [0, 1)$ and $\partial_x f \in L^p_x L^q_t$, then

$$\|\langle D_x \rangle^s f\|_{L^p_x L^q_t} \lesssim \|f\|_{L^p_x L^q_t} + \|\partial_x f\|_{L^p_x L^q_t} . \tag{3.A.21}$$

Proof. Assume $s > 0$. For the first direction of (3.A.19) it suffices to prove that the multiplier

$$m(\xi) = \frac{|\xi|^s}{\langle \xi \rangle^s} = \left(\frac{\xi^2}{1+\xi^2}\right)^{\frac{s}{2}} = \sum_{k=0}^{\infty} \binom{\frac{s}{2}}{k} (-1)^k \langle \xi \rangle^{-2k}$$

yields a bounded operator $m(D_x) : L^p_x L^q_t \to L^p_x L^q_t$. Suppose that $f \in \mathcal{S}(\mathbb{R}^2; \mathbb{C})$. By [57, Section 1.2.4] we have

$$\langle D_x \rangle^{-2k} f = G_{2k} *_x f$$

where the Bessel kernels G_{2k} satisfy $\|G_{2k}\|_{L^1} = 1$ for $k \geq 1$. Then the mixed-norm Young's inequality [84, Theorem 3.1] implies that $\langle D_x \rangle^{-2k} : L^p_x L^q_t \to L^p_x L^q_t$ is bounded with uniformly bounded operator norm. Since the series in the definition of $m(D_x)$ is absolutely convergent, we are finished.

For the second direction of (3.A.19) we decompose

$$\langle \xi \rangle^s = \chi(\xi) \langle \xi \rangle^s + \left(1 + \frac{(1-\chi(\xi))}{m(\xi)} - 1\right) |\xi|^s = m_1(\xi) + (1 + m_2(\xi))|\xi|^s ,$$

where $\chi \in C^\infty_c(\mathbb{R}; [0, 1])$ with $\chi = 1$ in a neighborhood of the origin. Since $m_1(\xi) = \chi(\xi)\langle \xi \rangle^s$ is Schwartz, its inverse Fourier transform belongs to $\mathcal{S}(\mathbb{R}) \subseteq L^1(\mathbb{R})$. Hence, as above, the operator $m_1(D_x) : L^p_x L^q_t \to L^p_x L^q_t$ is bounded. Similarly we have $m_2 \in H^1$, and hence its inverse Fourier transform is in $H^{0,1} \subseteq L^1$, which implies as before that $m_2(D_x) : L^p_x L^q_t \to L^p_x L^q_t$ is bounded.

The estimate (3.A.20) follows again from the fact that the Bessel kernel is in L^1. Similarly one obtains (3.A.21) by splitting low and high frequencies as before, and observing that the relevant kernels are in L^1. □

3.B. Appendix: Proof of Lemma 3.2.2

Definition 3.B.1. Let X be a finite-dimensional complex Banach space. We say that $f : D \subset \mathbb{R} \to X$ is

(i) **Smooth on** D if it is infinitely differentiable on D.

(ii) **Bounded smooth on** D if it is bounded, smooth on D, and all derivatives are bounded.

(iii) **Polynomially bounded on** D if $p^{-1}f$ is bounded on D for some polynomial p.

(iv) **Polynomially bounded smooth on** D if it is polynomially bounded, smooth on D, and all derivatives are polynomially bounded.

(v) **Rapidly decreasing at** $\pm\infty$ if pf is bounded on $D \cap \mathbb{R}_\pm$ for every polynomial p.

(vi) **Schwartz at** $\pm\infty$ if pf is bounded smooth on $D \cap \mathbb{R}_\pm$ for every polynomial p.

Proof of Lemma 3.2.2. Recall for $\Gamma^\pm = \Gamma^\pm_{\mathrm{d}} + \Gamma^\pm_{\mathrm{od}}$ the system

$$(\partial_x + \partial_y)\Gamma^\pm_{\mathrm{d}}(x,y) = \Big(Q(x) + \overrightarrow{Q_*(y)}\Big)\Gamma^\pm_{\mathrm{od}}(x,y) \tag{3.2.15}$$

$$(\partial_x - \partial_y)\Gamma^\pm_{\mathrm{od}}(x,y) = \Big(Q(x) - \overrightarrow{Q_*(y)}\Big)\Gamma^\pm_{\mathrm{d}}(x,y) \tag{3.2.16}$$

$$\lim_{y\to\pm\infty} \Gamma^\pm(x,y) = 0 \qquad\qquad \Gamma^\pm_{\mathrm{od}}(x,x) = \frac{1}{2}(Q - Q_*)_{\mathrm{od}}(x)\,. \tag{3.2.17}$$

We give the proof for the case $\pm = -$; the plus case is analogous.

Reformulation into integral equations. The first step is to write (3.2.15)–(3.2.17) as a system of integral equations:

$$\Gamma^-_{\mathrm{d}}(x,y) = \int_{-\infty}^{x} \Big(Q(s) + \overrightarrow{Q_*(s+y-x)}\Big)\Gamma^-_{\mathrm{od}}(s, s+y-x)\,\mathrm{d}s$$

$$\Gamma^-_{\mathrm{od}}(x,y) = \frac{1}{2}(Q - Q_*)\Big(\frac{x+y}{2}\Big) + \int_{\frac{x+y}{2}}^{x} \Big(Q(s) - \overrightarrow{Q_*(x+y-s)}\Big)\Gamma^-_{\mathrm{d}}(s, x+y-s)\,\mathrm{d}s\,.$$

We now fix $K \in \mathbb{N}$ and consider the variables

$$\Gamma^-_K(x,y) = (\partial_y^k \Gamma^-(x,y))_{0\le k\le K} \in \big(\mathbb{C}^{2\times 2}\big)^{K+1}$$
$$(\partial_x + \partial_y)\Gamma^-_K(x,y) = ((\partial_x + \partial_y)\partial_y^k \Gamma^-(x,y))_{0\le k\le K} \in \big(\mathbb{C}^{2\times 2}\big)^{K+1}$$

for which the extended system of integral equations reads

$$\partial_y^k \Gamma^-_{\mathrm{d}}(x,y) = \int_{-\infty}^{x} \sum_{l=0}^{k} \binom{k}{l} \partial_y^l \Big(Q(s) + \overrightarrow{Q_*(s+y-x)}\Big)(\partial_y^{k-l}\Gamma^-_{\mathrm{od}})(s, s+y-x)\,\mathrm{d}s \tag{3.B.1}$$

$$\begin{aligned}\partial_y^k \Gamma^-_{\mathrm{od}}(x,y) &= \frac{1}{2}\partial_y^k \Big((Q - Q_*)\Big(\frac{x+y}{2}\Big)\Big) \\ &+ \int_{\frac{x+y}{2}}^{x} \sum_{l=0}^{k} \binom{k}{l} \partial_y^l \Big(Q(s) - \overrightarrow{Q_*(x+y-s)}\Big)(\partial_y^{k-l}\Gamma^-_{\mathrm{d}})(s, x+y-s)\,\mathrm{d}s \\ &- \sum_{j=0}^{k-1}\sum_{l=0}^{k-1-j} \binom{k-1-j}{l} \frac{1}{2}\partial_y^j \Bigg(\partial_y^l\Big(Q(s) - \overrightarrow{Q_*(x+y-s)}\Big)(\partial_y^{k-1-j-l}\Gamma^-_{\mathrm{d}})(s, x+y-s)\Big|_{s=\frac{x+y}{2}}\Bigg)\end{aligned} \tag{3.B.2}$$

and

$$(\partial_x + \partial_y)\partial_y^k \Gamma^-_{\mathrm{d}}(x,y) = \sum_{l=0}^{k} \binom{k}{l} \partial_y^l \Big(Q(x) + \overrightarrow{Q_*(y)}\Big)(\partial_y^{k-l}\Gamma^-_{\mathrm{od}})(x,y) \tag{3.B.3}$$

$$\begin{aligned}(\partial_x + \partial_y)\partial_y^k \Gamma^-_{\mathrm{od}}(x,y) &= \partial_y^{k+1} \Big((Q - Q_*)\Big(\frac{x+y}{2}\Big)\Big) \\ &+ 2\int_{\frac{x+y}{2}}^{x} \sum_{l=0}^{k+1} \binom{k+1}{l} \partial_y^l \Big(Q(s) - \overrightarrow{Q_*(x+y-s)}\Big)(\partial_y^{k+1-l}\Gamma^-_{\mathrm{d}})(s, x+y-s)\,\mathrm{d}s\end{aligned} \tag{3.B.4}$$

$$- \sum_{j=1}^{k} \sum_{l=0}^{k-j} \binom{k-j}{l} \partial_y^j \left(\partial_y^l \Big(Q(s) - \overrightarrow{Q_*(x+y-s)} \Big) (\partial_y^{k-j-l} \Gamma_{\mathrm{d}}^-)(s, x+y-s) \Big|_{s=\frac{x+y}{2}} \right)$$
$$+ \sum_{l=0}^{k} \binom{k}{l} \Big[\partial_y^l \Big(Q(s) - \overrightarrow{Q_*(x+y-s)} \Big) (\partial_y^{k-l} \Gamma_{\mathrm{d}}^-)(s, x+y-s) \Big]_{s=\frac{x+y}{2}}^{s=x} .$$

Reduction to the proof of (3.2.18)**.** Note that for $k \geq 1$ (3.2.15)–(3.2.16) imply

$$(\partial_x + \partial_y)^k \Gamma_{\mathrm{d}}^-(x,y) = \sum_{j=0}^{k-1} \binom{k-1}{j} (\partial_x + \partial_y)^{k-1-j} \Big(Q(x) + \overrightarrow{Q_*(y)} \Big) (\partial_x + \partial_y)^j \Gamma_{\mathrm{od}}^-(x,y)$$
$$(\partial_x - \partial_y)^k \Gamma_{\mathrm{od}}^-(x,y) = \sum_{j=0}^{k-1} \binom{k-1}{j} (\partial_x - \partial_y)^{k-1-j} \Big(Q(x) - \overrightarrow{Q_*(y)} \Big) (\partial_x - \partial_y)^j \Gamma_{\mathrm{d}}^-(x,y) .$$

As a result, $\partial_x^k \Gamma^-$ can be written as a linear (in Γ^-) combination of $(\partial_x^j \Gamma^-)_{0 \leq j \leq k-1}$ and $(\partial_y^l \Gamma^-)_{0 \leq l \leq k}$. It therefore suffices to construct $\partial_y^k \Gamma^-$ and prove the estimates (3.2.18)–(3.2.20) for $m = 0$. Smoothness of Γ^- then follows by an inductive construction.

Furthermore, we only need to prove (3.2.18), as the observation that our construction yields $\lim_{s \to -\infty} \Gamma^-(x+s, y+s) = 0$ (see (3.B.10) below), together with (3.2.18), directly implies (3.2.20):

$$|\partial_x^m \partial_y^k \Gamma^-(x,y)| \leq 0 + \Big| \int_{-\infty}^{\frac{x+y}{2}} ((\partial_x + \partial_y) \partial_x^m \partial_y^k \Gamma^-)(s + \frac{x-y}{2}, s - \frac{x-y}{2})\, ds \Big|$$
$$\leq \Big\| ((\partial_x + \partial_y) \partial_x^m \partial_y^k \Gamma^-)(s + \frac{x-y}{2}, s - \frac{x-y}{2}) \Big\|_{L_s^1((-\infty, \frac{x+y}{2}])} \overset{(3.2.18)}{\leq} c_-(\frac{x+y}{2}) e^{\int_y^x c_-(s) ds} .$$

A claim on the properties at the diagonal. We begin by analyzing the behavior of (3.B.1)–(3.B.4) on the diagonal $y = x$. Recall that we assume $q - q_*$ to be Schwartz, and note that the following "canceling product" of multiplication operators

$$\Big(Q(x) - \overrightarrow{Q_*(y)} \Big) \Big(Q(x) + \overrightarrow{Q_*(y)} \Big) = Q(x)^2 - \overrightarrow{Q_*(y)^2} = (|q|^2 - 1)(x) - (|q_*|^2 - 1)(y) \tag{3.B.5}$$

is Schwartz if $x = y$. This is an essential source of integrability in our argument.

Claim 3.B.2. *The map $[x \mapsto (\partial_y^k \Gamma^-)(x,x)]$ is polynomially bounded smooth on $\mathbb{R}$ and Schwartz at $-\infty$. Furthermore, $((\partial_x + \partial_y) \partial_y^k \Gamma^-)(x,x)$ and $\Big(Q(x) + \overrightarrow{Q_*(x)} \Big) (\partial_y^k \Gamma_{\mathrm{od}}^-)(x,x)$ are Schwartz functions in $x \in \mathbb{R}$.*

Proof of claim. Specializing (3.B.1)–(3.B.4) to $y = x$ yields

$$(\partial_y^k \Gamma_{\mathrm{d}}^-)(x,x) = \int_{-\infty}^{x} \sum_{l=0}^{k} \binom{k}{l} \partial_y^l \Big(Q(s) + \overrightarrow{Q_*(s+y-x)} \Big) \Big|_{y=x} (\partial_y^{k-l} \Gamma_{\mathrm{od}}^-)(s,s)\, \mathrm{d}s \tag{3.B.6}$$

$$(\partial_y^k \Gamma_{\mathrm{od}}^-)(x,x) = \frac{1}{2^{k+1}} \partial_x^k (Q - Q_*)(x) - \sum_{j=0}^{k-1} \sum_{l=0}^{k-1-j} \binom{k-1-j}{l} \frac{1}{2} \tag{3.B.7}$$
$$\partial_y^j \left(\partial_y^l \Big(Q(s) - \overrightarrow{Q_*(x+y-s)} \Big) (\partial_y^{k-1-j-l} \Gamma_{\mathrm{d}}^-)(s, x+y-s) \Big|_{s=\frac{x+y}{2}} \right) \Big|_{y=x}$$

and

$$((\partial_x + \partial_y) \partial_y^k \Gamma_{\mathrm{d}}^-)(x,x) = \sum_{l=0}^{k} \binom{k}{l} \partial_y^l \Big(Q(x) + \overrightarrow{Q_*(y)} \Big) \Big|_{y=x} (\partial_y^{k-l} \Gamma_{\mathrm{od}}^-)(x,x) \tag{3.B.8}$$

$$((\partial_x + \partial_y) \partial_y^k \Gamma_{\mathrm{od}}^-)(x,x) = \frac{1}{2^{k+1}} \partial_x^{k+1} (Q - Q_*)(x) - \sum_{j=1}^{k} \sum_{l=0}^{k-j} \binom{k-j}{l} \tag{3.B.9}$$

$$\partial_y^j \left(\partial_y^l \Big(Q(s) - \overrightarrow{Q_*(x+y-s)} \Big) (\partial_y^{k-j-l} \Gamma_{\mathrm{d}}^-)(s, x+y-s) \bigg|_{s=\frac{x+y}{2}} \right) \bigg|_{y=x} .$$

Recall here that Q and Q_* are bounded smooth and $Q - Q_*$ is Schwartz, which implies by (3.2.17) that $\Gamma_{\mathrm{od}}^-(x,x) = \frac{1}{2}(Q - Q_*)(x)$ is Schwartz. Then $\Gamma_{\mathrm{d}}^-(x,x)$ is Schwartz at $-\infty$ and bounded smooth on $\mathbb{R}$ by (3.B.6). By induction over $k \in \mathbb{N}$, using (3.B.6)–(3.B.7), it follows that $(\partial_y^k \Gamma^-)(x,x)$ is Schwartz at $-\infty$ and polynomially bounded smooth on $\mathbb{R}$.

Even though we only know $(\partial_y^k \Gamma^-)(x,x)$ to be Schwartz at $-\infty$, we can show that $((\partial_x + \partial_y)\partial_y^k \Gamma^-)(x,x)$ is Schwartz. We consider first $((\partial_x + \partial_y)\partial_y^k \Gamma_{\mathrm{d}}^-)(x,x)$. If $l > 0$ in (3.B.8), a Schwartz factor $\partial_x^l Q_*(x)$ is present, so it suffices to consider $l = 0$ and show that $\Big(Q(x) + \overrightarrow{Q_*(x)}\Big)(\partial_y^k \Gamma_{\mathrm{od}}^-)(x,x)$ is Schwartz. Substituting with (3.B.7), we again obtain a Schwartz factor if any derivative from ∂_y^j or ∂_y^l falls onto $\Big(Q(s) - \overrightarrow{Q_*(x+y-s)}\Big)$. In the final case, where this does not happen, we have as a factor the canceling product (3.B.5), which is Schwartz. Therefore $((\partial_x + \partial_y)\partial_y^k \Gamma_{\mathrm{d}}^-)(x,x)$ is Schwartz.

Lastly, we show that $((\partial_x + \partial_y)\partial_y^k \Gamma_{\mathrm{od}}^-)(x,x)$ is also Schwartz. In (3.B.9) we see again that the only difficult term is the one where no derivative from ∂_y^j or ∂_y^l falls onto $\Big(Q(s) - \overrightarrow{Q_*(x+y-s)}\Big)$. Since $j \geq 1$, we only need to consider terms where a derivative from ∂_y^j has fallen onto $(\partial_y^{k-j-l}\Gamma_{\mathrm{d}}^-)(s, x+y-s)|_{s=\frac{x+y}{2}}$, i. e. there is a factor $((\partial_x + \partial_y)\partial_y^{k-j-l}\Gamma_{\mathrm{d}}^-)(\frac{x+y}{2}, \frac{x+y}{2})$. Substituting (3.B.8), we again obtain a Schwartz factor either from a derivative or the canceling product (3.B.5). $\square$ (Claim)

Well-definedness of (3.B.10) and proof of (3.2.18). The rest of the proof is more elegantly written using the variables

$$\Lambda_{K,\mathrm{d}}(x,y) = \Gamma_{K,\mathrm{d}}^-(x+y, x-y) \qquad\qquad \Lambda_{K,\mathrm{od}}(x,y) = \Gamma_{K,\mathrm{od}}^-(x+y, x-y)$$

and $\Lambda_K = \Lambda_{K,\mathrm{d}} + \Lambda_{K,\mathrm{od}}$. We are going to construct Λ_K via

$$\Lambda_K(x_0, y_0) = \int_{-\infty}^{x_0} (\partial_x \Lambda_K)(x_1, y_0)\,\mathrm{d}x_1\,, \tag{3.B.10}$$

which implies

$$\lim_{x_0 \to -\infty} \Lambda_K(x_0, y_0) = 0\,, \qquad \forall y_0 \geq 0. \tag{3.B.11}$$

We aim to estimate $\partial_x \Lambda_K$ and show the well-definedness of (3.B.10).

Reformulation of $\partial_x \Lambda_K$ in terms of integral equations. The integral equations (3.B.1)–(3.B.2) can be succinctly written as

$$\Lambda_{K,\mathrm{d}}(x_0, y_0) = \int_{-\infty}^{x_0} \Theta_K^+(x_1, y_0)\Lambda_{K,\mathrm{od}}(x_1, y_0)\,\mathrm{d}x_1 \tag{3.B.12}$$

$$\Lambda_{K,\mathrm{od}}(x_0, y_0) = \int_0^{y_0} \Theta_K^-(x_0, y_1)\Lambda_{K,\mathrm{d}}(x_0, y_1)\,\mathrm{d}y_1 + \Gamma_{K,\mathrm{od}}^-(x_0, x_0) \tag{3.B.13}$$

where

$$\Theta_K^\pm(x,y) = \left(\binom{k}{l} 2^{l-k} (\partial_x - \partial_y)^{k-l} \Big(Q(x+y) \pm \overrightarrow{Q_*(x-y)} \Big) \right)_{0 \leq k,l \leq K} .$$

The system of integral equations for $\partial_x \Lambda_K = \partial_x \Lambda_{K,\mathrm{d}} + \partial_x \Lambda_{K,\mathrm{od}}$ is

$$\begin{aligned}(\partial_x \Lambda_{K,\mathrm{d}})(x_0, y_0) = \int_0^{y_0} \int_{-\infty}^{x_0} &(\partial_y \Theta_K^+)(x_0, y_1)(\partial_x \Lambda_{K,\mathrm{od}})(x_1, y_1) \\ &+ (\Theta_K^+ \Theta_K^-)(x_0, y_1)(\partial_x \Lambda_{K,\mathrm{d}})(x_1, y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 + \Theta_K^+(x_0, 0)\Gamma_{K,\mathrm{od}}^-(x_0, x_0)\end{aligned} \tag{3.B.14}$$

$$
\begin{aligned}
(\partial_x\Lambda_{K,\mathrm{od}})(x_0,y_0) = \int_0^{y_0}\int_{-\infty}^{x_0} &(\partial_x\Theta_K^-)(x_0,y_1)(\partial_x\Lambda_{K,\mathrm{d}})(x_1,y_1) \\
&+ (\Theta_K^-\Theta_K^+)(x_0,y_1)(\partial_x\Lambda_{K,\mathrm{od}})(x_1,y_1)\,\mathrm{d}x_1\,\mathrm{d}y_1 + ((\partial_x+\partial_y)\Gamma_{K,\mathrm{od}}^-)(x_0,x_0)\,.
\end{aligned} \tag{3.B.15}
$$

To obtain (3.B.14), we compute ∂_{x_0}(3.B.12), apply the fundamental theorem of calculus (FTC) over the interval $(0,y_0)$, substitute a term with ∂_{y_0}(3.B.13) and apply once more the FTC over the interval $(-\infty,x_0)$. Similarly, for (3.B.15) we compute ∂_{x_0}(3.B.13) and apply the FTC over $(-\infty,x_0)$ for one term, while substituting with ∂_{x_0}(3.B.12) for the other. Here we used the vanishing property (3.B.11).

Construction of $\partial_x\Lambda_K$. We shall construct $\partial_x\Lambda_K = \partial_x\Lambda_{K,\mathrm{d}} + \partial_x\Lambda_{K,\mathrm{od}}$ via the Neumann series ansatz

$$
\partial_x\Lambda_K = \sum_{n=0}^{\infty}\partial_x\Lambda_{K,n}.
$$

We estimate the coefficients and inhomogeneous terms in (3.B.14)–(3.B.15) as follows:

- Observe first that there exists a positive rapidly decreasing function $g_K \in C_b(\mathbb{R};\mathbb{R}_+)$ such that the terms appearing in the integrals in (3.B.14)–(3.B.15), except $\partial_x\Lambda_{K,\mathrm{d}}$ and $\partial_x\Lambda_{K,\mathrm{od}}$, can be bounded using

$$
F_K(x,y) = \big(|\partial_y\Theta_K^+| + |\partial_x\Theta_K^-| + |\Theta_K^+\Theta_K^-| + |\Theta_K^-\Theta_K^+|\big)(x,y) \le g_K(x+y) + g_K(x-y)\,. \tag{3.B.16}
$$

  This uses the canceling product (3.B.5) when no derivatives are present. To be precise, there exists a constant $C_K > 0$ such that

$$
g_K = C_K\Big(1 + \|(q,q_*)\|_{C_b^K(\mathbb{R})}\Big)\left(\big||q|^2-1\big| + \big||q_*|^2-1\big| + \sum_{k=1}^{K+1}|\partial_x^k q| + |\partial_x^k q_*|\right)
$$

  is a valid choice for g_K.

- Recall now from Claim 3.B.2 that $\Theta_K^+(x_0,0)\Gamma_{K,\mathrm{od}}^-(x_0,x_0)$ and $((\partial_x+\partial_y)\Gamma_{K,\mathrm{od}}^-)(x_0,x_0)$ are Schwartz. Formally, let $\mathcal{F}_{K,0}$ denote the set consisting of two functions

$$
\mathcal{F}_{K,0} = \Big\{[(x,y)\mapsto\Theta_K^+(x,0)\Gamma_{K,\mathrm{od}}^-(x,x)],\quad [(x,y)\mapsto((\partial_x+\partial_y)\Gamma_{K,\mathrm{od}}^-)(x,x)]\Big\},
$$

  then any element $f_0(x,y)\in\mathcal{F}_{K,0}$ is a function independent of y and Schwartz in x, and hence there exists a bounded function h_K which rapidly decreases at $-\infty$ such that

$$
\|f_0(x,y)\|_{L_x^1L_y^\infty((-\infty,x_0]\times\mathbb{R}_+)} \lesssim h_K(x_0)\,. \tag{3.B.17}
$$

  If $K=0$ then a valid choice for h_0 is

$$
h_0(x_0) = C(1+\|q\|_{L^\infty}+\|q_*\|_{L^\infty})\int_{-\infty}^{x_0}(|q-q_*| + |\partial_x(q-q_*)|)\,\mathrm{d}x\,.
$$

  For $K\ge 1$ an explicit example may be constructed by tracing the proof of Claim 3.B.2.

We now study the iterated application of the integral operators corresponding to the Neumann series ansatz for (3.B.14)–(3.B.15). For $n\ge1$ we find that there exist sets

$$
\mathcal{F}_{K,n}\subset\mathcal{F}_{K,0}\times\big\{\partial_y\Theta_K^+,\partial_x\Theta_K^-,\Theta_K^+\Theta_K^-,\Theta_K^-\Theta_K^+\big\}^n,
$$

consisting of vectors $(f_0,f_1,\dots,f_n)$ of functions, such that

$$
\begin{aligned}
(\partial_x\Lambda_{K,n})(x_0,y_0) = \sum_{(f_0,f_1,\dots,f_n)\in\mathcal{F}_{K,n}}\int\cdots\int &\mathbb{1}_{\{x_0>\cdots>x_n\}}\mathbb{1}_{\{y_0>\cdots>y_n>0\}} \\
&\prod_{j=1}^{n}f_j(x_{j-1},y_j)f_0(x_n,y_n)\,\mathrm{d}x_n\dots\mathrm{d}x_1\,\mathrm{d}y_n\dots\mathrm{d}y_1\,.
\end{aligned}
$$

We use the integrability of $f_0(x_n,y_n)$ in (3.B.17) to estimate the integral over $\mathrm{d}x_n$ in L^1. Next, we use the $n!$-fold symmetry of the integral (resp. $(n-1)!$-fold, when the $L^\infty_{x_1}$-norm is used instead of the $L^1_{x_1}$-norm) to obtain

$$\begin{aligned}\|(\partial_x\Lambda_{K,n})(x_1,y_0)\|_{(L^1\cap L^\infty)_{x_1}((-\infty,x_0])} &\lesssim h_K(x_0)|\mathcal{F}_{K,n}|\|F_K\|^{n-1}_{L^1_xL^1_y((-\infty,x_0]\times[0,y_0])}\\ &\qquad\left(\frac{\|F_K\|_{L^1_xL^1_y((-\infty,x_0]\times[0,y_0])}}{n!}+\frac{\|F_K\|_{L^\infty_xL^1_y((-\infty,x_0]\times[0,y_0])}}{(n-1)!}\right).\end{aligned}$$

Define the integral operator $\mathcal{I}g_K(x)=\int_{-\infty}^x g_K(s)\,\mathrm{d}s$. The estimate (3.B.16) implies

$$\begin{aligned}\|F_K\|_{L^\infty_xL^1_y((-\infty,x_0]\times[0,y_0])} &\lesssim \mathcal{I}g_K(x_0+y_0)\\ \|F_K\|_{L^1_xL^1_y((-\infty,x_0]\times[0,y_0])} &\lesssim \mathcal{I}^2g_K(x_0+y_0)-\mathcal{I}^2g_K(x_0-y_0)\,.\end{aligned}$$

Including the trivial case $n=0$, we have shown for all $n\geq 0$ that

$$\begin{aligned}&\|(\partial_x\Lambda_{K,n})(x,y_0)\|_{(L^1\cap L^\infty)_x((-\infty,x_0])}\\ &\lesssim C^nh_K(x_0)\left(\frac{(\mathcal{I}^2g_K(x_0+y_0)-\mathcal{I}^2g_K(x_0-y_0))^n}{n!}+\mathbb{1}_{\{n\geq1\}}\mathcal{I}g_K(x_0+y_0)\frac{(\mathcal{I}^2g_K(x_0+y_0)-\mathcal{I}^2g_K(x_0-y_0))^{n-1}}{(n-1)!}\right).\end{aligned}$$

Summing over $n\geq 0$, we arrive at

$$\|(\partial_x\Lambda_K)(x,y_0)\|_{(L^1\cap L^\infty)_x((-\infty,x_0])}\lesssim h_K(x_0)(1+\mathcal{I}g_K(x_0+y_0))e^{C(\mathcal{I}^2g_K(x_0+y_0)-\mathcal{I}^2g_K(x_0-y_0))}\,.$$

Then Λ_K is well-defined in (3.B.10), and by the definition of Λ_K the estimate (3.2.18) follows.

$\square$

CHAPTER 4

Long-wave KdV Hierarchy Approximation of the NLS Hierarchy with Nonzero Boundary Conditions

Abstract

We study the approximation of certain renormalized conserved quantities for the NLS hierarchy with nonzero boundary conditions, in the long-wave regime, by the energies of the KdV hierarchy. We extend this to all $n \in \mathbb{N}$ by proving an approximation result for the transmission coefficient of the Lax operator of the NLS hierarchy, which is a Dirac operator in the nonrelativistic regime, by the transmission coefficient of a Schrödinger operator, which is the Lax operator of the KdV hierarchy. This yields a formal approximation result between the hierarchies, which we quantify using energy methods and previously established well-posedness results.

4.1. Introduction

We fix $N \in \mathbb{N}$ and consider the "N-truncated" defocusing NLS hierarchy [1]

$$i\partial_{t_n} q = \frac{\delta \mathcal{H}_n^{\mathsf{NLS}}(q)}{\delta \overline{q}} \qquad \forall\, n \in \{0, \dots, N\} \tag{NLS_n}$$

for a function $q = q(\mathbf{t}, x) = q(t_0, \dots, t_N, x) : \mathbb{R}^{N+1} \times \mathbb{R} \longrightarrow \mathbb{C}$. Here $\mathcal{H}_n^{\mathsf{NLS}}$ are the Hamiltonians of the NLS hierarchy, the first of which are

$$\mathcal{H}_0^{\mathsf{NLS}}(q) = \int_{\mathbb{R}} q\overline{q}\, \mathrm{d}x \tag{Mass}$$

$$\mathcal{H}_1^{\mathsf{NLS}}(q) = -i \int_{\mathbb{R}} q\overline{q}_x\, \mathrm{d}x \tag{Momentum}$$

$$\mathcal{H}_2^{\mathsf{NLS}}(q) = \int_{\mathbb{R}} q_x\overline{q}_x + q^2\overline{q}^2\, \mathrm{d}x \tag{Energy}$$

$$\mathcal{H}_3^{\mathsf{NLS}}(q) = i \int_{\mathbb{R}} q\overline{q}_{xxx} - 4q^2\overline{q}\overline{q}_x - q_x q\overline{q}^2\, \mathrm{d}x$$

$$\mathcal{H}_4^{\mathsf{NLS}}(q) = \int_{\mathbb{R}} q_{xx}\overline{q}_{xx} - 6q^2\overline{q}\overline{q}_{xx} - 5q^2\overline{q}_x^2 - 6qq_x\overline{q}\overline{q}_x - qq_{xx}\overline{q}^2 + 2q^3\overline{q}^3\, \mathrm{d}x\,.$$

We are interested in long-wave solutions q of the form

$$q(\mathbf{t}, x) = \left(1 + \frac{\varepsilon^2}{2}(W_+ - W_-)(\mathbf{T}, X)\right) \exp\left(i\frac{\varepsilon}{\sqrt{2}} \int_0^X (W_+ + W_-)(\mathbf{T}, Y)\, \mathrm{d}Y\right) e^{i\varphi(\mathbf{t})} \tag{4.1.1}$$

which satisfy the nonzero boundary conditions

$$\lim_{x\to\pm\infty} q(\mathbf{t}, x)e^{-i\psi(\mathbf{t})} = q_\pm \text{ where } |q_\pm| = 1\,. \tag{NZBC}$$

We use the rescaled space and time variables

$$X = \sqrt{2}\varepsilon x \qquad\qquad \mathbf{T} = (T_n)_{0\le n\le N} = \left(\frac{\mathrm{d}T_n}{\mathrm{d}t_n} t_n\right)_{0\le n\le N}$$

[1] Some authors define the NLS hierarchy to consist of only the even flows and refer to the odd flows as the mKdV hierarchy.

with

$$\frac{\mathrm{d}T_{2n}}{\mathrm{d}t_{2n}} = 2(\sqrt{2}\varepsilon)^{2n-1} \qquad \frac{\mathrm{d}T_{2n-1}}{\mathrm{d}t_{2n-1}} = (\sqrt{2}\varepsilon)^{2n-1}\,, \tag{4.1.2}$$

where $\varepsilon > 0$ is a small parameter. The functions $\varphi, \psi : \mathbb{R}^{N+1} \to \mathbb{R}$ are assumed to depend only on $\mathbf{t}$, representing uniform-in-x phase rotations. We assume $W_+, W_- : \mathbb{R}^{N+1} \times \mathbb{R} \longrightarrow \mathbb{R}$ and use the variable $W = (W_+, W_-)$ to denote the pair. In order to match NZBC for q, we choose for W the zero boundary condition

$$\lim_{|X|\to\infty} W(\mathbf{T}, X) = 0\,.$$

We are interested only in the dynamics of W. Define $M = \left\lfloor \frac{N-1}{2} \right\rfloor$. Our main result is Theorem 4.1.12 below, which states that the behavior of the N-truncated NLS hierarchy with NZBC in the long-wave regime can be described by the M-truncated KdV hierarchy

$$\partial_{T_n} U = \frac{1}{2}\partial_X \frac{\delta \mathcal{E}_n^{\mathsf{KdV}}(U)}{\delta U} \qquad \forall\, n \in \{0, \ldots, M\} \tag{KdV_n}$$

for a function $U = U(\mathbf{T}, X) = U(T_0, \ldots, T_M, X) : \mathbb{R}^{M+1} \times \mathbb{R} \to \mathbb{R}$. Here $\mathcal{E}_n^{\mathsf{KdV}}$ are the Hamiltonians of the KdV hierarchy, the first of which are

$$\mathcal{E}_{-1}^{\mathsf{KdV}}(U) = \int_{\mathbb{R}} U \,\mathrm{d}X \tag{Mass}$$

$$\mathcal{E}_{0}^{\mathsf{KdV}}(U) = \int_{\mathbb{R}} U^2 \,\mathrm{d}X \tag{Momentum}$$

$$\mathcal{E}_{1}^{\mathsf{KdV}}(U) = \int_{\mathbb{R}} U_X^2 + 2U^3 \,\mathrm{d}X \tag{Energy}$$

$$\mathcal{E}_{2}^{\mathsf{KdV}}(U) = \int_{\mathbb{R}} U_{XX}^2 + 10U_X^2 U + 5U^4 \,\mathrm{d}X$$

$$\mathcal{E}_{3}^{\mathsf{KdV}}(U) = \int_{\mathbb{R}} U_{XXX}^2 + 14U_{XX}^2 U + 70U_X^2 U^2 + 14U^5 \,\mathrm{d}X\,.$$

Theorem 4.1.12 is stated below after further preliminary discussion. In Examples 4.1.13–4.1.20 at the end of the introduction several explicit demonstrations of our main result may be found.

4.1.1. From NLS to GP and $\widetilde{\mathsf{GP}}$

Let us begin by recalling the defocusing nonlinear Schrödinger equation

$$iq_t + q_{xx} = 2q|q|^2 \tag{NLS}$$

for a wavefunction $q = q(t,x) : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{C}$ that satisfies the zero boundary condition

$$\lim_{|x|\to\infty} q(t,x) = 0\,. \tag{ZBC}$$

It is closely related to the defocusing Gross-Pitaevskii equation

$$iq_t + q_{xx} = 2q(|q|^2 - 1)\,, \tag{GP}$$

which we associate with the nonzero boundary conditions

$$\lim_{x\to\pm\infty} q(t,x) = q_\pm \text{ where } |q_\pm| = 1\,. \tag{NZBC}$$

Both equations are formally equivalent, as the transformation $q \mapsto e^{2it}q$ maps solutions of NLS to solutions of GP. It is preferable to work with the system GP–NZBC instead of NLS–NZBC because solutions to the former system are stationary at infinity, which is not the case for the latter system. When writing NLS and GP, we generally mean the systems NLS–ZBC and GP–NZBC.

The Hamiltonians $\mathcal{H}_n^{\mathsf{NLS}}$ are formally conserved quantities for both NLS and GP, but ill-defined in the setting of NZBC. We use instead an alternative sequence of Hamiltonians $(\mathcal{H}_n^{\mathsf{GP}})_{n\geq -1}$ which is compatible with NZBC. The initial ones are

$$\begin{aligned}
\mathcal{H}_0^{\mathsf{GP}}(q) &= \int_{\mathbb{R}} q\overline{q} - 1\,\mathrm{d}x && \text{(Mass)}\\
\mathcal{H}_1^{\mathsf{GP}}(q) &= -i\int_{\mathbb{R}} q\overline{q}_x\,\mathrm{d}x && \text{(Momentum)}\\
\mathcal{H}_2^{\mathsf{GP}}(q) &= \int_{\mathbb{R}} q_x\overline{q}_x + (q\overline{q}-1)^2\,\mathrm{d}x && \text{(Energy)}\\
\mathcal{H}_3^{\mathsf{GP}}(q) &= i\int_{\mathbb{R}} q\overline{q}_{xxx} - 4q^2\overline{q}\overline{q}_x - q_x q\overline{q}^2 + 4q\overline{q}_x\,\mathrm{d}x\\
\mathcal{H}_4^{\mathsf{GP}}(q) &= \int_{\mathbb{R}} q_{xx}\overline{q}_{xx} - 6q^2\overline{q}\overline{q}_{xx} - 5q^2\overline{q}_x^2 - 6qq_x\overline{q}\overline{q}_x - qq_{xx}\overline{q}^2\\
&\qquad - 6q_x\overline{q}_x + 2q^3\overline{q}^3 - 6q^2\overline{q}^2 + 6q\overline{q} - 2\,\mathrm{d}x\,.
\end{aligned}$$

In this setting there exists an additional non-trivial conserved quantity: the asymptotic phase change

$$\mathcal{H}_{-1}^{\mathsf{GP}}(q) = -i\log_{\mathrm{pr}}\left(\frac{q_-}{q_+}\right).$$

Here $\log_{\mathrm{pr}}$ denotes the principal logarithm.

We consider also another sequence of Hamiltonians $(\mathcal{H}_n^{\widetilde{\mathsf{GP}}})_{n\geq -1}$, which is defined by $\mathcal{H}_{2n}^{\widetilde{\mathsf{GP}}} = \mathcal{H}_{2n}^{\mathsf{GP}}$ and the equivalent relations

$$\mathcal{H}_{2n-1}^{\widetilde{\mathsf{GP}}} = \sum_{m=0}^{n}\binom{-\frac{1}{2}}{n-m}4^{n-m}2^{-\mathbb{1}_{\{m=0\}}}\mathcal{H}_{2m-1}^{\mathsf{GP}} \quad\text{and}\quad \mathcal{H}_{2n-1}^{\mathsf{GP}} = 2^{\mathbb{1}_{\{n=0\}}}\sum_{m=0}^{n}\binom{\frac{1}{2}}{n-m}4^{n-m}\mathcal{H}_{2m-1}^{\widetilde{\mathsf{GP}}}\,. \tag{4.1.3}$$

Here the initial odd Hamiltonians are

$$\begin{aligned}
\mathcal{H}_{-1}^{\widetilde{\mathsf{GP}}}(q) &= \frac{1}{2}\mathcal{H}_{-1}^{\mathsf{GP}}(q)\\
\mathcal{H}_{1}^{\widetilde{\mathsf{GP}}}(q) &= -i\int_{\mathbb{R}} q\overline{q}_x\,\mathrm{d}x - \mathcal{H}_{-1}^{\mathsf{GP}}(q)\\
\mathcal{H}_{3}^{\widetilde{\mathsf{GP}}}(q) &= i\int_{\mathbb{R}} q\overline{q}_{xxx} - 4q^2\overline{q}\overline{q}_x - q_x q\overline{q}^2 + 6q\overline{q}_x\,\mathrm{d}x + 3\mathcal{H}_{-1}^{\mathsf{GP}}(q)\,.
\end{aligned}$$

These Hamiltonians play a crucial role in Chapter 3, which concerns the well-posedness of the NLS hierarchy with NZBC. In principle, the Hamiltonians $\mathcal{H}_0^{\widetilde{\mathsf{GP}}},\dots,\mathcal{H}_8^{\widetilde{\mathsf{GP}}}$ have previously been used[2] in [27, 28], but to the best of our knowledge[3] the authors were not aware of the "correct" definition for all $n\in\mathbb{N}$. We elaborate on this in Section 4.1.3. They have also appeared in [120, End of §1.3.4].

Using the invertible coefficient matrices $\mathbf{c}_{\mathsf{GP}\mapsto\widetilde{\mathsf{GP}}}, \mathbf{c}_{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} \in \mathbb{R}^{(N+1)\times(N+1)}$ given by

$$\begin{aligned}
\mathbf{c}_{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}^{2n,2m} &= \mathbb{1}_{\{n=m\}} & \mathbf{c}_{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}^{2n,2m-1} &= 0\\
\mathbf{c}_{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}^{2n-1,2m} &= 0 & \mathbf{c}_{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}^{2n-1,2m-1} &= 2^{\mathbb{1}_{\{n=0\}}}\binom{\frac{1}{2}}{n-m}4^{n-m}
\end{aligned}$$

[2]They use the variables $\partial_{\mathrm{x}}\Theta_\varepsilon = 6(W_+ + W_-)$ and $N_\varepsilon = -6(W_+ - W_- + \frac{\varepsilon^2}{4}(W_+ - W_-)^2)$, and their choice of energies and momenta corresponds to ours with $\mathcal{E}_n(N_\varepsilon,\Theta_\varepsilon) = \frac{18}{(\sqrt{2}\varepsilon)^{2n+1}}\mathcal{H}_{2n}^{\varepsilon\widetilde{\mathsf{GP}}}(W)$ and $\mathcal{P}_n(N_\varepsilon,\Theta_\varepsilon) = \sqrt{2}\frac{18}{(\sqrt{2}\varepsilon)^{2n+1}}\mathcal{H}_{2n-1}^{\varepsilon\widetilde{\mathsf{GP}}}(W)$. We have checked this only for $n\in\{1,2,3\}$ due to computational restrictions. Note that (101) in the reference contains a term twice, which we understand to be unintentional.

[3]We base this on [28, Remark 4.7].

and $\mathbf{c}_{\mathsf{GP}\mapsto\widetilde{\mathsf{GP}}} = (\mathbf{c}_{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}})^{-1}$, we can write

$$\left(\frac{\delta\mathcal{H}_n^{\widetilde{\mathsf{GP}}}}{\delta\overline{q}}\right)_{0\le n\le N} = \mathbf{c}_{\mathsf{GP}\mapsto\widetilde{\mathsf{GP}}}\left(\frac{\delta\mathcal{H}_n^{\mathsf{GP}}}{\delta\overline{q}}\right)_{0\le n\le N} \qquad \left(\frac{\delta\mathcal{H}_n^{\mathsf{GP}}}{\delta\overline{q}}\right)_{0\le n\le N} = \mathbf{c}_{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}\left(\frac{\delta\mathcal{H}_n^{\widetilde{\mathsf{GP}}}}{\delta\overline{q}}\right)_{0\le n\le N}.$$

Let us also state at this point a relation between $\mathcal{H}_n^{\mathsf{NLS}}$ and $\mathcal{H}_n^{\mathsf{GP}}$, which is a consequence of (4.2.37)–(4.2.38) below (see also [71, Chapter 1, (10.25)]):

$$\frac{\delta\mathcal{H}_{2n}^{\mathsf{NLS}}(q)}{\delta\overline{q}} = \sum_{m=0}^{n}\binom{n-\frac12}{n-m}4^{n-m}\frac{\delta\mathcal{H}_{2m}^{\mathsf{GP}}(q)}{\delta\overline{q}} \tag{4.1.4}$$

$$\frac{\delta\mathcal{H}_{2n-1}^{\mathsf{NLS}}(q)}{\delta\overline{q}} = \sum_{m=0}^{n}\binom{n-1}{n-m}4^{n-m}\frac{\delta\mathcal{H}_{2m-1}^{\mathsf{GP}}(q)}{\delta\overline{q}}. \tag{4.1.5}$$

Similarly, we have

$$\frac{\delta\mathcal{H}_{2n}^{\mathsf{NLS}}(q)}{\delta\overline{q}} = \sum_{m=0}^{n}\binom{n-\frac12}{n-m}4^{n-m}\frac{\delta\mathcal{H}_{2m}^{\widetilde{\mathsf{GP}}}(q)}{\delta\overline{q}} \tag{4.1.6}$$

$$\frac{\delta\mathcal{H}_{2n-1}^{\mathsf{NLS}}(q)}{\delta\overline{q}} = \sum_{m=0}^{n}\binom{n-\frac12}{n-m}4^{n-m}\frac{\delta\mathcal{H}_{2m-1}^{\widetilde{\mathsf{GP}}}(q)}{\delta\overline{q}}. \tag{4.1.7}$$

We do not state these relations for the Hamiltonians themselves because $\mathcal{H}_n^{\mathsf{NLS}}(q)$ can only be well-defined if q satisfies ZBC, while $\mathcal{H}_n^{\mathsf{GP}}(q)$ and $\mathcal{H}_n^{\widetilde{\mathsf{GP}}}(q)$ require NZBC. We define again invertible coefficient matrices $\mathbf{c}_{\mathsf{NLS}\mapsto\mathsf{GP}}, \mathbf{c}_{\mathsf{GP}\mapsto\mathsf{NLS}}, \mathbf{c}_{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}, \mathbf{c}_{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} \in \mathbb{R}^{(N+1)\times(N+1)}$ by

$$\mathbf{c}_{\mathsf{NLS}\mapsto\mathsf{GP}}^{2n,2m} = \binom{n-\frac12}{n-m}(-4)^{n-m} \qquad \mathbf{c}_{\mathsf{NLS}\mapsto\mathsf{GP}}^{2n,2m-1} = 0$$

$$\mathbf{c}_{\mathsf{NLS}\mapsto\mathsf{GP}}^{2n-1,2m} = 0 \qquad \mathbf{c}_{\mathsf{NLS}\mapsto\mathsf{GP}}^{2n-1,2m-1} = \binom{n-1}{n-m}(-4)^{n-m}$$

$$\mathbf{c}_{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}^{2n,2m} = \binom{n-\frac12}{n-m}(-4)^{n-m} \qquad \mathbf{c}_{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}^{2n,2m-1} = 0$$

$$\mathbf{c}_{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}^{2n-1,2m} = 0 \qquad \mathbf{c}_{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}^{2n-1,2m-1} = \binom{n-\frac12}{n-m}(-4)^{n-m}$$

and $\mathbf{c}_{\mathsf{GP}\mapsto\mathsf{NLS}} = (\mathbf{c}_{\mathsf{NLS}\mapsto\mathsf{GP}})^{-1}$, $\mathbf{c}_{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} = (\mathbf{c}_{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}})^{-1}$.

In analogy to the definition of the NLS hierarchy, we consider the N-truncated GP and $\widetilde{\mathsf{GP}}$ hierarchies, which consist, respectively, of the Hamiltonian flows

$$i\partial_{t_n}q = \frac{\delta\mathcal{H}_n^{\mathsf{GP}}(q)}{\delta\overline{q}} \tag{GP_n}$$

$$i\partial_{t_n}q = \frac{\delta\mathcal{H}_n^{\widetilde{\mathsf{GP}}}(q)}{\delta\overline{q}} \tag{$\widetilde{\mathsf{GP}}_n$}$$

for $n \in \{0,\dots,N\}$. While the GP hierarchy[4] is a more immediate renormalization of the NLS hierarchy for working with NZBC, the $\widetilde{\mathsf{GP}}$ hierarchy is the "better" hierarchy to work with in both Chapter 3 and this work. Note, however, that by Proposition 4.1.11 below the theories of the NLS, GP, and $\widetilde{\mathsf{GP}}$ hierarchies are equivalent, up to the application of explicit transformations.

Let it be said that $\frac{\delta\mathcal{H}_{-1}^{\mathsf{GP}}}{\delta\overline{q}} = 0$ and hence the Hamiltonian flow generated by this conserved quantity is trivial, meaning the asymptotic phase change $\mathcal{H}_{-1}^{\mathsf{GP}}$ is not of particular importance in this work. Nevertheless, Theorem 4.1.1 below *does* require the involvement of $\mathcal{H}_{-1}^{\widetilde{\mathsf{GP}}}$, and its proof motivates our choice of normalization for this conserved quantity, which is different from $\mathcal{H}_{-1}^{\mathsf{GP}}$ that appears in [120].

[4] Not to be confused with the "Gross–Pitaevskii hierarchy" from kinetic theory (see e. g. [70]).

