ANALYSIS & PDE

Volume 5 No. 1 2012

mathematical sciences publishers

Analysis & PDE

msp.berkeley.edu/apde

See inside back cover or msp.berkeley.edu/apde for submission instructions.

The subscription price for 2012 is US $140/year for the electronic version, and $240/year for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA.

Analysis & PDE, at Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFLOW™ from Mathematical Sciences Publishers.

PUBLISHED BY
mathematical sciences publishers
http://msp.org/
A NON-PROFIT CORPORATION
Typeset in LaTeX

ANALYSIS AND PDE
Vol. 5, No. 1, 2012
dx.doi.org/10.2140/apde.2012.5.1

A CHARACTERIZATION OF TWO WEIGHT NORM INEQUALITIES FOR MAXIMAL SINGULAR INTEGRALS WITH ONE DOUBLING MEASURE

MICHAEL LACEY, ERIC T. SAWYER AND IGNACIO URIARTE-TUERO

Let σ and ω be positive Borel measures on $\mathbb{R}$ with σ doubling. Suppose first that $1 < p \le 2$. We characterize boundedness of certain maximal truncations of the Hilbert transform $T_\natural$ from $L^p(\sigma)$ to $L^p(\omega)$ in terms of the strengthened A_p condition

$$\Big(\int_{\mathbb{R}} s_Q(x)^p \, d\omega(x)\Big)^{1/p} \Big(\int_{\mathbb{R}} s_Q(x)^{p'} \, d\sigma(x)\Big)^{1/p'} \le C|Q|,$$

where $s_Q(x) = |Q|/(|Q| + |x - x_Q|)$, and two testing conditions. The first applies to a restricted class of functions and is a strong-type testing condition,

$$\int_Q T_\natural(\chi_E \sigma)(x)^p \, d\omega(x) \le C_1 \int_Q d\sigma(x) \quad \text{for all } E \subset Q,$$

and the second is a weak-type or dual interval testing condition,

$$\int_Q T_\natural(\chi_Q f \sigma)(x) \, d\omega(x) \le C_2 \Big(\int_Q |f(x)|^p \, d\sigma(x)\Big)^{1/p} \Big(\int_Q d\omega(x)\Big)^{1/p'}$$

for all intervals Q in $\mathbb{R}$ and all functions $f \in L^p(\sigma)$. In the case $p > 2$ the same result holds if we include an additional necessary condition, the Poisson condition

$$\int_{\mathbb{R}} \Big(\sum_{r=1}^{\infty} |I_r|_\sigma |I_r|^{p'-1} \sum_{\ell=0}^{\infty} \frac{2^{-\ell}}{|(I_r)^{(\ell)}|} |\chi_{(I_r)^{(\ell)}}(y)\Big)^p \, d\omega(y) \le C \sum_{r=1}^{\infty} |I_r|_\sigma |I_r|^{p'},$$

for all pairwise disjoint decompositions $Q = \bigcup_{r=1}^{\infty} I_r$ of the dyadic interval Q into dyadic intervals I_r. We prove that analogues of these conditions are sufficient for boundedness of certain maximal singular integrals in $\mathbb{R}^n$ when σ is doubling and $1 < p < \infty$. Finally, we characterize the weak-type two weight inequality for certain maximal singular integrals $T_\natural$ in $\mathbb{R}^n$ when $1 < p < \infty$, without the doubling assumption on σ, in terms of analogues of the second testing condition and the A_p condition.

1. Introduction

Sawyer [1984; 1982; 1988] characterized two weight inequalities for maximal functions and other positive operators, in terms of the obviously necessary conditions that the operators be uniformly bounded on a restricted class of functions, namely indicators of intervals and cubes. Thus, these characterizations have a form reminiscent of the $T1$ theorem of David and Journé.

Lacey is supported in part by the NSF, through grant DMS-0456538. Sawyer is supported in part by NSERC. Uriarte-Tuero is supported in part by the NSF, through grant DMS-0901524.
MSC2000: 42B20.
Keywords: two weight, singular integral, maximal function, maximal truncation.

Corresponding results for even the Hilbert transform have only recently been obtained [Nazarov et al. 2010; Lacey et al. 2011] and even then only for $p = 2$; evidently these are much harder to obtain. We comment in more detail on prior results below, including the innovative work of Nazarov, Treil and Volberg [1999; 2008; 2010; 2003].

Our focus is on providing characterizations of the boundedness of certain maximal truncations of a fixed operator of singular integral type. The singular integrals will be of the usual type, for example the Hilbert transform or paraproducts. Only size and smoothness conditions on the kernel are assumed; see (1-9). The characterizations are in terms of certain obviously necessary conditions, in which the class of functions being tested is simplified. For such examples, we prove unconditional characterizations of both strong-type and weak-type two weight inequalities for certain maximal truncations of the Hilbert transform, but with the additional assumption that σ is *doubling* for the strong-type inequality. A major point of our characterizations is that they hold for *all* $1 < p < \infty$. The methods in [Lacey et al. 2011] and those of Nazarov, Treil and Volberg apply only to the case $p = 2$, where the orthogonality of measure-adapted Haar bases prove critical. The doubling hypothesis on σ may not be needed in our theorems, but is required by the use of Calderón–Zygmund decompositions in our method.

As the precise statements of our general results are somewhat complicated, we illustrate them with an important case here. Let

$$Tf(x) = \lim_{\varepsilon \to 0} \int_{\mathbb{R} \setminus (-\varepsilon, \varepsilon)} \frac{1}{y} f(x - y)\, dy$$

denote the Hilbert transform, let

$$T_\flat f(x) = \sup_{0 < \varepsilon < \infty} \left| \int_{\mathbb{R} \setminus (-\varepsilon, \varepsilon)} \frac{1}{y} f(x - y)\, dy \right|$$

denote the usual maximal singular integral associated with T, and finally let

$$T_\natural f(x) = \sup_{\substack{0 < \varepsilon_1, \varepsilon_2 < \infty \\ 1/4 < \varepsilon_2 / \varepsilon_1 < 4}} \left| \int_{\mathbb{R} \setminus (-\varepsilon_1, \varepsilon_2)} \frac{1}{y} f(x - y)\, dy \right|$$

denote the new *strongly* (or *noncentered*) maximal singular integral associated with T that is defined more precisely below. Suppose σ and ω are two locally finite positive Borel measures on $\mathbb{R}$ that have no point masses in common. Then we have the following weak and strong-type characterizations, which we emphasize hold for *all* $1 < p < \infty$.

- The operator $T_\flat$ is *weak* type (p, p) with respect to (σ, ω), that is,

$$\| T_\flat (f \sigma) \|_{L^{p,\infty}(\omega)} \le C \| f \|_{L^p(\sigma)} \tag{1-1}$$

  for all f bounded with compact support *if and only if* the two weight A_p condition

$$\frac{1}{|Q|} \int_Q d\omega \Big(\frac{1}{|Q|} \int_Q d\sigma \Big)^{p-1} \le C,$$

holds for all intervals Q and the dual $T_\flat$ interval testing condition

$$\int_Q T_\flat(\chi_Q f\sigma)\,d\omega \le C\Big(\int_Q |f|^p\,d\sigma\Big)^{1/p}\Big(\int_Q d\omega\Big)^{1/p'}, \tag{1-2}$$

holds for all intervals Q and $f \in L^p_Q(\sigma)$ (part 4 of Theorem 1.8). The same is true for $T_\natural$. It is easy to see that (1-2) is equivalent to the more familiar dual interval testing condition

$$\int_Q |L^*(\chi_Q\omega)|^{p'}\,d\sigma \le C\int_Q d\omega, \tag{1-3}$$

for all intervals Q and linearizations L of the maximal singular integral $T_\flat$ (see (2-10)).

- Suppose in addition that σ is doubling and $1 < p < \infty$. Then the operator $T_\natural$ is strong-type (p, p) with respect to (σ, ω), that is,

  $$\|T_\natural(f\sigma)\|_{L^p(\omega)} \le C\|f\|_{L^p(\sigma)}$$

  for all f bounded with compact support *if and only if* these four conditions hold: (1) the strengthened A_p condition

  $$\Big(\int_Q s_Q(x)^p\,d\omega(x)\Big)^{1/p}\Big(\int_I s_Q(x)^{p'}\,d\sigma(x)\Big)^{1/p'} \le C|Q|,$$

  where $s_Q(x) = \frac{|Q|}{|Q|+|x-x_Q|}$, holds for all intervals Q; (2) the dual $T_\natural$ interval testing condition

  $$\int_Q T_\natural(\chi_Q f\sigma)\,d\omega \le C\Big(\int_Q |f|^p\,d\sigma\Big)^{1/p}\Big(\int_Q d\omega\Big)^{1/p'},$$

  holds for all intervals Q and $f \in L^p_Q(\sigma)$; (3) the forward $T_\natural$ testing condition

  $$\int_Q T_\natural(\chi_E\sigma)^p\,d\omega \le C\int_Q d\sigma, \tag{1-4}$$

  holds for all intervals Q and all compact subsets E of Q; and (4) the Poisson condition

  $$\int_{\mathbb{R}}\Big(\sum_{r=1}^{\infty}|I_r|_\sigma|I_r|^{p'-1}\sum_{\ell=0}^{\infty}\frac{2^{-\ell}}{|(I_r)^{(\ell)}|}\chi_{(I_r)^{(\ell)}}(y)\Big)^p\,d\omega(y) \le C\sum_{r=1}^{\infty}|I_r|_\sigma|I_r|^{p'},$$

  for all pairwise disjoint decompositions $Q = \bigcup_{r=1}^{\infty} I_r$ of the dyadic interval Q into dyadic intervals I_r for any fixed dyadic grid. In the case $1 < p \le 2$, only the first three conditions are needed (Theorem 1.10). Note that in (1-4) we are required to test over all compact subsets E of Q on the left side, but retain the upper bound over the (larger) cube Q on the right side.

As these results indicate, the imposition of the weight σ on both sides of (1-1) is a standard part of weighted theory, and is in general necessary for the testing conditions to be sufficient. Compare to the characterization of the two weight maximal function inequalities in Theorem 1.2 below.

Problem 1.1. In (1-4), our testing condition is more complicated than one would like, in that one must test over all compact $E \subset Q$ in (1-4). There is a corresponding feature of (1-2), seen after one unwinds the definition of the linearization L^*. We do not know if these testing conditions can be further simplified. The form of these testing conditions is dictated by our use of what we call the "maximum principle"; see Lemma 2.6.

We now recall the two weight inequalities for the maximal function as they are central to the new results of this paper. Define the maximal function

$$\mathcal{M}\nu(x) = \sup_{x \in Q} \frac{1}{|Q|} \int_Q |\nu| \quad \text{for } x \in \mathbb{R},$$

where the supremum is taken over all cubes Q (by which we mean cubes with sides parallel to the coordinate axes) containing x.

Theorem 1.2 (maximal function inequalities). *Suppose that σ and ω are positive locally finite Borel measures on $\mathbb{R}^n$, and that $1 < p < \infty$. The maximal operator $\mathcal{M}$ satisfies the two weight norm inequality* [Sawyer 1982]

$$\|\mathcal{M}(f\sigma)\|_{L^p(\omega)} \le C\|f\|_{L^p(\sigma)} \quad \textit{for } f \in L^p(\sigma), \tag{1-5}$$

if and only if *for all cubes $Q \subset \mathbb{R}^n$,*

$$\int_Q \mathcal{M}(\chi_Q\sigma)(x)^p \, d\omega(x) \le C_1 \int_Q d\sigma(x). \tag{1-6}$$

The maximal operator $\mathcal{M}$ satisfies the weak-type *two weight norm inequality* [Muckenhoupt 1972]

$$\|\mathcal{M}(f\sigma)\|_{L^{p,\infty}(\omega)} \equiv \sup_{\lambda>0} \lambda |\{\mathcal{M}(f\sigma) > \lambda\}|_\omega^{1/p} \le C\|f\|_{L^p(\sigma)} \quad \textit{for } f \in L^p(\sigma), \tag{1-7}$$

if and only if the two weight A_p condition holds for all cubes $Q \subset \mathbb{R}^n$:

$$\Bigl(\frac{1}{|Q|}\int_Q d\omega\Bigr)^{1/p}\Bigl(\frac{1}{|Q|}\int_Q d\sigma\Bigr)^{1/p'} \le C_2. \tag{1-8}$$

The necessary and sufficient condition (1-6) for the strong-type inequality (1-5) states that one need only test the strong-type inequality for functions of the form $\chi_Q\sigma$. Not only that, but the full $L^p(\omega)$ norm of $\mathcal{M}(\chi_Q\sigma)$ need not be evaluated. There is a corresponding weak-type interpretation of the A_p condition (1-8). Finally, the proofs given in [Sawyer 1982] and [Muckenhoupt 1972] for absolutely continuous weights carry over without difficulty for the locally finite measures considered here.

1.3. ***Two weight inequalities for singular integrals.*** Let us set notation for our theorems. Consider a kernel function $K(x, y)$ defined on $\mathbb{R}^n \times \mathbb{R}^n$ satisfying the size and smoothness conditions

$$\begin{aligned} &|K(x, y)| \le C|x - y|^{-n}, \\ &|K(x, y) - K(x', y)| \le C\delta\Bigl(\frac{|x - x'|}{|x - y|}\Bigr)|x - y|^{-n}, \qquad \frac{|x - x'|}{|x - y|} \le \frac{1}{2}, \end{aligned} \tag{1-9}$$

where δ is a Dini modulus of continuity, that is, a nondecreasing function on [0, 1] with $\delta(0) = 0$ and $\int_0^1 \delta(s)s^{-1}\,ds < \infty$.

Next we describe the truncations we consider. Let ζ, η be fixed smooth functions on the real line satisfying

$$\begin{gathered} \zeta(t) = 0 \quad \text{for } t \leq \tfrac{1}{2} \quad \text{and} \quad \zeta(t) = 1 \quad \text{for } t \geq 1, \\ \eta(t) = 0 \quad \text{for } t \geq 2 \quad \text{and} \quad \eta(t) = 1 \quad \text{for } t \leq 1, \\ \zeta \text{ is nondecreasing and } \eta \text{ is nonincreasing.} \end{gathered}$$

Given $0 < \varepsilon < R < \infty$, set $\zeta_\varepsilon(t) = \zeta(t/\varepsilon)$ and $\eta_R(t) = \eta(t/R)$ and define the smoothly truncated operator $T_{\varepsilon,R}$ on $L^1_{\text{loc}}(\mathbb{R}^n)$ by the absolutely convergent integrals

$$T_{\varepsilon,R} f(x) = \int K(x, y)\zeta_\varepsilon(|x - y|)\eta_R(|x - y|) f(y)\,dy \quad \text{for } f \in L^1_{\text{loc}}(\mathbb{R}^n).$$

Define the *maximal* singular integral operator $T_\flat$ on $L^1_{\text{loc}}(\mathbb{R}^n)$ by

$$T_\flat f(x) = \sup_{0<\varepsilon<R<\infty} |T_{\varepsilon,R} f(x)| \quad \text{for } x \in \mathbb{R}^n.$$

We also define a corresponding *new* notion of *strongly maximal* singular integral operator $T_\natural$ as follows. In dimension $n = 1$, we set

$$T_\natural f(x) = \sup_{\substack{0<\varepsilon_i<R<\infty \\ 1/4\leq\varepsilon_1/\varepsilon_2\leq 4}} |T_{\boldsymbol{\varepsilon},R} f(x)| \quad \text{for } x \in \mathbb{R},$$

where $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2)$ and

$$T_{\boldsymbol{\varepsilon},R} f(x) = \int K(x, y)\{\zeta_{\varepsilon_1}(x - y) + \zeta_{\varepsilon_2}(y - x)\}\eta_R(|x - y|) f(y)\,dy.$$

Thus the local singularity has been removed by a *noncentered* smooth cutoff — ε_1 to the left of x and ε_2 to the right of x, but with controlled eccentricity $\varepsilon_1/\varepsilon_2$. There is a similar definition of $T_\natural f$ in higher dimensions involving in place of $\zeta_\varepsilon(|x - y|)$, a product of smooth cutoffs,

$$\zeta_{\boldsymbol{\varepsilon}}(x - y) \equiv 1 - \prod_{k=1}^{n}(1 - \{\zeta_{\varepsilon_{2k-1}}(x_k - y_k) + \zeta_{\varepsilon_{2k}}(y_k - x_k)\}),$$

satisfying $1/4 \leq \varepsilon_{2k-1}/\varepsilon_{2k} \leq 4$ for $1 \leq k \leq n$. The advantage of this larger operator $T_\natural$ is that in many cases boundedness of $T_\natural$ (or collections thereof) implies boundedness of the maximal operator $\mathcal{M}$. Our method of proving boundedness of $T_\flat$ and $T_\natural$ requires boundedness of the maximal operator $\mathcal{M}$ anyway, and as a result we can in some cases give necessary and sufficient conditions for strong boundedness of $T_\natural$. As for weak-type boundedness, we can in many more cases give necessary and sufficient conditions for weak boundedness of the usual truncations $T_\flat$.

Definition 1.4. We say that T is a *standard singular integral operator with kernel* K if T is a bounded linear operator on $L^q(\mathbb{R}^n)$ for some fixed $1 < q < \infty$, that is

$$\|Tf\|_{L^q(\mathbb{R}^n)} \le C\|f\|_{L^q(\mathbb{R}^n)} \quad \text{for } f \in L^q(\mathbb{R}^n), \tag{1-10}$$

if $K(x, y)$ is defined on $\mathbb{R}^n \times \mathbb{R}^n$ and satisfies both (1-9) and the Hörmander condition,

$$\int_{B(y,2\varepsilon)^c} |K(x, y) - K(x, y')|\, dx \le C \quad \text{for } y' \in B(y, \varepsilon), \varepsilon > 0, \tag{1-11}$$

and finally if T and K are related by

$$Tf(x) = \int K(x, y) f(y)\, dy \quad \text{for a.e.-}x \notin \operatorname{supp} f, \tag{1-12}$$

whenever $f \in L^q(\mathbb{R}^n)$ has compact support in $\mathbb{R}^n$. We call a kernel $K(x, y)$ *standard* if it satisfies (1-9) and (1-11).

For standard singular integral operators, we have this classical result. (See the appendix on truncation of singular integrals on [Stein 1993, page 30] for the case $R = \infty$; the case $R < \infty$ is similar.)

Theorem 1.5. *Suppose that* T *is a standard singular integral operator. Then the map* $f \to T_\flat f$ *is of weak type* $(1, 1)$, *and bounded on* $L^p(\mathbb{R})$ *for* $1 < p < \infty$. *There exist sequences* $\varepsilon_j \to 0$ *and* $R_j \to \infty$ *such that for* $f \in L^p(\mathbb{R})$ *with* $1 \le p < \infty$,

$$\lim_{j\to\infty} T_{\varepsilon_j,R_j} f(x) \equiv T_{0,\infty} f(x)$$

exists for a.e. $x \in \mathbb{R}$. *Moreover, there is a bounded measurable function* $a(x)$ (*depending on the sequences*) *satisfying*

$$Tf(x) = T_{0,\infty} f(x) + a(x) f(x) \quad \text{for } x \in \mathbb{R}^n.$$

We state a conjecture, so that the overarching goals of this subject are clear.

Conjecture 1.6. *Suppose that* σ *and* ω *are positive Borel measures on* $\mathbb{R}^n$, *let* $1 < p < \infty$, *and suppose* T *is a standard singular integral operator on* $\mathbb{R}^n$. *Then the following two statements are equivalent*:

$$\int |T(f\sigma)|^p \omega \le C \int |f|^p \sigma \quad \text{for } f \in C_0^\infty,$$

$$\left.\begin{aligned} \Big(\frac{1}{|Q|}\int_Q d\omega\Big)^{1/p}\Big(\frac{1}{|Q|}\int_Q d\sigma\Big)^{1/p'} &\le C, \\ \int_Q |T\chi_Q\sigma|^p \le C' \int_Q \sigma&, \\ \int_Q |T^*\chi_Q\omega|^{p'}\sigma \le C'' \int_Q \omega&, \end{aligned}\right\} \quad \text{for all cubes } Q.$$

Remark 1.7. The first of the three testing conditions above is the two-weight A_p condition. We expect that this condition can be strengthened to a "Poisson two-weight A_p condition". See [Nazarov et al. 2010; Volberg 2003].

The most important instances of this conjecture occur when T is one of a few canonical singular integral operators, such as the Hilbert transform, the Beurling transform, or the Riesz transforms. This question occurs in different instances, such as the Sarason conjecture concerning the composition of Hankel operators, or the semicommutator of Toeplitz operators [Cruz-Uribe et al. 2007; Zheng 1996], mathematical physics [Peherstorfer et al. 2007], as well as perturbation theory of some self-adjoint operators. See references in [Volberg 2003].

To date, this has only been verified for positive operators, such as Poisson integrals and fractional integral operators [Sawyer 1984; 1982; 1988]. Recently the authors have used the methods of Nazarov, Treil and Volberg to prove a special case of the conjecture for the Hilbert transform when $p = 2$ and an energy hypothesis is assumed [Lacey et al. 2011]. Earlier in [2010] Nazarov, Treil and Volberg used a stronger pivotal condition in place of the energy hypothesis, but neither of these conditions are necessary [Lacey et al. 2011]. The two weight Helson–Szegö theorem was proved many years earlier by Cotlar and Sadosky [1979; 1983]; thus the L^2 case for the Hilbert transform is completely settled.

Nazarov, Treil and Volberg [1999; 2010] have characterized those weights for which the class of Haar multipliers is bounded when $p = 2$. They also have a result for an important special class of singular integral operators, the "well-localized" operators of [2008]. Citing the specific result here would carry us too far afield, but this class includes the important Haar shift examples, such as the one found by S. Petermichl [2000], and generalized in [2002]. Consequently, characterizations are given in [Volberg 2003] and [Nazarov et al. 2010] for the Hilbert transform and Riesz transforms in weighted L^2 spaces under various additional hypotheses. In particular they obtain an analogue of the case $p = 2$ of the strong-type theorem below. Our results can be reformulated in the context there, a theme we do not pursue further here.

We now characterize the weak-type two weight norm inequality for both maximal singular integrals and strongly maximal singular integrals.

Theorem 1.8 (maximal singular integral weak-type inequalities). *Suppose that σ and ω are positive locally finite Borel measures on $\mathbb{R}^n$, let $1 < p < \infty$, and let $T_\flat$ and $T_\natural$ be the maximal singular integral operators as above with kernel $K(x, y)$ satisfying* (1-9).

(1) *Suppose that the maximal operator $\mathcal{M}$ satisfies* (1-7). *Then $T_\natural$ satisfies the weak-type two weight norm inequality*
$$\|T_\natural(f\sigma)\|_{L^{p,\infty}(\omega)} \leq C\|f\|_{L^p(\sigma)} \quad \text{for } f \in L^p(\sigma), \tag{1-13}$$
if and only if
$$\int_Q T_\natural(\chi_Q f\sigma)(x)\, d\omega(x) \leq C_2\Big(\int_Q |f(x)|^p\, d\sigma(x)\Big)^{1/p}\Big(\int_Q d\omega(x)\Big)^{1/p'}, \tag{1-14}$$
for all cubes $Q \subset \mathbb{R}^n$ and all functions $f \in L^p(\sigma)$.

(2) *The same characterization as above holds for $T_\flat$ in place of $T_\natural$ everywhere.*

(3) *Suppose that σ and ω are absolutely continuous with respect to Lebesgue measure, that the maximal operator $\mathcal{M}$ satisfies* (1-7), *and that T is a standard singular integral operator with kernel K as*

above. If (1-13) *holds for* $T_\natural$ *or* $T_\flat$, *then it also holds for* T:

$$\|T(f\sigma)\|_{L^{p,\infty}(\omega)} \le C\|f\|_{L^p(\sigma)} \quad \text{for } f \in L^p(\sigma),\ f\sigma \in L^\infty \text{ with } \operatorname{supp} f\sigma \text{ compact}. \tag{1-15}$$

(4) *Suppose* $c > 0$ *and that* $\{K_j\}_{j=1}^J$ *is a collection of* standard *kernels such that for* each *unit vector* $\boldsymbol{u}$ *there is* j *satisfying*

$$|K_j(x, x + t\boldsymbol{u})| \ge ct^{-n} \quad \text{for } t \in \mathbb{R}. \tag{1-16}$$

Suppose also that σ *and* ω *have no common point masses*, *that is*, $\sigma(\{x\}) \cdot \omega(\{x\}) = 0$ *for all* $x \in \mathbb{R}^n$. *Then*

$$\|(T_j)_\flat(f\sigma)\|_{L^{p,\infty}(\omega)} \le C\|f\|_{L^p(\sigma)} \quad \text{for } f \in L^p(\sigma), \text{ with } 1 \le j \le J,$$

if and only if the two weight A_p *condition* (1-8) *holds and*

$$\int_Q (T_j)_\flat(\chi_Q f\sigma)(x)\, d\omega(x) \le C_2\Big(\int_Q |f(x)|^p\, d\sigma(x)\Big)^{1/p}\Big(\int_Q d\omega(x)\Big)^{1/p'}, \\ f \in L^p(\sigma),\ \text{cubes } Q \subset \mathbb{R}^n,\ 1 \le j \le J.$$

While in (1)–(3), we assume that the maximal function inequality holds, in point (4), we obtain an *unconditional* characterization of the weak-type inequality for a large class of families of (centered) maximal singular integral operators $T_\flat$. This class includes the individual maximal Hilbert transform in one dimension, the individual maximal Beurling transform in two dimensions, and the families of maximal Riesz transforms in higher dimensions; see Lemma 2.11.

Note that in (1) above, there is only size and smoothness assumptions placed on the kernel, so that it could for instance be a degenerate fractional integral operator, and therefore unbounded on $L^2(dx)$. But, the characterization still has content in this case, if ω and σ are not of full dimension.

In (3), we deduce a two weight inequality for standard singular integrals T without truncations when the measures are absolutely continuous. The proof of this is easy. From (1-13) and the pointwise inequality $T_{0,\infty} f\sigma(x) \le T_\flat f\sigma(x) \le T_\natural f\sigma(x)$, we obtain that for any limiting operator $T_{0,\infty}$ the map $f \to T_{0,\infty} f\sigma$ is bounded from $L^p(\sigma)$ to $L^{p,\infty}(\omega)$. By (1-7) $f \to \mathcal{M} f\sigma$ is bounded; hence $f \to f\sigma$ is bounded, and so Theorem 1.5 shows that $f \to Tf\sigma = T_{0,\infty} f\sigma + af\sigma$ is also bounded, provided we initially restrict attention to functions f for which $f\sigma$ is bounded with compact support.

The characterizing condition (1-14) is a weak-type condition, with the restriction that one only needs to test the weak-type condition for functions supported on a given cube, and test the weak-type norm over that given cube. It also has an interpretation as a dual inequality $\int_Q |L^*(\chi_Q\omega)|^{p'}\, d\sigma \le C_2 \int_Q d\omega$, which we return to below; see (2-10) and (2-11).

We now consider the two weight norm inequality for a strongly maximal singular integral $T_\natural$, but assuming that the measure σ is doubling.

Theorem 1.9 (maximal singular integral strong-type inequalities). *Suppose that* σ *and* ω *are positive locally finite Borel measures on* $\mathbb{R}^n$ *with* σ *doubling*, *let* $1 < p < \infty$, *and let* $T_\flat$ *and* $T_\natural$ *be the maximal singular integral operators as above with kernel* $K(x, y)$ *satisfying* (1-9).

(1) *Suppose that the maximal operator* $\mathcal{M}$ *satisfies* (1-5) *and also the "dual" inequality*

$$\|\mathcal{M}(g\omega)\|_{L^{p'}(\sigma)} \le C\|g\|_{L^{p'}(\omega)} \quad \text{for } g \in L^{p'}(\omega). \tag{1-17}$$

Then $T_\natural$ *satisfies the two weight norm inequality*

$$\int_{\mathbb{R}^n} T_\natural(f\sigma)(x)^p \, d\omega(x) \le C \int_{\mathbb{R}^n} |f(x)|^p \, d\sigma(x), \tag{1-18}$$

for all $f \in L^p(\sigma)$ *that are bounded with compact support in* $\mathbb{R}^n$, *if and only if both the dual cube testing condition* (1-14) *and the condition*

$$\int_Q T_\natural(\chi_Q g\sigma)(x)^p \, d\omega(x) \le C_1 \int_Q d\sigma(x), \tag{1-19}$$

holds for all cubes $Q \subset \mathbb{R}^n$ *and all functions* $|g| \le 1$.

(2) *The same characterization as above holds for* $T_\flat$ *in place of* $T_\natural$ *everywhere. In fact*

$$|T_\natural f\sigma(x) - T_\flat f\sigma(x)| \le C\mathcal{M}(f\sigma)(x).$$

(3) *Suppose that* σ *and* ω *are absolutely continuous with respect to Lebesgue measure, that the maximal operator* $\mathcal{M}$ *satisfies* (1-5), *and that* T *is a standard singular integral operator. If* (1-18) *holds for* $T_\natural$ *or* $T_\flat$, *then it also holds for* T:

$$\int_{\mathbb{R}^n} |T(f\sigma)(x)|^p \, d\omega(x) \le C \int_{\mathbb{R}^n} |f(x)|^p \, d\sigma(x) \quad \text{for } f \in L^p(\sigma),\ f\sigma \in L^\infty, \text{ with } \operatorname{supp}(f\sigma) \text{ compact}.$$

(4) *Suppose that* $\{K_j\}_{j=1}^n$ *is a collection of standard kernels satisfying for some* $c > 0$,

$$\pm \operatorname{Re} K_j(x,y) \ge \frac{c}{|x-y|^n} \quad \text{for } \pm(y_j - x_j) \ge \tfrac{1}{4}|x-y|, \tag{1-20}$$

where $x = (x_j)_{1 \le j \le n}$. *If both* ω *and* σ *are doubling, then* (1-18) *holds for* $(T_j)_\natural$ *and* $(T_j^*)_\natural$ *for all* $1 \le j \le n$ *if and only if both* (1-19) *and* (1-14) *hold for* $(T_j)_\natural$ *and* $(T_j^*)_\natural$ *for all* $1 \le j \le n$.

Note that the second condition (1-19) is a stronger condition than we would like: it is the L^p inequality, applied to functions *bounded by* 1 and supported on a cube Q, but with the $L^p(\sigma)$ norm of $\mathbf{1}_Q$ on the right side. It is easy to see that the bounded function g in (1-19) can be replaced by χ_E for every compact subset E of Q. Indeed if L ranges over all linearizations of $T_\natural$, then with

$$g_{h,Q,L} = L^*(\chi_Q h\omega)/|L^*(\chi_Q h\omega)|$$

we have

$$\begin{aligned}\sup_{|g|\le 1}\int_Q T_\natural(\chi_Q g\sigma)^p\omega &= \sup_{|g|\le 1}\sup_L \sup_{\|h\|_{L^{p'}(\omega)}\le 1}\left|\int_Q L(\chi_Q g\sigma)h\omega\right| \\ &= \sup_L \sup_{\|h\|_{L^{p'}(\omega)}\le 1}\sup_{|g|\le 1}\left|\int_Q L^*(\chi_Q h\omega)g\sigma\right| \\ &= \sup_L \sup_{\|h\|_{L^{p'}(\omega)}\le 1}\int_Q L^*(\chi_Q h\omega)g_{h,Q,L}\sigma \\ &= \sup_{\|h\|_{L^{p'}(\omega)}\le 1}\sup_L \int_Q L(\chi_Q g_{h,Q,L})h\omega\sigma \\ &\le \sup_{\|h\|_{L^{p'}(\omega)}\le 1}\sup_L \int_Q T_\natural(\chi_Q g_{h,Q,L}\sigma)^p\omega.\end{aligned}$$

Since $g_{h,Q,L}$ takes on only the values ± 1, it is easy to see that we can take $g=\chi_E$. Point (3) is again easy, just as in the previous weak-type theorem.

And in (4), we note that the truncations, in the way that we formulate them, dominate the maximal function, so that our assumption on $\mathcal{M}$ in (1)–(3) is not unreasonable. The main result of [Nazarov et al. 2010] assumes $p=2$ and that T is the Hilbert transform, and makes similar kinds of assumptions. In fact it is essentially the same as our result in the case $p=2$, but without doubling on σ and only for T and not $T_\flat$ or $T_\natural$. Finally, we observe that by our definition of the truncation $T_\natural$, we obtain in point (4) a characterization for doubling measures of the strong-type inequality for appropriate families of standard singular integrals and their adjoints, including the Hilbert and Riesz transforms; see Lemma 2.12.

We don't know if the bounded function g in condition (1-19) can be replaced by the constant function 1.

We now give a characterization of the strong-type weighted norm inequality for the *individual* strongly maximal Hilbert transform $T_\natural$ when $1<p<\infty$ and the measure σ is *doubling*. If $p>2$ we use an extra necessary condition (see (1-24)) that involves a "dyadic" Poisson function $\sum_{\ell=0}^{\infty}(2^{-\ell}/|I^{(\ell)}|)\chi_{I^{(\ell)}}(y)$, where I is a dyadic interval and $I^{(\ell)}$ denotes its ℓ-th ancestor in the dyadic grid, that is, the unique dyadic interval containing I with $|I^{(\ell)}|=2^\ell|I|$. This condition is a variant of the pivotal condition of Nazarov, Treil and Volberg in [2010]; when $1<p\le 2$ it is a consequence of the A_p condition (1-8).

Theorem 1.10. *Suppose that σ and ω are positive locally finite Borel measures on $\mathbb{R}$ with σ* doubling, *let $1<p<\infty$, and let $T_\natural$ be the strongly maximal Hilbert transform. Then $T_\natural$ is* strong *type (p,p) with respect to (σ,ω), that is,*

$$\|T_\natural(f\sigma)\|_{L^p(\omega)}\le C\|f\|_{L^p(\sigma)},$$

for all f bounded with compact support if and only if the following four conditions hold. In the case $1<p\le 2$, the fourth condition (1-24) *is implied by the A_p condition* (1-8), *and so in this case we only need the first three conditions below*:

(1) *The dual $T_\natural$ interval testing condition*

$$\int_Q T_\natural(\chi_Q f\sigma)\,d\omega \le C\Big(\int_Q |f|^p\,d\sigma\Big)^{1/p}\Big(\int_Q d\omega\Big)^{1/p'} \tag{1-21}$$

holds for all intervals Q and $f \in L^p_Q(\sigma)$.

(2) *The forward $T_\natural$ testing condition*

$$\int_Q T_\natural(\chi_E\sigma)^p\,d\omega \le C\int_Q d\sigma \tag{1-22}$$

holds for all intervals Q and all compact subsets E of Q.

(3) *The strengthened A_p condition*

$$\Big(\int_{\mathbb{R}}\Big(\frac{|Q|}{|Q|+|x-x_Q|}\Big)^p\,d\omega(x)\Big)^{1/p}\Big(\int_{\mathbb{R}}\Big(\frac{|Q|}{|Q|+|x-x_Q|}\Big)^{p'}\,d\sigma(x)\Big)^{1/p'} \le C|Q| \tag{1-23}$$

holds for all intervals Q.

(4) *The Poisson condition*

$$\int_{\mathbb{R}}\Big(\sum_{r=1}^{\infty}|I_r|_\sigma|I_r|^{p'-1}\sum_{\ell=0}^{\infty}\frac{2^{-\ell}}{|(I_r)^{(\ell)}|}\chi_{(I_r)^{(\ell)}}(y)\Big)^p\,d\omega(y) \le C\sum_{r=1}^{\infty}|I_r|_\sigma|I_r|^{p'} \tag{1-24}$$

holds for all pairwise disjoint decompositions $Q=\bigcup_{r=1}^{\infty} I_r$ of the dyadic interval Q into dyadic intervals I_r, for any fixed dyadic grid.

Remark 1.11. The strengthened A_p condition (1-23) can be replaced with the weaker "half" condition where the first factor on the left is replaced by $(\int_Q d\omega)^{1/p}$. We do not know if the first three conditions suffice when $p > 2$.

2. Overview of the proofs and general principles

If Q is a cube, then $\ell(Q)$ is its side length, $|Q|$ is its Lebesgue measure and for a positive Borel measure ν, $|Q|_\nu = \int_Q d\nu$ is its ν-measure.

2.1. ***Calderón–Zygmund decompositions.*** Our starting place is the argument in [Sawyer 1988] used to prove a two weight norm inequality for fractional integral operators on Euclidean space. Of course the fractional integral is a positive operator with a monotone kernel, properties we do not have in the current setting.

A central tool arises from the observation that for any positive Borel measure μ, one has the boundedness of a maximal function associated with μ. Define the dyadic μ-maximal operator $\mathcal{M}^{dy}_\mu$ by

$$\mathcal{M}^{dy}_\mu f(x) = \sup_{\substack{Q\in\mathcal{D}\\ x\in Q}}\frac{1}{|Q|_\mu}\int_Q |f|\mu, \tag{2-1}$$

with the supremum taken over all dyadic cubes $Q \in \mathcal{D}$ containing x. It is immediate to check that $\mathcal{M}_\mu^{dy}$ satisfies the weak-type (1, 1) inequality, and the $L^\infty(\mu)$ bound is obvious. Hence we have

$$\int (\mathcal{M}_\mu^{dy} f)^p \mu \le C \int f^p \mu \quad \text{for } f \ge 0 \text{ on } \mathbb{R}^n. \tag{2-2}$$

This observation places certain Calderón–Zygmund decompositions at our disposal. Exploitation of this brings in the testing condition (1-19) involving the bounded function g on a cube Q, and indeed, g turns out to be the "good" function in a Calderón–Zygmund decomposition of f on Q. The associated "bad" function requires the dual testing condition (1-14) as well.

2.2. ***Edge effects of dyadic grids.*** Our operators are not dyadic operators, nor — in contrast to the fractional integral operators — can they be easily obtained from dyadic operators. This leads to the necessity of considering for instance triples of dyadic cubes, which are not dyadic.

Also, dyadic grids distinguish points by for instance making some points on the boundary of many cubes. As our measures are arbitrary, they could conspire to assign extra mass to some of these points. To address this point, Nazarov, Treil and Volberg [2010; 2003; 1997] use a random shift of the grid.

A random approach would likely work for us as well, though the argument would be different from those in the cited papers above. Instead, we will use a nonrandom technique of shifted dyadic grid from [Muscalu et al. 2002], which goes back to P. Jones and J. Garnett. Define a *shifted dyadic grid* to be the collection of cubes

$$\mathcal{D}^\alpha = \left\{2^j(k + [0,1)^n + (-1)^j \alpha) : j \in \mathbb{Z}, k \in \mathbb{Z}^n\right\}, \quad \text{where } \alpha \in \{0, \tfrac{1}{3}, \tfrac{2}{3}\}^n. \tag{2-3}$$

The basic properties of these collections are these: In the first place, each $\mathcal{D}^\alpha$ is a grid, that is, for $Q, Q' \in \mathcal{D}^\alpha$ we have $Q \cap Q' \in \{\varnothing, Q, Q'\}$ and Q is a union of 2^n elements of $\mathcal{D}^\alpha$ of equal volume. In the second place (and this is the novel property for us), for any cube $Q \subset \mathbb{R}^n$ there is a choice of some α and some $Q' \in \mathcal{D}_\alpha$ such that $Q \subset (9/10)Q'$ and $|Q'| \le C|Q|$.

We define the analogues of the dyadic maximal operator in (2-1), namely

$$\mathcal{M}_\mu^\alpha f(x) = \sup_{\substack{Q \in \mathcal{D}^\alpha \\ x \in Q}} \frac{1}{|Q|_\mu} \int_Q |f| \mu. \tag{2-4}$$

These operators clearly satisfy (2-2). Shifted dyadic grids will return in Section 4.5.

2.3. ***A maximum principle.*** A second central tool is a "maximum principle" (or good λ inequality) that will permit one to localize large values of a singular integral, provided the maximal function is bounded. It is convenient for us to describe this in conjunction with another fundamental tool of this paper, a family of Whitney decompositions.

We begin with the Whitney decompositions. Fix a finite measure ν with compact support on $\mathbb{R}^n$ and for $k \in \mathbb{Z}$, let

$$\Omega_k = \{x \in \mathbb{R}^n : T_\natural \nu(x) > 2^k\}. \tag{2-5}$$

Note that $\Omega_k \neq \mathbb{R}^n$ has compact closure for such ν. Fix an integer $N \geq 3$. We can choose $R_W \geq 3$ sufficiently large, depending only on the dimension and N, such that there is a collection of cubes $\{Q_j^k\}_j$ that satisfy the following properties:

$$\begin{aligned}
&\text{(disjoint cover)} && \Omega_k = \textstyle\bigcup_j Q_j^k \ \text{ and } \ Q_j^k \cap Q_i^k = \varnothing \text{ if } i \neq j, \\
&\text{(Whitney condition)} && R_W Q_j^k \subset \Omega_k \ \text{ and } \ 3R_W Q_j^k \cap \Omega_k^c \neq \varnothing \ \text{ for all } k, j, \\
&\text{(bounded overlap)} && \textstyle\sum_j \chi_{NQ_j^k} \leq C\chi_{\Omega_k} \ \text{ for all } k, \\
&\text{(crowd control)} && \#\{Q_s^k : Q_s^k \cap NQ_j^k \neq \varnothing\} \leq C \ \text{ for all } k, j, \\
&\text{(nested property)} && Q_j^k \subsetneq Q_i^\ell \ \text{ implies } \ k > \ell.
\end{aligned} \tag{2-6}$$

Indeed, one should choose the $\{Q_j^k\}_j$ satisfying the Whitney condition, and then show that the other properties hold. The different combinatorial properties above are fundamental to the proof. And alternate Whitney decompositions are constructed in Section 4.9.1 below.

Remark 2.4. Our use of the Whitney decomposition and the maximum principle are derived from the two weight fractional integral argument of Sawyer; see [1988, Section 2]. In particular, the properties above are as Sawyer's, aside from the crowd control property above, which is $N = 3$ there.

Remark 2.5. In our notation for the Whitney cubes, the superscript indicates a "height" and the subscript an arbitrary enumeration of the cubes. We will use super- and subscripts below in this manner consistently throughout the paper. It is important to note that a fixed cube Q can arise in *many* Whitney decompositions: There are integers $K_-(Q) \leq K_+(Q)$ with $Q = Q_{j(k)}^k$ for some choice of $j(k)$ for all $K_-(Q) \leq k \leq K_+(Q)$. (The last point follows from the nested property.) There is no *a priori* upper bound on $K_+(Q) - K_-(Q)$.

Lemma 2.6 (maximum principle). *Let ν be a finite (signed) measure with compact support. For any cube Q_j^k as above, we have the pointwise inequality*

$$\sup_{x \in Q_j^k} T_\natural(\chi_{(3Q_j^k)^c}\nu)(x) \leq 2^k + C\boldsymbol{P}(Q_j^k, \nu) \leq 2^k + CM(Q_j^k, \nu), \tag{2-7}$$

where $\boldsymbol{P}(Q, \nu)$ and $M(Q, \nu)$ are defined by

$$\begin{aligned}
\boldsymbol{P}(Q, \nu) &\equiv \frac{1}{|Q|}\int_Q d|\nu| + \sum_{\ell=0}^{\infty} \frac{\delta(2^{-\ell})}{|2^{\ell+1}Q|}\int_{2^{\ell+1}Q \setminus 2^\ell Q} d|\nu|, \\
M(Q, \nu) &\equiv \sup_{Q' \supset Q} \frac{1}{|Q'|}\int_{Q'} d|\nu|.
\end{aligned} \tag{2-8}$$

The bound in terms of $\boldsymbol{P}(Q, \nu)$ should be regarded as one in terms of a modified Poisson integral. It is both slightly sharper than that of $M(Q, \nu)$, and a linear expression in $|\nu|$, a fact will be used in the proof of the strong-type estimates.

Proof. To see this, take $x \in Q_j^k$ and note that for each $\eta > 0$ there is $\boldsymbol{\varepsilon}$ with $\ell(Q_j^k) < \max_{1\le j\le n} \varepsilon_j < R < \infty$ and $\theta \in [0, 2\pi)$ such that

$$T_\natural(\chi_{(3Q_j^k)^c}\nu)(x) \le (1+\eta)\Big|\int_{(3Q_j^k)^c} K(x,y)\zeta_\varepsilon(x-y)\eta_R(x-y)d\nu(y)\Big|$$
$$= (1+\eta)e^{i\theta}T_{\boldsymbol{\varepsilon},R}(\chi_{(3Q_j^k)^c}\nu)(x).$$

For convenience we take $\eta = 0$ in the sequel. By the Whitney condition in (2-6), there is a point $z \in 3R_W Q_j^k \cap \Omega_k^c$ and it now follows that (remember that $\ell(Q_j^k) < \max_{1\le j\le n} \varepsilon_j$)

$$\begin{aligned}
&|T_{\boldsymbol{\varepsilon},R}(\chi_{(3Q_j^k)^c}\nu)(x) - T_{\boldsymbol{\varepsilon},R}\nu(z)| \\
&\quad\le C\frac{1}{|6R_W Q_j^k|}\int_{6R_W Q_j^k} d|\nu| + |T_{\varepsilon,R}(\chi_{(6R_W Q_j^k)^c}\nu)(x) - T_{\boldsymbol{\varepsilon},R}(\chi_{(6R_W Q_j^k)^c}\nu)(z)| \\
&\quad= C\frac{1}{|6R_W Q_j^k|}\int_{6R_W Q_j^k} d|\nu| \\
&\qquad\quad + \int_{(6R_W Q_j^k)^c} |K(x,y)\zeta_\varepsilon(x-y)\eta_R(x-y) - K(z,y)\zeta_{\boldsymbol{\varepsilon}}(z-y)\eta_R(z-y)|d|\nu|(y) \\
&\quad\le C\frac{1}{|6R_W Q_j^k|}\int_{6R_W Q_j^k} d|\nu| + C\int_{(6R_W Q_j^k)^c} \delta\Big(\frac{|x-z|}{|x-y|}\Big)\frac{1}{|x-y|^n}d|\nu|(y) \\
&\quad\le C\boldsymbol{P}(Q_j^k, \nu).
\end{aligned}$$

Thus

$$T_\natural(\chi_{(3Q_j^k)^c}\nu)(x) \le |T_\natural\nu(z)| + C\boldsymbol{P}(Q_j^k,\nu) \le 2^k + C\boldsymbol{P}(Q_j^k,\nu),$$

which yields (2-7) since $\boldsymbol{P}(Q,\nu) \le CM(Q,\nu)$. □

2.7. ***Linearizations.*** We now make comments on the linearizations of our maximal singular integral operators. We would like, at different points, to treat $T_\natural$ as a linear operator, which of course it is not. Nevertheless $T_\natural$ is a pointwise supremum of the linear truncation operators $T_{\varepsilon,R}$, and as such, the supremum can be linearized with measurable selection of the parameters $\boldsymbol{\varepsilon}$ and R, as was just done in the previous proof. We make this a definition.

Definition 2.8. We say that L is a linearization of $T_\natural$ if there are measurable functions $\boldsymbol{\varepsilon}(x) \in (0,\infty)^n$ and $R(x) \in (0,\infty)$ with $1/4 \le \varepsilon_i/\varepsilon_j \le 4$, $\max_{1\le i\le n} \varepsilon_i < R(x) < \infty$ and $\theta(x) \in [0, 2\pi)$ such that

$$Lf(x) = e^{i\theta(x)}T_{\boldsymbol{\varepsilon}(x),R(x)}f(x) \quad \text{for } x \in \mathbb{R}^n. \tag{2-9}$$

For fixed f and $\delta > 0$, we can always choose a linearization L so that $T_\natural f(x) \le (1+\delta)Lf(x)$ for all x. In a typical application of this lemma, one takes δ to be one.

Note that condition (1-19) is obtained from inequality (1-18) by testing over f of the form $f = \chi_Q g$ with $|g| \le 1$, and then restricting integration on the left to Q. By passing to linearizations L, we can "dualize" (1-14) to the testing conditions

$$\int_Q |L^*(\chi_Q \omega)(x)|^{p'}\, d\sigma(x) \le C_2 \int_Q d\omega(x), \tag{2-10}$$

or equivalently (note that in (1-19) the presence of g makes a difference, but not here),

$$\int_Q |L^*(\chi_Q g\omega)(x)|^{p'}\, d\sigma(x) \le C_2 \int_Q d\omega(x) \quad \text{for } |g| \le 1, \tag{2-11}$$

with the requirement that these inequalities hold *uniformly* in all linearizations L of $T_\natural$.

While the smooth truncation operators $T_{\boldsymbol{\varepsilon},R}$ are essentially self-adjoint, the dual of a linearization L is generally complicated. Nevertheless, the dual L^* does satisfy one important property, which plays a crucial role in the proof of Theorem 1.9, the L^p-norm inequalities.

Lemma 2.9. *$L^*\mu$ is δ-Hölder continuous (where δ is the Dini modulus of continuity of the kernel K) with constant $C\boldsymbol{P}(Q,\mu)$ on any cube Q satisfying $\int_{3Q} d|\mu| = 0$, that is,*

$$|L^*\mu(y) - L^*\mu(y')| \le C\boldsymbol{P}(Q,\mu)\delta\Big(\frac{|y-y'|}{\ell(Q)}\Big) \quad \textit{for } y, y' \in Q. \tag{2-12}$$

Here, recall the definition (2-8) *and that $\boldsymbol{P}(Q,\mu) \le CM(Q,\mu)$.*

Proof. Suppose L is as in (2-9). Then for any finite measure ν,

$$L\nu(x) = e^{i\theta(x)} \int \zeta_{\boldsymbol{\varepsilon}(x)}(x-y)\eta_{R(x)}(x-y)K(x,y)d\nu(y).$$

Fubini's theorem shows that the dual operator L^* is given on a finite measure μ by

$$L^*\mu(y) = \int \zeta_{\boldsymbol{\varepsilon}(x)}(x-y)\eta_{R(x)}(x-y)K(x,y)e^{i\theta(x)}d\mu(x). \tag{2-13}$$

For $y, y' \in Q$ and $|\mu|(3Q) = 0$, we thus have

$$\begin{aligned} L^*\mu(y) - L^*\mu(y') = \int \{(\zeta_{\boldsymbol{\varepsilon}(x)}\eta_{R(x)})(x-y) - (\zeta_{\boldsymbol{\varepsilon}(x)}\eta_{R(x)})(x-y')\}K(x,y)e^{i\theta(x)}d\mu(x) \\ + \int (\zeta_{\boldsymbol{\varepsilon}(x)}\eta_{R(x)})(x-y')(K(x,y)-K(x,y'))e^{i\theta(x)}d\mu(x), \end{aligned}$$

from which (2-12) follows easily if we split the two integrals in x over dyadic annuli centered at the center of Q. □

2.10. *Control of maximal functions.* Next we record the facts that T and $T_\natural$ control $\mathcal{M}$ for many (sets of) standard singular integrals T, including the Hilbert transform, the Beurling transform and the sets of Riesz transforms in higher dimensions.

Lemma 2.11. *Suppose that σ and ω have no point masses in common, and that $\{K_j\}_{j=1}^J$ is a collection of* standard *kernels satisfying* (1-9) *and* (1-16). *If the corresponding operators T_j given by* (1-12) *satisfy*

$$\|\chi_E T_j(f\sigma)\|_{L^{p,\infty}(\omega)} \le C\|f\|_{L^p(\sigma)} \quad \textit{where} \quad E = \mathbb{R}^n \setminus \operatorname{supp} f,$$

for $1 \le j \le J$, *then the two weight* A_p *condition* (1-8) *holds, and hence also the weak-type two weight inequality* (1-7).

Proof. Part of the "one weight" argument of [Stein and Shakarchi 2005, page 21] yields the *asymmetric* two weight A_p condition

$$|Q|_\omega |Q'|_\sigma^{p-1} \le C|Q|^p, \tag{2-14}$$

where Q and Q' are cubes of equal side length r and distance approximately $C_0 r$ apart for some fixed large positive constant C_0 (for this argument we choose the unit vector $\boldsymbol{u}$ in (1-16) to point in the direction from the center of Q to the center of Q', and then with j as in (1-16), C_0 is chosen large enough by (1-9) that (1-16) holds for all unit vectors $\boldsymbol{u}$ pointing from a point in Q to a point in Q'). In the one weight case treated in [Stein and Shakarchi 2005], it is easy to obtain from this (even for a *single* direction $\boldsymbol{u}$) the usual (symmetric) A_p condition (1-8). Here we will instead use our assumption that σ and ω have no point masses in common for this purpose.