4.1.2. From GP to the hydrodynamic formulation hGP and the long-wave rescaling εGP

The solutions we are interested in have no vacuum, meaning that there exists a constant $\delta > 0$ for which

$$\inf_{x\in\mathbb{R}} |q(\mathbf{t},x)| > \delta\,.$$

As a result, it is possible to apply the so-called Madelung transform, which maps q to the pair (a,v) with

$$a = |q| \qquad v = \operatorname{Im}\left(\frac{q_x}{q}\right) \qquad q = a\exp\left(i\int_0^x v(y)\,\mathrm{d}y\right)e^{i\varphi}\,.$$

Here φ is a uniform-in-x choice of phase, which is information that has to be provided in order to invert the Madelung transform. It is well-known that if q solves GP, then (a,v) solve a hydrodynamic system, which we call the hydrodynamic Gross-Pitaevskii equations:

$$\begin{aligned}\partial_t a + a\partial_x v + 2v\partial_x a &= 0\\ \partial_t v + \partial_x(v^2) + 4a\partial_x a &= \partial_x\left(\frac{\partial_{xx}a}{a}\right).\end{aligned}$$

We prefer to work with the slightly different pair of variables $w = (w_+, w_-)$ where

$$w_\pm = \frac{v}{2} \pm (a-1)\,.$$

We still refer to the system of PDEs that w solves as the hydrodynamic Gross-Pitaevskii equations:

$$\partial_t w_- = -(3w_- + w_+ - 2)\partial_x w_- + \frac{1}{2}\partial_x\left(\frac{\partial_{xx}(w_+ - w_-)}{2a}\right) \tag{hGP-1}$$

$$\partial_t w_+ = -(w_- + 3w_+ + 2)\partial_x w_+ + \frac{1}{2}\partial_x\left(\frac{\partial_{xx}(w_+ - w_-)}{2a}\right). \tag{hGP-2}$$

Of course, this system has an infinite family of conserved energies given by

$$\mathcal{H}_n^{\mathsf{hGP}}(w) = \mathcal{H}_n^{\mathsf{GP}}(q) \qquad \mathcal{H}_n^{\mathsf{h}\widetilde{\mathsf{GP}}}(w) = \mathcal{H}_n^{\widetilde{\mathsf{GP}}}(q)$$

for $n \geq -1$. We define the N-truncated hGP and $\mathsf{h}\widetilde{\mathsf{GP}}$ hierarchies as the hierarchies of Hamiltonian flows

$$\partial_{t_n} w = \frac{1}{4}\begin{pmatrix}-\partial_x a^{-1} - a^{-1}\partial_x & \partial_x a^{-1} - a^{-1}\partial_x\\ -\partial_x a^{-1} + a^{-1}\partial_x & \partial_x a^{-1} + a^{-1}\partial_x\end{pmatrix}\begin{pmatrix}\frac{\delta}{\delta w_+}\mathcal{H}_n^{\mathsf{hGP}}(w)\\ \frac{\delta}{\delta w_-}\mathcal{H}_n^{\mathsf{hGP}}(w)\end{pmatrix} \tag{hGP$_n$}$$

$$\partial_{t_n} w = \frac{1}{4}\begin{pmatrix}-\partial_x a^{-1} - a^{-1}\partial_x & \partial_x a^{-1} - a^{-1}\partial_x\\ -\partial_x a^{-1} + a^{-1}\partial_x & \partial_x a^{-1} + a^{-1}\partial_x\end{pmatrix}\begin{pmatrix}\frac{\delta}{\delta w_+}\mathcal{H}_n^{\mathsf{h}\widetilde{\mathsf{GP}}}(w)\\ \frac{\delta}{\delta w_-}\mathcal{H}_n^{\mathsf{h}\widetilde{\mathsf{GP}}}(w)\end{pmatrix} \tag{h$\widetilde{\mathsf{GP}}_n$}$$

for $n \in \{0,\ldots,N\}$. Note that here $a = 1 + \frac{1}{2}(w_+ - w_-)$, $|a| > \delta$, and a^{-1} should be understood as the operator of multiplication by a^{-1}. In Appendix 4.B.2 we detail the complete Hamiltonian structure of (hGP$_n$), which is chosen so that $q \leftrightarrow w$ is a symplecto-/Poisson morphism. It is indeed the case that (hGP$_1$), (h$\widetilde{\mathsf{GP}}_1$), and (hGP-1)–(hGP-2) are all identical.

For a small $\varepsilon > 0$ we now set

$$X = \sqrt{2}\varepsilon x \qquad T_2 = 2\sqrt{2}\varepsilon t$$

and consider the rescaled variables

$$\begin{aligned}W_\pm(T_2, X) &= \varepsilon^{-2}w_\pm(t,x)\\ A(T_2,X) &= 1 + \frac{\varepsilon^2}{2}(W_+(T_2,X) - W_-(T_2,X)) = a(t,x)\,.\end{aligned}$$

Transforming back to q yields the ansatz (4.1.1) above, up to the use of T_2 instead of the time vector $\mathbf{T}$, and the necessary reconstruction of the uniform-in-x phase function $\varphi(\mathbf{t})$. We refer to the system that $W=(W_+,W_-)$ solves as the ε-rescaled hydrodynamic Gross-Pitaevskii equations:

$$\partial_{T_2}W_+ = -\partial_X W_+ + \frac{\varepsilon^2}{4}\left(-(2W_- + 6W_+)\partial_X W_+ + 2\partial_X\left(\frac{\partial_{XX}(W_+-W_-)}{2A}\right)\right) \qquad (\varepsilon\mathsf{GP}\text{-1})$$

$$\partial_{T_2}W_- = \partial_X W_- + \frac{\varepsilon^2}{4}\left(-(6W_- + 2W_+)\partial_X W_- + 2\partial_X\left(\frac{\partial_{XX}(W_+-W_-)}{2A}\right)\right). \qquad (\varepsilon\mathsf{GP}\text{-2})$$

They have the infinite family of conserved energies

$$\mathcal{H}_n^{\varepsilon\mathsf{GP}}(W) = \mathcal{H}_n^{\mathsf{hGP}}(w) \qquad\qquad \mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}(W) = \mathcal{H}_n^{\mathsf{h}\widetilde{\mathsf{GP}}}(w)$$

for $n\geq -1$, the first of which are

$$\mathcal{H}_{-1}^{\varepsilon\mathsf{GP}}(W) = -\frac{\varepsilon}{\sqrt{2}}\int_{\mathbb{R}} W_+ + W_-\,\mathrm{d}X \qquad \text{(Phase Change)}$$

$$\mathcal{H}_{0}^{\varepsilon\mathsf{GP}}(W) = \frac{1}{\sqrt{2}\varepsilon}\int_{\mathbb{R}} \left(1+\varepsilon^2\frac{W_+-W_-}{2}\right)^2 - 1\,\mathrm{d}X \qquad \text{(Mass)}$$

$$\mathcal{H}_{1}^{\varepsilon\mathsf{GP}}(W) = -\frac{\varepsilon}{\sqrt{2}}\int_{\mathbb{R}} \left(1+\frac{\varepsilon^2}{2}(W_+-W_-)\right)^2 (W_++W_-)\,\mathrm{d}X \qquad \text{(Momentum)}$$

$$\begin{aligned}\mathcal{H}_{2}^{\varepsilon\mathsf{GP}}(W) = \frac{1}{\sqrt{2}\varepsilon}\int &\frac{\varepsilon^6}{2}(\partial_X(W_+-W_-))^2 + \varepsilon^4\left(1+\varepsilon^2\frac{W_+-W_-}{2}\right)^2(W_++W_-)^2 \\ &+\left(\left(1+\varepsilon^2\frac{W_+-W_-}{2}\right)^2 - 1\right)^2\mathrm{d}X\,.\end{aligned} \qquad \text{(Energy)}$$

We define the N-truncated $\varepsilon\mathsf{GP}$ and $\varepsilon\widetilde{\mathsf{GP}}$ hierarchies as the hierarchies of Hamiltonian flows

$$\partial_{T_n}W = \frac{\mathrm{d}t_n}{\mathrm{d}T_n}\frac{1}{2\varepsilon^2}\begin{pmatrix}-\partial_X A^{-1} - A^{-1}\partial_X & \partial_X A^{-1} - A^{-1}\partial_X\\ -\partial_X A^{-1} + A^{-1}\partial_X & \partial_X A^{-1} + A^{-1}\partial_X\end{pmatrix}\begin{pmatrix}\frac{\delta}{\delta W_+}\mathcal{H}_n^{\varepsilon\mathsf{GP}}(W)\\ \frac{\delta}{\delta W_-}\mathcal{H}_n^{\varepsilon\mathsf{GP}}(W)\end{pmatrix} \qquad (\varepsilon\mathsf{GP}_n)$$

$$\partial_{T_n}W = \frac{\mathrm{d}t_n}{\mathrm{d}T_n}\frac{1}{2\varepsilon^2}\begin{pmatrix}-\partial_X A^{-1} - A^{-1}\partial_X & \partial_X A^{-1} - A^{-1}\partial_X\\ -\partial_X A^{-1} + A^{-1}\partial_X & \partial_X A^{-1} + A^{-1}\partial_X\end{pmatrix}\begin{pmatrix}\frac{\delta}{\delta W_+}\mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}(W)\\ \frac{\delta}{\delta W_-}\mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}(W)\end{pmatrix} \qquad (\varepsilon\widetilde{\mathsf{GP}}_n)$$

for $n\in\{0,\dots,N\}$. Here

$$A(\mathbf{T},X) = 1+\frac{\varepsilon^2}{2}(W_+-W_-)(\mathbf{T},X)$$

is the rescaled amplitude. In Appendix 4.B.3 we detail the complete Hamiltonian structure of ($\varepsilon\mathsf{GP}_n$), which is chosen so that $w\leftrightarrow W$ is a symplecto-/Poisson morphism. Recall that the hGP, $\mathsf{h}\widetilde{\mathsf{GP}}$, $\varepsilon\mathsf{GP}$, and $\varepsilon\widetilde{\mathsf{GP}}$ hierarchies are equipped with the zero boundary conditions

$$\lim_{|x|\to\infty} w(\mathbf{t},x) = 0 \qquad\qquad \lim_{|X|\to\infty} W(\mathbf{T},X) = 0\,.$$

4.1.3. Approximation of εGP by transport and KdV equations from [27, 28]

Observe that the leading order behavior of (εGP-1)–(εGP-2) is given by two transport equations with constant velocities ±1. As such, we can think of W_+ and W_- as two largely decoupled waves moving in opposite directions. This was observed and rigorously studied in [25]. The next-to-leading order of the system is reminiscent of the KdV equation

$$\partial_T U = 6UU_X - U_{XXX}\,, \qquad \text{(KdV)}$$

although non-trivial interaction terms such as $W_+\partial_X W_-$ formally persist in the limit $\varepsilon\to0$. In [27, 28] F. Béthuel, P. Gravejat, J.-C. Saut, and D. Smets studied hGP in the long-wave regime described above and proved that the system (εGP-1)–(εGP-2) indeed behaves at leading order like two transport equations, and at next-to-leading order like two KdV equations. Specifically, they proved [27, Theorem 1], which states that for any solution q of GP in a Zhidkov space of sufficient regularity, for which $W(0)$ satisfies additional estimates, there exist two solutions $\pm U_\pm$ of KdV so that

$$\left\| 6W_\pm\left(\frac{8}{\varepsilon^2}\tau, X \mp \frac{8}{\varepsilon^2}\tau\right) - U_\pm(\tau, \pm X)\right\|_{H^s} \leq \varepsilon^2 C(W(0)) e^{C(W(0))|\tau|} \tag{4.1.8}$$

for all $\tau \in \mathbb{R}$ and $s \in \mathbb{N}$. Both the transport equation (KdV_0) and of course KdV = (KdV_1) are part of the KdV hierarchy. Note that after switching to a co-moving frame to remove the leading-order transport behavior, we need to accelerate the timescale by a factor of order ε^2 in order to see the KdV approximation.

It turns out to be the case that for the higher equations in the NLS, GP and $\widetilde{\mathsf{GP}}$ hierarchies, or more precisely the εGP and $\varepsilon\widetilde{\mathsf{GP}}$ hierarchies, the higher equations in the KdV hierarchy show up at leading order as approximating equations. We can demonstrate this by formally expanding

$$A^{-1} = \sum_{k=0}^{\infty}\left(-\frac{\varepsilon^2}{2}(W_+ - W_-)\right)^k$$

and computing the initial orders in powers of ε of the right-hand sides of ($\varepsilon\mathsf{GP}_n$) and ($\varepsilon\widetilde{\mathsf{GP}}_n$). For the flows with even n there is no difference between the εGP and $\varepsilon\widetilde{\mathsf{GP}}$ hierarchies, and we obtain

$$\partial_{T_0} W = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{$\varepsilon\widetilde{\mathsf{GP}}_0$}$$

$$\partial_{T_2} W = \begin{pmatrix} -(W_+)_X \\ (W_-)_X \end{pmatrix} + \frac{\varepsilon^2}{4}\begin{pmatrix} (W_+)_{XXX} - (W_-)_{XXX} - 6(W_+)_X W_+ - 2(W_+)_X W_- \\ (W_+)_{XXX} - (W_-)_{XXX} - 2(W_-)_X W_+ - 6(W_-)_X W_- \end{pmatrix} + O(\varepsilon^4) \tag{$\varepsilon\widetilde{\mathsf{GP}}_2$}$$

$$\begin{aligned}
\partial_{T_4} W = {} & \begin{pmatrix} (W_+)_{XXX} - 6(W_+)_X W_+ \\ -(W_-)_{XXX} - 6(W_-)_X W_- \end{pmatrix} \\
& + \frac{\varepsilon^2}{4}\begin{pmatrix} -(W_+)_{XXXXX} + (W_-)_{XXXXX} + 13(W_+)_{XXX} W_+ + 3(W_+)_{XXX} W_- + 29(W_+)_{XX}(W_+)_X \\ -(W_+)_{XXXXX} + (W_-)_{XXXXX} + 13(W_-)_{XXX} W_- + 3(W_-)_{XXX} W_+ + 29(W_-)_{XX}(W_-)_X \end{pmatrix} \\
& + \frac{\varepsilon^2}{4}\begin{pmatrix} 9(W_-)_{XX}(W_-)_X + 3(W_+)_{XX}(W_-)_X - (W_+)_X(W_-)_{XX} + 3(W_-)_{XXX} W_- - 3(W_-)_{XXX} W_+ \\ 9(W_+)_{XX}(W_+)_X + 3(W_+)_X(W_-)_{XX} - (W_+)_{XX}(W_-)_X + 3(W_+)_{XXX} W_+ - 3(W_+)_{XXX} W_- \end{pmatrix} \\
& + \frac{\varepsilon^2}{4}\begin{pmatrix} -45(W_+)_X W_+^2 - 6(W_+)_X W_+ W_- + 3(W_+)_X W_-^2 \\ 45(W_-)_X W_-^2 + 6(W_-)_X W_+ W_- - 3(W_-)_X W_+^2 \end{pmatrix} + O(\varepsilon^4)\,.
\end{aligned} \tag{$\varepsilon\widetilde{\mathsf{GP}}_4$}$$

Observe that the leading orders correspond precisely to the initial equations (KdV_n) in the KdV hierarchy, and this happens on a purely algebraic level. For the flows with odd n we first look at the εGP hierarchy:

$$\partial_{T_1} W = \begin{pmatrix} (W_+)_X \\ (W_-)_X \end{pmatrix} \tag{$\varepsilon\mathsf{GP}_1$}$$

$$\partial_{T_3} W = \frac{1}{\varepsilon^2}\begin{pmatrix} (W_+)_X \\ (W_-)_X \end{pmatrix} + \begin{pmatrix} -(W_+)_{XXX} + 6(W_+)_X W_+ \\ -(W_-)_{XXX} - 6(W_-)_X W_- \end{pmatrix} + O(\varepsilon^2) \tag{$\varepsilon\mathsf{GP}_3$}$$

$$\begin{aligned}
\partial_{T_5} W = {} & \frac{1}{2\varepsilon^4}\begin{pmatrix} -(W_+)_X \\ -(W_-)_X \end{pmatrix} + \frac{1}{\varepsilon^2}\begin{pmatrix} -(W_+)_{XXX} + 6(W_+)_X W_+ \\ -(W_-)_{XXX} - 6(W_-)_X W_- \end{pmatrix} \\
& + \begin{pmatrix} (W_+)_{XXXXX} - 10(W_+)_{XXX} W_+ - 20(W_+)_{XX}(W_+)_X + 30(W_+)_X W_+^2 \\ (W_-)_{XXXXX} + 10(W_-)_{XXX} W_- + 20(W_-)_{XX}(W_-)_X + 30(W_-)_X W_-^2 \end{pmatrix} + O(\varepsilon^2)\,.
\end{aligned} \tag{$\varepsilon\mathsf{GP}_5$}$$

Here the dynamics at leading order are mixed up. The $\varepsilon\widetilde{\mathsf{GP}}$ hierarchy is a renormalization which separates these dynamics:

$$\partial_{T_1} W = \begin{pmatrix} (W_+)_X \\ (W_-)_X \end{pmatrix} \tag{$\varepsilon\widetilde{\mathsf{GP}}_1$}$$

$$
\begin{aligned}
\partial_{T_3} W &= \begin{pmatrix} -(W_+)_{XXX} + 6(W_+)_X W_+ \\ -(W_-)_{XXX} - 6(W_-)_X W_- \end{pmatrix} + \frac{\varepsilon^2}{4} \begin{pmatrix} -3(W_+)_{XXX} W_+ - 3(W_+)_{XXX} W_- \\ 3(W_-)_{XXX} W_+ + 3(W_-)_{XXX} W_- \end{pmatrix} \qquad (\varepsilon\widetilde{\mathsf{GP}}_3) \\
&\quad + \frac{\varepsilon^2}{4} \begin{pmatrix} -9(W_+)_{XX}(W_+)_X - 3(W_+)_{XX}(W_-)_X + 3(W_-)_{XXX} W_- + 3(W_-)_{XXX} W_+ + 3(W_+)_X W_-^2 \\ 9(W_-)_{XX}(W_-)_X + 3(W_+)_X (W_-)_{XX} - 3(W_+)_{XXX} W_+ - 3(W_+)_{XXX} W_- + 3(W_-)_X W_+^2 \end{pmatrix} \\
&\quad + \frac{\varepsilon^2}{4} \begin{pmatrix} 3(W_+)_X (W_-)_{XX} + 9(W_-)_{XX}(W_-)_X + 15(W_+)_X W_+^2 + 6(W_+)_X W_+ W_- \\ -3(W_+)_{XX}(W_-)_X - 9(W_+)_{XX}(W_+)_X + 15(W_-)_X W_-^2 + 6(W_-)_X W_+ W_- \end{pmatrix} + O(\varepsilon^4) \\
\partial_{T_5} W &= \begin{pmatrix} (W_+)_{XXXXX} - 10(W_+)_{XXX} W_+ - 20(W_+)_{XX}(W_+)_X + 30(W_+)_X W_+^2 \\ (W_-)_{XXXXX} + 10(W_-)_{XXX} W_- + 20(W_-)_{XX}(W_-)_X + 30(W_-)_X W_-^2 \end{pmatrix} + O(\varepsilon^2) \,. \\
&\qquad (\varepsilon\widetilde{\mathsf{GP}}_5)
\end{aligned}
$$

The correspondence between the leading order of the equations in the $\varepsilon\widetilde{\mathsf{GP}}$ hierarchy and the equations in the KdV hierarchy, for all $n \in \mathbb{N}$, is the content of Theorems 4.1.1 and 4.1.8 below, and also our main result Theorem 4.1.12. Let us state clearly at this point that for the case of GP = (GP_2) these theorems *do not* generalize the next-to-leading order KdV approximation result, i. e. (4.1.8). They only generalize the leading order transport approximation result. Note also that the leading order contains no interaction terms, which simplifies the analysis.

This work was inspired by [28, Remarks 1.4, 1.5, 4.7], where the authors comment on their discovery of *"[...] a strong and somewhat striking relationship between the (GP) invariants and the (KdV) invariants"*. Specifically, they noticed that certain linear combinations of the conserved quantities $(\mathcal{H}_k^{\varepsilon\widetilde{\mathsf{GP}}}(W))_{-1 \leq k \leq 2n}$ approximate the energy $\mathcal{E}_{n-1}^{\mathsf{KdV}}(\pm W_\pm)$ up to an error of size $O(\varepsilon^2)$ for $0 \leq n \leq 4$. They subsequently remark that *"It would be of interest to investigate further the relationships between (GP) and (KdV), in particular at the level of the spectral problems associated to the corresponding inverse scattering methods."*, and that *"[the relationship mentioned above] might be extended to higher order (GP) and (KdV) invariants, provided one was first able to compute some expressions for them."*.

We carry this out, extending the relationship between $(\mathcal{H}_k^{\varepsilon\widetilde{\mathsf{GP}}}(W))_{-1 \leq k \leq 2n}$ and $\mathcal{E}_{n-1}^{\mathsf{KdV}}(\pm W_\pm)$ to all $n \in \mathbb{N}$.

4.1.4. APPROXIMATION OF HAMILTONIANS AND TRANSMISSION COEFFICIENTS

Theorem 4.1.1. *All identities below hold up to formal integration by parts. For all $n \in \mathbb{N}$ there exist polynomials $\mathbf{P}_n$, depending on ε, W_+, W_-, and up to $\lfloor \frac{n}{2} \rfloor$ derivatives thereof, such that*

$$
\mathcal{H}_{2n}^{\varepsilon\widetilde{\mathsf{GP}}}(W) \mp 2\mathcal{H}_{2n-1}^{\varepsilon\widetilde{\mathsf{GP}}}(W) = (\sqrt{2}\varepsilon)^{2n+1} \left(\mathcal{E}_{n-1}^{\mathsf{KdV}}(\pm W_\pm) + \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \mathbf{P}_{2n} \mp \mathbf{P}_{2n-1} \, \mathrm{d}X \right) \tag{4.1.9}
$$

In particular

$$
\mathcal{H}_{2n}^{\varepsilon\widetilde{\mathsf{GP}}}(W) = \frac{(\sqrt{2}\varepsilon)^{2n+1}}{2} \left(\mathcal{E}_{n-1}^{\mathsf{KdV}}(-W_-) + \mathcal{E}_{n-1}^{\mathsf{KdV}}(W_+) + \varepsilon^2 \int_{\mathbb{R}} \mathbf{P}_{2n} \, \mathrm{d}X \right) \tag{4.1.10}
$$

$$
\mathcal{H}_{2n-1}^{\varepsilon\widetilde{\mathsf{GP}}}(W) = \frac{(\sqrt{2}\varepsilon)^{2n+1}}{4} \left(\mathcal{E}_{n-1}^{\mathsf{KdV}}(-W_-) - \mathcal{E}_{n-1}^{\mathsf{KdV}}(W_+) + \varepsilon^2 \int_{\mathbb{R}} \mathbf{P}_{2n-1} \, \mathrm{d}X \right) . \tag{4.1.11}
$$

Furthermore, there are no constant or linear monomials in $\mathbf{P}_n$. Lastly, there exist polynomials $\mathbf{Q}_{2n}$, depending on ε, W_+, W_-, and up to $n-1$ derivatives thereof, such that

$$
\int_{\mathbb{R}} \mathbf{P}_{2n} \, \mathrm{d}X = \frac{1}{4} \int_{\mathbb{R}} (\partial_X^n (W_+ - W_-))^2 \, \mathrm{d}X + \int_{\mathbb{R}} \mathbf{Q}_{2n} \, \mathrm{d}X \,. \tag{4.1.12}
$$

Remark 4.1.2. *The initial values of the error polynomials $\mathbf{P}_n$ are as follows:*

$$
\mathbf{P}_{-1} = 0 \qquad \mathbf{P}_0 = \frac{1}{4}(W_+ - W_-)^2 \qquad \mathbf{P}_1 = \frac{1}{4}(-W_+^3 + W_+^2 W_- + W_+ W_-^2 - W_-^3)
$$

$$
\begin{aligned}
\mathbf{P}_2 &= \frac{1}{4}(W_+ - W_-)_X^2 + \frac{1}{4}(3W_+^3 - W_+^2 W_- + W_+ W_-^2 - 3W_-^3) \\
&\quad + \frac{\varepsilon^2}{32}(5W_+^4 - 4W_+^3 W_- - 2W_+^2 W_-^2 - 4W_+ W_-^3 + 5W_-^4)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{P}_3 = &\frac{1}{4}(-(W_+)_X^2(5W_+ + W_-) + 6(W_+)_X(W_-)_X(W_+ + W_-) - (W_-)_X^2(5W_- + W_+)) \\
&+ \frac{1}{4}(-5W_+^4 + 2W_+^3W_- - 2W_+W_-^3 + 5W_-^4) \\
&+ \frac{\varepsilon^2}{32}(-7W_+^5 + 5W_+^4W_- + 2W_+^3W_-^2 + 2W_+^2W_-^3 + 5W_+W_-^4 - 7W_-^5)\,.
\end{aligned}
$$

In order to explain the origin of this result, we need to consider the Lax operators

$$
\begin{aligned}
L^{\varepsilon\mathsf{GP}}(W) &= \begin{pmatrix} -\varepsilon^2 W_- + 1 & i\sqrt{2}\varepsilon\partial_X \\ i\sqrt{2}\varepsilon\partial_X & -\varepsilon^2 W_+ - 1 \end{pmatrix} & L^{\mathsf{KdV}}(\pm W_\pm) &= -\partial_X^2 \pm W_\pm \\
L_0^{\varepsilon\mathsf{GP}} &= \begin{pmatrix} 1 & i\sqrt{2}\varepsilon\partial_X \\ i\sqrt{2}\varepsilon\partial_X & -1 \end{pmatrix} & L_0^{\mathsf{KdV}} &= -\partial_X^2\,.
\end{aligned}
$$

Their relevance is explained in Section 4.2.1, where we give a brief overview of the Lax pair formalisms for the equations under consideration. In the following, we denote by $\mathcal{Q}_j$, $j \in \{1, 2, 3, 4\}$ the four open quadrants of the complex plane. We take two spectral parameters $\lambda \in \overline{\mathcal{Q}_1} \cup \overline{\mathcal{Q}_3}$ and $k \in \overline{\mathcal{Q}_1}$ such that $\lambda = \pm\sqrt{1 + 2\varepsilon^2 k^2}$, and consider the scattering problems associated to $L^{\varepsilon\mathsf{GP}} - \lambda$ and $L^{\mathsf{KdV}} - k^2$. This is discussed in detail in Section 4.2, where we define in (4.2.14) the corresponding transmission coefficients $a^{\varepsilon\mathsf{GP}}(\lambda, k; W)$ and $a^{\mathsf{KdV}}(k; \pm W_\pm)$. They are conserved quantities for every flow in their respective hierarchies and admit for $\lambda \in \overline{\mathcal{Q}_1}$ the following asymptotic expansions in powers of $2k$ at infinity:

$$
\log a^{\varepsilon\mathsf{GP}}(\lambda, k; W) \sim i \sum_{n=-1}^{\infty} \frac{\mathcal{H}_n^{\varepsilon\mathsf{GP}}(W)}{(2k)^{n+1}} \qquad \log a^{\mathsf{KdV}}(k; \pm W_\pm) \sim i \sum_{n=-1}^{\infty} \frac{\mathcal{E}_n^{\mathsf{KdV}}(\pm W_\pm)}{(2k)^{2n+3}}\,. \tag{4.1.13}
$$

The precise sense of expansion is given in Definition 4.2.2 below. The following corollary presents a different perspective on Theorem 4.1.1. Here we define for functions $f = f(\lambda) : \mathbb{C} \to \mathbb{C}$ the even and odd parts

$$
\mathrm{Ev}_\lambda\, f(\lambda) = \frac{f(\lambda) + f(-\lambda)}{2} \qquad \mathrm{Od}_\lambda\, f(\lambda) = \frac{f(\lambda) - f(-\lambda)}{2}\,. \tag{4.1.14}
$$

Proposition 4.1.3. *Let $W = (W_+, W_-) \in \mathcal{S}(\mathbb{R}; \mathbb{R}^2)$, $\varepsilon \in (0, 1)$, and $\lambda, k \in \overline{\mathcal{Q}_1}$ with $|k| > c(W)$ so that Lemma 4.2.3 below is applicable and hence the asymptotic expansions (4.1.13) exist. Then*

$$
\mathrm{Ev}_\lambda\big[\log a^{\varepsilon\mathsf{GP}}(\lambda, k; W)\big] \mp \frac{1}{\lambda}\,\mathrm{Od}_\lambda\big[\log a^{\varepsilon\mathsf{GP}}(\lambda, k; W)\big] \sim \log a^{\mathsf{KdV}}(k; (\pm W_\pm)) + \sum_{n=-1}^{\infty} \frac{O(\varepsilon^2)}{(2k)^{n+1}} \tag{4.1.15}
$$

in the sense of asymptotic expansions as $2ik \to \infty$. Here $O(\varepsilon^2)$ must be understood as in Theorem 4.1.1, i. e. in terms of polynomials in ε, W_+, W_-, and derivatives thereof.

If instead $\lambda \in (0, 1)$, $k \in i\,\mathbb{R}_+$, and $\kappa = -ik < \frac{1}{\sqrt{2}\varepsilon}$ satisfies

$$
\max\left\{\frac{1}{\kappa}, \frac{1}{\kappa^3}\right\}\|W\|_{L^2}^2 \le \frac{1}{30}\,,
$$

then

$$
\left|\mathrm{Ev}_\lambda\big[\log a^{\varepsilon\mathsf{GP}}(\lambda, k; W)\big] \mp \frac{1}{\lambda}\,\mathrm{Od}_\lambda\big[\log a^{\varepsilon\mathsf{GP}}(\lambda, k; W)\big] - \log a^{\mathsf{KdV}}(k; \pm W_\pm)\right| \le 8\varepsilon^2\kappa^{-1}\|W\|_{L^2}^2\,. \tag{4.1.16}
$$

Proof. The asymptotic expansion relation (4.1.15) is a reformulation of Lemma 4.2.7, or alternatively Theorem 4.1.1, involving the identities (4.1.3), (4.2.18), (4.2.19), and (4.2.49). The second approximation result (4.1.16) is taken from Lemma 4.A.6. □

Remark 4.1.4 (Splitting of the Dirac operator in the nonrelativistic limit)**.** *Setting $c = \frac{1}{\varepsilon^2}$, we have*

$$\frac{L^{\varepsilon\mathsf{GP}}}{\varepsilon^2} = c\begin{pmatrix} 0 & \sqrt{2}i\partial_X \\ \sqrt{2}i\partial_X & 0 \end{pmatrix} + c^2\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} -W_- & 0 \\ 0 & -W_+ \end{pmatrix},$$

which is a one-dimensional Dirac operator with speed of light c and mass $m = 1$. On the other hand, $L^{\mathsf{KdV}}(\pm W_\pm)$ is the Schrödinger operator with potential $\pm W_\pm$. The limit $\varepsilon \to 0$ that we are interested in corresponds to the nonrelativistic limit $c \to \infty$, where one expects to recover from the Dirac operator a combination of two Schrödinger operators. The problem of taking this limit is well-studied in mathematical physics. We refer to [171, Chapter 6] for an introduction in a general setting. As an example of such a result, in [79, Corollary 4.1] (see also [181, Theorem 4.1]) it is proven that

$$\left(\frac{L^{\varepsilon\mathsf{GP}} \mp 1 - \lambda}{\varepsilon^2}\right)^{-1} \xrightarrow{\varepsilon\to 0} \mp(L^{\mathsf{KdV}}(\pm W_\pm) + \kappa^2)^{-1}\begin{pmatrix} \frac{1\pm 1}{2} & 0 \\ 0 & \frac{1\mp 1}{2} \end{pmatrix}$$

uniformly strongly for any $\lambda \in \mathbb{C}\setminus\mathbb{R}$. More generally, one may study the asymptotic behavior of the Dirac operator in the nonrelativistic regime by considering various quantities as holomorphic functions of $\frac{1}{c^2}$ around $c = \infty$ and expanding them in a power series. For example, in [79, Theorem 2.1] such a holomorphic representation of the Dirac operator is used to prove the aforementioned result. Most relevant for us is [39, Theorem 4.2], where such a holomorphic representation of the scattering matrix associated to $\frac{L^{\varepsilon\mathsf{GP}}\mp 1}{\varepsilon^2}$ is given, and the initial expansion coefficients are explicitly calculated. In particular, the scattering matrix associated to $L^{\mathsf{KdV}}(\pm W_\pm)$ is recovered in the limit. Since the scattering matrices contain the transmission coefficients $a^{\varepsilon\mathsf{GP}}(\lambda; W)$ and $a^{\mathsf{KdV}}(\kappa, \pm W_\pm)$, this result is in correspondence with (4.1.15), although it is not identical. Admitting that we cannot speak with confidence about this broad field of mathematical physics, we hope that Proposition 4.1.3 may present a novel perspective on the nonrelativistic limit of the one-dimensional Dirac operator.

A consequence of Proposition 4.1.3 is that the energies $\mathcal{E}_n^{\mathsf{KdV}}$ on the right-hand side of (4.1.9) can be used to control the Sobolev norms of $\pm W_\pm$ *individually*, up to an error of size ε^2. This is the content of the following corollary, which is proven in Section 4.2.5 by combining Theorem 4.1.1 with [122, Theorem 9.1]. Here we use the functionals

$$\mathcal{E}_n^{\varepsilon\widetilde{\mathsf{GP}},\pm}(W) = \sum_{l=0}^{n}\binom{n}{l}(\sqrt{2}\varepsilon)^{-2l-3}\left(\mathcal{H}_{2l+2}^{\varepsilon\widetilde{\mathsf{GP}}}(W) \mp 2\mathcal{H}_{2l+1}^{\varepsilon\widetilde{\mathsf{GP}}}(W)\right),$$

which are of course conserved under every flow in the $\varepsilon\mathsf{GP}$ and $\varepsilon\widetilde{\mathsf{GP}}$ hierarchies.

Corollary 4.1.5. *Let $n \in \mathbb{N}$ with $n \geq 1$ and $\varepsilon \in (0,1)$. There exists some $\delta = \delta(n) \in (0,1)$ and $C = C(n) > 0$ such that for all $W \in H^{n+1}$ with $\|W\|_{H^{-1}} < \delta$ we have $\mathcal{E}_n^{\varepsilon\widetilde{\mathsf{GP}},\pm}(W) < \infty$ and*

$$\begin{aligned} \left|\mathcal{E}_n^{\varepsilon\widetilde{\mathsf{GP}},\pm}(W) - \|W_\pm\|_{H^n}^2 - \frac{\varepsilon^2}{8}\|\partial_X(W_+ - W_-)\|_{H^n}^2\right| &\leq C\|W_\pm\|_{H^n}^2\|W_\pm\|_{H^{-1}} \\ &\quad + \varepsilon^2 C\|W\|_{H^n}^2(1 + \|W\|_{H^n})^{n+1}. \end{aligned}$$

Remark 4.1.6. *We use Corollary 4.1.5 to prove the crucial estimates (4.1.23) and (4.1.24) below, which control the size of W globally in time. One may also attempt such estimates from (4.1.22), or more specifically the conserved energies $\sum_{l=0}^{h-1}\tau^{2(h-1-l)}\binom{h-1}{l}\mathcal{H}_{2l+2}^{\mathsf{GP}}$ constructed in [119, 120] for $\tau \geq 2$. They control the size of the solution as measured by the semiclassical Sobolev norm*

$$\|(|q|^2 - 1, \partial_x q)\|_{H_\tau^h} = \left(\int_{\mathbb{R}}(\tau^2 + \xi^2)^h\left(\left|\widehat{|q|^2 - 1}\right|^2 + \left|\widehat{\partial_x q}\right|^2\right)\mathrm{d}\xi\right)^{\frac{1}{2}}.$$

After obtaining an estimate, it is necessary to pass it through a rescaling and then the Madelung transform. The first step turns out to be problematic due to the inhomogeneity causing negative powers of ε to appear. This could be resolved by lifting the restriction $\tau \geq 2$.

4.1.5. Formal approximation of hierarchies

In the previous section we have established an approximative relationship between the Hamiltonians of the $\widetilde{\mathsf{GP}}$ and KdV hierarchies. This suggests a method of proof for the pattern we have observed in Section 4.1.3 to hold for the leading orders of the Hamiltonian PDEs in the $\varepsilon\widetilde{\mathsf{GP}}$ hierarchy. Such an argument requires the symplectic structures of GP and KdV to have a similar approximative relationship, which turns out to be the case. In Appendix 4.B the reader may find a detailed description of the symplectic/Poisson structures of the hierarchies in question. Let us write ($\widetilde{\mathsf{GP}}_n$) in the Hamiltonian formulation

$$i\partial_{t_n}\begin{pmatrix} q \\ \overline{q} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \frac{\delta}{\delta \overline{q}}\mathcal{H}_n^{\widetilde{\mathsf{GP}}}(q) \\ \frac{\delta}{\delta q}\mathcal{H}_n^{\widetilde{\mathsf{GP}}}(q) \end{pmatrix}.$$

Conjugation of the Hamiltonian operator with the map $(q,\overline{q}) \mapsto (w_+, w_-)$ and a careful change of variables for the functional derivative shows that the Hamiltonian formulation of ($\mathsf{h}\widetilde{\mathsf{GP}}_n$) is

$$\partial_{t_n}\begin{pmatrix} w_+ \\ w_- \end{pmatrix} = \frac{1}{4}\begin{pmatrix} -\partial_x a^{-1} - a^{-1}\partial_x & \partial_x a^{-1} - a^{-1}\partial_x \\ -\partial_x a^{-1} + a^{-1}\partial_x & \partial_x a^{-1} + a^{-1}\partial_x \end{pmatrix}\begin{pmatrix} \frac{\delta}{\delta w_+}\mathcal{H}_n^{\mathsf{h}\widetilde{\mathsf{GP}}}(w) \\ \frac{\delta}{\delta w_-}\mathcal{H}_n^{\mathsf{h}\widetilde{\mathsf{GP}}}(w) \end{pmatrix}.$$

Note that here $\partial_x a^{-1}$ and $a^{-1}\partial_x$ are compositions of operators. Then the Hamiltonian formulation of ($\varepsilon\widetilde{\mathsf{GP}}_n$) is

$$\begin{aligned}
\partial_{T_n}\begin{pmatrix} W_+ \\ W_- \end{pmatrix} &= \frac{\mathrm{d}t_n}{\mathrm{d}T_n}\frac{1}{2\varepsilon^2}\begin{pmatrix} -\partial_X A^{-1} - A^{-1}\partial_X & \partial_X A^{-1} - A^{-1}\partial_X \\ -\partial_X A^{-1} + A^{-1}\partial_X & \partial_X A^{-1} + A^{-1}\partial_X \end{pmatrix}\begin{pmatrix} \frac{\delta}{\delta W_+}\mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}(W) \\ \frac{\delta}{\delta W_-}\mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}(W) \end{pmatrix} \\
&= \frac{\mathrm{d}t_n}{\mathrm{d}T_n}\frac{2}{\varepsilon^2}\left(\begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix} + \varepsilon^2\mathbf{E}\right)\begin{pmatrix} \frac{\delta}{\delta W_+}\mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}(W) \\ \frac{\delta}{\delta W_-}\mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}(W) \end{pmatrix},
\end{aligned}$$

where the scaling factor $\frac{\mathrm{d}t_n}{\mathrm{d}T_n}$ was defined in (4.1.2) and

$$\begin{aligned}
\mathbf{E} &= \frac{A^{-1}-1}{2\varepsilon^2}\begin{pmatrix} -\partial_X & 0 \\ 0 & \partial_X \end{pmatrix} + \frac{A_X A^{-2}}{4\varepsilon^2}\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \\
&= -\frac{1}{4}\frac{W_+ - W_-}{1+\frac{\varepsilon^2}{2}(W_+ - W_-)}\begin{pmatrix} -\partial_X & 0 \\ 0 & \partial_X \end{pmatrix} + \frac{1}{8}\frac{(W_+ - W_-)_X}{\left(1+\frac{\varepsilon^2}{2}(W_+ - W_-)\right)^2}\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}.
\end{aligned}$$

Applying Theorem 4.1.1 yields

$$\begin{aligned}
\partial_{T_{2n}}\begin{pmatrix} W_+ \\ W_- \end{pmatrix} &= \frac{\mathrm{d}t_{2n}}{\mathrm{d}T_{2n}}\frac{2}{\varepsilon^2}\left(\begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix} + \varepsilon^2\mathbf{E}\right)\frac{(\sqrt{2}\varepsilon)^{2n+1}}{2} \\
&\quad \times\left(\begin{pmatrix} \frac{\delta}{\delta W_+}\mathcal{E}_{n-1}^{\mathsf{KdV}}(W_+) \\ \frac{\delta}{\delta W_-}\mathcal{E}_{n-1}^{\mathsf{KdV}}(-W_-) \end{pmatrix} + \varepsilon^2\begin{pmatrix} \frac{\delta}{\delta W_+}\int_{\mathbb{R}}\mathbf{P}_{2n}\,\mathrm{d}X \\ \frac{\delta}{\delta W_-}\int_{\mathbb{R}}\mathbf{P}_{2n}\,\mathrm{d}X \end{pmatrix}\right) \\
&= \begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix}\begin{pmatrix} \frac{\delta}{\delta W_+}\mathcal{E}_{n-1}^{\mathsf{KdV}}(W_+) \\ \frac{\delta}{\delta W_-}\mathcal{E}_{n-1}^{\mathsf{KdV}}(-W_-) \end{pmatrix} + \varepsilon^2\mathbf{R}_{2n} \\
\partial_{T_{2n-1}}\begin{pmatrix} W_+ \\ W_- \end{pmatrix} &= \frac{\mathrm{d}t_{2n-1}}{\mathrm{d}T_{2n-1}}\frac{2}{\varepsilon^2}\left(\begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix} + \varepsilon^2\mathbf{E}\right)\frac{(\sqrt{2}\varepsilon)^{2n+1}}{4} \\
&\quad \times\left(\begin{pmatrix} -\frac{\delta}{\delta W_+}\mathcal{E}_{n-1}^{\mathsf{KdV}}(W_+) \\ \frac{\delta}{\delta W_-}\mathcal{E}_{n-1}^{\mathsf{KdV}}(-W_-) \end{pmatrix} + \varepsilon^2\begin{pmatrix} \frac{\delta}{\delta W_+}\int_{\mathbb{R}}\mathbf{P}_{2n-1}\,\mathrm{d}X \\ \frac{\delta}{\delta W_-}\int_{\mathbb{R}}\mathbf{P}_{2n-1}\,\mathrm{d}X \end{pmatrix}\right) \\
&= \begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix}\begin{pmatrix} -\frac{\delta}{\delta W_+}\mathcal{E}_{n-1}^{\mathsf{KdV}}(W_+) \\ \frac{\delta}{\delta W_-}\mathcal{E}_{n-1}^{\mathsf{KdV}}(-W_-) \end{pmatrix} + \varepsilon^2\mathbf{R}_{2n-1}
\end{aligned}$$

with residual terms

$$\mathbf{R}_{2n} = \mathbf{E}\begin{pmatrix} \frac{\delta}{\delta W_+}\mathcal{E}_{n-1}^{\mathsf{KdV}}(W_+) \\ \frac{\delta}{\delta W_-}\mathcal{E}_{n-1}^{\mathsf{KdV}}(-W_-) \end{pmatrix} + \left(\begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix} + \varepsilon^2\mathbf{E}\right)\begin{pmatrix} \frac{\delta}{\delta W_+}\int_{\mathbb{R}}\mathbf{P}_{2n}\,\mathrm{d}X \\ \frac{\delta}{\delta W_-}\int_{\mathbb{R}}\mathbf{P}_{2n}\,\mathrm{d}X \end{pmatrix}$$

$$\mathbf{R}_{2n-1} = \mathbf{E}\begin{pmatrix} -\frac{\delta}{\delta W_+}\mathcal{E}^{\mathsf{KdV}}_{n-1}(W_+) \\ \frac{\delta}{\delta W_-}\mathcal{E}^{\mathsf{KdV}}_{n-1}(-W_-) \end{pmatrix} + \left(\begin{pmatrix} -\frac{1}{2}\partial_X & 0 \\ 0 & \frac{1}{2}\partial_X \end{pmatrix} + \varepsilon^2\mathbf{E}\right)\begin{pmatrix} \frac{\delta}{\delta W_+}\int_{\mathbb{R}}\mathbf{P}_{2n-1}\,\mathrm{d}X \\ \frac{\delta}{\delta W_-}\int_{\mathbb{R}}\mathbf{P}_{2n-1}\,\mathrm{d}X \end{pmatrix}.$$

Remark 4.1.7. *The initial values of the residual terms $\mathbf{R}_n$ are as follows:*

$$\mathbf{R}_0 = \mathbf{R}_1 = 0$$

$$\mathbf{R}_2 = \frac{1}{4}\begin{pmatrix} (W_+)_{XXX} - 6(W_+)_X W_+ - (W_-)_{XXX} - 2(W_+)_X W_- \\ -(W_-)_{XXX} - 6(W_-)_X W_- + (W_+)_{XXX} - 2(W_-)_X W_+ \end{pmatrix} + O(\varepsilon^2)$$

$$\begin{aligned}\mathbf{R}_3 = {}&\frac{1}{4}\begin{pmatrix} -3(W_+)_{XXX}W_+ - 3(W_+)_{XXX}W_- - 9(W_+)_{XX}(W_+)_X - 3(W_+)_{XX}(W_-)_X \\ -3(W_+)_{XXX}W_+ - 3(W_+)_{XXX}W_- - 9(W_+)_{XX}(W_+)_X - 3(W_+)_{XX}(W_-)_X \end{pmatrix} \\ &+ \frac{1}{4}\begin{pmatrix} 3(W_+)_X(W_-)_{XX} + 15(W_+)_X W_+^2 + 6(W_+)_X W_+ W_- + 3(W_+)_X W_-^2 \\ 3(W_+)_X(W_-)_{XX} + 15(W_-)_X W_-^2 + 6(W_-)_X W_+ W_- + 3(W_-)_X W_+^2 \end{pmatrix} \\ &+ \frac{1}{4}\begin{pmatrix} 3(W_-)_{XXX}W_+ + 3(W_-)_{XXX}W_- + 9(W_-)_{XX}(W_-)_X \\ 3(W_-)_{XXX}W_+ + 3(W_-)_{XXX}W_- + 9(W_-)_{XX}(W_-)_X \end{pmatrix} + O(\varepsilon^2)\end{aligned}$$

$$\begin{aligned}\mathbf{R}_4 = {}&\frac{1}{4}\begin{pmatrix} -(W_+)_{XXXXX} + 13(W_+)_{XXX}W_+ + 3(W_+)_{XXX}W_- + 29(W_+)_{XX}(W_+)_X + 3(W_+)_{XX}(W_-)_X \\ -(W_+)_{XXXXX} + 13(W_-)_{XXX}W_- + 3(W_+)_{XXX}W_+ + 29(W_-)_{XX}(W_-)_X + 3(W_+)_X(W_-)_{XX} \end{pmatrix} \\ &+ \frac{1}{4}\begin{pmatrix} -(W_+)_X(W_-)_{XX} - 45(W_+)_X W_+^2 - 6(W_+)_X W_+ W_- + 3(W_+)_X W_-^2 + (W_-)_{XXXXX} \\ -(W_+)_{XX}(W_-)_X + 45(W_-)_X W_-^2 + 6(W_-)_X W_+ W_- - 3(W_-)_X W_+^2 + (W_-)_{XXXXX} \end{pmatrix} \\ &+ \frac{1}{4}\begin{pmatrix} -3(W_-)_{XXX}W_+ + 3(W_-)_{XXX}W_- + 9(W_-)_{XX}(W_-)_X \\ -3(W_+)_{XXX}W_- + 3(W_-)_{XXX}W_+ + 9(W_+)_{XX}(W_+)_X \end{pmatrix} + O(\varepsilon^2)\,.\end{aligned}$$

It is the KdV-like structure of $\mathbf{R}_2$ that is proven in [27, 28] to yield a KdV approximation at next-to-leading order.

We define invertible coefficient matrices $\mathbf{c}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}, \mathbf{c}_{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} \in \mathbb{R}^{(N+1)\times(N+1)}$ by

$$\mathbf{c}^{2n,2m}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}} = \binom{n-\frac{1}{2}}{n-m}\left(-\frac{2}{\varepsilon^2}\right)^{n-m} \qquad \mathbf{c}^{2n,2m-1}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}} = 0$$

$$\mathbf{c}^{2n-1,2m}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}} = 0 \qquad \mathbf{c}^{2n-1,2m-1}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}} = \binom{n-\frac{1}{2}}{n-m}\left(-\frac{2}{\varepsilon^2}\right)^{n-m}$$

and $\mathbf{c}_{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} = (\mathbf{c}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}})^{-1}$. For any of the coefficient matrices $\mathbf{c}_{\dots\mapsto\dots}$ that we use, we define for their transposes the notation

$$\mathbf{c}^{\dots\mapsto\dots} = (\mathbf{c}_{\dots\mapsto\dots})^T \qquad \mathbf{c}^{\dots\mapsto\dots}_{m,n} = \mathbf{c}^{n,m}_{\dots\mapsto\dots}\,.$$

Then with

$$(\nabla_{\mathbf{t}}\mathbf{T}) = \operatorname{diag}\left(\left(\frac{\mathrm{d}T_n}{\mathrm{d}t_n}\right)_{0\le n\le N}\right) \qquad (\nabla_{\mathbf{T}}\mathbf{t}) = \operatorname{diag}\left(\left(\frac{\mathrm{d}t_n}{\mathrm{d}T_n}\right)_{0\le n\le N}\right)$$

we can write

$$\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} = (\nabla_{\mathbf{t}}\mathbf{T})\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}(\nabla_{\mathbf{T}}\mathbf{t}) \qquad \mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}} = (\nabla_{\mathbf{t}}\mathbf{T})\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}(\nabla_{\mathbf{T}}\mathbf{t})$$

$$\mathbf{c}_{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} = (\nabla_{\mathbf{t}}\mathbf{T})\mathbf{c}_{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}(\nabla_{\mathbf{T}}\mathbf{t}) \qquad \mathbf{c}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}} = (\nabla_{\mathbf{T}}\mathbf{t})\mathbf{c}_{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}(\nabla_{\mathbf{t}}\mathbf{T})\,.$$

For later, we define also

$$\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\mathsf{GP}} = (\nabla_{\mathbf{t}}\mathbf{T})\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}(\nabla_{\mathbf{T}}\mathbf{t}) \qquad \mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\varepsilon\mathsf{GP}} = \mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\mathsf{GP}}\,.$$

For the convenience of the reader, we note the following facts: $\mathbf{c}_{\dots\mapsto\dots}$ is always lower triangular and $\mathbf{c}^{\dots\mapsto\dots}$ always upper triangular. Furthermore, the diagonal values are all 1, these matrices are always invertible, and the row and column indices of any nonzero entry have matching parity.

Theorem 4.1.8 (Formal approximation of the $\varepsilon\widetilde{\mathsf{GP}}$ hierarchy by the KdV hierarchy)**.** *Suppose that q is formally a solution of the N-truncated* NLS *hierarchy. If $W(\mathbf{T})$ is defined by (4.1.1), then $W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T})$ solves the N-truncated $\varepsilon\widetilde{\mathsf{GP}}$ hierarchy and*

$$\frac{\mathrm{d}}{\mathrm{d}T_n}W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T}) = \begin{pmatrix} -(-1)^n\frac{1}{2}\partial_X\frac{\delta}{\delta W_+}\mathcal{E}^{\mathsf{KdV}}_{\lfloor\frac{n-1}{2}\rfloor}(W_+(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T})) \\ -\frac{1}{2}\partial_X\frac{\delta}{\delta(-W_-)}\mathcal{E}^{\mathsf{KdV}}_{\lfloor\frac{n-1}{2}\rfloor}(-W_-(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T})) \end{pmatrix} + \varepsilon^2\mathbf{R}_n(\mathbf{T})\,.$$

Correspondingly, $W(\mathbf{T})$ solves

$$\partial_{T_{2n}} W(\mathbf{T}) = \sum_{m=0}^{n} \binom{n-\frac{1}{2}}{n-m} \left(\frac{2}{\varepsilon^2}\right)^{n-m} \left(\begin{pmatrix} -\frac{1}{2}\partial_X \frac{\delta}{\delta W_+} \mathcal{E}^{\mathsf{KdV}}_{m-1}(W_+(\mathbf{T})) \\ -\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)} \mathcal{E}^{\mathsf{KdV}}_{m-1}(-W_-(\mathbf{T})) \end{pmatrix} + \varepsilon^2 \mathbf{R}_{2m}(\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{T}) \right)$$

$$\partial_{T_{2n-1}} W(\mathbf{T}) = \sum_{m=0}^{n} \binom{n-\frac{1}{2}}{n-m} \left(\frac{2}{\varepsilon^2}\right)^{n-m} \left(\begin{pmatrix} \frac{1}{2}\partial_X \frac{\delta}{\delta W_+} \mathcal{E}^{\mathsf{KdV}}_{m-1}(W_+(\mathbf{T})) \\ -\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)} \mathcal{E}^{\mathsf{KdV}}_{m-1}(-W_-(\mathbf{T})) \end{pmatrix} + \varepsilon^2 \mathbf{R}_{2m-1}(\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{T}) \right).$$

Proof. By the chain rule we have

$$\nabla_{\mathbf{T}}(W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T})) = \mathbf{c}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}(\nabla_{\mathbf{T}} W)(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T}).$$

Together with (4.1.6)–(4.1.7) and the definition of $\mathbf{c}_{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}$, this implies that $W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T})$ solves the N-truncated $\varepsilon\widetilde{\mathsf{GP}}$ hierarchy. Then the claimed formulas follow from the elaborations above the theorem. ☐

Remark 4.1.9 (Conjecture on the next-to-leading order of the $\varepsilon\widetilde{\mathsf{GP}}$ hierarchy)**.** *We have verified the following formula to hold for $n \in \{1,2,3,4\}$:*

$$4(\mathbf{R}_{2n}+\mathbf{R}_{2n-1}) = \begin{pmatrix} -\frac{1}{2}\partial_X \frac{\delta}{\delta W_+}\mathcal{E}^{\mathsf{KdV}}_{n}(W_+) + \partial_{XX}\left(\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)}\mathcal{E}^{\mathsf{KdV}}_{n-1}(-W_-)\right) + \textit{interaction terms} + O(\varepsilon^2) \\ O(\varepsilon^0) \end{pmatrix}$$

$$4(\mathbf{R}_{2n}-\mathbf{R}_{2n-1}) = \begin{pmatrix} O(\varepsilon^0) \\ -\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)}\mathcal{E}^{\mathsf{KdV}}_{n}(-W_-) + \partial_{XX}\left(\frac{1}{2}\partial_X \frac{\delta}{\delta W_+}\mathcal{E}^{\mathsf{KdV}}_{n-1}(W_+)\right) + \textit{interaction terms} + O(\varepsilon^2) \end{pmatrix}.$$

Here "interaction terms" are those which contain as factors both W_+ and W_-, or, respectively, some derivatives thereof. The significance of the above formulas is the following: suppose $W(\mathbf{T})$ is a solution to the $\varepsilon\widetilde{\mathsf{GP}}$ hierarchy. Then Theorem 4.1.8 implies

$$\frac{1}{2}\left(\frac{\mathrm{d}}{\mathrm{d}T_{2n}} + \frac{\mathrm{d}}{\mathrm{d}T_{2n-1}}\right) W(\mathbf{T}) = \begin{pmatrix} 0 \\ -\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)}\mathcal{E}^{\mathsf{KdV}}_{n-1}(-W_-(\mathbf{T})) \end{pmatrix}$$
$$+ \frac{\varepsilon^2}{8}\begin{pmatrix} -\frac{1}{2}\partial_X \frac{\delta}{\delta W_+}\mathcal{E}^{\mathsf{KdV}}_{n}(W_+(\mathbf{T})) + \partial_{XX}\left(\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)}\mathcal{E}^{\mathsf{KdV}}_{n-1}(-W_-(\mathbf{T}))\right) + \textit{interaction terms} + O(\varepsilon^2) \\ O(\varepsilon^0) \end{pmatrix}$$
$$\frac{1}{2}\left(\frac{\mathrm{d}}{\mathrm{d}T_{2n}} - \frac{\mathrm{d}}{\mathrm{d}T_{2n-1}}\right) W(\mathbf{T}) = \begin{pmatrix} -\frac{1}{2}\partial_X \frac{\delta}{\delta W_+}\mathcal{E}^{\mathsf{KdV}}_{n-1}(W_+(\mathbf{T})) \\ 0 \end{pmatrix}$$
$$+ \frac{\varepsilon^2}{8}\begin{pmatrix} O(\varepsilon^0) \\ -\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)}\mathcal{E}^{\mathsf{KdV}}_{n}(-W_-(\mathbf{T})) + \partial_{XX}\left(\frac{1}{2}\partial_X \frac{\delta}{\delta W_+}\mathcal{E}^{\mathsf{KdV}}_{n-1}(W_+(\mathbf{T}))\right) + \textit{interaction terms} + O(\varepsilon^2) \end{pmatrix}.$$

This suggests that an extension of Theorem 4.1.8 and Theorem 4.1.12 below to the next-to-leading order is plausible: while ($\varepsilon\widetilde{\mathsf{GP}}_{2n}$) has its leading order behavior given by (KdV_{n-1}), switching to a "co-($\varepsilon\widetilde{\mathsf{GP}}_{2n-1}$)-moving" frame allows us to see, on a longer/slower timescale, the dynamics of (KdV_n) in one component of our solution W. In the case $n = 1$, i. e. that of ($\varepsilon\mathsf{GP}_2$), this corresponds to the utilization of right- or left-moving frames in order to see the dynamics of (KdV_1) in one component of the solution. This is the content of [27]. Interestingly, this suggests the co-moving frame used in this result should be thought of as a renormalization using ($\varepsilon\widetilde{\mathsf{GP}}_1$), instead of ($\mathsf{KdV}_0$), although they are both transport equations.

In the general case, supposing the above formulas can be proven to hold, the term wrapped in two derivatives and the interaction terms, which are a priori not negligible, may still turn out to be of order $O(\varepsilon^2)$ in the energy estimates by the same method that is used in [27]: one component of the solution travels rapidly to infinity due to the dynamics of (KdV_{n-1}) happening on a faster timescale.

4.1.6. Review of well-posedness results for the hierarchies

The introduction of Chapter 3 contains an overview of well-posedness results for NLS, GP, hGP, KdV, and the associated hierarchies. Since this work concerns the hierarchies, we focus our review of the literature on well-posedness results specifically for hierarchies. We start with the KdV hierarchy, where initial existence and uniqueness results were given by J.-C. Saut in [159] and M. Schwarz in [99]. Later, C. E. Kenig, G. Ponce and L. Vega produced a number of works [105, 106, 107, 108, 109, 110, 111], developing dispersive estimates and using them to prove the local well-posedness of various generalized KdV-type hierarchies in weighted Sobolev spaces. Another local well-posedness result in weighted Sobolev spaces for small initial data is given in [152]. This work also contains an ill-posedness result for all of the higher equations in the KdV hierarchy, in the sense of failure of the flow map to be C^2 in the Sobolev spaces $H^s(\mathbb{R})$. Nevertheless, by employing subtle energy estimates and a parabolic regularization argument, C. E. Kenig and D. Pilod [104] proved local well-posedness in $H^s(\mathbb{R})$ at high regularity. We combine this with the conserved quantities that H. Koch and D. Tataru constructed in [122], which prevent the Sobolev norm of a local solution from blowing up in finite time, to formulate the following global well-posedness result. Here and thereafter we use the notation $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^{N+1}$.

Proposition 4.1.10 ([104, Theorem 1.1] and [122, Theorem 9.1])**.** *Let $N \in \mathbb{N}$, $s \geq 4N - \frac{9}{2}$. For every $U_0 \in H^s$ there exists a unique $U \in C(\mathbb{R}^{N+1}; H^s)$ with initial data $U(\mathbf{0}) = U_0$ that satisfies (KdV_n) in the sense of distributions for all $n \in \{0, \dots, N\}$. For every $s' \in [0, s]$ we have*

$$\|U\|_{L^\infty_{\mathbf{T}} H^{s'}_X} \leq C(s')(1 + \|U_0\|_{H^{-1}}^{6s'})\|U_0\|_{H^{s'}} . \tag{4.1.17}$$

Since we shall work with high regularity this formulation is all that we need, but it is far from the state of the art in the literature. In the low regularity regime A. Rybkin [158] and also R. Killip, M. Vişan, and X. Zhang [114, 115] achieved a breakthrough by using a perturbation determinant approach to construct conserved quantities that control the H^{-1}-norm of the solution. This allowed them to prove the global well-posedness of KdV in H^{-1} on the line and the torus, and their "method of commuting flows" has since been applied to several other integrable PDEs (see e. g. [38, 91, 112]). We use the uniform estimates they give for solutions to the KdV hierarchy in the proof of (4.1.17). Recently, the method of commuting flows has been used to prove the global well-posedness of the KdV hierarchy in $H^{-1}(\mathbb{R})$ by F. Klaus, H. Koch, and B. Liu [118].

Moving on to the NLS hierarchy, we mention the work [88] by A. Grünrock where the Fourier restriction norm method is used to prove global well-posedness and ill-posedness results for the flows in the NLS hierarchy, and also the KdV hierarchy, in Fourier-Lebesgue spaces at low regularity. This was extended by J. Adams in [5] to the even flows in the NLS hierarchy, and recently in [4] to the dNLS hierarchy. Note also [149], where orbital stability of the dark soliton solutions to the odd flows (NLS_{2n+1}) of the NLS hierarchy, with real-valued q, is proven.

Of direct relevance is Chapter 3, where the global well-posedness of the NLS hierarchy with NZBC is proven, i. e. precisely the hierarchy we study here. While the local well-posedness is based on the aforementioned methods from C. E. Kenig, G. Ponce, and L. Vega, the globalization argument uses certain conserved energies constructed in [119, 120] by H. Koch and X. Liao. This can be understood as the analogue of [122] in the setting of NZBC. Since this is essential for Proposition 4.1.11 below, we recall now some of their definitions.

We define for $s, s' \in \mathbb{N}$ and $\tau > 0$ the (semi-classical/weighted) Sobolev spaces

$$\big(H^s, \|\cdot\|_{H^s}\big) = \big(\{u \in L^2 : (1+|\xi|^2)^{\frac{s}{2}}\widehat{u} \in L^2\}, \|(1+|\xi|^2)^{\frac{s}{2}}\widehat{\cdot}\|_{L^2}\big) \tag{4.1.18}$$

$$\big(H^s_\tau, \|\cdot\|_{H^s_\tau}\big) = \big(\{u \in L^2 : (\tau^2+|\xi|^2)^{\frac{s}{2}}\widehat{u} \in L^2\}, \|(\tau^2+|\xi|^2)^{\frac{s}{2}}\widehat{\cdot}\|_{L^2}\big) \tag{4.1.19}$$

$$\big(H^{s',1}, \|\cdot\|_{H^{s',1}}\big) = \big(\{u \in H^{s'} : xu \in H^{s'}\}, \|x\cdot\|_{H^{s'}} + \|\cdot\|_{H^{s'}}\big) \tag{4.1.20}$$

and the energy functionals

$$E^s(q) = \||q|^2 - 1\|_{H^{s-1}} + \|q_x\|_{H^{s-1}} \qquad E^{s',1}(q) = \||q|^2 - 1\|_{H^{s'-1,1}} + \|q_x\|_{H^{s'-1,1}} .$$

We recall the complete metric space (X^s, d^s) whose definition is

$$X^s = \{q \in H^s_{loc} : E^{s-1}(q) < \infty\}/\mathbb{S}^1 \qquad d^s(q,p) = \left(\int_{\mathbb{R}} \inf_{\lambda\in\mathbb{S}^1} \|\operatorname{sech}(\cdot - y)(\lambda p - q)\|^2_{H^s}\, \mathrm{d}y\right)^{\frac{1}{2}} . \tag{4.1.21}$$

This space was used by H. Koch and X. Liao in [119, 120] to prove the global well-posedness of GP for $s > 0$. Their globalization argument relies on the aforementioned conserved energies, as they can be used to control the energy functionals $E^s(q)$.

We state now the well-posedness result for the NLS, GP, $\widetilde{\mathsf{GP}}$ and $\varepsilon\widetilde{\mathsf{GP}}$ hierarchies that we use. It is a consequence of Theorem 3.1.2 and the theory used in its proof.

Proposition 4.1.11 (Well-posedness of the NLS, GP, and $\widetilde{\mathsf{GP}}$ hierarchies)**.** *Let $N \geq 2$ and define $\omega = \frac{N-1}{2}$. Let $s, s' \in \mathbb{N}$ such that $s \geq 3N - 1 + \omega$ and $\frac{s+3\omega}{2} \leq s' \leq s - N$. For every $q_0 : \mathbb{R} \to \mathbb{C}$ of the form $q_0 = q_* + p_0$, where $q_* \in H^{s+1-\lceil\omega\rceil+N}_{\mathrm{loc}}$ with $E^{s+1-\lceil\omega\rceil+N,1}(q_*) < \infty$ and $p_0 \in H^s \cap H^{s',1}$, there exists a unique function $q : \mathbb{R}^{N+1} \times \mathbb{R} \to \mathbb{C}$ with $q(\mathbf{0}) = q_0$ satisfying the following:*

(i) *For every $n \in \{0, \dots, N\}$ the functions $q(\mathbf{t}, x)$, $q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t}, x)$, and $q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t}, x)$ solve respectively (NLS_n), (GP_n), and ($\widetilde{\mathsf{GP}}_n$) in the sense of distributions.*

(ii) *We have $q \in C_b(\mathbb{R}^{N+1}; X^s)$ and*

$$\sup_{\mathbf{t}\in\mathbb{R}^{N+1}} E^s(q(\mathbf{t})) \leq C(N, s, E^s(q_0)) . \tag{4.1.22}$$

(iii) *The function $p = q - e^{-i(\mathbf{c}^{\mathsf{GP}\mapsto\mathsf{NLS}}\mathbf{t})_0} q_*$ satisfies $p \in C(\mathbb{R}^{N+1}; H^s \cap H^{s',1})$. Furthermore, $p(\mathbf{t})$, $p(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})$, and $p(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t})$ solve the corresponding perturbative formulations of respectively (NLS_n), (GP_n), and ($\widetilde{\mathsf{GP}}_n$), for the respective fronts $e^{-i(\mathbf{c}^{\mathsf{GP}\mapsto\mathsf{NLS}}\mathbf{t})_0} q_*$, $e^{-it_0} q_*$, and $e^{-it_0} q_*$, strongly in $H^{s-n} \cap H^{s'-n,1}$.*

(iv) *If $\mathcal{H}^{\mathsf{GP}}_2(q_0) < b < \frac{8}{3}$ for some $b > 0$, then there exists $\delta = \delta(b) > 0$ such that*

$$\inf_{\mathbf{t}\in\mathbb{R}^{N+1}, x\in\mathbb{R}} |q(\mathbf{t}, x)| > \delta .$$

In this case we can define for some $\varepsilon > 0$ the function W via the ansatz (4.1.1), and find that $W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T}) \in C(\mathbb{R}^{N+1}; H^{s-1})$ solves ($\varepsilon\widetilde{\mathsf{GP}}_n$) strongly in $H^{s-2\lfloor\frac{n}{2}\rfloor-2}$. Furthermore, for all $h \in \{1, \dots, s-2\}$ there exists some $\varepsilon_0 = \varepsilon_0(h, \|W(\mathbf{0})\|_{H^{h+1}}) > 0$, $c = c(h) > 0$, and $C = C(h) > 0$ such that the following holds. Define

$$\tau = \max\{1, c\|W(\mathbf{0})\|^{\frac{2}{3}}_{H^{h+1}}\} .$$

Then $\varepsilon \in (0, \varepsilon_0)$ implies

$$\begin{aligned}\|W_\pm(\mathbf{T})\|^2_{H^h_\tau} + \frac{\varepsilon^2}{8}\|\partial_X(W_+ - W_-)(\mathbf{T})\|^2_{H^h_\tau} &\leq C\Big(\|W_\pm(\mathbf{0})\|^2_{H^h_\tau} + \frac{\varepsilon^2}{8}\|\partial_X(W_+ - W_-)(\mathbf{0})\|^2_{H^h_\tau}\Big) \\ &\quad + \varepsilon^2\tau^2 C\Big(\|W(\mathbf{0})\|^2_{H^h_\tau} + \frac{\varepsilon^2}{4}\|\partial_X(W_+ - W_-)(\mathbf{0})\|^2_{H^h_\tau}\Big)\end{aligned} \tag{4.1.23}$$

for all $\mathbf{T} \in \mathbb{R}^{N+1}$.

4.1.7. Quantitative approximation of hierarchies

Let us give some justification for why the study of integrable hierarchies is interesting. Due to Noether's theorem, the higher flows in an integrable hierarchy are symmetries of each equation in the hierarchy. Any statement about the NLS/KdV/GP/... hierarchy is therefore also a statement about the NLS/KdV/GP/... equation. These equations are all central to the study of nonlinear PDEs and ubiquitous in the physical modelling of nonlinear waves (see e. g. [23, 24, 92] for applications of NLS and [72, 73, 76, 83, 127, 138] for applications of GP). Besides, integrable hierarchies directly appear in mathematical physics in connection with algebraic geometry [14, 67, 68]. For the reader interested in more algebraic approaches to these integrable hierarchies, we refer to [66, 80, 81]. We are particularly interested in the relevance of these hierarchies in the theory of amplitude and modulation equations.

The transport equation (KdV_0) and the Korteweg-de Vries equation (KdV_1) play an important role as long-wave amplitude equations in the study of one-dimensional nonlinear waves. For example, they have been used in the description of long-wave solutions to the water wave problem [53, 54, 163], a Boussinesq equation [161], a capillary-gravity wave system [164], the NLS equation [48, 160, 47], the GP equation [27, 28], and the Euler-Poisson system [137]. For particular classes of dispersive nonlinear equations, KdV acts as a *universal* long-wave amplitude equation [37] (see also [162, §12.2]). Other examples of long-wave amplitude and modulation equations are NLS, the Ginzburg-Landau equation, and the Whitham modulation equations. For a broad perspective on the theory of such amplitude and modulation equations, we refer to the books [162, 182, 183] and the survey [69]. Many of these amplitude equations are completely integrable, an aspect which has been studied in [100] by L. A. Kalyakin. Another perspective on this phenomenon was given in [42] by F. Calogero, who argued that for certain large classes of PDEs the process of passing to an effective limit equation preserves integrability, and hence the resulting limit equations ought to be integrable if the original class contained any integrable equations. This idea was further developed toward the classification of integrable equations.

Returning specifically to (KdV_1) as a long-wave amplitude equation, we note that it is common for the (trivial transport equation) (KdV_0) to predict the leading order behavior, and (KdV_1) the next-to-leading order behavior. One is therefore tempted to conjecture that the higher equations (KdV_n) in the KdV hierarchy may be used in the description of the higher order asymptotic behavior of long-wave solutions to dispersive nonlinear equations. Yet, such results, and in general direct physical applications of the KdV hierarchy in the modelling of nonlinear waves, are few and far between. In this direction, there exist results by Y. Kodama, T. Taniuti, and Y. Hiraoka [94, 124]. They were used by D. Bambusi [22] to show that, indeed, (KdV_2) shows up after (KdV_0) and (KdV_1) in the aforementioned long-wave approximation of water waves. Furthermore, using the reductive perturbation method of T. Taniuti [169], R. A. Kraenkel, M. A. Manna, and J. G. Pereira [126] discovered that the full KdV hierarchy appears in the study of the asymptotic behavior of long, shallow water waves. Other results in a similar direction are [62, 97]. In [95, Remark 4.6.10] S. Hofbauer noted that a result of her dissertation, using the inverse scattering transform method to prove an approximation result for KdV by a linear Schrödinger equation, can be extended to the higher equations in the KdV hierarchy.

We state now our main result, which shows rigorously that the equations in the KdV hierarchy act as long-wave amplitude equations for the $\widetilde{\mathsf{GP}}$ hierarchy. Recall our notations $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^{N+1}$ and $M = \lfloor \frac{N-1}{2} \rfloor$. Define matrices $\mathbf{b}^+, \mathbf{b}^- \in \mathbb{R}^{(M+1)\times(N+1)}$ by

$$\mathbf{b}^+_{m,n} = -(-1)^n \mathbb{1}_{\{m=\lfloor \frac{n-1}{2} \rfloor\}} \qquad\qquad \mathbf{b}^-_{m,n} = \mathbb{1}_{\{m=\lfloor \frac{n-1}{2} \rfloor\}} .$$

The purpose of this matrix is to match the correct flows in the NLS hierarchy with the corresponding flows in the KdV hierarchy, as described in Theorem 4.1.8.

Theorem 4.1.12 (Quantitative approximation of the $\varepsilon\widetilde{\mathsf{GP}}$ hierarchy by the KdV hierarchy)**.** *Let $N \geq 2$, $s \in \mathbb{N}$, and $q_0 : \mathbb{R} \to \mathbb{C}$ satisfy the assumptions of Proposition 4.1.11 for some $s' \in \mathbb{N}$ and $q_*, p_0 : \mathbb{R} \to \mathbb{C}$. Assume in addition, as in (iv) of Proposition 4.1.11, that $\mathcal{H}_2^{\mathsf{GP}}(q_0) < \frac{8}{3}$ so that $|q(\mathbf{t}, x)| > \delta > 0$ and hence we can define for some $\varepsilon > 0$ the function W according to the ansatz (4.1.1). We have $W \in C(\mathbb{R}^{N+1}; H^{s-1})$, and if $\varepsilon < \varepsilon_0(s, \|W(\mathbf{0})\|_{H^{s-1}})$ then*

$$\|W\|_{L^\infty_{\mathbf{T}} H^{h-1}_X} \leq C(h, \|W(\mathbf{0})\|_{H^h}) \|W(\mathbf{0})\|_{H^h} \qquad \forall\, h \in \{2, \dots, s-1\} . \tag{4.1.24}$$

Furthermore, $W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T}) \in C(\mathbb{R}^{N+1}; H^{s-1})$ solves ($\varepsilon\widetilde{\mathsf{GP}}_n$) in the sense of distributions.

Let $\pm U_\pm \in C(\mathbb{R}^{M+1}; H^{s-1})$ be the solutions to the M-truncated KdV hierarchy, given by Proposition 4.1.10, with initial data $U_\pm(\mathbf{0}) = W_\pm(\mathbf{0})$. They satisfy

$$\|U_\pm\|_{L^\infty_{\mathbf{T}} H^h_X} \leq C(h, \|U_\pm(\mathbf{0})\|_{H^{-1}}) \|U_\pm(\mathbf{0})\|_{H^h} \qquad \forall\, h \in \{0, \dots, s-1\} . \tag{4.1.25}$$

Define $[\mathbf{T}] = |T_2| + \dots + |T_N|$. For all $h \in \{0, \dots, \lfloor \frac{s-N}{2} \rfloor\}$ there exists $C = C(N, h, \delta) > 0$ such that

$$\begin{aligned}\left\| \begin{pmatrix} U_+(\mathbf{b}^+\mathbf{T}) \\ U_-(\mathbf{b}^-\mathbf{T}) \end{pmatrix} - W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T}) \right\|_{H^h} &\leq \varepsilon^2 [\mathbf{T}]^{\frac{1}{2}} \exp\Big([\mathbf{T}] C \big(1 + \|(W,U)\|^M_{L^\infty_{\mathbf{T}} H^{2M}_X}\big)\Big) \\ &\quad \times C\big(1 + \|W\|^{N+4}_{L^\infty_{\mathbf{T}} H^{N+1+h}_X}\big)\big(1 + [\mathbf{T}] \|(W,U)\|^{N-1}_{L^\infty_{\mathbf{T}} H^{2(M+h)}_X}\big)^{\frac{h}{2}} .\end{aligned} \tag{4.1.26}$$

This result represents a collection of approximations, and the timescales involved have several edge cases. Note, for example, that $\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}$ contains negative powers of ε. We shall therefore take some time now to explain this result. It is convenient to introduce for a *time-independent* function $q:\mathbb{R}\to\mathbb{C}$ with $|q|>\delta>0$ the notation

$$[q\mapsto W](q)(X)=\frac{1}{\varepsilon^2}\left(\frac{1}{2}\operatorname{Im}\left[\frac{q_x}{q}\right]+(|q|-1),\frac{1}{2}\operatorname{Im}\left[\frac{q_x}{q}\right]-(|q|-1)\right)(x)\,.$$

This means that if $q:\mathbb{R}^{N+1}\times\mathbb{R}\to\mathbb{C}$ is a solution to the N-truncated NLS hierarchy, and we define $W:\mathbb{R}^{N+1}\times\mathbb{R}\to\mathbb{R}^2$ according to (4.1.1), then the following hold:

$$\begin{aligned}
&q(\mathbf{t})\text{ solves the }N\text{-truncated }\mathsf{NLS}\text{ hierarchy} &&\text{and} && [q\mapsto W](q(\mathbf{t}))=W(\mathbf{T})\\
&q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})\text{ solves the }N\text{-truncated }\mathsf{GP}\text{ hierarchy} &&\text{and} && [q\mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t}))=W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\mathsf{GP}}\mathbf{T})\\
&q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t})\text{ solves the }N\text{-truncated }\widetilde{\mathsf{GP}}\text{ hierarchy} &&\text{and} && [q\mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t}))=W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T})\,.
\end{aligned}$$

We see now that Theorem 4.1.12 is an approximation result written in terms of the $\widetilde{\mathsf{GP}}/\varepsilon\widetilde{\mathsf{GP}}$ hierarchy. We want to sketch the consequences of Theorem 4.1.12 for the equations (NLS_n), (GP_n), and ($\widetilde{\mathsf{GP}}_n$) for $n\in\{0,\dots,7\}$. To keep the notation simple, we shall use the formal statement

$$W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\widetilde{\mathsf{GP}}}\mathbf{T})=\begin{pmatrix}U_+(\mathbf{b}^+\mathbf{T})\\U_-(\mathbf{b}^-\mathbf{T})\end{pmatrix}+\begin{cases}O(\varepsilon^2)\quad\text{as long as }[\mathbf{T}]\lesssim 1 & ,[\mathbf{T}]>0\\ 0\quad\text{for all }\mathbf{T}\in\mathbb{R}^{N+1} & ,[\mathbf{T}]=0\end{cases}$$

to represent the precise statement (4.1.26), which as explained above is an approximation result for the $\widetilde{\mathsf{GP}}$ hierarchy. A corresponding approximation result for the NLS hierarchy reads

$$W(\mathbf{T})=\begin{pmatrix}U_+(\mathbf{b}^+\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{T})\\U_-(\mathbf{b}^-\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{T})\end{pmatrix}+\begin{cases}O(\varepsilon^2)\quad\text{as long as }[\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{T}]\lesssim 1 & ,[\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{T}]>0\\ 0\quad\text{for all }\mathbf{T}\in\mathbb{R}^{N+1} & ,[\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{T}]=0\end{cases},$$

and for the GP hierarchy it is

$$W(\mathbf{c}^{\mathsf{NLS}\mapsto\varepsilon\mathsf{GP}}\mathbf{T})=\begin{pmatrix}U_+(\mathbf{b}^+\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\varepsilon\mathsf{GP}}\mathbf{T})\\U_-(\mathbf{b}^-\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\varepsilon\mathsf{GP}}\mathbf{T})\end{pmatrix}+\begin{cases}O(\varepsilon^2)\quad\text{as long as }[\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\varepsilon\mathsf{GP}}\mathbf{T}]\lesssim 1 & ,[\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\varepsilon\mathsf{GP}}\mathbf{T}]>0\\ 0\quad\text{for all }\mathbf{T}\in\mathbb{R}^{N+1} & ,[\mathbf{c}^{\varepsilon\widetilde{\mathsf{GP}}\mapsto\varepsilon\mathsf{GP}}\mathbf{T}]=0\end{cases}.$$

Here the dependence of various time-related quantities on ε is not clearly visible, so let us write it in a more explicit form. We assume that N is even. The approximation results can be expressed as

$$\begin{aligned}
[q\mapsto W](q(\mathbf{t}))&=\begin{pmatrix}U_+\left(\left((\sqrt{2}\varepsilon)^{2m+1}\left((\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{t})_{2m+1}-2(\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{t})_{2m+2}\right)\right)_{0\le m\le M}\right)\\U_-\left(\left((\sqrt{2}\varepsilon)^{2m+1}\left((\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{t})_{2m+1}+2(\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{t})_{2m+2}\right)\right)_{0\le m\le M}\right)\end{pmatrix}\\
&\quad+\begin{cases}O(\varepsilon^2)\quad\text{as long as }[(\nabla_{\mathbf{t}}\mathbf{T})\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{t}]\lesssim 1 & ,[(\nabla_{\mathbf{t}}\mathbf{T})\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{t}]>0\\ 0\quad\text{for all }\mathbf{T}\in\mathbb{R}^{N+1} & ,[(\nabla_{\mathbf{t}}\mathbf{T})\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}\mathbf{t}]=0\end{cases}
\end{aligned}$$

for the NLS hierarchy,

$$\begin{aligned}
[q\mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t}))&=\begin{pmatrix}U_+\left(\left((\sqrt{2}\varepsilon)^{2m+1}\left((\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}\mathbf{t})_{2m+1}-2(\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}\mathbf{t})_{2m+2}\right)\right)_{0\le m\le M}\right)\\U_-\left(\left((\sqrt{2}\varepsilon)^{2m+1}\left((\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}\mathbf{t})_{2m+1}+2(\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}\mathbf{t})_{2m+2}\right)\right)_{0\le m\le M}\right)\end{pmatrix}\\
&\quad+\begin{cases}O(\varepsilon^2)\quad\text{as long as }[(\nabla_{\mathbf{t}}\mathbf{T})\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}\mathbf{t}]\lesssim 1 & ,[(\nabla_{\mathbf{t}}\mathbf{T})\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}\mathbf{t}]>0\\ 0\quad\text{for all }\mathbf{T}\in\mathbb{R}^{N+1} & ,[(\nabla_{\mathbf{t}}\mathbf{T})\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}}\mathbf{t}]=0\end{cases}
\end{aligned}$$

for the GP hierarchy, and

$$[q\mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t}))=\begin{pmatrix}U_+\left(\left((\sqrt{2}\varepsilon)^{2m+1}\left(t_{2m+1}-2t_{2m+2}\right)\right)_{0\le m\le M}\right)\\U_-\left(\left((\sqrt{2}\varepsilon)^{2m+1}\left(t_{2m+1}+2t_{2m+2}\right)\right)_{0\le m\le M}\right)\end{pmatrix}$$

$$+\begin{cases} O(\varepsilon^2) & \text{as long as } [(\nabla_{\mathbf{t}}\mathbf{T})\mathbf{t}] \lesssim 1 & ,[(\nabla_{\mathbf{t}}\mathbf{T})\mathbf{t}] > 0 \\ 0 & \text{for all } \mathbf{T} \in \mathbb{R}^{N+1} & ,[(\nabla_{\mathbf{t}}\mathbf{T})\mathbf{t}] = 0 \end{cases}$$

for the $\widetilde{\mathsf{GP}}$ hierarchy.

4.1.8. Examples

We use the above formulations to give explicit approximation results along one-dimensional time-rays in $\mathbb{R}^{N+1}$ for the cases $N = 0, \ldots, 7$.

Example 4.1.13 ($(\mathsf{NLS}_0) = (\mathsf{GP}_0) = (\widetilde{\mathsf{GP}}_0) =$ “phase rotation”)**.** Take

$$N = 0 \qquad M = -1 \qquad \mathbf{T} = (2(\sqrt{2}\varepsilon)^{-1}t_0)\,.$$

Then

$$[q \mapsto W](q(\mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t})) = \begin{pmatrix} U_+(0) \\ U_-(0) \end{pmatrix} \quad \text{for all } t_0 \in \mathbb{R}\,.$$

Here we use the convention $\mathbb{R}^{M+1} = \mathbb{R}^0 = \{0\}$.

Example 4.1.14 ($(\mathsf{NLS}_1) = (\mathsf{GP}_1) = (\widetilde{\mathsf{GP}}_1) =$ “transport”)**.** Take

$$N = 1 \qquad M = 0 \qquad \mathbf{T} = (0, \sqrt{2}\varepsilon t_1)\,.$$

Then

$$[q \mapsto W](q(\mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t})) = \begin{pmatrix} U_+(\sqrt{2}\varepsilon t_1) \\ U_-(\sqrt{2}\varepsilon t_1) \end{pmatrix} \quad \text{for all } t_1 \in \mathbb{R}\,.$$

Example 4.1.15 ($(\mathsf{NLS}_2) = \mathsf{NLS}$ and $(\mathsf{GP}_2) = (\widetilde{\mathsf{GP}}_2) = \mathsf{GP}$)**.** Take

$$N = 2 \qquad M = 0 \qquad \mathbf{T} = (0, 0, 2\sqrt{2}\varepsilon t_2)\,.$$

Then

$$\begin{aligned} [q \mapsto W](q(\mathbf{t})) &= \begin{pmatrix} U_+(-2\sqrt{2}\varepsilon t_2) \\ U_-(2\sqrt{2}\varepsilon t_2) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_2| \lesssim \varepsilon^{-1} \\ [q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})) &= \begin{pmatrix} U_+(-2\sqrt{2}\varepsilon t_2) \\ U_-(2\sqrt{2}\varepsilon t_2) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_2| \lesssim \varepsilon^{-1} \\ [q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t})) &= \begin{pmatrix} U_+(-2\sqrt{2}\varepsilon t_2) \\ U_-(2\sqrt{2}\varepsilon t_2) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_2| \lesssim \varepsilon^{-1}\,. \end{aligned}$$

In fact, we have $[q \mapsto W](q(\mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})) = [q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t}))$. This is a consequence of the fact that the difference between these flows for q is one of phase rotation, which does not show up in the variable W.

This result corresponds to [25, Theorem 2], i. e. it is an approximation of GP by two transport equations. As discussed above, it is known from [28, 27] that in the subsequent order $\mathsf{KdV} =$ (KdV_1) shows up as the approximating equation. It is natural to conjecture that this may be generalized for the cases $N \geq 3$, i. e. to characterize the effect of the terms $O(\varepsilon^2)$ up to $O(\varepsilon^4)$. Remark 4.1.9 above provides some evidence towards this.

Example 4.1.16 ($(\mathsf{NLS}_3) = \mathsf{mKdV}$, (GP_3), and $(\widetilde{\mathsf{GP}}_3)$)**.** Take

$$N = 3 \qquad M = 1 \qquad \mathbf{T} = (0, 0, 0, (\sqrt{2}\varepsilon)^3 t_3)\,.$$

Then

$$[q \mapsto W](q(\mathbf{t})) = \begin{pmatrix} U_+(\sqrt{2}\varepsilon\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}_{1,3}t_3, (\sqrt{2}\varepsilon)^3 t_3) \\ U_-(\sqrt{2}\varepsilon\mathbf{c}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}}_{1,3}t_3, (\sqrt{2}\varepsilon)^3 t_3) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_3| \lesssim \varepsilon^{-3}$$

$$[q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})) = \begin{pmatrix} U_+(\sqrt{2}\varepsilon \mathbf{c}_{1,3}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_3, (\sqrt{2}\varepsilon)^3 t_3) \\ U_-(\sqrt{2}\varepsilon \mathbf{c}_{1,3}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_3, (\sqrt{2}\varepsilon)^3 t_3) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_3| \lesssim \varepsilon^{-3}$$

$$[q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t})) = \begin{pmatrix} U_+(0, (\sqrt{2}\varepsilon)^3 t_3) \\ U_-(0, (\sqrt{2}\varepsilon)^3 t_3) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_3| \lesssim \varepsilon^{-3}\,.$$

This is the first new result that Theorem 4.1.12 yields. Here mKdV and renormalized versions of mKdV are approximated by two transport equations in leading order, and then by two KdV equations at next-to-leading order. Observe that for the NLS and GP hierarchies the timescale is ε^{-3}, which is better than the expected ε^{-1}. This is because the approximation of transport equations by transport equations has zero error.

Example 4.1.17 ((NLS$_4$) = "4-th order NLS", (GP$_4$) = ($\widetilde{\mathsf{GP}}_4$) = "4-th order GP")**.** Take

$$N = 4 \qquad M = 1 \qquad \mathbf{T} = (0, 0, 0, 0, 2(\sqrt{2}\varepsilon)^3 t_4)\,.$$

Then

$$[q \mapsto W](q(\mathbf{t})) = \begin{pmatrix} U_+(-2\sqrt{2}\varepsilon \mathbf{c}_{2,4}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_4, -2(\sqrt{2}\varepsilon)^3 t_4) \\ U_-(2\sqrt{2}\varepsilon \mathbf{c}_{2,4}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_4, 2(\sqrt{2}\varepsilon)^3 t_4) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_4| \lesssim \varepsilon^{-1}$$

$$[q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})) = \begin{pmatrix} U_+(0, -2(\sqrt{2}\varepsilon)^3 t_4) \\ U_-(0, 2(\sqrt{2}\varepsilon)^3 t_4) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_4| \lesssim \varepsilon^{-3}$$

$$[q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t})) = \begin{pmatrix} U_+(0, -2(\sqrt{2}\varepsilon)^3 t_4) \\ U_-(0, 2(\sqrt{2}\varepsilon)^3 t_4) \end{pmatrix} + O(\varepsilon^2) \quad \text{as long as } |t_4| \lesssim \varepsilon^{-3}\,.$$

For the NLS hierarchy the case $n = 2$ "dominates", leading to the shorter timescale ε^{-1}. For the $\widetilde{\mathsf{GP}}$ hierarchy the dynamics are cleanly separated, leading to a better timescale. Remark 4.1.9 suggests that studying ($\varepsilon\widetilde{\mathsf{GP}}_4$) in a "co-($\varepsilon\widetilde{\mathsf{GP}}_3$)-moving" frame allows us to see, on a longer timescale, the dynamics of (KdV$_2$) in one component of our solution W.

Example 4.1.18 ((NLS$_5$) = "5-th order mKdV", (GP$_5$), and ($\widetilde{\mathsf{GP}}_5$))**.** Take

$$N = 5 \qquad M = 2 \qquad \mathbf{T} = (0, 0, 0, 0, 0, (\sqrt{2}\varepsilon)^5 t_5)\,.$$

Then

$$\begin{aligned}[q \mapsto W](q(\mathbf{t})) &= \begin{pmatrix} U_+(\sqrt{2}\varepsilon \mathbf{c}_{1,5}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_5, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,5}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_5, (\sqrt{2}\varepsilon)^5 t_5) \\ U_-(\sqrt{2}\varepsilon \mathbf{c}_{1,5}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_5, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,5}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_5, (\sqrt{2}\varepsilon)^5 t_5) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_5| \lesssim \varepsilon^{-3} \\ [q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})) &= \begin{pmatrix} U_+(\sqrt{2}\varepsilon \mathbf{c}_{1,5}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_5, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,5}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_5, (\sqrt{2}\varepsilon)^5 t_5) \\ U_-(\sqrt{2}\varepsilon \mathbf{c}_{1,5}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_5, (\sqrt{2}\varepsilon)^3 \mathbf{c}_{3,5}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_5, (\sqrt{2}\varepsilon)^5 t_5) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_5| \lesssim \varepsilon^{-3} \\ [q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t})) &= \begin{pmatrix} U_+(0, 0, (\sqrt{2}\varepsilon)^5 t_5) \\ U_-(0, 0, (\sqrt{2}\varepsilon)^5 t_5) \end{pmatrix} \\ &\quad + O(\varepsilon^2) \quad \text{as long as } |t_5| \lesssim \varepsilon^{-5}\,. \end{aligned}$$

Here (KdV$_2$) shows up for the first time. Note that once again the approximation of transport equations by transport equations has no error, yielding the better-than-expected timescale ε^{-3} for the NLS and GP hierarchies.

Example 4.1.19 ((NLS$_6$) = "6-th order NLS", (GP$_6$) = ($\widetilde{\mathsf{GP}}_6$) = "6-th order GP")**.** Take

$$N = 6 \qquad M = 2 \qquad \mathbf{T} = (0, 0, 0, 0, 0, 0, 2(\sqrt{2}\varepsilon)^5 t_6)\,.$$

Then

$$[q \mapsto W](q(\mathbf{t})) = \begin{pmatrix} U_+(-2\sqrt{2}\varepsilon \mathbf{c}_{2,6}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_6, -2(\sqrt{2}\varepsilon)^3 \mathbf{c}_{4,6}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_6, -2(\sqrt{2}\varepsilon)^5 t_6) \\ U_-(2\sqrt{2}\varepsilon \mathbf{c}_{2,6}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_6, 2(\sqrt{2}\varepsilon)^3 \mathbf{c}_{4,6}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_6, 2(\sqrt{2}\varepsilon)^5 t_6) \end{pmatrix}$$

$$
\begin{aligned}
&+ O(\varepsilon^2) \quad \text{as long as } |t_6| \lesssim \varepsilon^{-1} \\
[q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})) = &\begin{pmatrix} U_+(0,0,-2(\sqrt{2}\varepsilon)^5 t_6) \\ U_-(0,0,2(\sqrt{2}\varepsilon)^5 t_6) \end{pmatrix} \\
&+ O(\varepsilon^2) \quad \text{as long as } |t_6| \lesssim \varepsilon^{-5} \\
[q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t})) = &\begin{pmatrix} U_+(0,0,-2(\sqrt{2}\varepsilon)^5 t_6) \\ U_-(0,0,2(\sqrt{2}\varepsilon)^5 t_6) \end{pmatrix} \\
&+ O(\varepsilon^2) \quad \text{as long as } |t_6| \lesssim \varepsilon^{-5} \,.
\end{aligned}
$$

Example 4.1.20 ((NLS_7) = "7-th order mKdV", (GP_7), and ($\widetilde{\mathsf{GP}}_7$))**.** Take

$$
N = 7 \qquad\qquad M = 3 \qquad\qquad \mathbf{T} = (0,0,0,0,0,0,0,(\sqrt{2}\varepsilon)^7 t_7)\,.
$$

Then

$$
\begin{aligned}
[q \mapsto W](q(\mathbf{t})) = &\begin{pmatrix} U_+(\sqrt{2}\varepsilon\mathbf{c}_{1,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_7, (\sqrt{2}\varepsilon)^3\mathbf{c}_{3,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_7, (\sqrt{2}\varepsilon)^5\mathbf{c}_{5,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_7, (\sqrt{2}\varepsilon)^7 t_7) \\ U_-(\sqrt{2}\varepsilon\mathbf{c}_{1,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_7, (\sqrt{2}\varepsilon)^3\mathbf{c}_{3,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_7, (\sqrt{2}\varepsilon)^5\mathbf{c}_{5,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{NLS}} t_7, (\sqrt{2}\varepsilon)^7 t_7) \end{pmatrix} \\
&+ O(\varepsilon^2) \quad \text{as long as } |t_7| \lesssim \varepsilon^{-3} \\
[q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\mathsf{GP}}\mathbf{t})) = &\begin{pmatrix} U_+(\sqrt{2}\varepsilon\mathbf{c}_{1,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_7, (\sqrt{2}\varepsilon)^3\mathbf{c}_{3,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_7, (\sqrt{2}\varepsilon)^5\mathbf{c}_{5,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_7, (\sqrt{2}\varepsilon)^7 t_7) \\ U_-(\sqrt{2}\varepsilon\mathbf{c}_{1,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_7, (\sqrt{2}\varepsilon)^3\mathbf{c}_{3,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_7, (\sqrt{2}\varepsilon)^5\mathbf{c}_{5,7}^{\widetilde{\mathsf{GP}}\mapsto\mathsf{GP}} t_7, (\sqrt{2}\varepsilon)^7 t_7) \end{pmatrix} \\
&+ O(\varepsilon^2) \quad \text{as long as } |t_7| \lesssim \varepsilon^{-3} \\
[q \mapsto W](q(\mathbf{c}^{\mathsf{NLS}\mapsto\widetilde{\mathsf{GP}}}\mathbf{t})) = &\begin{pmatrix} U_+(0,0,0,(\sqrt{2}\varepsilon)^7 t_7) \\ U_-(0,0,0,(\sqrt{2}\varepsilon)^7 t_7) \end{pmatrix} \\
&+ O(\varepsilon^2) \quad \text{as long as } |t_7| \lesssim \varepsilon^{-7} \,.
\end{aligned}
$$

At this point we stop because all the edge cases have been exhibited and the pattern continues as expected.

4.1.9. NOTATIONS AND DEFINITIONS

We use $\mathbb{N} = \{n \in \mathbb{Z} : n \geq 0\}$ but note that in some cases where we write $n \in \mathbb{N}$, and when the intended interpretation is clear from the context, $n = 0$ will be an inadmissible value.

We often write $A \leq C(\text{things})B$ to indicate that there exists a constant, depending only on the given parameters, for which the estimate holds. Such constants may change from one line to the next without being given a new name.

Whenever we consider a function space such as H^s without explicit domain, the domain is implicitly assumed to be $\mathbb{R}$. The target space may be $\mathbb{R}$, $\mathbb{C}$, or corresponding spaces of vectors or matrices, and must be determined from context.