So fix an open dyadic cube Q_0 in $\mathbb{R}^n$, say with side length 1, let $\mathsf{Q}_0 = Q_0 \times Q_0$ and set

$$\Omega = \{\mathsf{Q} = Q \times Q' \text{ dyadic} : \mathsf{Q} \subset \mathsf{Q}_0 \text{ and (2-14) holds for } Q \text{ and } Q'\}.$$

Note that with $\mathsf{Q} = Q \times Q'$, inequality (2-14) can be written

$$\mathcal{A}_p(\omega, \sigma; \mathsf{Q}) \le C|\mathsf{Q}|^{p/2}, \tag{2-15}$$

where

$$\mathcal{A}_p(\omega, \sigma; \mathsf{Q}) = |Q|_\omega |Q'|_\sigma^{p-1}.$$

Here $\mathcal{A}_2(\omega, \sigma; \mathsf{Q}) = |\mathsf{Q}|_{\omega\times\sigma}$, where $\omega \times \sigma$ denotes product measure on $\mathbb{R}^n \times \mathbb{R}^n$. For $1 < p < \infty$ we easily see that if $\mathsf{Q}_0 = \bigcup_\alpha \mathsf{Q}_\alpha$ is a pairwise disjoint union of cubes Q_α, then the Lebesgue measures satisfy

$$\sum_\alpha |\mathsf{Q}_\alpha|^{p/2} \le C|Q_0 \times Q_0|^{p/2} = C|Q_0|^p.$$

Suppose first that $1 < p \le 2$. Divide Q_0 into $2^n \times 2^n = 4^n$ congruent subcubes $\mathsf{Q}_0^1, \dots, \mathsf{Q}_0^{4^n}$ of side length $\frac{1}{2}$, and set aside those $\mathsf{Q}_0^j \in \Omega$ (those for which (2-14) holds) into a collection of *stopping cubes* Γ. Continue to divide the remaining Q_0^j into 4^n congruent subcubes $\mathsf{Q}_0^{j,1}, \dots, \mathsf{Q}_0^{j,4^n}$ of side length $\frac{1}{4}$, and again, set aside those $\mathsf{Q}_0^{j,i} \in \Omega$ into Γ, and continue subdividing those that remain. We continue with such subdivisions for N generations so that all the cubes *not* set aside into Γ have side length 2^{-N}. The important property these cubes have is that they all lie within distance $r2^{-N}$ of the diagonal $\mathcal{D} = \{(x, x) : (x, x) \in \mathsf{Q}_0\}$ in $\mathsf{Q}_0 = Q_0 \times Q_0$ since (2-14) holds for all pairs of cubes Q and Q' of equal side length r having distance approximately $C_0 r$ apart. Enumerate the cubes in Γ as $\{\mathsf{Q}_\alpha\}_\alpha$ and those remaining that are not in Γ as $\{\mathsf{P}_\beta\}_\beta$. Thus we have the pairwise disjoint decomposition

$$\mathsf{Q}_0 = \Big(\bigcup_\alpha \mathsf{Q}_\alpha\Big) \cup \Big(\bigcup_\beta \mathsf{Q}_\beta\Big).$$

In the case $p=2$, the countable additivity of the product measure $\omega\times\sigma$ shows that

$$\mathscr{A}_2(\omega,\sigma;\mathsf{Q}_0)=\sum_\alpha \mathscr{A}_2(\omega,\sigma;\mathsf{Q}_\alpha)+\sum_\beta \mathscr{A}_2(\omega,\sigma;\mathsf{P}_\beta).$$

For the more general case $1<p\le 2$, note that at each division described above, we have using $0<p-1\le 1$

$$\mathscr{A}_p(\omega,\sigma;\mathsf{Q}_0)=\Big(\sum_{i=1}^{2^n}|\mathsf{Q}_0^j|_\omega\Big)\Big(\sum_{i=1}^{2^n}|\mathsf{Q}_0^j|_\sigma\Big)^{p-1}\le\Big(\sum_{i=1}^{2^n}|\mathsf{Q}_0^j|_\omega\Big)\Big(\sum_{i=1}^{2^n}|\mathsf{Q}_0^j|_\sigma^{p-1}\Big)=\sum_{j=1}^{4^n}\mathscr{A}_p(\omega,\sigma;\mathsf{Q}_0^j),$$

$$\mathscr{A}_p(\omega,\sigma;\mathsf{Q}_0^j)\le\sum_{i=1}^{4^n}\mathscr{A}_p(\omega,\sigma;\mathsf{Q}_0^{j,i})\quad\text{for }\mathsf{Q}_0^j\notin\Gamma,$$

and so on. It follows that

$$\begin{aligned}\mathscr{A}_p(\omega,\sigma;\mathsf{Q}_0)&\le\sum_\alpha\mathscr{A}_p(\omega,\sigma;\mathsf{Q}_\alpha)+\sum_\beta\mathscr{A}_p(\omega,\sigma;\mathsf{P}_\beta)\\&\le C\sum_\alpha|\mathsf{Q}_\alpha|^{p/2}+\sum_\beta\mathscr{A}_p(\omega,\sigma;\mathsf{P}_\beta)\le C|Q_0|^p+\sum_\beta\mathscr{A}_p(\omega,\sigma;\mathsf{P}_\beta).\end{aligned}$$

Since ω and σ have no point masses in common, it is not hard to show, using that the side length of $\mathsf{P}_\beta=P_\beta\times P'_\beta$ is 2^{-N} and $\operatorname{dist}(\mathsf{P}_\beta,\mathfrak{D})\le C2^{-N}$, that we have the limit

$$\sum_\beta\mathscr{A}_p(\omega,\sigma;\mathsf{P}_\beta)\to 0\quad\text{as }N\to\infty.$$

Indeed, if σ has no point masses at all, then

$$\begin{aligned}\sum_\beta\mathscr{A}_p(\omega,\sigma;\mathsf{P}_\beta)&=\sum_\beta|P_\beta|_\omega|P'_\beta|_\sigma^{p-1}\\&\le\Big(\sum_\beta|P_\beta|_\omega\Big)\sup_\beta|P'_\beta|_\sigma^{p-1}\le C|Q_0|_\omega\sup_\beta|P'_\beta|_\sigma^{p-1}\to 0\quad\text{as }N\to\infty.\end{aligned}$$

If σ contains a point mass $c\delta_x$, then

$$\sum_{\beta:x\in P'_\beta}\mathscr{A}_p(\omega,\sigma;\mathsf{P}_\beta)\le\Big(\sum_{\beta:x\in P'_\beta}|P_\beta|_\omega\Big)\sup_{\beta:x\in P'_\beta}|P'_\beta|_\sigma^{p-1}\le C\Big(\sum_{\beta:x\in P'_\beta}|P_\beta|_\omega\Big)\to 0\quad\text{as }N\to\infty$$

since ω has no point mass at x. The argument in the general case is technical, but involves no new ideas, and we leave it to the reader. We thus conclude that

$$\mathscr{A}_p(\omega,\sigma;\mathsf{Q}_0)\le C|Q_0|^p,$$

which is (1-8). The case $2\le p<\infty$ is proved in the same way using that (2-14) can be written

$$\mathscr{A}_{p'}(\sigma,\omega;\mathsf{Q}_\alpha)\le C'|\mathsf{Q}_\alpha|^{p'/2}.$$

□

Lemma 2.12. *If* $\{T_j\}_{j=1}^n$ *satisfies* (1-20), *then*

$$\mathcal{M}\nu(x) \le C \sum_{j=1}^n (T_j)_\natural \nu(x) \quad \textit{for } x \in \mathbb{R}^n, \textit{ with } \nu \ge 0 \textit{ a finite measure with compact support.}$$

Proof. We prove the case $n = 1$, the general case being similar. Then with $T = T_1$ and $r > 0$ we have

$$\begin{aligned} \operatorname{Re}(T_{r,r/4,100r}\nu(x) - T_{r,4r,100r}\nu(x)) &= \int (\zeta_{r/4}(y-x) - \zeta_{4r}(y-x)) \operatorname{Re} K(x,y) d\nu(y) \\ &\ge \frac{c}{r} \int_{[x+r/2, x+2r]} d\nu(y). \end{aligned}$$

Thus

$$T_\natural \nu(x) \ge \max\{|T_{r,r/4,100r}\nu(x)|, |T_{r,4r,100r}\nu(x)|\} \ge \frac{c}{r} \int_{[x+r/2,x+2r]} d\nu(y),$$

and similarly

$$T_\natural \nu(x) \ge \frac{c}{r} \int_{[x-2r,x-r/2]} d\nu(y).$$

It follows that

$$\begin{aligned} \mathcal{M}\nu(x) &\le \sup_{r>0} \frac{1}{4r} \int_{[x-2r,x+2r]} d\nu(y) \\ &= \sup_{r>0} \sum_{k=0}^\infty 2^{-k} \frac{1}{2^{2-k}r} \int_{[x-2^{1-k}r, x-2^{-1-k}r] \cup [x+2^{-1-k}r, x+2^{1-k}r]} d\nu(y) \le C T_\natural \nu(x). \end{aligned}$$ □

Finally, we will use the following covering lemma of Besicovitch type for multiples of dyadic cubes (the case of triples of dyadic cubes arises in (4-50) below).

Lemma 2.13. *Let M be an odd positive integer, and suppose that Φ is a collection of cubes P with bounded diameters and having the form $P = MQ$, where Q is dyadic* (*a product of clopen dyadic intervals*). *If Φ^* is the collection of* maximal *cubes in Φ, that is, $P^* \in \Phi^*$ provided there is no strictly larger P in Φ that contains P^*, then the cubes in Φ^* have finite overlap at most M^n.*

Proof. Let $Q_0 = [0,1)^n$ and assign labels $1, 2, 3, \ldots, M^n$ to the dyadic subcubes of side length one of MQ_0. We say that the subcube labeled k is of type k, and we extend this definition by translation and dilation to the subcubes of MQ having side length that of Q. Now we simply observe that if $\{P_i^*\}_i$ is a set of cubes in Φ^* containing the point x, then for a given k, there is at most one P_i^* that contains x in its subcube of type k. The reason is that if P_j^* is another such cube and $\ell(P_j^*) \le \ell(P_i^*)$, we must have $P_j^* \subset P_i^*$ (draw a picture in the plane for example). □

2.14. ***Preliminary precaution.*** Given a positive locally finite Borel measure μ on $\mathbb{R}^n$, there exists a rotation such that all boundaries of rotated dyadic cubes have μ-measure zero (see [Mateu et al. 2000] where they actually prove a stronger assertion when μ has no point masses, but our conclusion is obvious for a sum of point mass measures). We will assume that such a rotation has been made so that all boundaries of rotated dyadic cubes have $(\omega+\sigma)$-measure zero, where ω and σ are the positive Borel measures appearing in the theorems above (of course σ doubling implies that σ cannot contain any point masses, but this argument works as well for general σ as in the weak type theorem). While this assumption is not essential for the proof, it relieves the reader of having to consider the possibility that boundaries of dyadic cubes have positive measure at each step of the argument below.

Recall also (see for example [Rudin 1987, Theorem 2.18]) that any positive locally finite Borel measure on $\mathbb{R}^n$ is both inner and outer regular.

3. The proof of Theorem 1.8: Weak-type inequalities

We begin with the necessity of condition (1-14):

$$\begin{aligned}\int_Q T_\natural(\chi_Q f\sigma)\omega &= \int_0^\infty \min\bigl\{|Q|_\omega, |\{T_\natural(\chi_Q f\sigma)>\lambda\}|_\omega\bigr\}d\lambda \\ &\le \Bigl(\int_0^A + \int_A^\infty\Bigr)\min\Bigl\{|Q|_\omega, C\lambda^{-p}\int |f|^p\,d\sigma\Bigr\}d\lambda \\ &\le A|Q|_\omega + CA^{1-p}\int |f|^p\,d\sigma = (C+1)|Q|_\omega^{1/p'}\Bigl(\int |f|^p\,d\sigma\Bigr)^{1/p},\end{aligned}$$

if we choose $A = (\int |f|^p\,d\sigma/|Q|_\omega)^{1/p}$.

Now we turn to proving (1-13), assuming both (1-14) and (1-7), and moreover that f is bounded with compact support. We will prove the quantitative estimate

$$\|T_\natural f\sigma\|_{L^{p,\infty}(\omega)} \le C\{\mathfrak{A}+\mathfrak{T}_*\}\|f\|_{L^p(\sigma)}, \tag{3-1}$$

$$\mathfrak{A} = \sup_Q \sup_{\|f\|_{L^p(\sigma)}=1} \sup_{\lambda>0} \lambda|\{\mathcal{M}(f\sigma)>\lambda\}|_\omega^{1/p}, \tag{3-2}$$

$$\mathfrak{T}_* = \sup_{\|f\|_{L^p(\sigma)}=1} \sup_Q |Q|_\omega^{-1/p'}\int_Q T_\natural(\chi_Q f\sigma)(x)\,d\omega(x). \tag{3-3}$$

We should emphasize that the term (3-2) is comparable to the two weight A_p condition (1-8).

Standard considerations [Sawyer 1984, Section 2] show that it suffices to prove the following good-λ inequality: There is a positive constant C such that for $\beta>0$ sufficiently small, and provided

$$\sup_{0<\lambda<\Lambda} \lambda^p|\{x\in\mathbb{R}^n : T_\natural f\sigma(x)>\lambda\}|_\omega < \infty \quad \text{for } \Lambda<\infty, \tag{3-4}$$

we have this inequality:

$$\begin{aligned}&|\{x\in\mathbb{R}^n : T_\natural f\sigma(x)>2\lambda \text{ and } \mathcal{M}f\sigma(x)\le\beta\lambda\}|_\omega \\ &\qquad\le C\beta\mathfrak{T}_*^p|\{x\in\mathbb{R}^n : T_\natural f\sigma(x)>\lambda\}|_\omega + C\beta^{-p}\lambda^{-p}\int|f|^p\,d\sigma.\end{aligned} \tag{3-5}$$

Our presumption (3-4) holds due to the A_p condition (1-8) and the fact that

$$\{x\in\mathbb{R}^n : T_\natural f\sigma(x)>\lambda\} \subset B(0, c\lambda^{-1/n}) \quad \text{for } \lambda>0 \text{ small},$$

Hence it is enough to prove (3-5).

To prove (3-5) we choose $\lambda = 2^k$ and apply the decomposition in (2-6). In this argument, we can take k to be fixed, so that we can suppress its appearance as a superscript in this section. (When we come to L^p estimates, we will not have this luxury.)

Define

$$E_j = \{x \in Q_j : T_\natural f\sigma(x) > 2\lambda \text{ and } \mathcal{M} f\sigma(x) \le \beta\lambda\}.$$

Then for $x \in E_j$, we can apply Lemma 2.6 to deduce

$$T_\natural(\chi_{(3Q_j)^c} f\sigma)(x) \le (1 + C\beta)\lambda. \tag{3-6}$$

If we take $\beta > 0$ so small that $1 + C\beta \le \frac{3}{2}$, then (3-6) implies that for $x \in E_j$

$$2\lambda < T_\natural f\sigma(x) \le T_\natural \chi_{3Q_j} f\sigma(x) + T_\natural \chi_{(3Q_j)^c} f\sigma(x) \le T_\natural \chi_{3Q_j^k} f\sigma(x) + \tfrac{3}{2}\lambda.$$

Integrating this inequality with respect to ω over E_j we obtain

$$\lambda |E_j|_\omega \le 2 \int_{E_j} (T_\natural \chi_{3Q_j} f\sigma)\omega. \tag{3-7}$$

The disjoint cover condition in (2-6) shows that the sets E_j are disjoint, and this suggests we should sum their ω-measures. We split this sum into two parts, according to the size of $|E_j|_\omega / |3Q_j|_\omega$. The left side of (3-5) satisfies

$$\sum_j |E_j|_\omega \le \beta \sum_{j : |E_j|_\omega \le \beta |3Q_j|_\omega} |3Q_j|_\omega + \beta^{-p} \sum_{j : |E_j|_\omega > \beta |3Q_j|_\omega} |E_j|_\omega \Big(\frac{2}{\lambda} \frac{1}{|3Q_j|_\omega} \int_{E_j} (T_\natural \chi_{3Q_j} f\sigma)\omega \Big)^p.$$

Call the added pieces of this I and II. Now

$$I \le \beta \sum_j |3Q_j^k|_\omega \le C\beta |\Omega|_\omega,$$

by the finite overlap condition in (2-6). From (1-14) with $Q = 3Q_j$ we have

$$\begin{aligned} II &\le \Big(\frac{2}{\beta\lambda}\Big)^p \sum_j |E_j|_\omega \Big(\frac{1}{|3Q_j|_\omega} \int_{E_j^k} (T_\natural \chi_{3Q_j} f\sigma)\omega \Big)^p \\ &\le C \Big(\frac{2}{\beta\lambda}\Big)^p \mathfrak{T}_*^p \sum_j |E_j|_\omega \frac{1}{|3Q_j|_\omega^p} |3Q_j|_\omega^{p-1} \int_{3Q_j} |f|^p \, d\sigma \\ &\le C \Big(\frac{2}{\beta\lambda}\Big)^p \mathfrak{T}_*^p \int \Big(\sum_j \chi_{3Q_j^k} \Big) |f|^p \, d\sigma \le C \Big(\frac{2}{\beta\lambda}\Big)^p \mathfrak{T}_*^p \int |f|^p \, d\sigma, \end{aligned}$$

by the finite overlap condition in (2-6) again. This completes the proof of the good-λ inequality (3-5).

The proof of assertion 2 regarding $T_\flat$ is similar. Assertion 3 was discussed earlier and assertion 4 follows readily from assertion 2 and Lemma 2.11. □

4. The proof of Theorem 1.9: Strong-type inequalities

Since conditions (1-19) and (1-14) are obviously necessary for (1-18), we turn to proving the weighted inequality (1-18) for the strongly maximal singular integral $T_\natural$.

4.1. *The quantitative estimate.* In particular, we will prove

$$\|T_\natural f\sigma\|_{L^p(\omega)} \le C\big(\mathfrak{M}+\gamma^2\mathfrak{M}_*+\gamma^2\mathfrak{T}+\mathfrak{T}_*\big)\|f\|_{L^p(\sigma)}, \tag{4-1}$$

$$\mathfrak{M} = \sup_{\|f\|_{L^p(\sigma)}=1} \|\mathcal{M}(f\sigma)\|_{L^p(\omega)}, \tag{4-2}$$

$$\mathfrak{M}_* = \sup_{\|g\|_{L^{p'}(\omega)}=1} \|\mathcal{M}(g\omega)\|_{L^{p'}(\sigma)}, \tag{4-3}$$

$$\mathfrak{T} = \sup_{Q} \sup_{\|f\|_{L^\infty}\le 1} |Q|_\sigma^{-1/p}\|\chi_Q T_\natural(\chi_Q f\sigma)\|_{L^p(\omega)}, \tag{4-4}$$

$$\mathfrak{T}_* = \sup_{\|f\|_{L^p(\sigma)}=1} \sup_{Q}|Q|_\omega^{-1/p'}\int_Q T_\natural(\chi_Q f\sigma)(x)\,d\omega(x), \tag{4-5}$$

where $\gamma \ge 2$ is a doubling constant for the measure σ; see (4-19) below. Note that γ appears only in conjunction with $\mathfrak{T}$ and $\mathfrak{M}_*$. The norm estimates on the maximal function (4-2) and (4-3) are equivalent to the testing conditions in (1-6) and its dual formulation. The term $\mathfrak{T}_*$ also appeared in (3-3).

4.2. *The initial construction.* We suppose that both (1-19) and (1-14) hold, that is, (4-4) and (4-5) are finite, and that f is bounded with compact support on $\mathbb{R}^n$. Moreover, in the case (1-20) holds, we see that (1-19) (the finiteness of (4-4)) implies (1-6) by Lemma 2.12, and so by Theorem 1.2 we may also assume that the maximal operator $\mathcal{M}$ satisfies the two weight norm inequality (1-5). It now follows that $\int (T_\natural f\sigma)^p\omega < \infty$ for f bounded with compact support. Indeed, $T_\natural f\sigma \le C\mathcal{M}f\sigma$ far away from the support of f, while $T_\natural f\sigma$ is controlled by the finiteness of the testing condition (4-4) near the support of f.

Let $\{Q_j^k\}$ be the cubes as in (2-5) and (2-6), with the measure ν that appears in there being $\nu = f\sigma$. We will use Lemma 2.6 with this choice of ν as well. Now define an "exceptional set" associated to Q_j^k to be

$$E_j^k = Q_j^k \cap (\Omega_{k+1}\setminus\Omega_{k+2}).$$

See Figure 4.1. One might anticipate the definition of the exceptional set to be more simply $Q_j^k\cap\Omega_{k+1}$. We are guided to this choice by the work on fractional integrals [Sawyer 1988]. And indeed, the choice of exceptional set above enters in a decisive way in the analysis of the bad function at the end of the proof.

We estimate the left side of (1-18) in terms of this family of dyadic cubes $\{Q_j^k\}_{k,j}$ by

$$\begin{aligned}\int (T_\natural f\sigma)^p\omega(dx) &\le \sum_{k\in\mathbb{Z}}(2^{k+2})^p|\Omega_{k+1}\setminus\Omega_{k+2}|_\omega \\ &\le \sum_{k,j}(2^{k+2})^p|E_j^k|_\omega.\end{aligned} \tag{4-6}$$

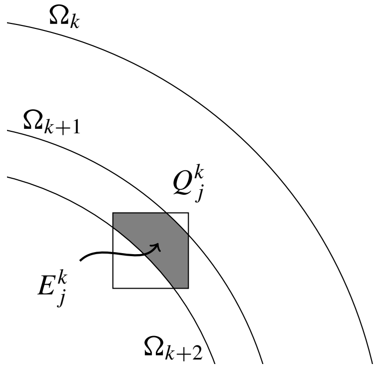

Figure 4.1. The set $E_j^k(Q)$.

Choose a linearization L of $T_\natural$ as in (2-9) so that (recall $R(x)$ is the upper limit of truncation)

$$R(x) \le \tfrac{1}{2}\ell(Q_j^k) \quad \text{for } x \in E_j^k, \tag{4-7}$$

$$\text{and} \quad T_\natural(\chi_{3Q_j^k} f\sigma)(x) \le 2L(\chi_{3Q_j^k} f\sigma)(x) + C\frac{1}{|3Q_j^k|}\int_{3Q_j^k} |f|\sigma \quad \text{for } x \in E_j^k.$$

For $x \in E_j^k$, the maximum principle (2-7) yields

$$T_\natural \chi_{3Q_j^k} f\sigma(x) \ge T_\natural f\sigma(x) - T_\natural \chi_{(3Q_j^k)^c} f\sigma(x) > 2^{k+1} - 2^k - C\boldsymbol{P}(Q_j^k, f\sigma) = 2^k - C\boldsymbol{P}(Q_j^k, f\sigma).$$

From (4-7) we conclude that

$$L\chi_{3Q_j^k} f\sigma(x) \ge 2^{k-1} - C\boldsymbol{P}(Q_j^k, f\sigma).$$

Thus either $2^k \le 4\inf_{E_j^k} L\chi_{3Q_j^k} f\sigma$ or $2^k \le 4C\boldsymbol{P}(Q_j^k, f\sigma) \le 4CM(Q_j^k, f\sigma)$. So we obtain either

$$|E_j^k|_\omega \le C2^{-k}\int_{E_j^k} (L\chi_{3Q_j^k} f\sigma)\omega(dx), \tag{4-8}$$

or

$$|E_j^k|_\omega \le C2^{-pk}|E_j^k|_\omega M(Q_j^k, f\sigma)^p \le C2^{-pk}\int_{E_j^k} (\mathcal{M} f\sigma)^p \omega(dx). \tag{4-9}$$

Now consider the following decomposition of the set of indices (k, j):

$$\begin{aligned} \mathbb{E} &= \{(k,j) : |E_j^k|_\omega \le \beta|NQ_j^k|_\omega\},\\ \mathbb{F} &= \{(k,j) : \text{(4-9) holds}\},\\ \mathbb{G} &= \{(k,j) : |E_j^k|_\omega > \beta|NQ_j^k|_\omega \text{ and (4-8) holds}\}, \end{aligned} \tag{4-10}$$

where $0 < \beta < 1$ will be chosen sufficiently small at the end of the argument. (It will be of the order of c^p for a small constant c.) By the "bounded overlap" condition of (2-6), we have

$$\sum_j \chi_{NQ_j^k} \le C \quad \text{for } k \in \mathbb{Z}. \tag{4-11}$$

We then have the corresponding decomposition:

$$\begin{aligned}
\int (T_\natural f\sigma)^p \omega &\le \Big(\sum_{(k,j)\in\mathbb{E}} + \sum_{(k,j)\in\mathbb{F}} + \sum_{(k,j)\in\mathbb{G}} \Big) (2^{k+2})^p |E_j^k|_\omega \qquad (4\text{-}12)\\
&\le \beta \sum_{(k,j)\in\mathbb{E}} (2^{k+2})^p |NQ_j^k|_\omega + C \sum_{(k,j)\in\mathbb{F}} \int_{E_j^k} (\mathcal{M} f\sigma)^p \omega \\
&\qquad + C \sum_{(k,j)\in\mathbb{G}} |E_j^k|_\omega \Big(\frac{1}{\beta |NQ_j^k|_\omega} \int_{E_j^k} (L\chi_{3Q_j^k} f\sigma)\omega \Big)^p \\
&= J(1) + J(2) + J(3) \\
&\le C_0 \Big(\beta \int (T_\natural f\sigma)^p \omega + \beta^{-p} \int |f|^p \sigma \Big), \qquad (4\text{-}13)
\end{aligned}$$

where $C_0 \le C(\mathfrak{M} + \gamma^2 \mathfrak{M}_* + \gamma^2 \mathfrak{T} + \mathfrak{T}_*)^p$. The last line is the claim that we take up in the remainder of the proof. Once it is proved, note that if we take $0 < C_0\beta < \frac{1}{2}$ and use the fact that $\int (T_\natural f\sigma)^p \omega < \infty$ for f bounded with compact support, we have proved assertion (1) of Theorem 1.9, and in particular (4-1).

The proof of the strong-type inequality requires a complicated series of decompositions of the dominating sums, which are illustrated for the reader's convenience as a schematic tree in Figure 4.2.

4.3. ***Two easy estimates.*** Note that the first term $J(1)$ in (4-12) satisfies

$$J(1) = \beta \sum_{(k,j)\in\mathbb{E}} (2^{k+2})^p |NQ_j^k|_\omega \le C\beta \int (T_\natural f\sigma)^p \omega,$$

by the finite overlap condition (4-11). The second term $J(2)$ is dominated by

$$C \sum_{(k,j)\in\mathbb{F}} \int_{E_j^k} (\mathcal{M} f\sigma)^p \omega \le C \mathfrak{M}^p \|f\|_{L^p(\sigma)}^p,$$

by our assumption (1-5). It is useful to note that this is the *only* time in the proof that we use the maximal function inequality (1-5) — from now on we use the *dual* maximal function inequality (1-17).

Remark 4.4. In the arguments below we can use [Sawyer 1988, Theorem 2] to replace the dual maximal function assumption $\mathfrak{M}_* < \infty$ with two assumptions, namely a "Poisson two weight A_p condition" and the analogue of the dual pivotal condition of Nazarov, Treil and Volberg [2010]. The Poisson two weight A_p condition is in fact necessary for the two weight inequality, but the pivotal conditions are *not* necessary for the Hilbert transform two weight inequality [Lacey et al. 2011]. On the other hand, the assumption $\mathfrak{M} < \infty$ cannot be weakened here, reflecting that our method requires the maximum principle in Lemma 2.6.

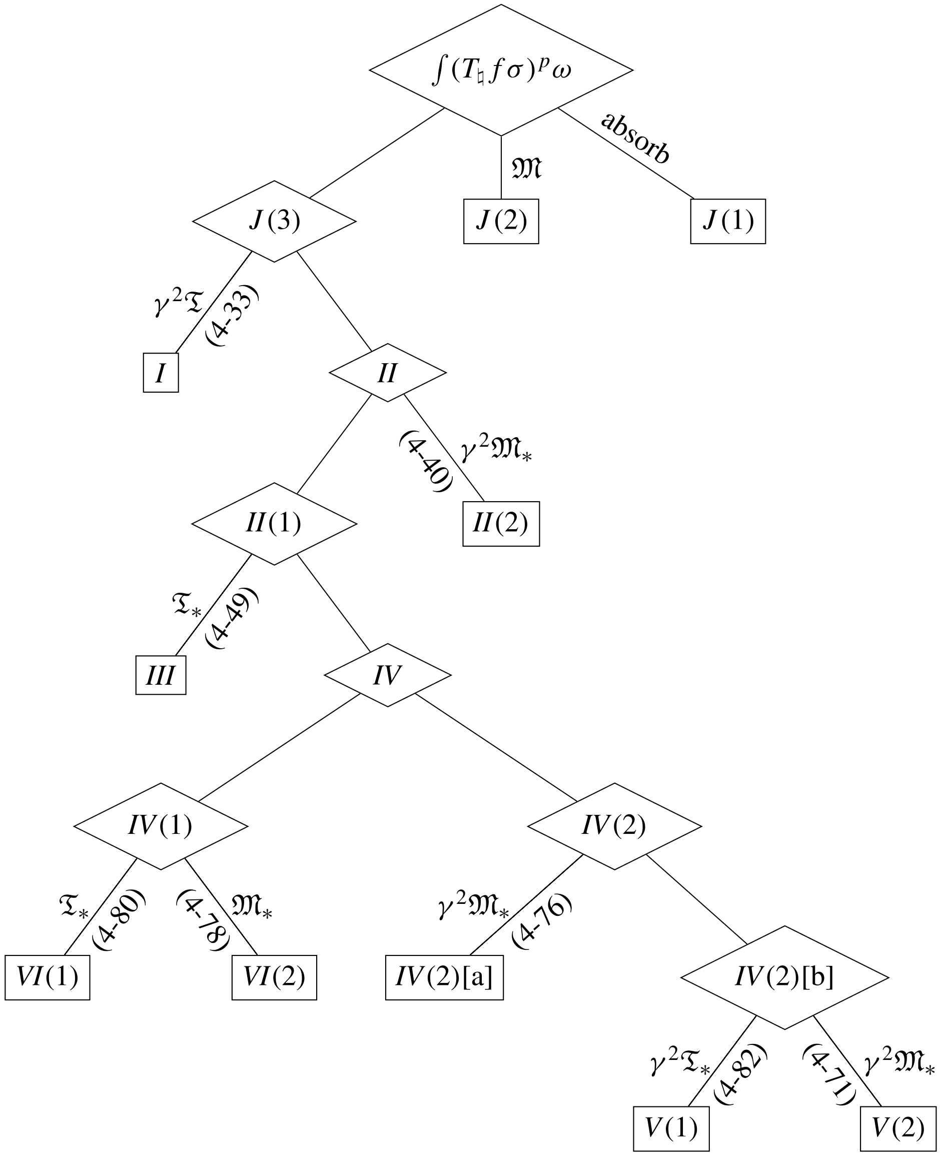

Figure 4.2. This is a schematic tree of how the integral $\int (T_\natural f\sigma)^p \omega$ has been, and will continue to be, decomposed. We have suppressed superscripts, subscripts and sums in the tree. Terms in diamonds are further decomposed, while terms in rectangles are final estimates. The edges leading into rectangles are labeled by $\mathfrak{M}$, $\mathfrak{M}_*$, $\mathfrak{T}$ or $\mathfrak{T}_*$ whose finiteness is used to control that term. Those terms controlled by the doubling constant γ are also indicated. Equation references are to where the final estimates on the term is obtained. The word "absorb" leading into $J(1)$ indicates that this term is a small multiple of $\int (T_\natural f\sigma)^p \omega$ and can be absorbed into the left-hand side of the inequality. As most of the terms involve the maximal theorem (Equation (2-2)), we do not indicate its use in the schematic tree.

It is the third term $J(3)$ that is the most involved; see Figure 4.2. The remainder of the proof is taken up with the proof of

$$\sum_{(k,j)\in\mathbb{G}} R_j^k \Big| \int_{E_j^k} (L\chi_{3Q_j^k} f\sigma)\omega \Big|^p \le C\{\gamma^{2p}\mathfrak{M}_*^p + \gamma^{2p}\mathfrak{T}^p + \mathfrak{T}_*^p\} \|f\|_{L^p\sigma(dx)}^p, \tag{4-14}$$

where

$$R_j^k = \frac{|E_j^k|_\omega}{|NQ_j^k|_\omega^p}. \tag{4-15}$$

Once this is done, the proof of (4-12) is complete, and the proof of assertion (1) is finished.

4.5. *The Calderón–Zygmund decompositions.* To carry out this proof, we make Calderón–Zygmund decompositions relative to the measure σ. These decompositions will be done at *all heights simultaneously*. We will use the shifted dyadic grids; see (2-3). Suppose that $\gamma \ge 2$ is a doubling constant for the measure σ:

$$|3Q|_\sigma \le \gamma |Q|_\sigma \quad \text{for all cubes } Q. \tag{4-16}$$

For $\alpha \in \{0, \frac{1}{3}, \frac{2}{3}\}^n$, let

$$\mathcal{M}_\sigma^\alpha f(x) = \sup_{x\in Q\in\mathcal{D}^\alpha} \frac{1}{|Q|_\sigma} \int_Q |f|\, d\sigma,$$
$$\Gamma_t^\alpha = \{x \in \mathbb{R} : \mathcal{M}_\sigma^\alpha f(x) > \gamma^t\} = \bigcup_s G_s^{\alpha,t}, \tag{4-17}$$

where $\{G_s^{\alpha,t}\}_{(t,s)\in\mathbb{L}^\alpha}$ are the maximal $\mathcal{D}^\alpha$ cubes in Γ_t^α, and $\mathbb{L}^\alpha$ is the set of pairs we use to label the cubes. This implies that we have the nested property: If $G_s^{\alpha,t} \subsetneqq G_{s'}^{\alpha,t'}$ then $t > t'$. Moreover, if $t > t'$ there is some s' with $G_s^{\alpha,t} \subset G_{s'}^{\alpha,t'}$. These are the cubes used to make a Calderón–Zygmund decomposition at height γ^t for the grid $\mathcal{D}^\alpha$ with respect to the measure σ. We will refer to the cubes $\{G_s^{\alpha,t}\}_{(t,s)\in\mathbb{L}^\alpha}$ as *principal cubes*.

Of course we have from the maximal inequality in (2-2)

$$\sum_{(t,s)\in\mathbb{L}^\alpha} \gamma^{pt} |G_s^{\alpha,t}|_\sigma \le C\|f\|_{L^p(\sigma)}^p. \tag{4-18}$$

The point of these next several definitions is to associate to each dyadic cube Q, a good shifted dyadic grid, and an appropriate height, at which we will build our Calderón–Zygmund decomposition.

We now use a consequence of the doubling condition (4-16) for the measure σ, that

$$|P(G)|_\sigma \le \gamma |G|_\sigma \quad \text{for } G \in \mathcal{D}^\alpha. \tag{4-19}$$

The average $|G_s^{\alpha,t}|_\sigma^{-1} \int_{G_s^{\alpha,t}} |f|\, d\sigma$ is thus at most γ^{t+1} by (4-19) and the maximality of the cubes in (4-17):

$$\gamma^t < \frac{1}{|G_s^{\alpha,t}|_\sigma} \int_{G_s^{\alpha,t1}} |f|\, d\sigma \le \frac{|P(G_s^{\alpha,t})|_\sigma}{|G_s^{\alpha,t}|_\sigma} \frac{1}{|P(G_s^{\alpha,t})|_\sigma} \int_{P(G_s^{\alpha,t})} |f|\, d\sigma \le \gamma\gamma^t = \gamma^{t+1}. \tag{4-20}$$

Select a shifted grid: Let $\vec{\alpha} : \mathcal{D} \longrightarrow \{0, \frac{1}{3}, \frac{2}{3}\}^n$ be a map such that for $Q \in \mathcal{D}$, there is a $\hat{Q} \in \mathcal{D}^{\vec{\alpha}(Q)}$ such that $3Q \subset \hat{Q}$ and $|\hat{Q}| \le C|Q|$. Here, C is an appropriate constant depending only on dimension. Thus, $\vec{\alpha}(Q)$ picks a "good" shifted dyadic grid for Q. Moreover we will assume that $\hat{Q}$ is the smallest such cube. Note that we are discarding the extra requirement that $3Q \subset \frac{9}{10}\hat{Q}$ since this property will not be used. Also we have

$$\hat{Q} \subset MQ, \tag{4-21}$$

for some positive dimensional constant M. The cubes $\hat{Q}_j^k$ will play a critical role below. See Figure 4.3

Select a principal cube: Define $\mathcal{A}(Q)$ to be the smallest cube from the collection $\{G_s^{\vec{\alpha}(Q),t} \mid (t, s) \in \mathbb{L}^\alpha\}$ that contains $3Q$; such $\mathcal{A}(Q)$ is uniquely determined by Q and the choice of function $\vec{\alpha}$. Define

$$\mathbb{H}_s^{\alpha,t} = \{(k, j) : \mathcal{A}(Q_j^k) = G_s^{\alpha,t}\} \quad \text{for } (s, t) \in \mathbb{L}^\alpha. \tag{4-22}$$

This is an important definition for us. The combinatorial structure this places on the corresponding cubes is essential for this proof to work. Note that $3Q_j^k \subset \hat{Q}_j^k \subset \mathcal{A}(Q_j^k)$.

Parents: For any of the shifted dyadic grids $\mathcal{D}^\alpha$, a $Q \in \mathcal{D}^\alpha$ has a unique parent denoted as $P(Q)$, the smallest member of $\mathcal{D}^\alpha$ that strictly contains Q. We suppress the dependence upon α here.

Indices: Let

$$\mathcal{K}_s^{\alpha,t} = \{r \mid G_r^{\alpha,t+1} \subset G_s^{\alpha,t}\}. \tag{4-23}$$

We use a calligraphic font $\mathcal{K}$ for sets of indices related to the grid $\{G_s^{\alpha,t}\}$, and a blackboard font $\mathbb{H}$ for sets of indices related to the grid $\{Q_j^k\}$.

The good and bad functions: Let $A_{G_r^{\alpha,t+1}} = |G_r^{\alpha,t+1}|_\sigma^{-1} \int_{G_r^{\alpha,t+1}} f\sigma$ be the σ-average of f on $G_r^{\alpha,t+1}$. Define functions $g_s^{\alpha,t}$ and $h_s^{\alpha,t}$ satisfying $f = g_s^{\alpha,t} + h_s^{\alpha,t}$ on $G_s^{\alpha,t}$ by

$$g_s^{\alpha,t}(x) = \begin{cases} A_{G_r^{\alpha,t+1}} & \text{for } x \in G_r^{\alpha,t+1} \text{ with } r \in \mathcal{K}_s^{\alpha,t}, \\ f(x) & \text{for } x \in G_s^{\alpha,t} \setminus \bigcup\{G_r^{\alpha,t+1} : r \in \mathcal{K}_s^{\alpha,t}\}, \end{cases} \tag{4-24}$$

$$h_s^{\alpha,t}(x) = \begin{cases} f(x) - A_{G_r^{\alpha,t+1}} & \text{for } x \in G_r^{\alpha,t+1} \text{ with } r \in \mathcal{K}_s^{\alpha,t}, \\ 0 & \text{for } x \in G_s^{\alpha,t} \setminus \bigcup\{G_r^{\alpha,t+1} : r \in \mathcal{K}_s^{\alpha,t}\}. \end{cases} \tag{4-25}$$

We extend both $g_s^{\alpha,t}$ and $h_s^{\alpha,t}$ to all of $\mathbb{R}^n$ by defining them to vanish outside $G_s^{\alpha,t}$.

Now $|A_{G_r^{\alpha,t+1}}| \le \gamma^{t+1}$ by (4-20). Thus Lebesgue's differentiation theorem shows that (any of the standard proofs can be adapted to the dyadic setting for positive locally finite Borel measures on $\mathbb{R}^n$)

$$|g_s^{\alpha,t}(x)| \le \gamma^{t+1} < \frac{\gamma}{|G_s^{\alpha,t}|_\sigma} \int_{G_s^{\alpha,t}} |f|\sigma \quad \text{for } \sigma\text{-a.e. } x \in G_s^{\alpha,t} \text{ and } (t, s) \in \mathbb{L}^\alpha. \tag{4-26}$$

That is, $g_s^{\alpha,t}$ is the "good" function and $h_s^{\alpha,t}$ is the "bad" function.

We can now refine the final sum on the left side of (4-14) according to the decomposition of $\mathcal{M}_\sigma^\alpha f$. We carry this out in three steps. In the first step, we fix an $\alpha \in \{0, \frac{1}{3}, \frac{2}{3}\}^n$, and for the remainder of the proof, we only consider Q_j^k for which $\vec{\alpha}(Q_j^k)=\alpha$. Namely, we will modify the important definition of $\mathbb{G}$ in (4-10) to

$$\mathbb{G}^\alpha = \big\{(k,j) : \vec{\alpha}(Q_j^k)=\alpha,\ |E_j^k|_\omega > \beta |N Q_j^k|_\omega \text{ and (4-8) holds}\big\}, \tag{4-27}$$

In the second step, we partition the indices (k,j) into the sets $\mathbb{H}_s^{\alpha,t}$ in (4-22) for $(t,s)\in\mathbb{L}^\alpha$. In the third step, for $(k,j)\in\mathbb{H}_s^{\alpha,t}$, we split f into the corresponding good and bad parts, yielding the decomposition

$$\sum_{(k,j)\in\mathbb{G}^\alpha} R_j^k \Big|\int_{E_j^k} (L\chi_{3Q_j^k} f\sigma)\omega\Big|^p \le C(I+II), \tag{4-28}$$

$$I=\sum_{(t,s)\in\mathbb{L}^\alpha} I_s^t, \qquad II=\sum_{(t,s)\in\mathbb{L}^\alpha} II_s^t, \tag{4-29}$$

$$I_s^t = \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big|\int_{E_j^k} (L\chi_{3Q_j^k} g_s^{\alpha,t}\sigma)\omega\Big|^p, \tag{4-30}$$

$$II_s^t = \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big|\int_{E_j^k} (L\chi_{3Q_j^k} h_s^{\alpha,t}\sigma)\omega\Big|^p, \tag{4-31}$$

$$\mathbb{I}_s^{\alpha,t} = \mathbb{G}\cap\mathbb{H}_s^{\alpha,t}. \tag{4-32}$$

Recall the definition of R_j^k in (4-15). In the definitions of I, I_s^t and II, II_s^t, we will suppress the dependence on $\alpha\in\{0,\frac{1}{3},\frac{2}{3}\}^n$. The same will be done for the subsequent decompositions of the (difficult) term II, although we usually retain the superscript α in the quantities arising in the estimates. In particular, the combinatorial properties of the cubes associated with $\mathbb{I}_s^{\alpha,t}$ are essential to completing this proof.

Term I requires only the forward testing condition (1-19) and the maximal theorem (2-2), while term II requires only the dual testing condition (1-14), along with the dual maximal function inequality (1-17) and the maximal theorem (2-2). The reader is again directed to Figure 4.2 for a map of the various decompositions of the terms and the conditions used to control them.

4.6. *The analysis of the good function.*

We claim that

$$I \le C\gamma^{2p}\mathfrak{T}^p \|f\|_{L^p(\sigma)}^p. \tag{4-33}$$

Proof. We use boundedness of the "good" function $g_s^{\alpha,t}$, as defined in (4-24), the testing condition (1-19) for $T_\natural$ (see also (4-4)), and finally the universal maximal function bound (2-2) with $\mu=\omega$. Here are the details. For $x\in E_j^k$, (4-7) implies that $L\chi_{3Q_j^k} g_s^{\alpha,t}\sigma(x) = Lg_s^{\alpha,t}\sigma(x)$ and so

$$\begin{aligned} I=\sum_{(t,s)\in\mathbb{L}^\alpha} I_s^t &= C\sum_{(t,s)\in\mathbb{L}^\alpha}\ \sum_{(k,j)\in\mathbb{G}^\alpha\cap\mathbb{H}_s^{\alpha,t}} R_j^k\Big|\int_{E_j^k}(Lg_s^{\alpha,t}\sigma)\omega\Big|^p \\ &\le C\sum_{(t,s)\in\mathbb{L}^\alpha}\int |\mathcal{M}_\omega^{dy}(\chi_{G_s^{\alpha,t}} Lg_s^{\alpha,t}\sigma)|^p\omega \le C\sum_{(t,s)\in\mathbb{L}^\alpha}\int_{G_s^{\alpha,t}} |Lg_s^{\alpha,t}\sigma|^p\omega \\ &\le C\gamma^{2p}\sum_{(t,s)\in\mathbb{L}^\alpha}\gamma^{pt}\int_{G_s^{\alpha,t}}\Big(T_\natural\frac{g_s^{\alpha,t}}{\gamma^{t+2}}\sigma\Big)^p\omega \le C\gamma^{2p}\mathfrak{T}^p\sum_{(t,s)\in\mathbb{L}^\alpha}\gamma^{pt}|G_s^{\alpha,t}|_\sigma, \end{aligned}$$

where we have used (4-26) and (1-19) with $g = g_s^{\alpha,t}/\gamma^{t+2}$ in the final inequality. This last sum is controlled by (4-18), and completes the proof of the claim. □

4.7. *The analysis of the bad function: Part 1.* It remains to estimate term II, as in (4-31), but this is in fact the harder term. Recall the definition of $\mathcal{H}_s^{\alpha,t}$ in (4-23). We now write

$$h_s^{\alpha,t} = \sum_{r\in\mathcal{H}_s^{\alpha,t}} (f - A_{G_r^{\alpha,t+1}})\chi_{G_r^{\alpha,t+1}} \equiv \sum_{r\in\mathcal{H}_s^{\alpha,t}} b_r, \tag{4-34}$$

where the "bad" functions b_r are supported in the cube $G_r^{\alpha,t+1}$ and have σ-mean zero, $\int_{G_r^{\alpha,t+1}} b_r\sigma = 0$. To take advantage of this, we will pass to the dual L^* below.

But first we must address the fact that the triples of the $\mathcal{D}^\alpha$ cubes $G_r^{\alpha,t+1}$ do not form a grid. Fix $(t, s) \in \mathbb{L}^\alpha$ and let

$$\mathfrak{C}_s^{\alpha,t} = \{3G_r^{\alpha,t+1} : r \in \mathcal{H}_s^{\alpha,t}\} \tag{4-35}$$

be the collection of triples of the $\mathcal{D}^\alpha$ cubes $G_r^{\alpha,t+1}$ with $r \in \mathcal{H}_s^{\alpha,t}$. We select the *maximal* triples

$$\{3G_{r_\ell}^{t+1}\}_{\ell\in\mathcal{L}_s^{\alpha,t}} \equiv \{T_\ell\}_{\ell\in\mathcal{L}_s^{\alpha,t}} \tag{4-36}$$

from the collection $\mathfrak{C}_s^{\alpha,t}$, and assign to each $r \in \mathcal{H}_s^{\alpha,t}$, the maximal triple $T_\ell = T_{\ell(r)}$ containing $3G_r^{\alpha,t+1}$ with least ℓ. Note that $T_{\ell(r)}$ extends outside $G_s^{\alpha,t}$ if $G_r^{\alpha,t+1}$ and $G_s^{\alpha,t}$ share a face. By Lemma 2.13 applied to $\mathcal{D}^\alpha$ the maximal triples $\{T_\ell\}_{\ell\in\mathcal{L}_s^{\alpha,t}}$ have finite overlap 3^n, and this will prove crucial in (4-49), (4-82) and (4-50) below.

We will pass to the dual of the linearization.

$$\int_{E_j^k} (Lh_s^{\alpha,t}\sigma)\omega = \sum_{r\in\mathcal{H}_s^{\alpha,t}} \int_{E_j^k} (Lb_r\sigma)\omega = \sum_{r\in\mathcal{H}_s^{\alpha,t}} \int_{G_r^{\alpha,t+1}\cap 3Q_j^k} (L^*\chi_{E_j^k}\omega)b_r\sigma \tag{4-37}$$

Note that (4-7) implies $L^*\nu$ is supported in $3Q_j^k$ if ν is supported in E_j^k, explaining the range of integration above. Continuing, we have for fixed $(k, j) \in \mathbb{I}_s^{\alpha,t}$,

$$|(4\text{-}37)| \le \Big|\sum_{r\in\mathcal{H}_s^{\alpha,t}} \int_{G_r^{\alpha,t+1}\cap 3Q_j^k} (L^*\chi_{E_j^k\cap T_{\ell(r)}}\omega)b_r\sigma\Big| + C\sum_{r\in\mathcal{H}_s^{\alpha,t}} \boldsymbol{P}(G_r^{\alpha,t+1}, \chi_{E_j^k\setminus 3G_r^{\alpha,t+1}}\omega)\int_{G_r^{\alpha,t+1}} |f|\sigma. \tag{4-38}$$

To see the inequality above, note that for $r \in \mathcal{H}_s^{\alpha,t}$ we are splitting the set E_j^k into $E_j^k \cap T_{\ell(r)}$ and $E_j^k \setminus T_{\ell(r)}$. On the latter set, the hypotheses of Lemma 2.9 are in force, namely the set $E_j^k \setminus T_{\ell(r)}$ does not intersect $3G_r^{\alpha,t+1}$, whence we have an estimate on the δ-Hölder modulus of continuity of $L^*\chi_{E_j^k} \setminus T_{\ell(r)}\omega$. Combine this with the fact that b_r has σ-mean zero on $G_r^{\alpha,t+1}$ to derive the estimate below, in which y_r^{t+1} is the center of the cube $G_r^{\alpha,t+1}$.

$$\begin{aligned}\Big|\int_{G_r^{\alpha,t+1}} (L^*\chi_{E_j^k\setminus T_{\ell(r)}}\omega)b_r\sigma\Big| &= \Big|\int_{G_r^{\alpha,t+1}} (L^*\chi_{E_j^k\setminus T_{\ell(r)}}\omega(y) - L^*\chi_{E_j^k\setminus T_{\ell(r)}}\omega(y_r^{t+1}))(b_r\sigma)\Big| \\ &\le \int_{G_r^{\alpha,t+1}\cap 3Q_j^k} C\boldsymbol{P}(G_r^{\alpha,t+1}, \chi_{E_j^k\setminus T_{\ell(r)}}\omega)\delta\Big(\frac{|y - y_r^{t+1}|}{\ell(G_r^{\alpha,t+1})}\Big)|b_r(y)|\, d\sigma(y)\end{aligned}$$

$$\leq C\,\boldsymbol{P}(G_r^{\alpha,t+1},\, \chi_{E_j^k \setminus 3G_r^{\alpha,t+1}}\omega) \int_{G_r^{\alpha,t+1}} |f|\, d\sigma.$$

We have after application of (4-38),

$$II_s^t = \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} R_j^k \Big(\int_{E_j^k} (L h_s^{\alpha,t}\sigma)\omega \Big)^p \leq II_s^t(1) + II_s^t(2),$$

where

$$II_s^t(1) = \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r\in \mathcal{H}_s^{\alpha,t}} \int_{G_r^{\alpha,t+1}} (L^* \chi_{E_j^k \cap T_{\ell(r)}}\omega) b_r \sigma \Big|^p, \tag{4-39}$$

$$II_s^t(2) = \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} R_j^k \Big(\sum_{r\in \mathcal{H}_s^{\alpha,t}} \boldsymbol{P}(G_r^{\alpha,t+1},\, \chi_{E_j^k}\omega) \int_{G_r^{\alpha,t+1}} |f|\sigma \Big)^p.$$

Note that we may further restrict the integration in (4-39) to $G_r^{\alpha,t+1} \cap 3Q_j^k$ since $L^* \chi_{E_j^k} \cap T_{\ell(r)}\omega$ is supported in $3Q_j^k$.

4.7.1. *Analysis of $II(2)$.* Recalling the definition of $\mathfrak{M}_*$ in (4-3), we claim that

$$\sum_{(t,s)\in \mathbb{L}^\alpha} II_s^t(2) \leq C\gamma^{2p} \mathfrak{M}_*^p \int |f|^p \sigma. \tag{4-40}$$

Proof. We begin by defining a linear operator by

$$\boldsymbol{P}_j^k(\mu) \equiv \sum_{r\in \mathcal{H}_s^{\alpha,t}} \boldsymbol{P}(G_r^{\alpha,t+1},\, \chi_{E_j^k}\mu)\chi_{G_r^{\alpha,t+1}}. \tag{4-41}$$

In this notation, we have for $(k,j) \in \mathbb{I}_s^{\alpha,t}$ (see (4-22) and (4-31)),

$$\begin{aligned}\sum_{r\in \mathcal{H}_s^{\alpha,t}} \boldsymbol{P}\big(G_r^{\alpha,t+1},\, \chi_{E_j^k}\omega(dx)\big) \int_{G_r^{\alpha,t+1}} |f|\sigma &= \sum_{r\in \mathcal{H}_s^{\alpha,t}} \boldsymbol{P}\big(G_r^{\alpha,t+1},\, \chi_{E_j^k}\omega\big) \int_{G_r^{\alpha,t+1}} \sigma \Big(\frac{1}{|G_r^{\alpha,t+1}|_\sigma} \int_{G_r^{\alpha,t+1}} |f|\sigma \Big) \\ &\leq \gamma^{t+2} \int_{G_s^{\alpha,t}} \boldsymbol{P}_j^k(\omega)\sigma = \gamma^{t+2} \int_{E_j^k} (\boldsymbol{P}_j^k)^*(\chi_{G_s^{\alpha,t}}\sigma)\omega.\end{aligned}$$

By assumption, the maximal function $\mathcal{M}(\omega\cdot)$ maps $L^{p'}(\omega)$ to $L^{p'}(\sigma)$, and we now note a particular consequence of this. In the definition (4-41) we were careful to insert $\chi_{E_j^k}$ on the right hand side. These sets are pairwise disjoint, whence we have the inequality below for measures μ.

$$\begin{aligned}\sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} \boldsymbol{P}_j^k(\mu)(x) &\leq \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} \sum_{r\in \mathcal{H}_s^{\alpha,t}} \sum_{\ell=0}^{\infty} \frac{\delta(2^{-\ell})}{|2^\ell G_r^{\alpha,t+1}|} \Big(\int_{2^\ell G_r^{\alpha,t+1}} \chi_{E_j^k}\mu \Big) \chi_{G_r^{\alpha,t+1}}(x) \\ &\leq \sum_{\ell=0}^{\infty} \sum_{r\in \mathcal{H}_s^{\alpha,t}} \frac{\delta(2^{-\ell})}{|2^\ell G_r^{\alpha,t+1}|} \Big(\int_{2^\ell G_r^{\alpha,t+1} \cap G_s^{\alpha,t}} \mu \Big) \chi_{G_r^{\alpha,t+1}}(x) \leq C\chi_{G_s^{\alpha,t}} \mathcal{M}(\chi_{G_s^{\alpha,t}}\mu)(x).\end{aligned} \tag{4-42}$$

Thus the inequality

$$\Big\| \chi_{G_s^{\alpha,t}} \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} \boldsymbol{P}_j^k(|g|\omega) \Big\|_{L^{p'}(\sigma)} \le C\mathfrak{M}_* \|\chi_{G_s^{\alpha,t}} g\|_{L^{p'}(\omega)} \tag{4-43}$$

follows immediately. By duality we then have

$$\Big\| \chi_{G_s^{\alpha,t}} \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} (\boldsymbol{P}_j^k)^*(|h|\sigma) \Big\|_{L^{p}(w)} \le C\mathfrak{M}_* \|\chi_{G_s^{\alpha,t}} h\|_{L^{p}(\sigma)}. \tag{4-44}$$

Note that it was the linearity that we wanted in (4-41), so that we could appeal to the dual maximal function assumption.

We thus obtain

$$II_s^t(2) \le \gamma^{p(t+2)} \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} R_j^k \Big(\int_{Q_j^k} (\boldsymbol{P}_j^k)^*(\chi_{G_s^{\alpha,t}}\sigma)\, d\omega \Big)^p.$$

Summing in (t,s) and using $(\boldsymbol{P}_j^k)^* \le \sum_{(\ell,i)\in \mathbb{I}_s^{\alpha,t}} (\boldsymbol{P}_i^\ell)^*$ for $(k,j)\in \mathbb{I}_s^{\alpha,t}$, we obtain

$$\begin{aligned}
\sum_{(t,s)\in \mathbb{L}^\alpha} II_s^t(2) &\le C\gamma^{2p} \sum_{(t,s)\in \mathbb{L}^\alpha} \gamma^{pt} \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} R_j^k \Big(\int_{Q_j^k} (\boldsymbol{P}_j^k)^*(\chi_{G_s^{\alpha,t}}\sigma)\, d\omega \Big)^p & \text{(4-45)}\\
&= C\gamma^{2p} \sum_{(t,s)\in \mathbb{L}^\alpha} \gamma^{pt} \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} |E_j^k|_\omega \Big(\frac{1}{|NQ_j^k|_w} \int_{Q_j^k} (\boldsymbol{P}_j^k)^*(\chi_{G_s^{\alpha,t}}\sigma)\omega \Big)^p &\\
&\le C\gamma^{2p} \sum_{(t,s)\in \mathbb{L}^\alpha} \gamma^{pt} \int \Big(\mathcal{M}_\omega^{dy}\big(\chi_{G_s^{\alpha,t}} \sum_{(\ell,i)\in \mathbb{I}_s^{\alpha,t}} (\boldsymbol{P}_i^\ell)^*(\chi_{G_s^{\alpha,t}}\sigma)\big) \Big)^p \omega & \text{(4-46)}\\
&\le C\gamma^{2p} \sum_{(t,s)\in \mathbb{L}^\alpha} \gamma^{pt} \int_{G_s^{\alpha,t}} \Big(\sum_{(\ell,i)\in \mathbb{I}_s^{\alpha,t}} (\boldsymbol{P}_i^\ell)^*(\chi_{G_s^{\alpha,t}}\sigma) \Big)^p \omega & \text{(4-47)}\\
&\le C\gamma^{2p}\mathfrak{M}_*^p \sum_{(t,s)\in \mathbb{L}^\alpha} \gamma^{pt} |G_s^{\alpha,t}|_\sigma, &
\end{aligned}$$

which is bounded by $C\gamma^{2p}\mathfrak{M}_*^p \int |f|^p\sigma$. In the last line we are applying (4-44) with $h \equiv 1$. □

4.7.2. *Decomposition of* $II(1)$. We note that the term $II_s^t(1)$ is dominated by $II_s^t(1) \le III_s^t + IV_s^t$, where

$$\begin{aligned}
III_s^t &= \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r\in \mathcal{K}_s^{\alpha,t}} \int_{G_r^{\alpha,t+1}\setminus \Omega_{k+2}} (L^*\chi_{E_j^k\cap T_{\ell(r)}}\omega) b_r \sigma \Big|^p,\\
IV_s^t &= \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r\in \mathcal{K}_s^{\alpha,t}} \int_{G_r^{\alpha,t+1}\cap \Omega_{k+2}} (L^*\chi_{E_j^k\cap T_{\ell(r)}}\omega) b_r \sigma \Big|^p. \qquad \text{(4-48)}
\end{aligned}$$

The term III_s^t includes that part of b_r supported on $G_r^{\alpha,t+1} \setminus \Omega_{k+2}$, and the term IV_s^t includes that part of b_r supported on $G_r^{\alpha,t+1} \cap \Omega_{k+2}$, which is the more delicate case.

Remark 4.8. The key difference between the terms III_s^t and IV_s^t is the range of integration: $G_r^{\alpha,t+1}\setminus\Omega_{k+2}$ for III_s^t and $G_r^{\alpha,t+1}\cap\Omega_{k+2}$ for IV_s^t. Just as for the fractional integral case, it is the latter case that is harder, requiring combinatorial facts, which we come to at the end of the argument. An additional fact that we return to in different forms is that the set $G_r^{\alpha,t+1} \cap \Omega_{k+2}$ can be further decomposed using Whitney decompositions of Ω_{k+2} in the grid $\mathcal{D}^\alpha$.

Recall the definition of $\mathfrak{T}_*$ in (4-5). We claim

$$\sum_{(t,s)\in\mathbb{L}^\alpha} III_s^t \le C\mathfrak{T}_*^p \int |f|^p \sigma. \tag{4-49}$$

Proof. Let $\widetilde{E}_j^k = 3Q_j^k \setminus \Omega_{k+2}$ (note that $\widetilde{E}_j^k$ is much larger than E_j^k). We will use the definition of R_j^k in (4-15), and the fact that

$$\sum_{\ell\in\mathcal{L}_s^{\alpha,t}} \chi_{T_\ell} \le 3^n \tag{4-50}$$

provided $N \ge 9$. We will apply the form (2-11) of (1-14) with $g = \chi_{E_j^k \cap T_\ell}$ — also see (4-5) — and with

$$Q \equiv T_\ell \cap \hat{Q}_j^k \quad \text{and} \quad Q \equiv T_\ell$$

in the cases $T_\ell \cap \hat{Q}_j^k$ is a cube and is not a cube, respectively (the latter is possible since T_ℓ is the *triple* of a $\mathcal{D}^\alpha$-cube). In each case we claim that

$$Q \subset T_\ell \cap 3\hat{Q}_j^k.$$

Indeed, recall that $\hat{Q}_j^k$ is the cube in the shifted grid $\mathcal{D}^\alpha$ that is selected by Q_j^k as in the definition "Select a shifted grid" above and satisfies $3\hat{Q}_j^k \subset MQ_j^k \subset NQ_j^k$, where N is as in Remark 2.4, by choosing R_W sufficiently large in (2-6). Now T_ℓ is a triple of a cube in the grid $\mathcal{D}^\alpha$ and $\hat{Q}_j^k$ is a cube in $\mathcal{D}^\alpha$. Thus if $T_\ell \cap \hat{Q}_j^k$ is *not* a cube, then we must have $T_\ell \subset 3\widetilde{Q}_j^k$ and this proves the claim. We then have

$$\begin{aligned}
III_s^t &\le \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big(\sum_{\ell\in\mathcal{L}_s^{\alpha,t}} \sum_{r\in\mathcal{H}_s^{\alpha,t}:\ell=\ell(r)} \int_{G_r^{\alpha,t+1}\cap\widetilde{E}_j^k} |L^*\chi_{E_j^k\cap T_{\ell(r)}}\omega|^{p'}\sigma \Big)^{p-1} \int_{\widetilde{E}_j^k} |h_s^{\alpha,t}|^p\sigma \\
&\le \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big(\sum_{\ell\in\mathcal{L}_s^{\alpha,t}} \int_{T_\ell\cap 3\hat{Q}_j^k} |L^*\chi_{E_j^k\cap T_\ell}\omega|^{p'}\sigma \Big)^{p-1} \int_{\widetilde{E}_j^k} |h_s^{\alpha,t}|^p\sigma \\
&\le \mathfrak{T}_*^p \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big(\sum_{\ell\in\mathcal{L}_s^{\alpha,t}} |T_\ell\cap 3\hat{Q}_j^k|_\omega \Big)^{p-1} \int_{\widetilde{E}_j^k} |h_s^{\alpha,t}|^p\sigma \\
&\le \mathfrak{T}_*^p \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} \frac{|E_j^k|_\omega}{|NQ_j^k|_\omega} |NQ_j^k|_\omega^{p-1} \int_{\widetilde{E}_j^k} |h_s^{\alpha,t}|^p\sigma \\
&\le C\mathfrak{T}_*^p \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} \int_{\widetilde{E}_j^k} |h_s^{\alpha,t}|^p\sigma \le C\mathfrak{T}_*^p \sum_{(k,j)\in\mathbb{G}^\alpha\cap\mathbb{H}_s^{\alpha,t}} \int_{\widetilde{E}_j^k} (|f|^p + |\mathcal{M}_\sigma^\alpha f|^p)\sigma.
\end{aligned}$$

Using

$$\sum_{(t,s)\in\mathbb{L}^\alpha} \sum_{(k,j)\in\mathbb{G}^\alpha\cap\mathbb{H}_s^{\alpha,t}} \chi_{\widetilde{E}_j^k} = \sum_{\text{all } k,j} \chi_{\widetilde{E}_j^k} \le C, \tag{4-51}$$

we thus obtain (4-49). □

4.9. ***The analysis of the bad function: Part 2.*** This is the most intricate and final case. We will prove

$$\sum_{(t,s)\in\mathbb{L}^\alpha} IV_s^t \le C(\gamma^{2p}\mathfrak{T}^p + \mathfrak{T}_*^p + \gamma^{2p}\mathfrak{M}_*^p)\int |f|^p\sigma, \tag{4-52}$$

where $\mathfrak{T}$, $\mathfrak{T}_*$ and $\mathfrak{M}_*$ are defined in (4-4), (4-5) and (4-3), respectively. The estimates (4-33), (4-40), (4-49), (4-52) prove (4-12), and so complete the proof of assertion 1 of the strong-type characterization in Theorem 1.9. Assertions 2 and 3 of Theorem 1.9 follow as in the weak-type Theorem 1.8. Finally, to prove assertion 4 we note that Lemma 2.12 and condition (1-19) imply (1-6), which by Theorem 1.2 yields (1-5).

4.9.1. *Whitney decompositions with shifted grids.* We now use the shifted grid $\mathcal{D}^\alpha$ in place of the dyadic grid $\mathcal{D}$ to form a Whitney decomposition of Ω_k in the spirit of (2-6). However, in order to fit the $\mathcal{D}^\alpha$-cubes $\hat{Q}_j^k$ defined above in "Select a shifted grid", it will be necessary to use a smaller constant than the constant R_W already used for the Whitney decomposition of Ω_k into $\mathcal{D}$-cubes. Recall the dimensional constant M defined in (4-21): it satisfies $\hat{Q} \subset MQ$. Define the new constant

$$R'_W = \frac{R_W}{M}.$$

We now use the decomposition of Ω_k in (2-6), but with $\mathcal{D}$ replaced by $\mathcal{D}^\alpha$ and with R_W replaced by R'_W. We have thus decomposed

$$\Omega_k = \bigsqcup_m B_m^k$$

into a Whitney decomposition of pairwise disjoint cubes B_m^k in $\mathcal{D}^\alpha$ satisfying

$$\begin{gathered} R'_W B_m^k \subset \Omega_k, \\ 3R'_W B_m^k \cap \Omega_k^c \neq \varnothing, \end{gathered} \tag{4-53}$$

and the following analogue of the nested property in (2-6):

$$B_j^k \subsetneq B_i^\ell \quad \text{implies} \quad k > \ell. \tag{4-54}$$

Now we introduce yet another construction. For every pair (k, j) let $\widetilde{Q}_j^k$ be the unique $\mathcal{D}^\alpha$-cube B_m^k containing $\hat{Q}_j^k$. Note that such a cube $\widetilde{Q}_j^k = B_m^k$ exists since $\hat{Q}_j^k \subset MQ_j^k$ by (4-21) and $R_W Q_j^k \subset \Omega_k$ by (2-6) implies that $R'_W \hat{Q}_j^k \subset \Omega_k$. Of course the cube $\widetilde{Q}_j^k = B_m^k$ satisfies

$$R'_W \widetilde{Q}_j^k \subset \Omega_k. \tag{4-55}$$

Moreover, we can arrange to have

$$3\widetilde{Q}_j^k \subset NQ_j^k, \tag{4-56}$$

where N is as in Remark 2.4, by choosing R_W sufficiently large in (2-6). See Figure 4.3.

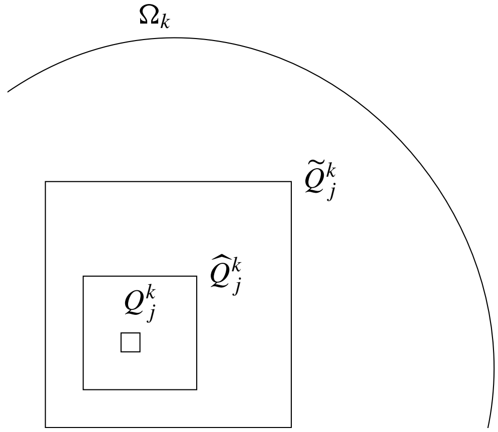

Figure 4.3. The relative positions of the cubes Q_j^k, $\widehat{Q}_j^k$, and $\widetilde{Q}_j^k$ inside a set Ω_k.

We will use this decomposition for the set $\Omega_{k+2} = \sqcup_m B_m^{k+2}$ in our arguments below. The corresponding cubes $\widetilde{Q}_i^{k+2}$ that arise as certain of the B_m^{k+2} satisfy the conditions

$$3Q_i^{k+2} \subset \widehat{Q}_i^{k+2} \subset \widetilde{Q}_i^{k+2} \subset 3\widetilde{Q}_i^{k+2} \subset N Q_i^{k+2} \subset \Omega_{k+2}. \tag{4-57}$$

Note that the set of indices m arising in the decomposition of Ω_{k+2} into $\mathcal{D}^\alpha$ cubes B_m^{k+2} is *not* the same as the set of indices i arising in the decomposition of Ω_{k+2} into $\mathcal{D}$ cubes Q_i^{k+2}, but this should not cause confusion. So we will usually write B_i^{k+2} with dummy index i unless it is important to distinguish the cubes $B_i^k + 2$ from the cubes Q_i^{k+2}. This distinction will be important in the proof of the "bounded occurrence of cubes" property in Section 4.14.7 below.

Now use $\Omega_{k+2} = \bigcup B_i^{k+2}$ to split the term IV_s^t in (4-48) into two pieces as follows:

$$\begin{aligned} IV_s^t \le \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r\in\mathcal{H}_s^{\alpha,t}} \sum_{i\in\mathcal{J}_s^t} \int_{G_r^{\alpha,t+1}\cap B_i^{k+2}} (L^* \chi_{E_j^k \cap T_{\ell(r)}}\omega) b_r \sigma \Big|^p \\ + \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r\in\mathcal{H}_s^{\alpha,t}} \sum_{i\in\mathfrak{J}_s^t} \int_{G_r^{\alpha,t+1}\cap B_i^{k+2}} (L^* \chi_{E_j^k \cap T_{\ell(r)}}\omega) b_r \sigma \Big|^p \\ = IV_s^t(1) + IV_s^t(2), \end{aligned} \tag{4-58}$$

where

$$\mathcal{J}_s^t = \{i : \mathsf{A}_i^{k+2} > \gamma^{t+2}\} \quad \text{and} \quad \mathfrak{J}_s^t = \{i : \mathsf{A}_i^{k+2} \le \gamma^{t+2}\}, \tag{4-59}$$

and where

$$\mathsf{A}_i^{k+2} = \frac{1}{|B_i^{k+2}|_\sigma} \int_{B_i^{k+2}} |f|\, d\sigma \tag{4-60}$$

denotes the σ-average of $|f|$ on the cube B_i^{k+2}. Thus $IV(1)$ corresponds to the case where the averages are "big" and $IV(2)$ where the averages are "small". The analysis of $IV_s^t(1)$ in (4-58) is the hard case, taken up later.

4.9.2. *A first combinatorial argument.*

Lemma 4.10 (bounded occurrence of cubes). *A given cube $B \in \mathcal{D}^\alpha$ can occur only a* bounded *number of times as B_i^{k+2}, where*

$$B_i^{k+2} \subset \widetilde{Q}_j^k \quad \text{with } (k, j) \in \mathbb{G}^\alpha.$$

Specifically, let $(k_1, j_1), \dots, (k_M, j_M) \in \mathbb{G}^\alpha$, as defined in (4-27), *be such that $B = B_{i_\sigma}^{k_\sigma+2}$ for some i_σ and $B \subset \widetilde{Q}_{j_\sigma}^{k_\sigma}$ for $1 \le \sigma \le M$. It follows that $M \le C\beta^{-1}$, where β is the small constant chosen in the definition of $\mathbb{G}^\alpha$. The constant C here depends only on dimension.*

The Whitney structure (see (2-6)) is decisive here, as well as the fact that $|E_j^k|_\omega \ge \beta |NQ_j^k|_\omega$ for $(k, j) \in \mathbb{G}^\alpha$. For this proof it will be useful to use m to index the cubes $B_m^k + 2$ and to use i to index the cubes Q_i^{k+2}. The following lemma captures the main essence of the Whitney structure, and will be applied to cubes B_m^{k+2} satisfying (4-53) and cubes Q_i^{k+2} satisfying (2-6).

Lemma 4.11. *Suppose that Q is a member of the Whitney decomposition of Ω with respect to the grid $\mathcal{D}$ and with Whitney constant R_W. Suppose also that a cube B is a member of a Whitney decomposition of the same open set Ω but with respect to the grid $\mathcal{D}^\alpha$ and with Whitney constant R'_W. If $N < \frac{1}{2} R_W$ and $B \subset NQ$, then the side lengths of Q and B are comparable*:

$$\ell(Q) \approx \ell(B).$$

Proof of Lemma 4.11. Since $N < \frac{1}{2} R_W$ and Q is a Whitney cube we have

$$\ell(Q) \approx \operatorname{dist}(Q, \partial\Omega) \approx \sup_{x \in NQ} \operatorname{dist}(x, \partial\Omega) \approx \inf_{x \in NQ} \operatorname{dist}(x, \partial\Omega).$$

Then since $B \subset NQ$ and B is a Whitney cube (for the other decomposition) we have

$$\ell(Q) \approx \operatorname{dist}(B, \partial\Omega) \approx \ell(B). \qquad \square$$

Proof of Lemma 4.10. So suppose that $(k_1, j_1), \dots, (k_M, j_M) \in \mathbb{G}^\alpha$ and $B = B_{i_\sigma}^{k_\sigma+2} \subset \widetilde{Q}_{j_\sigma}^{k_\sigma}$ for $1 \le \sigma \le M$, with the pairs of indices (k_σ, j_σ) being distinct. Observe that the finite overlap property in (2-6) applies to the cubes $\widetilde{Q}_{j_\sigma}^{k_\sigma}$ in the Whitney decomposition (4-53) of Ω_{k_σ} with grid $\mathcal{D}^\alpha$ and Whitney constant R'_W. Thus for fixed k, the number of (k_σ, j_σ) with $k_\sigma = k$ is bounded by the finite overlap constant since B is inside each $\widetilde{Q}_{j_\sigma}^{k_\sigma}$. This gives us the observation that a single integer k can occur only a *bounded* number C_b of times among the $k_1, \dots, k_M$.

After a relabeling, we can assume that all the k_σ for $1 \le \sigma \le M'$ are distinct, listed in increasing order, and that the number M' of k_σ satisfies $M \le C_b M'$. The nested property of (2-6) assures us that B is an element of the Whitney decomposition (4-53) of Ω_k for *all* $k_1 \le k \le k_{M'}$.

Remark 4.12. Note that the k_σ are not necessarily consecutive since we require that $(k_\sigma, j_\sigma) \in \mathbb{G}^\alpha$. Nevertheless, the cube B *does* occur among the B_i^{k+2} for any k that lies between k_σ and $k_{\sigma+1}$. These latter occurrences of B may be unbounded, but we are only concerned with bounding those for which $(k_\sigma, j_\sigma) \in \mathbb{G}^\alpha$, and it is these occurrences that our argument is treating.

Thus for $3 \le \sigma \le M'$, we have $k_1 \le k_\sigma - 2 \le k_{M'}$, and it follows from Remark 4.12 that the cube B is a member of the Whitney decomposition (4-53) of the open set Ω_{k_σ} with grid $\mathcal{D}^\alpha$ and Whitney constant R'_W. But we also have that $Q^{k_\sigma}_{j_\sigma}$ is a member of the Whitney decomposition (2-6) of Ω_{k_σ} with grid $\mathcal{D}$ and Whitney constant R_W. Thus Lemma 4.11 gives us the equivalence of side lengths $\ell(Q^{k_\sigma}_{j_\sigma}) \approx \ell(B)$. Combining this with the containment $NQ^{k_\sigma}_{j_\sigma} \supset B$, we see that the number of possible locations for the cubes $Q^{k_\sigma}_{j_\sigma} \in \mathcal{D}$ is bounded by a constant C'_b depending only on dimension.

Apply the pigeonhole principle to the possible locations of the $Q^{k_\sigma}_{j_\sigma}$. After a relabeling, we can argue under the assumption that all $Q^{k_\sigma}_{j_\sigma}$ equal the same cube Q' for all choices of $1 \le \sigma \le M''$, where $M' \le C'_b M''$. Now comes the crux of the argument where the condition that the indices (k_σ, j_σ) lie in $\mathbb{G}^\alpha$, as given in (4-27), proves critical. In particular we have $|E^{k_\sigma}_{j_\sigma}|_\omega \ge \beta |NQ'|_\omega$ where N is as in Remark 2.4. The k_σ are distinct, and the sets $E^{k_\sigma}_{j_\sigma} \subset Q'$ are pairwise disjoint; hence

$$M''\beta|NQ'|_\omega \le \sum_{\sigma=1}^{M''} |E^{k_\sigma}_{j_\sigma}|_\omega \le |Q'|_\omega \quad \text{implies} \quad M'' \le \beta^{-1}.$$

Thus $M \le C_b C'_b \beta^{-1}$ and our proof of the claim is complete. □

4.12.1. *Replace bad functions by averages.* The first task in the analysis of the terms $IV^t_s(1)$ and $IV^t_s(2)$ will be to replace part of the "bad functions" b_r by their averages over $B^k_i + 2$, or more exactly the averages A^{k+2}_i. We again appeal to the Hölder continuity of $L^*\chi_{E^k_j \cap T_\ell}\omega$. By construction, $3B^{k+2}_i$ does not meet E^k_j, so that Lemma 2.9 applies. If $B^{k+2}_i \subset G^{\alpha,t+1}_r$ for some r, then there is a constant c^{k+2}_i satisfying $|c^{k+2}_i| \le 1$ such that

$$\begin{aligned}\Big| \int_{B^{k+2}_i} (L^*\chi_{E^k_j \cap T_{\ell(r)}}\omega) b_r \sigma - \Big(c^{k+2}_i \int_{B^{k+2}_i} (L^*\chi_{E^k_j \cap T_{\ell(r)}}\omega)\sigma \Big)(|A^{\alpha,t+1}_r| + \mathsf{A}^{k+2}_i) \Big| \\ \le C\boldsymbol{P}(B^{k+2}_i, \chi_{E^k_j \cap T_{\ell(r)}}\omega) \int_{B^{k+2}_i} |b_r|\sigma. \quad \text{(4-61)}\end{aligned}$$

Indeed, if z^{k+2}_i is the center of the cube B^{k+2}_i, we have

$$\begin{aligned}&\int_{B^{k+2}_i} (L^*\chi_{E^k_j \cap T_{\ell(r)}}\omega) b_r \sigma \\ &\quad = L^*(\chi_{E^k_j \cap T_{\ell(r)}}\omega)(z^{k+2}_i) \int_{B^{k+2}_i} b_r \sigma + O\Big(\boldsymbol{P}(B^{k+2}_i, \chi_{E^k_j \cap T_{\ell(r)}}\omega) \int_{B^{k+2}_i} |b_r|\sigma \Big) \\ &\quad = \Big(\int_{B^{k+2}_i} (L^*\chi_{E^k_j \cap T_{\ell(r)}}\omega)\sigma \Big) \frac{1}{|B^{k+2}_i|_\sigma} \int_{B^{k+2}_i} b_r \sigma + O\Big(\boldsymbol{P}(B^{k+2}_i, \chi_{E^k_j \cap T_{\ell(r)}}\omega) \int_{B^{k+2}_i} |b_r|\sigma \Big).\end{aligned}$$

Now, the functions b_r are given in (4-34), and by construction, we note that

$$\frac{1}{|B^{k+2}_i|_\sigma} \Big| \int_{B^{k+2}_i} b_r \sigma \Big| \le \Big| \frac{1}{|G^{\alpha,t+1}_r|_\sigma} \int_{G^{\alpha,t+1}_r} f\sigma \Big| + \frac{1}{|B^{k+2}_i|_\sigma} \int_{B^{k+2}_i} |f|\sigma = |A^{\alpha,t+1}_r| + \mathsf{A}^{k+2}_i.$$

So with

$$c^{k+2}_i = \frac{1}{|A^{\alpha,t+1}_r| + \mathsf{A}^{k+2}_i} \frac{1}{|B^{k+2}_i|_\sigma} \int_{B^{k+2}_i} b_r \sigma,$$

we have $|c_i^{k+2}| \leq 1$ and

$$\int_{B_i^{k+2}} (L^* \chi_{E_j^k \cap T_{\ell(r)}} \omega) b_r \sigma = \Big(c_i^{k+2} \int_{B_i^{k+2}} (L^* \chi_{E_j^k \cap T_{\ell(r)}} \omega)\sigma \Big) (|A_r^{\alpha,t+1}| + \mathsf{A}_i^{k+2}) \\ + O\Big(\boldsymbol{P}(B_i^{k+2}, \chi_{E_j^k \cap T_{\ell(r)}} \omega) \int_{B_i^{k+2}} |b_r| \sigma \Big).$$

In the special case where B_i^{k+2} is equal to $G_r^{\alpha,t+1}$, we have $\int_{B_i^{k+2}} b_r \sigma = \int b_r \sigma = 0$ and the proof above shows that

$$\Big| \int_{G_r^{\alpha,t+1}} (L^* \chi_{E_j^k \cap T_{\ell(r)}} \omega) b_r \sigma \Big| \leq C \boldsymbol{P}(G_r^{\alpha,t+1}, \chi_{E_j^k \cap T_{\ell(r)}} \omega) \int_{G_r^{\alpha,t+1}} |f| \sigma \ , \tag{4-62}$$

since $\int_{G_r^{\alpha,t+1}} |b_r| \sigma = \int_{G_r^{\alpha,t+1}} |f - A_r^{\alpha,t+1}| \sigma \leq 2 \int_{G_r^{\alpha,t+1}} |f| \sigma$.

Our next task is to organize the sum over the cubes B_i^{k+2} relative to the cubes $G_r^{\alpha,t+1}$. This is needed because the cubes B_i^{k+2} are *not* pairwise disjoint in k, and we thank Tuomas Hytonen for bringing this point to our attention. The cube B_i^{k+2} must intersect $\bigcup_{r \in \mathcal{K}_s^{\alpha,t}} G_r^{\alpha,t+1}$ since otherwise

$$\int_{G_r^{\alpha,t+1} \cap B_i^{k+2}} (L^* \chi_{E_j^k \cap T_{\ell(r)}} \omega) b_r \sigma = 0 \quad \text{for } r \in \mathcal{K}_s^{\alpha,t}.$$

Thus B_i^{k+2} satisfies exactly one of the following two cases which we indicate by writing $i \in$ Case(a) or $i \in$ Case(b)

Case(a) B_i^{k+2} *strictly* contains at least one of the cubes $G_r^{\alpha,t+1}$ for $r \in \mathcal{K}_s^{\alpha,t}$.

Case(b) $B_i^{k+2} \subset G_r^{\alpha,t+1}$ for some $r \in \mathcal{K}_s^{\alpha,t}$.

Note that the cubes B_i^{k+2} with $i \in \mathcal{I}_s^t$ can only satisfy Case(b), while the cubes B_i^{k+2} with $i \in \mathcal{J}_s^t$ can satisfy either of the two cases above. However, we have the following claim.

Claim 4.13. *For each fixed $r \in \mathcal{K}_s^{\alpha,t}$, we have*

$$\sum_{(k+2,i,j) \text{ admissible}} \chi_{B_i^{k+2}} \leq C,$$

where the sum is taken over all admissible *index* triples $(k+2, i, j)$, *that is, those for which the cube B_i^{k+2} arises in term IV_s^t with both $B_i^{k+2} \subset G_r^{\alpha,t+1}$ and $B_i^{k+2} \subset \widetilde{Q}_j^k$.*

But we first establish a containment that will be useful later as well. Recall that Ω_{k+2} decomposes as a pairwise disjoint union of cubes B_i^{k+2}, and thus we have

$$\int_{G_r^{\alpha,t+1} \cap \Omega_{k+2}} (L^* \chi_{E_j^k \cap T_{\ell(r)}} \omega) b_r \sigma = \sum_{i : B_i^{k+2} \cap \widetilde{Q}_j^k \neq \varnothing} \int_{B_i^{k+2}} (L^* \chi_{E_j^k \cap T_{\ell(r)}} \omega) b_r \sigma,$$

since the support of $L^* \chi_{E_j^k \cap T_{\ell(r)}} \omega$ is contained in $2Q_j^k \subset \hat{Q}_j^k \subset \widetilde{Q}_j^k$ by (4-7). Since both B_i^{k+2} and $\widetilde{Q}_j^k$ lie in the grid $\mathcal{D}^\alpha$ and have nonempty intersection, one of these cubes is contained in the other. Now B_i^{k+2}

cannot *strictly* contain $\widetilde{Q}_j^k$ since $\widetilde{Q}_j^k = B_\ell^k$ for some ℓ and the cubes $\{B_j^k\}_{k,j}$ satisfy the nested property (4-54). It follows that we must have

$$B_i^{k+2} \subset \widetilde{Q}_j^k \quad \text{whenever } B_i^{k+2} \cap \widetilde{Q}_j^k \neq \varnothing. \tag{4-63}$$

Now we return to Claim 4.13, and note that for a fixed index pair $(k+2, i)$, the bounded overlap condition in (2-6) shows that there are only a bounded number of indices j such that $B_i^{k+2} \subset \widetilde{Q}_j^k \subset N Q_j^k$ — see (4-56). We record this observation here:

$$\#\{j : B_i^{k+2} \subset \widetilde{Q}_j^k\} \leq C \quad \text{for each pair } (k+2, i). \tag{4-64}$$

Thus Claim 4.13 is reduced to this one:

Claim 4.14. $\sum\{\chi_{B_i^{k+2}} : B_i^{k+2} \subset G_r^{\alpha,t+1}$ *for some* $(k, j) \in \mathbb{I}_s^{\alpha,t}$ *with* $B_i^{k+2} \subset \widetilde{Q}_j^k\} \leq C$ *for each* $r \in \mathcal{K}_s^{\alpha,t}$.

As is the case with similar assertions in this argument, a central obstacle is that a given cube B can arise in many different ways as a B_i^{k+2}.

Proof of Claim 4.14. We will appeal to the "bounded occurrence of cubes" in Section 4.9.2 above. This principle relies upon the definition of $\mathbb{G}^\alpha$ in (4-27), and applies in this setting due to the definition of $\mathbb{I}_s^{\alpha,t}$ in (4-28). We also appeal to the following fact:

$$G_r^{\alpha,t+1} \subset \widetilde{Q}_j^k \quad \text{whenever } B_i^{k+2} \subset G_r^{\alpha,t+1} \cap \widetilde{Q}_j^k \text{ with } (k, j) \in \mathbb{I}_s^{\alpha,t}. \tag{4-65}$$

To see (4-65), we note that both of the cubes $G_r^{\alpha,t+1}$ and $\widetilde{Q}_j^k$ lie in the grid $\mathcal{D}^\alpha$ and have nonempty intersection (they contain B_i^{k+2}), so that one of these cubes must be contained in the other. However, if $\widetilde{Q}_j^k \subset G_r^{\alpha,t+1}$, then $3Q_j^k \subset \hat{Q}_j^k \subset \widetilde{Q}_j^k$ implies $\mathcal{A}(Q_j^k) \subset G_r^{\alpha,t+1}$, which contradicts $(k, j) \in \mathbb{I}_s^{\alpha,t}$. Therefore we must have $G_r^{\alpha,t+1} \subset \widetilde{Q}_j^k$ as asserted in (4-65).

So to see that Claim 4.14 holds, suppose that $\mathcal{A}(Q_j^k) = G_s^{\alpha,t}$ and $B_i^{k+2} \subset G_r^{\alpha,t+1}$ with an associated cube $\widetilde{Q}_j^k$ as in (4-65). Then by (4-65) and (4-57) the side length $\ell(Q_j^k)$ of Q_j^k satisfies

$$\ell(Q_j^k) = \frac{1}{N}\ell(N Q_j^k) \geq \frac{1}{N}\ell(\widetilde{Q}_j^k) \geq \frac{1}{N}\ell(G_r^{\alpha,t+1}). \tag{4-66}$$

Also, if B_ℓ^k is *any* Whitney cube at level k that is contained in $G_r^{\alpha,t+1}$, then by (4-65) and (4-57) we have

$$B_\ell^k \subset G_r^{\alpha,t+1} \subset \widetilde{Q}_j^k \subset N Q_j^k,$$

so that Lemma 4.11 shows that B_ℓ^k and Q_j^k have comparable side lengths:

$$\ell(B_\ell^k) \approx \ell(Q_j^k). \tag{4-67}$$

Moreover, if $B_{\ell'}^{k'}$ is any Whitney cube at level $k' < k$ that is contained in $G_r^{\alpha,t+1}$, then there is some Whitney cube B_ℓ^k at level k such that $B_\ell^k \subset B_{\ell'}^{k'}$. Thus we have the containments $B_\ell^k \subset B_{\ell'}^{k'} \subset N Q_j^k$, and it follows from (4-67) that

$$\ell(B_{\ell'}^{k'}) \approx \ell(Q_j^k). \tag{4-68}$$

Now momentarily *fix* k_0 such that there is a cube $B_i^{k_0+2}$ satisfying the conditions in Claim 4.14. Then *all* of the cubes B_ℓ^{k+2} that arise in Claim 4.14 with $k \le k_0 - 2$ satisfy

$$\ell(B_\ell^{k+2}) \approx \ell(Q_j^{k_0}) \ge \frac{1}{N}\ell(G_r^{\alpha,t+1}).$$

Thus all of the cubes B_ℓ^{k+2} with $k \le k_0$, except perhaps those with $k \in \{k_0 - 1, k_0\}$, have side lengths bounded below by $c\ \ell(G_r^{\alpha,t+1})$, which bounds the number of possible locations for these cubes by a dimensional constant. However, those cubes $B_i^{k_0+1}$ at level $k_0 + 1$ are pairwise disjoint, as are those cubes $B_i^{k_0+2}$ at level k_0+2. Consequently, we can apply the "bounded occurrence of cubes" to show that the sum in Claim 4.14, when restricted to $k \le k_0$, is bounded by a constant C independent of k_0. Since k_0 is arbitrary, this completes the proof of Claim 4.14. □

As a result of Claim 4.14, for those i in either $\mathcal{I}_s^t$ or $\mathcal{J}_s^t$ that satisfy Case(b), we will be able to apply below the Poisson argument used to estimate term $II_s^t(2)$ in (4-40) above.

We now further split the sum over $i \in \mathcal{J}_s^t$ in term $IV_s^t(2)$ into two sums according to Case(a) and Case(b) above:

$$\begin{aligned} IV_s^t(2) \le &\sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r\in\mathcal{H}_s^{\alpha,t}} \sum_{\substack{i\in\mathcal{J}_s^t\\ i\in\text{Case(a)}}} \int_{G_r^{\alpha,t+1}\cap B_i^{k+2}} (L^*\chi_{E_j^k\cap T_{\ell(r)}}\omega) b_r\sigma \Big|^p \\ &+ \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r\in\mathcal{H}_s^{\alpha,t}} \sum_{\substack{i\in\mathcal{J}_s^t\\ i\in\text{Case(b)}}} \int_{G_r^{\alpha,t+1}\cap B_i^{k+2}} (L^*\chi_{E_j^k\cap T_{\ell(r)}}\omega) b_r\sigma \Big|^p \\ &\equiv IV_s^t(2)[\text{a}] + IV_s^t(2)[\text{b}]. \end{aligned} \tag{4-69}$$

We apply the definition of Case(b) and (4-61), to decompose $IV_s^t(2)[\text{b}]$ as follows:

$$\begin{aligned} IV_s^t(2)[\text{b}] &= \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r\in\mathcal{H}_s^{\alpha,t}} \sum_{\substack{i\in\mathcal{J}_s^t\\ B_i^{k+2}\subset G_r^{\alpha,t+1}}} \int_{B_i^{k+2}} (L^*\chi_{E_j^k\cap T_{\ell(r)}}\omega) b_r\sigma \Big|^p \\ &\le \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r\in\mathcal{H}_s^{\alpha,t}} \sum_{\substack{i\in\mathcal{J}_s^t\\ B_i^{k+2}\subset G_r^{\alpha,t+1}}} \Big(\int_{B_i^{k+2}} (L^*\chi_{E_j^k\cap T_{\ell(r)}}\omega)\sigma\Big) \times c_i^{k+2}(|A_r^{\alpha,t+1}| + \mathsf{A}_i^{k+2}) \Big|^p \\ &\quad + \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r\in\mathcal{H}_s^{\alpha,t}} \sum_{\substack{i\in\mathcal{J}_s^t\\ B_i^{k+2}\subset G_r^{\alpha,t+1}}} \boldsymbol{P}(B_i^{k+2}, \chi_{E_j^k\cap T_{\ell(r)}}\omega) \int_{B_i^{k+2}} |b_r|\sigma \Big|^p \\ &= V_s^t(1) + V_s^t(2). \end{aligned} \tag{4-70}$$

4.14.1. *The bound for* $V(2)$. We claim that

$$\sum_{(t,s)\in\mathbb{L}^\alpha} V_s^t(2) \le C\gamma^{2p}\mathfrak{M}_*^p \|f\|_{L^p(\sigma)}^p. \tag{4-71}$$

Here, $\mathfrak{M}_*$ is defined in (4-3), and $V_s^t(2)$ is defined in (4-70).

Proof. The estimate for term $V_s^t(2)$ is similar to that of $II_s^t(2)$ above (see (4-40)), except that this time we use Claim 4.13 to handle a complication arising from the extra sum in the cubes B_i^{k+2}. We define

$$\boldsymbol{P}_j^k(\mu) \equiv \sum_{\ell} \sum_{\substack{r \in \mathcal{H}_s^{\alpha,t} \\ \ell(r)=\ell}} \sum_{\substack{i \in \mathcal{I}_s^t \\ B_i^{k+2} \subset G_r^{\alpha,t+1}}} \boldsymbol{P}(B_i^{k+2}, \chi_{E_j^k \cap T_\ell} \mu) \chi_{B_i^{k+2}}. \tag{4-72}$$

We observe that by Claim 4.14 the sum of these operators satisfies

$$\sum_{(k,j) \in \mathbb{I}_s^{\alpha,t}} \boldsymbol{P}_j^k(\mu) \le C \chi_{G_s^{\alpha,t}} \mathcal{M}(\chi_{G_s^{\alpha,t}} \mu), \tag{4-73}$$

and hence the analogue of (4-44) holds with $\boldsymbol{P}_j^k$ defined as above:

$$\Big\| \chi_{G_s^{\alpha,t}} \sum_{(k,j) \in \mathbb{I}_s^{\alpha,t}} (\boldsymbol{P}_j^k)^*(|h|\sigma) \Big\|_{L^p(w)} \le C \mathfrak{M}_* \| \chi_{G_s^{\alpha,t}} h \|_{L^p(\sigma)}. \tag{4-74}$$

For our use below, we note that this conclusion holds independent of the assumption, imposed in (4-72), that $i \in \mathcal{I}_s^t$.

With this notation, the summands in the definition of $V_s^t(2)$, as given in (4-70), are

$$\begin{aligned} &\sum_{\ell} \sum_{\substack{r \in \mathcal{H}_s^{\alpha,t} \\ \ell(r)=\ell}} \sum_{\substack{i \in \mathcal{I}_s^t \\ B_i^{k+2} \subset G_r^{\alpha,t+1}}} \boldsymbol{P}(B_i^{k+2}, \chi_{E_j^k \cap T_\ell} \omega) \Big(\int_{B_i^{k+2}} \sigma \Big) \Big(\frac{1}{|B_i^{k+2}|_\sigma} \int_{B_i^{k+2}} |f|\sigma \Big) \\ &\quad \le \gamma^{t+2} \sum_{\ell} \int \sum_{\substack{r \in \mathcal{H}_s^{\alpha,t} \\ \ell(r)=\ell}} \sum_{\substack{i \in \mathcal{I}_s^t \\ B_i^{k+2} \subset G_r^{\alpha,t+1}}} \boldsymbol{P}(B_i^{k+2}, \chi_{E_j^k \cap T_\ell} \omega) \chi_{B_i^{k+2}} \sigma \quad (\text{since } i \in \mathcal{I}_s^t) \\ &\quad \le \gamma^{t+2} \int_{G_s^{\alpha,t}} \boldsymbol{P}_j^k(\omega) \sigma = \gamma^{t+2} \int_{E_j^k} (\boldsymbol{P}_j^k)^*(\chi_{G_s^{\alpha,t}} \sigma) \omega. \end{aligned} \tag{4-75}$$

We then have from (4-70) and (4-75) by the argument for term $II_s^t(2)$,

$$\begin{aligned} \sum_{(t,s) \in \mathbb{L}^\alpha} V_s^t(2) &\le C \gamma^{2p} \sum_{(t,s) \in \mathbb{L}^\alpha} \gamma^{pt} \sum_{(k,j) \in \mathbb{I}_s^{\alpha,t}} R_j^k \Big| \int_{Q_j^k} (\boldsymbol{P}_j^k)^*(\chi_{G_s^{\alpha,t}} \sigma) \, \omega \Big|^p \\ &\le C \gamma^{2p} \sum_{(t,s) \in \mathbb{L}^\alpha} \gamma^{pt} \int \Big| \mathcal{M}_\omega \Big(\chi_{G_s^{\alpha,t}} \sum_{(\ell,i) \in \mathbb{I}_s^{\alpha,t}} (\boldsymbol{P}_i^\ell)^*(\chi_{G_s^{\alpha,t}} \sigma) \Big) \Big|^p \omega \\ &\le C \gamma^{2p} \sum_{(t,s) \in \mathbb{L}^\alpha} \gamma^{pt} \int_{G_s^{\alpha,t}} \Big(\sum_{(\ell,i) \in \mathbb{I}_s^{\alpha,t}} (\boldsymbol{P}_i^\ell)^*(\chi_{G_s^{\alpha,t}} \sigma) \Big)^p \omega \\ &\le C \gamma^{2p} \mathfrak{M}_*^p \sum_{(t,s) \in \mathbb{L}^\alpha} \gamma^{pt} \sum_{\ell} |G_s^{\alpha,t}|_\sigma \le C \gamma^{2p} \mathfrak{M}_*^p \int |f|^p \, \sigma. \end{aligned}$$

In last lines we are using the boundedness (1-17) of the maximal operator. □

We will use the same method to treat term $V(1)$ and term $VI(1)$ below, and we postpone the argument for now.

4.14.2. *The bound for $IV(2)$[a].* We turn to the term defined in (4-69). In Case(a) the cubes B_i^{k+2} satisfy

$$G_r^{\alpha,t+1} \subset B_i^{k+2} \quad \text{whenever } G_r^{\alpha,t+1} \cap B_i^{k+2} \neq \varnothing.$$

and so recalling that $i \in \mathcal{J}_s^t$ and $i \in \text{Case(a)}$, we obtain from (4-62) that

$$\begin{aligned} IV_s^t(2)[\mathrm{a}] &= \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{\substack{i\in\mathcal{J}_s^t\\ i\in\text{Case(a)}}} \sum_{r:G_r^{\alpha,t+1}\subset B_i^{k+2}} \int_{G_r^{\alpha,t+1}} (L^*\chi_{E_j^k}\omega) b_r\sigma \Big|^p \\ &\le C \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{i\in\text{Case(a)}} \sum_{r:G_r^{\alpha,t+1}\subset B_i^{k+2}} \boldsymbol{P}(G_r^{\alpha,t+1}, \chi_{E_j^k}\omega) \int_{G_r^{\alpha,t+1}} |f|\sigma \Big|^p \\ &\le C\gamma^{p(t+2)} \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r:G_r^{\alpha,t+1}\subset 3Q_j^k} \boldsymbol{P}(G_r^{\alpha,t+1}, \chi_{E_j^k}\omega) |G_r^{\alpha,t+1}|_\sigma \Big|^p. \end{aligned}$$

But this last sum is identical to the estimate for the term $II_s^t(2)$ used in (4-45) above. The estimate there thus gives

$$\sum_{(t,s)\in\mathbb{L}^\alpha} IV_s^t(2)[a] \le C\gamma^{2p}\mathfrak{M}_*^p \sum_{(t,s)\in\mathbb{L}^\alpha} \gamma^{pt} |G_s^{\alpha,t}|_\sigma \le C\gamma^{2p}\mathfrak{M}_*^p \int |f|^p\sigma, \tag{4-76}$$

which is the desired estimate.

4.14.3. *The decomposition of $IV(1)$.* This term is the first term on the right hand side of (4-58). Recall that for $i \in \mathcal{J}_s^t$ we have $i \in \text{Case(b)}$ and so $B_i^{k+2} \subset G_r^{\alpha,t+1} \subset T_{\ell(r)}$ for some $r \in \mathcal{K}_s^{\alpha,t}$. From (4-63) we also have $B_i^{k+2} \subset \widetilde{Q}_j^k$. To estimate $IV_s^t(1)$ in (4-58), we again apply (4-61) to be able to write

$$\begin{aligned} IV_s^t(1) &\le C \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big(\sum_\ell \sum_{\substack{i\in\mathcal{J}_s^t\\ B_i^{k+2}\subset T_\ell\cap\widetilde{Q}_j^k}} \Big(\int_{B_i^{k+2}} |L^*\chi_{E_j^k\cap T_\ell}\omega|\sigma \Big) \mathsf{A}_i^{k+2} \Big)^p \\ &\quad + C \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big(\sum_\ell \sum_{\substack{i\in\mathcal{J}_s^t\\ B_i^{k+2}\subset T_\ell\cap\widetilde{Q}_j^k}} \boldsymbol{P}(B_i^{k+2}, \chi_{E_j^k\cap T_\ell}\omega) \int_{B_i^{k+2}} |f|\sigma \Big)^p \qquad (4\text{-}77) \\ &= VI_s^t(1) + VI_s^t(2). \end{aligned}$$

We can dominate the averages on B_i^{k+2} of the bad function b_r by $\mathsf{A}_i^{k+2} + |A_r^{\alpha,t+1}| \le 2\mathsf{A}_i^{k+2}$, since in this case $i \in \mathcal{J}_s^t$ (see (4-59)), and this implies that the average of $|b_r| = |f - A_r^{\alpha,t+1}|$ over the cube B_i^{k+2} is dominated by

$$\mathsf{A}_i^{k+2} + |A_r^{\alpha,t+1}| \le \mathsf{A}_i^{k+2} + \gamma^{t+2} < 2\mathsf{A}_i^{k+2}.$$

4.14.4. *The bound for VI*(2). We claim that

$$VI_s^t(2) \le C\mathfrak{M}_*^p \sum_{k,i}^{s,t,\mathcal{J}} |B_i^{k+2}|_\sigma (\mathsf{A}_i^{k+2})^p. \tag{4-78}$$

Here, the sum on the right is over all pairs of integers $k, i \in \mathcal{J}_s^t$ such that $B_i^{k+2} \subset T_\ell \cap \widetilde{Q}_j^k$ for some ℓ, j with $(k, j) \in \mathbb{I}_s^{\alpha,t}$. (Below, we will need a similar sum, with the condition $i \in \mathcal{J}_s^t$ replaced by $i \in \mathcal{J}_s^t$ and $i \in$ Case(b).) This is a provisional bound, one that requires additional combinatorial arguments in Section 4.14.7.

Proof. The term $VI_s^t(2)$ can be handled the same way as the term $V_s^t(2)$ (see (4-71)), with these two changes. First, in the definition of $\mathbf{P}_j^k$, we replace $\mathcal{J}_s^t$ by $\mathcal{J}_s^t$, and second, we use the function

$$h = \sum_{k,i}^{s,t,\mathcal{J}} \mathsf{A}_i^{k+2} \chi_{B_i^{k+2}}$$

in (4-74). That argument then obtains

$$\Big\| \chi_{G_s^{\alpha,t}} \sum_{k,j} (\boldsymbol{P}_j^k)^* (\chi_{G_s^{\alpha,t}} h\sigma) \Big\|_{L^p(\omega)}^p \le C\mathfrak{M}_*^p \sum_{k,i}^{s,t,\mathcal{J}} |B_i^{k+2}|_\sigma (\mathsf{A}_i^{k+2})^p. \tag{4-79}$$

Here we are using the bounded overlap of the cubes B_i^{k+2} given in Claim 4.13, along with the fact recorded in (4-64) that for fixed $(k+2, i)$, only a bounded number of j satisfy $B_i^{k+2} \subset \widetilde{Q}_j^k$. Claim 4.13 applies in this setting, as we are in a subcase of the analysis of IV. We then use the universal maximal function bound (2-2).

$$\begin{aligned} VI_s^t(2) &= \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big(\sum_\ell \sum_{\substack{i\in\mathcal{J}_s^t \\ B_i^{k+2}\subset T_\ell\cap\widetilde{Q}_j^k}} \boldsymbol{P}(B_i^{k+2}, \chi_{E_j^k\cap T_\ell}\omega) \Big(\int_{B_i^{k+2}} \sigma \Big) \mathsf{A}_i^{k+2} \Big)^p \\ &= C \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big| \int_{Q_j^k} (\boldsymbol{P}_j^k)^*(h\sigma)\omega \Big|^p \\ &\le C \int \Big(\mathcal{M}_\omega \Big(\chi_{G_s^{\alpha,t}} \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} (\boldsymbol{P}_j^k)^* (\chi_{G_s^{\alpha,t}} h\sigma) \Big) \Big)^p \omega \\ &\le C \int \Big(\chi_{G_s^{\alpha,t}} \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} (\boldsymbol{P}_j^k)^* (\chi_{G_s^{\alpha,t}} h\sigma) \Big)^p \omega. \end{aligned}$$

In view of (4-79), this completes the proof of the provisional estimate (4-78). □

4.14.5. *The bound for VI*(1). Recall the definition of $VI(1)$ from (4-77), and also from (4-63) the fact that $B_i^{k+2} \subset \widetilde{Q}_j^k$ whenever $B_i^{k+2} \cap \widetilde{Q}_j^k \ne \varnothing$. We claim that

$$VI_s^t(1) \le C\mathfrak{T}_*^p \sum_{k,i}^{s,t,\mathcal{J}} |B_i^{k+2}|_\sigma (\mathsf{A}_i^{k+2})^p. \tag{4-80}$$

The notation here is as in (4-78), but since $i \in \mathcal{I}_s^t$ implies i belongs to Case(b), the sum over the right is over $k, i \in \mathcal{I}_s^t$ such that $B_i^{k+2} \subset G_r^{\alpha,t+1} \subset T_{\ell(r)} \cap \widetilde{Q}_j^k$, for some integers j, r, with $(k, j) \in \mathbb{I}_s^{\alpha,t}$. As with (4-78), this is a provisional estimate.