We denote by $\mathbf{t} = (t_0, \ldots, t_N), \mathbf{T} = (T_0, \ldots, T_N), \mathbf{0} = (0, \ldots, 0) \in \mathbb{R}^{N+1}$ certain vectors of times. For matrices $\mathbf{c} \in \mathbb{C}^{(N+1)\times(N+1)}$ we denote the entry in the n-th row and m-th column by $\mathbf{c}_{n,m}$. The only exception is the matrix of Jost solutions $\Phi^{\dagger,\pm} = (\Phi_1^{\dagger,\pm}|\Phi_2^{\dagger,\pm})$.

We define the Pauli matrices

$$
\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad\qquad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
$$

For real numbers $x, y \in \mathbb{R}$, we define $x \wedge y = \min\{x, y\}$ and $x \vee y = \max\{x, y\}$, as well as the Japanese bracket $\langle x \rangle = \sqrt{x^2 + 1}$ and the floor and ceiling functions $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}$ and $\lceil x \rceil = \min\{n \in \mathbb{Z} : n \geq x\}$.

For functions $f : \mathbb{R} \to \mathbb{C}^n$ we use the following convention for the Fourier transform:

$$
\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ix\xi} f(x)\,\mathrm{d}x\,.
$$

We extend the Fourier transform to the tempered distributions $\mathcal{S}'(\mathbb{R};\mathbb{C}^n)$ as usual.

We denote by $\mathcal{Q}_j = i^{j-1}\{z \in \mathbb{C} : \Re z > 0, \operatorname{Im} z > 0\}$, $j \in \{1,2,3,4\}$ the four open quadrants of the complex plane.

Acknowledgements. I thank my supervisor Xian Liao for her support and guidance over the past few years, and for many enjoyable and insightful discussions.

Statements and Declarations.
This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173.

No datasets were generated or analysed during the current study.

The author has no competing interests to declare that are relevant to the content of this article.

SymPy [144] was used to carry out symbolic computations for the purposes of exploration and verification of various identities involving recurrence relations. Various versions of the OpenAI services ChatGPT and Codex were used for mathematical and grammatical proofreading, literature search, and exploratory discussion.

4.2. Definition and approximation of the Hamiltonians $\mathcal{H}_n^{\varepsilon\mathsf{GP}}$ by $\mathcal{E}_n^{\mathsf{KdV}}$

4.2.1. Lax pairs

Formally, a function $q : \mathbb{R}^2 \to \mathbb{C}$ is a solution to GP if and only if the time-dependent operators

$$L^{\mathsf{GP}} = \begin{pmatrix} i\partial_x & -iq \\ i\overline{q} & -i\partial_x \end{pmatrix}$$

and

$$P^{\mathsf{GP}} = i\begin{pmatrix} 2\partial_x^2 - q\overline{q} + 1 & -2q\partial_x - q_x \\ 2\overline{q}\partial_x + \overline{q}_x & -2\partial_x^2 + q\overline{q} - 1 \end{pmatrix}$$

solve the so-called Lax equation

$$\partial_t L^{\mathsf{GP}} = [P^{\mathsf{GP}}, L^{\mathsf{GP}}]\,.$$

Here L^{GP} is called the Lax operator of GP and $(L^{\mathsf{GP}}, P^{\mathsf{GP}})$ the Lax pair. Setting $L^{\mathsf{NLS}} = L^{\mathsf{GP}}$ and

$$P^{\mathsf{NLS}} = i\begin{pmatrix} 2\partial_x^2 - q\overline{q} & -2q\partial_x - q_x \\ 2\overline{q}\partial_x + \overline{q}_x & -2\partial_x^2 + q\overline{q} \end{pmatrix}$$

yields a Lax pair for NLS. A direct consequence of the Lax pair formalism is that eigenvalues of the Lax operator are conserved quantities. A deeper consequence is that the potential q can be reconstructed from the scattering data associated to the Lax operator L^{GP}, and this scattering data undergoes a linear evolution in time. As such, one can formally linearize NLS or GP by conjugating the flow with the direct and inverse scattering transforms. This is called the inverse scattering transform (IST) method, and it is the reason why NLS and GP are considered to be completely integrable PDEs. Note that understanding the mapping properties of the IST on standard function spaces is non-trivial, especially for nonzero boundary conditions such as in the case of GP. For NLS and GP the scattering data is obtained from the Zakharov-Shabat scattering problem, which is given as the cases $\dagger \in \{\mathsf{NLS}, \mathsf{GP}\}$ in Definition 4.2.1 below. Since the IST method is not the focus of this work, we shall merely highlight here the seminal papers [2, 189] and refer to the recent review [155] regarding the development of the IST method, with a focus on GP. We emphasize the connection to Section 3.2, where the theory that we use here is set up, with matching notation, and more background on complete integrability and the IST method is given.

We shall now give Lax pairs for hGP, $\varepsilon\mathsf{GP}$ and KdV. Set $q = ae^{i\varphi}$ and $\theta = \frac{i}{2}(\varphi - \frac{\pi}{2})$. In [134] X. Liao and M. Plum note that the explicit unitary matrix

$$M = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-\theta} & e^{\theta} \\ e^{-\theta} & -e^{\theta} \end{pmatrix} \qquad\qquad M^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{\theta} & e^{-\theta} \\ e^{\theta} & -e^{-\theta} \end{pmatrix}$$

yields the unitary equivalence

$$L^{\mathsf{GP}} = M^{-1} L^{\mathsf{hGP}} M \,,$$

where

$$L^{\mathsf{hGP}} = \begin{pmatrix} -w_- + 1 & i\partial_x \\ i\partial_x & -w_+ - 1 \end{pmatrix}.$$

Together with

$$P^{\mathsf{hGP}} = \begin{pmatrix} -\frac{1}{2}(3w_- + w_+)_x - (3w_- + w_+ - 2)\partial_x & 2i\partial_x^2 - \frac{i}{2}\frac{(w_+ - w_-)_{xx}}{w_+ - w_- + 2} \\ 2i\partial_x^2 - \frac{i}{2}\frac{(w_+ - w_-)_{xx}}{w_+ - w_- + 2} & -\frac{1}{2}(w_- + 3w_+)_x - (w_- + 3w_+ + 2)\partial_x \end{pmatrix}$$

this forms a Lax pair for the hydrodynamic formulation hGP of GP.

By rescaling, we can construct from $(L^{\mathsf{hGP}}, P^{\mathsf{hGP}})$ the Lax pair

$$L^{\varepsilon\mathsf{GP}} = \begin{pmatrix} -\varepsilon^2 W_- + 1 & i\sqrt{2}\varepsilon\partial_X \\ i\sqrt{2}\varepsilon\partial_X & -\varepsilon^2 W_+ - 1 \end{pmatrix}$$

and

$$P^{\varepsilon\mathsf{GP}} = \begin{pmatrix} P^{\varepsilon\mathsf{GP}}_{1,1} & P^{\varepsilon\mathsf{GP}}_{1,2} \\ P^{\varepsilon\mathsf{GP}}_{2,1} & P^{\varepsilon\mathsf{GP}}_{2,2} \end{pmatrix} \text{ where } \begin{pmatrix} P^{\varepsilon\mathsf{GP}}_{1,1} \\ P^{\varepsilon\mathsf{GP}}_{1,2} \\ P^{\varepsilon\mathsf{GP}}_{2,1} \\ P^{\varepsilon\mathsf{GP}}_{2,2} \end{pmatrix} = \begin{pmatrix} -\frac{\varepsilon}{\sqrt{2}}(3\varepsilon^2 W_- + \varepsilon^2 W_+)_X - (3\varepsilon^2 W_- + \varepsilon^2 W_+ - 2)\sqrt{2}\varepsilon\partial_X \\ 4i\varepsilon^2\partial_X^2 - i\varepsilon^2\frac{(\varepsilon^2 W_+ - \varepsilon^2 W_-)_{XX}}{\varepsilon^2 W_+ - \varepsilon^2 W_- + 2} \\ 4i\varepsilon^2\partial_X^2 - i\varepsilon^2\frac{(\varepsilon^2 W_+ - \varepsilon^2 W_-)_{XX}}{\varepsilon^2 W_+ - \varepsilon^2 W_- + 2} \\ -\frac{\varepsilon}{\sqrt{2}}(\varepsilon^2 W_- + 3\varepsilon^2 W_+)_X - (\varepsilon^2 W_- + 3\varepsilon^2 W_+ + 2)\sqrt{2}\varepsilon\partial_X \end{pmatrix}$$

for $\varepsilon\mathsf{GP}$. Lastly, we recall the well-known Lax pair

$$\widetilde{L}^{\mathsf{KdV}} = -\partial_X^2 + U \qquad\qquad P^{\mathsf{KdV}} = -4\partial_X^3 + 3U_X + 6U\partial_X$$

for KdV. For the sake of uniformity we treat the scattering problem for KdV as a first-order system, i. e. we use instead of $\widetilde{L}^{\mathsf{KdV}}$ the Lax operator

$$L^{\mathsf{KdV}} = \begin{pmatrix} -\partial_X^2 + U & 0 \\ U_X & -\partial_X^2 + U \end{pmatrix}.$$

We are interested in these Lax pairs because the conservation of the spectral data of the Lax operator implies the existence of infinitely many conserved quantities. More specifically, we shall construct in the subsequent section the Hamiltonians $\mathcal{H}^{\mathsf{NLS}}_n, \mathcal{H}^{\mathsf{GP}}_n, \mathcal{H}^{\mathsf{hGP}}_n, \mathcal{H}^{\varepsilon\mathsf{GP}}_n$, and $\mathcal{E}^{\mathsf{KdV}}_n$ from the transmission coefficients associated to each scattering problem.

4.2.2. JOST SOLUTIONS AND THEIR ASYMPTOTIC EXPANSION

The Jost solutions for each scattering problem depend on a complex parameter $\lambda \in \mathbb{C}$. We prefer to use for GP and hGP the alternative variable $z = \sqrt{\lambda^2 - 1}$, and for $\varepsilon\mathsf{GP}$ and KdV the variable $k = \frac{z}{\sqrt{2}\varepsilon}$. Due to the multivaluedness of the relation $\lambda \leftrightarrow z$, it is common to consider pairs $(\lambda, z) \in \mathbb{C}^2$ as lying on a Riemann surface. Since we also consider the additional parameter k, we define

$$\mathbb{K} = \{(\lambda, z, k) \in \mathbb{C}^3 : \lambda^2 - z^2 = 1 \text{ and } z = \sqrt{2}\varepsilon k\}\,. \tag{4.2.1}$$

We decompose $\mathbb{K}$ into two sheets $\mathbb{K} = \overline{\mathbb{K}_+} \cup \overline{\mathbb{K}_-}$ where

$$\mathbb{K}_\pm = \{(\lambda, z, k) \in \mathbb{K} : \lambda \in \mathbb{C} \setminus ((-\infty, -1] \cup [1, \infty)), \pm \operatorname{Im} z > 0\}\,.$$

By an abuse of notation, we may write $\lambda \in \overline{\mathbb{K}_\pm}$, $z \in \overline{\mathbb{K}_\pm}$, or $k \in \overline{\mathbb{K}_\pm}$ to mean $(\lambda, z, k) \in \overline{\mathbb{K}_\pm}$. Similarly, for functions f with domain $\overline{\mathbb{K}_\pm}$ we write $f(\lambda) = f(z) = f(k) = f(\lambda, z, k)$.

In order to give a uniform treatment of all the scattering problems under consideration, we define

$$\mathsf{Eqns} = \{\mathsf{NLS}, \mathsf{GP}, \mathsf{hGP}, \varepsilon\mathsf{GP}, \mathsf{KdV}\}$$

and write $\dagger \in \mathsf{Eqns}$ when we wish for $\dagger$ to refer to any of the equations under consideration. Using this notation, we shall characterize all five scattering problems under consideration using a single set of variables:

$$(q^\dagger, q_\pm^\dagger, x^\dagger, \lambda^\dagger, z^\dagger) = \begin{cases} (q, 0, x, \lambda, \lambda) & , \dagger = \mathsf{NLS} \\ (q, q_\pm, x, \lambda, z) & , \dagger = \mathsf{GP} \\ (w, 0, x, \lambda, z) & , \dagger = \mathsf{hGP} \\ (W, 0, X, \lambda, k) & , \dagger = \varepsilon\mathsf{GP} \\ (U, 0, X, k^2, k) & , \dagger = \mathsf{KdV} \end{cases} .$$

In the following we always take $j \in \{1, 2\}$ and set $\gamma_1 = 1$ and $\gamma_2 = -1$.

Definition 4.2.1 (Jost solutions of $\dagger$)**.** Given a potential $q^\dagger$ with $\lim_{x \to \pm\infty} q^\dagger = q_\pm^\dagger$, we define the Jost solution $\Phi_j^{\dagger,\pm}$ as the solution to the scattering problem

$$\Phi_j^{\dagger,\pm} : \mathbb{R} \times \overline{\mathbb{K}_{\mp\gamma_j}} \longrightarrow \mathbb{C}^2$$

$$L^\dagger\, \Phi_j^{\dagger,\pm}(x^\dagger, z^\dagger; q^\dagger) = \lambda^\dagger\, \Phi_j^{\dagger,\pm}(x^\dagger, z^\dagger; q^\dagger) \tag{4.2.2}$$

$$\lim_{x \to \pm\infty} \Phi_j^{\dagger,\pm}(x^\dagger, z^\dagger; q^\dagger) e^{x^\dagger i z^\dagger \gamma_j} = E_j^{\dagger,\pm}(z^\dagger)\, . \tag{4.2.3}$$

We define in addition the modified Jost solutions as $\Psi_j^{\dagger,\pm} = \Phi_j^{\dagger,\pm} e^{x^\dagger i z^\dagger \gamma_j}$. The eigenvalue problem (4.2.2) is equivalent to

$$\partial_{x^\dagger} \Phi_j^{\dagger,\pm}(x^\dagger, z^\dagger; q^\dagger) = A^\dagger(x^\dagger, z^\dagger; q^\dagger)\, \Phi_j^{\dagger,\pm}(x^\dagger, z^\dagger; q^\dagger) \tag{4.2.4}$$

$$\partial_{x^\dagger} \Psi_j^{\dagger,\pm}(x^\dagger, z^\dagger; q^\dagger) = (A^\dagger(x^\dagger, z^\dagger; q^\dagger) + i z^\dagger \gamma_j)\, \Psi_j^{\dagger,\pm}(x^\dagger, z^\dagger; q^\dagger) \tag{4.2.5}$$

The matrices $A^\dagger(x^\dagger, z^\dagger; q^\dagger), E^{\dagger,\pm}(z^\dagger) \in \mathbb{C}^{2\times 2}$ are defined in the following table:

$\dagger$	$A^\dagger$	$E^{\dagger,\pm}$
NLS	$\begin{pmatrix} -i\lambda & q \\ \bar{q} & i\lambda \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
GP	$\begin{pmatrix} -i\lambda & q \\ \bar{q} & i\lambda \end{pmatrix}$	$\begin{pmatrix} q_\pm & i(z-\lambda) \\ i(\lambda - z) & \bar{q}_\pm \end{pmatrix}$
hGP	$\begin{pmatrix} 0 & -i(\lambda + 1 + w_+) \\ -i(\lambda - 1 + w_-) & 0 \end{pmatrix}$	$\frac{i}{\sqrt{2}} \begin{pmatrix} \lambda - z + 1 & -\lambda + z - 1 \\ -\lambda + z + 1 & -\lambda + z + 1 \end{pmatrix} \begin{pmatrix} e^{\theta_\pm} & 0 \\ 0 & e^{-\theta_\pm} \end{pmatrix}$
$\varepsilon\mathsf{GP}$	$\begin{pmatrix} 0 & -\frac{i\varepsilon}{\sqrt{2}}\left(\frac{\lambda+1}{\varepsilon^2} + W_+\right) \\ -\frac{i\varepsilon}{\sqrt{2}}\left(\frac{\lambda-1}{\varepsilon^2} + W_-\right) & 0 \end{pmatrix}$	$\frac{i}{\sqrt{2}} \begin{pmatrix} \lambda - \sqrt{2}\varepsilon k + 1 & -\lambda + \sqrt{2}\varepsilon k - 1 \\ -\lambda + \sqrt{2}\varepsilon k + 1 & -\lambda + \sqrt{2}\varepsilon k + 1 \end{pmatrix} \begin{pmatrix} e^{\theta_\pm} & 0 \\ 0 & e^{-\theta_\pm} \end{pmatrix}$
KdV	$\begin{pmatrix} 0 & 1 \\ U - k^2 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 \\ -ik & ik \end{pmatrix}$

Here $\theta_\pm = \frac{i}{2}(\varphi_\pm - \frac{\pi}{2})$ where $e^{i\varphi_\pm} = q_\pm$. We also define $A^{\dagger,\pm}(z^\dagger; q^\dagger) = \lim_{x^\dagger \to \pm\infty} A^\dagger(x^\dagger, z^\dagger; q^\dagger)$. Observe that $E^{\dagger,\pm}$ is indeed a solution to (4.2.5) at $x^\dagger = \pm\infty$.

Our choice of normalization for these eigenvalue problems yields the identity

$$\Phi_j^{\mathsf{GP},\pm}(x, z) = M(x) \Phi_j^{\mathsf{hGP},\pm}(x, z) = M\left(\frac{X}{\sqrt{2}\varepsilon}\right) \Phi_j^{\varepsilon\mathsf{GP},\pm}(X, k)\, . \tag{4.2.6}$$

4.2.3. Transmission coefficients and their asymptotic expansion

We recall Definition 3.2.4.

Definition 4.2.2. Let $D \subset \mathbb{C}$, $d \in \mathbb{N}$ and $1 \leq p \leq \infty$.

(i) We say that a function $f = f(z) \in C^\infty(D; \mathbb{C}^d)$ has an **asymptotic expansion in powers of $2iz$ at infinity on D** if there exist $(f_n)_{n\in\mathbb{N}} \subset \mathbb{C}^d$ such that

$$\forall N \in \mathbb{N} \quad \lim_{|z|\to\infty} (2iz)^N \left(f(z) - \sum_{n=0}^{N} \frac{f_n}{(2iz)^n} \right) = 0\,.$$

We call f_n the **expansion coefficients** of f and note that they are unique. If f and g have such an expansion then $2izf$ and fg do as well. If $g(z) \neq 0$ for $|z|$ sufficiently large, then fg^{-1} also has such an expansion.

(ii) We say that a function $f = f(x,z) \in C^\infty(\mathbb{R} \times D; \mathbb{C}^d)$ has an **L^p-smooth asymptotic expansion in powers of $2iz$ at infinity on D** if there exist $(f_n)_{n\in\mathbb{N}} \subset (C^\infty \cap L^p)(\mathbb{R}; \mathbb{C}^d)$ such that

$$\forall\, k, N \in \mathbb{N} \quad \lim_{|z|\to\infty} \left\| (2iz)^N \left(\partial_x^k f(x,z) - \sum_{n=0}^{N} \frac{\partial_x^k f_n(x)}{(2iz)^n} \right) \right\|_{L^p_x(\mathbb{R})} = 0\,.$$

In this case $\partial_x f$ has such an expansion as well. If g has such an L^∞-smooth expansion and $\inf_{x\in\mathbb{R}} |g(x,z)| > 0$ for $|z|$ sufficiently large, then fg^{-1} also has an L^p-smooth asymptotic expansion. If $p = 1$ then $\int_{-\infty}^{\infty} f(x,z)\,\mathrm{d}x$ has an asymptotic expansion in powers of $2iz$ at infinity on D with expansion coefficients $\int_{-\infty}^{\infty} f_n(x)\,\mathrm{d}x$.

When f has an asymptotic expansion of some kind with expansion coefficients $(f_n)_{n\in\mathbb{N}}$, we write

$$f \sim \sum_{n=0}^{\infty} \frac{f_n}{(2iz)^n}\,.$$

We collect now a number of results on the Jost solutions which are known in the literature, although rarely arranged in this form. For the rest of the section we focus only on $\overline{\mathbb{K}_+}$.

Lemma 4.2.3. *Let $q^\dagger \in C^\infty(\mathbb{R})$ with $q^\dagger - q^\dagger_\pm \in \mathcal{S}(\mathbb{R}_\pm)$.*

(i) *The Jost solutions $\Phi_j^{\dagger,\pm}$ exist and are smooth in $x^\dagger$ and analytic in $z^\dagger$ up to a finite set of choices for $z^\dagger$.*

(ii) *There exists a constant $c(q^\dagger) > 0$ such that $\operatorname{Im} z^\dagger > c(q^\dagger)$ implies*

$$\text{the limits } \lim_{x^\dagger\to\infty} \Psi_1^{\dagger,-}(x^\dagger, z^\dagger),\ \lim_{x^\dagger\to-\infty} \Psi_2^{\dagger,+}(x^\dagger, z^\dagger) \text{ exist} \tag{4.2.7}$$

$$\lim_{x^\dagger\to\infty} \partial_{x^\dagger} \Psi_1^{\dagger,-}(x^\dagger, z^\dagger) = \lim_{x^\dagger\to-\infty} \partial_{x^\dagger} \Psi_2^{\dagger,+}(x^\dagger, z^\dagger) = 0 \tag{4.2.8}$$

$$\sup_{x^\dagger\in\mathbb{R}} \left| \Psi_1^{\dagger,-}(x^\dagger, z^\dagger) - E_1^{\dagger,-}(z^\dagger) \right| < \frac{|E_1^{\dagger,-}(z^\dagger)|}{c(q^\dagger)|z^\dagger|} \tag{4.2.9}$$

$$\sup_{x^\dagger\in\mathbb{R}} \left| \Psi_2^{\dagger,+}(x^\dagger, z^\dagger) - E_2^{\dagger,+}(z^\dagger) \right| < \frac{|E_2^{\dagger,+}(z^\dagger)|}{c(q^\dagger)|z^\dagger|}\,. \tag{4.2.10}$$

(iii) *Define on $\overline{\mathbb{K}_+} \cap \{\operatorname{Im} z^\dagger > c, \pm \operatorname{Im} \lambda > 0\}$ the quantities*

$$\sigma^\dagger = \frac{\partial_{x^\dagger} \Psi_{1,1}^{\dagger,-}}{\Psi_{1,1}^{\dagger,-}} \qquad \widetilde{\sigma}^\dagger = -\frac{\partial_{x^\dagger} \Psi_{2,2}^{\dagger,+}}{\Psi_{2,2}^{\dagger,+}} \tag{4.2.11}$$

$$r^\dagger = \frac{\Psi_{1,2}^{\dagger,-}}{\Psi_{1,1}^{\dagger,-}} \qquad \widetilde{r}^\dagger = \frac{\Psi_{2,1}^{\dagger,+}}{\Psi_{2,2}^{\dagger,+}}\,. \tag{4.2.12}$$

They are bounded and have L^∞-smooth asymptotic expansions in powers of $2iz^\dagger$ at infinity. The densities $\sigma^\dagger$ and $\widetilde{\sigma}^\dagger$ are in addition integrable, and their asymptotic expansions are also L^1-smooth.

Proof. Let us first discuss the case of GP. Here (i) and (ii) can be proven by the well-known method of writing the boundary value problem as a Volterra integral equation and using a Neumann series ansatz. For a demonstration we refer to [64, Proposition 3], and also the theory in [71, Chapter 1]. We refer also to Section 3.2 for a treatment of these matters with a different approach, and specifically Lemma 3.2.5 as a reference for (iii). These same methods are strictly easier to apply for NLS and KdV because the potential under consideration has zero boundary data and the structure of the scattering problem is similar. For more information on the direct scattering theory of KdV we refer to [1]. Using (4.2.6), the cases hGP and εGP can be obtained directly from GP. $\square$

If $\operatorname{Im} z^\dagger = 0$ then $\Phi^{\dagger,-} = (\Phi_1^{\dagger,-}, \Phi_2^{\dagger,-})$ and $\Phi^{\dagger,+} = (\Phi_1^{\dagger,+}, \Phi_2^{\dagger,+})$ are two fundamental solution matrices to (4.2.4), hence there exist $a^\dagger(z^\dagger;q^\dagger), b^\dagger(z^\dagger;q^\dagger) \in \mathbb{C}$ such that

$$\Phi_1^{\dagger,-}(x^\dagger,z^\dagger) = a^\dagger(z^\dagger;q^\dagger)\Phi_1^{\dagger,+}(x^\dagger,z^\dagger) + b^\dagger(z^\dagger;q^\dagger)\Phi_2^{\dagger,+}(x^\dagger,z^\dagger)\,. \tag{4.2.13}$$

We call $a^\dagger(z^\dagger;q^\dagger)$ the transmission coefficient of $\dagger$ with potential $q^\dagger$. It satisfies

$$a^\dagger = \frac{\det(\Phi_1^{\dagger,-}|\Phi_2^{\dagger,+})}{\det E^{\dagger,+}} = \frac{\Phi_{1,1}^{\dagger,-}\Phi_{2,2}^{\dagger,+} - \Phi_{1,2}^{\dagger,-}\Phi_{2,1}^{\dagger,+}}{\Phi_{1,1}^{\dagger,+}\Phi_{2,2}^{\dagger,+} - \Phi_{1,2}^{\dagger,+}\Phi_{2,1}^{\dagger,+}} = \frac{\Psi_{1,1}^{\dagger,-}\Psi_{2,2}^{\dagger,+} - \Psi_{1,2}^{\dagger,-}\Psi_{2,1}^{\dagger,+}}{\Psi_{1,1}^{\dagger,+}\Psi_{2,2}^{\dagger,+} - \Psi_{1,2}^{\dagger,+}\Psi_{2,1}^{\dagger,+}}\,. \tag{4.2.14}$$

Using this formula, we extend $a^\dagger(z^\dagger)$ analytically to $z^\dagger \in \overline{\mathbb{K}_+}$ up to a finite set of points. Since the determinant is multiplicative, and due to (4.2.6), we have

$$a^{\mathsf{GP}}(z;q) = a^{\mathsf{hGP}}(z;w) = a^{\varepsilon\mathsf{GP}}(k;W)\,.$$

Note that

$$\det E^{\dagger,\pm} = \begin{cases} 1 & ,\dagger = \mathsf{NLS} \\ 2z(\lambda - z) & ,\dagger = \mathsf{GP} \\ -2z(\lambda - z) & ,\dagger = \mathsf{hGP} \\ -2\sqrt{2}\varepsilon k(\lambda - \sqrt{2}\varepsilon k) & ,\dagger = \varepsilon\mathsf{GP} \\ 2ik & ,\dagger = \mathsf{KdV} \end{cases}.$$

Using (4.2.5) to substitute $\Phi_{1,2}^{\dagger,-}$ in (4.2.14) and taking the limit $x \to \infty$, or, respectively, by substituting $\Phi_{2,1}^{+}$ and taking the limit $x \to -\infty$, we find that

$$\lim_{x^\dagger\to\infty} \Psi_{1,1}^{\dagger,-} = a^\dagger\beta_1^\dagger \qquad\qquad \lim_{x^\dagger\to-\infty} \Psi_{2,2}^{\dagger,+} = a^\dagger\beta_2^\dagger$$

where

$$\beta_1^\dagger = \frac{\det E^{\dagger,+}}{E_{2,2}^{\dagger,+} + (A_{2,2}^{\dagger,+} + iz^\dagger)^{-1}A_{1,2}^{\dagger,+}E_{2,1}^{\dagger,+}} = \begin{cases} 1 & ,\dagger = \mathsf{NLS} \\ q_+ & ,\dagger = \mathsf{GP} \\ e^{\theta_+}\frac{i}{\sqrt{2}}(1+\lambda - z) & ,\dagger = \mathsf{hGP} \\ e^{\theta_+}\frac{i}{\sqrt{2}}(1+\lambda - \sqrt{2}\varepsilon k) & ,\dagger = \varepsilon\mathsf{GP} \\ 1 & ,\dagger = \mathsf{KdV} \end{cases}$$

$$\beta_2^\dagger = \frac{\det E^{\dagger,-}}{E_{1,1}^{\dagger,-} + (A_{1,1}^{\dagger,-} - iz^\dagger)^{-1}A_{2,1}^{\dagger,-}E_{1,2}^{\dagger,-}} = \begin{cases} 1 & ,\dagger = \mathsf{NLS} \\ \overline{q}_- & ,\dagger = \mathsf{GP} \\ e^{-\theta_-}\frac{i}{\sqrt{2}}(1+z-\lambda) & ,\dagger = \mathsf{hGP} \\ e^{-\theta_-}\frac{i}{\sqrt{2}}(1+\sqrt{2}\varepsilon k - \lambda) & ,\dagger = \varepsilon\mathsf{GP} \\ ik & ,\dagger = \mathsf{KdV} \end{cases}.$$

On the other hand, (4.2.3) implies

$$\lim_{x^\dagger\to-\infty} \Psi_{1,1}^{\dagger,-} = E_{1,1}^{\dagger,-} = \begin{cases} 1 & ,\dagger = \mathsf{NLS} \\ q_- & ,\dagger = \mathsf{GP} \\ e^{\theta_-}\frac{i}{\sqrt{2}}(1+\lambda - z) & ,\dagger = \mathsf{hGP} \\ e^{\theta_-}\frac{i}{\sqrt{2}}(1+\lambda - \sqrt{2}\varepsilon k) & ,\dagger = \varepsilon\mathsf{GP} \\ 1 & ,\dagger = \mathsf{KdV}\,. \end{cases}$$

$$\lim_{x^\dagger\to\infty} \Psi^{\dagger,+}_{2,2} = E^{\dagger,+}_{2,2} = \begin{cases} 1 & ,\dagger = \mathsf{NLS} \\ \overline{q}_+ & ,\dagger = \mathsf{GP} \\ e^{-\theta_+}\frac{i}{\sqrt{2}}(1+z-\lambda) & ,\dagger = \mathsf{hGP} \\ e^{-\theta_+}\frac{i}{\sqrt{2}}(1+\sqrt{2}\varepsilon k-\lambda) & ,\dagger = \varepsilon\mathsf{GP} \\ ik & ,\dagger = \mathsf{KdV}\,. \end{cases}$$

Due to (4.2.9)–(4.2.10), there exists for sufficiently large $\operatorname{Im} z > c = c(q)$ a branch $\log^\dagger$ of the logarithm, only depending on $q^\dagger_\pm$, for which $\log(\Psi^{\dagger,-}_{1,1}(\beta^\dagger_1)^{-1})$ and $\log(\Psi^{\dagger,+}_{2,2}(\beta^\dagger_2)^{-1})$ are continuously differentiable in $x^\dagger$ on all of $\mathbb{R}$. Specifically, writing $\log_{\mathrm{pr}}$ for the principal logarithm, we choose

$$\log^\dagger(\omega) = \begin{cases} \log_{\mathrm{pr}}(\omega) & ,\dagger \in \{\mathsf{NLS},\mathsf{KdV}\} \\ \log_{\mathrm{pr}}(\omega e^{i(\varphi_+-\varphi_-)}) - \log_{\mathrm{pr}}(e^{i(\varphi_+-\varphi_-)}) & ,\dagger = \mathsf{GP} \\ \log_{\mathrm{pr}}(\omega e^{\frac{i}{2}(\varphi_+-\varphi_-)}) - \log_{\mathrm{pr}}(e^{\frac{i}{2}(\varphi_+-\varphi_-)}) & ,\dagger \in \{\mathsf{hGP},\varepsilon\mathsf{GP}\} \end{cases} \qquad \forall\,\omega\in\mathbb{C}\,.$$

Then

$$\log^\dagger a^\dagger - \begin{cases} 0 & ,\dagger = \mathsf{NLS} \\ i(\varphi_- - \varphi_+) & ,\dagger = \mathsf{GP} \\ \frac{i}{2}(\varphi_- - \varphi_+) & ,\dagger \in \{\mathsf{hGP},\varepsilon\mathsf{GP}\} \\ 0 & ,\dagger = \mathsf{KdV} \end{cases} = \int_{\mathbb{R}} \sigma^\dagger\,\mathrm{d}x^\dagger = \int_{\mathbb{R}} \widetilde{\sigma}^\dagger\,\mathrm{d}x^\dagger\,. \tag{4.2.15}$$

We are interested in the asymptotic expansions

$$\sigma^\dagger(x^\dagger,z^\dagger;q^\dagger) \sim \sum_{n=0}^\infty \frac{\sigma^\dagger_n(x^\dagger;q^\dagger)}{(2iz^\dagger)^n} \qquad\qquad r^\dagger(x^\dagger,z^\dagger;q^\dagger) \sim \sum_{n=0}^\infty \frac{r^\dagger_n(x^\dagger;q^\dagger)}{(2iz^\dagger)^n}\,. \tag{4.2.16}$$

By (iii) from Lemma 4.2.3 these expansions indeed exist, and the expansion coefficients are bounded and integrable in $x^\dagger$. Furthermore, the asymptotic expansion of $\sigma^\dagger$ is L^1-smooth, meaning that we can take the integral and thereby obtain an asymptotic expansion in powers of $2iz^\dagger$ for the transmission coefficient of the form

$$\log^\dagger a^\dagger - \begin{cases} 0 & ,\dagger = \mathsf{NLS} \\ i(\varphi_- - \varphi_+) & ,\dagger = \mathsf{GP} \\ \frac{i}{2}(\varphi_- - \varphi_+) & ,\dagger \in \{\mathsf{hGP},\varepsilon\mathsf{GP}\} \\ 0 & ,\dagger = \mathsf{KdV} \end{cases} \sim \sum_{n=0}^\infty \frac{\int_{\mathbb{R}} \sigma^\dagger_n\,\mathrm{d}x^\dagger}{(2iz^\dagger)^n} \sim \sum_{n=0}^\infty \frac{\int_{\mathbb{R}} \widetilde{\sigma}^\dagger_n\,\mathrm{d}x^\dagger}{(2iz^\dagger)^n}\,. \tag{4.2.17}$$

For the cases $\dagger \in \{\mathsf{GP},\mathsf{hGP},\varepsilon\mathsf{GP}\}$ clarification is necessary, as the expansion coefficients both in (4.2.16) and (4.2.17) are not the same on all of $\overline{\mathbb{K}_+}$. Instead, they depend on the sign of $\operatorname{Im}\lambda$, or alternatively a choice of bijection $\lambda \leftrightarrow (z,k)$. We shall choose the expansion coefficients so that (4.2.16)–(4.2.17) hold true for $\operatorname{Im}\lambda > 0$.

Using these expansion coefficients, we define for $\dagger \neq \mathsf{KdV}$ the Hamiltonians

$$\mathcal{H}^\dagger_n(q^\dagger) = -(-i)^n\big((\sqrt{2}\varepsilon)^{(n+1)}\big)^{\mathbb{1}_{\{\dagger=\varepsilon\mathsf{GP}\}}}\int_{\mathbb{R}} \sigma^\dagger_{n+1}(x^\dagger;q^\dagger)\,\mathrm{d}x^\dagger + \mathbb{1}_{\{n=-1\}}\begin{cases} 0 & ,\dagger = \mathsf{NLS} \\ \varphi_- - \varphi_+ & ,\dagger = \mathsf{GP} \\ \frac{\varphi_- - \varphi_+}{2} & ,\dagger \in \{\mathsf{hGP},\varepsilon\mathsf{GP}\} \end{cases}, \tag{4.2.18}$$

and for $\dagger = \mathsf{KdV}$ the energies

$$\mathcal{E}^{\mathsf{KdV}}_n(U) = (-1)^n \int_{\mathbb{R}} \sigma^{\mathsf{KdV}}_{2n+3}(X;U)\,\mathrm{d}X\,. \tag{4.2.19}$$

We do not define conserved quantities corresponding to the even orders because they vanish (see Lemma 4.2.6 below). These Hamiltonians act as expansion coefficients for the logarithm of the transmission coefficient. Specifically, for $(\lambda,z,k) \in \mathbb{K}_+ \cap \overline{\mathcal{Q}_1}$ we have as $\operatorname{Im} z \to \infty$ the asymptotic expansions

$$\log a^\dagger(z^\dagger;q^\dagger) \sim i\sum_{n=-1}^\infty \frac{\mathcal{H}^\dagger_n(q^\dagger)}{(2z)^{n+1}} \qquad\qquad \forall\,\dagger \in \{\mathsf{NLS},\mathsf{GP},\mathsf{hGP},\varepsilon\mathsf{GP}\}$$

$$\log a^{\mathsf{KdV}}(k;U) \sim i \sum_{n=-1}^{\infty} \frac{\mathcal{E}_n^{\mathsf{KdV}}(U)}{(2k)^{2n+3}}\,.$$

In particular

$$\mathcal{H}_n^{\mathsf{GP}}(q) = \mathcal{H}_n^{\mathsf{hGP}}(w) = \mathcal{H}_n^{\varepsilon\mathsf{GP}}(W)\,. \tag{4.2.20}$$

Note that we can reformulate (4.1.3) as

$$\log a^{\mathsf{GP}}(z;q) \sim i \sum_{n=0}^{\infty} \frac{\mathcal{H}_{2n}^{\widetilde{\mathsf{GP}}}(q)}{(2z)^{2n+1}} + i\frac{\lambda}{z} \sum_{n=0}^{\infty} \frac{2^{\mathbb{1}_{\{n=0\}}}\mathcal{H}_{2n-1}^{\widetilde{\mathsf{GP}}}(q)}{(2z)^{2n}}\,,$$

which gives an alternative view on the definition of the $\widetilde{\mathsf{GP}}$ hierarchy and its relation to the GP hierarchy.

As mentioned above, for the cases $\dagger \in \{\mathsf{GP}, \mathsf{hGP}, \varepsilon\mathsf{GP}\}$ these expansion coefficients are specific to our choice $\operatorname{Im}\lambda > 0$. A crucial property of the Hamiltonians $\mathcal{H}_n^\dagger$ is that they are pairwise Poisson commuting. This is a well-known result and it implies that each Hamiltonian is a conserved quantity for every Hamiltonian flow.

Lemma 4.2.4. *Let $\dagger \in \mathsf{Eqns}$ and $q^\dagger \in C^\infty(\mathbb{R})$ with $q^\dagger - q_\pm^\dagger \in \mathcal{S}(\mathbb{R}_\pm)$. Below, we write $\mathcal{H}_n^{\mathsf{KdV}} = \mathcal{E}_n^{\mathsf{KdV}}$. For all $n, m \geq 0$ and $z_1^\dagger, z_2^\dagger \in \overline{\mathbb{K}_+}$ with $\operatorname{Im} z_1^\dagger, \operatorname{Im} z_2^\dagger > c(q^\dagger)$ we have*

$$\{\mathcal{H}_n^\dagger, \mathcal{H}_m^\dagger\}^\dagger(q^\dagger) = 0 \qquad\qquad \{\log a^\dagger(z_1^\dagger), \log a^\dagger(z_2^\dagger)\}(q^\dagger) = 0\,.$$

The Poisson brackets $\{\ldots,\ldots\}^\dagger$ are given in Appendix 4.B.

Proof. For NLS, GP and KdV, which are all part of the AKNS hierarchy, we refer to [118, Theorem B.7]. Specifically for GP we note also a different proof given in [71, III.§2]. With (4.2.20), we can derive from the case GP the cases hGP and $\varepsilon\mathsf{GP}$. □

4.2.4. Recurrence relations for the densities of the Hamiltonians

From (4.2.5) we derive Riccati equations for $\sigma^\dagger$ and $\widetilde{\sigma}^\dagger$ with $\dagger \in \{\mathsf{NLS}, \mathsf{GP}\}$:

$$\partial_x \sigma^{\mathsf{NLS}} = (2i\lambda)\sigma^{\mathsf{NLS}} + \frac{q_x}{q}\sigma^{\mathsf{NLS}} - (\sigma^{\mathsf{NLS}})^2 + q\overline{q} \tag{4.2.21}$$

$$-\partial_x \widetilde{\sigma}^{\mathsf{NLS}} = (2i\lambda)\widetilde{\sigma}^{\mathsf{NLS}} - \frac{\overline{q}_x}{\overline{q}}\widetilde{\sigma}^{\mathsf{NLS}} - (\widetilde{\sigma}^{\mathsf{NLS}})^2 + q\overline{q} \tag{4.2.22}$$

$$\partial_x \sigma^{\mathsf{GP}} = (2iz)\sigma^{\mathsf{GP}} - (\sigma^{\mathsf{GP}})^2 + q\overline{q} - 1 + q_x\frac{\sigma^{\mathsf{GP}} + i(\lambda - z)}{q} \tag{4.2.23}$$

$$-\partial_x \widetilde{\sigma}^{\mathsf{GP}} = (2iz)\widetilde{\sigma}^{\mathsf{GP}} - (\widetilde{\sigma}^{\mathsf{GP}})^2 + q\overline{q} - 1 - \overline{q}_x\frac{\widetilde{\sigma}^{\mathsf{GP}} + i(\lambda - z)}{\overline{q}} \tag{4.2.24}$$

They were first derived in [189]. For $\dagger \in \{\mathsf{hGP}, \varepsilon\mathsf{GP}\}$ we prefer to work with the systems for the pairs of variables $(\sigma^\dagger, r^\dagger)$ and $(\widetilde{\sigma}^\dagger, \widetilde{r}^\dagger)$, which are specifically

$$\tag{4.2.25}$$

$$\partial_x \sigma^{\mathsf{hGP}} = (2iz)\sigma^{\mathsf{hGP}} - (\sigma^{\mathsf{hGP}})^2 - w_-w_+ - w_-(1+\lambda) + w_+(1-\lambda) - i(w_+)_x r^{\mathsf{hGP}} \tag{4.2.26}$$

$$-\partial_x \widetilde{\sigma}^{\mathsf{hGP}} = (2iz)\widetilde{\sigma}^{\mathsf{hGP}} - (\widetilde{\sigma}^{\mathsf{hGP}})^2 - w_-w_+ - w_-(1+\lambda) + w_+(1-\lambda) - i(w_-)_x \widetilde{r}^{\mathsf{hGP}} \tag{4.2.27}$$

$$\partial_x r^{\mathsf{hGP}} = i(1+\lambda+w_+)(r^{\mathsf{hGP}})^2 + i(1-\lambda-w_-) \tag{4.2.28}$$

$$-\partial_x \widetilde{r}^{\mathsf{hGP}} = i(1-\lambda-w_-)(\widetilde{r}^{\mathsf{hGP}})^2 + i(1+\lambda+w_+) \tag{4.2.29}$$

$$\partial_X \sigma^{\varepsilon\mathsf{GP}} = (2ik)\sigma^{\varepsilon\mathsf{GP}} - (\sigma^{\varepsilon\mathsf{GP}})^2 + \frac{1}{2}\Big(-\varepsilon^2 W_-W_+ - W_-(1+\lambda) + W_+(1-\lambda) - i\sqrt{2}\varepsilon(W_+)_X r^{\varepsilon\mathsf{GP}}\Big) \tag{4.2.30}$$

$$-\partial_X \widetilde{\sigma}^{\varepsilon\mathsf{GP}} = (2ik)\widetilde{\sigma}^{\varepsilon\mathsf{GP}} - (\widetilde{\sigma}^{\varepsilon\mathsf{GP}})^2 + \frac{1}{2}\Big(-\varepsilon^2 W_-W_+ - W_-(1+\lambda) + W_+(1-\lambda) - i\sqrt{2}\varepsilon(W_-)_X \widetilde{r}^{\varepsilon\mathsf{GP}}\Big) \tag{4.2.31}$$

$$\partial_X r^{\varepsilon\mathsf{GP}} = \frac{i\varepsilon}{\sqrt{2}}(\varepsilon^{-2}(1+\lambda)+W_+)(r^{\varepsilon\mathsf{GP}})^2 + \frac{i\varepsilon}{\sqrt{2}}(\varepsilon^{-2}(1-\lambda)-W_-) \tag{4.2.32}$$

$$-\partial_X \widetilde{r}^{\varepsilon\mathsf{GP}} = \frac{i\varepsilon}{\sqrt{2}}(\varepsilon^{-2}(1-\lambda)-W_-)(\widetilde{r}^{\varepsilon\mathsf{GP}})^2 + \frac{i\varepsilon}{\sqrt{2}}(\varepsilon^{-2}(1+\lambda)+W_+)\,. \tag{4.2.33}$$

The case KdV is distinct in that we consider a single equation for σ^{KdV}, but also the system $(\widetilde{\sigma}^{\mathsf{KdV}}, \widetilde{r}^{\mathsf{KdV}})$:

$$\partial_X \sigma^{\mathsf{KdV}} = (2ik)\sigma^{\mathsf{KdV}} - (\sigma^{\mathsf{KdV}})^2 + U \tag{4.2.34}$$

$$-\partial_X \widetilde{\sigma}^{\mathsf{KdV}} = (2ik)\widetilde{\sigma}^{\mathsf{KdV}} - (\widetilde{\sigma}^{\mathsf{KdV}})^2 + U + U_X\widetilde{r}^{\mathsf{KdV}} \tag{4.2.35}$$

$$-\partial_X \widetilde{r}^{\mathsf{KdV}} = \left(U + \frac{1}{4}(2ik)^2\right)(\widetilde{r}^{\mathsf{KdV}})^2 - 1\,. \tag{4.2.36}$$

These equations prescribe recurrence relations for the expansion coefficients $\sigma_n^\dagger$. In order to avoid an abuse of notation, we shall use bold variables $\boldsymbol{\sigma}^\dagger$, $\widetilde{\boldsymbol{\sigma}}^\dagger$, $\mathbf{r}^\dagger$, $\widetilde{\mathbf{r}}^\dagger$, and so on to denote the formal power series corresponding to each function. By substituting the formal power series $\boldsymbol{\sigma}^{\mathsf{NLS}}, \widetilde{\boldsymbol{\sigma}}^{\mathsf{NLS}}, \boldsymbol{\sigma}^{\mathsf{GP}}$ and $\widetilde{\boldsymbol{\sigma}}^{\mathsf{GP}}$ into the Riccati equations (4.2.21)–(4.2.24), we can derive recurrence relations for the expansion coefficients σ_n^{NLS}, $\widetilde{\sigma}_n^{\mathsf{NLS}}$, σ_n^{GP} and $\widetilde{\sigma}_n^{\mathsf{GP}}$. For the case NLS this yields

$$\sigma_0^{\mathsf{NLS}} = 0 \qquad \sigma_1^{\mathsf{NLS}} = -q\overline{q} \qquad \sigma_{n+1}^{\mathsf{NLS}} = \partial_x\sigma_n^{\mathsf{NLS}} - \frac{q_x}{q}\sigma_n^{\mathsf{NLS}} + \sum_{k=0}^{n}\sigma_k^{\mathsf{NLS}}\sigma_{n-k}^{\mathsf{NLS}}\,.$$

Due to the presence of λ on the right-hand side of (4.2.23), the recurrence relation for σ_n^{GP} is awkward, so we use instead the relation

$$\boldsymbol{\sigma}^{\mathsf{NLS}}(x,\lambda) = \boldsymbol{\sigma}^{\mathsf{GP}}(x,z) + i(\lambda - z)\,,$$

which is proven by using it to substitute $\boldsymbol{\sigma}^{\mathsf{NLS}}$ in (4.2.21) and observing that this yields (4.2.23). We can then derive the expansion coefficients σ_n^{GP} from σ_n^{NLS} and vice versa. We assume now that $\lambda, z \in \overline{\mathcal{Q}_1}$. Then we can write $z = \sqrt{\lambda^2 - 1}$ and $\lambda = \sqrt{z^2+1}$ using the principal square root, and obtain the relations

$$\sigma_{2m}^{\mathsf{NLS}} = \sum_{k=0}^{m}\binom{m-1}{m-k}(-4)^{m-k}\sigma_{2k}^{\mathsf{GP}} \qquad \sigma_{2m+1}^{\mathsf{NLS}} = \sum_{k=0}^{m}\binom{m-\frac{1}{2}}{m-k}(-4)^{m-k}\sigma_{2k+1}^{\mathsf{GP}} + (-1)^{m+1}C_m \tag{4.2.37}$$

$$\sigma_{2m}^{\mathsf{GP}} = \sum_{k=0}^{m}\binom{m-1}{m-k}4^{m-k}\sigma_{2k}^{\mathsf{NLS}} \qquad \sigma_{2m+1}^{\mathsf{GP}} = \sum_{k=0}^{m}\binom{m-\frac{1}{2}}{m-k}4^{m-k}\sigma_{2k+1}^{\mathsf{NLS}} + C_m\,. \tag{4.2.38}$$

Here C_n are the Catalan numbers. They can be defined either by the recurrence relation they solve, or by their generating function:

$$C_0 = 1 \qquad C_{n+1} = \sum_{k=0}^{n}C_kC_{n-k} \qquad \sum_{n=0}^{\infty}X^nC_n = \frac{1-\sqrt{1-4X}}{2X}\,. \tag{4.2.39}$$

If $\lambda \notin \overline{\mathcal{Q}_1}$ or $z \notin \overline{\mathcal{Q}_1}$ then the relations (4.2.37)–(4.2.38) may need to be corrected by sign changes. Defining $q^\sharp(x) = q(-x)$, we can observe that $\boldsymbol{\sigma}^{\mathsf{NLS}}(x,\lambda;q)$ and $\widetilde{\boldsymbol{\sigma}}^{\mathsf{NLS}}(-x,\lambda;\overline{q}^\sharp)$ both solve (4.2.21), and $\boldsymbol{\sigma}^{\mathsf{GP}}(x,z;q)$ and $\widetilde{\boldsymbol{\sigma}}^{\mathsf{GP}}(-x,z;\overline{q}^\sharp)$ both solve (4.2.23). A consequence of the recurrence relations discussed above is that σ_n^{NLS} and σ_n^{GP} consist of monomials whose total number of derivatives has the opposite parity of n. Then $(\partial_x^k q^\sharp)(x) = (-1)^k(\partial_x^k q)(-x)$ together with (4.2.17) yields

$$\int \sigma_n^\dagger(x,z;q)\,\mathrm{d}x = (-1)^{n+1}\int \overline{\sigma_n^\dagger(x,z;q)}\,\mathrm{d}x \tag{4.2.40}$$

for $\dagger \in \{\mathsf{NLS},\mathsf{GP}\}$, implying that the functionals $\mathcal{H}_n^{\mathsf{NLS}}, \mathcal{H}_n^{\mathsf{GP}}, \mathcal{H}_n^{\mathsf{hGP}}$ and $\mathcal{H}_n^{\varepsilon\mathsf{GP}}$ are real-valued. We assume again that $\lambda, k \in \overline{\mathcal{Q}_1}$ are in the first quadrant of the complex plane in order to have an explicit expansion

$$\lambda = \sqrt{1+2\varepsilon^2k^2} = \sum_{n=-1}^{\infty}\frac{c_n}{(2ik)^n} \quad \text{where} \quad c_n = \begin{cases} -\frac{i\varepsilon}{\sqrt{2}} & ,n=-1 \\ \frac{i}{(\sqrt{2}\varepsilon)^n}\frac{1}{n}\binom{n}{\frac{n-1}{2}} & ,n\geq 0 \text{ odd} \\ 0 & ,\text{else}\end{cases}\,. \tag{4.2.41}$$

Then we can derive from (4.2.30)–(4.2.33) the recurrence relations

$$\sigma_0^{\varepsilon\mathsf{GP}} = \frac{c_{-1}}{2}(W_+ + W_-) \tag{4.2.42}$$

$$\sigma_1^{\varepsilon\mathsf{GP}} = \frac{c_{-1}}{2}\partial_X(W_+ + W_-) + \frac{c_{-1}^2}{4}(W_+ + W_-)^2 + \frac{\varepsilon^2}{2}W_-W_+ + \frac{1}{2}(W_- - W_+) + \frac{i\varepsilon}{\sqrt{2}}\partial_X W_+ r_0^{\varepsilon\mathsf{GP}} \tag{4.2.43}$$

$$\sigma_{n+1}^{\varepsilon\mathsf{GP}} = \partial_X \sigma_n^{\varepsilon\mathsf{GP}} + \sum_{k=0}^{n} \sigma_{n-k}^{\varepsilon\mathsf{GP}}\sigma_k^{\varepsilon\mathsf{GP}} + \frac{1}{2}c_n(W_- + W_+) + \frac{i\varepsilon}{\sqrt{2}}\partial_X W_+ r_n^{\varepsilon\mathsf{GP}}\,. \tag{4.2.44}$$

and

$$r_0^{\varepsilon\mathsf{GP}} = 1 \tag{4.2.45}$$

$$r_1^{\varepsilon\mathsf{GP}} = -\frac{i\varepsilon}{\sqrt{2}}\left(W_+ - W_- + \frac{2}{\varepsilon^2}\right) \tag{4.2.46}$$

$$\begin{aligned} r_{n+1}^{\varepsilon\mathsf{GP}} &= \partial_X r_n^{\varepsilon\mathsf{GP}} + \frac{ic_n}{\sqrt{2}\varepsilon} + \frac{i\varepsilon}{\sqrt{2}}\left(-W_+ - \frac{1}{\varepsilon^2}\right)\sum_{k=0}^{n} r_{n-k}^{\varepsilon\mathsf{GP}} r_k^{\varepsilon\mathsf{GP}} \\ &\quad - \frac{i}{\sqrt{2}\varepsilon}\sum_{m=0}^{n}\sum_{k=0}^{m} c_{n-m} r_{m-k}^{\varepsilon\mathsf{GP}} r_k^{\varepsilon\mathsf{GP}} - \frac{ic_{-1}}{\sqrt{2}\varepsilon}\sum_{k=1}^{n} r_{n+1-k}^{\varepsilon\mathsf{GP}} r_k^{\varepsilon\mathsf{GP}}\,. \end{aligned} \tag{4.2.47}$$

In the same way recurrence relations can be derived for $\widetilde{\sigma}^{\varepsilon\mathsf{GP}}$, $\widetilde{r}^{\varepsilon\mathsf{GP}}$, σ^{hGP}, $\widetilde{\sigma}^{\mathsf{hGP}}$, r^{hGP}, and $\widetilde{r}^{\mathsf{hGP}}$, which we omit here for brevity.