Proof. We first estimate the sum in i inside term $VI_s^t(1)$. Recall that the sum in i is over those i such that $B_i^{k+2} \subset G_r^{\alpha,t+1} \subset T_\ell$ for some r with $\ell = \ell(r)$, and where $\{T_\ell\}_\ell$ is the set of maximal cubes in the collection $\{3G_r^{\alpha,t+1} : r \in \mathcal{K}_s^{\alpha,t}\}$. See the discussion at (4-35), and (4-50). We will write $\ell(i) = \ell(r)$ when $B_i^{k+2} \subset G_r^{\alpha,t+1}$. It is also important to note that the sum in i deriving from term IV_s^t is also restricted to those i such that $B_i^{k+2} \subset \widetilde{Q}_j^k$ by (4-63), so that altogether, $B_i^{k+2} \subset T_\ell \cap \widetilde{Q}_j^k$. We have

$$\begin{aligned}\Big|\sum_i \Big(\int_{B_i^{k+2}} |L^* \chi_{E_j^k \cap T_{\ell(i)}} \omega|\sigma\Big) \mathsf{A}_i^{k+2}\Big|^p &\le \sum_i |B_i^{k+2}|_\sigma (\mathsf{A}_i^{k+2})^p \Big(\sum_i |B_i^{k+2}|_\sigma^{1-p'} \Big(\int_{B_i^{k+2}} |L^* \chi_{E_j^k \cap T_{\ell(i)}} \omega|\sigma\Big)^{p'}\Big)^{p-1} \\ &\le \sum_i |B_i^{k+2}|_\sigma (\mathsf{A}_i^{k+2})^p \Big(\sum_i \int_{B_i^{k+2}} |L^* \chi_{E_j^k \cap T_{\ell(i)}} \omega|^{p'}\sigma\Big)^{p-1} \\ &\le C \sum_i |B_i^{k+2}|_\sigma (\mathsf{A}_i^{k+2})^p \Big(\sum_\ell \sum_{i:\ell(i)=\ell} \int_{B_i^{k+2}} |L^* \chi_{E_j^k \cap T_{\ell(i)}} \omega|^{p'}\sigma\Big)^{p-1}.\end{aligned}$$

Now we will apply the form (2-11) of (1-14) with $g = \chi_{E_j^k \cap T_\ell}$ and Q chosen to be either T_ℓ or $\widetilde{Q}_j^k$ depending on the relative positions of T_ℓ and $\widetilde{Q}_j^k$. Since T_ℓ is a triple of a cube in the grid $\mathcal{D}^\alpha$ and $\widetilde{Q}_j^k$ is a cube in the grid $\mathcal{D}^\alpha$, we must have either

$$\widetilde{Q}_j^k \subset T_\ell \quad \text{or} \quad T_\ell \subset 3\widetilde{Q}_j^k.$$

If $\widetilde{Q}_j^k \subset T_\ell$ we choose Q in (2-11) to be $\widetilde{Q}_j^k$ and note that by bounded overlap of Whitney cubes, there are only a bounded number of such cases. If on the other hand $T_\ell \subset 3\widetilde{Q}_j^k$, then we choose Q to be T_ℓ and note that the cubes T_ℓ have bounded overlap. This gives

$$\sum_\ell \sum_{i:\ell(i)=\ell} \int_{B_i^{k+2}} |L^* \chi_{E_j^k \cap T_{\ell(i)}} \omega|^{p'}\sigma \lesssim \mathfrak{T}_*^p |3\widetilde{Q}_j^k|_\omega,$$

and hence

$$\Big|\sum_i \Big(\int_{B_i^{k+2}} |L^* \chi_{E_j^k \cap T_{\ell(i)}} \omega|\sigma\Big) \mathsf{A}_i^{k+2}\Big|^p \le C \mathfrak{T}_*^p \sum_i |B_i^{k+2}|_\sigma (\mathsf{A}_i^{k+2})^p |N Q_j^k|_\omega^{p-1},$$

since $3\widetilde{Q}_j^k \subset N Q_j^k$ by (4-56). With this we obtain

$$\begin{aligned} VI_s^t(1) &\le C \mathfrak{T}_*^p \sum_{(k,j)\in \mathbb{I}_s^{\alpha,t}} R_j^k \sum_{i \in \mathcal{I}_s^t} |B_i^{k+2}|_\sigma (\mathsf{A}_i^{k+2})^p |N Q_j^k|_\omega^{p-1} \\ &\le C \mathfrak{T}_*^p {\sum_{k,i}}^{*\,s,t,\mathcal{I}} |B_i^{k+2}|_\sigma (\mathsf{A}_i^{k+2})^p, \end{aligned} \tag{4-81}$$

where we are using $R_j^k |NQ_j^k|_\omega^{p-1} \leq 1$ and (4-64) in the final line. □

4.14.6. *The bound for* $V(1)$. We will use the same method as in the estimate for term $VI(1)$ above to obtain

$$\sum_{(t,s)\in\mathbb{L}^\alpha} V_s^t(1) \leq C\mathfrak{T}_*^p \gamma^{2p} \|f\|_{L^p(\sigma)}^p. \tag{4-82}$$

Recall from (4-70) that $V_s^t(1)$ is given by

$$\sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} R_j^k \Big| \sum_{r\in\mathcal{H}_s^{\alpha,t}} \sum_{\substack{i\in\mathcal{J}_s^t \\ B_i^{k+2}\subset G_r^{\alpha,t+1}}} \Big(\int_{B_i^{k+2}} (L^* \chi_{E_j^k \cap T_{\ell(r)}} \omega)\sigma\Big) c_i^{k+2} (|A_r^{\alpha,t+1}| + \mathsf{A}_i^{k+2}) \Big|^p.$$

The main difference here, as opposed to the previous estimate, is that $i \in \mathcal{J}_s^t$ rather than in $\mathcal{I}_s^t$; see (4-59). As a result, we have the estimate

$$|A_r^{\alpha,t+1}| + \mathsf{A}_i^{k+2} \lesssim \gamma^{t+2}, \tag{4-83}$$

instead of $|A_r^{\alpha,t+1}| + \mathsf{A}_i^{k+2} \lesssim \mathsf{A}_i^{k+2}$, which holds when $i \in \mathcal{I}_s^t$.

Proof of (4-82). We follow the argument leading up to and including (4-81) in the estimate for term $VI(1)$ above, but using instead (4-83). The result is as below, where we are using the notation of (4-78), with the condition $i \in \mathcal{I}_s^t$ replaced by $i \in \mathcal{J}_s^t$ and $i \in Case(b)$, and so we use an asterisk and $\mathcal{J}$ in the notation below.

$$V_s^t(1) \leq C\mathfrak{T}_*^p \,{}^*\!\!\sum_{k,i}^{s,t,\mathcal{J}} |B_i^{k+2}|_\sigma (\gamma^{t+2})^p.$$

Now we collect those cubes B_i^{k+2} that lie in a given cube $G_r^{\alpha,t+1}$ and write the right hand side above as a constant times

$$\mathfrak{T}_*^p \gamma^{(t+2)p} \sum_{r\in\mathcal{H}_s^{\alpha,t}} {}^*\!\!\sum_{k,i}^{s,t,\mathcal{J}} |B_i^{k+2}|_\sigma := \mathfrak{T}_*^p \gamma^{(t+2)p} \sum_{r\in\mathcal{H}_s^{\alpha,t}} \mathcal{S}_{s,r}^{\alpha,t}.$$

By Claim 4.13, which applies as we are in a subcase of IV, we have $\mathcal{S}_{s,r}^{\alpha,t} \leq C|G_r^{\alpha,t+1}|_\sigma$, and it follows that

$$V_s^t(1) \leq C\mathfrak{T}_*^p \gamma^{(t+2)p} \sum_{r\in\mathcal{H}_s^{\alpha,t}} |G_r^{\alpha,t+1}|_\sigma \leq C\mathfrak{T}_*^p \gamma^{(t+2)p} |G_s^{\alpha,t}|_\sigma,$$

and hence from (4-18) that

$$\sum_{(t,s)\in\mathbb{L}^\alpha} V_s^t(1) \leq C\mathfrak{T}_*^p \gamma^{2p} \sum_{(t,s)\in\mathbb{L}^\alpha} \gamma^{tp} |G_s^{\alpha,t}|_\sigma \leq C\mathfrak{T}_*^p \gamma^{2p} \|f\|_{L^p(\sigma)}^p.$$ □

4.14.7. *The final combinatorial arguments.* Our final estimate in the proof of (4-52) is to dominate by $C\int|f|^p\,d\sigma$ the sum of the right hand sides of (4-78) and (4-80) over $(t,s)\in\mathbb{L}^\alpha$, namely

$$\sum_{(t,s)\in\mathbb{L}^\alpha} \sum_{k,i}^{s,t,\mathcal{J}} \leq C\int |f|^p\,d\sigma. \tag{4-84}$$

The proof of (4-84) will require combinatorial facts related to the principal cubes, and the definition of the collection $\mathbb{G}^\alpha$ in (4-27). Also essential is the implementation of the shifted dyadic grids. We now detail the arguments.

Definition 4.15. We say that a cube B_i^{k+2} satisfying the defining condition in $VI_s^t(1)$, namely

there is $(k, j) \in \mathbb{I}_s^{\alpha,t} = \mathbb{G}^\alpha \cap \mathbb{H}_s^{\alpha,t}$ such that
$B_i^{k+2} \subset \widetilde{Q}_j^k$ and $B_i^{k+2} \subset$ some $G_r^{\alpha,t+1} \subset G_s^{\alpha,t}$ satisfying $\mathsf{A}_i^{k+2} > \gamma^{t+2}$,

is a *final type* cube for the pair $(t, s) \in \mathbb{L}^\alpha$ generated from Q_j^k.

The collection $\mathcal{F}$ of cubes B_i^{k+2} such that B_i^{k+2} is a *final type* cube generated from some Q_j^k with $(k, j) \in \mathbb{I}_s^{\alpha,t}$ for some pair $(t, s) \in \mathbb{L}^\alpha$ satisfies the following three properties:

Property 1. *$\mathcal{F}$ is a nested grid in the sense that given any two distinct cubes in $\mathcal{F}$, either one is strictly contained in the other, or they are disjoint (ignoring boundaries).*

Property 2. *If B_i^{k+2} and $B_{i'}^{k'+2}$ are two distinct cubes in $\mathcal{F}$ with $B_{i'}^{k'+2} \subsetneqq B_i^{k+2}$, and k and k' have the same parity, then*

$$\mathsf{A}_{i'}^{k'+2} > \gamma \mathsf{A}_i^{k+2}.$$

Property 3. *A given cube B_i^{k+2} can occur at most a bounded number of times in the grid $\mathcal{F}$.*

Proof of Properties 1, 2 and 3. Property 1 is obvious from the properties of the dyadic shifted grid $\mathcal{D}^\alpha$. Property 3 follows from the "bounded occurrence of cubes" noted above. So we turn to Property 2. It is this property that prompted the use of the shifted dyadic grids.

Indeed, since $B_{i'}^{k'+2} \subsetneqq B_i^{k+2}$, it follows from the nested property (4-54) that $k' > k$. By Definition 4.15 there are cubes

$$Q_{j'}^{k'} \quad \text{and} \quad Q_j^k \quad \text{satisfying} \quad B_{i'}^{k'+2} \subset \widetilde{Q}_{j'}^{k'} \quad \text{and} \quad B_i^{k+2} \subset \widetilde{Q}_j^k,$$

and also cubes $G_{s'}^{\alpha,t'} \subset G_s^{\alpha,t}$ such that $(k', j') \in \mathbb{I}_{s'}^{\alpha,t'}$ and $(k, j) \in \mathbb{I}_s^{\alpha,t}$ with $(t', s'), (t, s) \in \mathbb{L}^\alpha$, so that in particular,

$$\widetilde{Q}_{j'}^{k'} \subset G_{s'}^{\alpha,t'} \quad \text{and} \quad \widetilde{Q}_j^k \subset G_s^{\alpha,t}.$$

Now $k' \geq k+2$ and in the extreme case where $k' = k+2$, it follows that the $\mathcal{D}^\alpha$-cube $\widetilde{Q}_{j'}^{k'}$ is one of the cubes B_ℓ^{k+2}, so in fact it must be B_i^{k+2} since $B_{i'}^{k'+2} \subset B_i^{k+2}$. Thus we have

$$B_{i'}^{k'+2} \subset \widetilde{Q}_{j'}^{k'} = B_i^{k+2}.$$

In the general case $k' \geq k+2$ we have instead

$$B_{i'}^{k'+2} \subset \widetilde{Q}_{j'}^{k'} \subset B_i^{k+2}.$$

Now $\mathsf{A}_i^{k+2} > \gamma^{t+2}$ by Definition 4.15, and so there is $t_0 \geq t+2$ determined by the condition

$$\gamma^{t_0} < \mathsf{A}_i^{k+2} \leq \gamma^{t_0+1}, \tag{4-85}$$

and also s_0 such that

$$B_i^{k+2} \subset G_{s_0}^{\alpha,t_0} \subset G_s^{\alpha,t},$$

where the label (t_0, s_0) need not be principal. Combining inclusions we have

$$\widetilde{Q}_{j'}^{k'} \subset B_i^{k+2} \subset G_{s_0}^{\alpha,t_0},$$

and since $(k', j') \in \mathbb{I}_{s'}^{\alpha,t'}$, we obtain $G_{s'}^{\alpha,t'} \subset G_{s_0}^{\alpha,t_0}$. Since $(t', s') \in \mathbb{L}^\alpha$ is a principal label, we have the key property that

$$t' \geq t_0. \tag{4-86}$$

Indeed, if $G_{s'}^{\alpha,t'} = G_{s_0}^{\alpha,t_0}$ then (4-86) holds because $(t', s') \in \mathbb{L}^\alpha$ is a principal label, and otherwise the maximality of $G_{s'}^{\alpha,t'}$ shows that

$$\gamma^{t_0} < \frac{1}{|G_{s_0}^{\alpha,t_0}|_\sigma} \int_{G_{s_0}^{\alpha,t_0}} |f| \, d\sigma \leq \gamma^{t'+1} \quad \text{that is,} \quad t_0 < t' + 1.$$

Thus using (4-86) and (4-85) we obtain Property 2:

$$\mathsf{A}_{i'}^{k'+2} > \gamma^{t'+2} \geq \gamma^{t_0+2} \geq \gamma \mathsf{A}_i^{k+2}.$$ □

Proof of (4-84). Now for $Q = B_i^{k+2} \in \mathcal{F}$ set

$$\mathsf{A}(Q) = \frac{1}{|Q|_\sigma} \int_Q |f|\sigma = \mathsf{A}_i^{k+2} = \frac{1}{|B_i^{k+2}|_\sigma} \int_{B_i^{k+2}} |f|\sigma.$$

With the three properties above we can now prove (4-84) as follows. Recall that in term $IV(1)$ we have $i \in \mathcal{I}_s^t$ which implies B_i^{k+2} satisfies Case(b). In the display below by $\sum_i^*$ we mean the sum over i such that B_i^{k+2} is contained in some $G_r^{\alpha,t+1} \subset G_s^{\alpha,t}$, and also in some $\widetilde{Q}_j^k$ with $(k, j) \in \mathbb{I}_s^{\alpha,t}$, and satisfying $\mathsf{A}_i^{k+2} > 2^{t+2}$. The left side of (4-84) is dominated by

$$\begin{aligned}
\sum_{(t,s)\in\mathbb{L}^\alpha} \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} \sum_i^* |B_i^{k+2}|_\sigma (\mathsf{A}_i^{k+2})^p &= \sum_{Q\in\mathcal{F}} |Q|_\sigma \mathsf{A}(Q)^p = \sum_{Q\in\mathcal{F}} |Q|_\sigma \Big(\frac{1}{|Q|_\sigma} \int_Q |f|\sigma \Big)^p \\
&= \int_{\mathbb{R}^n} \sum_{Q\in\mathcal{F}} \chi_Q(x) \Big(\frac{1}{|Q|_\sigma} \int_Q |f|\sigma \Big)^p \, d\sigma(x) \\
&\leq C \int_{\mathbb{R}^n} \sup_{x\in Q: Q\in\mathcal{F}} \Big(\frac{1}{|Q|_\sigma} \int_Q |f|\sigma \Big)^p \, d\sigma(x) \\
&\leq C \int_{\mathbb{R}^n} \mathcal{M}_\sigma^\alpha f(x)^p \sigma(dx) \leq C \int_{\mathbb{R}^n} |f(x)|^p \, d\sigma(x),
\end{aligned}$$

where the second to last line follows since for fixed $x \in \mathbb{R}^n$, the sum

$$\sum_{Q\in\mathcal{F}} \chi_Q(x) \Big(\frac{1}{|Q|_\sigma} \int_Q |f|\sigma \Big)^p$$

is supergeometric by Properties 1, 2 and 3 above, that is, for any two distinct cubes Q and Q' in $\mathcal{F}$ each containing x, the ratio of the corresponding values is bounded away from 1, more precisely,

$$\frac{\left(\frac{1}{|Q|_\sigma}\int_Q |f|\sigma\right)^p}{\left(\frac{1}{|Q'|_\sigma}\int_{Q'} |f|\sigma\right)^p} \notin [\gamma^{-p}, \gamma^p) \quad \text{for } \gamma \geq 2.$$

This completes the proof of (4-84). □

5. The proof of Theorem 1.10 on the strongly maximal Hilbert transform

To prove Theorem 1.10 we first show that in the proof of Theorem 1.9 above, we can replace the use of the dual maximal function inequality (1-17) with the dual weighted Poisson inequality (5-5) defined below. After that we will show that in the case of standard kernels satisfying (1-9) with $\delta(s) = s$ in dimension $n = 1$, the dual weighted Poisson inequality (5-5) is implied by the *half-strengthened* A_p condition

$$\left(\int_{\mathbb{R}} \left(\frac{|Q|}{|Q| + |x - x_Q|}\right)^{p'} d\sigma(x)\right)^{1/p'} \left(\int_Q d\omega(x)\right)^{1/p} \leq \mathcal{A}_p(\omega, \sigma)|Q|, \tag{5-1}$$

for all intervals Q, together with the dual pivotal condition (5-2) of Nazarov, Treil and Volberg [2010], namely that

$$\sum_{r=1}^{\infty} |Q_r|_\sigma \mathsf{P}(Q_r, \chi_{Q_0}\omega)^{p'} \leq \mathfrak{C}_*^{p'} |Q_0|_\omega, \tag{5-2}$$

holds for all decompositions of an interval Q_0 into a union of pairwise disjoint intervals $Q_0 = \bigcup_{r=1}^{\infty} Q_r$. We will assume $1 < p \leq 2$ for this latter implication. Finally, for $p > 2$, we show that (5-5) is implied by (5-1), (5-2) and the Poisson condition (1-24).

It follows from work in [Nazarov et al. 2010] and [Lacey et al. 2011] that the strengthened A_2 condition (5-16) is necessary for the two weight inequality for the Hilbert transform, and also from [Lacey et al. 2011] that the dual pivotal condition (5-2) is necessary for the dual testing condition

$$\int_Q T(\chi_Q\omega)^2\, d\sigma \leq C \int_Q d\omega,$$

for T when $p = 2$ and σ is doubling. We show below that these results extend to $1 < p < \infty$. A slightly weaker result was known earlier from work of Nazarov, Treil and Volberg — namely that the pivotal conditions are necessary for the Hilbert transform H when *both* of the weights ω and σ are doubling and $p = 2$. However, [Lacey et al. 2011] gives an example that shows that (5-2) is *not* in general necessary for boundedness of the Hilbert transform T when $p = 2$.

Finally, we show below that when σ is doubling, the dual weighted Poisson inequality (5-5) is implied by the two weight inequality for the Hilbert transform. Since the Poisson condition (1-24) is a special case of the inequality dual to (5-5), we obtain the necessity of (1-24) for the two weight inequality for the Hilbert transform.

5.1. ***The Poisson inequalities.*** We begin working in $\mathbb{R}^n$ with $1 < p < \infty$. Recall the definition of the Poisson integral $\mathsf{P}(Q, \nu)$ of a measure ν relative to a cube Q, given by

$$\mathsf{P}(Q, \nu) \equiv \sum_{\ell=0}^{\infty} \frac{\delta(2^{-\ell})}{|2^\ell Q|} \int_{2^\ell Q} d|\nu|. \tag{5-3}$$

We will consider here only the standard Poisson integral with $\delta(s) = s$ in (5-3), and so we also suppose that $\delta(s) = s$ in (1-9) above. We now fix a cube Q_0 and a collection of pairwise disjoint subcubes $\{Q_r\}_{r=1}^{\infty}$. Corresponding to these cubes we define a positive linear operator

$$\mathbb{P}\nu(x) = \sum_{r=1}^{\infty} \mathsf{P}(Q_r, \nu)\chi_{Q_r}(x). \tag{5-4}$$

We wish to obtain *sufficient* conditions for the following "dual" weighted Poisson inequality,

$$\int_{\mathbb{R}^n} \mathbb{P}(f\omega)(x)^{p'}\, d\sigma(x) \leq C \int_{\mathbb{R}^n} f^{p'}\, d\omega(x) \quad \text{for } f \geq 0. \tag{5-5}$$

uniformly in Q_0 and pairwise disjoint subcubes $\{Q_r\}_{r=1}^{\infty}$. As we will see below, this inequality is necessary for the two weight Hilbert transform inequality when σ is doubling.

The reason for wanting the dual Poisson inequality (5-5) is that in Theorem 1.9 above, we can replace the assumption (1-17) on dual boundedness of the maximal operator $\mathcal{M}$ by the dual Poisson inequality (5-5). Indeed, this will be revealed by simple modifications of the proof of Theorem 1.9 above. In fact (5-5) can replace (1-17) in estimating term $II_s^t(2)$, as well as in the similar estimates for terms $V_s^t(2)$ and $VI_s^t(2)$. We turn now to the proofs of these assertions before addressing the question of sufficient conditions for the dual Poisson inequality (5-5).

5.1.1. *Sufficiency of the dual Poisson inequality.* We begin by demonstrating that the term $II_s^t(2)$ in (4-40) can be handled using the "dual" Poisson inequality (5-5) in place of the maximal inequality (1-17). We are working here in $\mathbb{R}^n$ with $1 < p < \infty$. In fact we claim that

$$\sum_{(t,s)\in\mathbb{L}^\alpha} II_s^t(2) \leq C\gamma^{2p}\mathfrak{P}_*^p \int |f|^p \sigma, \tag{5-6}$$

where $\mathfrak{P}_*$ is the norm of the dual Poisson inequality (5-5) if we take Q_0 and its collection of pairwise disjoint subcubes $\{Q_r\}_{r=1}^{\infty}$ to be $G_s^{\alpha,t}$ and $\{G_r^{\alpha,t+1}\}_{r\in\mathcal{K}_s^{\alpha,t}}$. Now the maximal inequality (1-17) was used in the proof of (4-40) only in establishing (4-43), which says

$$\Big\| \chi_{G_s^{\alpha,t}} \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} \boldsymbol{P}_j^k(|g|\omega) \Big\|_{L^{p'}(\sigma)} \leq C\mathfrak{M}_* \|\chi_{G_s^{\alpha,t}} g\|_{L^{p'}(\omega)},$$

where

$$\boldsymbol{P}_j^k(\mu) \equiv \sum_{r\in\mathcal{K}_s^{\alpha,t}} \boldsymbol{P}(G_r^{\alpha,t+1}, \chi_{E_j^k}\mu)\chi_{G_r^{\alpha,t+1}}.$$

We now note that

$$\sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} \boldsymbol{P}_j^k(|g|\omega) = \sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} \sum_{r\in\mathcal{K}_s^{\alpha,t}} \boldsymbol{P}(G_r^{\alpha,t+1}, \chi_{E_j^k}|g|\omega)\chi_{G_r^{\alpha,t+1}}$$
$$\leq \sum_{r\in\mathcal{K}_s^{\alpha,t}} \boldsymbol{P}(G_r^{\alpha,t+1}, \chi_{G_s^{\alpha,t}}|g|\omega)\chi_{G_r^{\alpha,t+1}} = \mathbb{P}(\chi_{G_s^{\alpha,t}}|g|\omega)(x),$$

which proves

$$\Big\|\chi_{G_s^{\alpha,t}} \sum_{k,j} \boldsymbol{P}_j^k(|g|\omega)\Big\|_{L^{p'}(\sigma)} \leq C\mathfrak{P}_* \|\chi_{G_s^{\alpha,t}} g\|_{L^{p'}(\omega)},$$

which yields (5-6) as before.

The terms $V(2)$ and $VI(2)$ are handled similarly. Indeed, Claim 4.14 yields the following analogue of (4-73):

$$\sum_{(k,j)\in\mathbb{I}_s^{\alpha,t}} \boldsymbol{P}_j^k(\mu) \leq C\chi_{G_s^{\alpha,t}}\mathbb{P}(\chi_{G_s^{\alpha,t}}\mu),$$

from which the arguments above yield both (4-71) and (4-78) with $\mathfrak{M}_*$ replaced by $\mathfrak{P}_*$.

5.1.2. *Sufficient conditions for Poisson inequalities.* We continue to work in $\mathbb{R}^n$ with $1 < p < \infty$. We note that (5-5) can be rewritten

$$\sum_{r=1}^{\infty} |Q_r|_\sigma \mathrm{P}(Q_r, f\omega)^{p'} \leq C \int_{\mathbb{R}^n} f^{p'}\, d\omega \quad \text{for } f \geq 0,$$

and this latter inequality can then be expressed in terms of the Poisson operator $\mathbb{P}_+$ in the upper half space $\mathbb{R}_+^{n+1}$ given by

$$\mathbb{P}_+(f\omega)(x,t) = \int_{\mathbb{R}^n} P_t(x-y) f(y)\, d\omega(y).$$

Indeed, let $Z_r = (x_{Q_r}, \ell(Q_r))$ be the point in $\mathbb{R}_+^{n+1}$ that lies above the center x_{Q_r} of Q_r at a height equal to the side length $\ell(Q_r)$ of Q_r. Define an atomic measure ds in $\mathbb{R}_+^{n+1}$ by

$$ds(x,t) = \sum_{r=1}^{\infty} |Q_r|_\sigma \delta_{Z_r}(x,t). \tag{5-7}$$

Then (5-5) is equivalent to the inequality (this is where we use $\delta(s) = s$),

$$\int_{\mathbb{R}_+^{n+1}} \mathbb{P}_+(f\omega)(x,t)^{p'}\, ds(x,t) \leq C \int_{\mathbb{R}^n} f^{p'}\, d\omega(x) \quad \text{for } f \geq 0. \tag{5-8}$$

We can use [Sawyer 1988, Theorem 2] to characterize this latter inequality in terms of testing conditions over $\mathbb{P}_+$ and its dual $\mathbb{P}_+^*$ given by

$$\mathbb{P}_+^*(gw)(x,t) = \int_{\mathbb{R}_+^{n+1}} P_t(y-x) g(x,t) dw(x,t).$$

Let $\hat{Q}$ denote the cube in $\mathbb{R}_+^{n+1}$ with Q as a face. Then [ibid., Theorem 2] yields the following.

Theorem 5.2. *The Poisson inequality* (5-5) *holds for given data Q_0 and $\{Q_r\}_{r=1}^{\infty}$* if and only if *the measure s in* (5-7) *satisfies*

$$\int_{\mathbb{R}_+^{n+1}} \mathbb{P}_+(\chi_Q\omega)^{p'}\,ds \le C\int_Q d\omega \qquad \text{for all cubes } Q\in\mathcal{D},$$

$$\int_{\mathbb{R}^n} \mathbb{P}_+^*(t^{p'-1}\chi_{\hat{Q}}\,ds)^p\,d\omega \le C\int_{\hat{Q}} t^{p'}\,ds \quad \text{for all cubes } Q\in\mathcal{D}.$$

Note that

$$\int_{\mathbb{R}_+^{n+1}} \mathbb{P}_+(\chi_Q\omega)^{p'}\,ds \approx \sum_{r=1}^{\infty}|Q_r|_\sigma \mathsf{P}(Q_r,\chi_Q\omega)^{p'}.$$

Claim 5.3. *Let $n=1$ and suppose that σ is doubling. First assume that $1<p<\infty$. Then for the special measure s in* (5-7), *inequality* (5-8) *follows from the dual pivotal condition* (5-2), *the Poisson condition* (1-24), *and the half-strengthened A_p condition* (5-1). *Now assume that $1<p\le 2$. Then for the special measure s in* (5-7), *inequality* (5-8) *follows from* (5-2) *and* (5-1) *without* (1-24).

With Claim 5.3 proved, the discussion above yields the following result.

Theorem 5.4. *Let $n=1$ and suppose that σ is doubling. First assume that $1<p<\infty$. Then the dual Poisson inequality* (5-5) *holds uniformly in Q_0 and $\{Q_r\}_{r=1}^{\infty}$ satisfying $\bigcup_{r=1}^{\infty} Q_r\subset Q_0$ provided the half-strengthened A_p condition* (5-1), *the dual pivotal condition* (5-2), *and the Poisson condition* (1-24) *all hold. Now assume that $1<p\le 2$. Then* (5-5) *holds uniformly in Q_0 and $\{Q_r\}_{r=1}^{\infty}$ satisfying $\bigcup_{r=1}^{\infty} Q_r\subset Q_0$ provided* (5-1) *and* (5-2) *both hold.*

Remark 5.5. We do not know if Claim 5.3 and Theorem 5.4 hold without the assumption that σ is doubling, nor do we know if the Poisson condition (1-24) is implied by (5-1) and (5-2) when $p>2$.

We work exclusively in dimension $n=1$ from now on.

5.5.1. *Proof of Claim 5.3.* Instead of applying Theorem 5.2 directly, we first reduce matters to proving that certain $\mathcal{D}^\alpha$-dyadic analogues hold of the two conditions in Theorem 5.2. For $\alpha\in\{0,\frac{1}{3},\frac{2}{3}\}$ we use the following atomic measures ds_α on $\mathbb{R}_+^2$, along with the following $\mathcal{D}^\alpha$-dyadic analogues of the Poisson operators $\mathbb{P}$ and $\mathbb{P}_+$ (with $\delta(s)=s$),

$$\begin{aligned}\mathbb{P}_\alpha^{dy}\nu(x)&=\sum_{r=1}^{\infty}\mathsf{P}_\alpha^{dy}(I_r^\alpha,\nu)\chi_{I_r^\alpha}(x),\quad \mathbb{P}_{+,\alpha}^{dy}\nu(x,t)=\sum_{\substack{Q\in\mathcal{D}^\alpha\\ x\in Q\text{ and }\ell(Q)\ge t}}\frac{t}{\ell(Q)}\frac{1}{|Q|}\int_Q d\nu,\\ ds_\alpha(x,t)&=\sum_{r=1}^{\infty}|I_r^\alpha|_\sigma\delta_{Z_r^\alpha}(x,t),\end{aligned}\tag{5-9}$$

where

(1) the interval I_r^α is chosen to be a *maximal* $\mathcal{D}^\alpha$-interval contained in Q_r with *maximum* length (there can be at most two such intervals, in which case we choose the leftmost one),

(2) the $\mathcal{D}^\alpha$-Poisson integral $\mathsf{P}^{dy}_\alpha(Q,\nu)$ is given by

$$\mathsf{P}^{dy}_\alpha(Q,\nu) \equiv \sum_{\ell=0}^{\infty} \frac{2^{-\ell}}{|Q^{(\ell)}|}\int_{Q^{(\ell)}} d\nu \quad \text{for } Q\in\mathcal{D}^\alpha,$$

where $Q^{(\ell)}$ denotes the ℓ-th dyadic parent of Q in $\mathcal{D}^\alpha$, and

(3) the point $Z^\alpha_r = (x_{I^\alpha_r}, \ell(I^\alpha_r))$ in $\mathbb{R}^2_+$ lies above the center $x_{I^\alpha_r}$ of I^α_r at a height equal to the side length $\ell(I^\alpha_r)$ of I^α_r.

We will use the following dyadic analogue of Theorem 5.2, whose proof is the obvious dyadic analogue of the proof of Theorem 5.2 as given in [Sawyer 1988].

Theorem 5.6. *The $\mathcal{D}^\alpha$-Poisson inequality*

$$\int_{\mathbb{R}^2_+} \mathbb{P}^{dy}_{+,\alpha}(f\omega)^{p'}\, ds_\alpha \le C\int_Q f^{p'}\,d\omega \quad \textit{for } f\ge 0,$$

holds if and only if

$$\begin{aligned}\int_{\mathbb{R}^2_+} \mathbb{P}^{dy}_{+,\alpha}(\chi_Q\omega)^{p'}\, ds_\alpha &\le C\int_Q d\omega \qquad \textit{for all intervals } Q\in\mathcal{D}^\alpha,\\ \int_{\mathbb{R}} (\mathbb{P}^{dy}_{+,\alpha})^*(t^{p'-1}\chi_{\hat{Q}}\, ds_\alpha)^p\, d\omega &\le C\int_{\hat{Q}} t^{p'}\,ds_\alpha \quad \textit{for all intervals } Q\in\mathcal{D}^\alpha.\end{aligned} \tag{5-10}$$

We claim that for any positive measure ν, the set of shifted dyadic grids $\{\mathcal{D}^\alpha\}_{\alpha\in\{0,1/3,2/3\}}$ satisfies

$$\mathsf{P}(Q_r,\nu) = \sum_{\ell=0}^{\infty}\frac{2^{-\ell}}{|2^\ell Q_r|}\int_{2^\ell Q_r} d\nu \approx \sum_{\alpha\in\{0,1/3,2/3\}}\sum_{\ell=0}^{\infty}\frac{2^{-\ell}}{|(I^\alpha_r)^{(\ell)}|}\int_{(I^\alpha_r)^{(\ell)}} d\nu = \sum_{\alpha\in\{0,1/3,2/3\}} \mathsf{P}^{dy}_\alpha(I^\alpha_r,\nu)$$

for all r. Indeed, for each interval $2^\ell Q_r$, there is $\alpha\in\{0,1/3,2/3\}$ and an interval $Q\in\mathcal{D}^\alpha$ containing $2^\ell Q_r$ whose length is comparable to that of $2^\ell Q_r$. Thus $Q=(I^\alpha_r)^{(\ell+c)}$ for some universal positive integer c. Now

$$\mathbb{P}_+(\nu)(x_{Q_r},\ell(Q_r)) = \int_{\mathbb{R}} P_{\ell(Q_r)}(x_{Q_r}-y)\,d\nu(y) \approx \sum_{\ell=0}^{\infty} 2^{-\ell}\frac{1}{|2^\ell Q_r|}\int_{2^\ell Q_r} d\nu = \mathsf{P}(Q_r,\nu).$$

Since σ is *doubling* and I^α_r is a maximal $\mathcal{D}^\alpha$-interval in Q_r with maximum length, we have $|Q_r|_\sigma \lesssim |I^\alpha_r|_\sigma$ and

$$\begin{aligned}\int_{\mathbb{R}^{n+1}_+} \mathbb{P}_+\nu(x,t)^p\,ds &= \sum_{r=1}^{\infty}|Q_r|_\sigma \mathbb{P}_+\nu(x_{Q_r},\ell(Q_r))^p \approx \sum_{r=1}^{\infty}|Q_r|_\sigma \mathsf{P}(Q_r,\nu)^p\\ &\approx \sum_{\alpha\in\{0,1/3,2/3\}}\sum_{r=1}^{\infty}|I^\alpha_r|_\sigma \mathsf{P}^{dy}_\alpha(I^\alpha_r,\nu)^p = \sum_{\alpha\in\{0,1/3,2/3\}}\int_{\mathbb{R}^2_+}\mathbb{P}^{dy}_{+,\alpha}\nu(x,t)^p\,ds_\alpha.\end{aligned}$$

This together with Theorem 5.6 reduces the proof of Claim 5.3 to showing that (5-10) holds for all $\alpha\in\{0,1/3,2/3\}$.

Now the definition of s_α in (5-9) shows that the left side of the first line in (5-10) is

$$\int_{\mathbb{R}^2_+} \mathbb{P}^{dy}_{+,\alpha}(\chi_Q\omega)^{p'}\,ds_\alpha = \sum_{r=1}^{\infty} |I^\alpha_r|_\sigma \mathsf{P}^{dy}_\alpha(I^\alpha_r, \chi_Q\omega)^{p'}.$$

Recall that $I^\alpha_r, Q \in \mathscr{D}^\alpha$. Now if $Q \subset I^\alpha_r$ for some r, then the sum above consists of just one term that satisfies

$$|I^\alpha_r|_\sigma \mathsf{P}^{dy}_\alpha(I^\alpha_r, \chi_Q\omega)^{p'} \le C\frac{|I^\alpha_r|_\sigma |Q|^{p'-1}_\omega}{|I^\alpha_r|^{p'}}|Q|_\omega \le C\mathscr{A}_p(\omega,\sigma)^{p'}|Q|_\omega.$$

Otherwise we have

$$\begin{aligned}\int_{\mathbb{R}^2_+} \mathbb{P}^{dy}_{+,\alpha}(\chi_Q\omega)^{p'}\,ds_\alpha &\lesssim \sum_{I^\alpha_r\subset Q} |I^\alpha_r|_\sigma \mathsf{P}^{dy}_\alpha(I^\alpha_r, \chi_Q\omega)^{p'} + \sum_{I^\alpha_r\cap Q=\varnothing} |I^\alpha_r|_\sigma \mathsf{P}^{dy}_\alpha(I^\alpha_r, \chi_Q\omega)^{p'} \\ &\le \mathfrak{C}^{p'}_* \int_Q d\omega + \sum_{I^\alpha_r\cap Q=\varnothing} |I^\alpha_r|_\sigma \Big(\sum_{\ell=0}^{\infty} \frac{2^{-\ell}}{|(I^\alpha_r)^{(\ell)}|}\int_{Q\cap (I^\alpha_r)^{(\ell)}} d\omega\Big)^{p'},\end{aligned}$$

where the local term has been estimated by the dual pivotal condition (5-2) applied to Q.

Now if $I^\alpha_r \subset Q^{(m)} \setminus Q^{(m-1)}$, then $Q \cap Q^{(\ell)}_r \ne \varnothing$ only if $Q^{(m)} \subset (I^\alpha_r)^{(\ell)}$. Thus the second term on the right can be estimated by

$$\begin{aligned}&\sum_{m=1}^{\infty} \sum_{I^\alpha_r\subset Q^{(m)}\setminus Q^{(m-1)}} |I^\alpha_r|_\sigma \Big(\sum_{\ell=0}^{\infty} \frac{2^{-\ell}}{|(I^\alpha_r)^{(\ell)}|}\int_{Q\cap (I^\alpha_r)^{(\ell)}} d\omega\Big)^{p'} \\ &\quad\le \sum_{m=1}^{\infty} \sum_{I^\alpha_r\subset Q^{(m)}\setminus Q^{(m-1)}} |I^\alpha_r|_\sigma \sum_{\ell=0}^{\infty} 2^{-\ell}\Big(\frac{\int_{Q\cap (I^\alpha_r)^{(\ell)}} d\omega}{|(I^\alpha_r)^{(\ell)}|}\Big)^{p'} \\ &\quad\le C\sum_{m=1}^{\infty} \sum_{I^\alpha_r\subset Q^{(m)}\setminus Q^{(m-1)}} |I^\alpha_r|_\sigma \sum_{\ell=0}^{\infty} 2^{-\ell}\Big(\frac{\int_Q d\omega}{|Q^{(m)}|}\Big)^{p'} \\ &\quad\le \Big(\sum_{m=1}^{\infty} \frac{|Q^{(m)}|_\sigma}{|Q^{(m)}|^{p'}}\Big)|Q|^{p'-1}_\omega \int_Q d\omega \\ &\quad= \Big(\frac{1}{|Q|^{p'}}\Big(\int s^{dy}_{Q,\alpha}(x)^{p'}\,d\sigma(x)\Big)|Q|^{p'-1}_\omega\Big)\int_Q d\omega \le C\mathscr{A}_p(\omega,\sigma)^{p'}\int_Q d\omega,\end{aligned}$$

where we have used

$$s^{dy}_{Q,\alpha}(x) \equiv \sum_{m=0}^{\infty} \frac{|Q|}{|Q^{(m)}|}\chi_{Q^{(m)}}(x) \lesssim s_Q(x),$$

and the half-strengthened A_p condition (5-1) in the final inequality.

Now we turn to showing that the second line in (5-10) holds using only the A_p condition (1-8). First we compute the dual operator $(\mathbb{P}^{dy}_{+,\alpha})^*$. Since the kernel of $\mathbb{P}^{dy}_{+,\alpha}$ is

$$\mathbb{P}^{dy}_{+,\alpha}[(x,t),y] \equiv \sum_{I\in\mathscr{D}^\alpha:\ell(I)\ge t} \chi_I(x)\frac{t}{\ell(I)}\frac{1}{|I|}\chi_I(y),$$

we have for any positive measure $\mu(x,t)$ on the upper half space $\mathbb{R}^2_+$,

$$(\mathbb{P}^{dy}_{+,\alpha})^*\mu(y)=\int_{\mathbb{R}^2_+}\Big(\sum_{I\in\mathcal{D}^\alpha:\ell(I)\geq t}\chi_I(x)\frac{t}{\ell(I)}\frac{1}{|I|}\chi_I(y)\Big)d\mu(x,t)=\sum_{I\in\mathcal{D}^\alpha:y\in I}\frac{1}{|I|}\int_{\hat{I}}\frac{t}{\ell(I)}d\mu(x,t).$$

Using the third line in (5-9) we compute that

$$\int_{\hat{Q}}t^{p'}\,ds_\alpha=\sum_{I_r^\alpha\subset Q}|I_r^\alpha|_\sigma|I_r^\alpha|^{p'},$$

and

$$\begin{aligned}(\mathbb{P}^{dy}_{+,\alpha})^*(t^{p'-1}\chi_{\hat{Q}}\,ds_\alpha)(y)&=\sum_{I\in\mathcal{D}^\alpha:y\in I}\frac{1}{|I|}\int_{\hat{I}\cap\hat{Q}}\frac{t}{\ell(I)}t^{p'-1}\,ds_\alpha(x,t)\\&=\sum_{I_r^\alpha\subset Q}|I_r^\alpha|_\sigma|I_r^\alpha|^{p'-1}\sum_{\ell=0}^\infty\frac{2^{-\ell}}{|(I_r^\alpha)^{(\ell)}|}\chi_{(I_r^\alpha)^{(\ell)}}(y).\end{aligned}$$

Thus we must prove

$$\int_{\mathbb{R}}\Big(\sum_{I_r^\alpha\subset Q}|I_r^\alpha|_\sigma|I_r^\alpha|^{p'-1}\sum_{\ell=0}^\infty\frac{2^{-\ell}}{|(I_r^\alpha)^{(\ell)}|}\chi_{(I_r^\alpha)^{(\ell)}}(y)\Big)^p\,d\omega(y)\leq C\mathcal{A}_p(\omega,\sigma)^p\sum_{I_r^\alpha\subset Q}|I_r^\alpha|_\sigma|I_r^\alpha|^{p'};\tag{5-11}$$

but this is the Poisson condition (1-24) in Theorem 1.10 for the shifted dyadic grid $\mathcal{D}^\alpha$. This completes the proof of the first assertion in Claim 5.3 regarding the case $1<p<\infty$. We now assume that $1<p\leq 2$ for the remainder of the proof.

To obtain (5-11) it suffices to show that for each $\ell\geq 0$

$$\int_{\mathbb{R}}\Big(\sum_{I_r^\alpha\subset Q}|I_r^\alpha|_\sigma|I_r^\alpha|^{p'-2}2^{-2\ell}\chi_{(I_r^\alpha)^{(\ell)}}(y)\Big)^p\,d\omega(y)\leq C2^{-p\ell}\mathcal{A}_p(\omega,\sigma)^p\sum_{I_r^\alpha\subset Q}|I_r^\alpha|_\sigma|I_r^\alpha|^{p'}.\tag{5-12}$$

Indeed, with this in hand, Minkowski's inequality yields

$$\begin{aligned}\|(\mathbb{P}^{dy}_{+,\alpha})^*(t\chi_{\hat{Q}}\,ds_\alpha)\|_{L^p(\omega)}&=\Big\|\sum_{\ell=0}^\infty\sum_{I_r^\alpha\subset Q}|I_r^\alpha|_\sigma|I_r^\alpha|^{p'-2}2^{-2\ell}\chi_{(I_r^\alpha)^{(\ell)}}\Big\|_{L^p(\omega)}\\&\leq\sum_{\ell=0}^\infty\Big\|\sum_{I_r^\alpha\subset Q}|I_r^\alpha|_\sigma|I_r^\alpha|^{p'-2}2^{-2\ell}\chi_{(I_r^\alpha)^{(\ell)}}\Big\|_{L^p(\omega)}\\&\leq C\sum_{\ell=0}^\infty 2^{-\ell}\mathcal{A}_p(\omega,\sigma)\Big(\sum_{I_r^\alpha\subset Q}|I_r^\alpha|_\sigma|I_r^\alpha|^{p'}\Big)^{1/p},\end{aligned}\tag{5-13}$$

as required.

Note that for $a>0$ and $p>1$,

$$h(x)\equiv(a+x)^p-a^p-p(a+x)^{p-1}x,$$

is decreasing on $[0,\infty)$ since $h'(x) = -p(p-1)(a+x)^{p-2}x < 0$ for $x > 0$. Since $h(0) = 0$ we have $h(x) \le 0$ for $x \ge 0$, that is,

$$(a+x)^p - a^p \le p(a+x)^{p-1}x \quad \text{for } a, x > 0 \text{ and } p > 1. \tag{5-14}$$

Now fix an interval Q in (5-12) and arrange the intervals I_r^α that are contained in Q into a sequence $\{I_r^\alpha\}_{r=1}^N$ in which the lengths $|I_r^\alpha|$ are increasing (we may suppose without loss of generality that N is finite). Recall we are now assuming $1 < p \le 2$. Integrate by parts and apply (5-14) to estimate the left side of (5-12) by

$$\begin{aligned}
&2^{-2p\ell}\int_{\mathbb{R}}\Big(\sum_{r=1}^N |I_r^\alpha|_\sigma |I_r^\alpha|^{p'-2}\chi_{(I_r^\alpha)^{(\ell)}}(y)\Big)^p d\omega(y)\\
&\quad= 2^{-2p\ell}\int_{\mathbb{R}}\sum_{n=1}^N\Big(\Big(\sum_{r=1}^n |I_r^\alpha|_\sigma |I_r^\alpha|^{p'-2}\chi_{(I_r^\alpha)^{(\ell)}}(y)\Big)^p - \Big(\sum_{r=1}^{n-1} |I_r^\alpha|_\sigma |I_r^\alpha|^{p'-2}\chi_{(I_r^\alpha)^{(\ell)}}(y)\Big)^p\Big) d\omega(y)\\
&\quad\le 2^{-2p\ell}\int_{\mathbb{R}}\sum_{n=1}^N\Big(p\Big(\sum_{r=1}^n |I_r^\alpha|_\sigma |I_r^\alpha|^{p'-2}\chi_{(I_r^\alpha)^{(\ell)}}(y)\Big)^{p-1} |I_n^\alpha|_\sigma |I_n^\alpha|^{p'-2}\chi_{(I_n^\alpha)^{(\ell)}}(y)\Big) d\omega(y)\\
&\quad\le 2^{-2p\ell}p\sum_{n=1}^N\int_{\mathbb{R}}\Big(\Big(\sum_{r=1}^n |I_r^\alpha|_\sigma \chi_{(I_r^\alpha)^{(\ell)}}(y)\Big)^{p-1} |I_n^\alpha|_\sigma |I_n^\alpha|^{p'-2}|I_n^\alpha|^{(p'-2)(p-1)}\chi_{(I_n^\alpha)^{(\ell)}}(y)\Big) d\omega(y),
\end{aligned}$$

where we have used (5-14) with

$$a = \sum_{r=1}^{n-1} |I_r^\alpha|_\sigma |I_r^\alpha|^{p'-2}\chi_{(I_r^\alpha)^{(\ell)}}(y) \quad\text{and}\quad x = |I_n^\alpha|_\sigma |I_n^\alpha|^{p'-2}\chi_{(I_n^\alpha)^{(\ell)}}(y),$$

and then used $|I_r^\alpha|^{p'-2} \le |I_n^\alpha|^{p'-2}$ for $1 \le r \le n$, which follows from $|I_r^\alpha| \le |I_n^\alpha|$ and $p' \ge 2$. If $(I_r^\alpha)^{(\ell)} \cap (I_n^\alpha)^{(\ell)} \ne \varnothing$ and $1 \le r \le n$, then $I_r^\alpha \subset (I_n^\alpha)^{(\ell)}$ and so

$$\begin{aligned}
&\int_{\mathbb{R}}\Big(\sum_{I_r^\alpha\subset Q} |I_r^\alpha|_\sigma |I_r^\alpha|^{p'-2}2^{-2\ell}\chi_{(I_r^\alpha)^{(\ell)}}(y)\Big)^p d\omega(y)\\
&\qquad\le 2^{-2p\ell}p\sum_{n=1}^N |I_n^\alpha|_\sigma |I_n^\alpha|^{p'p-2p}\int_{\mathbb{R}}\Big(\sum\nolimits^n_{\substack{1\le r\le n\\ I_r^\alpha\subset (I_n^\alpha)^{(\ell)}}} |I_r^\alpha|_\sigma\Big)^{p-1}\chi_{(I_n^\alpha)^{(\ell)}}(y)\, d\omega(y)\\
&\qquad\le 2^{-2p\ell}p\sum_{n=1}^N |I_n^\alpha|_\sigma |I_n^\alpha|^{p'p-2p}|(I_n^\alpha)^{(\ell)}|_\sigma^{p-1}|(I_n^\alpha)^{(\ell)}|_\omega\\
&\qquad\le 2^{-2p\ell}p\mathcal{A}_p(\omega,\sigma)^p\sum_{n=1}^N |I_n^\alpha|_\sigma |I_n^\alpha|^{p'p-2p}|(I_n^\alpha)^{(\ell)}|^p\\
&\qquad= 2^{-p\ell}p\mathcal{A}_p(\omega,\sigma)^p\sum_{n=1}^N |I_n^\alpha|_\sigma |I_n^\alpha|^{p'} = 2^{-p\ell}p\mathcal{A}_p(\omega,\sigma)^p\sum_{I_r^\alpha\subset Q} |I_n^\alpha|_\sigma |I_n^\alpha|^{p'}.
\end{aligned}$$

Thus we have proved (5-12) for $p \in (1,2]$, which completes the proof of (5-10). This finishes the proof of Claim 5.3, and hence also that of Theorem 5.4.

5.7. *Necessity of the conditions.* Here we consider the two weight Hilbert transform inequality for $1 < p < \infty$. We show the necessity of the strengthened A_p condition for general weights, as well as the necessity of the dual pivotal condition for the dual testing condition, and the dual Poisson inequality for the dual Hilbert transform inequality, when σ is doubling.

5.7.1. *The strengthened A_p condition.* Here we derive a necessary condition for the weighted inequality (1-18) but with the Hilbert transform T in place of $T_\natural$, that is,

$$\int_{\mathbb{R}\setminus \operatorname{supp} f} T(f\sigma)(x)^p \, d\omega(x) \le C \int_{\mathbb{R}^n} |f(x)|^p \, d\sigma(x). \tag{5-15}$$

The condition,

$$\Big(\int_{\mathbb{R}} \Big(\frac{|Q|}{|Q|+|x-x_Q|}\Big)^p d\omega(x)\Big)^{1/p} \Big(\int_{\mathbb{R}} \Big(\frac{|Q|}{|Q|+|x-x_Q|}\Big)^{p'} d\sigma(x)\Big)^{1/p'} \le C|Q| \tag{5-16}$$

for all intervals Q, is stronger than the two weight A_p condition (1-8), and we call it the *strengthened* A_p condition.

Preliminary results in this direction were obtained by Muckenhoupt and Wheeden, and in the setting of fractional integrals by Gabidzashvili and Kokilashvili, and here we follow the argument proving [Sawyer and Wheeden 1992, (1.9)], where "two-tailed" inequalities of the type (5-16) originated in the fractional integral setting. A somewhat different approach to this for the conjugate operator in the disk when $p=2$ uses conformal invariance and appears in [Nazarov et al. 2010], and provides the first instance of a strengthened A_2 condition being proved necessary for a two weight inequality for a singular integral.

Fix an interval Q and for $a \in \mathbb{R}$ and $r > 0$, let

$$s_Q(x) = \frac{|Q|}{|Q|+|x-x_Q|} \quad \text{and} \quad f_{a,r}(y) = \chi_{(a-r,a)}(y) s_Q(y)^{p'-1},$$

where x_Q is the center of the interval Q. For convenience we assume that neither ω nor σ have any point masses — see [Lacey et al. 2011] for the modifications necessary when point masses are present. For $y < x$ we have

$$|Q|(x-y) = |Q|(x-x_Q) + |Q|(x_Q - y) \le (|Q|+|x-x_Q|)(|Q|+|x_Q-y|),$$

and so

$$\frac{1}{x-y} \ge |Q|^{-1} s_Q(x) s_Q(y) \quad \text{for } y < x.$$

Thus for $x > a$ we obtain that

$$H(f_{a,r}\sigma)(x) = \int_{a-r}^{a} \frac{1}{x-y} s_Q(y)^{p'-1} \, d\sigma(y) \ge |Q|^{-1} s_Q(x) \int_{a-r}^{a} s_Q(y)^{p'} \, d\sigma(y),$$

and hence by (5-15) for the Hilbert transform H,

$$|Q|^{-p}\int_a^\infty s_Q(x)^p\Big(\int_{a-r}^a s_Q(y)^{p'}\,d\sigma(y)\Big)^p\,d\omega(x)$$
$$\le \int |H(f_{a,r}\sigma)(x)|^p\,d\omega(x)\le C\int |f_{a,r}(y)|^p\,d\sigma(y)=C\int_{a-r}^a s_Q(y)^{p'}\,d\sigma(y).$$

From this we obtain

$$|Q|^{-p}\Big(\int_a^\infty s_Q(x)^p\,d\omega(x)\Big)\Big(\int_{a-r}^a s_Q(y)^{p'}\,d\sigma(y)\Big)^{p-1}\le C,$$

and upon letting $r\to\infty$ and taking p-th roots, we get

$$\Big(\int_a^\infty s_Q(x)^p\,d\omega(x)\Big)^{1/p}\Big(\int_{-\infty}^a s_Q(y)^{p'}\,d\sigma(y)\Big)^{1/p'}\le C|Q|.$$

Similarly we have

$$\Big(\int_{-\infty}^a s_Q(x)^p\,d\omega(x)\Big)^{1/p}\Big(\int_a^\infty s_Q(y)^{p'}\,d\sigma(y)\Big)^{1/p'}\le C|Q|.$$

Now we choose a so that

$$\int_{-\infty}^a s_Q(y)^{p'}\,d\sigma(y)=\int_a^\infty s_Q(y)^{p'}\,d\sigma(y)=\frac{1}{2}\int s_Q(y)^{p'}\,d\sigma(y),$$

and conclude that

$$\Big(\int s_Q(x)^p\,d\omega(x)\Big)^{1/p}\Big(\int s_Q(y)^{p'}d\sigma(y)\Big)^{1/p'}$$
$$\le\Big(\int_{-\infty}^a s_Q(x)^p\,d\omega(x)\Big)^{1/p}\Big(\int s_Q(y)^{p'}d\sigma(y)\Big)^{1/p'}+\Big(\int_a^\infty s_Q(x)^p\,d\omega(x)\Big)^{1/p}\Big(\int s_Q(y)^{p'}d\sigma(y)\Big)^{1/p'}$$
$$\le 2^{1/p'}\Big(\int_{-\infty}^a s_Q(x)^p\,d\omega(x)\Big)^{1/p}\Big(\int_a^\infty s_Q(y)^{p'}d\sigma(y)\Big)^{1/p'}$$
$$+2^{1/p'}\Big(\int_a^\infty s_Q(x)^p\,d\omega(x)\Big)^{1/p}\Big(\int_{-\infty}^a s_Q(y)^{p'}d\sigma(y)\Big)^{1/p'}$$
$$\le 2^{1+1/p'}C|Q|.$$

5.7.2. *Necessity of the dual pivotal condition and the dual Poisson inequality for a doubling measure.* Here we show first that if σ is a *doubling* measure, then the dual pivotal condition (5-2) with $\delta(s)=s$ is implied by the A_p condition (1-8) and the dual testing condition for the Hilbert transform H, that is,

$$\int_I |H(\chi_I\omega)(x)|^{p'}\,d\sigma(x)\le C_{\omega,\sigma,p}|I|_\omega\quad\text{for all intervals } I. \tag{5-17}$$

After this we show that the dual Poisson inequality (5-5) is implied by the A_p condition (1-8) and the dual Hilbert transform inequality,

$$\int_I |H(\chi_I g\omega)(x)|^{p'}\,d\sigma(x)\le C_{\omega,\sigma,p}\int_I g(x)^{p'}\,d\omega(x)\quad\text{for all } g\ge 0 \text{ and intervals } I. \tag{5-18}$$

Lemma 5.8. *Suppose that σ is doubling and $T = H$ is the Hilbert transform. Then the dual pivotal condition* (5-2) *is implied by the A_p condition* (1-8) *and the dual testing condition* (5-17).

Proof. We begin by proving that for any interval I and any positive measure ν supported in $\mathbb{R}\setminus I$, we have

$$\mathbb{P}(I;\nu) \le \frac{1}{|I|}\int_I d\nu + 2|I| \inf_{x,y\in I} \frac{H(\chi_{I^c}\nu)(x) - H(\chi_{I^c}\nu)(y)}{x-y}, \tag{5-19}$$

where we here redefine

$$\mathbb{P}(I;\nu) \equiv \frac{1}{|I|}\int_I d\nu + \frac{|I|}{2}\int_{\mathbb{R}\setminus I} \frac{1}{|z-z_I|^2} d\nu(z), \tag{5-20}$$

with z_I the center of I. Note that this definition of $\mathbb{P}(I;\nu)$ is comparable to that in (5-3) with $\delta(s)=s$. Note also that $H(\chi_{I^c}\nu)$ is defined by (5-15) on I, and increasing on I when ν is positive, so that the infimum in (5-19) is nonnegative.

To see (5-19), we suppose without loss of generality that $I=(-a,a)$, and a calculation then shows that for $-a \le x < y \le a$,

$$\begin{aligned} &H(\chi_{I^c}\nu)(y) - H(\chi_{I^c}\nu)(x) \\ &\quad = \int_{\mathbb{R}\setminus I}\Big(\frac{1}{z-y}-\frac{1}{z-x}\Big)d\nu(z) = (y-x)\int_{\mathbb{R}\setminus I}\frac{1}{(z-y)(z-x)}d\nu(z) \ge \tfrac14 (y-x)\int_{\mathbb{R}\setminus I}\frac{1}{z^2}d\nu(z), \end{aligned}$$

since $((z-y)(z-x))^{-1}$ is positive and satisfies

$$\frac{1}{(z-y)(z-x)} \ge \frac{1}{4z^2}$$

on each interval $(-\infty,-a)$ and (a,∞) in $\mathbb{R}\setminus I$ when $-a\le x<y\le a$. Thus we have from (5-20)

$$\mathbb{P}(I;\nu) = \frac{1}{|I|}\int_I d\nu + \frac{|I|}{2}\int_{\mathbb{R}\setminus I}\frac{1}{z^2}d\nu(z) \le \frac{1}{|I|}\int_I d\nu + 2|I|\inf_{x,y\in I}\frac{H(\chi_{I^c}\nu)(y)-H(\chi_{I^c}\nu)(x)}{y-x}.$$

Now we return to the dual pivotal condition (5-2), and let $C_{\omega,\sigma,p}$ be the best constant in the dual testing condition (5-17) for H. Let $Q_0 = \bigcup_{r=1}^\infty Q_r$ be a pairwise disjoint decomposition of Q_0 and consider $\varepsilon,\delta>0$, which will be chosen at the end of the proof (we will take $\delta=\frac12$ and $\varepsilon>0$ very small). For each interval Q_r, let $\alpha_r\in Q_r$ minimize $|H(\chi_{Q_r^c}\omega)|$ on Q_r, that is,

$$|H(\chi_{Q_r^c}\omega)(\alpha_r)| = \min_{x\in I}|H(\chi_{Q_r^c}\omega)(x)|,$$

and set

$$J_{r,\varepsilon} \equiv (\alpha_r - \varepsilon|Q_r|, \alpha_r+\varepsilon|Q_r|)\cap Q_r.$$

Now for each interval Q_r, consider the following three mutually exclusive and exhaustive cases:

Case 1: $\displaystyle \frac{1}{|Q_r|}\int_{Q_r} d\omega > \frac{|Q_r|}{4}\int_{\mathbb{R}\setminus Q_r}\frac{1}{|z-z_{Q_r}|^2}\,d\omega(z),$

Case 2: $\displaystyle \frac{1}{|Q_r|}\int_{Q_r} d\omega \le \frac{|Q_r|}{4}\int_{\mathbb{R}\setminus Q_r}\frac{1}{|z-z_{Q_r}|^2}\,d\omega(z)$ and $|Q_r\setminus J_{r,\varepsilon}|_\sigma \ge \delta|Q_r|_\sigma,$

Case 3: $\frac{1}{|Q_r|}\int_{Q_r} d\omega \le \frac{|Q_r|}{4}\int_{\mathbb{R}\setminus Q_r}\frac{1}{|z-z_{Q_r}|^2}\,d\omega(z)$ and $|J_{r,\varepsilon}|_\sigma > (1-\delta)|Q_r|_\sigma$.

If Q_r is a Case 1 interval we have $\mathbb{P}(Q_r,\chi_{Q_0}\omega) \le 3|Q_r|^{-1}\int_{Q_r} d\omega$ and so

$$\sum_{Q_r \text{ satisfies Case 1}} |Q_r|_\sigma \mathbb{P}(Q_r,\chi_{Q_0}\omega)^{p'} \le 3^p \sum_{r=1}^\infty |Q_r|_\sigma \Big(\frac{1}{|Q_r|}\int_{Q_r} d\omega\Big)^{p'}$$
$$\le C_p \sum_{r=1}^\infty \frac{|Q_r|_\sigma |Q_r|_\omega^{p'-1}}{|Q_r|^{p'}}\int_{Q_r} d\omega \le C_p \|(\omega,\sigma)\|_{A_p}^{p'}\int_{Q_0} d\omega.$$

If Q_r is a Case 2 or Case 3 interval we have from (5-19) with $\nu=\chi_{Q_0}\omega$ that for all $x\in Q_r\setminus J_{r,\varepsilon}$,

$$\mathbb{P}(Q_r;\chi_{Q_0}\omega) \le 6|Q_r|\frac{H(\chi_{Q_0\cap Q_r^c}\omega)(x)-H(\chi_{Q_0\cap Q_r^c}\omega)(\alpha_r)}{x-\alpha_r}$$
$$\le 6|Q_r|\frac{1}{\varepsilon|Q_r|}\big(|H(\chi_{Q_0\cap Q_r^c}\omega)(x)|+|H(\chi_{Q_0\cap Q_r^c}\omega)(\alpha_r)|\big) \le \frac{12}{\varepsilon}|H(\chi_{Q_0\cap Q_r^c}\omega)(x)|.$$

If now Q_r is a Case 2 interval, we also have $|Q_r|_\sigma \le \delta^{-1}|Q_r\setminus J_{r,\varepsilon}|_\sigma$ and so

$$\sum_{Q_r \text{ satisfies Case 2}} |Q_r|_\sigma \mathbb{P}(Q_r,\chi_{Q_0}\omega)^{p'}$$
$$\le \frac{1}{\delta}\sum_{Q_r \text{ satisfies Case 2}} |Q_r\setminus J_{r,\varepsilon}|_\sigma \mathbb{P}(Q_r,\chi_{Q_0}\omega)^{p'}$$
$$\le \frac{1}{\delta}\sum_{r=1}^\infty \Big(\frac{12}{\varepsilon}\Big)^{p'}\int_{Q_r\setminus J_{r,\varepsilon}} |H(\chi_{Q_0\cap Q_r^c}\omega)(x)|^{p'}\,d\sigma(x)$$
$$\le C_{\varepsilon,\delta,p}\sum_{r=1}^\infty \int_{Q_r\setminus J_{r,\varepsilon}} \big(|H(\chi_{Q_0}\omega)(x)|^{p'}+|H(\chi_{Q_r}\omega)(x)|^{p'}\big)\,d\sigma(x)$$
$$\le C_{\varepsilon,\delta,p}\Big(\int_{Q_0}|H(\chi_{Q_0}\omega)(x)|^{p'}\,d\sigma(x)+\sum_{r=1}^\infty\int_{Q_r}|H(\chi_{Q_r}\omega)(x)|^{p'}\,d\sigma(x)\Big)$$
$$\le C_{\varepsilon,\delta,p}\Big(C|Q_0|_\omega+\sum_{r=1}^\infty C|Q_r|_\omega\Big) = C_{\varepsilon,\delta,p}|Q_0|_\omega,$$

where the final inequality follows from (5-17) with $I=Q_0$ and then $I=Q_r$.

Now we use our assumption that σ is doubling. There are $C,\eta>0$ such that

$$|J|_\sigma \le C\Big(\frac{|J|}{|Q|}\Big)^\eta |Q|_\sigma$$

whenever J is a subinterval of an interval Q. If Q_r is a Case 3 interval we have both

$$\frac{|J_{r,\varepsilon}|}{|Q_r|} \le 2\varepsilon \quad\text{and}\quad |J_{r,\varepsilon}|_\sigma > (1-\delta)|Q_r|_\sigma,$$

which altogether yields

$$(1-\delta)|Q_r|_\sigma < |J_{r,\varepsilon}|_\sigma \le C\Big(\frac{|J_{r,\varepsilon}|}{|Q_r|}\Big)^\eta |Q_r|_\sigma \le C(2\varepsilon)^\eta |Q_r|_\sigma,$$

which is a contradiction if $\delta = 1/2$ and $\varepsilon > 0$ is chosen sufficiently small, so that $\varepsilon < 1/2(1/(2C))^{1/\eta}$. With this choice, there are no Case 3 intervals, and so we are done. □

Lemma 5.9. *Suppose that σ is doubling and $T = H$ is the Hilbert transform. Then the dual Poisson inequality* (5-5) *is implied by the A_p condition* (1-8) *and the dual Hilbert transform inequality* (5-18).

Proof. The proof is virtually identical to that of Lemma 5.8 but with $d\nu = \chi_{Q_0} g\, d\omega$ in place of $\chi_{Q_0}\, d\omega$ where $g \ge 0$. Indeed, if Q_r is a Case 1 interval we then have $\mathbb{P}(Q_r, \chi_{Q_0} g\omega) \le 3|Q_r|^{-1} \int_{Q_r} g\, d\omega$ and so

$$\begin{aligned}\sum_{Q_r \text{ satisfies Case 1}} |Q_r|_\sigma \mathbb{P}(Q_r, \chi_{Q_0} g\omega)^{p'} &\le 3^p \sum_{r=1}^\infty |Q_r|_\sigma \Big(\frac{1}{|Q_r|}\int_{Q_r} g\, d\omega\Big)^{p'} \\ &\le C_p \sum_{r=1}^\infty \frac{|Q_r|_\sigma |Q_r|_\omega^{p'-1}}{|Q_r|^{p'}} \int_{Q_r} g^{p'}\, d\omega \le C_p \|(\omega,\sigma)\|_{A_p}^{p'} \int_{Q_0} g^{p'}\, d\omega.\end{aligned}$$

If Q_r is a Case 2 interval, then $|Q_r|_\sigma \le \delta^{-1}|Q_r \setminus J_{r,\varepsilon}|_\sigma$ and

$$\begin{aligned}&\sum_{Q_r \text{ satisfies Case 2}} |Q_r|_\sigma \mathbb{P}(Q_r, \chi_{Q_0} g\omega)^{p'} \\ &\quad\le \frac{1}{\delta} \sum_{Q_r \text{ satisfies Case 2}} |Q_r \setminus J_{r,\varepsilon}|_\sigma \mathbb{P}(Q_r, \chi_{Q_0} g\omega)^{p'} \\ &\quad\le \frac{1}{\delta}\sum_{r=1}^\infty \Big(\frac{12}{\varepsilon}\Big)^{p'} \int_{Q_r\setminus J_{r,\varepsilon}} |H(\chi_{Q_0 \cap Q_r^c} g\omega)(x)|^{p'}\, d\sigma(x) \\ &\quad\le C_{\varepsilon,\delta,p} \sum_{r=1}^\infty \int_{Q_r\setminus J_{r,\varepsilon}} (|H(\chi_{Q_0} g\omega)(x)|^{p'} + |H(\chi_{Q_r} g\omega)(x)|^{p'})\, d\sigma(x) \\ &\quad\le C_{\varepsilon,\delta,p}\Big(\int_{Q_0} |H(\chi_{Q_0} g\omega)(x)|^{p'}\, d\sigma(x) + \sum_{r=1}^\infty \int_{Q_r} |H(\chi_{Q_r} g\omega)(x)|^{p'}\, d\sigma(x)\Big) \\ &\quad\le C_{\varepsilon,\delta,p}\Big(C\int_{Q_0} g^{p'}\, d\omega + \sum_{r=1}^\infty C\int_{Q_r} g^{p'}\, d\omega\Big) = C_{\varepsilon,\delta,p}\int_{Q_0} g^{p'}\, d\omega,\end{aligned}$$

upon using (5-18) with Q_0 and Q_r, which is (5-5). As before, Case 3 intervals don't exist if σ is doubling and $\varepsilon > 0$ is sufficiently small. □

Proof of Theorem 1.10. Theorem 5.4 shows that the dual Poisson inequality (5-5) holds uniformly in Q_0 and pairwise disjoint $\{Q_r\}_{r=1}^\infty$ satisfying $\bigcup_{r=1}^\infty Q_r \subset Q_0$, provided both the half-strengthened A_p condition (5-1) and the dual pivotal condition (5-2) hold when $1 < p \le 2$ — and provided (5-1), (5-2) and the Poisson condition (1-24) hold when $p > 2$. Since σ is doubling, Lemma 5.8 shows that the dual pivotal condition (5-2) follows from the dual testing condition (1-21) — and Lemma 5.9 shows that the dual Poisson inequality (5-5), and hence also the Poisson condition (1-24), follows from the dual Hilbert transform inequality (5-18). Thus Theorem 1.10 now follows from the claim proved in Section 5.1.1 that (5-5) can be substituted for (1-17) in the proof of Theorem 1.9. □

Acknowledgment

We began this work during research stays at the Fields Institute in Toronto, Canada, and continued at the Centre de Recerca Matemàtica in Barcelona, Spain. We thank these institutions for their generous hospitality. In addition, this paper has been substantially improved by the careful attention of the referee, for which we are particularly grateful.

References

[Cotlar and Sadosky 1979] M. Cotlar and C. Sadosky, "On the Helson–Szegő theorem and a related class of modified Toeplitz kernels", pp. 383–407 in *Harmonic analysis in Euclidean spaces* (Williamstown, MA 1978), vol. 1, edited by G. Weiss and S. Wainger, Proc. Sympos. Pure Math. **35**, Amer. Math. Soc., Providence, R.I., 1979. MR 81j:42022 Zbl 0448.42008

[Cotlar and Sadosky 1983] M. Cotlar and C. Sadosky, "On some L^p versions of the Helson–Szegő theorem", pp. 306–317 in *Conference on harmonic analysis in honor of Antoni Zygmund* (Chicago, 1981), vol. 1, edited by W. Beckner et al., Wadsworth, Belmont, CA, 1983. MR 85i:42015

[Cruz-Uribe et al. 2007] D. Cruz-Uribe, J. M. Martell, and C. Pérez, "Sharp two-weight inequalities for singular integrals, with applications to the Hilbert transform and the Sarason conjecture", *Adv. Math.* **216**:2 (2007), 647–676. MR 2008k:42029 Zbl 1129.42007

[Lacey et al. 2011] M. T. Lacey, E. T. Sawyer, and I. Uriarte-Tuero, "A two weight inequality for the Hilbert transform assuming an energy hypothesis", preprint, version 7, 2011. arXiv 1001.4043v7

[Mateu et al. 2000] J. Mateu, P. Mattila, A. Nicolau, and J. Orobitg, "BMO for nondoubling measures", *Duke Math. J.* **102**:3 (2000), 533–565. MR 2001e:26019 Zbl 0964.42009

[Muckenhoupt 1972] B. Muckenhoupt, "Weighted norm inequalities for the Hardy maximal function", *Trans. Amer. Math. Soc.* **165** (1972), 207–226. MR 45 #2461 Zbl 0236.26016

[Muscalu et al. 2002] C. Muscalu, T. Tao, and C. Thiele, "Multi-linear operators given by singular multipliers", *J. Amer. Math. Soc.* **15**:2 (2002), 469–496. MR 2003b:42017 Zbl 0994.42015

[Nazarov et al. 1997] F. Nazarov, S. Treil, and A. Volberg, "Cauchy integral and Calderón–Zygmund operators on nonhomogeneous spaces", *Internat. Math. Res. Notices* 15 (1997), 703–726. MR 99e:42028 Zbl 0889.42013

[Nazarov et al. 1999] F. Nazarov, S. Treil, and A. Volberg, "The Bellman functions and two-weight inequalities for Haar multipliers", *J. Amer. Math. Soc.* **12**:4 (1999), 909–928. MR 2000k:42009 Zbl 0951.42007

[Nazarov et al. 2003] F. Nazarov, S. Treil, and A. Volberg, "The Tb-theorem on non-homogeneous spaces", *Acta Math.* **190**:2 (2003), 151–239. MR 2005d:30053 Zbl 1065.42014

[Nazarov et al. 2008] F. Nazarov, S. Treil, and A. Volberg, "Two weight inequalities for individual Haar multipliers and other well localized operators", *Math. Res. Lett.* **15**:3 (2008), 583–597. MR 2009e:42031 Zbl 05310656

[Nazarov et al. 2010] F. Nazarov, S. Treil, and A. Volberg, "Two weight estimate for the Hilbert transform and corona decomposition for non-doubling measures", preprint, 2005 and arXiv, 2010. arXiv 1003.1596

[Peherstorfer et al. 2007] F. Peherstorfer, A. Volberg, and P. Yuditskii, "Two-weight Hilbert transform and Lipschitz property of Jacobi matrices associated to hyperbolic polynomials", *J. Funct. Anal.* **246**:1 (2007), 1–30. MR 2008j:47024 Zbl 1125.47023

[Petermichl 2000] S. Petermichl, "Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol", *C. R. Acad. Sci. Paris Sér. I Math.* **330**:6 (2000), 455–460. MR 2000m:42016 Zbl 0991.42003

[Petermichl et al. 2002] S. Petermichl, S. Treil, and A. Volberg, "Why the Riesz transforms are averages of the dyadic shifts?", pp. 209–228 in *Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations* (El Escorial, 2000), vol. extra, 2002. MR 2003m:42028 Zbl 1031.47021

[Rudin 1987] W. Rudin, *Real and complex analysis*, 3rd ed., McGraw-Hill, New York, 1987. MR 88k:00002 Zbl 0925.00005

[Sawyer 1982] E. T. Sawyer, "A characterization of a two-weight norm inequality for maximal operators", *Studia Math.* **75**:1 (1982), 1–11. MR 84i:42032 Zbl 0508.42023

[Sawyer 1984] E. Sawyer, "A two weight weak type inequality for fractional integrals", *Trans. Amer. Math. Soc.* **281**:1 (1984), 339–345. MR 85j:26010 Zbl 0539.42008

[Sawyer 1988] E. T. Sawyer, "A characterization of two weight norm inequalities for fractional and Poisson integrals", *Trans. Amer. Math. Soc.* **308**:2 (1988), 533–545. MR 89d:26009 Zbl 0665.42023

[Sawyer and Wheeden 1992] E. Sawyer and R. L. Wheeden, "Weighted inequalities for fractional integrals on Euclidean and homogeneous spaces", *Amer. J. Math.* **114**:4 (1992), 813–874. MR 94i:42024 Zbl 0783.42011

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series **43**, Princeton University Press, 1993. MR 95c:42002 Zbl 0821.42001

[Stein and Shakarchi 2005] E. M. Stein and R. Shakarchi, *Real analysis: Measure theory, integration, and Hilbert spaces*, Princeton Lectures in Analysis **3**, Princeton University Press, 2005. MR 2005k:28024 Zbl 1081.28001

[Volberg 2003] A. Volberg, *Calderón–Zygmund capacities and operators on nonhomogeneous spaces*, CBMS Regional Conference Series in Mathematics **100**, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2003. MR 2005c:42015 Zbl 1053.42022

[Zheng 1996] D. Zheng, "The distribution function inequality and products of Toeplitz operators and Hankel operators", *J. Funct. Anal.* **138**:2 (1996), 477–501. MR 97e:47040 Zbl 0865.47019

Received 7 Oct 2009. Revised 2 Feb 2011. Accepted 2 Mar 2011.

MICHAEL LACEY: `lacey@math.gatech.edu`
School of Mathematics, Georgia Institute of Technology, 686 Cherry Street NW, Atlanta, GA 30332-0160, United States
http://www.math.gatech.edu/~lacey

ERIC T. SAWYER: `sawyer@mcmaster.ca`
Department of Mathematics and Statistics, McMaster University, 1280 Main St. West, Hamilton, ON L8S 4K1, Canada
http://www.math.mcmaster.ca/~sawyer/

IGNACIO URIARTE-TUERO: `ignacio@math.msu.edu`
Department of Mathematics, Michigan State University, East Lansing, MI 48824, United States

ANALYSIS AND PDE
Vol. 5, No. 1, 2012
dx.doi.org/10.2140/apde.2012.5.61

ENERGY IDENTITY FOR INTRINSICALLY BIHARMONIC MAPS IN FOUR DIMENSIONS

PETER HORNUNG AND ROGER MOSER

Let u be a mapping from a bounded domain $S \subset \mathbb{R}^4$ into a compact Riemannian manifold N. Its intrinsic biharmonic energy $E_2(u)$ is given by the squared L^2-norm of the intrinsic Hessian of u. We consider weakly converging sequences of critical points of E_2. Our main result is that the energy dissipation along such a sequence is fully due to energy concentration on a finite set and that the dissipated energy equals a sum over the energies of finitely many entire critical points of E_2.

1. Introduction and main result

Let $S \subset \mathbb{R}^4$ be a bounded Lipschitz domain and let N be a compact Riemannian manifold without boundary. For convenience we assume that N is embedded in $\mathbb{R}^n$ for some $n \geq 2$. We denote the second fundamental form of this embedding by A and we denote the Riemannian curvature tensor of N by R. For $u \in C^\infty(S, N)$ define the pull-back vector bundle $u^{-1}TN$ in the usual way and denote the norm on it and on related bundles by $|\cdot|$. Together with the Levi-Cività connection on the tangent bundle TN, the mapping u induces a covariant derivative ∇^u on $u^{-1}TN$. We extend this covariant derivative to tensor fields in the usual way. Denote by π_N the nearest point projection from a neighborhood of N onto N and set $P_u(x) = D\pi_N(u(x))$. Then $P_u(x)$ is the orthogonal projection from $\mathbb{R}^4$ onto the tangent space $T_{u(x)}N$ to N at $u(x)$. Let $X \in L^2(S, \mathbb{R}^n)$ be a section of $u^{-1}TN$. Following [Moser 2008] we define

$$\nabla^u X = (P_u \partial_\alpha X) \otimes dx^\alpha$$

Denote the derivative of u by $Du = (\partial_\alpha u) \otimes dx^\alpha$. The intrinsic Hessian $\nabla^u Du$ is a section of $(TS)^* \otimes (TS)^* \otimes u^{-1}TN$. By a standard fact about $D\pi_N$, it is given by

$$\begin{aligned}\nabla^u Du &= (P_u \partial_\alpha \partial_\beta u) \otimes dx^\alpha \otimes dx^\beta \\ &= \big(\partial_\alpha \partial_\beta u + A(u)(\partial_\alpha u, \partial_\beta u)\big) \otimes dx^\alpha \otimes dx^\beta.\end{aligned}$$

We define the Sobolev spaces

$$W^{k,p}(S, N) = \{u \in W^{k,p}(S, \mathbb{R}^n) : u(x) \in N \text{ for almost all } x \in S\}$$

Supported by EPSRC grant EP/F048769/1. Part of this work was carried out while Hornung held a postdoc position in the group of Sergio Conti in Bonn.
Hornung is the corresponding author.
MSC2000: 58E20, 35J35.
Keywords: biharmonic map, energy identity, bubbling.

and we introduce the energy functional $E_2 : W^{2,2}(S, N) \to \mathbb{R}_+$ given by

$$E_2(u) = \frac{1}{4}\int_S |\nabla^u Du|^2.$$

Critical points of E_2 are called intrinsically biharmonic mappings. There are also other kinds of second order functionals whose critical points are called "biharmonic" mappings. The functional E_2 is defined intrinsically, that is, it does not depend on the embedding of N into $\mathbb{R}^n$. Another intrinsically defined second order functional that is naturally associated with u is $F_2(u) = \frac{1}{4}\int_S |\tau(u)|^2$, where $\tau(u) := \operatorname{trace} \nabla^u Du$ denotes the tension field of u. Critical points of F_2 are usually called intrinsically biharmonic mappings. Another functional that can be associated with u is the energy $\tilde{E}_2(u) = \frac{1}{4}\int_S |D^2 u|^2$. Its critical points are usually called extrinsically biharmonic mappings. The functional $\tilde{E}_2$ enjoys better analytical properties than E_2 and F_2, but it has the drawback of depending on the particular embedding of N into $\mathbb{R}^n$.

Biharmonic mappings, being the next higher order equivalent of harmonic mappings, have attracted a lot of attention in the differential geometry literature; see [Montaldo and Oniciuc 2006] for an overview. Analytic aspects of the problem are less well understood, and on questions other than regularity (see [Chang et al. 1999; Wang 2004b; Wang 2004a; Wang 2004c; Lamm and Rivière 2008; Struwe 2008]) not much work has been done. This is the case in particular for intrinsic biharmonic mappings, because the problem is difficult due to a lack of coercivity of the corresponding functions in the Sobolev spaces traditionally used. Thus despite the fact that the intrinsic case is geometrically more interesting, the problem has not widely been studied from the analysis point of view.

Recent progress has been made, however, based on the observation that the lack of coercivity can be removed for one type of intrinsic biharmonic mappings (the type studied in the present paper), provided that one works in a geometrically motivated variant of Sobolev spaces [Moser 2008; Scheven 2009]. This approach permits methods analogous to what has been used for harmonic mappings. But since we have a fourth order equation for biharmonic mappings (in contrast to second order for harmonic mappings), and since we have to work in different spaces, such an approach still requires additional ideas and arguments. In this paper, we develop the theory a step further.

The existence of minimizers of E_2 under given boundary conditions on the mapping itself and on its first derivatives was established in [Moser 2008] using the direct method of the calculus of variations. For simplicity, from now on we will omit the adverb "intrinsically":

In the present paper, a mapping $u \in W^{2,2}(S, N)$ will be called *biharmonic* if it is critical for E_2 under outer variations, that is,

$$\frac{d}{dt}\bigg|_{t=0} E_2(\pi_N(u + t\phi)) = 0 \quad \text{for all } \phi \in C_0^\infty(S, \mathbb{R}^n);$$

see [Scheven 2009; Moser 2008]. In [Scheven 2009] it is shown that a mapping $u \in W^{2,2}(S, N)$ is biharmonic precisely if it satisfies

$$\int_S \nabla_\alpha \partial_\beta u \cdot \big(\nabla_\alpha \nabla_\beta \phi + R(u)(\phi, \partial_\alpha u)\partial_\beta u\big) = 0 \tag{1}$$

for every section $\phi \in W_0^{2,2}(S, \mathbb{R}^n) \cap L^\infty(S, \mathbb{R}^n)$ of $u^{-1}TN$.

We will study sequences of biharmonic mappings $(u_k) \subset W^{2,2}(S, N)$ with uniformly bounded energy, that is, $\limsup_{k\to\infty} E_2(u_k) < \infty$. Since our results are analogous to known facts about harmonic mappings, we describe the situation encountered in that context: Let $\Omega \subset \mathbb{R}^2$ be a bounded Lipschitz domain in $\mathbb{R}^2$. A mapping $u \in W^{1,2}(\Omega, N)$ is said to be (weakly) harmonic if it is a critical point for the Dirichlet energy

$$E_1(u) = \frac{1}{2}\int_\Omega |Du|^2.$$

A given sequence $(u_k) \subset W^{1,2}(\Omega, N)$ of harmonic mappings with uniformly bounded Dirichlet energy has a subsequence that converges weakly in $W^{1,2}$ to some mapping $u \in W^{1,2}(\Omega, N)$. This convergence in general fails to be strong, that is, in general $\liminf_{k\to\infty} E_1(u_k) > E_1(u)$. The only reason for this loss is that the energy can concentrate on a lower dimensional subset $\Sigma_0 \subset \Omega$. In particular, $u_k \to u$ in $C^1_{\mathrm{loc}}(\Omega \setminus \Sigma_0, \mathbb{R}^n)$. By the results in [Hélein 1991; Hélein 1990], the mappings u_k and u are smooth. In addition, the set Σ_0 is finite. Moreover, for each point $x \in \Sigma_0$ there exist $M_x \in \mathbb{N}$ and entire harmonic mappings $v_1^x, \dots, v_{M_x}^x \in C^\infty(\mathbb{R}^2, N)$ such that, after passing to a subsequence,

$$\lim_{k\to\infty}\int_\Omega |Du_k|^2 \ge \int_\Omega |Du|^2 + \sum_{x\in\Sigma_0}\sum_{j=1}^{M_x}\int_{\mathbb{R}^2} |v_j^x|^2.$$

Later the converse inequality was shown to hold as well [Jost 1991; Parker 1996; Ding and Tian 1995]. Our main result is the analogue of these facts for critical points of the functional E_2. It is summarized in the following theorem:

Theorem 1.1. *Let $S \subset \mathbb{R}^4$ be a bounded Lipschitz domain and let N be a smooth compact manifold without boundary embedded in $\mathbb{R}^n$. Let $(u_k) \subset W^{2,2}(S, N)$ be a sequence of biharmonic mappings and assume that*

$$\limsup_{k\to\infty}\int_S |\nabla^{u_k} Du_k|^2 + |Du_k|^4 < \infty. \tag{2}$$

Then $u_k \in C^\infty(S, N)$ and we may pass to a subsequence in k (again called (u_k)) and find a biharmonic map $u \in C^\infty(S, N)$ and a finite set $\Sigma_0 \subset S$ such that

(i) $u_k \rightharpoonup u$ *weakly in* $(W^{2,2} \cap W^{1,4})(S, \mathbb{R}^n)$,

(ii) $u_k \to u$ *in* $C^2_{\mathrm{loc}}(S \setminus \Sigma_0, \mathbb{R}^n)$.

Moreover, for each $x \in \Sigma_0$ there exist $M_x \in \mathbb{N}$ and biharmonic mappings $v_1^x, \dots, v_{M_x}^x \in C^\infty(\mathbb{R}^4, N)$ such that

$$\begin{aligned}\lim_{k\to\infty}\int_S |\nabla^{u_k} Du_k|^2 &= \int_S |\nabla^u Du|^2 + \sum_{x\in\Sigma_0}\sum_{j=1}^{M_x}\int_{\mathbb{R}^4} |\nabla^{v_j^x} Dv_j^x|^2,\\ \lim_{k\to\infty}\int_S |Du_k|^4 &= \int_S |Du|^4 + \sum_{x\in\Sigma_0}\sum_{j=1}^{M_x}\int_{\mathbb{R}^4} |Dv_j^x|^4.\end{aligned} \tag{3}$$

Remarks. (i) By [Moser 2008, Theorem 2.1] the hypothesis (2) is equivalent to the seemingly weaker hypothesis $\limsup_{k\to\infty}\int_S |\nabla^{u_k} Du_k|^2 + |Du_k|^2 < \infty$ and also to the seemingly stronger hypothesis

$$\limsup_{k\to\infty} \|u_k\|_{W^{2,2}(S,N)} < \infty.$$

(ii) Moser [2008] showed that every biharmonic mapping $v \in W^{2,2}(S, N)$ in fact satisfies $v \in C^\infty(S, N)$.

(iii) To obtain smoothness of the limiting mapping u as well, one needs a removability result for isolated singularities of biharmonic mappings. This is derived in Lemma 2.3 below. Another auxiliary result is the existence of a uniform lower bound on the energy of entire nonconstant biharmonic mappings, given in Lemma 2.6 below. Analogues of these facts are well known for harmonic mappings and also for critical points of other higher order functionals; see for example [Wang 2004b].

(iv) The main contribution of Theorem 1.1 are the energy identities of (3). To obtain an equality (and not just a lower bound for the left hand sides), one has to show that no energy concentrates in a "neck" region around a concentration point $x \in \Sigma_0$. This is proven in Section 3 below. Similar results are known in the context of harmonic mappings; see for example [Jost 1991; Parker 1996; Ding and Tian 1995; Lin and Rivière 2002]. They are also known for other kinds of biharmonic mappings, but only if the target manifold is a round sphere, since then the Euler–Lagrange equations enjoy a special structure [Wang 2004b]. Under the general hypotheses of Theorem 1.1 no such structure seems available, so a different approach is needed.

Notation. By $e_1, \dots, e_4$ we denote the standard basis of $\mathbb{R}^4$. We also set $e_r(x) = x/|x|$ for all $x \in \mathbb{R}^4$. By $B_r(x)$ we denote the open ball in $\mathbb{R}^4$ with center x and radius r. We set $B_r = B_r(0)$. If A and B are tensors of the same type, then $A \cdot B$ denotes their scalar product. We will often write ∇Du instead of $\nabla^u Du$, and we identify $\mathbb{R}^k$ with its dual $(\mathbb{R}^k)^*$, writing, for example, e_α instead of dx^α.

2. Proof of Theorem 1.1

We define the energy densities

$$e_1(u) = |Du|^4 \quad\text{and}\quad e_2(u) = |\nabla Du|^2.$$

(These should not be confused with the unit vectors in $\mathbb{R}^4$.) We also set $e(u) = e_1(u) + e_2(u)$. For $U \subset S$ we define $\mathscr{E}_i(u; U) = \int_U e_i(u)$, where $i = 1, 2$, and we define $\mathscr{E}(u; U) = \mathscr{E}_1(u; U) + \mathscr{E}_2(u; U)$.

Theorem 1.1 is a consequence of the following two propositions.

Proposition 2.1. *There exists an $\varepsilon_1 > 0$ such that the following holds: Let $(u_k) \subset W^{2,2}(S, N)$ be a sequence of biharmonic mappings (so $u_k \in C^\infty(S, N)$) and assume that $u \in W^{2,2}(S, N)$ is such that*

$$u_k \rightharpoonup u \quad \textit{weakly in } (W^{2,2} \cap W^{1,4})(S, \mathbb{R}^n). \tag{4}$$

Define

$$\Sigma_0 = \{x \in S : \liminf_{k\to\infty} \mathscr{E}(u_k; B_r(x)) \ge \varepsilon_1/2 \textit{ for all } r > 0\}.$$

Then $u \in C^\infty(S, N)$ is biharmonic and $u_k \to u$ in $C^2_{\mathrm{loc}}(S \setminus \Sigma_0, N)$. Moreover, there exist functions θ_1, $\theta_2 : \Sigma_0 \to (0, \infty)$ such that $\theta_1(x) \geq \varepsilon_1$ for all $x \in \Sigma_0$ and

$$\mathscr{L}^4 \lfloor e_i(u_k) \overset{*}{\rightharpoonup} \mathscr{L}^4 \lfloor e_i(u) + \sum_{x \in \Sigma_0} \theta_i(x) \delta_{\{x\}} \quad \text{for } i = 1, 2 \tag{5}$$

weakly-$*$ in the dual space of $C^0_0(S)$.

Remarks. (i) By Remark (i) following Theorem 1.1, the hypothesis (2) implies (4) for a subsequence.

(ii) The measures $\sum_{x \in \Sigma_0} \theta_i(x) \delta_{\{x\}}$ are called defect measures. Their common support Σ_0 is empty if and only if the convergence (4) is strong. In that case the last sum in (5) is defined to be zero.

Proposition 2.2. *Let u_k, u, Σ_0 and θ_i be as in Proposition 2.1. Then, for each $x \in \Sigma_0$, there exists $M_x \in \mathbb{N}$ and biharmonic mappings $v^x_1, \ldots, v^x_{M_x} \in C^\infty(\mathbb{R}^4, N)$ such that $\theta_i(x) = \sum_{j=1}^{M_x} \mathscr{E}_i(v^x_j; \mathbb{R}^4)$. In particular,*

$$\lim_{k \to \infty} \mathscr{E}_i(u_k; S) = \mathscr{E}_i(u; S) + \sum_{x \in \Sigma_0} \sum_{j=1}^{M_x} \mathscr{E}_i(v^x_j; \mathbb{R}^4) \quad \text{for } i = 1, 2.$$

For the proof of Proposition 2.1 we need three auxiliary results. The following lemma is a simple consequence of [Moser 2008, Theorem 2.1]:

Lemma 2.1. *There exists a universal constant C such that the following holds: Let $r > 0$, let $u \in W^{2,2}(B_r, N)$ and let $X \in L^2(B_r, \mathbb{R}^n)$ be a section of $u^{-1}TN$. If $\nabla^u X \in L^2(B_r)$ then $X \in L^4(B_r)$, and*

$$\|X\|_{L^4(B_r)} \leq C(\|\nabla^u X\|_{L^2(B_r)} + r^{-1}\|X\|_{L^2(B_r)}).$$

For $u \in C^k$ we introduce the notation $[u]_{C^k}(x) = \sum_{j=1}^k |D^j u(x)|^{1/j}$. An obvious consequence of [Scheven 2009, Lemma 5.3] is the following:

Lemma 2.2. *There exists $\varepsilon_1 > 0$ such that, for all $r > 0$ and for all biharmonic $u \in C^\infty(B_r, N)$ satisfying*

$$\int_{B_r} |Du|^4 \leq \varepsilon_1 \quad \text{we have} \quad \sup_{x \in B_{r/2}} |x|[u]_{C^3}(x) \leq 1.$$

The following lemma shows that isolated singularities of biharmonic mappings are removable.

Lemma 2.3. *Let $\Sigma \subset S$ be finite and let $u \in W^{2,2}(S, N)$ be biharmonic on $S \setminus \Sigma$. Then u is biharmonic on S. In particular, $u \in C^\infty(S, N)$.*

Proof. This proof closely follows that of [Jost 2005, Lemma 8.5.3]. We assume without loss of generality that $S = B_1$ and that $\Sigma = \{0\}$. Then (1) is equivalent to

$$\int_{B_1} \nabla_\alpha \partial_\beta u \cdot \nabla_\alpha \nabla_\beta \phi = \int_{B_1} f(u, Du \otimes Du \otimes D^2 u) \cdot \phi \tag{6}$$

for some $\mathbb{R}^n$-valued mapping f that is smooth in the first argument and linear in the second argument. Since u is biharmonic on $B_1 \setminus \{0\}$, Equation (6) is satisfied for all $\phi \in (L^\infty \cap W^{2,2}_0)(B_1 \setminus \{0\}, \mathbb{R}^n)$ that are sections of $u^{-1}TN$. From the properties of f we deduce that

$$|f(u, Du \otimes Du \otimes D^2 u)| \leq C(|D^2 u|^2 + |Du|^4). \tag{7}$$

Hence $f(u, Du \otimes Du \otimes D^2u) \in L^1(B_1, \mathbb{R}^n)$. For small $R \in (0, 1)$ we set

$$\tau_R(t) = \begin{cases} 0 & \text{for } t \in [0, R^2], \\ 1 - \log(t/R)/|\log R| & \text{for } t \in [R^2, R], \\ 1 & \text{for } t \in [R, 1). \end{cases}$$

One readily checks that

$$\lim_{R \to 0} \int_{B_1} |D^2 \tau_R(|x|)|^2 + |D\tau_R(|x|)|^4 \, dx = 0. \tag{8}$$

Now let $\phi \in (L^\infty \cap W^{2,2})(B, \mathbb{R}^n)$ be a section of $u^{-1}TN$. Then, for all $R \in (0, 1)$,

$$\phi_R(x) = \tau(|x|)\phi(x)$$

is still a section of $u^{-1}TN$, and $\phi_R \in (L^\infty \cap W_0^{2,2})(B_1 \setminus \{0\}, \mathbb{R}^n)$. Hence it is an admissible test function for (6). Using (7) and (8) it is easy to check that (6) holds for all ϕ as above, that is, u is biharmonic. Since $u \in W^{2,2}(S, N)$, Remark (ii) to Theorem 1.1 implies that $u \in C^\infty(S, N)$. □

Proof of Proposition 2.1. Clearly (4) implies $\limsup_{k\to\infty} \mathscr{E}(u_k; S) < \infty$. Hence Σ_0 is finite whatever the choice of ε_1. We choose ε_1 as in the statement of Lemma 2.2. Then the Arzèla–Ascoli theorem implies that $u_k \to u$ in $C^2_{\text{loc}}(S \setminus \Sigma_0, N)$. Hence u is biharmonic on $S \setminus \Sigma_0$. Lemma 2.3 therefore implies that $u \in C^\infty(S, N)$ and that u is biharmonic on S.

Weak lower semicontinuity of the L^2-norm and (4) imply the existence of (positive) Radon measures μ_1 and μ_2 on S such that

$$\mathscr{L}^4 \lfloor e_i(u_k) \overset{*}{\rightharpoonup} \mathscr{L}^4 \lfloor e_i(u) + \mu_i \quad \text{for } i = 1, 2. \tag{9}$$

We claim that

$$\mu_1(\{x\}) \geq \varepsilon_1 \quad \text{for all } x \in \operatorname{spt} \mu_1. \tag{10}$$

In fact, let $x \in S$ be such that $\mu_1(\{x\}) < \varepsilon_1$. Then by (9) there exists $r > 0$ such that

$$\limsup_{k\to\infty} \int_{B_r(x)} e_1(u_k) \leq \int_{B_r(x)} e_1(u) + \mu_1(\overline{B}_r(x)) < \varepsilon_1.$$

Thus $u_k \to u$ in $C^2(B_{r/2}(x))$ by Lemma 2.2 and the Arzèla–Ascoli theorem. (First only for a subsequence, but all subsequences must converge to the same limit u because $u_k \rightharpoonup u$ in $W^{2,2}(S, \mathbb{R}^n)$.) Thus $\mu_1(B_r/2(x)) = 0$, so $x \notin \operatorname{spt} \mu_1$. This proves (10), which in turn implies that $\operatorname{spt} \mu_1$ is finite and that $\mu_1 = \sum_{x \in \operatorname{spt} \mu_1} \theta_1(x)\delta_{\{x\}}$ for a function $\theta_1 : \operatorname{spt} \mu_1 \to [\varepsilon_1, \infty)$.

If $x \notin \operatorname{spt} \mu_1$, then (9) implies that

$$\inf_{r>0} \lim_{k\to\infty} \int_{B_r(x)} e_1(u_k) = \inf_{r>0} \int_{B_r(x)} e_1(u) = 0. \tag{11}$$

On the other hand, if $x \in \operatorname{spt} \mu_1$ then there exists $r > 0$ such that $B_{2r}(x) \cap \operatorname{spt} \mu_1 = \{x\}$ because $\operatorname{spt} \mu_1$ is finite. Thus $\mu(\partial B_r(x)) = 0$, and so (9) implies

$$\lim_{k\to\infty} \int_{B_r(x)} e_1(u_k) = \int_{B_r(x)} e_1(u) + \mu_1(\{x\}).$$

We conclude that

$$\inf_{r>0} \lim_{k\to\infty} \int_{B_r(x)} e_1(u_k) = \mu_1(\{x\}) \quad \text{for all } x \in S. \tag{12}$$

Now (12) together with (10) imply that $\operatorname{spt}\mu_1 \subset \Sigma_0$. On the other hand, if $x \notin \operatorname{spt}\mu_1$ then (11) and Lemma 2.2 imply that there is $r > 0$ such that $u_k \to u$ on $C^2(B_r(x), N)$; hence $x \notin \operatorname{spt}\mu_2$ and $x \notin \Sigma_0$. Thus $\operatorname{spt}\mu_2 \subset \operatorname{spt}\mu_1 = \Sigma_0$. It remains to check that $\operatorname{spt}\mu_1 \subset \operatorname{spt}\mu_2$. But (9) implies that, for $r \in (0, \operatorname{dist}_{\partial S}(x))$,

$$\limsup_{k\to\infty} \int_{B_r(x)} \Big(\frac{|Du_k|^2}{r^2} + e_2(u_k)\Big) \le \int_{B_r(x)} \Big(\frac{|Du|^2}{r^2} + e_2(u)\Big) + \mu_2(\overline{B}_r(x)), \tag{13}$$

because by Sobolev embedding we have $Du_k \to Du$ strongly in L^2. If $x \notin \operatorname{spt}\mu_2$, then the infimum over $r > 0$ of the right side of (13) is zero, since $Du \in L^4$. Hence Lemma 2.1 implies that $x \notin \Sigma_0$. □

For the proof of Proposition 2.2 we will need the following three lemmas:

Lemma 2.4. *There exists a modulus of continuity ω (that is, $\omega \in C^0([0,\infty))$ is nondecreasing and $\omega(0) = 0$) such that, whenever $r > 0$ and $u \in W^{2,2}(B_r, N)$ is biharmonic, then*

$$\operatorname{dist}_{\partial B_r}(x)[u]_{C^3}(x) \le \omega\Big(\int_{B_r} |Du|^4\Big) \quad \textit{for all } x \in B_r.$$

Proof. Notice that $u \in C^\infty(B_r, N)$ by Remark (ii) to Theorem 1.1. The claim follows from a scaled version of [Scheven 2009, Lemma 5.3] and from the fact that, by Jensen's inequality,

$$\Big(\rho^{-2}\int_{B_\rho(a)} |Du|^2\Big)^2 \le \int_{B_\rho(a)} |Du|^4. \qquad \square$$

We will also need the following crucial estimate.

Lemma 2.5. *There exists a constant C_3 such that the following holds: For all $R \in (0, 3/8)$ and for all biharmonic $u \in C^\infty(B_1, N)$ satisfying*

$$\varepsilon := \sup_{\rho\in(R,1/2)} \mathscr{E}(u; B_{2\rho} \setminus B_\rho) \le C_3^{-1}$$

we have

$$\mathscr{E}(u; B_1 \setminus B_R) \le C_3\omega(\varepsilon) + 2\varepsilon. \tag{14}$$

Here, ω is as in the conclusion of Lemma 2.4.

The proof of Lemma 2.5 will be given in Section 3.

Finally, we will need the existence of a uniform lower bound on the energy of nonconstant entire biharmonic mappings. An analogous fact is well known for harmonic mappings and also for other kinds of biharmonic mappings; see for example [Wang 2004b].

Lemma 2.6. *There exists a constant $\alpha > 0$ such that $\mathscr{E}(u; \mathbb{R}^4) \ge \alpha$ for every nonconstant biharmonic mapping $u \in C^\infty(\mathbb{R}^4, N)$.*

Proof. If the claim were false then there would exist nonconstant biharmonic $u_m \in C^\infty(\mathbb{R}^4, N)$ such that $\lim_{m\to\infty} \mathscr{E}(u_m; \mathbb{R}^4) = 0$. After passing to a subsequence we have $Du_m \to 0$ pointwise almost everywhere. Therefore, since u_m is nonconstant and since Du_m is continuous, there exist $x_m \in \mathbb{R}^4$ such that $r_m := |Du_m(x_m)|$ are nonzero but $\lim_{m\to\infty} r_m = 0$. Define $\tilde{u}_m(x) = u_m(x_m + x/r_m)$. Then $\mathscr{E}(\tilde{u}_m; \mathbb{R}^4) = \mathscr{E}(u_m; \mathbb{R}^4)$ converges to zero as $m \to \infty$. By Lemma 2.2 this implies the existence of a constant mapping u such that $\tilde{u}_m \to u$ in $C^2_{\mathrm{loc}}(\mathbb{R}^4, N)$. But on the other hand, $|D\tilde{u}_m(0)| = 1$ for all m, so $|Du(0)| = 1$. This contradiction finishes the proof. □

Proof of Proposition 2.2. By Proposition 2.1 we have $u_k, u \in C^\infty(S, N)$. Since the case $\Sigma_0 = \varnothing$ is trivial, we assume that Σ_0 is nonempty. After translating, rescaling (the energy $\mathscr{E}$ is scaling invariant) and restricting, we may assume that $\Sigma_0 = \{0\}$ and that $S = B_1$. By Proposition 2.1 we have $u_k \rightharpoonup u$ weakly in $(W^{2,2} \cap W^{1,4})(B_1, \mathbb{R}^n)$ and $u_k \to u$ in $C^2_{\mathrm{loc}}(B_1 \setminus \{0\}, N)$. Moreover, there is some

$$\theta \geq \varepsilon_1 \tag{15}$$

such that

$$\mathscr{L}^4 \lfloor e(u_k) \overset{*}{\rightharpoonup} \mathscr{L}^4 \lfloor e(u) + \theta \delta_{\{0\}}. \tag{16}$$

Let $\varepsilon \in (0, 1)$ be such that $C_3\omega(\varepsilon) + 3\varepsilon \leq \min\{\alpha/4, \varepsilon_1/4\}$, where ω is as in Lemma 2.4, C_3 is as in Lemma 2.5 and ε_1 is as in Lemma 2.2. Since $u \in W^{2,2}(B_1, \mathbb{R}^n)$, there exists $Q \in (0, 1)$ such that

$$\int_{B_Q} e(u) \leq \varepsilon/2. \tag{17}$$

We claim that there exists a sequence $R_k \to 0$ such that, for all k large enough,

$$\mathscr{E}(u_k; B_{2\rho} \setminus B_\rho) \leq \varepsilon \quad \text{for all } \rho \in [R_k, Q/2], \tag{18}$$

$$\mathscr{E}(u_k; B_{2R_k} \setminus B_{R_k}) = \varepsilon. \tag{19}$$

In fact, set

$$\mathscr{R}_k = \{r \in (0, Q/2) : \mathscr{E}(u_k; B_{2r} \setminus B_r) > \varepsilon\}.$$

If infinitely many of the $\mathscr{R}_k$ were empty, Lemma 2.5 would imply that there exists $k_i \to \infty$ such that $\mathscr{E}(u_{k_i}; B_Q \setminus B_{r_i}) \leq C_3\omega(\varepsilon) + 2\varepsilon$ for any sequence $r_i \to 0$. Choosing this sequence in such a way that $\mathscr{E}(u_{k_i}; B_{r_i}) \leq \varepsilon$ for all i, we would conclude that $\mathscr{E}(u_{k_i}; B_Q) \leq C_3\omega(\varepsilon) + 3\varepsilon \leq \varepsilon_1/4$, contradicting (15).