Lastly, we derive from (4.2.34)–(4.2.36) the recurrence relations

$$\sigma_0^{\mathsf{KdV}} = 0 \qquad \sigma_1^{\mathsf{KdV}} = -U \qquad \sigma_{n+1}^{\mathsf{KdV}} = \partial_X \sigma_n^{\mathsf{KdV}} + \sum_{k=0}^{n} \sigma_k^{\mathsf{KdV}} \sigma_{n-k}^{\mathsf{KdV}}\,. \tag{4.2.48}$$

and

$$\widetilde{\sigma}_0^{\mathsf{KdV}} = 0 \qquad \widetilde{\sigma}_1^{\mathsf{KdV}} = -U \qquad \widetilde{\sigma}_{n+1}^{\mathsf{KdV}} = -\partial_X \widetilde{\sigma}_n^{\mathsf{KdV}} + \sum_{k=0}^{n} \widetilde{\sigma}_k^{\mathsf{KdV}} \widetilde{\sigma}_{n-k}^{\mathsf{KdV}} - U_X \widetilde{r}_n^{\mathsf{KdV}}$$

as well as

$$\widetilde{r}_0^{\mathsf{KdV}} = 0 \qquad \widetilde{r}_1^{\mathsf{KdV}} = 2 \qquad \widetilde{r}_{n+1}^{\mathsf{KdV}} = -\partial_X \widetilde{r}_n^{\mathsf{KdV}} - U\sum_{k=0}^{n} \widetilde{r}_k^{\mathsf{KdV}} \widetilde{r}_{n-k}^{\mathsf{KdV}} - \frac{1}{4}\sum_{k=2}^{n} \widetilde{r}_k^{\mathsf{KdV}} \widetilde{r}_{n+2-k}^{\mathsf{KdV}}\,.$$

Note that something subtle has happened in (4.2.45): the equation (4.2.30) only implies $(r_0^{\varepsilon\mathsf{GP}})^2 = 1$, so we need to use the additional information

$$r_0^{\varepsilon\mathsf{GP}} = \lim_{k\to i\infty} \frac{E_{1,2}^{\varepsilon\mathsf{GP},-}(k)}{E_{1,1}^{\varepsilon\mathsf{GP},-}(k)} = \lim_{k,\lambda\to i\infty} \frac{1+\sqrt{2}\varepsilon k - \lambda}{1-\sqrt{2}\varepsilon k + \lambda} = 1\,.$$

On the other hand, we have $\widetilde{r}_0^{\varepsilon\mathsf{GP}} = -1$. These statements are true only in the case $\operatorname{Im}\lambda > 0$, which we have assumed by considering the asymptotic expansion for $\lambda, k \in \overline{\mathcal{Q}_1}$. On other regions of $\overline{\mathbb{K}_+}$, where instead $\operatorname{Im}\lambda < 0$ while of course $\operatorname{Im} k > 0$, we have $r_0^{\varepsilon\mathsf{GP}} = -1$, $\widetilde{r}_0^{\varepsilon\mathsf{GP}} = 1$. In addition we must also replace c_n by $-c_n$, hence the expansion coefficients change significantly. This is relevant for the following lemma, which uses the operators Ev_λ and Od_λ defined in (4.1.14).

Lemma 4.2.5. *For $\lambda, k \in \overline{\mathcal{Q}_1}$ we have the following asymptotic expansions in powers of $2ik$ at infinity:*

$$\mathrm{Ev}_\lambda[\log a^{\varepsilon\mathsf{GP}}(k)] \sim \sum_{n=0}^{\infty} \frac{\int_{\mathbb{R}} \operatorname{Re}\sigma_{2n+1}^{\varepsilon\mathsf{GP}}\,\mathrm{d}X}{(2ik)^{2n+1}} \qquad \mathrm{Od}_\lambda[\log a^{\varepsilon\mathsf{GP}}(k)] \sim \sum_{n=0}^{\infty} \frac{\int_{\mathbb{R}} i\operatorname{Im}\sigma_{2n}^{\varepsilon\mathsf{GP}}\,\mathrm{d}X}{(2ik)^{2n}}\,. \tag{4.2.49}$$

Proof. Due to (4.2.40), the real and imaginary parts on the right-hand side are superfluous. Define now $W_{\pm}^{\sharp}(X) = W_{\pm}(-X)$. An analysis of (4.2.30)–(4.2.33) reveals that

$$r^{\varepsilon\mathsf{GP}}(X,(-\lambda,k);(W_+,W_-)) = \widetilde{r}^{\varepsilon\mathsf{GP}}(-X,(\lambda,k);(-W_-^{\sharp},-W_+^{\sharp}))$$
$$\sigma^{\varepsilon\mathsf{GP}}(X,(-\lambda,k);(W_+,W_-)) = \widetilde{\sigma}^{\varepsilon\mathsf{GP}}(-X,(\lambda,k);(-W_-^{\sharp},-W_+^{\sharp}))\,.$$

Observe that $(W_+,W_-) \mapsto (-W_-^{\sharp},-W_+^{\sharp})$ corresponds to $q \mapsto q^{\sharp}$. Then one can proceed as in the proof of (4.2.40) to obtain

$$\int_{\mathbb{R}} \widetilde{\sigma}_n^{\varepsilon\mathsf{GP}}(-X,(\lambda,k);(-W_-^{\sharp},-W_+^{\sharp}))\,\mathrm{d}X = (-1)^{n+1}\int_{\mathbb{R}} \sigma_n^{\varepsilon\mathsf{GP}}(X,(\lambda,k);(W_+,W_-))\,\mathrm{d}X\,,$$

which yields the claimed expansions. □

4.2.5. PROOFS OF THEOREM 4.1.1 AND COROLLARY 4.1.5

Let us emphasize that these recurrence relations are convenient for the explicit computation of the expansion coefficients, but difficult to work with in proofs by induction. Instead, it is preferable to work with their generating functions, which solve the Riccati equations derived above, as operations on these generating functions elegantly represent complicated steps of a proof by induction. As an example, consider the following lemma.

Lemma 4.2.6. *For all $U \in \mathcal{S}(\mathbb{R};\mathbb{R})$ and $n \geq 0$*

$$\int_{\mathbb{R}} \sigma_{2n}^{\mathsf{KdV}}(X;U)\,\mathrm{d}X = 0\,.$$

Proof. We decompose $\boldsymbol{\sigma}^{\mathsf{KdV}} = E + O$ into the even and odd orders of the formal power series. Then (4.2.34) implies

$$O_X = (2ik)E - 2EO\,.$$

In particular

$$E = \frac{O_X}{2ik-2O} = -\frac{(2ik-2O)_X}{2ik-2O} = -\frac{1}{2}(\log(2ik-2O))_X\,.$$

Since the right-hand side is a formal power series where each coefficient is a derivative of a polynomial in U and its derivatives, its integral vanishes. □

This is essentially a proof by induction, but all the work of the induction step is hidden in the calculus of formal power series. In order to transform it into an actual proof by induction one needs to compare the coefficients at each order, which yields recurrence relations. The calculations with generating functions then represent the inductive step for these recurrence relations. We generally skip the base case, which sometimes also falls out of the generating function calculations, but otherwise can always be checked manually.

Using the generating-function approach, we prove our approximation result for the densities of the Hamiltonians. In the following one should understand $k \in i\,\mathbb{R}_+$ to be sufficiently large so that $\lambda \in i\,\mathbb{R}$.

Lemma 4.2.7. *Define $W_{\pm}^{\sharp}(X) = W_{\pm}(-X)$. We assume that $k \in i\,\mathbb{R}$ for the definition of the real and imaginary part of formal power series in the variable $2ik$. Let $\boldsymbol{\lambda}$ be a purely imaginary formal power series in the variable $2ik$ which satisfies $\boldsymbol{\lambda}^2 = 1 - \frac{\varepsilon^2}{2}(2ik)^2$. Then*

$$\mathrm{Re}[\boldsymbol{\sigma}^{\varepsilon\mathsf{GP}}](X,k;W) - \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}[\boldsymbol{\sigma}^{\varepsilon\mathsf{GP}}](X,k;W) = \boldsymbol{\sigma}^{\mathsf{KdV}}(X,k;-W_-) + O(\varepsilon^2) \tag{4.2.50}$$

$$\mathrm{Re}[\boldsymbol{\sigma}^{\varepsilon\mathsf{GP}}](X,k;W) + \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}[\boldsymbol{\sigma}^{\varepsilon\mathsf{GP}}](X,k;W) = \widetilde{\boldsymbol{\sigma}}^{\mathsf{KdV}}(-X,k;W_+^{\sharp}) + O(\varepsilon^2) \tag{4.2.51}$$

$$\mathrm{Re}[\widetilde{\boldsymbol{\sigma}}^{\varepsilon\mathsf{GP}}](X,k;W) - \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}[\widetilde{\boldsymbol{\sigma}}^{\varepsilon\mathsf{GP}}](X,k;W) = \widetilde{\boldsymbol{\sigma}}^{\mathsf{KdV}}(X,k;-W_-) + O(\varepsilon^2) \tag{4.2.52}$$

$$\mathrm{Re}[\widetilde{\boldsymbol{\sigma}}^{\varepsilon\mathsf{GP}}](X,k;W) + \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}[\widetilde{\boldsymbol{\sigma}}^{\varepsilon\mathsf{GP}}](X,k;W) = \boldsymbol{\sigma}^{\mathsf{KdV}}(-X,k;W_+^{\sharp}) + O(\varepsilon^2) \tag{4.2.53}$$

and

$$\begin{aligned}
&\mathrm{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon\mathsf{GP}}\right](X,k;W) - \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon\mathsf{GP}}\right](X,k;W) = O(\varepsilon^2)\\
&\mathrm{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon\mathsf{GP}}\right](X,k;W) + \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon\mathsf{GP}}\right](X,k;W) = \widetilde{\mathbf{r}}^{\mathsf{KdV}}(-X,k;W_+^\sharp) + O(\varepsilon^2)\\
&\mathrm{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\widetilde{\mathbf{r}}^{\varepsilon\mathsf{GP}}\right](X,k;W) - \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\widetilde{\mathbf{r}}^{\varepsilon\mathsf{GP}}\right](X,k;W) = \widetilde{\mathbf{r}}^{\mathsf{KdV}}(X,k;-W_-) + O(\varepsilon^2)\\
&\mathrm{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\widetilde{\mathbf{r}}^{\varepsilon\mathsf{GP}}\right](X,k;W) + \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\widetilde{\mathbf{r}}^{\varepsilon\mathsf{GP}}\right](X,k;W) = O(\varepsilon^2)\,.
\end{aligned}$$

Proof. We define

$$\begin{aligned}
\boldsymbol{\mu} &= \mathrm{Re}[\boldsymbol{\sigma}^{\varepsilon\mathsf{GP}}] + \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}[\boldsymbol{\sigma}^{\varepsilon\mathsf{GP}}] &\qquad \boldsymbol{\nu} &= \mathrm{Re}[\boldsymbol{\sigma}^{\varepsilon\mathsf{GP}}] - \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}[\boldsymbol{\sigma}^{\varepsilon\mathsf{GP}}]\\
\widetilde{\boldsymbol{\mu}} &= \mathrm{Re}[\widetilde{\boldsymbol{\sigma}}^{\varepsilon\mathsf{GP}}] + \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}[\widetilde{\boldsymbol{\sigma}}^{\varepsilon\mathsf{GP}}] &\qquad \widetilde{\boldsymbol{\nu}} &= \mathrm{Re}[\widetilde{\boldsymbol{\sigma}}^{\varepsilon\mathsf{GP}}] - \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}[\widetilde{\boldsymbol{\sigma}}^{\varepsilon\mathsf{GP}}]\\
\boldsymbol{\alpha} &= \mathrm{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon\mathsf{GP}}\right] + \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon\mathsf{GP}}\right] &\qquad \boldsymbol{\beta} &= \mathrm{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon\mathsf{GP}}\right] - \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\mathbf{r}^{\varepsilon\mathsf{GP}}\right]\\
\widetilde{\boldsymbol{\alpha}} &= \mathrm{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\widetilde{\mathbf{r}}^{\varepsilon\mathsf{GP}}\right] + \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\widetilde{\mathbf{r}}^{\varepsilon\mathsf{GP}}\right] &\qquad \widetilde{\boldsymbol{\beta}} &= \mathrm{Re}\left[\frac{i\varepsilon}{\sqrt{2}}\widetilde{\mathbf{r}}^{\varepsilon\mathsf{GP}}\right] - \frac{1}{i\boldsymbol{\lambda}}\,\mathrm{Im}\left[\frac{i\varepsilon}{\sqrt{2}}\widetilde{\mathbf{r}}^{\varepsilon\mathsf{GP}}\right]
\end{aligned}$$

From (4.2.32)–(4.2.33) we derive the system

$$\begin{aligned}
\boldsymbol{\alpha}_X - W_+\boldsymbol{\alpha}^2 + (2ik)^2\left(\frac{\varepsilon^2}{2}W_+\left(\frac{\boldsymbol{\alpha}-\boldsymbol{\beta}}{2}\right)^2 - \left(\frac{\boldsymbol{\alpha}+\boldsymbol{\beta}}{2}\right)\left(\frac{\boldsymbol{\alpha}-\boldsymbol{\beta}}{2}\right)\right) + 1 - \frac{\varepsilon^2}{2}W_- &= 0\\
\boldsymbol{\beta}_X - (W_+ + 2\varepsilon^{-2})\boldsymbol{\beta}^2 + (2ik)^2\left(\frac{\varepsilon^2}{2}W_+\left(\frac{\boldsymbol{\alpha}-\boldsymbol{\beta}}{2}\right)^2 - \boldsymbol{\beta}\left(\frac{\boldsymbol{\alpha}-\boldsymbol{\beta}}{2}\right)\right) - \frac{\varepsilon^2}{2}W_- &= 0\\
-\widetilde{\boldsymbol{\alpha}}_X + (W_- - 2\varepsilon^{-2})\widetilde{\boldsymbol{\alpha}}^2 - (2ik)^2\left(\frac{\varepsilon^2}{2}W_-\left(\frac{\widetilde{\boldsymbol{\alpha}}-\widetilde{\boldsymbol{\beta}}}{2}\right)^2 - \widetilde{\boldsymbol{\alpha}}\left(\frac{\widetilde{\boldsymbol{\alpha}}-\widetilde{\boldsymbol{\beta}}}{2}\right)\right) + \frac{\varepsilon^2}{2}W_+ &= 0\\
-\widetilde{\boldsymbol{\beta}}_X + W_-\widetilde{\boldsymbol{\alpha}}^2 - (2ik)^2\left(\frac{\varepsilon^2}{2}W_-\left(\frac{\widetilde{\boldsymbol{\alpha}}-\widetilde{\boldsymbol{\beta}}}{2}\right)^2 - \left(\frac{\widetilde{\boldsymbol{\alpha}}+\widetilde{\boldsymbol{\beta}}}{2}\right)\left(\frac{\widetilde{\boldsymbol{\alpha}}-\widetilde{\boldsymbol{\beta}}}{2}\right)\right) + 1 + \frac{\varepsilon^2}{2}W_+ &= 0\,.
\end{aligned}$$

This implies

$$\boldsymbol{\alpha}, \widetilde{\boldsymbol{\beta}} = O(\varepsilon^0) \qquad\qquad \widetilde{\boldsymbol{\alpha}}, \boldsymbol{\beta} = O(\varepsilon^2) \tag{4.2.54}$$

by induction. More precisely, we have

$$\begin{aligned}
\boldsymbol{\alpha}_X - \left(W_+ + \frac{1}{4}(2ik)^2\right)\boldsymbol{\alpha}^2 + 1 &= O(\varepsilon^2)\\
-\widetilde{\boldsymbol{\beta}}_X - \left(-W_- + \frac{1}{4}(2ik)^2\right)\widetilde{\boldsymbol{\beta}}^2 + 1 &= O(\varepsilon^2)\,.
\end{aligned}$$

As a result, the claimed approximations for $\mathbf{r}^{\varepsilon\mathsf{GP}}$ and $\widetilde{\mathbf{r}}^{\varepsilon\mathsf{GP}}$ hold. On the other hand, we obtain from (4.2.30)–(4.2.31) the system

$$\begin{aligned}
\boldsymbol{\mu}_X - (2ik)\boldsymbol{\mu} + \boldsymbol{\mu}^2 - \frac{\varepsilon^2}{2}(2ik)^2\left(\frac{\boldsymbol{\mu}-\boldsymbol{\nu}}{2}\right)^2 - W_+ + \varepsilon^2\frac{W_-W_+}{2} + \partial_X W_+\boldsymbol{\alpha} &= 0\\
\boldsymbol{\nu}_X - (2ik)\boldsymbol{\nu} + \boldsymbol{\nu}^2 - \frac{\varepsilon^2}{2}(2ik)^2\left(\frac{\boldsymbol{\mu}-\boldsymbol{\nu}}{2}\right)^2 + W_- + \varepsilon^2\frac{W_-W_+}{2} + \partial_X W_+\boldsymbol{\beta} &= 0\\
-\widetilde{\boldsymbol{\mu}}_X - (2ik)\widetilde{\boldsymbol{\mu}} + \widetilde{\boldsymbol{\mu}}^2 - \frac{\varepsilon^2}{2}(2ik)^2\left(\frac{\widetilde{\boldsymbol{\mu}}-\widetilde{\boldsymbol{\nu}}}{2}\right)^2 - W_+ + \varepsilon^2\frac{W_-W_+}{2} + \partial_X W_-\widetilde{\boldsymbol{\alpha}} &= 0
\end{aligned}$$

$$-\widetilde{\boldsymbol{\nu}}_X - (2ik)\widetilde{\boldsymbol{\nu}} + \widetilde{\boldsymbol{\nu}}^2 - \frac{\varepsilon^2}{2}(2ik)^2\left(\frac{\widetilde{\boldsymbol{\mu}} - \widetilde{\boldsymbol{\nu}}}{2}\right)^2 + W_- + \varepsilon^2\frac{W_-W_+}{2} + \partial_X W_-\widetilde{\boldsymbol{\beta}} = 0\,.$$

Again with induction, we find that

$$\begin{aligned}\boldsymbol{\mu}_X - (2ik)\boldsymbol{\mu} + \boldsymbol{\mu}^2 - W_+ + (W_+)_X\boldsymbol{\alpha} &= O(\varepsilon^2)\\ \boldsymbol{\nu}_X - (2ik)\boldsymbol{\nu} + \boldsymbol{\nu}^2 + W_- &= O(\varepsilon^2)\\ -\widetilde{\boldsymbol{\mu}}_X - (2ik)\widetilde{\boldsymbol{\mu}} + \widetilde{\boldsymbol{\mu}}^2 - W_+ &= O(\varepsilon^2)\\ -\widetilde{\boldsymbol{\nu}}_X - (2ik)\widetilde{\boldsymbol{\nu}} + \widetilde{\boldsymbol{\nu}}^2 + W_- + (W_-)_X\widetilde{\boldsymbol{\beta}} &= O(\varepsilon^2)\,.\end{aligned}$$

This concludes the proof of the approximations for $\boldsymbol{\sigma}^{\varepsilon\mathsf{GP}}$ and $\widetilde{\boldsymbol{\sigma}}^{\varepsilon\mathsf{GP}}$. □

Proof of Theorem 4.1.1. We assume again a setting where $\lambda, k \in \overline{\mathcal{Q}_1}$. From the definition (4.2.18) and the fact that the Hamiltonians are real, we obtain

$$\begin{aligned}\mathcal{H}_{2n}^{\varepsilon\mathsf{GP}}(W) &= (-1)^{n+1}(\sqrt{2}\varepsilon)^{2n+1}\int_{\mathbb{R}}\operatorname{Re}\sigma_{2n+1}^{\varepsilon\mathsf{GP}}(X;W)\,\mathrm{d}X\\ \mathcal{H}_{2n-1}^{\varepsilon\mathsf{GP}}(W) &= (-1)^{n}(\sqrt{2}\varepsilon)^{2n}\int_{\mathbb{R}}\operatorname{Im}\sigma_{2n}^{\varepsilon\mathsf{GP}}(X;W)\,\mathrm{d}X\,.\end{aligned}$$

Now combining (4.1.3) with (4.2.19) and Lemma 4.2.7 yields (4.1.9). Note that $\sigma_n^{\mathsf{KdV}}(\pm W_\pm)$ is a polynomial in U and its derivatives with an even number of derivatives present in each monomial. Therefore $\int_{\mathbb{R}}\sigma_n^{\mathsf{KdV}}(\mp X;\pm W_\pm^\sharp)\,\mathrm{d}X = \int_{\mathbb{R}}\sigma_n^{\mathsf{KdV}}(X;\pm W_\pm)\,\mathrm{d}X$.

The main work is done, but we would like to gain a precise understanding of the number of derivatives present in $\mathbf{P}_n$ and prove (4.1.12). Below we prove results about $\mathcal{H}^{\varepsilon\mathsf{GP}}$, which then directly imply the claims about $\mathcal{H}^{\varepsilon\widetilde{\mathsf{GP}}}$. For $k \in \mathbb{N}$ and a polynomial P in W_+, W_-, and their derivatives, we define $\pi_k P$ to be the sum of all monomials in P whose total number of derivatives is k. In the following we identify such polynomials P up to formal integration by parts, as the claims of Theorem 4.1.1 concern only integrals $\int_{\mathbb{R}} P\,\mathrm{d}X$.

We set now $\dagger = \varepsilon\mathsf{GP}$ in order to simplify the notation. From (4.2.42)–(4.2.44) we derive

$$\begin{aligned}\operatorname{Re}\sigma_0^\dagger &= 0\\ \operatorname{Re}\sigma_1^\dagger &= -\frac{\varepsilon^2}{8}(W_+ - W_-)^2 + \frac{1}{2}(W_- - W_+)\\ \operatorname{Re}\sigma_{n+1}^\dagger &= \partial_X\operatorname{Re}\sigma_n^\dagger + \sum_{k=0}^{n}(\operatorname{Re}\sigma_{n-k}^\dagger\operatorname{Re}\sigma_k^\dagger - \operatorname{Im}\sigma_{n-k}^\dagger\operatorname{Im}\sigma_k^\dagger) - \frac{\varepsilon}{\sqrt{2}}\partial_X W_+\operatorname{Im}r_n^\dagger\end{aligned}$$

and

$$\begin{aligned}\operatorname{Im}\sigma_0^\dagger &= -\frac{\varepsilon}{2\sqrt{2}}(W_+ + W_-)\\ \operatorname{Im}\sigma_1^\dagger &= \frac{\varepsilon}{2\sqrt{2}}\partial_X(W_+ - W_-)\\ \operatorname{Im}\sigma_{n+1}^\dagger &= \partial_X\operatorname{Im}\sigma_n^\dagger + \sum_{k=0}^{n}2\operatorname{Re}\sigma_{n-k}^\dagger\operatorname{Im}\sigma_k^\dagger + \frac{1}{2}\operatorname{Im}c_n(W_- + W_+) + \frac{\varepsilon}{\sqrt{2}}\partial_X W_+\operatorname{Re}r_n^\dagger\end{aligned}$$

as well as

$$\begin{aligned}\operatorname{Im}r_0^\dagger &= 0\\ \operatorname{Im}r_1^\dagger &= -\frac{\varepsilon}{\sqrt{2}}\left(W_+ - W_- + \frac{2}{\varepsilon^2}\right)\\ \operatorname{Im}r_{n+1}^\dagger &= \partial_X\operatorname{Im}r_n^\dagger + \frac{\varepsilon}{\sqrt{2}}\left(-W_+ - \frac{1}{\varepsilon^2}\right)\sum_{k=0}^{n}(\operatorname{Re}r_{n-k}^\dagger\operatorname{Re}r_k^\dagger - \operatorname{Im}r_{n-k}^\dagger\operatorname{Im}r_k^\dagger)\\ &\quad + \frac{1}{\sqrt{2}\varepsilon}\sum_{m=0}^{n}\sum_{k=0}^{m}\operatorname{Im}c_{n-m}2\operatorname{Re}r_{m-k}^\dagger\operatorname{Im}r_k^\dagger + \frac{\operatorname{Im}c_{-1}}{\sqrt{2}\varepsilon}\sum_{k=1}^{n}2\operatorname{Re}r_{n+1-k}^\dagger\operatorname{Im}r_k^\dagger\,.\end{aligned}$$

We do not write out the recurrence relation for $\operatorname{Re} r_n^\dagger$ since it plays no significant role. The first fact that we read from these recurrence relations is that $\operatorname{Re}\sigma_{2n+1}^\dagger$ and $\operatorname{Im}\sigma_{2n}^\dagger$ are sums of monomials with an even number of derivatives, while all monomials in $\operatorname{Re}\sigma_{2n}^\dagger$ and $\operatorname{Im}\sigma_{2n+1}^\dagger$ have an odd number of derivatives. Studying the recurrence relation for $\operatorname{Im} r_n^\dagger$ explicitly, we see that $\pi_k r_n^\dagger = 0$ for all $k \geq n-1$ and

$$\begin{aligned}\pi_0 \operatorname{Im} r_1^\dagger &= -\frac{\varepsilon}{\sqrt{2}}\left(W_+ - W_- + \frac{2}{\varepsilon^2}\right)\\ \pi_n \operatorname{Im} r_{n+1}^\dagger &= \partial_X \pi_{n-1} \operatorname{Im} r_n^\dagger\,,\end{aligned}$$

from which we deduce for all $n \geq 1$ that

$$\pi_{n-1} \operatorname{Im} r_n^\dagger = -\frac{\varepsilon}{\sqrt{2}}\partial_X^{n-1}\left(W_+ - W_- + \frac{2}{\varepsilon^2}\right).$$

Similarly, we find that $\pi_k \operatorname{Im}\sigma_n^\dagger = 0$ for all $k \geq n$ and

$$\begin{aligned}\pi_0 \operatorname{Im}\sigma_0^\dagger &= -\frac{\varepsilon}{2\sqrt{2}}(W_+ + W_-)\\ \pi_1 \operatorname{Im}\sigma_1^\dagger &= \frac{\varepsilon}{2\sqrt{2}}\partial_X(W_+ - W_-)\\ \pi_{n+1} \operatorname{Im}\sigma_{n+1}^\dagger &= \pi_n \partial_X \operatorname{Im}\sigma_n^\dagger\,,\end{aligned}$$

which implies

$$\pi_n \operatorname{Im}\sigma_n^\dagger = \begin{cases} -\frac{\varepsilon}{2\sqrt{2}}(W_+ + W_-) & , n = 0\\ \frac{\varepsilon}{2\sqrt{2}}\partial_X^n(W_+ - W_-) & , n \geq 1\end{cases}.$$

Lastly, we have $\pi_k \operatorname{Re}\sigma_n^\dagger = 0$ for $k \geq n$ and

$$\begin{aligned}\pi_0 \operatorname{Re}\sigma_1^\dagger &= -\frac{\varepsilon^2}{8}(W_+ - W_-)^2 + \frac{1}{2}(W_- - W_+)\\ \pi_n \operatorname{Re}\sigma_{n+1}^\dagger &= \partial_X \pi_{n-1}\operatorname{Re}\sigma_n^\dagger - \sum_{k=0}^{n}(\pi_k \operatorname{Im}\sigma_k^\dagger)(\pi_{n-k}\operatorname{Im}\sigma_{n-k}^\dagger) - \frac{\varepsilon}{\sqrt{2}}\partial_X W_+ \pi_{n-1}\operatorname{Im} r_n^\dagger\\ &= \partial_X \pi_{n-1}\operatorname{Re}\sigma_n^\dagger - \sum_{k=1}^{n-1}\frac{\varepsilon}{2\sqrt{2}}\partial_X^k(W_+ - W_-)\frac{\varepsilon}{2\sqrt{2}}\partial_X^{n-k}(W_+ - W_-)\\ &+ \frac{\varepsilon}{\sqrt{2}}(W_+ + W_-)\frac{\varepsilon}{2\sqrt{2}}\partial_X^n(W_+ - W_-) + \frac{\varepsilon}{\sqrt{2}}\partial_X W_+ \frac{\varepsilon}{\sqrt{2}}\partial_X^{n-1}\left(W_+ - W_- + \frac{2}{\varepsilon^2}\right)\\ &= \partial_X \pi_{n-1}\operatorname{Re}\sigma_n^\dagger + \frac{\varepsilon^2}{8}\Bigg(-\sum_{k=1}^{n-1}\partial_X^k(W_+ - W_-)\partial_X^{n-k}(W_+ - W_-)\\ &+ 2(W_+ + W_-)\partial_X^n(W_+ - W_-) + 4\partial_X W_+ \partial_X^{n-1}(W_+ - W_-)\Bigg)\\ &+ \mathbb{1}_{\{n=1\}}\partial_X W_+\,.\end{aligned}$$

This implies

$$\int_{\mathbb{R}} \pi_{2n}\operatorname{Re}\sigma_{2n+1}^\dagger \,\mathrm{d}X = (-1)^{n+1}\frac{\varepsilon^2}{8}\int_{\mathbb{R}}(\partial_X^n(W_+ - W_-))^2\,\mathrm{d}X\,,$$

from which we obtain (4.1.12). So far we have shown that $\operatorname{Re}\sigma_{2n+1}^\dagger$ has up to $2n$ derivatives present in each monomial. Via repeated integration by parts, we can reduce the number of derivatives on the factor which has the most, and thereby ensure that each monomial in $\mathbf{P}_{2n}$ has no more than n derivatives on any factor. For $\operatorname{Im}\sigma_{2n}^\dagger$ we observe that the terms $\pi_{2n}\operatorname{Im}\sigma_{2n}^\dagger$ with the most derivatives vanish under integration by parts, so we obtain at most $n-1$ derivatives on any factor in $\mathbf{P}_{2n-1}$. As a result of these elaborations, we also know that $\mathbf{P}_n$ has at most $\lfloor \frac{n}{2} \rfloor$ derivatives in each monomial.

It remains to prove the absence of constant or linear monomials in $\mathbf{P}_n$ and $\mathbf{Q}_n$. This follows by direct computation for $n \in \{-1, 0\}$. For $n \geq 1$ we shall obtain it from the absence of constant or linear monomials in $\mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}$ and $\mathcal{E}_{n-1}^{\mathsf{KdV}}$, which for the energies $\mathcal{E}_{n-1}^{\mathsf{KdV}}$ follows directly from (4.2.48). If P is a polynomial as above, we denote by $\tau_0 P$ its constant part and by $\tau_1 P$ its linear part. First, we note that (4.2.42)–(4.2.44) directly imply $\tau_0 \sigma_n^\dagger = 0$. Since $\pi_k \tau_1 \sigma_n^\dagger$ vanishes under integration by parts for all $k \geq 1$, it suffices to determine $\pi_0 \tau_1 \sigma_n^\dagger$. Here (4.2.42)–(4.2.44) imply

$$\pi_0 \tau_1 \sigma_0^\dagger = \frac{c_{-1}}{2}(W_+ + W_-) \qquad \pi_0 \tau_1 \sigma_1^\dagger = \frac{1}{2}(W_- - W_+) \qquad \pi_0 \tau_1 \sigma_{n+1}^\dagger = \frac{c_n}{2}(W_+ + W_-) \,.$$

Since $c_n = 0$ for even n it remains to prove the claim for the odd Hamiltonians

$$\mathcal{H}_{2n-1}^{\varepsilon\widetilde{\mathsf{GP}}}(W) = \int_{\mathbb{R}} \sum_{m=0}^{n} \binom{-\frac{1}{2}}{n-m} 4^{n-m}(-1)^m (\sqrt{2}\varepsilon)^{2m} \operatorname{Im} \sigma_{2m}^{\varepsilon\mathsf{GP}}(X; W) \, \mathrm{d}X \,.$$

Here we have

$$\begin{aligned}\pi_0 \tau_1 \sum_{m=0}^{n} \binom{-\frac{1}{2}}{n-m} 4^{n-m}(-1)^m (\sqrt{2}\varepsilon)^{2m} \operatorname{Im} \sigma_{2m}^{\varepsilon\mathsf{GP}} &= \sum_{m=0}^{n} \binom{-\frac{1}{2}}{n-m} 4^{n-m}(-1)^m (\sqrt{2}\varepsilon)^{2m} \frac{1}{2} \operatorname{Im} c_{2m-1}(W_+ + W_-) \\ &= -\mathbb{1}_{\{n=0\}} \frac{\varepsilon}{2\sqrt{2}}(W_+ + W_-) \,.\end{aligned}$$

□

Proof of Corollary 4.1.5. Define for $n \in \mathbb{N}$ the energies

$$\widetilde{E}^n(U) = \sum_{l=0}^{n} \binom{n}{l} \mathcal{E}_l^{\mathsf{KdV}}(U) \,.$$

According to [122, Theorem 9.1] there exists $\delta \in (0,1)$ such that for every $n \in \mathbb{N}$ the smallness condition $\|U\|_{H^{-1}} \leq \delta$ implies

$$|\widetilde{E}^n(U) - \|U\|_{H^n}^2| \leq c(n) \|U\|_{H^{-1}} \|U\|_{H^n}^2 \,.$$

We calculate

$$\begin{aligned}\mathcal{E}_n^{\varepsilon\widetilde{\mathsf{GP}},\pm}(W) &= \sum_{l=0}^{n} \binom{n}{l} (\sqrt{2}\varepsilon)^{-2l-3} \big(\mathcal{H}_{2l+2}^{\varepsilon\widetilde{\mathsf{GP}}}(W) \mp 2\mathcal{H}_{2l+1}^{\varepsilon\widetilde{\mathsf{GP}}}(W)\big) \\ &= \sum_{l=0}^{n} \binom{n}{l} \left(\mathcal{E}_l^{\mathsf{KdV}}(\pm W_\pm) + \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \mathbf{P}_{2l+2} \mp \mathbf{P}_{2l+1} \, \mathrm{d}X \right) \\ &= \sum_{l=0}^{n} \binom{n}{l} \left(\mathcal{E}_l^{\mathsf{KdV}}(\pm W_\pm) + \frac{\varepsilon^2}{8} \|\partial_X^{l+1}(W_+ - W_-)\|_{L^2}^2 + \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \mathbf{Q}_{2l+2} \mp \mathbf{P}_{2l+1} \, \mathrm{d}X \right) \\ &= \widetilde{E}^n(\pm W_\pm) + \frac{\varepsilon^2}{8} \|\partial_X(W_+ - W_-)\|_{H^n}^2 + \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \sum_{l=0}^{n} \binom{n}{l} (\mathbf{Q}_{2l+2} \mp \mathbf{P}_{2l+1}) \, \mathrm{d}X \\ &= \|W_\pm\|_{H^n}^2 + \frac{\varepsilon^2}{8} \|\partial_X(W_+ - W_-)\|_{H^n}^2 \\ &\quad + \widetilde{E}^n(\pm W_\pm) - \|W_\pm\|_{H^n}^2 + \frac{\varepsilon^2}{2} \int_{\mathbb{R}} \sum_{l=0}^{n} \binom{n}{l} (\mathbf{Q}_{2l+2} \mp \mathbf{P}_{2l+1}) \, \mathrm{d}X\end{aligned}$$

An analysis of the recurrence relations (4.2.42)–(4.2.44), (4.2.45)–(4.2.47) reveals that the polynomials $\mathbf{Q}_{2l+2}$ and $\mathbf{P}_{2l+1}$ for $0 \leq l \leq n$ have degree less than or equal to $n+3$, and at most n derivatives on a single factor. Furthermore, the total number of derivatives in any monomial is strictly less than $n+1$, and we know from Theorem 4.1.1 that they contain no constant or linear monomials. For $n \geq 1$ this allows us to combine the Cauchy–Schwarz inequality and the Sobolev embedding $H^1 \hookrightarrow L^\infty$ to estimate

$$\left| \int_{\mathbb{R}} \sum_{l=0}^{n} (\mathbf{Q}_{2l+2} \mp \mathbf{P}_{2l+1})\, \mathrm{d}X \right| \leq C\|W\|_{H^n}^2 (1 + \|W\|_{H^n})^{n+1} .$$

By a similar analysis of $\mathcal{E}_n^{\mathsf{KdV}}$ we find in particular that $\mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}(W) < \infty$ if $W \in H^{\lfloor \frac{n}{2} \rfloor}$. □

4.3. Approximation of solutions to the $\varepsilon\widetilde{\mathsf{GP}}$ hierarchy by solutions to the KdV hierarchy

4.3.1. Proof of Proposition 4.1.10

Proposition 4.1.10 is an assembly of known results. Most importantly, [104, Theorem 1.1] by C. E. Kenig and D. Pilod states local well-posedness of (KdV_n). For any initial data, the local solution to the n-th flow (KdV_n) can be globalized by the standard argument of combining a blow-up alternative with the existence of conserved energies that control the Sobolev norm of the solution. Then the solution to the N-truncated KdV hierarchy is constructed by using this well-posedness result to move for a time t_0 along the flow of (KdV_0), then for a time t_1 along the flow of (KdV_1), and so on. Due to the commutativity of the flows, this is independent of the choice of ordering. The commutativity of the flows is a well-known result which can be proven, for example, by proceeding as in Section 3.4, using the weighted Sobolev spaces $H^s \cap H^{s',1}$. Then by density the commutativity of the flows on H^s follows.

It remains to use certain conserved quantities to prove the uniform bound (4.1.17). While the energies $\mathcal{E}_n^{\mathsf{KdV}}$ are of course conserved, they do not have the desired coercivity property. A first result controlling the Sobolev norm of the solution by certain combinations of $\mathcal{E}_n^{\mathsf{KdV}}$ in the periodic case was given in [133, Theorem 3.1] by P. D. Lax. We shall use the recent results [115, Theorem 1.1] by R. Killip, M. Vişan, and X. Zhang, and [122, Theorem 9.1] by H. Koch and D. Tataru. The former result is only stated for Schwartz solutions, but the continuity of the local flow maps on Sobolev spaces allows us to approximate.

Proof of Proposition 4.1.10. Since the low regularity conserved quantities used in [115, Theorem 1.1] are conserved for every flow in the KdV hierarchy, we know that

$$\sup_{\mathbf{T} \in \mathbb{R}^{N+1}} \|U(\mathbf{T})\|_{H^{-1}} \leq C(1 + \|U_0\|_{H^{-1}}^2)\|U_0\|_{H^{-1}} ,$$

so it suffices to prove for every $T > 0$ that

$$\sup_{|\mathbf{T}| \leq T} \|U(\mathbf{T})\|_{H^{s'}} \leq C(s') \sup_{|\mathbf{T}| \leq T} \big(1 + \|U(\mathbf{T})\|_{H^{-1}}^{2s'}\big) \|U_0\|_{H^{s'}} .$$

To show this we use [122, Theorem 9.1], which provides a smallness constant $\delta \in (0,1)$ and for every $s' > -1$ a functional $\widetilde{E}^{s'} : H^{s'} \to \mathbb{R}$ such that $\|U\|_{H^{-1}} \leq \delta$ implies

$$|\widetilde{E}^{s'}(U) - \|U\|_{H^{s'}}^2| \leq c(s')\|U\|_{H^{-1}}\|U\|_{H^{s'}}^2 .$$

Crucially, the functionals $\widetilde{E}^{s'}$ are conserved along every flow of the KdV hierarchy. In particular, if

$$\sup_{|\mathbf{T}| \leq T} \|U(\mathbf{T})\|_{H^{-1}} \leq \delta_0 = \min\left\{\delta, \frac{1}{2c(s')}\right\} ,$$

then

$$\|U(\mathbf{T})\|_{H^{s'}}^2 \leq \widetilde{E}^{s'}(U(\mathbf{T})) + c(s')\|U(\mathbf{T})\|_{H^{-1}}\|U(\mathbf{T})\|_{H^{s'}}^2 \leq \widetilde{E}^{s'}(U(\mathbf{T})) + \frac{1}{2}\|U(\mathbf{T})\|_{H^{s'}}^2 ,$$

which implies

$$\sup_{|\mathbf{T}| \leq T} \|U(\mathbf{T})\|_{H^{s'}}^2 \leq 2\widetilde{E}^{s'}(U_0) \leq 2\big(\|U_0\|_{H^{s'}}^2 + c(s')\|U_0\|_{H^{-1}}\|U_0\|_{H^{s'}}^2\big) \leq 3\|U_0\|_{H^{s'}}^2 .$$

We therefore assume $\sup_{|\mathbf{T}|\le T}\|U(\mathbf{T})\|_{H^{-1}}>\delta_0$ and use the scaling invariance of the KdV hierarchy to regain the required smallness. Specifically, we define $u(\mathbf{T},X)=\omega^2U(\Omega\mathbf{T},\omega X)$ where

$$\omega=\frac{\delta_0^2}{\sup_{|\mathbf{T}|\le T}\|U(\mathbf{T})\|_{H^{-1}}^2}\in(0,1)\text{ and }\Omega\in\mathbb{R}^{(N+1)\times(N+1)}\text{ with }\Omega_{n,m}=\mathbb{1}_{\{n=m\}}\omega^{2n+1}\,.$$

By the representation (4.3.1) below, u is still a solution to the KdV hierarchy in the sense of this theorem. It satisfies

$$\sup_{|\Omega\mathbf{T}|\le T}\|u(\mathbf{T})\|_{H^{-1}}^2=\sup_{|\Omega\mathbf{T}|\le T}\int_{\mathbb{R}}\frac{\omega}{\frac{1}{\omega^2}+\xi^2}|\widehat{U}(\Omega\mathbf{T},\xi)|^2\,\mathrm{d}\xi\le\omega\sup_{|\mathbf{T}|\le T}\|U(\mathbf{T})\|_{H^{-1}}^2=\delta_0^2\,.$$

The above reasoning now implies

$$\sup_{|\Omega\mathbf{T}|\le T}\|u(\mathbf{T})\|_{H^{s'}}^2\le 3\|u(\mathbf{0})\|_{H^{s'}}^2\,,$$

which after scaling yields

$$\sup_{|\mathbf{T}|\le T}\|U(\mathbf{T})\|_{H^{s'}}^2\le 3\omega^{-2s'}\|U_0\|_{H^{s'}}^2\,.$$

□

4.3.2. Proof of Proposition 4.1.11

Lemma 4.3.1 (Continuity argument)**.** *For every $C_0>0$ and $h\in\mathbb{N}$ there exists some $\delta_0,\widetilde{\varepsilon}_0,C_1>0$ such that the following holds. Let $r^*\in(0,\infty)\cup\{\infty\}$ and consider two functions $g_\pm\in C([0,r^*);[0,\infty))$ that satisfy for some $\widetilde{\varepsilon}\in(0,\widetilde{\varepsilon}_0)$ and all $r\in[0,r^*)$*

$$\begin{aligned}|g_\pm(r)-g_\pm(0)|&\le C_0(g_\pm(r)+g_\pm(0))(g_+(r)+g_-(r))^{\frac12}\\&\quad+\widetilde{\varepsilon}^2C_0(g_+(r)+g_-(r)+g_+(0)+g_-(0))(1+g_+(r)+g_-(r)+g_+(0)+g_-(0))^{h+1}\,.\end{aligned}$$

Then $g_+(0)+g_-(0)<\delta_0$ implies

$$g_\pm(r)\le\begin{cases}3g_\pm(0) & ,g_\pm(0)>0\\ 3g_\mp(0) & ,g_\pm(0)=0\end{cases}+\widetilde{\varepsilon}^2C_1(g_+(0)+g_-(0))$$

for all $r\in[0,r^*)$.