Thus, for k large, $\mathscr{R}_k \neq \varnothing$ and we can define $R_k = \sup \mathscr{R}_k$. Clearly $R_k > 0$ because $\int_{B_{2r}\setminus B_r} e(u_k) \leq \int_{B_{2r}} e(u_k) \to 0$ as $r \to 0$. On the other hand, $R_k \to 0$, since otherwise $\rho = \frac{1}{2} \liminf_{k\to\infty} R_k$ is positive, so

$$\limsup_{k\to 0} \int_{B_{2R_k}\setminus B_{R_k}} e(u_k) \leq \lim_{k\to 0} \int_{B_Q\setminus B_\rho} e(u_k) = \int_{B_Q\setminus B_\rho} e(u) \leq \varepsilon/2$$

by (17). This contradicts the fact that R_k is contained in the closure of $\mathscr{R}_k$, which by continuity of $r \mapsto \int_{B_{2r}\setminus B_r} e(u_k)$ implies that $\int_{B_{2R_k}\setminus B_{R_k}} e(u_k) \geq \varepsilon$. This also proves (19). Then (18) follows from the definition of R_k.

Combining (18) with (a scaled version of) Lemma 2.5, we conclude that

$$\mathscr{E}(u_k;\, B_Q \setminus B_{R_k}) \le C_3\omega(\varepsilon) + 2\varepsilon \le \alpha/4. \tag{20}$$

Set $v_k(x) = u_k(R_k x)$. Then by (16)

$$\limsup_{k\to\infty} \mathscr{E}(v_k;\, B_R) = \limsup_{k\to\infty} \mathscr{E}(u_k;\, B_{RR_k}) \le \inf_{\rho>0} \limsup_{k\to\infty} \mathscr{E}(u_k;\, B_\rho) = \theta \tag{21}$$

for all $R > 0$. Set

$$\Sigma^{(1)} = \{x \in \mathbb{R}^4 : \liminf_{k\to\infty} \mathscr{E}(v_k;\, B_r(x)) \ge \varepsilon_1/2 \ \text{ for all } r > 0\}.$$

By (21) we can apply Proposition 2.1 to each B_R. We conclude that $\Sigma^{(1)}$ is locally finite and that there exists a biharmonic mapping $v \in C^\infty(\mathbb{R}^4, N)$ such that, after passing to a subsequence, $v_k \rightharpoonup v$ weakly in $(W^{1,4}_{\text{loc}} \cap W^{2,2}_{\text{loc}})(\mathbb{R}^4, \mathbb{R}^n)$ and

$$v_k \to v \quad \text{in } C^2_{\text{loc}}(\mathbb{R}^4 \setminus \Sigma^{(1)}, \mathbb{R}^n), \tag{22}$$

and we find that there are a functions $\theta^{(1)}_1, \theta^{(1)}_2 : \Sigma^{(1)} \to (0, \infty)$ such that

$$\mathscr{L}^4 \lfloor e_i(v_k) \overset{*}{\rightharpoonup} \mathscr{L}^4 \lfloor e_i(v) + \sum_{x\in\Sigma^{(1)}} \theta^{(1)}_i(x)\delta_{\{x\}} \quad \text{for } i = 1, 2. \tag{23}$$

On the other hand, the bound (20) implies that

$$\limsup_{k\to\infty} \mathscr{E}(v_k;\, B_R \setminus \overline{B}_1) \le C_3\omega(\varepsilon) + 2\varepsilon \quad \text{for all } R > 1.$$

Thus $\Sigma^{(1)} \subset \overline{B}_1$ (so $\Sigma^{(1)}$ is finite) and therefore

$$v_k \to v \text{ in } C^2_{\text{loc}}(\mathbb{R}^4 \setminus \overline{B}_1, \mathbb{R}^n)$$

by (22). From this and since $\mathscr{E}(v_k;\, B_2 \setminus \overline{B}_1) = \mathscr{E}(u_k;\, B_{2R_k} \setminus \overline{B}_{R_k}) = \varepsilon$ for all k by (19), we conclude that $\mathscr{E}(v;\, \mathbb{R}^4) \ge \varepsilon$. Hence Lemma 2.6 implies that $\mathscr{E}(v;\, \mathbb{R}^4) \ge \alpha$.

Claim #1. *For all $\eta > 0$, there exist $R > 1$ and $\rho \in (0, 1)$ such that*

$$\liminf_{k\to\infty} \mathscr{E}(u_k;\, B_\rho \setminus B_{RR_k}) \le \eta.$$

To prove this claim, let us first show that for all $\delta > 0$ there exist R and ρ and a sequence $k_i \to \infty$ such that

$$\mathscr{E}(u_{k_i};\, B_{2r} \setminus B_r) \le \delta \quad \text{for all } i \in \mathbb{N} \text{ and all } r \in [RR_{k_i}, \rho/2]. \tag{24}$$

In fact, assume that this were not the case. Then there would exist $\delta \in (0, \varepsilon)$ such that for all R and ρ, the set

$$\hat{\mathscr{R}}_k = \{r \in [RR_k, \rho/2] : \mathscr{E}(u_k;\, B_{2r} \setminus B_r) > \delta\}$$

is nonempty for all k large enough. We choose $R > 2$ so large and $\rho \in (0, Q)$ so small that

$$\mathcal{E}(v;\, B_{4\hat{R}} \setminus B_{\hat{R}/2}) \le \delta/4 \quad \text{for all } \hat{R} \ge R, \text{ and} \tag{25}$$

$$\mathcal{E}(u;\, B_\rho) \le \delta/4. \tag{26}$$

This is clearly possible because $e(v) \in L^1(\mathbb{R}^4)$. Let $\hat{R}_k = \sup \hat{\mathcal{R}}_k$, hence $\hat{R}_k \in [RR_k, \rho/2]$. Arguing as above for R_k, using (26) one readily checks that $\hat{R}_k \to 0$. We claim that

$$\hat{R}_k / R_k \to \infty. \tag{27}$$

Indeed, if this were not the case then (after passing to a subsequence) there would exist $\hat{R} \in [R, \infty)$ such that $\hat{R}_k / R_k \in [\hat{R}/2, 2\hat{R}]$ for k large enough. Thus by the definition of $\hat{R}_k$ and since $\hat{R} \ge R > 2$ and $\Sigma^{(1)} \subset \overline{B}_1$,

$$\delta \le \limsup_{k\to\infty} \mathcal{E}(u_k;\, B_{2\hat{R}_k} \setminus B_{\hat{R}_k}) \le \limsup_{k\to\infty} \mathcal{E}(v_k;\, B_{4\hat{R}} \setminus B_{\hat{R}/2}).$$

This contradiction to (25) shows that (27) must be true.

Now define $\hat{v}_k(x) = u_k(\hat{R}_k x)$. As done above for R_k and v_k, using the fact that $\delta \le \varepsilon$, one shows that there exists a nontrivial biharmonic mapping $\hat{v} \in C^\infty(\mathbb{R}^4, N)$ such that, after passing to a subsequence, $\hat{v}_k \rightharpoonup v$ in $(W^{2,2}_{\mathrm{loc}} \cap W^{1,4}_{\mathrm{loc}})(\mathbb{R}^4, \mathbb{R}^n)$. Since $\hat{v}$ is nontrivial, Lemma 2.6 implies that $\mathcal{E}(\hat{v}; \mathbb{R}^4) \ge \alpha$. Hence by (27) and since $\hat{R}_k \to 0$, for all $\hat{R} > 1$ we have

$$\begin{aligned}\liminf_{k\to\infty} \mathcal{E}(u_k;\, B_\rho \setminus B_{RR_k}) &\ge \liminf_{k\to\infty} \mathcal{E}(u_k;\, B_{\hat{R}\hat{R}_k} \setminus B_{RR_k}) \\ &= \liminf_{k\to\infty} \mathcal{E}(\hat{v}_k;\, B_{\hat{R}} \setminus B_{R(R_k/\hat{R}_k)}) \\ &\ge \sup_{r>0} \liminf_{k\to\infty} \mathcal{E}(\hat{v}_k;\, B_{\hat{R}} \setminus B_r) \ge \mathcal{E}(\hat{v};\, B_{\hat{R}})\end{aligned}$$

because $\hat{v}_k \rightharpoonup \hat{v}$ on $B_{\hat{R}}$. Taking the supremum over all $\hat{R} > 1$ and recalling that $\mathcal{E}(\hat{v}; \mathbb{R}^4) \ge \alpha$, we conclude that $\liminf_{k\to\infty} \mathcal{E}(u_k;\, B_\rho \setminus B_{RR_k}) \ge \alpha$. This contradiction to (20) concludes the proof of (24).

Combining Lemma 2.5 with (24) and choosing δ small enough shows that Claim #1 is true.

The results obtained so far apply to any $\theta > 0$. Now we argue by induction: Assume that $m \in \mathbb{N}$ is such that $\theta \in ((m-1)\alpha, m\alpha]$. If $m \ge 2$ then assume, in addition, that Proposition 2.2 is true for all $\theta \in (0, (m-1)\alpha]$. On one hand, for $i = 1, 2$, for all $R \in (1, \infty)$ and for all $\rho \in (0, 1)$ we have

$$\begin{aligned}\theta_i + \mathcal{E}_i(u;\, B_\rho) = \lim_{k\to\infty} \big(\mathcal{E}_i(u_k;\, B_\rho \setminus B_{RR_k}) + \mathcal{E}_i(u_k;\, B_{RR_k})\big) &\ge \lim_{k\to\infty} \mathcal{E}_i(v_k;\, B_R) \\ &= \mathcal{E}_i(v;\, B_R) + \sum_{x\in\Sigma^{(1)}} \theta_i^{(1)}(x).\end{aligned}$$

(First we used (5) and that $\mu_i(\partial B_\rho) = 0$ for all $\rho \in (0, 1)$, and then we used (23) together with the fact that $\Sigma^{(1)} \subset \overline{B}_1$.) Taking $\rho \to 0$ and $R \to \infty$ we conclude

$$\theta_i \ge \mathcal{E}_i(v;\, \mathbb{R}^4) + \sum_{x\in\Sigma^{(1)}} \theta_i^{(1)}(x) \quad \text{for both } i = 1, 2. \tag{28}$$

Hence

$$\theta \geq \mathscr{E}(v; \mathbb{R}^4) + \sum_{x \in \Sigma^{(1)}} \theta^{(1)}(x). \tag{29}$$

Since $\mathscr{E}(v; \mathbb{R}^4) \geq \alpha$ this implies that $\theta^{(1)}(x) \leq \theta - \alpha$ for all $x \in \Sigma^{(1)}$. If $m \geq 2$ we can thus apply the induction hypothesis to conclude that

$$\theta_i^{(1)}(x) = \sum_{j=1}^{M_x} \mathscr{E}_i(v_x^j; \mathbb{R}^4) \quad \text{for both } i = 1, 2. \tag{30}$$

Here $v_x^1, \dots, v_x^{M_x} \in C^\infty(\mathbb{R}^4, N)$ are biharmonic and $M_x \in (0, m-1]$ is a natural number. (If $m = 1$, then (29) implies that $\Sigma^{(1)} = \varnothing$ and that $\theta = \alpha = \mathscr{E}(v; \mathbb{R}^4)$. This concludes the proof of the case $m = 1$.)

On the other hand, for all $\rho \in (0, 1)$ and all $R > 1$,

$$\begin{aligned} \theta &\leq \lim_{k\to\infty} \big(\mathscr{E}(u_k; B_\rho \setminus B_{RR_k}) + \mathscr{E}(u_k; B_{RR_k})\big) \\ &\leq \liminf_{k\to\infty} \mathscr{E}(u_k; B_\rho \setminus B_{RR_k}) + \lim_{k\to\infty} \mathscr{E}(v_k; B_R) \\ &= \liminf_{k\to\infty} \mathscr{E}(u_k; B_\rho \setminus B_{RR_k}) + \mathscr{E}(v; B_R) + \sum_{x \in \Sigma^{(1)}} \theta^{(1)}(x)\delta_{\{x\}}. \end{aligned} \tag{31}$$

We used that $\Sigma^{(1)} \subset \overline{B}_1$, so $\lim_{k\to\infty} \mathscr{E}(v_k; B_R) = \mathscr{E}(v; B_R) + \sum_{x\in\Sigma^{(1)}} \theta^{(1)}(x)\delta_{\{x\}}$. Now let $\rho \to 0$ and $R \to \infty$ in (31) using Claim #1. We conclude that $\theta \leq \mathscr{E}(v; \mathbb{R}^4) + \sum_{x\in\Sigma^{(1)}} \theta^{(1)}(x)$. Thus by (29) and (30),

$$\theta = \mathscr{E}(v; \mathbb{R}^4) + \sum_{x \in \Sigma^{(1)}} \sum_{j=1}^{M_x} \mathscr{E}(v_x^j; \mathbb{R}^4).$$

Combining this with the inequalities (28) immediately implies that

$$\theta_i = \mathscr{E}(v; \mathbb{R}^4) + \sum_{x \in \Sigma^{(1)}} \sum_{j=1}^{M_x} \mathscr{E}_i(v_x^j; \mathbb{R}^4) \quad \text{for both } i = 1, 2. \qquad \square$$

3. Energy estimates on the "neck" region

The purpose of this section is to prove the following proposition.

Proposition 3.1. *There exists a constant C_1 such that the following holds: For all $R \in (0, 1/2)$ and for all biharmonic $u \in C^\infty(B_1, N)$ satisfying*

$$\varepsilon := \sup_{x \in B_1 \setminus \overline{B}_R} |x|[u]_{C^3}(x) < 1, \tag{32}$$

we have

$$\int_{B_1 \setminus B_R} |\nabla^u Du|^2 \leq C_1(\varepsilon + \mathscr{E}(u; B_1 \setminus B_R))\varepsilon. \tag{33}$$

Corollary 3.1. *There exists a constant C_2 such that the following holds: For all $R \in (0, 1/2)$ and for all biharmonic $u \in C^\infty(B_1, N)$ satisfying* (32), *we have*

$$\int_{B_1 \setminus B_R} \frac{|Du|^2}{|x|^2} \le C_2(\varepsilon + \mathscr{E}(u;\, B_1 \setminus B_R))\varepsilon. \tag{34}$$

If, in addition, $\varepsilon \le 1/(2(C_1 + C_2))$, then

$$\mathscr{E}(u;\, B_1 \setminus B_R) \le 2(C_1 + C_2)\varepsilon^2. \tag{35}$$

Proof. Set $\varepsilon = \sup_{x \in B_1 \setminus \overline{B}_R} |x| [u]_{C^3}(x)$. By (33) and by (63) from Lemma 5.2, we have

$$\int_{B_1 \setminus \overline{B}_R} \frac{|Du|^2}{|x|^2} \le C_1(\varepsilon + \mathscr{E}(u;\, B_1 \setminus \overline{B}_R))\varepsilon + 2\mathscr{H}^3(\partial B_1)\varepsilon^2.$$

This implies (34) because $\varepsilon < 1$. We clearly have

$$\int_{B_1 \setminus \overline{B}_R} |Du|^4 \le \varepsilon^2 \int_{B_1 \setminus \overline{B}_R} \frac{|Du|^2}{|x|^2}.$$

Thus (34) implies that

$$\int_{B_1 \setminus \overline{B}_R} |Du|^4 \le C_2(\varepsilon + \mathscr{E}(u;\, B_1 \setminus \overline{B}_R))\varepsilon^3.$$

Adding this to (33) yields

$$\mathscr{E}(u;\, B_1 \setminus \overline{B}_R) \le (C_1 + C_2)\varepsilon^2 + (C_1 + C_2)\mathscr{E}(u;\, B_1 \setminus \overline{B}_R)\varepsilon,$$

because $\varepsilon < 1$. Since $\varepsilon \le 1/(2(C_1 + C_2))$, we can absorb the second term into the left hand side. This yields (35). □

As a consequence of Corollary 3.1 we obtain Lemma 2.5:

Proof of Lemma 2.5. Set $\varepsilon = \sup_{\rho \in (R, 1/2)} \mathscr{E}(u;\, B_{2\rho} \setminus B_\rho)$. We claim that

$$|x| [u]_{C^3}(x) \le 4\omega(\varepsilon) \quad \text{for all } x \in B_{1/2} \setminus \overline{B}_{4R/3}. \tag{36}$$

In fact, let $x \in B_{1/2} \setminus \overline{B}_{4R/3}$ and apply Lemma 2.4 to the ball $B_{|x|/4}(x)$. This yields

$$\operatorname{dist}_{\partial B_{|x|/4}(x)}(x) [u]_{C^3}(x) \le \omega\Big(\int_{B_{|x|/4}(x)} |Du|^4\Big).$$

Since $B_{|x|/4}(x) \subset B_{3|x|/2} \setminus \overline{B}_{3|x|/4}$, this implies (36).

Applying (35) (with $B_{1/2}$ instead of B_1 and $B_{4R/3}$ instead of B_R) to (36) implies

$$\mathscr{E}(u;\, B_{1/2} \setminus B_{4R/3}) \le C\omega^2(\varepsilon) \tag{37}$$

for some constant C, provided that ε is small enough (since then $\omega(\varepsilon)$ is small, and so $|x| [u]_{C^3}(x)$ is small by (36)). Finally, note that by definition of ε we have $\mathscr{E}(u;\, B_1 \setminus B_{1/2}) + \mathscr{E}(u;\, B_{2R} \setminus B_R) \le 2\varepsilon$. Together with (37) and smallness of $\omega(\varepsilon)$ this implies (14). □

The rest of this section will be devoted to the proof of Proposition 3.1. We will use the notation

$$\partial_r u = e_r^\alpha \partial_\alpha u, \quad D_r u = \partial_r u \otimes e_r, \quad D_{S^3} u = Du - D_r u, \quad D^2 u = (\partial_\alpha \partial_\beta u) \otimes e_\alpha \otimes e_\beta.$$

Above and in what follows we tacitly sum over repeated indices. A short calculation shows that

$$D_{S^3} u = \big(|x| \partial_{\partial_\alpha e_r} u\big) \otimes e_\alpha. \tag{38}$$

Proof of Proposition 3.1. Since $u \in C^\infty(B_1, N)$, [Scheven 2009, Lemma 4.2] implies that (1) is equivalent to

$$\Delta^2 u = -\partial_\alpha E_\alpha[u] + G[u], \tag{39}$$

where $E_\alpha[u] = -\partial_\beta(A(u)(\partial_\alpha u, \partial_\beta u)) + F_\alpha[u]$, and $F_\alpha[u] : S \to (\mathbb{R}^4)^* \otimes \mathbb{R}^n$ and $G[u] : S \to \mathbb{R}^n$ are as in [Scheven 2009, Lemma 4.2], that is, $F_\alpha[u] = f_\alpha(u, \nabla Du \otimes Du)$ for functions f_α that are smooth in the first and linear in the second argument, and $G[u] = g_1(u, \nabla Du \otimes \nabla Du) + g_2(u, \nabla Du \otimes Du \otimes Du)$ for functions g_1 and g_2 that again are smooth in the first and linear in the second argument. Therefore,

$$|G[u]| \le C(|D^2 u|^2 + |Du|^4), \tag{40}$$

$$|E_\alpha[u]| \le C(|D^2 u||Du| + |Du|^3). \tag{41}$$

For $r_1 < r_2$ define the open annulus $A(r_1, r_2) = B_{r_2} \setminus \overline{B}_{r_1}$ and set $A = A(R, 1)$. (This should not be confused with the second fundamental form of N.) As we will show at the end of this proof, we may assume without loss of generality that $R = 2^{-L}$ for some integer $L > 1$.

Define $R_k = 2^k R$ and set $A_k = A(R_k, R_{k+1})$. Set

$$\varepsilon = \sup_{x \in B_1 \setminus \overline{B}_R} |x| [u]_{C^3}(x). \tag{42}$$

Following an idea used in [Sacks and Uhlenbeck 1981] and [Ding and Tian 1995] in the context of harmonic mappings, we introduce the unique radial mapping $q : A \to \mathbb{R}^n$ solving the following boundary value problem for all $k = 0, \dots, L$:

$$\Delta^2 q = 0 \quad \text{on } A_k, \tag{43}$$

$$q(R_k) = \frac{1}{\mathcal{H}^3(\partial B_{R_k})} \int_{\partial B_{R_k}} u \quad \text{and} \quad q'(R_k) = \frac{1}{\mathcal{H}^3(\partial B_{R_k})} \int_{\partial B_{R_k}} \partial_r u. \tag{44}$$

(For a radial function of the form $q(x) = \tilde{q}(|x|)$, we often write q instead of $\tilde{q}$.) Notice that q is indeed well and uniquely defined on each A_k by (43) and (44) because (43) is simply a fourth order ordinary differential equation on (R_k, R_{k+1}), since q is radial. (See Lemma 5.1 below for details.) The rest of this proof is divided into Lemma 3.1 and Lemma 3.2 below. Combining their conclusions one obtains that of Proposition 3.1.

Let us finally check that the case of arbitrary $R \in (0, 1)$ follows from the case when $R = 2^{-L}$. In fact, for general R let L be such that $2^L R \in [\frac{1}{2}, 1)$. The definition of ε implies that

$$\int_{A(2^L R, 1)} |\nabla Du|^2 \le \varepsilon^2 \int_{A(2^L R, 1)} |x|^{-4} \le \varepsilon^2 \mathcal{H}^3(\partial B_1) \log 2.$$

Applying Proposition 3.1 with $B_{2^L R}$ instead of B_1, the estimate (33) follows. □

Lemma 3.1. *For u, q and R as in the proof of Proposition 3.1 we have*

$$\int_A |D^2(u-q)|^2 \le C\Big(\varepsilon + \int_A |\nabla^u Du|^2 + |Du|^4\Big)\varepsilon, \tag{45}$$

$$\int_A \frac{|D(u-q)|^2}{|x|^2} \le C\Big(\varepsilon + \int_A |\nabla^u Du|^2 + |Du|^4\Big)\varepsilon. \tag{46}$$

Proof. Since $q|_{A_k}$ is a solution of a linear ordinary differential equation with smooth coefficients, it is C^∞ up to the boundary of A_k. Moreover, for $r \in (R_k, R_{k+1})$, by Lemma 5.1 there exists a universal constant C such that

$$|q'(r)| \le C(|q'(R_k)| + |q'(R_{k+1})| + R_k^{-1}|q(R_{k+1}) - q(R_k)|). \tag{47}$$

By (44) and by (42) this implies that $|u(x) - q(R_k)| \le \|Du\|_{L^\infty(\partial B_{R_k})} \cdot \mathrm{diam}(\partial B_{R_k})$ for all $x \in \partial B_{R_k}$ and all k. Therefore,

$$|q(R_{k+1}) - q(R_k)| \le \|Du\|_{L^\infty(A_k)}\,\mathrm{diam}\,A_k \le C\varepsilon \tag{48}$$

by (42) and because $\mathrm{diam}\,A_k \le CR_k$. Since $|x|$ is comparable to R_k on A_k and since k was arbitrary, we conclude from (47) and (48) and from (44) and (42) that $|x||Dq(x)| \le C\varepsilon$ for all $x \in A$. By (44) and by (42) this implies that $|u - q| \le C\varepsilon$. Summarizing, we have shown that

$$|(u-q)(x)| + |x||D(u-q)(x)| \le C\varepsilon \quad \text{for all } x \in A. \tag{49}$$

Notice that while (44) implies that $q \in C^1(A, \mathbb{R}^n)$ and that $q|_{A_k} \in C^\infty(\bar{A}_k, \mathbb{R}^n)$ for all k, in general $q \notin C^2(A; \mathbb{R}^n)$.

By partial integration one obtains, for arbitrary $v \in C^2(\bar{A}_k, \mathbb{R}^n)$,

$$\int_{A_k} |D^2 v|^2 = \int_{A_k} (\partial_\alpha\partial_\beta v)\cdot(\partial_\alpha\partial_\beta v) = \int_{A_k} (\Delta^2 v)\cdot v + \Big[\int_{\partial A_k} (\partial_r\partial_\beta v)\cdot\partial_\beta v - (\partial_r \Delta v)\cdot v\Big]_{r=R_k}^{R_{k+1}}.$$

Here and below we use the notation

$$\big[f(r)\big]_{r=t_1}^{t_2} := f(t_2) - f(t_1)$$

for functions $f \in C^0([t_1, t_2])$. Inserting $v = u - q$ and summing over $k = 0, \ldots, L$ yields

$$\begin{aligned}\int_A |D^2(u-q)|^2 &= \int_A (\Delta^2 u)\cdot(u-q) \\ &\quad + \sum_{k=0}^{L}\Big[\int_{\partial B_\rho} (\partial_r\partial_\beta(u-q))\cdot\partial_\beta(u-q) - (\partial_r\Delta(u-q))\cdot(u-q)\Big]_{\rho=R_k}^{R_{k+1}} \\ &= \int_A (\Delta^2 u)\cdot(u-q) + \Big[\int_{\partial B_\rho} \partial_r\partial_\beta u\cdot\partial_\beta(u-q) - \partial_r\Delta u\cdot(u-q)\Big]_{\rho=R}^{1} \\ &\quad - \sum_{k=0}^{L}\Big[\int_{\partial B_\rho} (\partial_r\partial_r q)(\rho)\cdot\partial_r(u-q)(x) - (\partial_r\Delta q)(\rho)\cdot(u-q)(x)\,d\mathcal{H}^3(x)\Big]_{\rho=R_k}^{R_{k+1}}.\end{aligned} \tag{50}$$

In the first step we used that $\Delta^2 q = 0$ on A_k. In the last step we used that the boundary integrals with continuous integrands cancel successively, and we used that q is radial. Since q is radial, the same is true for $\partial_r\partial_r q$ and $\partial_r \Delta q$; see (60). The choice of boundary conditions (44) implies that

$$(\partial_r\partial_r q)(\rho)\cdot\int_{\partial B_\rho}\partial_r(u-q)(x)\,d\mathcal{H}^3(x)=0 \quad\text{and}\quad (\partial_r\Delta q)(\rho)\cdot\int_{\partial B_\rho}(u-q)(x)\,d\mathcal{H}^3(x)=0$$

for all $\rho\in\{R_0, R_1, \dots, R_L\}$. So the sum in the last term in (50) is zero. (The discontinuous expressions $q''=\partial_r\partial_r q$ and q''' occurring in $\partial_r\Delta q$ must be understood in the trace sense: If ∂B_{R_k} belongs to ∂A_k then $q''(R_k)=\lim_{r\uparrow R_k} q''(r)$ and if ∂B_{R_k} belongs to ∂A_{k+1} then $q''(R_k)=\lim_{r\downarrow R_k} q''(r)$. These limits exists because, as noted above, $q|_{A_k}$ is smooth up to the boundary of A_k.)

To estimate the second term in (50) we use (49) and (42). This gives

$$\int_{\partial B_r}|\partial_r\partial_\beta u||\partial_\beta(u-q)|\le C\mathcal{H}^3(\partial B_r)\frac{\varepsilon}{r^2}\frac{\varepsilon}{r}\le C\varepsilon^2.$$

Similarly, $\int_{\partial B_r}|\partial_r\Delta u||u-q|\le C\varepsilon^2$. Thus (50) implies

$$\int_A|D^2(u-q)|^2\le\left|\int_A(\Delta^2u)\cdot(u-q)\right|+C\varepsilon^2. \tag{51}$$

To estimate the term $\left|\int_A(\Delta^2 u)\cdot(u-q)\right|$ in (51), we use (39) to replace $\Delta^2 u$. We obtain

$$\begin{aligned}\int_A(\Delta^2u)\cdot(u-q)&=\int_A(-\partial_\alpha E_\alpha[u])\cdot(u-q)+G[u]\cdot(u-q)\\&=\int_A E_\alpha[u]\cdot\partial_\alpha(u-q)+\int_A G[u]\cdot(u-q)-\left[\int_{\partial B_r}\frac{x_\alpha}{|x|}E_\alpha[u]\cdot(u-q)\right]_{r=R}^1.\end{aligned} \tag{52}$$

To estimate the last term in (52) we simply use that $|E_\alpha[u]|\le|D^2u||Du|+|Du|^3\le C\varepsilon^2/|x|^3$ pointwise by (41). Thus

$$\int_{\partial B_r}|E_\alpha[u]||u-q|\le C\varepsilon^3\mathcal{H}^3(\partial B_r)r^{-3}\le C\varepsilon^3$$

for both $r=1$ and $r=R$.

To estimate the second term in (52), we use (40) and (49) to find

$$\int_A|G[u]||u-q|\le C\varepsilon\int_A(|D^2u|^2+|Du|^4).$$

To estimate the first term in (52) notice that by (41) and by (49) we have

$$\int_A|E_\alpha[u]||D(u-q)|\le C\varepsilon\int_A|D^2u|\frac{|Du|}{|x|}+\frac{|Du|^3}{|x|}\le C\varepsilon\int_A\Big(|D^2u|^2+|Du|^4+\frac{|Du|^2}{|x|^2}\Big). \tag{53}$$

Applying Lemma 5.2 to $v=u$ with $r_1=R$ and $r_2=1$, we have

$$\int_A\frac{|Du|^2}{|x|^2}\le\int_A|D^2u|^2+\left[\frac{1}{r}\int_{\partial B_r}|Du|^2\right]_{r=R}^1.$$

The boundary terms can be estimated as above using the definition of ε, Thus

$$\int_A \frac{|Du|^2}{|x|^2} \le \int_A |D^2u|^2 + C\varepsilon^2.$$

So (53) implies

$$\int_A |E_\alpha[u]||D(u-q)| \le C\varepsilon\Big(\varepsilon^2 + \int_A |D^2u|^2 + |Du|^4\Big).$$

Since $|D^2u|^2 \le C(N)(|\nabla Du|^2 + |Du|^4)$ for some constant $C(N)$ depending only on the immersion $N \hookrightarrow \mathbb{R}^n$, this concludes the proof of (45).

To prove (46) we apply Lemma 5.2 to each restriction $(u-q)|_{A_k}$. This yields

$$\int_{A_k} \frac{|D(u-q)|^2}{|x|^2} \le \int_{A_k} |D^2(u-q)|^2 + \left[\frac{1}{r}\int_{\partial B_r} |D(u-q)|^2\right]_{r=R_k}^{R_{k+1}}.$$

When we sum over $k = 0, \dots, L$, the terms in square brackets cancel successively because $D(u-q)$ is continuous. After estimating the boundary terms on ∂B_1 and on ∂B_R using (42), this yields (46). ☐

Lemma 3.2. *For u, q and R as in the proof of Proposition 3.1 we have*

$$\int_{A(R,1)} |D^2(u-q)|^2 \ge \Big(\frac{1}{2} - \frac{\sqrt{2}}{3}\Big)\int_{A(R,1)} |\nabla^u Du|^2 - C\Big(\varepsilon + \int_A |\nabla^u Du|^2 + |Du|^4\Big)\varepsilon.$$

Proof. For $v \in C^\infty(S, \mathbb{R}^n)$ we have

$$D^2 v = DD_{S^3}v + DD_r v,$$

where $D_{S^3}v = Dv - D_r v$. Thus

$$|D^2v|^2 \ge |DD_{S^3}v|^2 + 2D(Dv - D_rv)\cdot DD_rv. \tag{54}$$

Now $D(Dv - D_rv)\cdot DD_rv$ equals

$$\begin{aligned}
&\partial_\alpha((\partial_\beta v)\otimes(e_\beta - e_r^\beta e_r))\cdot\partial_\alpha(\partial_\gamma v\otimes e_r^\gamma e_r)\\
&\quad= \big((\partial_\alpha\partial_\beta v)\otimes(e_\beta - e_r^\beta e_r) - (\partial_\beta v)\otimes\partial_\alpha(e_r^\beta e_r)\big)\cdot\big((\partial_\alpha\partial_\gamma v)\otimes e_r^\gamma e_r + (\partial_\gamma v)\otimes\partial_\alpha(e_r^\gamma e_r)\big)\\
&\quad= \big((\partial_\alpha\partial_\beta v)\otimes(e_\beta - e_r^\beta e_r)\big)\cdot\big((\partial_\gamma v)\otimes\partial_\alpha(e_r^\gamma e_r)\big)\\
&\qquad\qquad - |(\partial_\beta v)\otimes\partial_\alpha(e_r^\beta e_r)|^2 - (\partial_\beta v)\otimes\partial_\alpha(e_r^\beta e_r)\cdot(\partial_\alpha\partial_\gamma v)\otimes e_r^\gamma e_r\\
&\quad= \big((\partial_\alpha\partial_\beta v)\otimes(e_\beta - e_r^\beta e_r)\big)\cdot\big((\partial_r v)\otimes(\partial_\alpha e_r)\big)\\
&\qquad\qquad - |(\partial_\beta v)\otimes\partial_\alpha(e_r^\beta e_r)|^2 - (\partial_\beta v)\otimes(\partial_\alpha e_r^\beta)e_r\cdot(\partial_\alpha\partial_\gamma v)\otimes e_r^\gamma e_r\\
&\quad= (\partial_{\partial_\alpha e_r}\partial_\alpha v)\cdot(\partial_r v) - |\partial_\beta v|^2\,|\partial_\alpha(e_r^\beta e_r)|^2 - \partial_{\partial_\alpha e_r}v\cdot(\partial_r\partial_\alpha v).
\end{aligned}$$

This shows that

$$\begin{aligned}
D(Dv - D_rv)\cdot DD_rv &\ge -2|De_r||D^2v||Dv| - 2|Dv|^2|De_r|^2\\
&\ge -(|D^2v|^2 + C|De_r|^2|Dv|^2)
\end{aligned} \tag{55}$$

for some universal constant $C > 0$. Since $|De_r(x)|^2 = 3/|x|^2$, inserting (55) into the estimate (54) yields

$$3|D^2 v|^2 \geq |DD_{S^3} v|^2 - C|Dv|^2/|x|^2.$$

Inserting $v = u - q$, integrating and using that $D_{S^3} q = 0$ gives

$$\begin{aligned} 3\int |D^2(u-q)|^2 &\geq \int |DD_{S^3} u|^2 - C\int \frac{|D(u-q)|^2}{|x|^2} \\ &\geq \int |\nabla Du - \nabla D_r u|^2 - C\int \frac{|D(u-q)|^2}{|x|^2} \\ &\geq \Big(1 - \frac{1}{\sqrt{2}}\Big)\int |\nabla Du|^2 + (1-\sqrt{2})\int |\nabla D_r u|^2 - C\int \frac{|D(u-q)|^2}{|x|^2}. \end{aligned}$$

In the second step we used that

$$Du = D_{S^3} u + D_r u$$

and the trivial estimate $|Df| \geq |\nabla^u f|$. By (58) the last line equals

$$\begin{aligned} \big(\tfrac{3}{2} - \sqrt{2}\big)\int |\nabla Du|^2 + (\sqrt{2}-1)\int \frac{|\nabla^u(|x|\partial_r u)|^2}{|x|^2} - C\int \frac{|D(u-q)|^2}{|x|^2} \\ + \frac{1-\sqrt{2}}{2}\bigg[\int_{\partial B_r}\Big(\frac{3}{r}|Du|^2 + 2(\nabla_r^u \partial_r u)\cdot \partial_r u - \frac{2}{r}|\partial_r u|^2\Big)d\mathscr{H}^3\bigg]_{r=R}^{1}. \end{aligned}$$

The claim follows by dropping the second term, which is nonnegative, and noticing that the fourth term is dominated by ε^2 by (42) while, by (46), the third term is dominated by

$$\varepsilon\Big(\varepsilon + \int_A |\nabla Du|^2 + |Du|^4\Big). \qquad \square$$

4. An equality for stationary biharmonic mappings

The following lemma is true for mappings that are stationary with respect to the energy E_2 in the sense of [Moser 2008]. We do not need the precise definition here. We only remark that every smooth biharmonic mapping is also stationary. Therefore by Remark (ii) to Theorem 1.1, every $u \in W^{2,2}(S, N)$ that is biharmonic is also stationary. To recall the monotonicity formula from [Moser 2008], for $u \in W^{2,2}(B_1, N)$ we define

$$\mathscr{F}(r) = \frac{1}{4}\int_{B_r} |\nabla Du|^2 + \frac{1}{4}\int_{\partial B_r}\Big(\frac{3}{r}|Du|^2 + 2(D_r\partial_r u \cdot \partial_r u)\Big)d\mathscr{H}^3.$$

Theorem 3.1 in [Moser 2008] (see also [Hornung and Moser 2012]) then states that, if $u \in W^{2,2}(S, N)$ is stationary, then

$$\mathscr{F}(r_2) - \mathscr{F}(r_1) = \int_{B_{r_2}\setminus B_{r_1}}\Big(\frac{|\nabla^u |x|\partial_r u(x)|^2}{|x|^2} + 2\frac{|\partial_r u(x)|^2}{|x|^2}\,dx\Big) \tag{56}$$

for almost all r_1, r_2 with $0 < r_1 \leq r_2 \leq 1$. As a corollary to this fact we obtain the following lemma:

Lemma 4.1. *Let $u \in W^{2,2}(B_1, N)$ be stationary and let $R \in (0,1)$. Then*

$$\int_{B_1\setminus B_R} |\nabla^u D_r u|^2 = \int_{B_1\setminus B_R} \Big(\frac{1}{4}|\nabla^u Du|^2 + 2\frac{|\partial_r u|^2}{|x|^2}\Big) + \frac{1}{4}\Big[\int_{\partial B_r}\Big(\frac{3}{r}|Du|^2 - \frac{4}{r}|\partial_r u|^2 + 2(\nabla_r^u \partial_r u)\cdot \partial_r u\Big) d\mathcal{H}^3\Big]_{r=R}^1 \tag{57}$$

$$= \int_{B_1\setminus B_R}\Big(\frac{1}{2}|\nabla^u Du|^2 - \frac{|\nabla^u(|x|\partial_r u)|^2}{|x|^2}\Big) + \frac{1}{2}\Big[\int_{\partial B_r}\Big(\frac{3}{r}|Du|^2 + 2(\nabla_r^u\partial_r u)\cdot\partial_r u - \frac{2}{r}|\partial_r u|^2\Big)d\mathcal{H}^3\Big]_{r=R}^1. \tag{58}$$

We remark that Lemma 4.1 can be regarded as a biharmonic counterpart of [Sacks and Uhlenbeck 1981, Lemma 3.5].

Proof. First notice that $|\nabla D_r u|^2 = |\nabla\partial_r u|^2 + |De_r|^2|\partial_r u|^2$ and that $|De_r|^2 = 3/|x|^2$. Moreover, a short calculation using (38) shows that $|x|\nabla\partial_r u = \nabla(|x|\partial_r u) - D_r u$. Using these facts we calculate

$$\begin{aligned}|\nabla D_r u|^2 &= \Big|\frac{\nabla(|x|\partial_r u)}{|x|} - \frac{D_r u}{|x|}\Big|^2 + |De_r|^2|\partial_r u|^2\\ &= \frac{|\nabla(|x|\partial_r u)|^2}{|x|^2} + 4\frac{|\partial_r u|^2}{|x|^2} - \frac{2}{|x|^2}D(|x|\partial_r u)\cdot D_r u\\ &= \frac{|\nabla(|x|\partial_r u)|^2}{|x|^2} + 4\frac{|\partial_r u|^2}{|x|^2} - \operatorname{div}\Big(\frac{|\partial_r u|^2}{|x|^2}x\Big).\end{aligned} \tag{59}$$

Integrating over $B_1 \setminus B_R$ and using (56) we obtain (57). On the other hand, (59) clearly equals

$$2\Big(\frac{|\nabla(|x|\partial_r u)|^2}{|x|^2} + 2\frac{|\partial_r u|^2}{|x|^2}\Big) - \frac{|\nabla(|x|\partial_r u)|^2}{|x|^2} - \operatorname{div}\Big(\frac{|\partial_r u|^2}{|x|^2}x\Big).$$

Integrating this over $B_1 \setminus B_R$ and using (56) we obtain (58). □

5. Appendix

Lemma 5.1. *There exists a universal constant C_4 such that for all $R > 0$ and for all radial solutions $q \in C^\infty(B_{2R}\setminus \overline{B}_R, \mathbb{R}^n)$ of the equation $\Delta^2 q = 0$ on $B_{2R}\setminus\overline{B}_R$, the following estimate holds*:

$$\|q'\|_{C^0(B_{2R}\setminus\overline{B}_R,\mathbb{R}^n)} \le C_4\big(|q'(R)| + |q'(2R)| + R^{-1}|q(2R) - q(R)|\big).$$

Proof. After rescaling we may assume without loss of generality that $R = 1$. Since

$$\Delta q(x) = 3\frac{q'(|x|)}{|x|} + q''(|x|), \tag{60}$$

we see that $\Delta^2 q = 0$ is equivalent to q' being a solution of the third order system

$$\frac{3}{t}\Big(\frac{3f(t)}{t} + f'(t)\Big)' + \Big(\frac{3f(t)}{t} + f'(t)\Big)'' = 0. \tag{61}$$

Denote by $X \subset C^\infty(B_2 \setminus B_1, \mathbb{R}^n)$ the (at most three dimensional) subspace of solutions to (61). Denote by $L : X \to \mathbb{R}^3$ the functional given by $Lf = (f(1), f(2), \int_1^2 f)$. We claim that L is surjective.

In fact, let $a \in \mathbb{R}^3$. By the direct method it is easy to see that the functional $v \mapsto \int_{B_2\setminus B_1} |\nabla^2 v|^2$ has a minimizer in the class of all radial $v \in W^{2,2}$ satisfying $v'(1) = a_1$ and $v'(2) = a_2$ and $v(2) - v(1) = a_3$. This minimizer q satisfies the Euler–Lagrange equation $\Delta^2 q = 0$, so its radial derivative q' solves the ODE (61). Thus $q' \in X$ and $Lq' = a$. This proves surjectivity of L.

Hence X is three dimensional and L is in fact bijective. Since all norms on X are equivalent and since the inverse of L is of course bounded, we conclude that $\|f\|_{C^0((1,2),\mathbb{R}^n)} \le C|Lf|$ for all $f \in X$. This implies the claim. □

Lemma 5.2. *Let* $0 < r_1 < r_2 \le 1$ *and assume that* $v \in W^{2,2}(B_{r_2} \setminus \overline{B}_{r_1}, \mathbb{R}^n)$. *Then*

$$\int_{B_{r_2}\setminus B_{r_1}} \frac{|Dv|^2}{|x|^2} \le \int_{B_{r_2}\setminus B_{r_1}} |D^2 v|^2 + \left[\frac{1}{r}\int_{\partial B_r} |Dv|^2\right]_{r=r_1}^{r_2}. \tag{62}$$

If $v \in W^{2,2}(B_{r_2} \setminus \overline{B}_{r_1}, N)$ *then*

$$\int_{B_{r_2}\setminus B_{r_1}} \frac{|Dv|^2}{|x|^2} \le \int_{B_{r_2}\setminus B_{r_1}} |\nabla^v Dv|^2 + \left[\frac{1}{r}\int_{\partial B_r} |Dv|^2\right]_{r=r_1}^{r_2}. \tag{63}$$

Proof. For $v \in C^2(A(r_1, r_2), \mathbb{R}^n)$ we have

$$2\frac{|Dv|^2}{|x|^2} = \operatorname{div}\left(\frac{|Dv|^2}{|x|^2}\, x\right) - \frac{\partial_r |Dv|^2}{|x|}.$$

Hence if Dv is continuous up to the boundary of $A(r_1, r_2)$ then

$$\begin{aligned} 2\int_{A(r_1,r_2)} \frac{|Dv|^2}{|x|^2} &= -\int_{A(r_1,r_2)} \frac{\partial_r |Dv|^2}{|x|} + \left[\int_{\partial B_r} \frac{|Dv|^2}{|x|^2} x \cdot \frac{x}{|x|}\right]_{r=r_1}^{r_2} \\ &= -2\int_{A(r_1,r_2)} (\partial_r \partial_\alpha v) \cdot \frac{\partial_\alpha v}{|x|} + \left[\frac{1}{r}\int_{\partial B_r} |Dv|^2\right]_{r=r_1}^{r_2}. \end{aligned} \tag{64}$$

By density and by continuity of the trace operator, this equality remains true for $v \in W^{2,2}(A(r_1, r_2), \mathbb{R}^n)$. We conclude that

$$2\int_{A(r_1,r_2)} \frac{|Dv|^2}{|x|^2} \le \int_{A(r_1,r_2)} |D^2 v|^2 + \int_{A(r_1,r_2)} \frac{|Dv|^2}{|x|^2} + \left[\frac{1}{r}\int_{\partial B_r} |Dv|^2\right]_{r=r_1}^{r_2}.$$

Absorbing the second term on the right into the left hand side yields (62).

If v takes values in N then the first term on the right hand side of (64) equals

$$-2\int_{A(r_1,r_2)} (\nabla_r^v \partial_\alpha v) \cdot \frac{\partial_\alpha v}{|x|}$$

because $\partial_\alpha v(x) \in T_{v(x)} N$ for all x. Estimating as above yields (63). □

References

[Chang et al. 1999] S.-Y. A. Chang, L. Wang, and P. C. Yang, "A regularity theory of biharmonic maps", *Comm. Pure Appl. Math.* **52**:9 (1999), 1113–1137. MR 2000j:58025 Zbl 0953.58013

[Ding and Tian 1995] W. Ding and G. Tian, "Energy identity for a class of approximate harmonic maps from surfaces", *Comm. Anal. Geom.* **3**:3-4 (1995), 543–554. MR 97e:58055 Zbl 0855.58016

[Hélein 1990] F. Hélein, "Régularité des applications faiblement harmoniques entre une surface et une sphère", *C. R. Acad. Sci. Paris Sér. I Math.* **311**:9 (1990), 519–524. MR 92a:58034 Zbl 0728.35014

[Hélein 1991] F. Hélein, "Régularité des applications faiblement harmoniques entre une surface et une variété riemannienne", *C. R. Acad. Sci. Paris Sér. I Math.* **312**:8 (1991), 591–596. MR 92e:58055 Zbl 0728.35015

[Hornung and Moser 2012] P. Hornung and R. Moser, "Intrinsically p-biharmonic maps", in preparation, 2012.

[Jost 1991] J. Jost, *Two-dimensional geometric variational problems*, Wiley, Chichester, 1991. MR 92h:58045 Zbl 0729.49001

[Jost 2005] J. Jost, *Riemannian geometry and geometric analysis*, 4th ed., Springer, Berlin, 2005. MR 2006c:53002 Zbl 1083.53001

[Lamm and Rivière 2008] T. Lamm and T. Rivière, "Conservation laws for fourth order systems in four dimensions", *Comm. Partial Differential Equations* **33**:1-3 (2008), 245–262. MR 2009h:35095 Zbl 1139.35328

[Lin and Rivière 2002] F.-H. Lin and T. Rivière, "Energy quantization for harmonic maps", *Duke Math. J.* **111**:1 (2002), 177–193. MR 2002k:58036 Zbl 1014.58008

[Montaldo and Oniciuc 2006] S. Montaldo and C. Oniciuc, "A short survey on biharmonic maps between Riemannian manifolds", *Rev. Un. Mat. Argentina* **47**:2 (2006), 1–22. MR 2008a:53063 Zbl 1140.58004

[Moser 2008] R. Moser, "A variational problem pertaining to biharmonic maps", *Comm. Partial Differential Equations* **33**:7-9 (2008), 1654–1689. MR 2009h:58034 Zbl 1154.58007

[Parker 1996] T. H. Parker, "Bubble tree convergence for harmonic maps", *J. Differential Geom.* **44**:3 (1996), 595–633. MR 98k:58069 Zbl 0874.58012

[Sacks and Uhlenbeck 1981] J. Sacks and K. Uhlenbeck, "The existence of minimal immersions of 2-spheres", *Ann. of Math.* (2) **113**:1 (1981), 1–24. MR 82f:58035 Zbl 0462.58014

[Scheven 2009] C. Scheven, "An optimal partial regularity result for minimizers of an intrinsically defined second-order functional", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26**:5 (2009), 1585–1605. MR 2010i:49061 Zbl 05612918

[Struwe 2008] M. Struwe, "Partial regularity for biharmonic maps, revisited", *Calc. Var. Partial Differential Equations* **33**:2 (2008), 249–262. MR 2009b:35068 Zbl 1151.58011

[Wang 2004a] C. Wang, "Biharmonic maps from $\mathbb{R}^4$ into a Riemannian manifold", *Math. Z.* **247**:1 (2004), 65–87. MR 2005c:58030 Zbl 1064.58016

[Wang 2004b] C. Wang, "Remarks on biharmonic maps into spheres", *Calc. Var. Partial Differential Equations* **21**:3 (2004), 221–242. MR 2005e:58026 Zbl 1060.58011

[Wang 2004c] C. Wang, "Stationary biharmonic maps from $\mathbb{R}^m$ into a Riemannian manifold", *Comm. Pure Appl. Math.* **57**:4 (2004), 419–444. MR 2005e:58027 Zbl 1055.58008

Received 4 Nov 2009. Revised 24 Nov 2010. Accepted 25 Jan 2011.

PETER HORNUNG: hornung@mis.mpg.de
Max-Planck-Institut für Mathematik in den Naturwissenschaften, Inselstraße. 22, 04103 Leipzig, Germany

ROGER MOSER: r.moser@bath.ac.uk
Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, United Kingdom

ANALYSIS AND PDE
Vol. 5, No. 1, 2012
dx.doi.org/10.2140/apde.2012.5.81

THE WAVE EQUATION ON ASYMPTOTICALLY ANTI DE SITTER SPACES

ANDRÁS VASY

In this paper we describe the behavior of solutions of the Klein–Gordon equation, $(\Box_g + \lambda)u = f$, on Lorentzian manifolds (X°, g) that are anti de Sitter-like (AdS-like) at infinity. Such manifolds are Lorentzian analogues of the so-called Riemannian conformally compact (or asymptotically hyperbolic) spaces, in the sense that the metric is conformal to a smooth Lorentzian metric $\hat{g}$ on X, where X has a nontrivial boundary, in the sense that $g = x^{-2}\hat{g}$, with x a boundary defining function. The boundary is conformally timelike for these spaces, unlike asymptotically de Sitter spaces studied before by Vasy and Baskin, which are similar but with the boundary being conformally spacelike.

Here we show local well-posedness for the Klein–Gordon equation, and also global well-posedness under global assumptions on the (null)bicharacteristic flow, for λ below the Breitenlohner–Freedman bound, $(n-1)^2/4$. These have been known before under additional assumptions. Further, we describe the propagation of singularities of solutions and obtain the asymptotic behavior (at ∂X) of regular solutions. We also define the scattering operator, which in this case is an analogue of the hyperbolic Dirichlet-to-Neumann map. Thus, it is shown that below the Breitenlohner–Freedman bound, the Klein–Gordon equation behaves much like it would for the conformally related metric, $\hat{g}$, with Dirichlet boundary conditions, for which propagation of singularities was shown by Melrose, Sjöstrand and Taylor, though the precise form of the asymptotics is different.

1. Introduction

In this paper we consider asymptotically anti de Sitter (AdS) type metrics on n-dimensional manifolds with boundary X for $n \geq 2$. We recall the actual definition of AdS space below, but for our purposes the most important feature is the asymptotic form of the metric on these spaces, so we start by making a bold general definition. Thus, an asymptotically AdS type space is a manifold with boundary X such that X° is equipped with a pseudo-Riemannian metric g of signature $(1, n-1)$ that near the boundary Y of X is of the form

$$g = \frac{-dx^2 + h}{x^2}, \tag{1-1}$$

where h is a smooth symmetric 2-cotensor on X such that $X = Y \times [0, \epsilon)_x$ with respect to some product decomposition of X near Y, and $h|_Y$ is a section of $T^*Y \otimes T^*Y$ (rather than merely[1] $T^*_Y X \otimes T^*_Y X$) and is a

This work is partially supported by the National Science Foundation under grant DMS-0801226, and a Chambers Fellowship from Stanford University.
MSC2000: 35L05, 58J45.
Keywords: asymptotics, wave equation, anti de Sitter space, propagation of singularities.

[1]In fact, even this most general setting would necessitate only minor changes, except that the "smooth asymptotics" of Proposition 8.10 would have variable order, and the restrictions on λ that arise here, $\lambda < (n-1)^2/4$, would have to be modified.

Lorentzian metric on Y (with signature $(1, n-2)$). Note that Y is timelike with respect to the conformal metric

$$\hat{g} = x^2 g, \quad \text{so } \hat{g} = -dx^2 + h \text{ near } Y,$$

that is, the dual metric $\hat{G}$ of $\hat{g}$ is negative definite on N^*Y, that is, on $\operatorname{span}\{dx\}$, in contrast with the asymptotically de Sitter-like setting studied in [Vasy 2010b] when the boundary is spacelike. Moreover, Y is *not* assumed to be compact; indeed, under the assumption (TF) below, which is useful for global well-posedness of the wave equation, it never is. Let the wave operator $\Box = \Box_g$ be the Laplace–Beltrami operator associated to this metric, and let

$$P = P(\lambda) = \Box_g + \lambda$$

be the Klein–Gordon operator, where $\lambda \in \mathbb{C}$. The convention with the positive sign for the "spectral parameter" λ preserves the sign of λ relative to the dx^2 component of the metric in both the Riemannian conformally compact and the Lorentzian de Sitter-like cases, and hence is convenient when describing the asymptotics. We remark that if $n = 2$ then up to a change of the (overall) sign of the metric, these spaces are asymptotically de Sitter, and hence the results of [Vasy 2010b] apply. However, some of the results are different even then, since in the two settings the role of the time variable is reversed, so the formulation of the results differs as the role of "initial" and "boundary" conditions changes.

These asymptotically AdS metrics are also analogues of the Riemannian 'conformally compact', or asymptotically hyperbolic, metrics, introduced by Mazzeo and Melrose [1987] in this form, which are of the form $x^{-2}(dx^2 + h)$ with $dx^2 + h$ smooth Riemannian on X, and $h|_Y$ a section of $T^*Y \otimes T^*Y$. These have been studied extensively, in part due to the connection to AdS metrics (so some phenomena might be expected to be similar for AdS and asymptotically hyperbolic metrics) and their Riemannian signature, which makes the analysis of related PDE easier. We point out that hyperbolic space actually solves the Riemannian version of Einstein's equations, while de Sitter and anti de Sitter space satisfy the actual hyperbolic Einstein equations. We refer to [Fefferman and Graham 1985; Graham and Lee 1991; Anderson 2008] among others for analysis on conformally compact spaces. We also refer to [Witten 1998; Graham and Witten 1999; Graham and Zworski 2003] and references therein for results in the Riemannian setting that are of physical relevance. There is also a large body of literature on asymptotically de Sitter spaces. Among others, Anderson and Chruściel studied the geometry of asymptotically de Sitter spaces [Anderson 2004; 2005; Anderson and Chruściel 2005], while in [Vasy 2010b] the asymptotics of solutions of the Klein–Gordon equation were obtained, and in [Baskin 2010] the forward fundamental solution was constructed as a Fourier integral operator. It should be pointed out that the de Sitter–Schwarzschild metric in fact has many similar features with asymptotically de Sitter spaces (in an appropriate sense, it simply has two de Sitter-like ends). A weaker version of the asymptotics in this case is contained in the works of Dafermos and Rodnianski [2005; 2009; 2007] (they also study a nonlinear problem), and local energy decay was studied by Bony and Häfner [2008], in part based on the stationary resonance analysis of Sá Barreto and Zworski [1997]; stronger asymptotics (exponential decay to constants) was shown in a series of papers with Antônio Sá Barreto, Richard Melrose and the author [Melrose et al. 2011; 2008].

For the universal cover of AdS space itself, the Klein–Gordon equation was studied by Breitenlohner and Freedman [1982a; 1982b], who showed its solvability for $\lambda < (n-1)^2/4$, $n = 4$, and uniqueness for $\lambda < 5/4$, in our normalization. Analogues of these results were extended to the Dirac equation by Bachelot [2008]; and on exact AdS space there is an explicit solution due to Yagdjian and Galstian [2009]. Finally, for a class of perturbations of the universal cover of AdS, which still possess a suitable Killing vector field, Holzegel [2010] showed well-posedness for $\lambda < (n-1)^2/4$ by imposing a boundary condition; see [Holzegel 2010, Definition 3.1]. He also obtained certain estimates on the derivatives of the solution, as well as pointwise bounds.

Below we consider solutions of $Pu = 0$, or indeed $Pu = f$ with f given. Before describing our results, first we recall a formulation of the conformal problem, namely $\hat{g} = x^2 g$, so $\hat{g}$ is Lorentzian smooth on X, and Y is timelike — at the end of the introduction we give a full summary of basic results in the "compact" and "conformally compact" Riemannian and Lorentzian settings, with spacelike as well as timelike boundaries in the latter case. Let

$$\hat{P} = \Box_{\hat{g}};$$

adding λ to the operator makes no difference in this case (unlike for P). Suppose that $\mathcal{S}$ is a spacelike hypersurface in X intersecting Y (automatically transversally). Then the Cauchy problem for the Dirichlet boundary condition,

$$\hat{P}u = f, \quad u|_Y = 0, \quad u|_{\mathcal{S}} = \psi_0, \quad Vu|_{\mathcal{S}} = \psi_1,$$

with f, ψ_0, ψ_1 given, V a vector field transversal to $\mathcal{S}$, is locally well-posed (in appropriate function spaces) near $\mathcal{S}$. Moreover, under a global condition on the generalized broken bicharacteristic (or GBB) flow and $\mathcal{S}$, which we recall below in Definition 1.1, the equation is globally well-posed.

Namely, the global geometric assumption is that

$$\begin{gathered}\text{there exists } t \in \mathscr{C}^\infty(X) \text{ such that for every GBB } \gamma, \text{ the map } t \circ \rho \circ \gamma : \mathbb{R} \to \mathbb{R} \\ \text{is either strictly increasing or strictly decreasing and has range } \mathbb{R},\end{gathered} \tag{TF}$$

where $\rho : T^*X \to X$ is the bundle projection. In the formulation above of the problem, we would assume that $\mathcal{S}$ is a level set, $t = t_0$; note that locally this is always true in view of the Lorentzian nature of the metric and the conditions on Y and $\mathcal{S}$. As is often the case in the presence of boundaries — see for example [Hörmander 1985, Theorem 24.1.1] and the subsequent remark — it is convenient to consider the special case of the Cauchy problem with vanishing initial data and f supported to one side of $\mathcal{S}$, say in $t \geq t_0$; one can phrase this as solving

$$\hat{P}u = f, \quad u|_Y = 0, \quad \operatorname{supp} u \subset \{t \geq t_0\}.$$

This forward Cauchy problem is globally well-posed for $f \in L^2_{\text{loc}}(X)$ and $u \in \dot{H}^1_{\text{loc}}(X)$, and the analogous statement also holds for the backward Cauchy problem. Here we use Hörmander's notation $\dot{H}^1(X)$ [1985, Appendix B] to avoid confusion with the "zero Sobolev spaces" $H^s_0(X)$, which we recall momentarily. In addition, (without any global assumptions) singularities of solutions, as measured by the b-wave front set, WF_{b}, relative to either $L^2_{\text{loc}}(X)$ or $\dot{H}^1_{\text{loc}}(X)$, propagate along GBB as was shown by Melrose, Sjöstrand and Taylor [Melrose and Sjöstrand 1978; 1982; Taylor 1976; Melrose and Taylor 1985]; see also [Sjöstrand 1980] in the analytic setting. Here recall that in X°, bicharacteristics are integral curves of the Hamilton vector field H_p (on $T^*X^\circ \setminus o$) of the principal symbol $\hat{p} = \sigma_2(\hat{P})$ inside the characteristic set,

$$\Sigma = \hat{p}^{-1}(\{0\}).$$

We also recall that the notion of a $\mathscr{C}^\infty$ and an analytic GBB is somewhat different due to the behavior at diffractive points, with the analytic definition being more permissive (that is, weaker). Throughout this paper we use the analytic definition, which we now recall.

First, we need the notion of the compressed characteristic set $\dot{\Sigma}$ of $\hat{P}$. This can be obtained by replacing $T^*_Y X$ in T^*X by its quotient $T^*_Y X/N^*Y$, where N^*Y is the conormal bundle of Y in X. One denotes then by $\dot{\Sigma}$ the image $\hat{\pi}(\Sigma)$ of Σ in this quotient. One can give a topology to $\dot{\Sigma}$, making a set O open if and only if $\hat{\pi}^{-1}(O)$ is open in Σ. This notion of the compressed characteristic set is rather intuitive, since working with the quotient encodes the law of reflection: Points with the same tangential but different normal momentum at Y are identified, which, when combined with the conservation of kinetic energy (that is, working on the characteristic set) gives the standard law of reflection. However, it is very useful to introduce another (equivalent) definition already at this point since it arises from structures that we also need.

The alternative point of view (which is what one needs in the proofs) is that the analysis of solutions of the wave equation takes place on the b-cotangent bundle, ${}^{\mathrm{b}}T^*X$ (‘b’ stands for boundary), introduced by Melrose. See [Melrose 1993] for a very detailed description, and [Vasy 2008c] for a concise discussion. Invariantly one can define ${}^{\mathrm{b}}T^*X$ as follows. First, let $\mathscr{V}_{\mathrm{b}}(X)$ be the set of all $\mathscr{C}^\infty$ vector fields on X tangent to the boundary. If $(x, y_1, \dots, y_{n-1})$ are local coordinates on X, with x defining Y, then elements of $\mathscr{V}_{\mathrm{b}}(X)$ have the form

$$a x \partial_x + \sum_{j=1}^{n-1} b_j \partial_{y_j}, \tag{1-2}$$

with a and b_j smooth. It follows immediately that $\mathscr{V}_{\mathrm{b}}(X)$ is the set of all smooth sections of a vector bundle ${}^{\mathrm{b}}TX$, and x, y_j, a, b_j for $j = 1, \dots, n-1$ give local coordinates in terms of (1-2). Then ${}^{\mathrm{b}}T^*X$ is defined as the dual bundle of ${}^{\mathrm{b}}TX$. Thus, points in the b-cotangent bundle, ${}^{\mathrm{b}}T^*X$, of X are of the form

$$\underline{\xi}\, \frac{dx}{x} + \sum_{j=1}^{n-1} \underline{\zeta}_j \, dy_j,$$

so $(x, y, \underline{\xi}, \underline{\zeta})$ give coordinates on ${}^{\mathrm{b}}T^*X$. There is a natural map $\pi : T^*X \to {}^{\mathrm{b}}T^*X$ induced by the corresponding map between sections,

$$\xi\, dx + \sum_{j=1}^{n-1} \zeta_j \, dy_j = (x\xi)\, \frac{dx}{x} + \sum_{j=1}^{n-1} \zeta_j \, dy_j.$$

Thus

$$\pi(x, y, \xi, \zeta) = (x, y, x\xi, \zeta), \tag{1-3}$$

that is, $\underline{\xi} = x\xi$ and $\underline{\zeta} = \zeta$. Over the interior of X we can identify $T^*_{X^\circ}X$ with ${}^{\mathrm{b}}T^*_{X^\circ}X$, but this identification π becomes singular (no longer a diffeomorphism) at Y. We denote the image of Σ under π by

$$\dot{\Sigma} = \pi(\Sigma),$$

called the compressed characteristic set. Thus, $\dot{\Sigma}$ is a subset of the vector bundle ${}^{\mathrm{b}}T^*X$, and hence is equipped with a topology that is equivalent to the one define by the quotient; see [Vasy 2008c, Section 5]. The definition of *analytic* GBB is then as follows:

Definition 1.1. *Generalized broken bicharacteristics,* or GBB, are continuous maps $\gamma : I \to \dot{\Sigma}$, where I is an interval, satisfying that for all $f \in \mathscr{C}^\infty({}^{\mathrm{b}}T^*X)$ real valued,

$$\liminf_{s \to s_0} \frac{(f \circ \gamma)(s) - (f \circ \gamma)(s_0)}{s - s_0} \geq \inf\{\mathsf{H}_p(\pi^* f)(q) : q \in \pi^{-1}(\gamma(s_0)) \cap \Sigma\}.$$

Since the map $p \mapsto \mathsf{H}_p$ is a derivation, $\mathsf{H}_{ap} = a\mathsf{H}_p$ at Σ, so bicharacteristics are merely reparametrized if p is replaced by a conformal multiple. In particular, if P is the Klein–Gordon operator $\Box_g + \lambda$ for an asymptotically AdS-metric g, the bicharacteristics over X° are, up to reparametrization, those of $\hat{g}$. We make this into our definition of GBB.

Definition 1.2. The compressed characteristic set $\dot{\Sigma}$ of P is that of $\Box_{\hat{g}}$.

Generalized broken bicharacteristics, or GBB, of P are GBB *in the analytic sense* of the smooth Lorentzian metric $\hat{g}$.

We now give a formulation for the global problem. For this purpose we need to recall one more class of differential operators in addition to $\mathcal{V}_{\mathrm{b}}(X)$ (which is the set of $\mathscr{C}^\infty$ vector fields *tangent to the boundary*). Namely, we denote the set of $\mathscr{C}^\infty$ vector fields *vanishing at the boundary* by $\mathcal{V}_0(X)$. In local coordinates (x, y), these have the form

$$a\, x\partial_x + \sum_{j=1}^{n} b_j (x\partial_{y_j}), \quad \text{with } a, b_j \in \mathscr{C}^\infty(X); \tag{1-4}$$

see (1-2). Again, $\mathcal{V}_0(X)$ is the set of all $\mathscr{C}^\infty$ sections of a vector bundle ${}^0TX$, which over X° can be naturally identified with $T_{X^\circ}X$; see [Mazzeo and Melrose 1987] for a detailed discussion of 0-geometry and analysis and [Vasy 2010b] for a summary. We then let $\mathrm{Diff}_{\mathrm{b}}(X)$ and $\mathrm{Diff}_0(X)$ be the set of differential operators generated by $\mathcal{V}_{\mathrm{b}}(X)$ and $\mathcal{V}_0(X)$, respectively, that is, they are locally finite sums of products of these vector fields with $\mathscr{C}^\infty(X)$-coefficients. In particular,

$$P = \Box_g + \lambda \in \mathrm{Diff}_0^2(X),$$

which explains the relevance of $\mathrm{Diff}_0(X)$. This can be seen easily from g being in fact a nondegenerate smooth symmetric bilinear form on ${}^0TX$; the conformal factor x^{-2} compensates for the vanishing factors of x in (1-4), so in fact this is *exactly* the same statement as $\hat{g}$ being Lorentzian on TX.

Let $H_0^k(X)$ denote the zero-Sobolev space relative to

$$L^2(X) = L^2_0(X) = L^2(X, dg) = L^2(X, x^{-n} d\hat{g});$$

so if $k \geq 0$ is an integer then

$$u \in H_0^k(X) \quad \text{if and only if} \quad Lu \in L^2(X) \text{ for all } L \in \mathrm{Diff}_0^k(X);$$

negative values of k give Sobolev spaces by dualization. For our problem, we need a space of "very nice" functions corresponding to $\mathrm{Diff}_b(X)$. We obtain this by replacing $\mathscr{C}^\infty(X)$ with the space of conormal functions to the boundary relative to a fixed space of functions, in this case $H_0^k(X)$, that is, functions $v \in H^k_{0,\mathrm{loc}}(X)$ such that $Qv \in H^k_{0,\mathrm{loc}}(X)$ for every $Q \in \mathrm{Diff}_b(X)$ (of any order). The finite order regularity version of this is $H^{k,m}_{0,b}(X)$, which is given for $m \geq 0$ integer by

$$u \in H^{k,m}_{0,b}(X) \quad \text{if and only if} \quad u \in H_0^k(X) \text{ and } Qu \in H_0^k(X) \text{ for all } Q \in \mathrm{Diff}_b^m(X),$$

while for $m < 0$ integer, $u \in H^{k,m}_{0,b}(X)$ if $u = \sum Q_j u_j$, $u_j \in H^{k,0}_{0,b}(X)$, and $Q_j \in \mathrm{Diff}_b^m(X)$. Thus, $H^{-k,-m}_{0,b}(X)$ is the dual space of $H^{k,m}_{0,b}(X)$, relative to $L^2_0(X)$. Note that in X°, there is no distinction between $\mathscr{V}_b(X)$, $\mathscr{V}_0(X)$, or indeed simply $\mathscr{V}(X)$ (smooth vector fields on X), so over compact subsets K of X°, $H^{k,m}_{0,b}(X)$ is the same as $H^{k+m}(K)$. On the other hand, at $Y = \partial X$, $H^{k,m}_{0,b}(X)$ distinguishes precisely between regularity relative to $\mathscr{V}_0(X)$ and $\mathscr{V}_b(X)$.

Although the finite speed of propagation means that the wave equation has a local character in X, and thus compactness of the slices $t = t_0$ is immaterial, it is convenient to assume that

$$\text{the map } t : X \to \mathbb{R} \text{ is proper.} \tag{PT}$$

Even as stated, the propagation of singularities results (which form the heart of the paper) do not assume this, and the assumption is made elsewhere merely to make the formulation and proof of the energy estimates and existence slightly simpler, in that one does not have to localize in spatial slices this way.

Suppose $\lambda < (n-1)^2/4$. Suppose

$$f \in H^{-1,1}_{0,b,\mathrm{loc}}(X) \quad \text{and} \quad \operatorname{supp} f \subset \{t \geq t_0\}. \tag{1-5}$$

We want to find $u \in H^1_{0,\mathrm{loc}}(X)$ such that

$$Pu = f \quad \text{and} \quad \operatorname{supp} u \subset \{t \geq t_0\}. \tag{1-6}$$

We show that this is locally well-posed near $\mathscr{S}$. Moreover, under the previous global assumption on GBB, this problem is globally well-posed:

Theorem 1.3 (see Theorem 4.16). *Assume that* (*TF*) *and* (*PT*) *hold. Suppose* $\lambda < (n-1)^2/4$. *The forward Dirichlet problem* (1-6) *has a unique global solution* $u \in H^1_{0,\mathrm{loc}}(X)$, *and for all compact* $K \subset X$ *there exists a compact* $K' \subset X$ *and a constant* $C > 0$ *such that for all* f *as in* (1-5), *the solution* u *satisfies*

$$\|u\|_{H^1_0(K)} \leq C \|f\|_{H^{-1,1}_{0,b}(K')}.$$

Remark 1.4. In fact, one can be quite explicit about K' in view of (PT), since $u|_{t\in[t_0,t_1]}$ can be estimated by $f|_{t\in I}$, with I open and containing $[t_0, t_1]$.

We also prove microlocal elliptic regularity and describe the propagation of singularities of solutions, as measured by WF_b relative to $H^1_{0,\mathrm{loc}}(X)$. We define this notion in Definition 5.9 and discuss it there in more detail. However, we recall the definition of the standard wave front set WF on manifolds without boundary X that immediately generalizes to the b-wave front set WF_b. Thus, one says that $q \in T^*X \setminus o$ is *not* in the wave front set of a distribution u if there exists $A \in \Psi^0(X)$ such $\sigma_0(A)(q)$ is invertible and $QAu \in L^2(X)$ for all $Q \in \mathrm{Diff}(X)$ — this is equivalent to $Au \in \mathscr{C}^\infty(X)$ by the Sobolev embedding theorem. Here $L^2(X)$ can be replaced by $H^m(X)$ instead, with m arbitrary. Moreover, WF^m can also be defined analogously, by requiring $Au \in L^2(X)$ for $A \in \Psi^m(X)$ elliptic at q. Thus, $q \notin \mathrm{WF}(u)$ means that u is 'microlocally $\mathscr{C}^\infty$ at q', while $q \notin \mathrm{WF}^m(u)$ means that u is 'microlocally H^m at q'.

In order to microlocalize $H^{k,m}_{0,\mathrm{b}}(X)$, we need pseudodifferential operators, here extending $\mathrm{Diff}_\mathrm{b}(X)$ (as that is how we measure regularity). These are the b-pseudodifferential operators $A \in \Psi^m_\mathrm{b}(X)$ introduced by Melrose; their principal symbol $\sigma_{\mathrm{b},m}(A)$ is a homogeneous degree m function on ${}^\mathrm{b}T^*X \setminus o$. See again [Melrose 1993; Vasy 2008c]. Then we say that $q \in {}^\mathrm{b}T^*X \setminus o$ is *not* in $\mathrm{WF}^{k,\infty}_\mathrm{b}(u)$ if there exists $A \in \Psi^0_\mathrm{b}(X)$ with $\sigma_{\mathrm{b},0}(A)(q)$ invertible and such that Au is H^k_0-conormal to the boundary. One also defines $\mathrm{WF}^{k,m}_\mathrm{b}(u)$: We say $q \notin \mathrm{WF}^m_\mathrm{b}(u)$ if there exists $A \in \Psi^m_\mathrm{b}(X)$ with $\sigma_{\mathrm{b},0}(A)(q)$ invertible and such that $Au \in H^k_{0,\mathrm{loc}}(X)$. One can also extend these definitions to $m < 0$.

With this definition we have the following theorem:

Theorem 1.5 (see Proposition 7.7 and Theorem 8.8). *Suppose that $P = \Box_g + \lambda$, where $\lambda < (n-1)^2/4$. Let $m \in \mathbb{R}$ or $m = \infty$. Suppose $u \in H^{1,k}_{0,\mathrm{b},\mathrm{loc}}(X)$ for some $k \in \mathbb{R}$. Then*

$$\mathrm{WF}^{1,m}_\mathrm{b}(u) \setminus \dot{\Sigma} \subset \mathrm{WF}^{-1,m}_\mathrm{b}(Pu).$$

Moreover,

$$(\mathrm{WF}^{1,m}_\mathrm{b}(u) \cap \dot{\Sigma}) \setminus \mathrm{WF}^{-1,m+1}_\mathrm{b}(Pu)$$

is a union of maximally extended generalized broken bicharacteristics of the conformal metric $\hat{g}$ in

$$\dot{\Sigma} \setminus \mathrm{WF}^{-1,m+1}_\mathrm{b}(Pu).$$

In particular, if $Pu = 0$, then $\mathrm{WF}^{1,\infty}_\mathrm{b}(u) \subset \dot{\Sigma}$ is a union of maximally extended generalized broken bicharacteristics of $\hat{g}$.

As a consequence of this theorem, we get a more general, and precise, well-posedness result:

Theorem 1.6 (see Theorem 8.12). *Assume that (TF) and (PT) hold. Suppose that $P = \Box_g + \lambda$, where $\lambda < (n-1)^2/4$. Let $m \in \mathbb{R}$ and suppose $m' \le m$. Suppose $f \in H^{-1,m+1}_{0,\mathrm{b},\mathrm{loc}}(X)$. Then* (1-6) *has a unique solution in $H^{1,m'}_{0,\mathrm{b},\mathrm{loc}}(X)$, which in fact lies in $H^{1,m}_{0,\mathrm{b},\mathrm{loc}}(X)$, and for all compact $K \subset X$ there exists a compact $K' \subset X$ and a constant $C > 0$ such that*

$$\|u\|_{H^{1,m}_0(K)} \le C \|f\|_{H^{-1,m+1}_{0,\mathrm{b}}(K')}.$$

While we prove this result using the relatively sophisticated technique of propagation of singularities, it could also be derived without full microlocalization, that is, without localizing the propagation of energy in phase space.

We also generalize propagation of singularities to the case $\operatorname{Im}\lambda \neq 0$ ($\operatorname{Re}\lambda$ arbitrary), in which case we prove one-sided propagation depending on the sign of $\operatorname{Im}\lambda$. Namely, if $\operatorname{Im}\lambda > 0$ respectively $\operatorname{Im}\lambda < 0$, then

$$(\mathrm{WF}_{\mathrm{b}}^{1,m}(u) \cap \dot{\Sigma}) \setminus \mathrm{WF}_{\mathrm{b}}^{-1,m+1}(Pu)$$

is a union of *maximally forward* respectively *maximally backward* extended generalized broken bicharacteristics of the conformal metric $\hat{g}$. There is no difference between the case $\operatorname{Im}\lambda = 0$ and $\operatorname{Re}\lambda < (n-1)^2/4$, respectively $\operatorname{Im}\lambda \neq 0$, at the elliptic set, that is, the statement

$$\mathrm{WF}_{\mathrm{b}}^{1,m}(u) \setminus \dot{\Sigma} \subset \mathrm{WF}_{\mathrm{b}}^{-1,m}(Pu).$$

holds even if $\operatorname{Im}\lambda \neq 0$. We refer to Proposition 7.7 and Theorem 8.9 for details.

These results indicate already that for $\operatorname{Im}\lambda \neq 0$ there are many interesting questions to answer, and in particular that one cannot think of λ as 'small'; this will be the focus of future work.

In particular, if f is conormal relative to $H_0^1(X)$ then $\mathrm{WF}_{\mathrm{b}}^{1,\infty}(u) = \varnothing$. Let $\sqrt{\cdot}$ denote the branch square root function on $\mathbb{C} \setminus (-\infty, 0]$ chosen so that takes positive values on $(0, \infty)$. The simplest conormal functions are those in $\mathscr{C}^\infty(X)$ that vanish to infinite order (that is, with all derivatives) at the boundary; the set of these is denoted by $\dot{\mathscr{C}}^\infty(X)$. If we assume $f \in \dot{\mathscr{C}}^\infty(X)$ then

$$u = x^{s_+(\lambda)} v, \quad v \in \mathscr{C}^\infty(X), \quad s_+(\lambda) = \tfrac{1}{2}(n-1) + \sqrt{\tfrac{1}{4}(n-1)^2 - \lambda},$$

as we show in Proposition 8.10. Since the indicial roots of $\Box_g + \lambda$ are

$$s_\pm(\lambda) = \tfrac{1}{2}(n-1) \pm \sqrt{\tfrac{1}{4}(n-1)^2 - \lambda}, \tag{1-7}$$

this explains the interpretation of this problem as a "Dirichlet problem", much like it was done in the Riemannian conformally compact case by Mazzeo and Melrose [1987]: Asymptotics $x^{s_-(\lambda)} v_-$, with $v_- \in \mathscr{C}^\infty(X)$, corresponding to the growing indicial root $s_-(\lambda)$ is ruled out.

For $\lambda < (n-1)^2/4$, one can then easily solve the problem with inhomogeneous "Dirichlet" boundary condition, that is, given $v_0 \in \mathscr{C}^\infty(Y)$ and $f \in \dot{\mathscr{C}}^\infty(X)$, both supported in $\{t \geq t_0\}$,

$$Pu = f, \quad u|_{t<t_0} = 0, \quad u = x^{s_-(\lambda)} v_- + x^{s_+(\lambda)} v_+, \quad v_\pm \in \mathscr{C}^\infty(X), \quad v_-|_Y = v_0$$

if $s_+(\lambda) - s_-(\lambda) = 2\sqrt{(n-1)^2/4 - \lambda}$ is not an integer. If $s_+(\lambda) - s_-(\lambda)$ is an integer, the same conclusion holds if we replace $v_- \in \mathscr{C}^\infty(X)$ by $v_- \in \mathscr{C}^\infty(X) + x^{s_+(\lambda) - s_-(\lambda)} \log x \, \mathscr{C}^\infty(X)$; see Theorem 8.11.

The operator $v_-|_Y \to v_+|_Y$ is the analogue of the Dirichlet-to-Neumann map, or the scattering operator. In the De Sitter setting the setup is somewhat different as both pieces of scattering data are specified either at past or future infinity; see [Vasy 2010b]. Nonetheless, one expects that the result of [ibid., Section 7], that the scattering operator is a Fourier integral operator associated to the GBB relation, can be extended to the present setting, at least if the boundary is totally geodesic with respect to the conformal metric $\hat{g}$

and the metric is even with respect to the boundary in an appropriate sense. Indeed, in an ongoing project, Baskin and the author are extending Baskin's construction [2010] of the forward fundamental solution on asymptotically De Sitter spaces to the even totally geodesic asymptotically AdS setting. In addition, it is interesting to ask what the "best" problem to pose is when $\operatorname{Im}\lambda \neq 0$; the results of this paper suggest that the global problem (rather than local, Cauchy data versions) is the best behaved. One virtue of the parametrix construction is that we expect to be able answer Lorentzian analogues of questions related to [Mazzeo and Melrose 1987], which would bring the Lorentzian world of AdS spaces significantly closer (in terms of results) to the Riemannian world of conformally compact spaces. We singled out the totally geodesic condition and evenness since they hold on actual AdS space, which we now discuss.

We now recall the structure of the actual AdS space to justify our terminology. Consider $\mathbb{R}^{n+1}$ with the pseudo-Riemannian metric of signature $(2, n-1)$ given by

$$-dz_1^2 - \cdots - dz_{n-1}^2 + dz_n^2 + dz_{n+1}^2,$$

with $(z_1, \ldots, z_{n+1})$ denoting coordinates on $\mathbb{R}^{n+1}$, and the hyperboloid

$$z_1^2 + \cdots + z_{n-1}^2 - z_n^2 - z_{n+1}^2 = -1$$

inside it. Note that $z_n^2 + z_{n+1}^2 \geq 1$ on the hyperboloid, so we can (diffeomorphically) introduce polar coordinates in these two variables, that is, we let $(z_n, z_{n+1}) = R\theta$, with $R \geq 1$ and $\theta \in \mathbb{S}^1$. Then the hyperboloid is of the form

$$z_1^2 + \cdots + z_{n-1}^2 - R^2 = -1$$

inside $\mathbb{R}^{n-1} \times (0, \infty)_R \times \mathbb{S}^1_\theta$. Since dz_j for $j = 1, \ldots, n-1$, $d\theta$ and $d(z_1^2 + \cdots + z_{n-1}^2 - R^2)$ are linearly independent at the hyperboloid,

$$z_1, \ldots, z_{n-1}, \theta$$

give local coordinates on it, and indeed these are global in the sense that the hyperboloid X° is identified with $\mathbb{R}^{n-1} \times \mathbb{S}^1$ via these. A straightforward calculation shows that the metric on $\mathbb{R}^{n+1}$ restricts to give a Lorentzian metric g on the hyperboloid. Indeed, away from $\{0\} \times \mathbb{S}^1$, we obtain a convenient form of the metric by using polar coordinates (r, ω) in $\mathbb{R}^{n-1}$, so $R^2 = r^2 + 1$:

$$g = -(dr)^2 - r^2\, d\omega^2 + (dR)^2 + R^2\, d\theta^2 = -(1 + r^2)^{-1}\, dr^2 - r^2\, d\omega^2 + (1 + r^2)\, d\theta^2,$$

where $d\omega^2$ is the standard round metric; a similar description is easily obtained near $\{0\} \times \mathbb{S}^1$ by using the standard Euclidean variables.

We can compactify the hyperboloid by compactifying $\mathbb{R}^{n-1}$ to a ball $\overline{\mathbb{B}^{n-1}}$ via inverse polar coordinates (x, ω), where $x = r^{-1}$,

$$(z_1, \ldots, z_{n-1}) = x^{-1}\omega, \quad 0 < x < \infty, \quad \omega \in \mathbb{S}^{n-2}.$$

Thus, the interior of $\overline{\mathbb{B}^{n-1}}$ is identified with $\mathbb{R}^{n-1}$, and the boundary $\mathbb{S}^{n-2}$ of $\overline{\mathbb{B}^{n-1}}$ is added at $x = 0$ to compactify $\mathbb{R}^{n-1}$. We let

$$X = \overline{\mathbb{B}^{n-1}} \times \mathbb{S}^1$$

be this compactification of X°; a collar neighborhood of ∂X is identified with

$$[0, 1)_x \times \mathbb{S}^{n-2}_\omega \times \mathbb{S}^1_\theta.$$

In this collar neighborhood, the Lorentzian metric takes the form

$$g = \frac{1}{x^2}\big(-(1+x^2)^{-1}\,dx^2 - d\omega^2 + (1+x^2)\,d\theta^2\big),$$

which is of the desired form, and the conformal metric is

$$\hat{g} = -(1+x^2)^{-1}\,dx^2 - d\omega^2 + (1+x^2)\,d\theta^2$$

with respect to which the boundary $\{x = 0\}$ is indeed timelike. Note that the induced metric on the boundary is $-d\omega^2 + d\theta^2$ up to a conformal multiple.

As already remarked, $\hat{g}$ has the special feature that Y is totally geodesic, unlike for example the case of $\mathbb{B}^{n-1} \times \mathbb{S}^1$ equipped with a product Lorentzian metric, with $\mathbb{B}^{n-1}$ carrying the standard Euclidean metric.

For global results, it is useful to work on the universal cover $\tilde{X} = \overline{\mathbb{B}^{n-1}} \times \mathbb{R}_t$ of X, where $\mathbb{R}_t$ is the universal cover of $\mathbb{S}^1_\theta$; we use t to emphasize the timelike nature of this coordinate. The local geometry is unchanged, but now t provides a global parameter along generalized broken bicharacteristics, and satisfies the assumptions (TF) and (PT) for our theorems.

We use this opportunity to summarize the results, already referred to earlier, for analysis on conformally compact Riemannian or Lorentzian spaces, including a comparison with the conformally related problem, that is, for $\Delta_{\hat{g}}$ or $\Box_{\hat{g}}$. We assume Dirichlet boundary conditions (DBC) when relevant for the sake of definiteness, and global hyperbolicity for the hyperbolic equations, and do not state the function spaces or optimal forms of regularity results.

(i) Riemannian: $(\Delta_{\hat{g}} - \lambda)u = f$ with DBC is well-posed for $\lambda \in \mathbb{C} \setminus [0, \infty)$; moreover, if $f \in \dot{\mathscr{C}}^\infty(X)$, then $u \in \mathscr{C}^\infty(X)$. (This also works outside a discrete set of poles λ in $[0, \infty)$.)

(ii) Lorentzian, $\partial X = Y_+ \cup Y_-$ is spacelike, f is supported in $t \geq t_0$, and $\lambda \in \mathbb{C}$: $(\Box_{\hat{g}} - \lambda)u = f$, for u supported in $t \geq t_0$, is well-posed. If $f \in \dot{\mathscr{C}}^\infty(X)$, the solution is $\mathscr{C}^\infty$ up to $Y_\pm$.

(iii) Lorentzian, ∂X is timelike, f is supported in $t \geq t_0$, and $\lambda \in \mathbb{C}$: $(\Box_{\hat{g}} - \lambda)u = f$, with DBC at Y and u supported in $t \geq t_0$, is well-posed. If $f \in \dot{\mathscr{C}}^\infty(X)$, the solution is $\mathscr{C}^\infty$ up to $Y_\pm$.

We now go through the original problems. Let $s_\pm(\lambda)$ be as in (1-7).

(i) Asymptotically hyperbolic, $\lambda \in \mathbb{C} \setminus [0, +\infty)$: There is a unique solution of $(\Delta_g - \lambda)u = f$, with $f \in \dot{\mathscr{C}}^\infty(X)$, such that $u = x^{s_+(\lambda)}v$, $v \in \mathscr{C}^\infty(X)$. (Analogue of DBC [Mazzeo and Melrose 1987].) (Indeed, $u = (\Delta_g - \lambda)^{-1} f$, and this can be extended to $\lambda \in [0, +\infty)$, apart from finitely many poles in $[0, (n-1)^2/4]$, and analytically continued further.)

(ii) Asymptotically de Sitter, $\lambda \in \mathbb{C}$: For f supported in $t \geq t_0$, there is a unique solution of $(\Box_g - \lambda)u = f$ supported in $t \geq t_0$. Moreover, for $f \in \dot{\mathscr{C}}^\infty(X)$,

$$u = x^{s_+(\lambda)}v_+ + x^{s_-(\lambda)}v_-, \quad v_\pm \in \mathscr{C}^\infty(X), \quad \text{and } v_\pm|_{Y_-} \text{ is specified,}$$

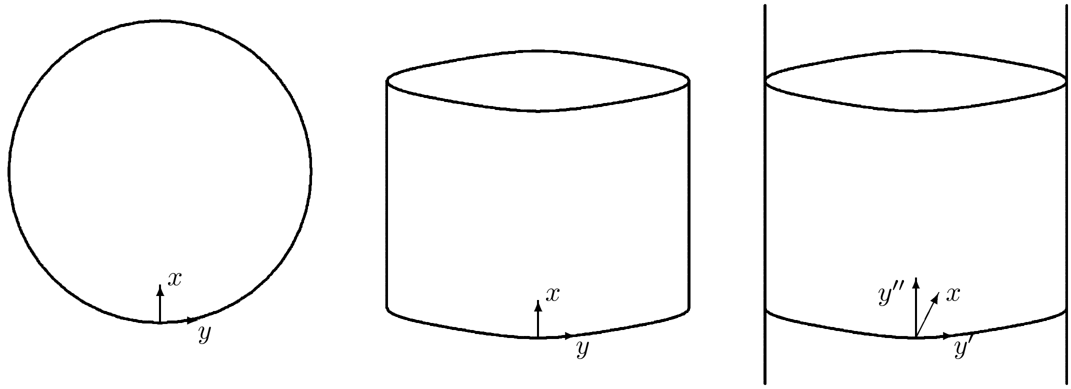

Figure 1. On the left, a Riemannian example, $\overline{\mathbb{B}^2}$. In the middle, an example of spacelike boundary, $[0, 1]_x \times \mathbb{S}^1_y$ with x timelike. On the right, the case of timelike boundary, $\overline{\mathbb{B}^2_{x,y'}} \times \mathbb{R}_{y''}$, with y'' timelike.

provided that $s_+(\lambda) - s_-(\lambda) \notin \mathbb{Z}$. (See [Vasy 2010b].)

(iii) Asymptotically anti de Sitter, $\lambda \in \mathbb{R} \setminus [(n-1)^2/4, +\infty)$: For $f \in \dot{\mathscr{C}}^\infty(X)$ supported in $t \geq t_0$, there is a unique solution of $(\Box_g - \lambda)u = f$ such that $u = x^{s_+(\lambda)} v$, $v \in \mathscr{C}^\infty(X)$ and $\operatorname{supp} u \subset \{t \geq t_0\}$.

The structure of this paper is as follows. In Section 2 we prove a Poincaré inequality that we use to allow the sharp range $\lambda < (n-1)^2/4$ for λ real. Then in Section 3 we recall the structure of energy estimates on manifolds without boundary as these are then adapted to our "zero geometry" in Section 4. In Section 5 we introduce microlocal tools to study operators such as P, namely the zero-differential-b-pseudodifferential calculus, $\mathrm{Diff}_0\, \Psi_{\mathrm{b}}(X)$. In Section 6 the structure of GBB is recalled. In Section 7 we study the Dirichlet form and prove microlocal elliptic regularity. Finally, in Section 8, we prove the propagation of singularities for P.

2. Poincaré inequality

Let h be a conformally compact Riemannian metric, that is, a positive definite inner product on ${}^0TX$ and hence by duality on ${}^0T^*X$; we denote the latter by H. We denote the corresponding space of L^2 sections of ${}^0T^*X$ by $L^2(X; {}^0T^*X) = L^2_0(X; {}^0T^*X)$. While the inner product on $L^2(X; {}^0T^*X)$ depends on the choice of h, the corresponding norms are independent of h, at least over compact subsets K of X. We first prove a Hardy-type inequality:

Lemma 2.1. *Suppose $V_0 \in \mathcal{V}(X)$ is real with $V_0 x|_{x=0} = 1$, and let $V \in \mathcal{V}_{\mathrm{b}}(X)$ be given by $V = xV_0$. Given any compact subset K of X and $\tilde{C} < (n-1)/2$, there exists $x_0 > 0$ such that if $u \in \dot{\mathscr{C}}^\infty(X)$ is supported in K, then for $\psi \in \mathscr{C}^\infty(X)$ supported in $x < x_0$,*

$$\tilde{C} \|\psi u\|_{L^2_0(X)} \leq \|\psi V u\|_{L^2_0(X)}. \tag{2-1}$$

Recall here that $\dot{\mathscr{C}}^\infty(X)$ denotes elements of $\mathscr{C}^\infty(X)$ that vanish at $Y = \partial X$ to infinite order, and the subscript comp on $\dot{\mathscr{C}}^\infty_{\text{comp}}(X)$ below indicates that in addition the support of the function under consideration is compact.

Proof. For any $V \in \mathcal{V}_{\text{b}}(X)$ real, and $\chi \in \mathscr{C}^\infty_{\text{comp}}(X)$, $u \in \dot{\mathscr{C}}^\infty_{\text{comp}}(X)$, we have, using $V^* = -V - \operatorname{div} V$,

$$\langle (V\chi)u, u\rangle = \langle [V, \chi]u, u\rangle = \langle \chi u, V^* u\rangle - \langle Vu, \chi u\rangle = -\langle \chi u, Vu\rangle - \langle Vu, \chi u\rangle - \langle \chi u, (\operatorname{div} V)u\rangle.$$

Now, if $V = xV_0$, with $V_0 \in \mathcal{V}(X)$ transversal to ∂X, and if we write $dg = x^{-n} d\hat{g}$ for $d\hat{g}$ a smooth nondegenerate density, then in local coordinates z_j such that $d\hat{g} = J|dz|$ and $V_0 = \sum V_0^j \partial_j$,

$$\begin{aligned}\operatorname{div} V &= x^n J^{-1} \sum \partial_j (x^{-n} J x V_0^j) \\ &= -(n-1) \sum_j V_0^j (\partial_j x) + x J^{-1} \sum \partial_j (J V_0^j) = -(n-1)(V_0 x) + x \operatorname{div}_{\hat{g}} V_0,\end{aligned}$$

where the subscript $\hat{g}$ in $\operatorname{div}_{\hat{g}} V_0$ denotes that the divergence is with respect to $\hat{g}$. Thus, assuming that $V_0 \in \mathcal{V}(X)$ with $V_0 x|_{x=0} = 1$, we have

$$\operatorname{div} V = -(n-1) + xa, \quad \text{where } a \in \mathscr{C}^\infty(X).$$

Let $x_0' > 0$ be such that $V_0 x > 1/2$ in $x \le x_0'$. Thus, if $0 \le \chi_0 \le 1$, $\chi_0 \equiv 1$ near 0, $\chi_0' \le 0$, χ_0 is supported in $x \le x_0'$, and $\chi = \chi_0 \circ x$, then

$$V\chi = x(V_0 x)(\chi_0' \circ x) \le 0;$$

hence $\langle (V\chi)u, u\rangle \le 0$ and

$$\langle \chi((n-1) + xa)u, u\rangle \le 2\|\chi^{1/2} u\| \|\chi^{1/2} Vu\|.$$

Thus given any $\tilde{C} < (n-1)/2$, there is $x_0 > 0$ such that for u supported in K,

$$\tilde{C} \|\chi^{1/2} u\| \le \|\chi^{1/2} Vu\|;$$

namely we take $x_0 < x_0'/2$ such that $(n-1)/2 - \tilde{C} > (\sup_K |a|) x_0$, and choose $\chi_0 \equiv 1$ on $[0, x_0]$ and supported in $[0, 2x_0)$. This completes the proof of the lemma. □

The basic Poincaré estimate is this:

Proposition 2.2. *Suppose $K \subset X$ compact, $K \cap \partial X \neq \varnothing$, O is open with $K \subset O$, O is arcwise connected to ∂X, and $K' = \overline{O}$ compact. There exists $C > 0$ such that for $u \in H^1_{0,\text{loc}}(X)$, one has*

$$\|u\|_{L^2_0(K)} \le C \|du\|_{L^2_0(O;\,{}^0T^*X)}, \tag{2-2}$$

where the norms are relative to the metric h.

Proof. It suffices to prove the estimate for $u \in \dot{\mathscr{C}}^\infty(X)$, for then the proposition follows by the density of $\dot{\mathscr{C}}^\infty(X)$ in $H^1_{0,\text{loc}}(X)$ and the continuity of both sides in the $H^1_{0,\text{loc}}(X)$ topology.

Let V_0 and V be as in Lemma 2.1, and let $\phi_0 \in \mathscr{C}^\infty_{\text{comp}}(Y)$ be identically 1 on a neighborhood of $K \cap Y$, supported in O, and let $x_0 > 0$ be as in the lemma with K replaced by K'. We pull back ϕ_0 to a function ϕ defined on a neighborhood of Y by the V_0 flow; thus, $V_0 \phi = 0$. By decreasing x_0 if needed, we may assume that ϕ is defined and is $\mathcal{C}^\infty$ in $x < x_0$, and $\operatorname{supp}\phi \cap \{x < x_0\} \subset O$. Now, let $\psi \in \mathcal{C}^\infty(X)$ be identically 1 where $x < x_0/2$, supported where $x < 3x_0/4$, and let $\psi_0 \in \mathcal{C}^\infty(X)$ be identically 1 where $x < 3x_0/4$, supported in $x < x_0$; thus $\psi_0\phi \in \mathcal{C}^\infty_{\text{comp}}(X)$. Then, by Lemma 2.1 applied to $\psi_0\phi u$,

$$\tilde{C}\|\psi\phi u\|_{L^2_0(X)} = \tilde{C}\|\psi\psi_0\phi u\|_{L^2_0(X)} \le \|\psi V(\psi_0\phi u)\|_{L^2_0(X)} = \|\psi\phi V u\|_{L^2_0(X)}. \tag{2-3}$$

The proposition follows by the standard Poincaré estimate and arcwise connectedness of K to Y (hence to $x < x_0/2$), since one can estimate $u|_{x>x_0/2}$ in L^2 in terms of $du|_{x>x_0/2}$ in L^2 and $u|_{x_0/4<x<x_0/2}$. □

We can get a more precise estimate of the constants if we restrict to a neighborhood of a spacelike hypersurface $\mathcal{S}$; it is convenient to state the result under our global assumptions. *Thus*, (*TF*) *and* (*PT*) *are assumed to hold from here on in this section.*

Proposition 2.3. *Suppose $V_0 \in \mathcal{V}(X)$ is real with $V_0 x|_{x=0} = 1$ and $V_0 t \equiv 0$ near Y and let $V \in \mathcal{V}_b(X)$ be given by $V = xV_0$. Let I be a compact interval. Let $C < (n-1)/2$ and $\gamma > 0$. Then there exist $\epsilon > 0$, $x_0 > 0$ and $C' > 0$ such that the following holds.*

For $t_0 \in I$, $0 < \delta < \epsilon$ and for $u \in H^1_{0,\text{loc}}(X)$, one has

$$\begin{aligned}\|u\|_{L^2_0(\{p:t(p)\in[t_0,t_0+\epsilon]\})} \le C^{-1}\|Vu\|_{L^2_0(\{p:t(p)\in[t_0-\delta,t_0+\epsilon],\, x(p)\le x_0\})} \\ + \gamma\|du\|_{L^2_0(\{p:t(p)\in[t_0-\delta,t_0+\epsilon]\})} + C'\|u\|_{L^2_0(\{p:t(p)\in[t_0-\delta,t_0]\})},\end{aligned} \tag{2-4}$$

where the norms are relative to the metric h.

Proof. We proceed as in the proof of Proposition 2.2, using that the t-preimage of the enlargement of the interval by distance ≤ 1 points is still compact by (PT); we always use $\epsilon < 1$ correspondingly. We simply let $\phi = \tilde{\phi} \circ t$, where $\tilde{\phi}$ is the characteristic function of $[t_0, t_0+\epsilon]$. Thus $V_0\phi$ vanishes near Y; at the cost of possibly decreasing x_0, we may assume that it vanishes in $x < x_0$. By (2-3), with $C = \tilde{C} < (n-1)/2$, if ψ is identically 1 on $[0, x_0/4)$ and is supported in $[0, x_0/2)$, then

$$\|\psi\phi u\|_{L^2_0(X)} \le C^{-1}\|\psi V\phi u\| = C^{-1}\|\psi\phi V u\|. \tag{2-5}$$

Thus, it remains to give a bound for $\|(1-\psi)u\|_{L^2_0(\{p:t(p)\in[t_0,t_0+\epsilon]\})}$.

Let $\mathcal{S}$ be the spacelike hypersurface in X given by $t = t_0$, with $t_0 \in I$. Now let $W \in \mathcal{V}_b(X)$ be transversal to $\mathcal{S}$. The standard Poincaré estimate (whose weighted version we prove below in Lemma 2.4) obtained by integrating from $t = t_0 - \delta$ yields that for $u \in \dot{\mathcal{C}}^\infty(X)$ with $u|_{t=t_0-\delta} = 0$,

$$\|u\|_{L^2_0(\{p:t(p)\in[t_0-\delta,t_0+\epsilon]\})} \le C'(\epsilon+\delta)^{1/2}\|Wu\|_{L^2_0(\{p:t(p)\in[t_0-\delta,t_0+\epsilon]\})}, \tag{2-6}$$

with $C'(\epsilon+\delta) \to 0$ as $\epsilon+\delta \to 0$. Applying this with u supported where $x \in (x_0/8, \infty)$, we have

$$\|u\|_{L^2_0(\{p:t(p)\in[t_0-\delta,t_0+\epsilon]\})} \le C''(\epsilon+\delta)^{1/2}\|xWu\|_{L^2_0(\{p:t(p)\in[t_0-\delta,t_0+\epsilon]\})}, \tag{2-7}$$

with $C''(\epsilon+\delta) \to 0$ as $\epsilon+\delta \to 0$. As we want $0 < \delta < \epsilon$, we choose $\epsilon > 0$ so that

$$C''(2\epsilon)^{1/2} < \gamma.$$

Let $\chi \in \mathscr{C}^\infty_{\mathrm{comp}}(\mathbb{R};[0,1])$ be identically 1 on $[t_0,\infty)$ and be supported in $(t_0-\delta,\infty)$. Applying (2-6) to $\chi(t)u$, we have

$$\|u\|_{L^2_0(\{p:\,t(p)\in[t_0,t_0+\epsilon]\})} \le C''(\epsilon+\delta)^{1/2}\|xWu\|_{L^2_0(\{p:\,t(p)\in[t_0-\delta,t_0+\epsilon]\})} + C''(\epsilon+\delta)^{1/2}\|x\chi'(t)(Wt)u\|_{L^2_0(\{p:\,t(p)\in[t_0-\delta,t_0]\})}.$$

In particular, this can be applied with u replaced by $(1-\psi)u$. □

We also need a weighted version of this result. We first recall a Poincaré inequality with weights.

Lemma 2.4. *Let $C_0>0$. Suppose that $W\in\mathcal{V}_{\mathrm{b}}(X)$ real, $|\operatorname{div} W|\le C_0$, $0\le\chi\in\mathscr{C}^\infty_{\mathrm{comp}}(X)$, and $\chi\le-\gamma(W\chi)$ for $t\ge t_0$, with $0<\gamma<1/(2C_0)$. Then there exists $C>0$ such that for $u\in H^1_{0,\mathrm{loc}}(X)$ with $t\ge t_0$ on* $\operatorname{supp} u$,

$$\int |W\chi||u|^2\,dg \le C\gamma\int \chi|Wu|^2\,dg.$$

Proof. We compute, using $W^*=-W-\operatorname{div} W$,

$$\langle (W\chi)u,u\rangle = \langle [W,\chi]u,u\rangle = \langle \chi u, W^*u\rangle - \langle Wu,\chi u\rangle = -\langle \chi u, Wu\rangle - \langle Wu,\chi u\rangle - \langle \chi u,(\operatorname{div} W)u\rangle,$$

so

$$\begin{aligned}\int |W\chi||u|^2\,dg = -\langle (W\chi)u,u\rangle &\le 2\|\chi^{1/2}u\|_{L^2}\|\chi^{1/2}Wu\|_{L^2} + C_0\|\chi^{1/2}u\|^2_{L^2}\\ &\le 2\Big(\int \gamma|W\chi||u|^2\,dg\Big)^{1/2}\|\chi^{1/2}Wu\|_{L^2} + C_0\int \gamma|W\chi||u|^2\,dg.\end{aligned}$$

Dividing through by $(\int|W\chi||u|^2\,dg)^{1/2}$ and rearranging yields

$$(1-C_0\gamma)\Big(\int |W\chi||u|^2\,dg\Big)^{1/2} \le 2\gamma^{1/2}\|\chi^{1/2}Wu\|_{L^2};$$

hence the claim follows. □

Our Poincaré inequality (which could also be named Hardy, in view of the relationship of (2-1) to the Hardy inequality) is then as follows:

Proposition 2.5. *Suppose $V_0\in\mathcal{V}(X)$ is real with $V_0x|_{x=0}=1$ and $V_0t\equiv 0$ near Y, and let $V\in\mathcal{V}_{\mathrm{b}}(X)$ be given by $V=xV_0$. Let I be a compact interval. Let $C<(n-1)/2$. Then there exist $\epsilon>0$, $x_0>0$, $C'>0$ and $\gamma_0>0$ such that the following holds.*

Suppose $t_0\in I$ and $0<\gamma<\gamma_0$. Let $\chi_0\in\mathscr{C}^\infty_{\mathrm{comp}}(\mathbb{R})$, $\chi=\chi_0\circ t$ and $0\le\chi_0\le-\gamma\chi_0'$ on $[t_0,t_0+\epsilon]$, with χ_0 supported in $(-\infty,t_0+\epsilon]$ and $\delta<\epsilon$. For $u\in H^1_{0,\mathrm{loc}}(X)$, one has

$$\||\chi'|^{1/2}u\|_{L^2_0(\{p:\,t(p)\in[t_0,t_0+\epsilon]\})} \le C^{-1}\||\chi'|^{1/2}Vu\|_{L^2_0(\{p:\,t(p)\in[t_0-\delta,t_0+\epsilon],\,x(p)\le x_0\})} + C'\gamma\|\chi^{1/2}du\|_{L^2_0(\{p:\,t(p)\in[t_0-\delta,t_0+\epsilon]\})} + C'\|u\|_{L^2_0(\{p:\,t(p)\in[t_0-\delta,t_0]\})}, \quad (2\text{-}8)$$

where the norms are relative to the metric h.