Proof. Set $E=\widetilde{\varepsilon}^2C_1(g_+(0)+g_-(0))$. We prove by contradiction, so we assume that the estimate fails before r^*. More specifically, we may assume without loss of generality that the estimate for g_+ becomes saturated at some time $r_+\in(0,r^*)$, and holds for both g_+ and g_- on the interval $(0,r_+)$. This means either $g_+(r_+)=3g_+(0)+E$ and $g_+(0)>0$, or $g_+(r_+)=3g_-(0)+E$ and $g_+(0)=0$, and furthermore for all $r\in(0,r_+)$ we have $g_\pm(r)\le 3g_\pm(0)+E$ if $g_\pm(0)>0$, and $g_\pm(r)\le 3g_\mp(0)+E$ if $g_\pm(0)=0$. In the case $g_+(0)>0$ this implies

$$\begin{aligned}2g_+(0)+E&\le C_0(8g_+(0)+E)(6g_+(0)+6g_-(0)+2E)^{\frac12}\\&\quad+\widetilde{\varepsilon}^2C_0(8g_+(0)+8g_-(0)+2E)(1+8g_+(0)+8g_-(0)+2E)^{h+1}\,,\end{aligned}$$

which leads to a contradiction if δ_0 and $\widetilde{\varepsilon}_0$ are sufficiently small and $g_+(0)+g_-(0)<\delta_0$. The next case is $g_+(0)=0$ and $g_-(0)>0$. Here we have

$$\begin{aligned}g_+(r)&\le C_0g_+(r)(g_+(r)+g_-(r))^{\frac12}\\&\quad+\widetilde{\varepsilon}^2C_0(g_+(r)+g_-(r)+g_-(0))(1+g_+(r)+g_-(r)+g_-(0))^{h+1}\,.\end{aligned}$$

and $\max\{g_+(r),g_-(r)\}\le 3g_-(0)+E$ for all $r\in(0,r_+)$. In particular

$$\begin{aligned}3g_-(0)+E=g_+(r_+)=\lim_{r\nearrow r_+}g_+(r)&\le C_0(3g_-(0)+E)(6g_-(0)+2E)^{\frac12}\\&\quad+\widetilde{\varepsilon}^2C_0(8g_-(0)+E)(1+8g_-(0)+E)^{h+1}\,,\end{aligned}$$

which again leads to a contradiction for sufficiently small δ_0 and $\widetilde{\varepsilon}_0$. The final case is $g_+(0) = g_-(0) = 0$, where

$$\begin{aligned} g_\pm(r) &\le C_0 g_\pm(r)(g_+(r) + g_-(r))^{\frac{1}{2}} \\ &\quad + \widetilde{\varepsilon}^2 C_0 (g_+(r) + g_-(r))(1 + g_+(r) + g_-(r))^{h+1} . \end{aligned}$$

If either g_+ or g_- are ever positive on $[0, r^*)$, then they assume arbitrarily small values for some $r \in (0, r^*)$, which is not compatible with the estimate when δ_0 and $\widetilde{\varepsilon}_0$ are sufficiently small. $\square$

We can now prove the proposition.

Proof of Proposition 4.1.11. This proposition is Corollary 3.1.7 with the addition of (iv), which we now prove. A lower bound for the distance to vacuum using the energy is given, for example, in [26, Lemma 1] and Lemma 2.1.2. Then continuity and boundedness of W in H^{s-1} follow from the local bilipschitz property of the Madelung transform $q \mapsto \left(\left(1 + \frac{w_+ - w_-}{2}\right)^2, w_+ + w_-\right)$ established in Corollary 2.1.7 and Lemma 2.2.4 together with $q \in C(\mathbb{R}^{N+1}; X^s)$. For a detailed proof of the fact that ($\varepsilon\mathsf{GP}_n$) is solved in the sense of distributions in the case $n = 2$, we refer to Section 2.3. The cases $n \ge 3$ are analogous.

In fact the solutions are strong in the sense that $T_n \mapsto W(\mathbf{c}^{\mathsf{NLS} \mapsto \varepsilon\widetilde{\mathsf{GP}}}\mathbf{T}) \in C^1(H^{s-2\lfloor \frac{n}{2} \rfloor - 2})$ satisfies ($\varepsilon\widetilde{\mathsf{GP}}_n$) in that space. We obtain this by transferring the differentiability of the solutions from (iii) of Proposition 4.1.11 through the Madelung transform. Away from vacuum, this boils down to combining product estimates and the boundedness of certain smooth nonlinear mappings on Sobolev spaces, which can be performed as in Chapter 2. Note that the highest order of derivative appearing in ($\varepsilon\widetilde{\mathsf{GP}}_n$) is $2\lfloor \frac{n}{2} \rfloor + 1$.

It remains to prove (4.1.23), so we let $h \in \{1, \dots, s-2\}$. Here we use Corollary 4.1.5. In order to obtain an estimate for $\|W_\pm\|_{H^h}$ from this bound we need smallness of $\|W_\pm\|_{H^{-1}}$. We shall obtain smallness for $\|W\|_{H^{h+1}}$ globally in time by slightly relaxing our scaling. Specifically, we define $\widetilde{\mathbf{T}}$, $\widetilde{X}$, and $\widetilde{W}$ in analogy to $\mathbf{T}$, X, and W, with ε replaced by

$$\widetilde{\varepsilon} = \varepsilon \max\{1, \delta^{-\frac{1}{3}} \|W(\mathbf{0})\|_{H^{h+1}}^{\frac{2}{3}}\}$$

for some $\delta = \delta(h) > 0$ to be chosen later. Then $\widetilde{W}$ has the same properties as described above, except ε is replaced by $\widetilde{\varepsilon}$. Since

$$\widetilde{W}(\widetilde{\mathbf{T}}, \widetilde{X}) = \frac{1}{\widetilde{\varepsilon}^2} w(\mathbf{t}, x) = \left(\frac{\varepsilon}{\widetilde{\varepsilon}}\right)^2 \frac{1}{\varepsilon^2} w(\mathbf{t}, x) = \left(\frac{\varepsilon}{\widetilde{\varepsilon}}\right)^2 W(\mathbf{T}, X) = \left(\frac{\varepsilon}{\widetilde{\varepsilon}}\right)^2 W\left(\mathbf{T}, \frac{\varepsilon}{\widetilde{\varepsilon}} \widetilde{X}\right)$$

we have for some constant $C = C(h) > 0$

$$C^{-1} \left(\frac{\varepsilon}{\widetilde{\varepsilon}}\right)^{\frac{3}{2} + h} \|W(\mathbf{T})\|_{H^h} \le \|\widetilde{W}(\widetilde{\mathbf{T}})\|_{H^h} \le C \left(\frac{\varepsilon}{\widetilde{\varepsilon}}\right)^{\frac{3}{2}} \|W(\mathbf{T})\|_{H^h} ,$$

in particular $\|\widetilde{W}(\mathbf{0})\|_{H^{h+1}}^2 < \delta$. By Corollary 4.1.5 there exists a choice of δ and $C = C(h) > 0$ so that $\|\widetilde{W}\|_{H^{-1}}^2 < 4\delta$ implies

$$\begin{aligned} \left| \mathcal{E}_h^{\varepsilon\widetilde{\mathsf{GP}}, \pm}(\widetilde{W}) - \|\widetilde{W}_\pm\|_{H^h}^2 - \frac{\widetilde{\varepsilon}^2}{8} \|\partial_X (\widetilde{W}_+ - \widetilde{W}_-)\|_{H^h}^2 \right| &\le C \|\widetilde{W}_\pm\|_{H^h}^2 \|\widetilde{W}_\pm\|_{H^{-1}} \\ &\quad + \widetilde{\varepsilon}^2 C \|\widetilde{W}\|_{H^h}^2 (1 + \|\widetilde{W}\|_{H^h})^{h+1} . \end{aligned}$$

We use this to estimate

$$\begin{aligned} &\left| \|\widetilde{W}_\pm(\widetilde{\mathbf{T}})\|_{H^h}^2 + \frac{\widetilde{\varepsilon}^2}{8} \|\partial_X (\widetilde{W}_+ - \widetilde{W}_-)(\widetilde{\mathbf{T}})\|_{H^h}^2 - \|\widetilde{W}_\pm(\mathbf{0})\|_{H^h}^2 - \frac{\widetilde{\varepsilon}^2}{8} \|\partial_X (\widetilde{W}_+ - \widetilde{W}_-)(\mathbf{0})\|_{H^h}^2 \right| \\ &\le C \Big(\|\widetilde{W}_\pm(\widetilde{\mathbf{T}})\|_{H^h}^2 \|\widetilde{W}_\pm(\widetilde{\mathbf{T}})\|_{H^{-1}} + \|\widetilde{W}_\pm(\mathbf{0})\|_{H^h}^2 \|\widetilde{W}_\pm(\mathbf{0})\|_{H^{-1}} \Big) \\ &+ \widetilde{\varepsilon}^2 C \Big(\|\widetilde{W}(\widetilde{\mathbf{T}})\|_{H^h}^2 (1 + \|\widetilde{W}(\widetilde{\mathbf{T}})\|_{H^h})^{h+1} + \|\widetilde{W}(\mathbf{0})\|_{H^h}^2 (1 + \|\widetilde{W}(\mathbf{0})\|_{H^h})^{h+1} \Big) \end{aligned}$$

as long as $\|\widetilde{W}(\mathbf{0})\|_{H^{-1}}^2, \|\widetilde{W}(\widetilde{\mathbf{T}})\|_{H^{-1}}^2 < 4\delta$. Define

$$g_\pm(r) = \sup_{\{|\widetilde{\mathbf{T}}|\le r\}} \left(\|\widetilde{W}_\pm(\widetilde{\mathbf{T}})\|_{H^h}^2 + \frac{\widetilde{\varepsilon}^2}{8} \|\partial_X (\widetilde{W}_+ - \widetilde{W}_-)(\widetilde{\mathbf{T}})\|_{H^h}^2 \right)$$
$$r^* = \inf\{r > 0 : g_+(r) + g_-(r) \ge 4\delta\}$$

and note that $g_\pm$ is continuous.

By choosing $\varepsilon < \varepsilon_0$ for some appropriate $\varepsilon_0 = \varepsilon_0(\delta, \|W(\mathbf{0})\|_{H^{h+1}})$, we can ensure that $\widetilde{\varepsilon} \in (\varepsilon, 1)$. Then the above inequality implies for $r \in [0, r^*)$ that for some $C_0 = C_0(h)$ we have

$$\begin{aligned} |g_\pm(r) - g_\pm(0)| &\le C_0 (g_\pm(r) + g_\pm(0))(g_+(r) + g_-(r))^{\frac{1}{2}} \\ &\quad + \widetilde{\varepsilon}^2 C_0 (g_+(r) + g_-(r) + g_+(0) + g_-(0))(1 + g_+(r) + g_-(r) + g_+(0) + g_-(0))^{h+1} . \end{aligned}$$

By Lemma 4.3.1 there exists some $\delta_0 = \delta_0(h) > 0$, $\widetilde{\varepsilon}_0 = \widetilde{\varepsilon}_0(h) > 0$, and $C_1 = C_1(h) > 0$ such that $g_+(0) + g_-(0) < \delta_0$ and $\widetilde{\varepsilon} < \widetilde{\varepsilon}_0$ imply

$$g_\pm(r) \le \begin{cases} 3 g_\pm(0) & , g_\pm(0) > 0 \\ 3 g_\mp(0) & , g_\pm(0) = 0 \end{cases} + \widetilde{\varepsilon}^2 C_1 (g_+(0) + g_-(0))$$

for all $r \in [0, r^*)$. Since $g_\pm(0) \lesssim \|\widetilde{W}(\mathbf{0})\|_{H^{h+1}}^2$ we can obtain $g_+(0) + g_-(0) < \min\{\delta_0, \delta\}$ and also $\widetilde{\varepsilon} < \widetilde{\varepsilon}_0$ from an appropriately small choice of $\varepsilon_0 = \varepsilon_0(h, \delta, \delta_0, \|W(\mathbf{0})\|_{H^{h+1}})$. Then the estimate just proved implies $g_+(r) + g_-(r) < 4\delta$, so $r^* = \infty$.

If $g_\pm(0) = 0$ then $(\widetilde{W}_+ - \widetilde{W}_-)(\mathbf{0})$ is constant in $\widetilde{X}$, hence $W_+(\mathbf{0}) = W_-(\mathbf{0})$. Therefore $g_\pm(0) = 0$ implies $W(\mathbf{0}) = 0$, which is a trivial case. If $g_\pm(0) > 0$ then we finally obtain (4.1.23). $\square$

4.3.3. PROOF OF THEOREM 4.1.12

We are now ready to give the proof of our main result. To simplify the notation, we set $\mathbf{c} = \mathbf{c}^{\mathsf{NLS} \mapsto \varepsilon \widetilde{\mathsf{GP}}}$.

Proof of Theorem 4.1.12. Let $\mathbf{T} \in \mathbb{R}^{N+1}$ and recall that $\mathbf{T} = (T_0, \dots, T_N)$. We shall use an energy inequality with Grönwall's lemma. Note, however, that we do not wish to apply Grönwall's lemma several times, i. e. for each flow in the hierarchy, as this would require reusing solutions obtained by our well-posedness theory as initial data and lead to worse estimates. Instead, we shall apply Grönwall's lemma only once.

Continuous path from $\mathbf{0}$ to $\mathbf{T}$. Let $n \in \{0, \dots, N\}$ and define $m = m(n) = \lfloor \frac{n-1}{2} \rfloor \in \{-1, \dots, M\}$ where $M = \lfloor \frac{N-1}{2} \rfloor$. In order to simplify the argument below we assume that $|T_0|, \dots, |T_N| > 0$. The degenerate cases only require minor modifications. Define $\tau_n = |T_0| + \dots + |T_n|$ as well as $\tau_{-1} = 0$ and set

$$(\mathbf{T}^\tau)_n = \begin{cases} 0 & , \tau \in [0, \tau_{n-1}) \\ \frac{\tau - \tau_{n-1}}{\tau_n - \tau_{n-1}} T_n & , \tau \in [\tau_{n-1}, \tau_n) \\ T_n & , \tau \in [\tau_n, \infty) \end{cases} .$$

Then $(\mathbf{T}^\tau)_{\tau \in [0,\infty)}$ is a continuous path with $\mathbf{T}^0 = \mathbf{0}$ and $(\mathbf{T}^\tau)_n = T_n$ if $\tau \ge \tau_n$. In particular $\mathbf{T}^{\tau_N} = \mathbf{T}$. We use Theorem 4.1.8 to write

$$\begin{aligned} W(\mathbf{c}\mathbf{T}^\tau) &= W(\mathbf{0}) + \int_0^{\tau_N \wedge \tau} \frac{\mathrm{d}}{\mathrm{d}\tau'} W(\mathbf{c}\mathbf{T}^{\tau'}) \,\mathrm{d}\tau' \\ &= W(\mathbf{0}) + \sum_{n=0}^N \operatorname{sign}(T_n) \int_{\tau_{n-1} \wedge \tau}^{\tau_n \wedge \tau} \frac{\mathrm{d}}{\mathrm{d}(\mathbf{T}^{\tau'})_n} W(\mathbf{c}\mathbf{T}^{\tau'}) \,\mathrm{d}\tau' \\ &= W(\mathbf{0}) + \sum_{n=0}^N \operatorname{sign}(T_n) \int_{\tau_{n-1} \wedge \tau}^{\tau_n \wedge \tau} \begin{pmatrix} -(-1)^n \frac{1}{2} \partial_X \frac{\delta}{\delta W_+} \mathcal{E}_m^{\mathsf{KdV}}(W_+(\mathbf{c}\mathbf{T}^{\tau'})) \\ -\frac{1}{2} \partial_X \frac{\delta}{\delta(-W_-)} \mathcal{E}_m^{\mathsf{KdV}}(-W_-(\mathbf{c}\mathbf{T}^{\tau'})) \end{pmatrix} + \varepsilon^2 \mathbf{R}_n(\mathbf{T}^{\tau'}) \,\mathrm{d}\tau' . \end{aligned}$$

Similarly, we can write

$$\begin{aligned}
\begin{pmatrix} U_+(\mathbf{b}^+\mathbf{T}^\tau) \\ U_-(\mathbf{b}^-\mathbf{T}^\tau) \end{pmatrix} &= U(\mathbf{0}) + \int_0^{\tau_N\wedge\tau} \frac{\mathrm{d}}{\mathrm{d}\tau'} \begin{pmatrix} U_+(\mathbf{b}^+\mathbf{T}^{\tau'}) \\ U_-(\mathbf{b}^-\mathbf{T}^{\tau'}) \end{pmatrix} \mathrm{d}\tau' \\
&= U(\mathbf{0}) + \sum_{n=0}^N \mathrm{sign}(T_n) \int_{\tau_{n-1}\wedge\tau}^{\tau_n\wedge\tau} \frac{\mathrm{d}}{\mathrm{d}(\mathbf{T}^{\tau'})_n} \begin{pmatrix} U_+(\mathbf{b}^+\mathbf{T}^{\tau'}) \\ U_-(\mathbf{b}^-\mathbf{T}^{\tau'}) \end{pmatrix} \mathrm{d}\tau' \\
&= U(\mathbf{0}) + \sum_{n=0}^N \mathrm{sign}(T_n) \int_{\tau_{n-1}\wedge\tau}^{\tau_n\wedge\tau} \begin{pmatrix} -(-1)^n \frac{1}{2}\partial_X \frac{\delta}{\delta U_+} \mathcal{E}_m^{\mathsf{KdV}}(U_+(\mathbf{b}^+\mathbf{T}^{\tau'})) \\ -\frac{1}{2}\partial_X \frac{\delta}{\delta(-U_-)} \mathcal{E}_m^{\mathsf{KdV}}(-U_-(\mathbf{b}^-\mathbf{T}^{\tau'})) \end{pmatrix} \mathrm{d}\tau' .
\end{aligned}$$

Energy estimate. Define

$$V(\mathbf{T}) = \begin{pmatrix} V_+(\mathbf{T}) \\ V_-(\mathbf{T}) \end{pmatrix} = \begin{pmatrix} U_+(\mathbf{b}^+\mathbf{T}) \\ U_-(\mathbf{b}^-\mathbf{T}) \end{pmatrix} - W(\mathbf{c}\mathbf{T}) .$$

Since $W(\mathbf{0}) = U(\mathbf{0})$ and $\mathbf{R}_0 = \mathbf{R}_1 = 0$, we know that $V(\mathbf{T}^\tau) = 0$ for $\tau \in [0, \tau_1]$. We therefore have

$$V(\mathbf{T}^\tau) = \sum_{n=2}^N \mathrm{sign}(T_n) \int_{\tau_{n-1}\wedge\tau}^{\tau_n\wedge\tau} (\mathbf{F}_n(\tau') - \varepsilon^2 \mathbf{R}_n(\mathbf{T}^{\tau'}))\, \mathrm{d}\tau' .$$

with

$$\mathbf{F}_n(\tau) = \begin{pmatrix} -(-1)^n \frac{1}{2}\partial_X \frac{\delta}{\delta W_+}\big(\mathcal{E}_m^{\mathsf{KdV}}(W_+(\mathbf{c}\mathbf{T}^\tau) + V_+(\mathbf{T}^\tau)) - \mathcal{E}_m^{\mathsf{KdV}}(W_+(\mathbf{c}\mathbf{T}^\tau))\big) \\ -\frac{1}{2}\partial_X \frac{\delta}{\delta(-W_-)}\big(\mathcal{E}_m^{\mathsf{KdV}}(-W_-(\mathbf{c}\mathbf{T}^\tau) - V_-(\mathbf{T}^\tau)) - \mathcal{E}_m^{\mathsf{KdV}}(-W_-(\mathbf{c}\mathbf{T}^\tau))\big) \end{pmatrix} \text{ for } \tau \in [\tau_{n-1}, \tau_n) .$$

In particular, for all $h \in \{0, \dots, \lfloor \frac{s-N}{2} \rfloor\}$ we have

$$\int_{\mathbb{R}} (\partial_X^h V(\mathbf{T}^\tau))^2\, \mathrm{d}X = \sum_{n=2}^N \mathrm{sign}(T_n) \int_{\tau_{n-1}\wedge\tau}^{\tau_n\wedge\tau} \int_{\mathbb{R}} 2\partial_X^h V(\mathbf{T}^{\tau'}) \partial_X^h (\mathbf{F}_n(\tau') - \varepsilon^2 \mathbf{R}_n(\mathbf{T}^{\tau'}))\, \mathrm{d}X\, \mathrm{d}\tau' .$$

For notational convenience, let us define for a set of functions F in the variable X and parameters $K, L \in \mathbb{N}$ the sets

$$\mathcal{O}^{K,L}(F) = \left\{ \sum_{k=0}^K \sum_{\substack{l=(l_1,\dots,l_k)\in\mathbb{N}^k \\ |l|\le L}} \sum_{f=(f_1,\dots,f_k)\in F^k} c_f^{k,l} \prod_{j=1}^k \partial_X^{l_j} f_j : c_f^{k,l} \in \mathbb{C}[\varepsilon] \right\} .$$

Studying the recurrence relations (4.2.42)–(4.2.44), (4.2.45)–(4.2.47), as we do in the proof of Theorem 4.1.1 at the end of Section 4.2, reveals that $\frac{\delta}{\delta W_\pm} \mathcal{H}_n^{\varepsilon\widetilde{\mathsf{GP}}}(W) \in \mathcal{O}^{n+1, 2\lfloor \frac{n}{2} \rfloor}(W_+, W_-)$. A similar analysis of (4.2.48) yields $\frac{\delta}{\delta U}\mathcal{E}_m^{\mathsf{KdV}}(U) \in \mathcal{O}^{m+1,2m}(U)$. Then

$$\mathbf{R}_n \in \mathcal{O}^{n+2, 2\lfloor \frac{n}{2} \rfloor + 1}(W_+, W_-) + \frac{\mathcal{O}^{n+2, 2\lfloor \frac{n}{2} \rfloor + 1}(W_+, W_-)}{1 + \frac{\varepsilon^2}{2}(W_+ - W_-)} + \frac{\mathcal{O}^{n+2, 2\lfloor \frac{n}{2} \rfloor + 1}(W_+, W_-)}{(1 + \frac{\varepsilon^2}{2}(W_+ - W_-))^2} .$$

We use Sobolev embedding $H^1 \hookrightarrow L^\infty$ and the H^h-algebra property for $h \geq 1$ as well as [20, Theorem 2.87] to find a constant $C(n, h, \delta) > 0$ such that

$$\|\mathbf{R}_n\|_{H^h} \le C(n, h, \delta)\big(1 + \|W\|_{H^{n+1+h}}^{n+4}\big) .$$

Then

$$\int_{\mathbb{R}} \partial_X^h V \varepsilon^2 \partial_X^h \mathbf{R}_n\, \mathrm{d}X \le \mathbb{1}_{\{n\ge 2\}}\big(\|V\|_{H^h}^2 + \varepsilon^4 C(N, h, \delta)^2 \big(1 + \|W\|_{H^{n+1+h}}^{2n+8}\big)\big) .$$

We are able to include the indicator function because $\mathbf{R}_0 = \mathbf{R}_1 = 0$. A more detailed analysis of (4.2.48) yields the representation

$$\frac{1}{2}\partial_X \frac{\delta}{\delta U} \mathcal{E}_m^{\mathsf{KdV}} = (-1)^m \partial_X^{2m+1} U + \sum_{k=2}^{m+1} \sum_{\substack{l=(l_1,\dots,l_k)\in\mathbb{N}^k \\ |l|+2k=2m+3}} c^{k,l} \prod_{j=1}^k \partial_X^{l_j} U \tag{4.3.1}$$

with coefficients $c^{k,l} \in \mathbb{C}$. Then

$$
\begin{aligned}
&\partial_X \frac{\delta}{\delta(\pm W_\pm)} \big(\mathcal{E}_m^{\mathsf{KdV}}(\pm W_\pm \pm V_\pm) - \mathcal{E}_m^{\mathsf{KdV}}(\pm W_\pm)\big) \\
&= \pm(-1)^m \partial_X^{2m+1} V_\pm + \sum_{k=2}^{m+1} \sum_{\substack{l=(l_1,\dots,l_k)\in\mathbb{N}^k \\ |l|\le 2m-1}} c^{k,l} \sum_{i=1}^k \partial_X^{l_i}(\pm V_\pm) \prod_{j=1}^{i-1} \partial_X^{l_j}(\pm W_\pm) \prod_{j=i+1}^{k} \partial_X^{l_j}(\pm W_\pm \pm V_\pm)\,.
\end{aligned}
$$

Using integration by parts and $\partial_X^{m+h} V_\pm \in H^1(\mathbb{R})$, we find that

$$
\int_{\mathbb{R}} \partial_X^h V_\pm \partial_X^h \partial_X^{2m+1} V_\pm \,\mathrm{d}X = \int_{\mathbb{R}} (-1)^m \partial_X \left(\frac{1}{2}\big(\partial_X^{m+h} V_\pm\big)^2\right) \mathrm{d}X = 0\,.
$$

Similarly, using repeated integration by parts and the Sobolev embedding $H^1 \longhookrightarrow L^\infty$, we estimate

$$
\begin{aligned}
&\left| \sum_{k=2}^{m+1} \sum_{\substack{l=(l_1,\dots,l_k)\in\mathbb{N}^k \\ |l|\le 2m-1}} c^{k,l} \sum_{i=1}^k \int_{\mathbb{R}} \partial_X^h V_\pm \partial_X^h \left(\partial_X^{l_i} V_\pm \prod_{j=1}^{i-1} \partial_X^{l_j}(\pm W_\pm) \prod_{j=i+1}^{k} \partial_X^{l_j}(\pm U_\pm) \right) \mathrm{d}X \right| \\
&\le C^{m+1} \sum_{r=0}^h \sum_{k=2}^{m+1} \sum_{\substack{l=(l_1,\dots,l_k)\in\mathbb{N}^k \\ |l|\le 2m-1+2(h-r)}} \sum_{i=1}^k \int_{\mathbb{R}} (\partial_X^r V_\pm)^2 \prod_{j=1}^{i-1} \partial_X^{l_j}(\pm W_\pm) \prod_{j=i+1}^{k} \partial_X^{l_j}(\pm U_\pm)\,\mathrm{d}X \\
&\le C^{m+1} \sum_{r=0}^h \int_{\mathbb{R}} (\partial_X^r V_\pm)^2 \,\mathrm{d}X \sum_{k=2}^{m+1} \sum_{\substack{l=(l_1,\dots,l_k)\in\mathbb{N}^k \\ |l|\le 2m-1+2(h-r)}} \sum_{i=1}^k \prod_{\substack{j=1 \\ j\ne i}}^{k} \big(\|\partial_X^{l_j} W_\pm\|_{L^\infty} + \|\partial_X^{l_j} U_\pm\|_{L^\infty}\big) \\
&\le C^{m+1} \sum_{r=0}^h \int_{\mathbb{R}} (\partial_X^r V_\pm)^2 \,\mathrm{d}X \big(1 + \|(W_\pm, U_\pm)\|_{H^{2(m+h-r)}}^m\big)\,.
\end{aligned}
$$

Grönwall's lemma for linear systems. In summary, for all $l \in \{0, \dots, h\}$ and $\tau \in [0, \tau_N]$ we have

$$
\int_{\mathbb{R}} (\partial_X^l V(\mathbf{T}^\tau))^2 \,\mathrm{d}X \le \mathbb{1}_{\{\tau\ge\tau_1\}} C(N,h,\delta) \left(a_l + \int_0^\tau \sum_{r=0}^h H_{l,r} \int_{\mathbb{R}} (\partial_X^r V(\mathbf{T}^{\tau'}))^2 \,\mathrm{d}X \,\mathrm{d}\tau' \right),
$$

where we consider the real numbers

$$
\begin{aligned}
a_l &= (\tau_N - \tau_1)\varepsilon^4 \big(1 + \|W\|_{L^\infty_{\mathbf{T}} H_X^{N+1+l}}^{2N+8}\big) \\
H_{l,r} &= \mathbb{1}_{r\le l} \big(1 + \|(W,U)\|_{L^\infty_{\mathbf{T}} H_X^{2(M+l-r)}}^{M}\big)
\end{aligned}
$$

as the entries of a vector $a \in \mathbb{R}^{h+1}$ and a matrix $H \in \mathbb{R}^{(h+1)\times(h+1)}$. Grönwall's lemma for linear systems (see [45, Theorem]) yields

$$
\begin{aligned}
\sup_{\tau\in[0,\tau_N]} \int_{\mathbb{R}} (\partial_X^h V(\mathbf{T}^\tau))^2 \,\mathrm{d}X &\le C \left(\left(1 + \int_0^{\tau_N-\tau_1} e^{C(\tau_N-\tau_1-\tau)H} H \,\mathrm{d}\tau \right) a \right)_h \\
&= C(e^{C(\tau_N-\tau_1)H} a)_h
\end{aligned}
$$

for some constant $C = C(N,h,\delta) > 0$. It remains to estimate $(e^{C(\tau_N-\tau_1)H} a)_h$. We define some notation for the part $\widetilde{H}$ of H which does not lie on the diagonal:

$$
\begin{aligned}
\widetilde{H} &= H - \left(1 + \|(W,U)\|_{L^\infty_{\mathbf{T}} H_X^{2M}}^M\right) \\
\widetilde{H}_{l,r} &= \mathbb{1}_{\{0\le r<l\}} g_{l-r} \quad \text{where} \quad g_j = 1 + \|(W,U)\|_{L^\infty_{\mathbf{T}} H_X^{2(M+j)}}^M\,.
\end{aligned}
$$

Then since $\widetilde{H}$ and $H-\widetilde{H}$ commute, we have

$$|(e^{(\tau_N-\tau_1)H}a)_h| \le |(e^{(\tau_N-\tau_1)\widetilde{H}}a)_h| e^{(\tau_N-\tau_1)\left(1+\|(W,U)\|^M_{L^\infty_{\mathbf{T}}H^{2M}_X}\right)} .$$

Next, we estimate for $l \ge 1$

$$|(\widetilde{H}^l)_{h,r}| = \sum_{h=j_0>\dots>j_l=r} g_{j_0-j_1}\cdots g_{j_{l-1}-j_l} \le \binom{h-r-1}{l-1} g^l_{h-r}$$

and then

$$(e^{(\tau_N-\tau_1)\widetilde{H}}a)_h = \sum_{l=0}^{\infty} \frac{(\tau_N-\tau_1)^l}{l!}(\widetilde{H}^l a)_h = \sum_{l=0}^{\infty} \frac{(\tau_N-\tau_1)^l}{l!} \sum_{r=0}^{h} (\widetilde{H}^l)_{h,r} a_r \le (h+1)a_h(1+(\tau_N-\tau_1)g_h)^h .$$

This yields

$$\begin{aligned} |(e^{(\tau_N-\tau_1)H}a)_h| &\le \varepsilon^4(\tau_N-\tau_1)C\big(1+\|W\|^{2N+8}_{L^\infty_{\mathbf{T}}H^{N+1+h}_X}\big)\big(1+(\tau_N-\tau_1)\|(W,U)\|^M_{L^\infty_{\mathbf{T}}H^{2(M+h)}_X}\big)^h \\ &\quad \times \exp\Big(C(\tau_N-\tau_1)\big(1+\|(W,U)\|^M_{L^\infty_{\mathbf{T}}H^{2M}_X}\big)\Big) . \end{aligned}$$

□

4.A. Appendix: Approximation of the perturbation determinant $\alpha^{\varepsilon\mathsf{GP}}$ by α^{KdV}

4.A.1. Definitions

Let $n \in \{1,2\}$. We use the convention of identifying functions $f \in L^2(\mathbb{R};\mathbb{C}^n)$ with the operator of multiplication by f. We write $(L(L^2(\mathbb{R};\mathbb{C}^n)), \|\cdot\|_{L^2\to L^2})$ for the space of bounded linear operators on $L^2(\mathbb{R};\mathbb{C}^n)$ equipped with the operator norm, and say that an operator $A \in L(L^2(\mathbb{R};\mathbb{C}^n))$ is Hilbert-Schmidt if there exists a kernel $K \in L^2(\mathbb{R}\times\mathbb{R};\mathbb{C}^{n\times n})$ such that

$$Af(X) = \int_{\mathbb{R}} K(X,Y)f(Y)\,\mathrm{d}Y .$$

In this case we define the Hilbert-Schmidt norm as

$$\|A\|^2_{\mathrm{HS}} = \int_{\mathbb{R}^2} |K(X,Y)|^2\,\mathrm{d}X\,\mathrm{d}Y .$$

We say that an operator $A \in L(L^2(\mathbb{R};\mathbb{C}^n))$ is trace-class if $A = A_1A_2$ for two Hilbert-Schmidt operators A_1 and A_2. Writing K_1 and K_2 for their kernels, we define the trace

$$\mathrm{tr}[A] = \mathrm{tr}[A_1A_2] = \int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{tr}[K_1(X,Y)K_2(Y,X)]\,\mathrm{d}X\,\mathrm{d}Y .$$

In particular

$$\|A\|^2_{\mathrm{HS}} = \mathrm{tr}[A^*A] .$$

If the trace $\mathrm{tr}[K]$ of the kernel K of A is continuous, then

$$\mathrm{tr}[A] = \int_{\mathbb{R}} \mathrm{tr}[K(X,X)]\,\mathrm{d}X .$$

The following lemma collects some results regarding the trace and the Hilbert-Schmidt norm.

Lemma 4.A.1 ([115, Lemmas 1.4, 1.5]). *Let $A, A_j : L^2(\mathbb{R};\mathbb{C}^n) \to L^2(\mathbb{R};\mathbb{C}^n)$, $1 \le j \le l$, $2 \le l \in \mathbb{N}$ be linear operators with finite Hilbert-Schmidt norm. Then*

$$\|A\|_{L^2\to L^2} \le \|A\|_{\mathrm{HS}} \tag{4.A.1}$$

$$|\operatorname{tr}[A_1 \dots A_l]| \le \|A_1\|_{\mathrm{HS}} \dots \|A_l\|_{\mathrm{HS}}\,. \tag{4.A.2}$$

If $A = A(s) \in C_b^1(\mathbb{R};\mathrm{HS})$ with $\sup_{s\in\mathbb{R}} \|A(s)\|_{\mathrm{HS}} < \frac{1}{3}$, then there exists an open interval $I \ni 0$ on which the series

$$\alpha(s) = \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \operatorname{tr}[A(s)^l]$$

is absolutely convergent and defines a function $\alpha \in C^1(I;\mathbb{C})$ with derivative

$$\partial_s\alpha(s) = \sum_{l=2}^{\infty} (-1)^l \operatorname{tr}[A(s)^{l-1}\partial_s A(s)]\,. \tag{4.A.3}$$

We have the bound

$$\sup_{s\in I} |\alpha(s)| \le \frac{2}{3} \sup_{s\in I} \|A(s)\|_{\mathrm{HS}}^2\,.$$

Recall from (4.2.1) the spectral parameters $(\lambda, z, k) \in \mathbb{K} \subset \mathbb{C}^3$ which satisfy $z^2 = (\sqrt{2}\varepsilon k)^2 = \lambda^2 - 1$. We assume for the rest of this section that $z, k \in i\,\mathbb{R}_+$ are large and choose $\lambda = \sqrt{1 + 2\varepsilon^2 k^2}$ as the principal square root. We define

$$\kappa = \frac{z}{\sqrt{2}\varepsilon i} = \frac{k}{i} \in \mathbb{R}_+ \qquad\qquad \kappa^2 = \frac{1-\lambda^2}{2\varepsilon^2}\,.$$

For notational convenience, we use the κ-dependent Sobolev norms

$$\|f\|_{H_\kappa^{-1}} = \left(\int_{\mathbb{R}} \frac{1}{\xi^2 + 4\kappa^2} |\widehat{f}(\xi)|^2 \,\mathrm{d}\xi\right)^{\frac{1}{2}} \le \min\left\{\frac{1}{\min\{1, 2\kappa\}} \|f\|_{H^{-1}}, \frac{1}{2\kappa} \|f\|_{L^2}\right\}$$

$$\|f\|_{L_\kappa^2} = \left(\int_{\mathbb{R}} \frac{\xi^2 + 2\kappa^2}{\xi^2 + 4\kappa^2} |\widehat{f}(\xi)|^2 \,\mathrm{d}\xi\right)^{\frac{1}{2}} \le \|f\|_{L^2}\,.$$

Here the Fourier transform is defined by

$$\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ix\xi} f(x)\,\mathrm{d}x\,.$$

We assume $W = (W_+, W_-) \in L^2(\mathbb{R};\mathbb{R}^2)$ or $W \in \mathcal{S}(\mathbb{R};\mathbb{R}^2)$ depending on the context. We are interested in the Lax operators

$$L^{\varepsilon\mathsf{GP}}(W) = \begin{pmatrix} -\varepsilon^2 W_- + 1 & i\sqrt{2}\varepsilon\partial_X \\ i\sqrt{2}\varepsilon\partial_X & -\varepsilon^2 W_+ - 1 \end{pmatrix} \qquad L^{\mathsf{KdV}}(\pm W_\pm) = -\partial_X^2 \pm W_\pm$$

$$L_0^{\varepsilon\mathsf{GP}} = \begin{pmatrix} 1 & i\sqrt{2}\varepsilon\partial_X \\ i\sqrt{2}\varepsilon\partial_X & -1 \end{pmatrix} \qquad L_0^{\mathsf{KdV}} = -\partial_X^2$$

and their resolvents

$$R^{\varepsilon\mathsf{GP}}(W) = \left(\frac{L^{\varepsilon\mathsf{GP}}(W) - \lambda}{\varepsilon^2}\right)^{-1} \qquad R^{\mathsf{KdV}}(\pm W_\pm) = (L^{\mathsf{KdV}}(\pm W_\pm) + \kappa^2)^{-1}$$

$$R_0^{\varepsilon\mathsf{GP}} = \left(\frac{L_0^{\varepsilon\mathsf{GP}} - \lambda}{\varepsilon^2}\right)^{-1} \qquad R_0^{\mathsf{KdV}} = (L_0^{\mathsf{KdV}} + \kappa^2)^{-1}\,.$$

Define

$$\mathcal{W} = \begin{pmatrix} -W_- & 0 \\ 0 & -W_+ \end{pmatrix}$$

and note that $L^{\varepsilon\mathsf{GP}}(W) = L_0^{\varepsilon\mathsf{GP}} + \varepsilon^2\mathcal{W}$ and $L^{\mathsf{KdV}}(\pm W_\pm) = L_0^{\mathsf{KdV}} \pm W_\pm$. In the following, we often drop the dependence of the Lax operators and resolvents on W and $(\pm W_\pm)$ from our notation. With

$$\mathcal{U} = \begin{pmatrix} \frac{\lambda+1}{2} & \frac{i\varepsilon}{\sqrt{2}}\partial_X \\ \frac{i\varepsilon}{\sqrt{2}}\partial_X & \frac{\lambda-1}{2} \end{pmatrix}$$

we can write

$$R_0^{\varepsilon\mathsf{GP}} = R_0^{\mathsf{KdV}}\mathcal{U} = \mathcal{U}R_0^{\mathsf{KdV}} = \sqrt{R_0^{\mathsf{KdV}}}\,\mathcal{U}\sqrt{R_0^{\mathsf{KdV}}}\,.$$

Next, we use the Green's functions

$$\begin{aligned} G_0^{\mathsf{KdV}} &= \frac{1}{\sqrt{2\pi}}\mathcal{F}^{-1}\left[R_0^{\mathsf{KdV}}\Big|_{\partial_X = i\xi}\right] = \frac{1}{2\kappa}e^{-\kappa|X|} \\ G_0^{\varepsilon\mathsf{GP}} &= \frac{1}{\sqrt{2\pi}}\mathcal{F}^{-1}\left[R_0^{\varepsilon\mathsf{GP}}\Big|_{\partial_X = i\xi}\right] = \frac{1}{2\kappa}e^{-\kappa|X|}\begin{pmatrix} \frac{\lambda+1}{2} & -\frac{i\varepsilon}{\sqrt{2}}\kappa\,\mathrm{sign}(X) \\ -\frac{i\varepsilon}{\sqrt{2}}\kappa\,\mathrm{sign}(X) & \frac{\lambda-1}{2} \end{pmatrix} \end{aligned}$$

to define the kernels

$$K_0^{\mathsf{KdV}}(X,Y) = G_0^{\mathsf{KdV}}(X-Y) \qquad\qquad K_0^{\varepsilon\mathsf{GP}}(X,Y) = G_0^{\varepsilon\mathsf{GP}}(X-Y)\,.$$

Then $K_0^{\mathsf{KdV}} \in L^\infty(\mathbb{R}\times\mathbb{R})$ and $K_0^{\varepsilon\mathsf{GP}} \in (L^\infty(\mathbb{R}\times\mathbb{R}))^{2\times 2}$ are the integral kernels of R_0^{KdV} and $R_0^{\varepsilon\mathsf{GP}}$ respectively. In order to give a formula for the Green's function $G^{\varepsilon\mathsf{GP}}$, for which $K^{\varepsilon\mathsf{GP}}(X,Y) = G^{\varepsilon\mathsf{GP}}(X,Y)$ is the kernel of $R^{\varepsilon\mathsf{GP}}$, we use the Jost solutions $\Phi_j^{\varepsilon\mathsf{GP},\pm}$, their boundary data $E^{\varepsilon\mathsf{GP},\pm}$, and the transmission coefficient $a^{\varepsilon\mathsf{GP}}$. Their definitions may be found in Section 4.2. For simplicity, we suppress the dependence of these scattering quantities on k from our notation in the following lemma.

Lemma 4.A.2. *The Green's function $G^{\varepsilon\mathsf{GP}}$ of the operator*

$$\frac{L^{\varepsilon\mathsf{GP}} - \lambda}{\varepsilon^2} = \begin{pmatrix} -W_- - \frac{\lambda-1}{\varepsilon^2} & \frac{i\sqrt{2}}{\varepsilon}\partial_X \\ \frac{i\sqrt{2}}{\varepsilon}\partial_X & -W_+ - \frac{\lambda+1}{\varepsilon^2} \end{pmatrix},$$

defined by the property that $\left(\frac{L^{\varepsilon\mathsf{GP}}-\lambda}{\varepsilon^2}\right)G^{\varepsilon\mathsf{GP}}(X,Y) = \delta_0(X-Y)$, is given by

$$G^{\varepsilon\mathsf{GP}}(X,Y) = \frac{\varepsilon}{i\sqrt{2}}\frac{1}{a^{\varepsilon\mathsf{GP}}(k)\det E^{\varepsilon\mathsf{GP},+}}\begin{cases} \begin{pmatrix} -\Phi_{1,1}^{\varepsilon\mathsf{GP},-}(X)\Phi_{2,1}^{\varepsilon\mathsf{GP},+}(Y) & \Phi_{1,1}^{\varepsilon\mathsf{GP},-}(X)\Phi_{2,2}^{\varepsilon\mathsf{GP},+}(Y) \\ -\Phi_{1,2}^{\varepsilon\mathsf{GP},-}(X)\Phi_{2,1}^{\varepsilon\mathsf{GP},+}(Y) & \Phi_{1,2}^{\varepsilon\mathsf{GP},-}(X)\Phi_{2,2}^{\varepsilon\mathsf{GP},+}(Y) \end{pmatrix} & ,X > Y \\ \begin{pmatrix} -\Phi_{2,1}^{\varepsilon\mathsf{GP},+}(X)\Phi_{1,1}^{\varepsilon\mathsf{GP},-}(Y) & \Phi_{2,1}^{\varepsilon\mathsf{GP},+}(X)\Phi_{1,2}^{\varepsilon\mathsf{GP},-}(Y) \\ -\Phi_{2,2}^{\varepsilon\mathsf{GP},+}(X)\Phi_{1,1}^{\varepsilon\mathsf{GP},-}(Y) & \Phi_{2,2}^{\varepsilon\mathsf{GP},+}(X)\Phi_{1,2}^{\varepsilon\mathsf{GP},-}(Y) \end{pmatrix} & ,X < Y \end{cases}.$$

Proof. For every $Y \in \mathbb{R}$ and $X \in \mathbb{R}\setminus\{Y\}$ each column of $G^{\varepsilon\mathsf{GP}}(X,Y)$ is a solution to the scattering problem $\left(\frac{L^{\varepsilon\mathsf{GP}}-\lambda}{\varepsilon^2}\right)G^{\varepsilon\mathsf{GP}}(X,Y) = 0$. Furthermore, the jump condition

$$\begin{aligned} &\lim_{X\searrow Y} G^{\varepsilon\mathsf{GP}}(X,Y) - \lim_{X\nearrow Y} G^{\varepsilon\mathsf{GP}}(X,Y) \\ &= \frac{\varepsilon}{i\sqrt{2}}\frac{1}{a^{\varepsilon\mathsf{GP}}(k)\det E^{\varepsilon\mathsf{GP},+}}\begin{pmatrix} \Phi_{2,1}^{\varepsilon\mathsf{GP},+}\Phi_{1,1}^{\varepsilon\mathsf{GP},-} - \Phi_{1,1}^{\varepsilon\mathsf{GP},-}\Phi_{2,1}^{\varepsilon\mathsf{GP},+} & \Phi_{1,1}^{\varepsilon\mathsf{GP},-}\Phi_{2,2}^{\varepsilon\mathsf{GP},+} - \Phi_{2,1}^{\varepsilon\mathsf{GP},+}\Phi_{1,2}^{\varepsilon\mathsf{GP},-} \\ \Phi_{2,2}^{\varepsilon\mathsf{GP},+}\Phi_{1,1}^{\varepsilon\mathsf{GP},-} - \Phi_{1,2}^{\varepsilon\mathsf{GP},-}\Phi_{2,1}^{\varepsilon\mathsf{GP},+} & \Phi_{1,2}^{\varepsilon\mathsf{GP},-}\Phi_{2,2}^{\varepsilon\mathsf{GP},+} - \Phi_{2,2}^{\varepsilon\mathsf{GP},+}\Phi_{1,2}^{\varepsilon\mathsf{GP},-} \end{pmatrix}(Y) \\ &= \frac{\varepsilon}{i\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

holds, which implies that $\left(\frac{L^{\varepsilon\mathsf{GP}}-\lambda}{\varepsilon^2}\right)G^{\varepsilon\mathsf{GP}}(X,Y) = \delta_0(X-Y)$ in the sense of distributions. ☐

An analogous representation for the Green's function of L^{KdV} exists, but is not necessary for our purposes.

4.A.2. Approximation of perturbation determinants

In this section we develop the theory of perturbation determinants of the Lax operators $L^{\varepsilon\mathsf{GP}}$ and L^{KdV}, following in structure the seminal papers [91, 114, 115] which concern L^{KdV} and L^{NLS}. These techniques have recently been used effectively to construct low-regularity conserved quantities and prove well-posedness of various completely integrable PDEs (see also [38, 112]). Our contribution is a proof that an analogue of the approximation property observed in Theorem 4.1.1 and Lemma 4.2.7 also holds for the perturbation determinants. This is the content of Lemma 4.A.6 below. The perturbation determinants in question are the quantities

$$\log a^{\varepsilon\mathsf{GP}}(k;W) = \alpha^{\varepsilon\mathsf{GP}}(\kappa;W) - \operatorname{tr}[R_0^{\varepsilon\mathsf{GP}}\mathcal{W}] = -\log\det\left(\left(\frac{L_0^{\varepsilon\mathsf{GP}}-\lambda}{\varepsilon^2}\right)^{-1}\frac{L^{\varepsilon\mathsf{GP}}(W)-\lambda}{\varepsilon^2}\right)$$

$$\log a^{\mathsf{KdV}}(k;\pm W_\pm) = \alpha^{\mathsf{KdV}}(\kappa;\pm W_\pm) - \operatorname{tr}[R_0^{\mathsf{KdV}}(\pm W_\pm)] = -\log\det\Big(\big(L_0^{\mathsf{KdV}}+\kappa^2\big)^{-1}\big(L^{\mathsf{KdV}}(\pm W_\pm)+\kappa^2\big)\Big)\,.$$

Here the formulation on the right-hand side should only be understood in a formal sense. This can be rigorously defined via the absolutely convergent series

$$\alpha^{\varepsilon\mathsf{GP}}(\kappa;W) - \operatorname{tr}[R_0^{\varepsilon\mathsf{GP}}\mathcal{W}] = \sum_{l=1}^{\infty}\frac{(-1)^l}{l}\operatorname{tr}[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l]$$

$$\alpha^{\mathsf{KdV}}(\kappa;\pm W_\pm) - \operatorname{tr}[R_0^{\mathsf{KdV}}(\pm W_\pm)] = \sum_{l=1}^{\infty}\frac{(-1)^l}{l}\operatorname{tr}[(R_0^{\mathsf{KdV}}(\pm W_\pm))^l]\,,$$

which is the content of Lemma 4.A.4 below. Subsequently, in Lemma 4.A.5, we prove the correspondence between these series and the transmission coefficient that we have stated above. We shall use the smallness of certain Hilbert-Schmidt norms to conclude the absolute convergence of these series as geometric series.

Lemma 4.A.3. *We explicitly compute the following Hilbert-Schmidt norms:*

$$\left\|\sqrt{R_0^{\mathsf{KdV}}}W_\pm\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}}^2 = \frac{1}{\kappa}\|W_\pm\|_{H_\kappa^{-1}}^2 \tag{4.A.4}$$

$$\left\|\sqrt{R_0^{\mathsf{KdV}}}i\partial_X W_\pm\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}}^2 = \frac{1}{2\kappa}\|W_\pm\|_{L_\kappa^2}^2 \tag{4.A.5}$$

$$\left\|\sqrt{R_0^{\mathsf{KdV}}}\mathcal{U}\mathcal{W}\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}}^2 = \left|\frac{\lambda+1}{2}\right|^2\frac{1}{\kappa}\|W_-\|_{H_\kappa^{-1}}^2 + \left|\frac{\lambda-1}{2}\right|^2\frac{1}{\kappa}\|W_+\|_{H_\kappa^{-1}}^2 \tag{4.A.6}$$

$$+\frac{\varepsilon^2}{2}\frac{1}{2\kappa}\|W_-\|_{L_\kappa^2}^2 + \frac{\varepsilon^2}{2}\frac{1}{2\kappa}\|W_+\|_{L_\kappa^2}^2\,. \tag{4.A.7}$$

Proof. For (4.A.4), we refer to [115, Proposition 2.1]. We shall reduce (4.A.5) and (4.A.6) to applications of (4.A.4). Towards this, we decompose

$$\begin{aligned}
\left\|\sqrt{R_0^{\mathsf{KdV}}}\mathcal{U}\mathcal{W}\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}}^2 &= \operatorname{tr}\left[\sqrt{R_0^{\mathsf{KdV}}}\mathcal{W}\mathcal{U}^*\sqrt{R_0^{\mathsf{KdV}}}\sqrt{R_0^{\mathsf{KdV}}}\mathcal{U}\mathcal{W}\sqrt{R_0^{\mathsf{KdV}}}\right]\\
&= \operatorname{tr}\left[R_0^{\mathsf{KdV}}\begin{pmatrix} W_-\frac{\overline{\lambda}+1}{2} & W_-\frac{i\varepsilon}{\sqrt{2}}\partial_X\\ W_+\frac{i\varepsilon}{\sqrt{2}}\partial_X & W_+\frac{\overline{\lambda}-1}{2}\end{pmatrix}R_0^{\mathsf{KdV}}\begin{pmatrix}\frac{\lambda+1}{2}W_- & \frac{i\varepsilon}{\sqrt{2}}\partial_X W_+\\ \frac{i\varepsilon}{\sqrt{2}}\partial_X W_- & \frac{\lambda-1}{2}W_+\end{pmatrix}\right]\\
&= \operatorname{tr}\left[R_0^{\mathsf{KdV}}W_-\frac{\overline{\lambda}+1}{2}R_0^{\mathsf{KdV}}\frac{\lambda+1}{2}W_-\right] + \operatorname{tr}\left[R_0^{\mathsf{KdV}}W_-\frac{i\varepsilon}{\sqrt{2}}\partial_X R_0^{\mathsf{KdV}}\frac{i\varepsilon}{\sqrt{2}}\partial_X W_-\right]\\
&\quad+\operatorname{tr}\left[R_0^{\mathsf{KdV}}W_+\frac{i\varepsilon}{\sqrt{2}}\partial_X R_0^{\mathsf{KdV}}\frac{i\varepsilon}{\sqrt{2}}\partial_X W_+\right] + \operatorname{tr}\left[R_0^{\mathsf{KdV}}W_+\frac{\overline{\lambda}-1}{2}R_0^{\mathsf{KdV}}\frac{\lambda-1}{2}W_+\right]\\
&= \left|\frac{\lambda+1}{2}\right|^2\left\|\sqrt{R_0^{\mathsf{KdV}}}W_-\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}}^2 + \left|\frac{\lambda-1}{2}\right|^2\left\|\sqrt{R_0^{\mathsf{KdV}}}W_+\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}}^2\\
&\quad+\frac{\varepsilon^2}{2}\left\|\sqrt{R_0^{\mathsf{KdV}}}W_-i\partial_X\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}}^2 + \frac{\varepsilon^2}{2}\left\|\sqrt{R_0^{\mathsf{KdV}}}W_+i\partial_X\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}}^2\,.
\end{aligned}$$

It remains to compute

$$\left\|\sqrt{R_0^{\mathsf{KdV}}}i\partial_X W_\pm\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}}^2 = \left\|\sqrt{R_0^{\mathsf{KdV}}}W_\pm i\partial_X\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}}^2 = \mathrm{tr}[R_0^{\mathsf{KdV}}W_\pm i\partial_X R_0^{\mathsf{KdV}}i\partial_X W_\pm]\,,$$

which can be written, using the kernels

$$\begin{aligned}
R_0^{\mathsf{KdV}}W_\pm i\partial_X f(X) &= \int_{\mathbb{R}}\frac{-i}{2\kappa}e^{-\kappa|X-Y|}(\kappa\,\mathrm{sign}(X-Y)W_\pm(Y)+\partial_Y W_\pm(Y))f(Y)\,\mathrm{d}Y\\
R_0^{\mathsf{KdV}}i\partial_X W_\pm f(X) &= \int_{\mathbb{R}}\frac{-i}{2\kappa}e^{-\kappa|X-Y|}\kappa\,\mathrm{sign}(X-Y)W_\pm(Y)f(Y)\,\mathrm{d}Y\,,
\end{aligned}$$

as

$$\begin{aligned}
&\mathrm{tr}[R_0^{\mathsf{KdV}}W_\pm i\partial_X R_0^{\mathsf{KdV}}i\partial_X W_\pm]\\
&=\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{-1}{4\kappa^2}e^{-\kappa|X-Y|}(\kappa\,\mathrm{sign}(X-Y)W_\pm(Y)+\partial_Y W_\pm(Y))e^{-\kappa|Y-X|}\kappa\,\mathrm{sign}(Y-X)W_\pm(X)\,\mathrm{d}X\,\mathrm{d}Y\\
&=\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{1}{4\kappa^2}e^{-2\kappa|X-Y|}\kappa^2W_\pm(X)W_\pm(Y)\,\mathrm{d}X\,\mathrm{d}Y\\
&+\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{1}{4\kappa^2}e^{-2\kappa|X-Y|}\kappa\,\mathrm{sign}(X-Y)W_\pm(X)\partial_Y W_\pm(Y)\,\mathrm{d}X\,\mathrm{d}Y\\
&=\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{1}{4\kappa^2}e^{-2\kappa|X-Y|}\kappa^2W_\pm(X)W_\pm(Y)\,\mathrm{d}X\,\mathrm{d}Y\\
&-\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{1}{4\kappa^2}e^{-2\kappa|X-Y|}(2\kappa^2-2\kappa\delta_0(X-Y))W_\pm(X)W_\pm(Y)\,\mathrm{d}X\,\mathrm{d}Y\\
&=-\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{1}{4\kappa^2}e^{-2\kappa|X-Y|}\kappa^2W_\pm(X)W_\pm(Y)\,\mathrm{d}X\,\mathrm{d}Y+2\kappa\int_{\mathbb{R}}\frac{1}{4\kappa^2}|W_\pm(X)|^2\,\mathrm{d}X\,.
\end{aligned}$$

We already know from (4.A.4) that

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{1}{4\kappa^2}e^{-2\kappa|X-Y|}\kappa^2W_\pm(X)W_\pm(Y)\,\mathrm{d}X\,\mathrm{d}Y=\kappa\|W_\pm\|_{H_\kappa^{-1}}^2\,,$$

which finishes the proof. □

As a consequence of this lemma we have $\sqrt{R_0^{\mathsf{KdV}}}\mathcal{W}\sqrt{R_0^{\mathsf{KdV}}}$, $\sqrt{R_0^{\mathsf{KdV}}}\mathcal{U}\mathcal{W}\sqrt{R_0^{\mathsf{KdV}}}\in L(L^2(\mathbb{R};\mathbb{C}^2))$. In the following, we sometimes use the operators $R_0^{\mathsf{KdV}}\mathcal{W}$ and $R_0^{\varepsilon\mathsf{GP}}\mathcal{W}$ instead for the sake of clarity in formal calculations, but always in a context where eventually a trace or Hilbert-Schmidt norm is taken, returning us to the well-behaved versions of these operators.

Lemma 4.A.4. *Let $\varepsilon\in(0,1)$. For all $0<\kappa<\frac{1}{\sqrt{2}\varepsilon}$ such that*

$$\frac{1}{\kappa}\|W\|_{H_\kappa^{-1}}^2\le\frac{1}{30}\qquad\qquad\frac{1}{\kappa}\|W\|_{L_\kappa^2}^2\le\frac{1}{15}$$

the series

$$\begin{aligned}
\alpha^{\mathsf{KdV}}(\kappa;\pm W_\pm)&=\sum_{l=2}^{\infty}\frac{(-1)^l}{l}\,\mathrm{tr}[(R_0^{\mathsf{KdV}}(\pm W_\pm))^l]\\
\alpha^{\varepsilon\mathsf{GP}}(\kappa;W)&=\sum_{l=2}^{\infty}\frac{(-1)^l}{l}\,\mathrm{tr}[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l]
\end{aligned}$$

are absolutely convergent and satisfy

$$\begin{aligned}
&\frac{3}{2}|\alpha^{\mathsf{KdV}}(\kappa;\pm W_\pm)|\le\frac{1}{\kappa}\|W_\pm\|_{H_\kappa^{-1}}^2\\
&\quad\frac{3}{2}|\alpha^{\varepsilon\mathsf{GP}}(\kappa;W)|\le\left|\frac{\lambda+1}{2}\right|^2\frac{1}{\kappa}\|W_-\|_{H_\kappa^{-1}}^2+\left|\frac{\lambda-1}{2}\right|^2\frac{1}{\kappa}\|W_+\|_{H_\kappa^{-1}}^2+\frac{\varepsilon^2}{2}\frac{1}{2\kappa}\|W_-\|_{L_\kappa^2}^2+\frac{\varepsilon^2}{2}\frac{1}{2\kappa}\|W_+\|_{L_\kappa^2}^2\,.
\end{aligned}$$

Moreover, the resolvent

$$R^{\varepsilon\mathsf{GP}} = \left(\frac{L^{\varepsilon\mathsf{GP}} - \lambda}{\varepsilon^2}\right)^{-1} \in L(L^2(\mathbb{R};\mathbb{C}^2))$$

is represented by the series

$$R^{\varepsilon\mathsf{GP}} = \sum_{l=0}^{\infty}(-1)^l (R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l R_0^{\varepsilon\mathsf{GP}} = \sum_{l=0}^{\infty}(-1)^l \sqrt{R_0^{\mathsf{KdV}}}\left(\sqrt{R_0^{\mathsf{KdV}}}\mathcal{U}\mathcal{W}\sqrt{R_0^{\mathsf{KdV}}}\right)^l \sqrt{R_0^{\mathsf{KdV}}}\mathcal{U}\,, \quad (4.A.8)$$

which is absolutely convergent with respect to the operator norm $\|\cdot\|_{L^2\to L^2}$*.*

Proof. This follows from Lemma 4.A.1 together with the Hilbert-Schmidt norms computed in Lemma 4.A.3. □

We have adapted the method of proof of the following lemma from [116, Theorem 2.6]. We refer to the paragraph above this theorem for some context on the history of this argument, and other approaches to proving that the perturbation determinant equals the logarithm of the transmission coefficient.

Lemma 4.A.5. *Assume the setting of Lemma 4.A.4, and in addition that* $W \in \mathcal{S}(\mathbb{R};\mathbb{R}^2)$ *so that we are in a setting where we have defined the transmission coefficients* $a^{\varepsilon\mathsf{GP}}(k;W)$ *and* $a^{\mathsf{KdV}}(k;\pm W_\pm)$ *(see (4.2.14)). We have*

$$\alpha^{\varepsilon\mathsf{GP}}(\kappa;W) - \log a^{\varepsilon\mathsf{GP}}(k;W) = \mathrm{tr}[R_0^{\varepsilon\mathsf{GP}}\mathcal{W}] = \frac{1}{2\kappa}\int_{\mathbb{R}}\frac{1}{2}(W_+ - W_-) - \frac{\lambda}{2}(W_+ + W_-)\,\mathrm{d}X \quad (4.A.9)$$

and similarly

$$\alpha^{\mathsf{KdV}}(\kappa;\pm W_\pm) - \log a^{\mathsf{KdV}}(k;\pm W_\pm) = \mathrm{tr}[R_0^{\mathsf{KdV}}(\pm W_\pm)] = \frac{1}{2\kappa}\int_{\mathbb{R}} \pm W_\pm\,\mathrm{d}X\,. \quad (4.A.10)$$

Proof. We compute

$$\begin{aligned}
\mathrm{tr}[R_0^{\varepsilon\mathsf{GP}}\mathcal{W}] &= \mathrm{tr}\left[R_0^{\mathsf{KdV}}\begin{pmatrix}\frac{\lambda+1}{2} & \frac{i\varepsilon}{\sqrt{2}}\partial_X \\ \frac{i\varepsilon}{\sqrt{2}}\partial_X & \frac{\lambda-1}{2}\end{pmatrix}\mathcal{W}\right] \\
&= \mathrm{tr}\left[R_0^{\mathsf{KdV}}\frac{\lambda+1}{2}(-W_-)\right] + \mathrm{tr}\left[R_0^{\mathsf{KdV}}\frac{\lambda-1}{2}(-W_+)\right] \\
&= \frac{\lambda+1}{2}\int_{\mathbb{R}}\frac{1}{2\kappa}e^{-\kappa|X-X|}(-W_-(X))\,\mathrm{d}X + \frac{\lambda-1}{2}\int_{\mathbb{R}}\frac{1}{2\kappa}e^{-\kappa|X-X|}(-W_+(X))\,\mathrm{d}X \\
&= \frac{1}{2\kappa}\int_{\mathbb{R}}\frac{1}{2}(W_+ - W_-) - \frac{\lambda}{2}(W_+ + W_-)\,\mathrm{d}X
\end{aligned}$$

and

$$\mathrm{tr}[R_0^{\mathsf{KdV}}(\pm W_\pm)] = \frac{1}{2\kappa}\int_{\mathbb{R}} \pm W_\pm\,\mathrm{d}X\,.$$

For (4.A.10) we refer to [114, Proposition 2.4], so it remains to prove (4.A.9).

By explicitly computing the Jost solution $\Psi_{1,1}^{\varepsilon\mathsf{GP},-}(X;0)$, which is constant, we find that $\log a^{\varepsilon\mathsf{GP}}(k;0) = 0$, which matches $\alpha^{\varepsilon\mathsf{GP}}(\kappa;0) = 0$. It therefore suffices to prove for any $W, V \in \mathcal{S}(\mathbb{R};\mathbb{R}^2)$ with

$$V = (V_+, V_-) \qquad\qquad \mathcal{V} = \begin{pmatrix}-V_- & 0 \\ 0 & -V_+\end{pmatrix}$$

that the following limits exist and are identical:

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0}\log a^{\varepsilon\mathsf{GP}}(k;W+sV) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0}(\alpha^{\varepsilon\mathsf{GP}}(\kappa;W+sV) - \mathrm{tr}[R_0^{\varepsilon\mathsf{GP}}(\mathcal{W}+s\mathcal{V})])\,.$$

Set $\Phi^{\varepsilon\mathsf{GP}} = (\Phi_1^{\varepsilon\mathsf{GP},-}|\Phi_2^{\varepsilon\mathsf{GP},+})$. We start with the formal calculation

$$\begin{aligned}0 &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\left(\frac{L^{\varepsilon\mathsf{GP}}(W+sV)-\lambda}{\varepsilon^2}\Phi_j^{\varepsilon\mathsf{GP},\pm}(k;W+sV)\right)\\ &= \mathcal{V}\Phi_j^{\varepsilon\mathsf{GP},\pm}(k;W) + \frac{L^{\varepsilon\mathsf{GP}}(W)-\lambda}{\varepsilon^2}\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\Phi_j^{\varepsilon\mathsf{GP},\pm}(k;W+sV)\,,\end{aligned}$$

which implies

$$\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\Phi^{\varepsilon\mathsf{GP}}(X,k;W+sV) = \int_{\mathbb{R}} G^{\varepsilon\mathsf{GP}}(X,Y)(-\mathcal{V}(Y))\Phi^{\varepsilon\mathsf{GP}}(Y,k;W)\,\mathrm{d}Y\,.$$

For the derivative of the transmission coefficient this yields

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\log a^{\varepsilon\mathsf{GP}}(k;W+sV)\\ &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\log\left(\frac{\det\Phi^{\varepsilon\mathsf{GP}}(X,k;W+sV)}{\det E^{\varepsilon\mathsf{GP},+}(k)}\right)\\ &= \mathrm{tr}\left[\Phi^{\varepsilon\mathsf{GP}}(X,k;W)^{-1}\int_{\mathbb{R}} G^{\varepsilon\mathsf{GP}}(X,Y)(-\mathcal{V}(Y))\Phi^{\varepsilon\mathsf{GP}}(Y,k;W)\,\mathrm{d}Y\right]\\ &= \int_{\mathbb{R}}\mathrm{tr}\big[\Phi^{\varepsilon\mathsf{GP}}(Y,k;W)\Phi^{\varepsilon\mathsf{GP}}(X,k;W)^{-1}G^{\varepsilon\mathsf{GP}}(X,Y)(-\mathcal{V}(Y))\big]\,\mathrm{d}Y\\ &= \int_{\mathbb{R}}\mathrm{tr}\left[\left(\mathbb{1}_{\{X<Y\}}\lim_{X'\nearrow Y}G^{\varepsilon\mathsf{GP}}(X',Y)+\mathbb{1}_{\{X>Y\}}\lim_{X'\searrow Y}G^{\varepsilon\mathsf{GP}}(X',Y)\right)(-\mathcal{V}(Y))\right]\mathrm{d}Y\\ &= \int_{\mathbb{R}} G_{1,1}^{\varepsilon\mathsf{GP}}(X,X)V_-(X)+G_{2,2}^{\varepsilon\mathsf{GP}}(X,X)V_+(X)\,\mathrm{d}X\,.\end{aligned}$$

Here we have used that $G_{1,1}^{\varepsilon\mathsf{GP}}$ and $G_{2,2}^{\varepsilon\mathsf{GP}}$ are continuous. Since the Green's function is bounded and V is Schwartz, the integrals are absolutely convergent and the derivative actually exists. On the other hand, we know from Lemma 4.A.1 that $\alpha^{\varepsilon\mathsf{GP}}(\kappa;W+sV)$ is continuously differentiable. Specifically (4.A.3) yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\alpha^{\varepsilon\mathsf{GP}}(\kappa;W+sV) &= \sum_{l=2}^{\infty}(-1)^l\,\mathrm{tr}[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^{l-1}R_0^{\varepsilon\mathsf{GP}}\mathcal{V}]\\ &= \mathrm{tr}[(R_0^{\varepsilon\mathsf{GP}}-R^{\varepsilon\mathsf{GP}})\mathcal{V}]\\ &= \int_{\mathbb{R}}\mathrm{tr}[(G^{\varepsilon\mathsf{GP}}-G_0^{\varepsilon\mathsf{GP}})(X,X)(-\mathcal{V})(X)]\,\mathrm{d}X\\ &= \int_{\mathbb{R}}(G^{\varepsilon\mathsf{GP}}-G_0^{\varepsilon\mathsf{GP}})_{1,1}(X,X)V_-(X)+(G^{\varepsilon\mathsf{GP}}-G_0^{\varepsilon\mathsf{GP}})_{2,2}(X,X)V_+(X)\,\mathrm{d}X\,.\end{aligned}$$

Here we have used that $G^{\varepsilon\mathsf{GP}}$ and $G_0^{\varepsilon\mathsf{GP}}$ have continuous diagonal parts. The proof is finished with the additional observation that

$$-\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\mathrm{tr}[R_0^{\varepsilon\mathsf{GP}}(\mathcal{W}+s\mathcal{V})] = \int_{\mathbb{R}}(G_0^{\varepsilon\mathsf{GP}})_{1,1}(X,X)V_-(X)+(G_0^{\varepsilon\mathsf{GP}})_{2,2}(X,X)V_+(X)\,\mathrm{d}X\,.$$

□

Recall from (4.1.14) the definition of the operators Ev_λ and Od_λ. In the following, we suppress the dependence of various quantities on λ from the notation, but for the application of these operators this dependence is of course relevant.

Lemma 4.A.6. *Let $W=(W_+,W_-)\in\mathcal{S}(\mathbb{R};\mathbb{R}^2)$, $\varepsilon\in(0,1)$, and $0<\kappa<\frac{1}{\sqrt{2}\varepsilon}$ with $\lambda\in(0,1)$. Assume that*

$$\max\left\{\frac{1}{\kappa},\frac{1}{\kappa^3}\right\}\|W\|_{L^2}^2\le\frac{1}{30}\,.$$

Then the series $\alpha^{\varepsilon\mathsf{GP}}(\kappa;W)$ and $\alpha^{\mathsf{KdV}}(\kappa;\pm W_\pm)$ converge. Furthermore, we have

$$\mathrm{Ev}_\lambda\big[\mathrm{tr}[R_0^{\varepsilon\mathsf{GP}}\mathcal{W}]\big] \mp \frac{1}{\lambda}\,\mathrm{Od}_\lambda\big[\mathrm{tr}[R_0^{\varepsilon\mathsf{GP}}\mathcal{W}]\big] = \mathrm{tr}[R_0^{\mathsf{KdV}}(\pm W_\pm)] \tag{4.A.11}$$

$$\mathrm{Ev}_\lambda\big[\alpha^{\varepsilon\mathsf{GP}}(\kappa;W)\big] \mp \frac{1}{\lambda}\,\mathrm{Od}_\lambda\big[\alpha^{\varepsilon\mathsf{GP}}(\kappa;W)\big] = \alpha^{\mathsf{KdV}}(\kappa;(\pm W_\pm)) + \mathbf{E}(\kappa;W) \tag{4.A.12}$$

with

$$|\mathbf{E}(\kappa;W)| \le 8\frac{\varepsilon^2}{\kappa}\|W\|_{L^2}^2\,.$$

Proof. Combining (4.A.9) and (4.A.10) directly yields (4.A.11). To prove (4.A.12) we need to understand

$$\mathrm{Ev}_\lambda\big[\mathrm{tr}[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l]\big] \mp \frac{1}{\lambda}\,\mathrm{Od}_\lambda\big[\mathrm{tr}[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l]\big]\,.$$

Using the Pauli matrices

$$\sigma_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \qquad\qquad \sigma_1 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},$$

we write

$$\begin{aligned}\mathrm{Ev}_\lambda[R_0^{\varepsilon\mathsf{GP}}\mathcal{W}] &= \frac{1}{2}R_0^{\mathsf{KdV}}(\sigma_3 + i\sqrt{2}\varepsilon\sigma_1\partial_X)\mathcal{W}\\ \mathrm{Od}_\lambda[R_0^{\varepsilon\mathsf{GP}}\mathcal{W}] &= \frac{1}{2}R_0^{\mathsf{KdV}}\lambda\mathcal{W}\,.\end{aligned}$$

Then

$$\mathrm{Ev}_\lambda\big[\mathrm{tr}[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l]\big] \mp \frac{1}{\lambda}\,\mathrm{Od}_\lambda\big[\mathrm{tr}[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l]\big] = \mathrm{tr}\Big[\mathrm{Ev}_\lambda[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l] \mp \frac{1}{\lambda}\,\mathrm{Od}_\lambda[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l]\Big]\,.$$

For $l \in \mathbb{N}$ and $\omega \in \{0,1,2\}^l$ we define

$$\overline{\omega} = \sum_{j=1}^{l} \mathbb{1}_{\{\omega_j=1\}}\,.$$

One can show directly by induction that

$$\begin{aligned}\mathrm{Ev}_\lambda[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l] &= \sum_{\substack{\omega\in\{0,1\}^l\\ \overline{\omega}\text{ even}}} 2^{-l}\prod_{j=1}^{l}\begin{cases}R_0^{\mathsf{KdV}}(\sigma_3 + \sqrt{2}\varepsilon\sigma_1 i\partial_X)\mathcal{W} & ,\omega_j = 0\\ R_0^{\mathsf{KdV}}\lambda\mathcal{W} & ,\omega_j = 1\end{cases}\\ &= \sum_{\substack{\omega\in\{0,1,2\}^l\\ \overline{\omega}\text{ even}}} 2^{-l}\prod_{j=1}^{l}\begin{cases}R_0^{\mathsf{KdV}}\sigma_3\mathcal{W} & ,\omega_j = 0\\ R_0^{\mathsf{KdV}}\lambda\mathcal{W} & ,\omega_j = 1\\ R_0^{\mathsf{KdV}}\sqrt{2}\varepsilon\sigma_1 i\partial_X\mathcal{W} & ,\omega_j = 2\end{cases}\end{aligned}$$

and

$$\begin{aligned}\mathrm{Od}_\lambda[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l] &= \sum_{\substack{\omega\in\{0,1\}^l\\ \overline{\omega}\text{ odd}}} 2^{-l}\prod_{j=1}^{l}\begin{cases}R_0^{\mathsf{KdV}}(\sigma_3 + \sqrt{2}\varepsilon\sigma_1 i\partial_X)\mathcal{W} & ,\omega_j = 0\\ R_0^{\mathsf{KdV}}\lambda\mathcal{W} & ,\omega_j = 1\end{cases}\\ &= \sum_{\substack{\omega\in\{0,1,2\}^l\\ \overline{\omega}\text{ odd}}} 2^{-l}\prod_{j=1}^{l}\begin{cases}R_0^{\mathsf{KdV}}\sigma_3\mathcal{W} & ,\omega_j = 0\\ R_0^{\mathsf{KdV}}\lambda\mathcal{W} & ,\omega_j = 1\,.\\ R_0^{\mathsf{KdV}}\sqrt{2}\varepsilon\sigma_1 i\partial_X\mathcal{W} & ,\omega_j = 2\end{cases}\end{aligned}$$

Here the ordering of the product of possibly non-commuting operators does not matter, as long as it is consistent for all $\omega \in \{0,1\}^l$. We shall fix as a convention that the first term in the product,

i. e. the term with $j=1$, is multiplied on the left, while the last term is multiplied on the right. Then

$$
\begin{aligned}
\mathrm{Ev}_\lambda[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l] \mp \frac{1}{\lambda}\mathrm{Od}_\lambda[(R_0^{\varepsilon\mathsf{GP}}\mathcal{W})^l] &= \sum_{\omega\in\{0,1,2\}^l} 2^{-l}\left(\frac{\mp 1}{\lambda}\right)^{\mathbb{1}_{\{\varpi\text{ odd}\}}} \prod_{j=1}^{l}\begin{cases} R_0^{\mathsf{KdV}}\sigma_3\mathcal{W} & ,\omega_j=0\\ R_0^{\mathsf{KdV}}\lambda\mathcal{W} & ,\omega_j=1\\ R_0^{\mathsf{KdV}}\sqrt{2}\varepsilon\sigma_1 i\partial_X\mathcal{W} & ,\omega_j=2\end{cases}\\
&= \sum_{\substack{\omega\in\{0,1,2\}^l\\ \exists j:\omega_j=2}} 2^{-l}\left(\frac{\mp 1}{\lambda}\right)^{\mathbb{1}_{\{\varpi\text{ odd}\}}} \prod_{j=1}^{l}\begin{cases} R_0^{\mathsf{KdV}}\sigma_3\mathcal{W} & ,\omega_j=0\\ R_0^{\mathsf{KdV}}\lambda\mathcal{W} & ,\omega_j=1\\ R_0^{\mathsf{KdV}}\sqrt{2}\varepsilon\sigma_1 i\partial_X\mathcal{W} & ,\omega_j=2\end{cases}\\
&\quad+ \sum_{\omega\in\{0,1\}^l} 2^{-l}\left(\frac{\mp 1}{\lambda}\right)^{\mathbb{1}_{\{\varpi\text{ odd}\}}} \prod_{j=1}^{l}\begin{cases} R_0^{\mathsf{KdV}}\sigma_3\mathcal{W} & ,\omega_j=0\\ R_0^{\mathsf{KdV}}\lambda\mathcal{W} & ,\omega_j=1\end{cases}\\
&= I + II\,.
\end{aligned}
$$

Using that the trace of off-diagonal matrices vanishes, we estimate

$$
\begin{aligned}
\left|\sum_{l=2}^{\infty}\frac{(-1)^l}{l}\operatorname{tr}[I]\right| &\le \sum_{l=2}^{\infty}\sum_{\substack{\omega\in\{0,1,2\}^l\\ |\{j:\omega_j=2\}|\ge 2\text{ even}}} |\lambda|^{-\mathbb{1}_{\{\varpi\text{ odd}\}}}2^{-l}\prod_{j=1}^{l}\begin{cases} \left\|\sqrt{R_0^{\mathsf{KdV}}}\sigma_3\mathcal{W}\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}} & ,\omega_j=0\\ \left\|\sqrt{R_0^{\mathsf{KdV}}}\mathcal{W}\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}}|\lambda| & ,\omega_j=1\\ \left\|\sqrt{R_0^{\mathsf{KdV}}}\sqrt{2}\varepsilon\sigma_1 i\partial_X\mathcal{W}\sqrt{R_0^{\mathsf{KdV}}}\right\|_{\mathrm{HS}} & ,\omega_j=2\end{cases}\\
&\le \sum_{l=2}^{\infty}\sum_{\substack{\omega\in\{0,1,2\}^l\\ |\{j:\omega_j=2\}|\ge 2\text{ even}}} |\lambda|^{-\mathbb{1}_{\{\varpi\text{ odd}\}}}2^{-l}\kappa^{-\frac{l}{2}}\prod_{j=1}^{l}\begin{cases} \|W\|_{H_\kappa^{-1}} & ,\omega_j=0\\ \|W\|_{H_\kappa^{-1}}|\lambda| & ,\omega_j=1\\ \|W\|_{L_\kappa^2}\varepsilon & ,\omega_j=2\end{cases}
\end{aligned}
$$

Here we have used (4.A.2), (4.A.4), and (4.A.5). Noting that $\|W\|_{H_\kappa^{-1}} \le \frac{1}{\sqrt{2\kappa}}\|W\|_{L_\kappa^2}$ and $|\lambda|\le 1$, we finish our estimate:

$$
\begin{aligned}
\left|\sum_{l=2}^{\infty}\frac{(-1)^l}{l}\operatorname{tr}[I]\right| &\le \sum_{l=2}^{\infty}\sum_{\substack{\omega\in\{0,1,2\}^l\\ |\{j:\omega_j=2\}|\ge 2\text{ even}}} \varepsilon^{|\{\omega_j=2\}|}2^{-l}\kappa^{-\frac{l}{2}}|\lambda|^{|\{\omega_j=1\}|-\mathbb{1}_{\{\varpi\text{ odd}\}}}\|W\|_{H_\kappa^{-1}}^{|\{\omega_j=0\}|+|\{\omega_j=1\}|}\|W\|_{L_\kappa^2}^{|\{\omega_j=2\}|}\\
&\le \sum_{l=2}^{\infty}\|W\|_{L_\kappa^2}^l 2^{-l}\sum_{\substack{\omega\in\{0,1,2\}^l\\ |\{j:\omega_j=2\}|\ge 2\text{ even}}} \varepsilon^{|\{\omega_j=2\}|}(\sqrt{2}\kappa)^{-|\{\omega_j=0\}|-|\{\omega_j=1\}|-\frac{l}{2}}\\
&\le 2\varepsilon^2\kappa^2\sum_{l=2}^{\infty}\kappa^{-\frac{l}{2}}\|W\|_{L_\kappa^2}^l 2^{-l}\sum_{\substack{n=2\\ n\text{ even}}}^{l}\sum_{j=0}^{l}\binom{l}{n,j}(\sqrt{2}\varepsilon\kappa)^{n-2}(\sqrt{2}\kappa)^{-\frac{3}{2}l}\\
&\le 2\varepsilon^2\kappa^2\sum_{l=2}^{\infty}\kappa^{-\frac{3}{2}l}\|W\|_{L_\kappa^2}^l 2^{-l}3^l\sqrt{2}^{-\frac{3}{2}l} \le \frac{\varepsilon^2\kappa^{-1}\|W\|_{L^2}^2}{1-\kappa^{-\frac{3}{2}}\|W\|_{L^2}}\,.
\end{aligned}
$$

In order to compute II, we decompose

$$
\begin{aligned}
\{0,1\}^l = \{\omega\in\{0,1\}^l:\omega_1=0\text{ and }\varpi\text{ even}\}&\cup\{\omega\in\{0,1\}^l:\omega_1=1\text{ and }\varpi\text{ odd}\}\\
\cup\{\omega\in\{0,1\}^l:\omega_1=0\text{ and }\varpi\text{ odd}\}&\cup\{\omega\in\{0,1\}^l:\omega_1=1\text{ and }\varpi\text{ even}\}
\end{aligned}
$$

and observe that

$$
\omega\longmapsto(0,\omega_2,\dots,\omega_l) \qquad\qquad \omega\longmapsto(1,\omega_2,\dots,\omega_l)
$$

are bijections between the first two and last two sets. We may sum only over the sets on the left and correspondingly adjust each term as follows:

$$
II = \sum_{\substack{\omega\in\{0,1\}^l\\ \omega_1=0\\ \varpi\text{ even}}} 2^{-l}R_0^{\mathsf{KdV}}(\sigma_3\mp 1)\mathcal{W}\prod_{j=2}^{l}\begin{cases} R_0^{\mathsf{KdV}}\sigma_3\mathcal{W} & ,\omega_j=0\\ R_0^{\mathsf{KdV}}\lambda\mathcal{W} & ,\omega_j=1\end{cases}
$$

$$
\begin{aligned}
&+ \sum_{\substack{\omega\in\{0,1\}^l \\ \omega_1=0 \\ \overline{\omega}\text{ odd}}} 2^{-l} R_0^{\mathsf{KdV}} \left(\frac{\mp\sigma_3}{\lambda} + \lambda\right) \mathcal{W} \prod_{j=2}^{l} \begin{cases} R_0^{\mathsf{KdV}} \sigma_3 \mathcal{W} & ,\omega_j = 0 \\ R_0^{\mathsf{KdV}} \lambda \mathcal{W} & ,\omega_j = 1 \end{cases} \\
&= \sum_{\substack{\omega\in\{0,1\}^l \\ \omega_1=0 \\ \overline{\omega}\text{ even}}} 2^{-l}(\sigma_3 \mp 1)\lambda^{\overline{\omega}}\sigma_3^{l-1-\overline{\omega}}(R_0^{\mathsf{KdV}}\mathcal{W})^l + \sum_{\substack{\omega\in\{0,1\}^l \\ \omega_1=0 \\ \overline{\omega}\text{ odd}}} 2^{-l}(\sigma_3 \mp 1 \pm 2\varepsilon^2\kappa^2)\left(\frac{\mp 1}{\lambda}\right)\lambda^{\overline{\omega}}\sigma_3^{l-1-\overline{\omega}}(R_0^{\mathsf{KdV}}\mathcal{W})^l \\
&= (\sigma_3 \mp 1)(R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l\sigma_3 \sum_{\substack{\omega\in\{0,1\}^l \\ \omega_1=0}} 2^{-l}\left(\frac{\mp 1}{\lambda}\right)^{\mathbb{1}_{\{\overline{\omega}\text{ odd}\}}} \lambda^{\overline{\omega}}\sigma_3^{\overline{\omega}} + 2\varepsilon^2\kappa^2(R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l\sigma_3 \sum_{\substack{\omega\in\{0,1\}^l \\ \omega_1=0 \\ \overline{\omega}\text{ odd}}} 2^{-l}\left(\frac{-1}{\lambda}\right)\lambda^{\overline{\omega}}\sigma_3^{\overline{\omega}} \\
&= III + IV\,.
\end{aligned}
$$

The further treatment of III and IV is a matter of computation. We have

$$
\begin{aligned}
III &= (\sigma_3 \mp 1)(R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l 2^{-l}\sigma_3 \left(\sum_{\substack{\overline{\omega}=0 \\ \overline{\omega}\text{ even}}}^{l-1} \binom{l-1}{\overline{\omega}}\lambda^{\overline{\omega}} \mp \sigma_3 \sum_{\substack{\overline{\omega}=0 \\ \overline{\omega}\text{ odd}}}^{l-1} \binom{l-1}{\overline{\omega}}\lambda^{\overline{\omega}-1} \right) \\
&= \frac{\sigma_3 \mp 1}{2}(R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l 2^{1-l} \left(\sigma_3 \sum_{j=0}^{\lfloor\frac{l-1}{2}\rfloor} \binom{l-1}{2j}(1-2\varepsilon^2\kappa^2)^j \mp \sum_{j=0}^{\lfloor\frac{l}{2}\rfloor-1} \binom{l-1}{2j+1}(1-2\varepsilon^2\kappa^2)^j \right) \\
&= \frac{(\sigma_3 \mp 1)^2}{4}(R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l + \frac{\sigma_3 \mp 1}{2}(R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l 2^{1-l} \sum_{j=0}^{\infty}\left(\sigma_3\binom{l-1}{2j} \mp \binom{l-1}{2j+1}\right)((1-2\varepsilon^2\kappa^2)^j - 1) \\
&= V + VI\,.
\end{aligned}
$$

The most important term is V, because

$$
\sum_{l=2}^{\infty} \frac{(-1)^l}{l}\operatorname{tr}[V] = \sum_{l=2}^{\infty} \frac{(-1)^l}{l}\left(\frac{(1\mp 1)^2}{4}\operatorname{tr}[(R_0^{\mathsf{KdV}}(-W_-))^l] + \frac{(-1\mp 1)^2}{4}\operatorname{tr}[(R_0^{\mathsf{KdV}}W_+)^l] \right) = \alpha^{\mathsf{KdV}}(\kappa; \pm W_\pm)\,.
$$

We calculate further

$$
\begin{aligned}
IV &= -2\varepsilon^2\kappa^2(R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l 2^{-l} \sum_{\substack{\overline{\omega}=0 \\ \overline{\omega}\text{ odd}}}^{l-1} \binom{l-1}{\overline{\omega}}\lambda^{\overline{\omega}-1} \\
&= -2\varepsilon^2\kappa^2(R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l 2^{-l} \sum_{j=0}^{\lfloor\frac{l}{2}\rfloor-1} \binom{l-1}{2j+1}(1-2\varepsilon^2\kappa^2)^j \\
&= (R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l 2^{-l} \sum_{m=1}^{\infty}(-2\varepsilon^2\kappa^2)^m \sum_{j=0}^{\infty}\binom{l-1}{2j+1}\binom{j}{m-1}
\end{aligned}
$$

and

$$
\begin{aligned}
VI &= (R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l 2^{-l}(\sigma_3 \mp 1)\sum_{j=0}^{\infty}\left(\sigma_3\binom{l-1}{2j} \mp \binom{l-1}{2j+1}\right)\sum_{m=1}^{\infty}\binom{j}{m}(-2\varepsilon^2\kappa^2)^m \\
&= (R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l 2^{-l}\sum_{m=1}^{\infty}(-2\varepsilon^2\kappa^2)^m \sum_{j=0}^{\infty}(1\mp\sigma_3)\binom{l}{2j+1}\binom{j}{m}\,.
\end{aligned}
$$

Then a binomial identity, which may be verified by a calculation with generating functions, yields

$$
IV + VI = (R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l 2^{-l}\sum_{m=1}^{\infty}(-2\varepsilon^2\kappa^2)^m \sum_{j=0}^{\infty}\left((1\mp\sigma_3)\binom{l}{2j+1}\binom{j}{m} + \binom{l-1}{2j+1}\binom{j}{m-1}\right)
$$

$$
\begin{aligned}
&= (R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l 2^{-l} \sum_{m=1}^{\infty} (-2\varepsilon^2\kappa^2)^m \left((1 \mp \sigma_3) 2^{l-1-2m} \binom{l-1-m}{m} + 2^{l-2m} \binom{l-1-m}{m-1} \right) \\
&= (R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l \sum_{m=1}^{\infty} (-2\varepsilon^2\kappa^2)^m 4^{-m} \left(\frac{1 \mp \sigma_3}{2} \binom{l-1-m}{m} + \binom{l-1-m}{m-1} \right).
\end{aligned}
$$

We find that

$$
\begin{aligned}
&\sum_{l=2}^{\infty} \frac{(-1)^l}{l} \operatorname{tr}[IV + VI] \\
&= \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \sum_{m=1}^{\infty} (-2\varepsilon^2\kappa^2)^m 4^{-m} \left(\operatorname{tr}\left[(R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l \frac{1 \mp \sigma_3}{2} \right] \binom{l-1-m}{m} + \operatorname{tr}[(R_0^{\mathsf{KdV}}\sigma_3\mathcal{W})^l] \binom{l-1-m}{m-1} \right) \\
&= \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \sum_{m=1}^{\infty} (-2\varepsilon^2\kappa^2)^m 4^{-m} \left(\operatorname{tr}[(R_0^{\mathsf{KdV}}(\pm W_\pm))^l] \binom{l-1-m}{m} \right. \\
&\qquad\qquad \left. + \left(\operatorname{tr}[(R_0^{\mathsf{KdV}}(-W_-))^l] + \operatorname{tr}[(R_0^{\mathsf{KdV}} W_+)^l] \right) \frac{1}{2} \binom{l-1-m}{m-1} \right),
\end{aligned}
$$

and finally

$$
\begin{aligned}
\left| \sum_{l=2}^{\infty} \frac{(-1)^l}{l} \operatorname{tr}[IV + VI] \right| &\leq \sum_{m=1}^{\infty} \varepsilon^{2m}\kappa^{2m} 2^{-m} \sum_{l=2m}^{\infty} 2^{l-m} \kappa^{-\frac{l}{2}} \left(\|W_-\|_{H_\kappa^{-1}}^l + \|W_+\|_{H_\kappa^{-1}}^l \right) \\
&\leq \sum_{m=1}^{\infty} \varepsilon^{2m}\kappa^m \|W\|_{H_\kappa^{-1}}^{2m} \frac{2}{1 - 2\kappa^{-\frac{1}{2}} \|W\|_{H_\kappa^{-1}}} \\
&\leq \frac{2\varepsilon^2\kappa\|W\|_{H_\kappa^{-1}}^2}{1 - 2\varepsilon^2\kappa\|W\|_{H_\kappa^{-1}}^2} \frac{2}{1 - 2\kappa^{-\frac{1}{2}}\|W\|_{H_\kappa^{-1}}} \\
&\leq \frac{\varepsilon^2\kappa^{-1}\|W\|_{L^2}^2}{1 - \varepsilon^2\kappa^{-1}\|W\|_{L^2}^2} \frac{1}{1 - \kappa^{-\frac{3}{2}}\|W\|_{L^2}} \\
&\leq \frac{\varepsilon^2\kappa^{-1}\|W\|_{L^2}^2}{(1 - \kappa^{-\frac{3}{2}}\|W\|_{L^2})^2}.
\end{aligned}
$$

Finally, we estimate

$$
\frac{1}{(1 - \kappa^{-\frac{3}{2}}\|W\|_{L^2})^2} + \frac{1}{1 - \kappa^{-\frac{3}{2}}\|W\|_{L^2}} \leq \frac{1}{(1 - \frac{1}{\sqrt{3}})^2} + \frac{1}{1 - \frac{1}{\sqrt{3}}} \leq 8\,.
$$

□

4.B. Appendix: Overview of Hamiltonian structures

In this section we present concisely and formally the Hamiltonian structures of GP, hGP, εGP, and KdV.

4.B.1. GP

We consider formally a manifold where we denote points by $\mathbf{q} = (q, \overline{q})$ and tangent vectors by $\mathbf{p} = (p, \overline{p})$ and $\mathbf{r} = (r, \overline{r})$. We define the Hamiltonian operator J^{GP}, the Riemannian metric g^{GP} and the symplectic form ω^{GP} by

$$
\begin{aligned}
J_{\mathbf{q}}^{\mathsf{GP}} &= -\frac{i}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
g_{\mathbf{q}}^{\mathsf{GP}}(\mathbf{p}, \mathbf{r}) &= \frac{1}{2} \int_{\mathbb{R}} (p\overline{r} + \overline{p}r) \, \mathrm{d}x \\
\omega_{\mathbf{q}}^{\mathsf{GP}}(\mathbf{p}, \mathbf{r}) &= -i \int_{\mathbb{R}} (p\overline{r} - \overline{p}r) \, \mathrm{d}x
\end{aligned}
$$

They satisfy

$$\omega^{\mathsf{GP}}_{\mathbf{q}}(\mathbf{p}, J^{\mathsf{GP}}_{\mathbf{q}}\mathbf{r}) = g^{\mathsf{GP}}_{\mathbf{q}}(\mathbf{p}, \mathbf{r})\,.$$

We define the gradient $\nabla^{g^{\mathsf{GP}}}_{\mathbf{q}}$ with respect to the metric g^{GP} as well as the functional derivative via

$$g^{\mathsf{GP}}_{\mathbf{q}}(\mathbf{p}, \nabla^{g^{\mathsf{GP}}}_{\mathbf{q}}\mathcal{H}) = \int_{\mathbb{R}} \mathbf{p}\left(\frac{\delta\mathcal{H}(\mathbf{q})}{\delta q}, \frac{\delta\mathcal{H}(\mathbf{q})}{\delta\overline{q}}\right)^T \mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\mathcal{H}(\mathbf{q}+s\mathbf{p})\,.$$

This implies

$$\nabla^{g^{\mathsf{GP}}}_{\mathbf{q}}\mathcal{H} = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}\left(\frac{\delta\mathcal{H}(\mathbf{q})}{\delta q}, \frac{\delta\mathcal{H}(\mathbf{q})}{\delta\overline{q}}\right)^T.$$

Then (GP_n) can be written as

$$i\partial_{t_n}\mathbf{q} = iJ^{\mathsf{GP}}_{\mathbf{q}}\nabla^{g^{\mathsf{GP}}}_{\mathbf{q}}\mathcal{H}^{\mathsf{GP}}_n = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\left(\frac{\delta\mathcal{H}^{\mathsf{GP}}_n(\mathbf{q})}{\delta q}, \frac{\delta\mathcal{H}^{\mathsf{GP}}_n(\mathbf{q})}{\delta\overline{q}}\right)^T.$$

The Poisson bracket is

$$\{\mathcal{F},\mathcal{G}\}^{\mathsf{GP}}(\mathbf{q}) = \int_{\mathbb{R}}\left(\frac{\delta\mathcal{F}}{\delta q}\frac{\delta\mathcal{G}}{\delta\overline{q}} - \frac{\delta\mathcal{F}}{\delta\overline{q}}\frac{\delta\mathcal{G}}{\delta q}\right)\mathrm{d}x\,.$$

4.B.2. hGP

We consider formally a manifold where we denote points by $w = (w_+, w_-)$ and tangent vectors by $u = (u_+, u_-)$ and $v = (v_+, v_-)$. We define the Hamiltonian operator J^{hGP}, the Riemannian metric g^{hGP} and the symplectic form ω^{hGP} by

$$\begin{aligned}
J^{\mathsf{hGP}}_w &= \frac{1}{8}\begin{pmatrix} -\partial_x a^{-1} + 4a\partial_x^{-1} & \partial_x a^{-1} + 4a\partial_x^{-1} \\ -\partial_x a^{-1} - 4a\partial_x^{-1} & \partial_x a^{-1} - 4a\partial_x^{-1} \end{pmatrix} \\
g^{\mathsf{hGP}}_w(u,v) &= \int_{\mathbb{R}} \frac{1}{4}(u_+ - u_-)(v_+ - v_-) + a^2\partial_x^{-1}(u_+ + u_-)\partial_x^{-1}(v_+ + v_-)\,\mathrm{d}x \\
\omega^{\mathsf{hGP}}_w(u,v) &= \int_{\mathbb{R}} a(v_+ - v_-)\partial_x^{-1}(u_+ + u_-) - a(u_+ - u_-)\partial_x^{-1}(v_+ + v_-)\,\mathrm{d}x\,.
\end{aligned}$$

They satisfy

$$\omega^{\mathsf{hGP}}_w(u, J^{\mathsf{hGP}}_w v) = g^{\mathsf{hGP}}_w(u,v)\,.$$

We define the gradient $\nabla^{g^{\mathsf{hGP}}}_w$ with respect to the metric g^{hGP} as well as the functional derivative via

$$g^{\mathsf{hGP}}_w(v, \nabla^{g^{\mathsf{hGP}}}_w\mathcal{H}) = \int_{\mathbb{R}}(v_+, v_-)\left(\frac{\delta\mathcal{H}(w)}{\delta w_+}, \frac{\delta\mathcal{H}(w)}{\delta w_-}\right)^T \mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\mathcal{H}(w+sv)\,.$$

This implies

$$\nabla^{g^{\mathsf{hGP}}}_w\mathcal{H} = \frac{1}{4}\begin{pmatrix} 4 - \partial_x a^{-2}\partial_x & -4 - \partial_x a^{-2}\partial_x \\ -4 - \partial_x a^{-2}\partial_x & 4 - \partial_x a^{-2}\partial_x \end{pmatrix}\left(\frac{\delta\mathcal{H}(w)}{\delta w_+}, \frac{\delta\mathcal{H}(w)}{\delta w_-}\right)^T.$$

Then (hGP_n) can be written as

$$\partial_{t_n}w = J^{\mathsf{hGP}}_w\nabla^{g^{\mathsf{hGP}}}_w\mathcal{H}^{\mathsf{hGP}}_n = \frac{1}{4}\begin{pmatrix} -\partial_x a^{-1} - a^{-1}\partial_x & \partial_x a^{-1} - a^{-1}\partial_x \\ -\partial_x a^{-1} + a^{-1}\partial_x & \partial_x a^{-1} + a^{-1}\partial_x \end{pmatrix}\left(\frac{\delta\mathcal{H}^{\mathsf{hGP}}_n(w)}{\delta w_+}, \frac{\delta\mathcal{H}^{\mathsf{hGP}}_n(w)}{\delta w_-}\right)^T.$$

The Poisson bracket is

$$\{\mathcal{F},\mathcal{G}\}^{\mathsf{hGP}}(w) = \frac{1}{4}\int_{\mathbb{R}}\begin{pmatrix} \frac{\delta\mathcal{F}}{\delta w_+} & \frac{\delta\mathcal{F}}{\delta w_-} \end{pmatrix}\begin{pmatrix} -\partial_x a^{-1} - a^{-1}\partial_x & \partial_x a^{-1} - a^{-1}\partial_x \\ -\partial_x a^{-1} + a^{-1}\partial_x & \partial_x a^{-1} + a^{-1}\partial_x \end{pmatrix}\begin{pmatrix} \frac{\delta\mathcal{G}}{\delta w_+} \\ \frac{\delta\mathcal{G}}{\delta w_-} \end{pmatrix}\mathrm{d}x\,.$$

4.B.3. $\varepsilon\mathsf{GP}$

We consider formally a manifold where we denote points by $W=(W_+,W_-)$ and tangent vectors by $U=(U_+,U_-)$ and $V=(V_+,V_-)$. We define the Hamiltonian operator $J^{\varepsilon\mathsf{GP}}$, the Riemannian metric $g^{\varepsilon\mathsf{GP}}$ and the symplectic form $\omega^{\varepsilon\mathsf{GP}}$ by

$$\begin{aligned} J_W^{\varepsilon\mathsf{GP}} &= \frac{1}{8}\begin{pmatrix} -\sqrt{2}\varepsilon\partial_X A^{-1} + \frac{4}{\sqrt{2}\varepsilon}A\partial_X^{-1} & \sqrt{2}\varepsilon\partial_X A^{-1} + \frac{4}{\sqrt{2}\varepsilon}A\partial_X^{-1} \\ -\sqrt{2}\varepsilon\partial_X A^{-1} - \frac{4}{\sqrt{2}\varepsilon}A\partial_X^{-1} & \sqrt{2}\varepsilon\partial_X A^{-1} - \frac{4}{\sqrt{2}\varepsilon}A\partial_X^{-1} \end{pmatrix} \\ g_W^{\varepsilon\mathsf{GP}}(U,V) &= \int_{\mathbb{R}} \frac{\varepsilon^3}{4\sqrt{2}}(U_+-U_-)(V_+-V_-) + \frac{\varepsilon}{2\sqrt{2}}A^2\partial_X^{-1}(U_++U_-)\partial_X^{-1}(V_++V_-)\,\mathrm{d}X \\ \omega_W^{\varepsilon\mathsf{GP}}(U,V) &= \frac{\varepsilon^2}{2}\int A\big((V_+-V_-)\partial_X^{-1}(U_++U_-) - (U_+-U_-)\partial_X^{-1}(V_++V_-)\big)\,\mathrm{d}X\,. \end{aligned}$$

They satisfy

$$\omega_W^{\varepsilon\mathsf{GP}}(U, J_W^{\varepsilon\mathsf{GP}}V) = g_W^{\varepsilon\mathsf{GP}}(U,V)\,.$$

We define the gradient $\nabla_W^{g^{\varepsilon\mathsf{GP}}}$ with respect to the metric $g^{\varepsilon\mathsf{GP}}$ as well as the functional derivative via

$$g_W^{\varepsilon\mathsf{GP}}(V,\nabla_W^{g^{\varepsilon\mathsf{GP}}}\mathcal{H}) = \int_{\mathbb{R}}(V_+,V_-)\left(\frac{\delta\mathcal{H}(W)}{\delta W_+},\frac{\delta\mathcal{H}(W)}{\delta W_-}\right)^T\mathrm{d}X = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\mathcal{H}(W+sV)\,.$$

This implies

$$\nabla_W^{g^{\varepsilon\mathsf{GP}}}\mathcal{H} = \frac{1}{\sqrt{2}\varepsilon^3}\begin{pmatrix} 2-\varepsilon^2\partial_X A^{-2}\partial_X & -2-\varepsilon^2\partial_X A^{-2}\partial_X \\ -2-\varepsilon^2\partial_X A^{-2}\partial_X & 2-\varepsilon^2\partial_X A^{-2}\partial_X \end{pmatrix}\left(\frac{\delta\mathcal{H}(W)}{\delta W_+},\frac{\delta\mathcal{H}(W)}{\delta W_-}\right)^T.$$

Then ($\varepsilon\mathsf{GP}_n$) can be written as

$$\partial_{T_n}W = \frac{\mathrm{d}t_n}{\mathrm{d}T_n}J_W^{\varepsilon\mathsf{GP}}\nabla_W^{g^{\varepsilon\mathsf{GP}}}\mathcal{H}_n^{\varepsilon\mathsf{GP}} = \frac{\mathrm{d}t_n}{\mathrm{d}T_n}\frac{1}{2\varepsilon^2}\begin{pmatrix} -\partial_X A^{-1} - A^{-1}\partial_X & \partial_X A^{-1} - A^{-1}\partial_X \\ -\partial_X A^{-1} + A^{-1}\partial_X & \partial_X A^{-1} + A^{-1}\partial_X \end{pmatrix}\left(\frac{\delta\mathcal{H}_n^{\varepsilon\mathsf{GP}}(W)}{\delta W_+},\frac{\delta\mathcal{H}_n^{\varepsilon\mathsf{GP}}(W)}{\delta W_-}\right)^T.$$

The Poisson bracket is

$$\{\mathcal{F},\mathcal{G}\}^{\varepsilon\mathsf{GP}}(W) = \frac{1}{2\varepsilon^2}\int_{\mathbb{R}}\begin{pmatrix}\frac{\delta\mathcal{F}}{\delta W_+} & \frac{\delta\mathcal{F}}{\delta W_-}\end{pmatrix}\begin{pmatrix} -\partial_X A^{-1} - A^{-1}\partial_X & \partial_X A^{-1} - A^{-1}\partial_X \\ -\partial_X A^{-1} + A^{-1}\partial_X & \partial_X A^{-1} + A^{-1}\partial_X \end{pmatrix}\begin{pmatrix}\frac{\delta\mathcal{G}}{\delta W_+} \\ \frac{\delta\mathcal{G}}{\delta W_-}\end{pmatrix}\mathrm{d}X\,.$$

4.B.4. KdV

We consider formally a manifold where we denote points by U and tangent vectors by F and G. We define the Hamiltonian operator J^{KdV}, the Riemannian metric g^{KdV} and the symplectic form ω^{KdV} by

$$\begin{aligned} J_U^{\mathsf{KdV}} &= \frac{1}{2}\partial_X \\ g_U^{\mathsf{KdV}}(F,G) &= \int_{\mathbb{R}} FG\,\mathrm{d}X \\ \omega_U^{\mathsf{KdV}} &= \frac{1}{2}\int_{\mathbb{R}} F\partial_X^{-1}G - G\partial_X^{-1}F\,\mathrm{d}X\,. \end{aligned}$$

They satisfy

$$\omega_U^{\mathsf{KdV}}(F, J_U^{\mathsf{KdV}}G) = g_U^{\mathsf{KdV}}(F,G)\,.$$

We define the gradient $\nabla_U^{g^{\mathsf{KdV}}}$ with respect to the metric g^{KdV} as well as the functional derivative via

$$g_U^{\mathsf{KdV}}(F,\nabla_U^{g^{\mathsf{KdV}}}\mathcal{H}) = \int_{\mathbb{R}} F\frac{\delta\mathcal{H}(U)}{\delta U}\,\mathrm{d}X = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\mathcal{H}(U+sF)\,.$$

This implies

$$\nabla_U^{g^{\mathsf{KdV}}} \mathcal{H} = \frac{\delta \mathcal{H}(U)}{\delta U} \, .$$

Then (KdV_n) can be written as

$$\partial_{T_n} U = J_U^{\mathsf{KdV}} \nabla_U^{g^{\mathsf{KdV}}} \mathcal{E}_n^{\mathsf{KdV}} = \frac{1}{2} \partial_X \frac{\delta \mathcal{E}_n^{\mathsf{KdV}}(U)}{\delta U} \, .$$

The Poisson bracket is

$$\{\mathcal{F}, \mathcal{G}\}^{\mathsf{KdV}}(U) = \frac{1}{2} \int_{\mathbb{R}} \frac{\delta \mathcal{F}}{\delta U} \partial_X \frac{\delta \mathcal{G}}{\delta U} \, \mathrm{d}X \, .$$

Bibliography

[1] M. J. Ablowitz and P. A. Clarkson. *Solitons, nonlinear evolution equations and inverse scattering*. Vol. 149. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 1991, pp. xii+516.

[2] M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur. "The inverse scattering transform-Fourier analysis for nonlinear problems". In: *Stud. Appl. Math.* 53.4 (1974), pp. 249–315.

[3] M. J. Ablowitz, B. Prinari, and A. D. Trubatch. "Discrete vector solitons: composite solitons, Yang-Baxter maps and computation". In: *Stud. Appl. Math.* 116.1 (2006), pp. 97–133.

[4] J. Adams. "Local and global well-posedness for the higher-order NLS & dNLS hierarchy equations". English. Dissertation (Ph.D.)–Heinrich Heine Universität Düsseldorf; supervisors: A. Grünrock and J. Saal. PhD thesis. Düsseldorf: Heinrich Heine Universität Düsseldorf, 2025.

[5] J. Adams. "Well-posedness for the NLS hierarchy". In: *J. Evol. Equ.* 24.4 (2024), p. 88.

[6] P. Antonelli. "Remarks on the derivation of finite energy weak solutions to the QHD system". In: *Proc. Amer. Math. Soc.* 149.5 (2021), pp. 1985–1997.

[7] P. Antonelli, L.E. Hientzsch, and P. Marcati. "Finite energy well-posedness for nonlinear Schrödinger equations with non-vanishing conditions at infinity". In: *arXiv e-prints* (2023).

[8] P. Antonelli, L.E. Hientzsch, P. Marcati, and H. Zheng. "On some results for quantum hydrodynamical models". In: *Mathematical Analysis in Fluid and Gas Dynamics*. Ed. by T. Kobayashi. Vol. 2070. RIMS Kokyuroku, 2018, pp. 107–129.

[9] P. Antonelli and P. Marcati. "On the finite energy weak solutions to a system in quantum fluid dynamics". In: *Comm. Math. Phys.* 287.2 (2009), pp. 657–686.

[10] P. Antonelli and P. Marcati. "The quantum hydrodynamics system in two space dimensions". In: *Arch. Ration. Mech. Anal.* 203.2 (2012), pp. 499–527.

[11] P. Antonelli, P. Marcati, and R. Scandone. "Existence and stability of almost finite energy weak solutions to the quantum Euler-Maxwell system". In: *J. Math. Pures Appl. (9)* 191, 103629 (2024).

[12] P. Antonelli, P. Marcati, and H. Zheng. "An intrinsically hydrodynamic approach to multidimensional QHD systems". In: *Arch. Ration. Mech. Anal.* 247.2 (2023), p. 24.

[13] P. Antonelli, P. Marcati, and H. Zheng. "Genuine hydrodynamic analysis to the 1-D QHD system: existence, dispersion and stability". In: *Comm. Math. Phys.* 383.3 (2021), pp. 2113–2161.

[14] H. Aratyn and A. S. Sorin, eds. *Integrable hierarchies and modern physical theories*. Vol. 18. NATO Science Series II: Mathematics, Physics and Chemistry. Kluwer Academic Publishers, Dordrecht, 2001, pp. x+435.

[15] V. I. Arnold. *Mathematical methods of classical mechanics*. Vol. 60. Graduate Texts in Mathematics. Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second (1989) edition. Springer-Verlag, New York, 1989, pp. xvi+516.

[16] C. Audiard. "Global well-posedness of a system from quantum hydrodynamics for small data". In: *Confluentes Math.* 7.2 (2015), pp. 7–16.

[17] C. Audiard. "On the time of existence of solutions of the Euler-Korteweg system". In: *Ann. Fac. Sci. Toulouse Math. (6)* 30.5 (2021), pp. 1139–1183.

[18] C. Audiard and B. Haspot. "Global well-posedness of the Euler-Korteweg system for small irrotational data". In: *Comm. Math. Phys.* 351.1 (2017), pp. 201–247.

[19] I. Avramidi and R. Schimming. "A new explicit expression for the Korteweg-de Vries hierarchy". In: *Math. Nachr.* 219 (2000), pp. 45–64.

[20] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier analysis and nonlinear partial differential equations*. Vol. 343. Grundlehren der mathematischen Wissenschaften. Springer, Heidelberg, 2011, pp. xvi+523.

[21] J.-B. Baillon, T. Cazenave, and M. Figueira. “Équation de Schrödinger non linéaire”. In: *C. R. Acad. Sci. Paris Sér. A-B* 284.15 (1977), A869–A872.

[22] D. Bambusi. “Hamiltonian studies on counter-propagating water waves”. In: *Water Waves* 3.1 (2021), pp. 49–83.

[23] W. Bao. “The nonlinear Schrödinger equation and applications in Bose-Einstein condensation and plasma physics”. In: *Dynamics in models of coarsening, coagulation, condensation and quantization*. Vol. 9. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. World Sci. Publ., Hackensack, NJ, 2007, pp. 141–239.

[24] L. Bergé. “Wave collapse in physics: principles and applications to light and plasma waves”. In: *Phys. Rep.* 303.5-6 (1998), pp. 259–370.

[25] F. Béthuel, R. Danchin, and D. Smets. “On the linear wave regime of the Gross-Pitaevskii equation”. In: *J. Anal. Math.* 110 (2010), pp. 297–338.

[26] F. Béthuel, P. Gravejat, and J.-C. Saut. “Existence and properties of travelling waves for the Gross-Pitaevskii equation”. In: *Stationary and time dependent Gross-Pitaevskii equations*. Vol. 473. Contemp. Math. Amer. Math. Soc., Providence, RI, 2008, pp. 55–103.

[27] F. Béthuel, P. Gravejat, J.-C. Saut, and D. Smets. “On the Korteweg-de Vries long-wave approximation of the Gross-Pitaevskii equation II”. In: *Comm. Partial Differ. Equ.* 35.1 (2010), pp. 113–164.

[28] F. Béthuel, P. Gravejat, J.-C. Saut, and D. Smets. “On the Korteweg-de Vries long-wave approximation of the Gross-Pitaevskii equation. I”. In: *Int. Math. Res. Not. IMRN* 14 (2009), pp. 2700–2748.

[29] S. Bianchini. “Exact integrability conditions for cotangent vector fields”. In: *Manuscr. Math.* (2023).

[30] G. Biondini and G. Kovačič. “Inverse scattering transform for the focusing nonlinear Schrödinger equation with nonzero boundary conditions”. In: *J. Math. Phys.* 55.3, 031506 (2014).

[31] D. Bohm. “A suggested interpretation of the quantum theory in terms of “hidden” variables”. In: *Phsy. Rev.* 85.2 (1952), pp. 166–179.

[32] J. L. Bona and R. Smith. “The initial-value problem for the Korteweg-de Vries equation”. In: *Philos. Trans. Roy. Soc. London Ser. A* 278.1287 (1975), pp. 555–601.

[33] J.-M. Bony. “Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires”. In: *Ann. Sci. École Norm. Sup. (4)* 14.2 (1981), pp. 209–246.

[34] M. Borghese, R. Jenkins, and K. D. T.-R. McLaughlin. “Long time asymptotic behavior of the focusing nonlinear Schrödinger equation”. In: *Ann. Inst. Henri Poincaré (C) Anal. Non Linéaire* 35.4 (2018), pp. 887–920.

[35] J. Bourgain. “Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. II. The KdV-equation”. In: *Geom. Funct. Anal.* 3.3 (1993), pp. 209–262.

[36] D. Bresch, M. Gisclon, and I. Lacroix-Violet. “On Navier-Stokes-Korteweg and Euler-Korteweg systems: application to quantum fluids models”. In: *Arch. Ration. Mech. Anal.* 233.3 (2019), pp. 975–1025.

[37] T. J. Bridges. “A universal form for the emergence of the Korteweg-de Vries equation”. In: *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 469.2153, 20120707 (2013), p. 15.

[38] B. Bringmann, R. Killip, and M. Visan. “Global well-posedness for the fifth-order KdV equation in $H^{-1}(\mathbb{R})$”. In: *Ann. PDE* 7.2 (2021), p. 21.

[39] W. Bulla, F. Gesztesy, and K. Unterkofler. “Holomorphy of the scattering matrix with respect to c^{-2} for Dirac operators and an explicit treatment of relativistic corrections”. In: *Comm. Math. Phys.* 144.2 (1992), pp. 391–416.

[40] D. L. Burkholder. “A geometric condition that implies the existence of certain singular integrals of Banach-space-valued functions”. In: *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)*. Wadsworth Math. Ser. Wadsworth, Belmont, CA, 1983, pp. 270–286.

[41] N. Burq and F. Planchon. "Smoothing and dispersive estimates for 1D Schrödinger equations with BV coefficients and applications". In: *J. Funct. Anal.* 236.1 (2006), pp. 265–298.

[42] F. Calogero. "Why Are Certain Nonlinear PDEs Both Widely Applicable and Integrable?" In: *What Is Integrability?* Ed. by V. E. Zakharov. Berlin, Heidelberg: Springer, 1991, pp. 1–62.

[43] R. Carles, R. Danchin, and J.-C. Saut. "Madelung, Gross-Pitaevskii and Korteweg". In: *Nonlinearity* 25.10 (2012), pp. 2843–2873.

[44] T. Cazenave. *Semilinear Schrödinger equations*. Vol. 10. Courant Lecture Notes in Mathematics. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2003, pp. xiv+323.

[45] J. Chandra and P. W. Davis. "Linear generalizations of Gronwall's inequality". In: *Proc. Amer. Math. Soc.* 60 (1976), pp. 157–160.

[46] X.-J. Chen, Z.-D. Chen, and N.-N. Huang. "A direct perturbation theory for dark solitons based on a complete set of the squared Jost solutions". In: *J. phys. A: Math. Nucl. Gen.* 31.33 (1998), pp. 6929–6947.

[47] M. Chirilus-Bruckner, W.-P. Düll, and G. Schneider. "Validity of the KdV equation for the modulation of periodic traveling waves in the NLS equation". In: *J. Math. Anal. Appl.* 414.1 (2014), pp. 166–175.

[48] D. Chiron and F. Rousset. "The KdV/KP-I limit of the nonlinear Schrödinger equation". In: *SIAM J. Math. Anal.* 42.1 (2010), pp. 64–96.

[49] M. Christ, J. Colliander, and T. Tao. "A priori bounds and weak solutions for the nonlinear Schrödinger equation in Sobolev spaces of negative order". In: *J. Funct. Anal.* 254.2 (2008), pp. 368–395.

[50] M. Christ and A. Kiselev. "Maximal functions associated to filtrations". In: *J. Funct. Anal.* 179.2 (2001), pp. 409–425.

[51] A. Church. "A Note on the Entscheidungsproblem". In: *Journal of Symbolic Logic* 1.1 (1936), pp. 40–41.

[52] A. Church. "An Unsolvable Problem of Elementary Number Theory". In: *Amer. J. Math.* 58.2 (1936), pp. 345–363.

[53] W. Craig. "An existence theory for water waves and the Boussinesq and Korteweg-de Vries scaling limits". In: *Comm. Partial Differential Equations* 10.8 (1985), pp. 787–1003.

[54] W. Craig and M. D. Groves. "Hamiltonian long-wave approximations to the water-wave problem". In: *Wave Motion* 19.4 (1994), pp. 367–389.

[55] J. P. Crutchfield and K. Young. "Computation at the onset of chaos". In: *Complexity, entropy and the physics of information (Santa Fe, NM, 1989)*. Vol. VIII. Santa Fe Inst. Stud. Sci. Complexity Proc. Addison-Wesley, Redwood City, CA, 1990, pp. 223–269.

[56] S. Cuccagna and R. Jenkins. "On the asymptotic stability of N-soliton solutions of the defocusing nonlinear Schrödinger equation". In: *Comm. Math. Phys.* 343.3 (2016), pp. 921–969.

[57] L. I. Hedberg D. R. Adams. *Function Spaces and Potential Theory*. Grundlehren der mathematischen Wissenschaften. Berlin, Heidelberg: Springer-Verlag, 1996.

[58] R. Danchin and X. Liao. "On the well-posedness of the full low Mach number limit system in general critical Besov spaces". In: *Commun. Contemp. Math.* 14.3, 1250022 (2012).

[59] R. Darren and K. Steiglitz. "Computing with Solitons". In: *Encyclopedia of Complexity and Systems Science*. Ed. by Robert A. Meyers. New York, NY: Springer New York, 2009, pp. 1376–1395.

[60] P. Deift and E. Trubowitz. "Inverse scattering on the line". In: *Comm. Pure Appl. Math.* 32.2 (1979), pp. 121–251.

[61] P. A. Deift. "Three lectures on "Fifty years of KdV: an integrable system"". In: *Nonlinear dispersive partial differential equations and inverse scattering*. Vol. 83. Fields Inst. Commun. Springer, New York, 2019, pp. 3–38.

[62] H. Demiray. “Modified reductive perturbation method as applied to long water-waves: the Korteweg-de Vries hierarchy”. In: *Int. J. Nonlinear Sci.* 6.1 (2008), pp. 11–20.

[63] F. Demontis and C. van der Mee. “From the AKNS system to the matrix Schrödinger equation with vanishing potentials: direct and inverse problems”. In: *Stud. Appl. Math.* 150.2 (2023), pp. 481–519.

[64] F. Demontis, B. Prinari, C. van der Mee, and F. Vitale. “The inverse scattering transform for the defocusing nonlinear Schrödinger equations with nonzero boundary conditions”. In: *Stud. Appl. Math.* 131.1 (2013), pp. 1–40.

[65] F. Demontis, B. Prinari, C. van der Mee, and F. Vitale. “The inverse scattering transform for the focusing nonlinear Schrödinger equation with asymmetric boundary conditions”. In: *J. Math. Phys.* 55.10, 101505 (2014).

[66] L. A. Dickey. *Soliton Equations and Hamiltonian Systems*. Vol. 12. Advanced Series in Mathematical Physics. Singapore: World Scientific, 1991, p. 310.

[67] R. Donagi and T. Shaska, eds. *Integrable systems and algebraic geometry. Vol. 1*. Vol. 458. London Mathematical Society Lecture Note Series. A celebration of Emma Previato’s 65th birthday. Cambridge University Press, Cambridge, 2020, pp. viii+412.

[68] R. Donagi and T. Shaska, eds. *Integrable systems and algebraic geometry. Vol. 2*. Vol. 459. London Mathematical Society Lecture Note Series. A celebration of Emma Previato’s 65th birthday. Cambridge University Press, Cambridge, 2020, pp. viii+528.

[69] G. A. El and M. A. Hoefer. “Dispersive shock waves and modulation theory”. In: *Phys. D* 333 (2016), pp. 11–65.

[70] L. Erdős, B. Schlein, and H.-T. Yau. “Derivation of the Gross-Pitaevskii hierarchy for the dynamics of Bose-Einstein condensate”. In: *Comm. Pure Appl. Math.* 59.12 (2006), pp. 1659–1741.

[71] L. D. Faddeev and L. A. Takhtajan. *Hamiltonian methods in the theory of solitons*. English. Classics in Mathematics. Translated from the 1986 Russian original by Alexey G. Reyman. Springer, Berlin, 2007, pp. x+592.

[72] R.P. Feynman. “Superfluidity and superconductivity”. In: *Rev. Mod. Phys.* 29.2 (1957), pp. 205–212.

[73] D. Franco, S. Giorgini, L.P. Pitaevskii, and S. Stringari. “Theory of Bose-Einstein condensation in trapped gases”. In: *Rev. Mod. Phys.* 71.3 (1999), pp. 463–512.

[74] C. Gallo. “Schrödinger group on Zhidkov spaces”. In: *Adv. Differ. Equ.* 9.5-6 (2004), pp. 509–538.

[75] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura. “Method for Solving the Korteweg-de Vries Equation”. In: *Phys. Rev. Lett.* 19.19 (Nov. 1967), pp. 1095–1097.

[76] C.L. Gardner. “The quantum hydrodynamic model for semiconductor devices”. In: *SIAM J. Appl. Math.* 54.2 (1994), pp. 409–427.

[77] P. Gérard. “The Cauchy problem for the Gross-Pitaevskii equation”. In: *Ann. Inst. Henri Poincaré (C) Anal. Non Linéaire* 23.5 (2006), pp. 765–779.

[78] P. Gérard. “The Gross-Pitaevskii equation in the energy space”. In: Contemp. Math. 473 (2008), pp. 129–148.

[79] F. Gesztesy, H. Grosse, and B. Thaller. “A rigorous approach to relativistic corrections of bound state energies for spin-1/2 particles”. In: *Ann. Inst. Henri Poincaré Phys. Théor.* 40.2 (1984), pp. 159–174.

[80] F. Gesztesy and H. Holden. *Soliton equations and their algebro-geometric solutions. Vol. I*. Vol. 79. Cambridge Studies in Advanced Mathematics. $(1+1)$-dimensional continuous models. Cambridge University Press, Cambridge, 2003, pp. xii+505.

[81] F. Gesztesy, H. Holden, J. Michor, and G. Teschl. *Soliton equations and their algebro-geometric solutions. Vol. II*. Vol. 114. Cambridge Studies in Advanced Mathematics. $(1+1)$-dimensional discrete models. Cambridge University Press, Cambridge, 2008, pp. x+438.

[82] J. Ginibre and G. Velo. “On a class of nonlinear Schrödinger equations. I. The Cauchy problem, general case”. In: *J. Funct. Anal.* 32.1 (1979), pp. 1–32.

[83] J. Grant. “Pressure and stress tensor expressions in the fluid mechanical formulation of the Bose condensate equations”. In: *J. phys. A: Math. Nucl. Gen.* 6.11 (1973), pp. L151–L153.

[84] W. Grey and G. Sinnamon. “Product operators on mixed norm spaces”. In: *Linear and Nonlinear Analysis* 2.2 (2016), pp. 189–197.

[85] E.P. Gross. “Hydrodynamics of a superfluid condensate”. In: *J. Math. Phys.* 4.2 (1963), pp. 195–207.

[86] A. Grünrock. “An improved local well-posedness result for the modified KdV equation”. In: *Int. Math. Res. Not.* 61 (2004), pp. 3287–3308.

[87] A. Grünrock. “Bi- and trilinear Schrödinger estimates in one space dimension with applications to cubic NLS and DNLS”. In: *Int. Math. Res. Not.* 41 (2005), pp. 2525–2558.

[88] A. Grünrock. “On the hierarchies of higher order mKdV and KdV equations”. In: *Cent. Eur. J. Math.* 8.3 (2010), pp. 500–536.

[89] S. Gustafson, K. Nakanishi, and T.-P. Tsai. “Scattering theory for the Gross-Pitaevskii equation in three dimensions”. In: *Commun. Contemp. Math.* 11.4 (2009), pp. 657–707.

[90] J. Hadamard. “Les surfaces à courbures opposées et leurs lignes géodésiques”. In: *Journal de mathématiques pures et appliquées.* 5th ser. 4 (1898), pp. 27–73.

[91] B. Harrop-Griffiths, R. Killip, and M. Vişan. “Sharp well-posedness for the cubic NLS and mKdV in $H^s(\mathbb{R})$”. In: *Forum Math. Pi* 12 (2024), e6.

[92] A. Hasegawa. “An historical review of application of optical solitons for high speed communications”. In: vol. 10. 3. Optical solitons—perspectives and applications. 2000, pp. 475–485.

[93] L.E. Hientzsch. “Nonlinear Schrödinger equations and quantum fluids non vanishing at infinity: incompressible limit and quantum vortices”. PhD thesis. Gran Sasso Science Insitute, 2019.

[94] Y. Hiraoka and Y. Kodama. “Normal form and solitons”. In: *Integrability.* Vol. 767. Lecture Notes in Phys. Springer, Berlin, 2009, pp. 175–214.

[95] S. Hofbauer. “Long wave approximation over and beyond the natural time scale”. PhD thesis. University of Stuttgart, Apr. 2024.

[96] M. Ifrim and D. Tataru. “Global solutions for 1D cubic defocusing dispersive equations: Part I”. In: *Forum Math. Pi* 11 (2023), e31.

[97] R. I. Ivanov. “Hamiltonian models for the propagation of long gravity waves, higher-order KdV-type equations and integrability”. In: *Nonlinear dispersive waves.* Adv. Math. Fluid Mech. Birkhäuser/Springer, Cham, 2024, pp. 81–97.

[98] M. H. Jakubowski, K. Steiglitz, and R. Squier. “State transformations of colliding optical solitons and possible application to computation in bulk media”. In: *Phys. Rev. E* 58.5 (Nov. 1998), pp. 6752–6758.

[99] M. Schwarz Jr. “The initial value problem for the sequence of generalized Korteweg-de Vries equations”. In: *Adv. Math.* 54.1 (1984), pp. 22–56.

[100] L. A. Kalyakin. “Long-wave asymptotics. Integrable equations as the asymptotic limit of nonlinear systems”. In: *Uspekhi Mat. Nauk* 44.1(265) (1989), pp. 5–34, 247.

[101] T. Kappeler and E. Trubowitz. “Properties of the scattering map”. In: *Comment. Math. Helv.* 61.3 (1986), pp. 442–480.

[102] T. Kappeler and E. Trubowitz. “Properties of the scattering map. II”. In: *Comment. Math. Helv.* 63.1 (1988), pp. 150–167.

[103] T. Kato. “The Cauchy problem for the Korteweg-de Vries equation”. In: *Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. I (Paris, 1978/1979).* Vol. 53. Res. Notes in Math. Pitman, Boston, Mass.-London, 1981, pp. 293–307.

[104] C. E. Kenig and D. Pilod. “Local well-posedness for the KdV hierarchy at high regularity”. In: *Adv. Differ. Equ.* 21.9-10 (2016), pp. 801–836.

[105] C. E. Kenig, G. Ponce, and L. Vega. “Higher-order nonlinear dispersive equations”. In: *Proc. Amer. Math. Soc.* 122.1 (1994), pp. 157–166.

[106] C. E. Kenig, G. Ponce, and L. Vega. "On the generalized Benjamin-Ono equation". In: *Trans. Amer. Math. Soc.* 342.1 (1994), pp. 155–172.

[107] C. E. Kenig, G. Ponce, and L. Vega. "On the hierarchy of the generalized KdV equations". In: *Singular limits of dispersive waves (Lyon, 1991)*. Vol. 320. NATO Adv. Sci. Inst. Ser. B: Phys. Plenum, New York, 1994, pp. 347–356.

[108] C. E. Kenig, G. Ponce, and L. Vega. "Oscillatory integrals and regularity of dispersive equations". In: *Indiana Univ. Math. J.* 40.1 (1991), pp. 33–69.

[109] C. E. Kenig, G. Ponce, and L. Vega. "Small solutions to nonlinear Schrödinger equations". In: *Ann. Inst. Henri Poincaré (C) Anal. Non Linéaire* 10.3 (1993), pp. 255–288.

[110] C. E. Kenig, G. Ponce, and L. Vega. "Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle". In: *Comm. Pure Appl. Math.* 46.4 (1993), pp. 527–620.

[111] C. E. Kenig, G. Ponce, and L. Vega. "Well-posedness of the initial value problem for the Korteweg-de Vries equation". In: *J. Amer. Math. Soc.* 4.2 (1991), pp. 323–347.

[112] R. Killip, T. Laurens, and M. Vişan. "Sharp well-posedness for the Benjamin-Ono equation". In: *Invent. Math.* 236.3 (2024), pp. 999–1054.

[113] R. Killip, T. Oh, O. Pocovnicu, and M. Vişan. "Global well-posedness of the Gross-Pitaevskii and cubic-quintic nonlinear Schrödinger equations with non-vanishing boundary conditions". In: *Math. Res. Lett.* 19.5 (2012), pp. 969–986.

[114] R. Killip and M. Vișan. "KdV is well-posed in H^{-1}". In: *Ann. of Math. (2)* 190.1 (2019), pp. 249–305.

[115] R. Killip, M. Vişan, and X. Zhang. "Low regularity conservation laws for integrable PDE". In: *Geom. Funct. Anal.* 28.4 (2018), pp. 1062–1090.

[116] F. Klaus. "Nonlinear Schrödinger Equations with Rough Data". PhD thesis. Karlsruher Institut für Technologie (KIT), 2022. 143 pp.

[117] F. Klaus. "Wellposedness of NLS in modulation spaces". In: *J. Fourier Anal. Appl.* 29.1 (2023), p. 9.

[118] F. Klaus, H. Koch, and B. Liu. "Well-posedness for the KdV hierarchy". In: *arXiv e-prints* (2023), pp. 1–139.

[119] H. Koch and X. Liao. "Conserved energies for the one dimensional Gross-Pitaevskii equation". In: *Adv. Math.* 377, 107467 (2021).

[120] H. Koch and X. Liao. "Conserved energies for the one dimensional Gross-Pitaevskii equation: low regularity case". In: *Adv. Math.* 420, 108996 (2023).

[121] H. Koch and D. Tataru. "A priori bounds for the 1D cubic NLS in negative Sobolev spaces". In: *Int. Math. Res. Not. IMRN* 16 (2007), Art. ID rnm053, 36.

[122] H. Koch and D. Tataru. "Conserved energies for the cubic nonlinear Schrödinger equation in one dimension". In: *Duke Math. J.* 167.17 (2018), pp. 3207–3313.

[123] H. Koch and D. Tataru. "Energy and local energy bounds for the 1-d cubic NLS equation in $H^{-\frac{1}{4}}$". In: *Ann. Inst. Henri Poincaré (C) Anal. Non Linéaire* 29.6 (2012), pp. 955–988.

[124] Y. Kodama and T. Taniuti. "Higher order approximation in the reductive perturbation method. III. The weakly dissipative system". In: *J. Phys. Soc. Japan* 47.5 (1979), pp. 1706–1716.

[125] D. J. Korteweg and G. de Vries. "On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves". In: *Philos. Mag. (5)* 39.240 (1895), pp. 422–443.

[126] R. A. Kraenkel, M. A. Manna, and J. G. Pereira. "The Korteweg-de Vries hierarchy and long water-waves". In: *J. Math. Phys.* 36.1 (1995), pp. 307–320.

[127] L. Landau. "Theory of the superfluidity of Helium II". In: *Phys. Rev.* 60.4 (1941), pp. 356–358.

[128] S. Lang. *Fundamentals of differential geometry*. Vol. 191. Graduate Texts in Mathematics. Springer-Verlag, New York, 1999, pp. xviii+535.

[129] C. G. Langton. “Computation at the edge of chaos: Phase transitions and emergent computation”. In: *Physica D: Nonlinear Phenomena* 42.1 (1990), pp. 12–37.

[130] P.-S. Laplace. *A Philosophical Essay on Probabilities.* Trans. by Frederick Wilson Truscott and Frederick Lincoln Emory. First. Translated from the sixth French edition. New York: John Wiley & Sons, 1902.

[131] P.-S. Laplace. *Essai philosophique sur les probabilités.* Paris Bachelier, 1840.

[132] P. D. Lax. “Integrals of nonlinear equations of evolution and solitary waves”. In: *Communications on Pure and Applied Mathematics* 21.5 (1968), pp. 467–490.

[133] P. D. Lax. “Periodic solutions of the KdV equation”. In: *Comm. Pure Appl. Math.* 28 (1975), pp. 141–188.

[134] X. Liao and M. Plum. “Eigenvalue analysis of the lax operator for the one-dimensional cubic nonlinear defocusing Schrödinger equation”. In: *SIAM J. Math. Anal.* 55.6 (2023), pp. 7399–7425.

[135] X. Liao and R. Wegner. *Global well-posedness of the NLS hierarchy with nonzero boundary condition.* CRC 1173 Preprint. 2025.

[136] J. Liouville. “Note sur l’intégration des équations différentielles de la Dynamique, présentée au Bureau des Longitudes le 29 juin 1853”. In: *J. Math. Pures Appl.* (1855), pp. 137–138.

[137] Y. Liu and X. Yang. “On the long-wave approximation for the Euler-Poisson system”. In: *Adv. Math.* 372 (2020), pp. 107300, 41.

[138] M. I. Loffredo and L. M. Morato. “On the creation of quantized vortex lines in rotating He II”. In: *Il Nuovo Cimento B (1971-1996)* 108.2 (1993), pp. 205–215.

[139] P. Markowich and J. Sierra. “Non-uniqueness of weak solutions of the quantum-hydrodynamic system”. In: *Kinet. Relat. Models.* 12.2 (2019), pp. 347–356.

[140] W. McLean. *Strongly elliptic systems and boundary integral equations.* Cambridge University Press, Cambridge, 2000, pp. xiv+357.

[141] C. van der Mee. “Focusing NLS equations with nonzero boundary conditions: triangular representations and direct scattering”. In: *J. Nonlinear Math. Phys.* 28.1 (2021), pp. 68–89.

[142] C. van der Mee. “Inverse scattering transform for the discrete focusing nonlinear Schrödinger equation with nonvanishing boundary conditions”. In: *J. Nonlinear Math. Phys.* 22.2 (2015), pp. 233–264.

[143] D. Merlini, R. Sprugnoli, and M. C. Verri. “Lagrange inversion: when and how”. In: *Acta Appl. Math.* 94.3 (2006), pp. 233–249.

[144] A. Meurer, C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K. Moore, S. Singh, T. Rathnayake, S. Vig, B. E. Granger, R. P. Muller, F. Bonazzi, H. Gupta, S. Vats, F. Johansson, F. Pedregosa, M. J. Curry, A. R. Terrel, Š. Roučka, A. Saboo, I. Fernando, S. Kulal, R. Cimrman, and A. Scopatz. “SymPy: symbolic computing in Python”. In: *PeerJ Comput. Sci.* 3 (Jan. 2017), e103.

[145] H. Mohamad. “Hydrodynamical form for the one-dimensional Gross-Pitaevskii equation”. In: *Electron. J. Differ. Equ.* (2014), No. 141, 27.

[146] L. Molinet. “A note on ill posedness for the KdV equation”. In: *Differential Integral Equations* 24.7-8 (2011), pp. 759–765.

[147] E. Nelson. “Derivation of the Schrödinger equation from Newtonian mechanics”. In: *Phys. Rev.* 150.4 (1966), pp. 1079–1085.

[148] T. Oh and Y. Wang. “Global well-posedness of the one-dimensional cubic nonlinear Schrödinger equation in almost critical spaces”. In: *J. Differ. Equ.* 269.1 (2020), pp. 612–640.

[149] T. Ozawa, S.-F. Tian, and Z.-J. Wu. “Orbital stability of dark solitons for the modified Korteweg–de Vries hierarchy with nonzero boundary conditions”. In: *Math. Ann.* 394.1 (2026), Paper No. 23, 51.

[150] R. S. Palais. “The symmetries of solitons”. In: *Bull. Amer. Math. Soc. (N.S.)* 34.4 (1997), pp. 339–403.

[151] H. Pecher. “Unconditional global well-posedness for the 3D Gross-Pitaevskii equation for data without finite energy”. In: *Nonlinear Differ. Equ. Appl.* 20.6 (2013), pp. 1851–1877.

[152] D. Pilod. “On the Cauchy problem for higher-order nonlinear dispersive equations”. In: *J. Differ. Equ.* 245.8 (2008), pp. 2055–2077.

[153] L.P. Pitaevskii. “Vortex lines in an imperfect Bose gas”. In: *Sov. Phys. JETP* 13 (1961), pp. 451–454.

[154] H. Poincaré. “Sur le problème des trois corps et les équations de la dynamique”. In: *Acta Mathematica* 13.1–2 (1890), pp. 1–270.

[155] B. Prinari. “Inverse scattering transform for nonlinear Schrödinger systems on a nontrivial background: a survey of classical results, new developments and future directions”. In: *J. Nonlinear Math. Phys.* 30.2 (2023), pp. 317–383.

[156] R. R. Beals and R. R. Coifman. “Scattering and inverse scattering for first order systems”. In: *Comm. Pure Appl. Math.* 37.1 (1984), pp. 39–90.

[157] M. Reddiger and B. Poirier. “Towards a mathematical theory of the Madelung equations: Takabayasi’s quantization condition, quantum quasi-irrotationality, weak formulations, and the Wallstrom phenomenon”. In: *J. phys. A: Math. Nucl. Gen.* 56.19 (2023), p. 193001.

[158] A. Rybkin. “Regularized perturbation determinants and KdV conservation laws for irregular initial profiles”. In: *Topics in operator theory. Volume 2. Systems and mathematical physics.* Vol. 203. Oper. Theory Adv. Appl. Birkhäuser Verlag, Basel, 2010, pp. 427–444.

[159] J.-C. Saut. “Quelques généralisations de l’équation de Korteweg-de Vries. II”. In: *J. Differential Equations* 33.3 (1979), pp. 320–335.

[160] G. Schneider. “Approximation of the Korteweg-de Vries equation by the nonlinear Schrödinger equation”. In: *J. Differential Equations* 147.2 (1998), pp. 333–354.

[161] G. Schneider. “The long wave limit for a Boussinesq equation”. In: *SIAM J. Appl. Math.* 58.4 (1998), pp. 1237–1245.

[162] G. Schneider and H. Uecker. *Nonlinear PDEs*. Vol. 182. Graduate Studies in Mathematics. A dynamical systems approach. American Mathematical Society, Providence, RI, 2017, pp. xiii+575.

[163] G. Schneider and C. E. Wayne. “The long-wave limit for the water wave problem. I. The case of zero surface tension”. In: *Comm. Pure Appl. Math.* 53.12 (2000), pp. 1475–1535.

[164] G. Schneider and C. E. Wayne. “The rigorous approximation of long-wavelength capillary-gravity waves”. In: *Arch. Ration. Mech. Anal.* 162.3 (2002), pp. 247–285.

[165] K. Steiglitz. “Time-gated Manakov spatial solitons are computationally universal”. In: *Phys. Rev. E* 63.1 (Dec. 2000), p. 016608.

[166] E. M. Stein. “Interpolation of linear operators”. In: *Trans. Amer. Math. Soc.* 83 (1956), pp. 482–492.

[167] E. M. Stein. *Oscillatory integrals in Fourier analysis.* Vol. 112. Beijing Lectures in Harmonic Analysis. Princeton Univ. Press, 1986, pp. 307–355.

[168] T. Takabayasi. “On the formulation of quantum mechanics associated with classical pictures”. In: *Progr. Theoret. Phys.* 8 (1952), pp. 143–182.

[169] T. Taniuti. “Reductive Perturbation Method and Far Fields of Wave Equations”. In: *Progress of Theoretical Physics Supplement* 55 (Jan. 1974), pp. 1–35.

[170] T. Tao. *Nonlinear dispersive equations.* Vol. 106. CBMS Regional Conference Series in Mathematics. Local and global analysis. Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006, pp. xvi+373.

[171] B. Thaller. *The Dirac equation.* Texts and Monographs in Physics. Springer-Verlag, Berlin, 1992, pp. xviii+357.

[172] N. Totz and S. Wu. “A rigorous justification of the modulation approximation to the 2D full water wave problem”. In: *Comm. Math. Phys.* 310.3 (2012), pp. 817–883.

[173] Y. Tsutsumi. “L^2-solutions for nonlinear Schrödinger equations and nonlinear groups”. In: *Funkcial. Ekvac.* 30.1 (1987), pp. 115–125.

[174] A. M. Turing. "On computable numbers, with an application to the Entscheidungsproblem (1936)". In: *Ideas that created the future—classic papers of computer science*. Reprinted from [1577030]. MIT Press, Cambridge, MA, 2021, pp. 51–60.

[175] V. E. Zakharov S. V. and Manakov. "On the complete integrability of a nonlinear Schrödinger equation". In: *Theoretical and Mathematical Physics* 19.3 (1974), pp. 551–559.

[176] A. Vargas and L. Vega. "Global wellposedness for 1D non-linear Schrödinger equation for data with an infinite L^2 norm". In: *J. Math. Pures Appl. (9)* 80.10 (2001), pp. 1029–1044.

[177] T.C. Wallstrom. "Inequivalence between the Schrödinger equation and the Madelung hydrodynamic equations". In: *Phys. Rev. A (3)* 49.3 (1994), pp. 1613–1617.

[178] T.C. Wallstrom. "On the derivation of the Schrödinger equation from stochastic mechanics". In: *Found. Phys. Lett.* 2.2 (1989), pp. 113–126.

[179] T.C. Wallstrom. "On the initial value problem for the Madelung hydrodynamic equations". In: *Phys. Lett. A* 184.3 (1994), pp. 229–233.

[180] R. Wegner. "Global-in-time well-posedness of the one-dimensional hydrodynamic Gross-Pitaevskii equations without vacuum". In: *Z. Angew. Math. Phys.* 74.5 (2023), p. 194.

[181] G. B. White. "Splitting of the Dirac operator in the nonrelativistic limit". In: *Ann. Inst. H. Poincaré Phys. Théor.* 53.1 (1990), pp. 109–121.

[182] G. B. Whitham. *Linear and nonlinear waves*. Pure and Applied Mathematics (New York). Reprint of the 1974 original, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999, pp. xviii+636.

[183] G. B. Whitham. "Non-linear dispersive waves". In: *Proc. Roy. Soc. Lond. Ser. A* 283 (1965), pp. 238–261.

[184] S. Wolfram. "Statistical mechanics of cellular automata". In: *Rev. Mod. Phys.* 55.3 (July 1983), pp. 601–644.

[185] V. E. Zaharov and L. D. Faddeev. "The Korteweg-de Vries equation is a fully integrable Hamiltonian system". In: *Funkcional. Anal. i Priložen.* 5.4 (1971), pp. 18–27.

[186] V. E. Zaharov and L. D. Faddeev. "The Korteweg-de Vries equation is a fully integrable Hamiltonian system". In: *Funkcional. Anal. i Priložen.* 5.4 (1971), pp. 18–27.

[187] V. E. Zakharov. "Stability of periodic waves of finite amplitude on the surface of a deep fluid". In: *J. Appl. Mech. Tech. Phys.* 9.2 (1968), pp. 190–194.

[188] V. E. Zakharov and A. B. Shabat. "Exact Theory of Two-dimensional Self-focusing and One-dimensional Self-modulation of Waves in Nonlinear Media". In: *Journal of Experimental and Theoretical Physics* 34 (1970), pp. 62–69.

[189] V. E. Zakharov and A. B. Shabat. "Interactions between solitons in a stable medium". In: *Sov. Phys. JETP* 37 (1973), pp. 823–828.

[190] P.E. Zhidkov. *Korteweg-de Vries and nonlinear Schrödinger equations: qualitative theory*. Vol. 1756. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2001, pp. vi+147.

[191] P.E. Zhidkov. "The Cauchy problem for the nonlinear Schrödinger equation". In: *Joint Inst. Nuclear Res., Dubna* (1987), p. 15.

[192] X. Zhou. "L^2-Sobolev space bijectivity of the scattering and inverse scattering transforms". In: *Comm. Pure Appl. Math.* 51.7 (1998), pp. 697–731.

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und nur unter Zuhilfenahme der ausgewiesenen Hilfsmittel angefertigt habe. Sämtliche Stellen der Arbeit, die im Wortlaut oder dem Sinn nach anderen gedruckten oder im Internet veröffentlichten Werken entnommen sind, habe ich durch genaue Quellenangaben kenntlich gemacht.

Karlsruhe, den 01.12.2025