Proof. Let $\mathcal{S}$ be the spacelike hypersurface in X given by $t = t_0$, where $t_0 \in I$. We apply Lemma 2.4 with $W \in \mathcal{V}_b(X)$ transversal to $\mathcal{S}$ as follows.

One has from (2-5) applied with ϕ replaced by $|\chi'|^{1/2}$ that

$$\|\psi|\chi'|^{1/2}u\|_{L^2_0(X)} \leq \tilde{C}^{-1}\|\psi|\chi'|^{1/2}Vu\|.$$

We now use Lemma 2.4 with χ replaced by $\chi\rho^2$, with $\rho \equiv 1$ on $\operatorname{supp}(1-\psi)$ and $\rho \in \mathcal{C}^\infty_{\mathrm{comp}}(X^\circ)$, to estimate $\|(1-\psi)|W\chi|^{1/2}u\|_{L^2_0(X)}$. We choose ρ so that in addition $W\rho = 0$; this can be done by pulling back a function ρ_0 from $\mathcal{S}$ under the W-flow. We may also assume that ρ is supported where $x \geq x_0/8$ in view of $x \geq x_0/4$ on $\operatorname{supp}(1-\psi)$ (we might need to shorten the time interval we consider, that is, $\epsilon > 0$, to accomplish this). Thus, $W(\rho^2\chi) = \rho^2 W\chi$, and hence

$$\int \rho^2|W\chi||u|^2\,dg \leq C\gamma \int \rho^2\chi|Wu|^2\,dg.$$

Since $x \geq x_0/8$ on $\operatorname{supp}\rho$, one can estimate $\int \chi\rho^2|Wu|^2\,dg$ in terms of $\int \chi|du|^2_H\,dg$ (even though h is a Riemannian 0-metric!), giving the desired result. □

3. Energy estimates

We recall energy estimates on manifolds without boundary in a form that will be particularly convenient in the next sections. Thus, we work on X°, equipped with a Lorentz metric g and dual metric G; let $\Box = \Box_g$ be the d'Alembertian, so $\sigma_2(\Box) = G$. We consider a "twisted commutator" with a vector field $V = -\imath Z$, where Z is a real vector field, typically of the form $Z = \chi W$, with χ a cutoff function. Thus, we compute $\langle -\imath(V^*\Box - \Box V)u, u\rangle$ — the point being that the use of V^* eliminates zeroth order terms and hence is useful when we work not merely modulo lower order terms.

Note that $-\imath(V^*\Box - \Box V)$ is a second order, real, self-adjoint operator, so if its principal symbol agrees with that of d^*Cd for some real self-adjoint bundle endomorphism C, then in fact both operators are the same as the difference is zeroth order and vanishes on constants. Correspondingly, there are no zeroth order terms to estimate, which is useful as the latter tend to involve higher derivatives of χ, which in turn tend to be large relative to $d\chi$. The principal symbol in turn is easy to calculate, for the operator is

$$-\imath(V^*\Box - \Box V) = -\imath(V^* - V)\Box + \imath[\Box, V], \tag{3-1}$$

whose principal symbol is

$$-\imath\sigma_0(V^* - V)G + H_G\sigma_1(V).$$

In fact, it is easy to perform this calculation explicitly in local coordinates z_j and dual coordinates ζ_j. Let $dg = J|dz|$, so $J = |\det g|^{1/2}$. We write the components of the metric tensors as g_{ij} and G^{ij}, and $\partial_j = \partial_{z_j}$ when this does not cause confusion. We also write $Z = \chi W = \sum_j Z^j\partial_j$. *In the remainder of this section only, we adopt the standard summation convention.* Then

$$(-\imath Z)^* = \imath Z^* = -\imath J^{-1}\partial_j J Z^j \quad \text{and} \quad -\Box = J^{-1}\partial_i J G^{ij}\partial_j,$$

so

$$\begin{aligned}-\imath(V^*-V)u &= -\imath((-\imath Z)^*+\imath Z)u=(Z^*+Z)u=(-J^{-1}\partial_j J Z^j+Z^j\partial_j)u\\ &= -J^{-1}(\partial_j J Z^j)u=-(\operatorname{div} Z)u,\\ H_G &= G^{ij}\zeta_i\partial_{z_j}+G^{ij}\zeta_j\partial_{z_i}-(\partial_{z_k}G^{ij})\zeta_i\zeta_j\partial_{\zeta_k},\end{aligned}$$

(the first two terms of H_G are the same after summation, but it is convenient to keep them separate); hence

$$H_G\sigma_1(V)=G^{ij}(\partial_{z_j}Z^k)\zeta_i\zeta_k+G^{ij}(\partial_{z_i}Z^k)\zeta_j\zeta_k-Z^k(\partial_{z_k}G^{ij})\zeta_i\zeta_j.$$

Relabeling the indices, we deduce that

$$-\imath\sigma_0(V^*-V)G+H_G\sigma_1(V)=(-J^{-1}(\partial_k J Z^k)G^{ij}+G^{ik}(\partial_k Z^j)+G^{jk}(\partial_k Z^i)-Z^k\partial_k G^{ij})\zeta_i\zeta_j,$$

with the first and fourth terms combining into $-J^{-1}\partial_k(JZ^kG^{ij})\zeta_i\zeta_j$, so

$$\begin{aligned}-\imath(V^*\Box-\Box V)&=d^*Cd,\quad C_{ij}=g_{i\ell}B_{\ell j}\\ B_{ij}&=-J^{-1}\partial_k(JZ^kG^{ij})+G^{ik}(\partial_k Z^j)+G^{jk}(\partial_k Z^i),\end{aligned}\tag{3-2}$$

where C_{ij} are the matrix entries of C relative to the basis $\{dz_s\}$ of the fibers of the cotangent bundle.

We now want to expand B using $Z=\chi W$, and separate the terms with χ derivatives, with the idea being that we choose the derivative of χ large enough relative to χ to dominate the other terms. Thus,

$$\begin{aligned}B_{ij}&=G^{ik}(\partial_k Z^j)+G^{jk}(\partial_k Z^i)-J^{-1}\partial_k(JZ^kG^{ij})\\ &=(\partial_k\chi)(G^{ik}W^j+G^{jk}W^i-G^{ij}W^k)+\chi(G^{ik}(\partial_k Z^j)+G^{jk}(\partial_k Z^i)-J^{-1}\partial_k(JZ^kG^{ij}))\end{aligned}\tag{3-3}$$

and multiplying the first term on the right hand side by $\partial_i u\,\overline{\partial_j u}$ (and summing over i, j) gives

$$\begin{aligned}E_{W,d\chi}(du)&=(\partial_k\chi)(G^{ik}W^j+G^{jk}W^i-G^{ij}W^k)\partial_i u\,\overline{\partial_j u}\\ &=(du,d\chi)_G\,\overline{du(W)}+du(W)\,(d\chi,du)_G-d\chi(W)(du,du)_G,\end{aligned}\tag{3-4}$$

which is twice the sesquilinear stress-energy tensor associated to the wave u. This is well known to be positive definite in du, that is, for covectors α, $E_{W,d\chi}(\alpha)\geq 0$ and vanishing if and only if $\alpha=0$, when W and $d\chi$ are both forward timelike for smooth Lorentz metrics, see for example [Taylor 1996, Section 2.7] or [Hörmander 1985, Lemma 24.1.2]. In the present setting, the metric is degenerate at the boundary, but the analogous result still holds, as we show below.

If we replace the wave operator by the Klein–Gordon operator $P=\Box+\lambda$, $\lambda\in\mathbb{C}$, we obtain an additional term

$$-\imath\lambda(V^*-V)+2\operatorname{Im}\lambda V=-\imath\operatorname{Re}\lambda(V^*-V)+\operatorname{Im}\lambda(V+V^*)=-\imath\operatorname{Re}\lambda\operatorname{div}V+\operatorname{Im}\lambda(V+V^*)$$

in $-\imath(V^*P-P^*V)$ as compared to (3-1). With $V=-\imath Z$, $Z=\chi W$, as above, this contributes $-\operatorname{Re}\lambda(W\chi)$ in terms containing derivatives of χ to $-\imath(V^*P-P^*V)$. In particular, we have

$$\begin{aligned}\langle-\imath(V^*P-P^*V)u,u\rangle&=\int E_{W,d\chi}(du)\,dg-\operatorname{Re}\lambda\langle(W\chi)u,u\rangle\\ &\quad+\operatorname{Im}\lambda(\langle\chi Wu,u\rangle+\langle u,\chi Wu\rangle)+\langle\chi R\,du,du\rangle+\langle\chi R'u,u\rangle,\end{aligned}\tag{3-5}$$

where $R \in \mathcal{C}^\infty(X^\circ; \mathrm{End}(T^*X^\circ))$ and $R' \in \mathcal{C}^\infty(X^\circ)$.

Now suppose that W and $d\chi$ are both timelike (either forward or backward; this merely changes an overall sign). The point of (3-5) is that one controls the left side if one controls Pu (in the extreme case, when $Pu = 0$, it simply vanishes), and one can regard all terms on the right side after $E_{W,d\chi}(du)$ as terms one can control by a small multiple of the positive definite quantity $\int E_{W,d\chi}(du)\,dg$ due to the Poincaré inequality if one arranges that χ' is large relative to χ, and thus one can control $\int E_{W,d\chi}(du)\,dg$ in terms of Pu.

In fact, one does not expect that $d\chi$ will be nondegenerate timelike everywhere: Then one decomposes the energy terms into a region Ω_+ where one has the desired definiteness, and a region Ω_- where this need not hold, and then one can estimate $\int E_{W,d\tilde{\chi}}(du)\,dg$ in Ω_+ in terms of its behavior in Ω_- and Pu. Thus one propagates energy estimates (from Ω_- to Ω_+), provided one controls Pu. Of course, if u is supported in Ω_+, then one automatically controls u in Ω_-, so we are back to the setting that u is controlled by Pu. This easily gives uniqueness of solutions, and a standard functional analytic argument by duality gives solvability.

It turns out that in the asymptotically AdS case one can proceed similarly, except that the term $\mathrm{Re}\,\lambda\langle (W\chi)u, u\rangle$ is not negligible any more at ∂X, and neither is $\mathrm{Im}\,\lambda(\langle \chi W u, u\rangle + \langle u, \chi W u\rangle)$. In fact, the $\mathrm{Re}\,\lambda$ term is the "same size" as the stress energy tensor at ∂X; hence the need for an upper bound for it. Meanwhile the $\mathrm{Im}\,\lambda$ term is even larger; hence the need for the assumption $\mathrm{Im}\,\lambda = 0$ because although χ is not differentiated (hence in some sense 'small'), W is a vector field that is too large compared to the vector fields the stress energy tensor can estimate at ∂X. It is a b-vector field, rather than a 0-vector field. We explain these concepts now.

4. Zero-differential operators and b-differential operators

We start by recalling that $\mathcal{V}_b(X)$ is the Lie algebra of $\mathcal{C}^\infty$ vector fields on X tangent to ∂X, while $\mathcal{V}_0(X)$ is the Lie algebra of $\mathcal{C}^\infty$ vector fields vanishing at ∂X. Thus, $\mathcal{V}_0(X)$ is a Lie subalgebra of $\mathcal{V}_b(X)$. Note also that both $\mathcal{V}_0(X)$ and $\mathcal{V}_b(X)$ are $\mathcal{C}^\infty(X)$-modules under multiplication from the left, and they act on $x^k\mathcal{C}^\infty(X)$, in the case of $\mathcal{V}_0(X)$ in addition mapping $\mathcal{C}^\infty(X)$ into $x\mathcal{C}^\infty(X)$. The Lie subalgebra property can be strengthened as follows.

Lemma 4.1. *$\mathcal{V}_0(X)$ is an ideal in $\mathcal{V}_b(X)$.*

Proof. Suppose $V \in \mathcal{V}_0(X)$ and $W \in \mathcal{V}_b(X)$. Then, since V vanishes at ∂X, there exists $V' \in \mathcal{V}(X)$ such that $V = xV'$. Thus,

$$[V, W] = [xV', W] = [x, W]V' + x[V', W].$$

Now, $[x, W] = -Wx \in x\mathcal{C}^\infty(X)$ since W is tangent to Y, and $[V', W] \in \mathcal{V}(X)$ since $V', W \in \mathcal{V}(X)$; so $[V, W] \in x\mathcal{V}(X) = \mathcal{V}_0(X)$. □

As usual, $\mathrm{Diff}_0(X)$ is the algebra generated by $\mathcal{V}_0(X)$, while $\mathrm{Diff}_b(X)$ is the algebra generated by $\mathcal{V}_b(X)$. We combine these in the following definition, originally introduced in [Vasy 2010b] (indeed, even weights x^r were allowed there).

Definition 4.2. Let $\mathrm{Diff}_0^k \mathrm{Diff}_b^m(X)$ be the (complex) vector space of operators on $\dot{\mathcal{C}}^\infty(X)$ of the form

$$\sum P_j Q_j, \quad P_j \in \mathrm{Diff}_0^k(X), \quad Q_j \in \mathrm{Diff}_b^k(X),$$

where the sum is locally finite, and let

$$\mathrm{Diff}_0 \mathrm{Diff}_b(X) = \bigcup_{k=0}^{\infty} \bigcup_{m=0}^{\infty} \mathrm{Diff}_0^k \mathrm{Diff}_b^m(X).$$

We recall that this space is closed under composition, and that commutators have one lower order in the 0-sense than products [Vasy 2010b, Lemma 4.5]:

Lemma 4.3. $\mathrm{Diff}_0 \mathrm{Diff}_b(X)$ *is a filtered ring under composition with*

$$AB \in \mathrm{Diff}_0^{k+k'} \mathrm{Diff}_b^{m+m'}(X) \quad \textit{if } A \in \mathrm{Diff}_0^k \mathrm{Diff}_b^m(X) \textit{ and } B \in \mathrm{Diff}_0^{k'} \mathrm{Diff}_b^{m'}(X)$$

Composition is commutative to leading order in Diff_0, *that is, for A and B as above, with* $k+k' \geq 1$,

$$[A, B] \in \mathrm{Diff}_0^{k+k'-1} \mathrm{Diff}_b^{m+m'}(X).$$

Here we need an improved property regarding commutators with $\mathrm{Diff}_b(X)$ (which would a priori only gain in the 0-sense by the preceding lemma). It is this lemma that necessitates the lack of weights on the $\mathrm{Diff}_b(X)$-commutant.

Lemma 4.4. *For* $A \in \mathrm{Diff}_b^s(X)$ *and* $B \in \mathrm{Diff}_0^k \mathrm{Diff}_b^m(X)$, *with* $s \geq 1$,

$$[A, B] \in \mathrm{Diff}_0^k \mathrm{Diff}_b^{s+m-1}(X).$$

Proof. Only the leading terms in terms of Diff_b order in both commutants matter for the conclusion, for otherwise the composition result Lemma 4.3 gives the desired conclusion. We again write elements of $\mathrm{Diff}_0 \mathrm{Diff}_b(X)$ as locally finite sums of products of vector fields and functions, and then, using Lemma 4.3 and expanding the commutators, we are reduced to checking that

(i) $[W, V] = -[V, W] \in \mathrm{Diff}_0^1(X)$ for $V \in \mathcal{V}_0(X)$ and $W \in \mathcal{V}_b(X)$, which follows from Lemma 4.1, and

(ii) $[W, f] = Wf \in \mathcal{C}^\infty(X) = \mathrm{Diff}_b^0(X)$ for $W \in \mathcal{V}_b(X)$ and $f \in \mathcal{C}^\infty(X)$.

In both cases thus, the commutator drops b-order by 1 as compared to the product. □

Lemma 4.5. *For each nonnegative integer l with* $l \leq m$,

$$x^l \mathrm{Diff}_0^k \mathrm{Diff}_b^m(X) \subset \mathrm{Diff}_0^{k+l} \mathrm{Diff}_b^{m-l}(X).$$

Proof. This result is an immediate consequence of $x\mathcal{V}_b(X) \subset x\mathcal{V}(X) = \mathcal{V}_0(X)$. □

Integer ordered Sobolev spaces, $H_{0,b}^{k,m}(X)$ were defined in the introduction. It is immediate from our definitions that for $P \in \mathrm{Diff}_0^r \mathrm{Diff}_b^s(X)$,

$$P : H_{0,b}^{k,m}(X) \to H_{0,b}^{k-r,s-m}(X)$$

is continuous.

A particular consequence of Lemma 4.4 is that if $V \in \mathcal{V}_b(X)$, $P \in \mathrm{Diff}_0^m(X)$, then $[P, V] \in \mathrm{Diff}_0^m(X)$.

We also note that for $Q \in \mathcal{V}_b(X)$, with $Q = -\imath Z$ and Z real, we have $Q^* - Q \in \mathscr{C}^\infty(X)$, where the adjoint is taken with respect to the $L^2 = L_0^2(X)$ inner product. Namely:

Lemma 4.6. *Suppose $Q \in \mathcal{V}_b(X)$, with $Q = -\imath Z$ and Z real. Then $Q^* - Q \in \mathscr{C}^\infty(X)$, and with*

$$Q = a_0(xD_x) + \sum a_j D_{y_j},$$

we have

$$Q^* - Q = \operatorname{div} Q = J^{-1}(D_x(xa_0J) + \sum D_{y_j}(a_jJ)).$$

with the metric density given by $J|dx\,dy|$, where $J \in x^{-n}\mathscr{C}^\infty(X)$.

Proposition 4.7. *Suppose $Q \in \mathcal{V}_b(X)$, with $Q = -\imath Z$ and Z real. Then*

$$-\imath(Q^*\Box - \Box Q) = d^*Cd, \tag{4-1}$$

where $C \in \mathscr{C}^\infty(X; \mathrm{End}({}^0T^*X))$. In the basis $\{dx/x, dy_1/x, \ldots, dy_{n-1}/x\}$, we have

$$C_{ij} = \sum_\ell g_{i\ell} \sum_k (-J^{-1}\partial_k(Ja_k\hat{G}^{\ell j}) + \hat{G}^{\ell k}(\partial_k a_j) + \hat{G}^{jk}(\partial_k a_\ell)).$$

Proof. We write

$$-\imath(Q^*\Box - \Box Q) = -\imath(Q^* - Q)\Box - \imath[Q, \Box] \in \mathrm{Diff}_0^2(X),$$

and compute the principal symbol, which we check agrees with that of d^*Cd. One way of achieving this is to do the computation over X°; by continuity if the symbols agree here, they agree on ${}^0T^*X$. But over the interior this is the standard computation leading to (3-2); in coordinates z_j, with dual coordinates ζ_j, writing $Z = \sum Z^j \partial_{z_j}$ and $G = \sum G^{ij}\partial_{z_i}\partial_{z_j}$, we find both sides have principal symbol

$$\sum_{ij} B_{ij}\zeta_i\zeta_j,\ B_{ij} = \sum_k (-J^{-1}\partial_k(JZ^kG^{ij}) + G^{ik}(\partial_k Z^j) + G^{jk}(\partial_k Z^i)).$$

Now both sides of (4-1) are elements of $\mathrm{Diff}_0^2(X)$, are formally self-adjoint, real, and have the same principal symbol. Thus, their difference is a first order, self-adjoint and real operator; it follows that its principal symbol vanishes, so in fact this difference is zeroth order. Since it annihilates constants (as both sides do), it actually vanishes. □

We particularly care about the terms in which the coefficients a_j are differentiated, with the idea being that we write $Z = \chi W$, and choose the derivative of χ large enough relative to χ to dominate the other terms. Thus, as in (3-4),

$$B_{ij} = \sum_k (\partial_k\chi)(G^{ik}W^j + G^{jk}W^i - G^{ij}W^k) + \chi\sum_k (G^{ik}(\partial_k Z^j) + G^{jk}(\partial_k Z^i) - J^{-1}\partial_k(JZ^kG^{ij})) \tag{4-2}$$

and multiplying the first term on the right hand side by $\partial_i u\,\overline{\partial_j u}$ (and summing over i, j) gives

$$\sum_{i,j,k}(\partial_k\chi)(G^{ik}W^j+G^{jk}W^i-G^{ij}W^k)\partial_i u\,\overline{\partial_j u},$$

which is twice the sesquilinear stress-energy tensor $\frac{1}{2}E_{W,d\chi}(du)$ associated to the wave u. As we mentioned before, this is positive definite when W and $d\chi$ are both forward timelike for smooth Lorentz metrics. In the present setting, the metric is degenerate at the boundary, but the analogous result still holds since

$$\begin{aligned} E_{W,d\chi}(du) &= \sum_{i,j,k}(\partial_k\chi)(\hat{G}^{ik}W^j+\hat{G}^{jk}W^i-\hat{G}^{ij}W^k)(x\partial_i u)\overline{x\partial_j u} \\ &= (x\,du,d\chi)_{\hat{G}}\overline{x\,du(W)}+x\,du(W)\,(d\chi,x\,du)_{\hat{G}}-d\chi(W)(x\,du,x\,du)_{\hat{G}}, \end{aligned} \tag{4-3}$$

so the Lorentzian nondegenerate nature of $\hat{G}$ proves the (uniform) positive definiteness in $x\,du$, considered as an element of T_q^*X, and hence in du, regarded as an element of ${}^0T_q^*X$. Indeed, we recall the quick proof here since we need to improve on this statement to get an optimal result below.

Thus, we wish to show that for $\alpha\in T_q^*X$, $W\in T_qX$, α and W forward timelike,

$$\hat{E}_{W,\alpha}(\beta)=(\beta,\alpha)_{\hat{G}}\overline{\beta(W)}+\beta(W)\,(\alpha,\beta)_{\hat{G}}-\alpha(W)(\beta,\beta)_{\hat{G}}$$

is positive definite as a quadratic form in β. Since replacing W by a positive multiple does not change the positive definiteness, we may assume, as below, that $(W,W)_{\hat{G}}=1$. Then we may choose local coordinates $(z_1,\dots,z_n)$ such that $W=\partial_{z_n}$ and $\hat{g}|_q=dz_n^2-(dz_1^2+\dots+dz_{n-1}^2)$; thus $\hat{G}|_q=\partial_{z_n}^2-(\partial_{z_1}^2+\dots+\partial_{z_{n-1}}^2)$. Then $\alpha=\sum\alpha_j\,dz_j$ being forward timelike means that $\alpha_n>0$ and $\alpha_n^2>\alpha_1^2+\dots+\alpha_{n-1}^2$. Thus,

$$\begin{aligned} \hat{E}_{W,\alpha}(\beta) &= \Big(\beta_n\alpha_n-\sum_{j=1}^{n-1}\beta_j\alpha_j\Big)\overline{\beta}_n+\beta_n\Big(\alpha_n\overline{\beta}_n-\sum_{j=1}^{n-1}\alpha_j\overline{\beta}_j\Big)-\alpha_n\Big(|\beta_n|^2-\sum_{j=1}^{n-1}|\beta_j|^2\Big) \\ &= \alpha_n\sum_{j=1}^{n}|\beta_j|^2-\beta_n\sum_{j=1}^{n-1}\alpha_j\overline{\beta}_j-\sum_{j=1}^{n-1}\beta_j\alpha_j\overline{\beta}_n \\ &\geq \alpha_n\sum_{j=1}^{n}|\beta_j|^2-2|\beta_n|\Big(\sum_{j=1}^{n-1}\alpha_j^2\Big)^{1/2}\Big(\sum_{j=1}^{n-1}|\beta_j|^2\Big)^{1/2} \\ &\geq \alpha_n\sum_{j=1}^{n}|\beta_j|^2-2|\beta_n|\alpha_n\Big(\sum_{j=1}^{n-1}|\beta_j|^2\Big)^{1/2}=\alpha_n\Big(|\beta_n|-\Big(\sum_{j=1}^{n-1}|\beta_j|^2\Big)^{1/2}\Big)^2\geq 0, \end{aligned} \tag{4-4}$$

with the last inequality strict if $|\beta_n|\neq(\sum_{j=1}^{n-1}|\beta_j|^2)^{1/2}$, and the preceding one (by the strict forward timelike character of α) strict if $\beta_n\neq 0$ and $\sum_{j=1}^{n-1}|\beta_j|^2\neq 0$. It is then immediate that at least one of these inequalities is strict unless $\beta=0$, which is the claimed positive definiteness.

We claim that we can make a stronger statement if $U \in T_qX$ and $\alpha(U) = 0$ and $(U, W)_{\hat{g}} = 0$ (thus U is necessarily spacelike, that is, $(U, U)_{\hat{g}} < 0$):

$$\hat{E}_{W,\alpha}(\beta) + c\,\frac{\alpha(W)}{(U, U)_{\hat{g}}}|\beta(U)|^2 \quad \text{for } c < 1$$

is positive definite in β. Indeed, in this case (again assuming $(W, W)_{\hat{g}} = 1$) we can choose coordinates as above so that $W = \partial_{z_n}$, and so that U is a multiple of ∂_{z_1}, namely $U = (-(U, U)_{\hat{g}})^{1/2}\partial_{z_1}$, where $\hat{g}|_q = dz_n^2 - (dz_1^2 + \cdots + dz_{n-1}^2)$. To achieve this, we complete $e_n = W$ and $e_1 = (-(U, U)_{\hat{g}})^{-1/2}U$ (which are orthogonal by assumption) to a $\hat{g}$ normalized orthogonal basis $(e_1, e_2, \ldots, e_n)$ of T_qX, and then choose coordinates so that the coordinate vector fields are given by the e_j at q. Then α forward timelike means that $\alpha_n > 0$ and $\alpha_n^2 > \alpha_1^2 + \cdots + \alpha_{n-1}^2$, and $\alpha(U) = 0$ means that $\alpha_1 = 0$. Thus, with $c < 1$,

$$\begin{aligned}
&\hat{E}_{W,\alpha}(\beta) + c\,\frac{\alpha(W)}{(U, U)_{\hat{g}}}|\beta(U)|^2 \\
&\quad = \Big(\beta_n\alpha_n - \sum_{j=2}^{n-1}\beta_j\alpha_j\Big)\overline{\beta}_n + \beta_n\Big(\alpha_n\overline{\beta}_n - \sum_{j=2}^{n-1}\alpha_j\overline{\beta}_j\Big) - \alpha_n\Big(|\beta_n|^2 - \sum_{j=1}^{n-1}|\beta_j|^2\Big) - c\alpha_n|\beta_1|^2 \\
&\quad \geq (1-c)\alpha_n|\beta_1|^2 + \Big(\Big(\beta_n\alpha_n - \sum_{j=2}^{n-1}\beta_j\alpha_j\Big)\overline{\beta}_n + \beta_n\Big(\alpha_n\overline{\beta}_n - \sum_{j=2}^{n-1}\alpha_j\overline{\beta}_j\Big) - \alpha_n\Big(|\beta_n|^2 - \sum_{j=2}^{n-1}|\beta_j|^2\Big)\Big).
\end{aligned}$$

On the right hand side the term in the large parentheses is the same kind of expression as in (4-4), with the terms with $j = 1$ dropped, and is thus positive definite in $(\beta_2, \ldots, \beta_n)$, For $c < 1$, the first term is positive definite in β_1, so the left hand side is indeed positive definite as claimed. Rewriting this in terms of G in our setting, we obtain that for $c < 1$

$$E_{W,d\chi}(du) - c(W\chi)|xUu|^2$$

is positive definite in du, considered an element of ${}^0T_q^*X$, when $q \in \partial X$, and hence is positive definite sufficiently close to ∂X.

We restate the result:

Lemma 4.8. *Suppose $q \in \partial X$, $U, W \in T_qX$, $\alpha \in T_q^*X$, $\alpha(U) = 0$ and $(U, W)_{\hat{g}} = 0$. Then*

$$E_{W,\alpha}(\beta) + c\,\frac{\alpha(W)}{(U, U)_{\hat{g}}}|\beta(xU)|^2 \quad \textit{for } c < 1$$

is positive definite in $\beta \in {}^0T_q^*X$.

At this point we modify the choice of our time function t so that we can construct U and W satisfying the requirements of the lemma.

Lemma 4.9. *Assume (TF) and (PT). Given $\delta_0 > 0$ and a compact interval I, there exists a function $\tau \in \mathscr{C}^\infty(X)$ such that $|t - \tau| < \delta_0$ for $t \in I$, $d\tau$ is timelike in the same component of the timelike cone as dt, and $\hat{G}(d\tau, dx) = 0$ at $x = 0$.*

Proof. Let $\chi \in \mathscr{C}^\infty_{\mathrm{comp}}([0,\infty))$ be identically 1 near 0, with $0 \le \chi \le 1$ and $\chi' \le 0$, and supported in $[0,1]$. For $\epsilon, \delta > 0$ to be specified, let

$$\tau = t - x\chi(x^\delta/\epsilon)\frac{\hat{G}(dt,dx)}{\hat{G}(dx,dx)}.$$

Note that $x \le \epsilon^{1/\delta}$ on the support of $\chi(x^\delta/\epsilon)$, so if $\epsilon^{1/\delta}$ is sufficiently small, then $\hat{G}(dx,dx)$ is negative and bounded away from 0, in view of (PT) and because $\hat{G}(dx,dx) < 0$ at Y.

At $x = 0$,

$$d\tau = dt - \frac{\hat{G}(dt,dx)}{\hat{G}(dx,dx)}\,dx,$$

so $\hat{G}(d\tau, dx) = 0$. As already noted, $x \le \epsilon^{1/\delta}$ on the support of $\chi(x^\delta/\epsilon)$, so for $t \in I$ with I compact, we have in view of (PT)

$$|\tau - t| \le C\epsilon^{1/\delta}, \tag{4-5}$$

with C independent of ϵ and δ. Next,

$$d\tau = dt - \alpha\gamma\,dx - \tilde{\alpha}\gamma\,dx - \beta\mu,$$

where

$$\alpha = \chi\Big(\frac{x^\delta}{\epsilon}\Big), \quad \gamma = \frac{\hat{G}(dt,dx)}{\hat{G}(dx,dx)}, \quad \tilde{\alpha} = \delta\frac{x^\delta}{\epsilon}\chi'\Big(\frac{x^\delta}{\epsilon}\Big), \quad \beta = x\chi\Big(\frac{x^\delta}{\epsilon}\Big), \quad \mu = d\Big(\frac{\hat{G}(dt,dx)}{\hat{G}(dx,dx)}\Big).$$

Now,

$$\begin{aligned}\hat{G}(dt - \alpha\gamma dx, dt - \alpha\gamma dx) &= \hat{G}(dt,dt) - 2\alpha\gamma\hat{G}(dt,dx) + \alpha^2\gamma^2\hat{G}(dx,dx)\\ &= \hat{G}(dt,dt) - (2\alpha - \alpha^2)\frac{\hat{G}(dt,dx)^2}{\hat{G}(dx,dx)},\end{aligned}$$

which is $\ge \hat{G}(dt,dt)$ if $2\alpha - \alpha^2 \ge 0$, that is, if $\alpha \in [0,2]$. But $0 \le \alpha \le 1$, so

$$\hat{G}(dt - \alpha\gamma dx, dt - \alpha\gamma dx) \ge \hat{G}(dt,dt) > 0$$

that is, $dt - \alpha\gamma dx$ is timelike. Since $dt - \rho\alpha\gamma\,dx$ is still timelike for $0 \le \rho \le 1$, $dt - \alpha\gamma dx$ is in the same component of timelike covectors as dt, that is, it is forward oriented. Next, observe that with $C' = \sup s|\chi'(s)|$,

$$|\tilde{\alpha}| \le C'\delta, \quad \text{and} \quad |\beta| \le \epsilon^{1/\delta},$$

so over compact sets, $\tilde{\alpha}\gamma\,dx + \beta\mu$ can be made arbitrarily small by first choosing $\delta > 0$ sufficiently small and then $\epsilon > 0$ sufficiently small. Thus, $\hat{G}(d\tau, d\tau)$ is forward timelike as well. Reducing $\epsilon > 0$ further if needed, (4-5) completes the proof. □

This lemma can easily be made global.

Lemma 4.10. *Assume (TF) and (PT). Given $\delta_0 > 0$ there exists a function $\tau \in \mathscr{C}^\infty(X)$ such that $|t - \tau| < \delta_0$ for $t \in \mathbb{R}$, $d\tau$ is timelike in the same component of the timelike cone as dt, and $\hat{G}(d\tau, dx) = 0$ at $x = 0$.*

In particular, τ also satisfies (TF) and (PT).

Proof. We proceed as above, but let

$$\tau = t - x\chi\Big(\frac{x^{\delta(t)}}{\epsilon(t)}\Big)\frac{\hat{G}(dt, dx)}{\hat{G}(dx, dx)}.$$

We then have two additional terms,

$$-x^{1-\delta(t)}\delta'(t)\log x\,\frac{x^{\delta(t)}}{\epsilon(t)}\chi'\Big(\frac{x^{\delta}(t)}{\epsilon(t)}\Big)\frac{\hat{G}(dt, dx)}{\hat{G}(dx, dx)}\,dt \quad \text{and} \quad x\frac{\epsilon'(t)}{\epsilon(t)}\frac{x^{\delta(t)}}{\epsilon(t)}\chi'\Big(\frac{x^{\delta(t)}}{\epsilon(t)}\Big)\frac{\hat{G}(dt, dx)}{\hat{G}(dx, dx)}\,dt$$

in $d\tau$. Note that $x \le \epsilon(t)^{1/\delta(t)}$ on the support of both terms, while $(x^{\delta(t)}/\epsilon(t))\chi'(x^{\delta(t)}/\epsilon(t))$ is uniformly bounded. Thus, if $\delta(t) < 1/3$, $|\delta'(t)| \le 1$, and $|\epsilon'(t)| \le 1$, the factor in front of dt in both terms is bounded in absolute value by $C\epsilon(t)\hat{G}(dt, dx)/\hat{G}(dx, dx)$. Now for any k there are $\delta_k, \epsilon_k > 0$, which we may assume are in $(0, 1/3)$ and are decreasing with k, such that τ so defined satisfies on $I = [-k, k]$ all the requirements if $0 < \epsilon(t) < \epsilon_k$, $0 < \delta(t) < \delta_k$ on I, $|\epsilon'(t)| \le 1$ and $|\delta'(t)| \le 1$. But now in view of the bounds on ϵ_k and δ_k it is straightforward to write down $\epsilon(t)$ and $\delta(t)$ with the desired properties, for example, by approximating the piecewise linear function that takes the value ϵ_k at $\pm(k-1)$ for $k \ge 2$, to get $\epsilon(t)$, and similarly with δ, finishing the proof. □

From the remainder of this section, we assume that (*TF*) *and* (*PT*) *hold. From now on we simply replace t by τ.* We let $W = \hat{G}(dt, \cdot)$ and $U_0 = \hat{G}(dx, \cdot)$. Thus, *at* $x = 0$,

$$dt(U_0) = \hat{G}(dx, dt) = 0 \quad \text{and} \quad (U_0, W)_{\hat{g}} = \hat{G}(dx, dt) = 0.$$

We extend $U_0|_Y$ to a vector field U such that $Ut = 0$, that is, U is tangent to the level surfaces of t. Then we have on all of X,

$$W(dt) = \hat{G}(dt, dt) > 0 \quad \text{and} \quad U(dx) = \hat{G}(dx, dx) < 0 \tag{4-6}$$

on a neighborhood of Y, with uniform upper and lower bounds (bounding away from 0) for both bounds (4-6) on compact subsets of X.

Using Lemma 4.8 and the equations just above, we thus deduce for $\chi = \tilde{\chi} \circ t$ and $c < 1$, for ρ in $\mathscr{C}^\infty(X)$ identically 1 near Y, and supported sufficiently close to Y, for $Q = -\imath Z$ and $Z = \chi W$,

$$\begin{aligned}\langle -\imath(Q^*P - P^*Q)u, u\rangle &= \int E_{W,d\chi}(du)\,dg - \operatorname{Re}\lambda\langle (W\chi)u, u\rangle \\ &\qquad + \operatorname{Im}\lambda(\langle \chi Wu, u\rangle + \langle u, \chi Wu\rangle) + \langle \chi R du, du\rangle + \langle \chi R'u, u\rangle \\ &= \langle (\chi' A + \chi R)du, du\rangle + \langle c\rho(W\chi)xUu, xUu\rangle - \operatorname{Re}\lambda\langle (W\chi)u, u\rangle \\ &\qquad + \operatorname{Im}\lambda(\langle \chi Wu, u\rangle + \langle u, \chi Wu\rangle) + \langle \chi R'u, u\rangle\end{aligned} \tag{4-7}$$

with $A, R \in \mathscr{C}^\infty(X; \operatorname{End}({}^0T^*X))$, $R' \in \mathscr{C}^\infty(X)$ and A positive definite, all independent of χ. Here ρ is used since $E_{W,d\chi}(du) - c(W\chi)|xUu|^2$ is only positive definite near Y.

Fix $t_0 < t_0 + \epsilon < t_1$. Let $\chi_0(s) = e^{-1/s}$ for $s > 0$ and $\chi_0(s) = 0$ for $s < 0$. Let χ_1 be in $\mathscr{C}^\infty(\mathbb{R})$, be identically 1 on $[1, \infty)$, and vanish on $(-\infty, 0]$. Thus, $s^2\chi_0'(s) = \chi_0(s)$ for $s \in \mathbb{R}$. Now consider

$$\tilde{\chi}(s) = \chi_0(-F^{-1}(s - t_1))\chi_1((s - t_0)/\epsilon),$$

so $\operatorname{supp}\tilde{\chi} \subset [t_0, t_1]$, and for $s \in [t_0+\epsilon, t_1]$ we have

$$\begin{aligned}\tilde{\chi}' &= -\digamma^{-1}\chi_0'(-\digamma^{-1}(s-t_1)), \quad \text{so}\\ \tilde{\chi} &= -\digamma^{-1}(s-t_1)^2\tilde{\chi}'.\end{aligned}$$

For $\digamma > 0$ sufficiently large, this is bounded by a small multiple of $\tilde{\chi}'$, namely on $[t_0+\epsilon, t_1]$

$$\tilde{\chi} = -\gamma\tilde{\chi}' \quad \text{where } \gamma = (t_1-t_0)^2\digamma^{-1}. \tag{4-8}$$

In particular, for sufficiently large $\digamma$, we have on $[t_0+\epsilon, t_1]$

$$-(\chi' A + \chi R) \geq -\chi' A/2.$$

In addition, by (2-8) and (4-8), for $\operatorname{Re}\lambda < (n-1)^2/4$ and $c' > 0$ sufficiently close to 1

$$-\langle \operatorname{Re}\lambda(W\chi)u, u\rangle \leq c'\langle \rho(-W\chi)xUu, xUu\rangle + C'\digamma^{-1}\|\chi^{1/2}du\|^2,$$

while

$$\begin{aligned}&|\langle \chi R' u, u\rangle| \leq C'\|\chi^{1/2}u\|^2 \quad \text{and}\\ &\|\chi^{1/2}u\|^2 \leq C'\digamma^{-1}\langle(-W\chi)u, u\rangle \leq C''\digamma^{-1}\langle(-W\chi)xUu, xUu\rangle + C''\digamma^{-2}\|\chi^{1/2}du\|^2.\end{aligned} \tag{4-9}$$

However, $\operatorname{Im}\lambda(\langle \chi Wu, u\rangle + \langle u, \chi Wu\rangle)$ is too large to be controlled by the stress energy tensor since W is a b-vector field, but not a 0-vector field. Thus, to control the $\operatorname{Im}\lambda$ term for $t \in [t_0+\epsilon, t_1]$, we need to assume that $\operatorname{Im}\lambda = 0$. Then, writing $Qu = Q^*u + (Q-Q^*)u$ and choosing $\digamma > 0$ sufficiently large to absorb the first term on the right hand side of (4-9), we have

$$\begin{aligned}\langle -\chi' A\,du, du\rangle/2 &\leq -\langle -\imath Pu, Qu\rangle + \langle \imath Pu, Qu\rangle + \gamma\langle(-\chi')du, du\rangle\\ &\leq 2C\|\chi^{1/2}WPu\|_{H_0^{-1}(X)}\|\chi^{1/2}u\|_{H_0^1(X)}\\ &\qquad + 2C\|(-\chi')^{1/2}Pu\|_{L_0^2(X)}\|(-\chi')^{1/2}u\|_{L_0^2(X)} + C\gamma\|(-\chi')^{1/2}du\|^2\\ &\leq 2C\delta^{-1}\big(\|WPu\|^2_{H_0^{-1}(X)} + \|Pu\|^2_{L_0^2(X)}\big)\\ &\qquad + 2C\delta\big(\|\chi^{1/2}u\|^2_{H_0^1(X)} + \|(-\chi')^{1/2}u\|^2_{L^2(X)}\big) + C\digamma^{-1}\|(-\chi')^{1/2}du\|^2.\end{aligned} \tag{4-10}$$

For sufficiently small $\delta > 0$ and sufficiently large $\digamma > 0$ we absorb all but the first parenthesized term on the right hand side into the left hand side by the positive definiteness of A and the Poincaré inequality, Proposition 2.5, to conclude that for u supported in $[t_0+\epsilon, t_1]$,

$$\|(-\chi')^{1/2}du\|_{L_0^2(X;{}^0T^*X)} \leq C\|Pu\|_{H_{0,\mathrm{b}}^{-1,1}(X)}. \tag{4-11}$$

In view of the Poincaré inequality, we have this result:

Lemma 4.11. *Suppose $\lambda < (n-1)^2/4$, $t_0 < t_0+\epsilon < t_1$ and χ is as above. For $u \in \dot{\mathscr{C}}^\infty(X)$ supported in $[t_0+\epsilon, t_1]$, one has*

$$\|(-\chi')^{1/2}u\|_{H_0^1(X)} \leq C\|Pu\|_{H_{0,\mathrm{b}}^{-1,1}(X)}. \tag{4-12}$$

Remark 4.12. If I is compact then there is $T > 0$ such that for $t_0 \in I$ we can take any $t_1 \in (t_0, t_0 + T]$, that is, the time interval over which we can make the estimate is uniform over such compact intervals I.

This lemma gives local in time uniqueness immediately; hence iterative application of the lemma, together with Remark 4.12, yields this:

Corollary 4.13. *Suppose $\lambda < (n-1)^2/4$. For $f \in H^{-1,1}_{0,\mathrm{b},\mathrm{loc}}(X)$ supported in $t > t_0$, there is at most one $u \in H^1_{0,\mathrm{loc}}(X)$ such that* $\operatorname{supp} u \subset \{p : t(p) \geq t_0\}$ *and $Pu = f$.*

Estimate (4-11) has another consequence via the standard functional analytic argument.

Lemma 4.14. *Suppose $\lambda < (n-1)^2/4$ and I is a compact interval. There is $\sigma > 0$ such that for $t_0 \in I$, and for $f \in H^{-1}_{0,\mathrm{loc}}(X)$ supported in $t > t_0$, there exists $u \in H^{1,-1}_{0,\mathrm{b},\mathrm{loc}}(X)$, such that*

$$\operatorname{supp} u \subset \{p : t(p) \geq t_0\} \quad \textit{and} \quad Pu = f \quad \textit{in } t < t_0 + \sigma.$$

Proof. For any subspace $\mathfrak{X}$ of $\mathscr{C}^{-\infty}(X)$, let $\mathfrak{X}|_{[\tau_0,\tau_1]}$ consist of elements of $\mathfrak{X}$ restricted to $t \in [\tau_0, \tau_1]$, and let $\mathfrak{X}^\bullet_{[\tau_0,\tau_1]}$ consist of elements of $\mathfrak{X}$ supported in $t \in [\tau_0, \tau_1]$. In particular, an element of $\dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[\tau_0,\tau_1]}$ vanishes to infinite order at $t = \tau_0, \tau_1$. Thus, the dot over $\mathscr{C}^\infty$ denotes the infinite order vanishing at ∂X, while the $\bullet$ denotes the infinite order vanishing at the time boundaries we artificially imposed.

We assume that f is supported in $t > t_0 + \delta_0$. We use Lemma 4.11, with the role of t_0 and t_1 reversed (backward in time propagation), and our requirement on σ is that it is small enough that the backward version of the lemma is valid with $t_1 = t_0 + 2\sigma$. (This can be done uniformly over I by Remark 4.12.) Let $T_1 = t_1 - \epsilon$ and t_1 be such that $t_0 + \sigma = T_1' < T_1 < t_1 < t_0 + 2\sigma$. Applying the estimate (4-11), using $P = P^*$, with u replaced by $\phi \in \dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]}$ with t_1 in the role of t_0 there (backward estimate), and with $\tau_0 \in [t_0, T_1)$ in the role of t_0, we obtain

$$\|(\chi')^{1/2}\phi\|_{H^1_0(X)|_{[\tau_0,T_1]}} \leq C\|P^*\phi\|_{H^{-1,1}_{0,\mathrm{b}}(X)|_{[\tau_0,T_1]}} \quad \text{for } \phi \in \dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[\tau_0,T_1]}. \tag{4-13}$$

It is also useful to rephrase this as

$$\|\phi\|_{H^1_0(X)|_{[\tau_0',T_1]}} \leq C\|P^*\phi\|_{H^{-1,1}_{0,\mathrm{b}}(X)|_{[\tau_0,T_1]}} \quad \text{for } \phi \in \dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[\tau_0,T_1]}, \tag{4-14}$$

when $\tau_0' > \tau_0$. By (4-13), $P^* : \dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]} \to \dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]}$ is injective. Define

$$(P^*)^{-1} : \operatorname{Ran}_{\dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]}} P^* \to \dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]}$$

by $(P^*)^{-1}\psi$ being the unique $\phi \in \dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]}$ such that $P^*\phi = \psi$. Now consider the conjugate linear functional on $\operatorname{Ran}_{\dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]}} P^*$ given by

$$\ell : \psi \mapsto \langle f, (P^*)^{-1}\psi\rangle. \tag{4-15}$$

In view of (4-13) and the support condition on f (namely, the support is in $t > t_0+\delta_0$) and ψ (the support is in $t \le T_1$),[2]

$$|\langle f, (P^*)^{-1}\psi\rangle| \le \|f\|_{H_0^{-1}(X)|_{[t_0+\delta_0,T_1]}} \|(P^*)^{-1}\psi\|_{H_0^1(X)|_{[t_0+\delta_0,T_1]}} \le C\|f\|_{H_0^{-1}(X)|_{[t_0+\delta_0,T_1]}} \|\psi\|_{H_{0,\mathrm{b}}^{-1,1}(X)|_{[t_0,T_1]}},$$

so ℓ is a continuous conjugate linear functional if we equip $\mathrm{Ran}_{\dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]}} P^*$ with the $H_{0,\mathrm{b}}^{-1,1}(X)|_{[t_0,T_1]}$ norm.

If we did not care about the solution vanishing in $t < t_0 + \delta_0$, we could simply use Hahn–Banach to extend this to a continuous conjugate linear functional u on $H_{0,\mathrm{b}}^{-1,1}(X)^\bullet_{[t_0,T_1]}$, which can thus be identified with an element of $H_{0,\mathrm{b}}^{1,-1}(X)|_{[t_0,T_1]}$. This would give

$$Pu(\phi) = \langle Pu, \phi\rangle = \langle u, P^*\phi\rangle = \ell(P^*\phi) = \langle f, (P^*)^{-1}P^*\phi\rangle = \langle f, \phi\rangle$$

for $\phi \in \dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]}$, so $Pu = f$.

We do want the vanishing of u in $(t_0, t_0+\delta_0)$, that is, when applied to ϕ supported in this region. As a first step in this direction, let $\delta_0' \in (0, \delta_0)$, and note that if

$$\phi \in \dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,t_0+\delta_0')} \cap \mathrm{Ran}_{\dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]}} P^*,$$

then $\ell(\phi) = 0$ directly by (4-15), namely, the right hand side vanishes by the support condition on f. Correspondingly, the conjugate linear map L is well defined on the algebraic sum

$$\dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,t_0+\delta_0')} + \mathrm{Ran}_{\dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]}} P^* \tag{4-16}$$

by

$$L(\phi+\psi) = \ell(\psi) \quad \text{for } \phi \in \dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,t_0+\delta_0')} \text{ and } \psi \in \mathrm{Ran}_{\dot{\mathscr{C}}^\infty_{\mathrm{comp}}(X)^\bullet_{[t_0,T_1]}} P^*.$$

We claim that the functional L is actually continuous when (4-16) is equipped with the $H_{0,\mathrm{b}}^{-1,1}(X)|_{[t_0,T_1]}$ norm. This follows from

$$|\langle f, (P^*)^{-1}\psi\rangle| \le C\|f\|_{H_0^{-1}(X)|_{[t_0+\delta_0,T_1]}} \|\psi\|_{H_{0,\mathrm{b}}^{-1,1}(X)|_{[t_0+\delta_0',T_1]}}$$

together with

$$\|\psi\|_{H_{0,\mathrm{b}}^{-1,1}(X)|_{[t_0+\delta_0',T_1]}} \le \|\phi+\psi\|_{H_{0,\mathrm{b}}^{-1,1}(X)|_{[t_0,T_1]}}$$

since ϕ vanishes on $[t_0+\delta_0', T_1]$. Correspondingly, by the Hahn–Banach theorem, we can extend L to a continuous conjugate linear map

$$u : H_{0,\mathrm{b}}^{-1,1}(X)^\bullet_{[t_0,T_1]} \to \mathbb{C},$$

which can thus by identified with an element of $H_{0,\mathrm{b}}^{1,-1}(X)|_{[t_0,T_1]}$. This gives

$$Pu(\phi) = \langle Pu, \phi\rangle = \langle u, P^*\phi\rangle = \ell(P^*\phi) = \langle f, (P^*)^{-1}P^*\phi\rangle = \langle f, \phi\rangle$$

[2]We use below that we can regard f as an element of $H_0^{-1}(X)^\bullet_{[t_0+\delta_0,\infty)}$ and $(P^*)^{-1}\psi$ as an element of $H_0^1(X)^\bullet_{(-\infty,T_1]}$, so these can be naturally paired, with the pairing bounded in the appropriate norms. We then write these norms as $H_0^{-1}(X)|_{[t_0+\delta_0,T_1]}$ and $H_0^1(X)|_{[t_0+\delta_0,T_1]}$.

for $\phi \in \dot{\mathscr{C}}^\infty_{\text{comp}}(X)^\bullet_{[t_0,T_1]}$ supported in (t_0, T_1), so $Pu = f$, and in addition

$$u(\phi) = 0 \quad \text{for } \phi \in \dot{\mathscr{C}}^\infty_{\text{comp}}(X)^\bullet_{[t_0,t_0+\delta'_0]},$$

so

$$t \geq t_0 + \delta'_0 \quad \text{on } \operatorname{supp} u. \tag{4-17}$$

In particular, extending u to vanish on $(-\infty, t_0 + \delta'_0)$, which is compatible with the existing definition in view of (4-17), we have a distribution solving the PDE, defined on $t < T_1$, with the desired support condition. In particular, we use a cutoff function χ that is identically 1 for $t \in (-\infty, T'_1]$ and supported on $t \in (-\infty, T_1]$, one has that $\chi u \in H^{1,-1}_{0,\text{b}}(X)$ and χu vanishes for $t < t_0 + \delta'_0$ and for $t \geq T_1$. Then $Pu = f$ on $(-\infty, T'_1)$, thus completing the proof. □

Proposition 4.15. *Suppose $\lambda < (n-1)^2/4$. For $f \in H^{-1}_{0,\text{loc}}(X)$ supported in $t > t_0$, there exists u in $H^{1,-1}_{0,\text{b},\text{loc}}(X)$ such that* $\operatorname{supp} u \subset \{p : t(p) \geq t_0\}$ *and $Pu = f$.*

Proof. We subdivide the timeline into intervals $[t_j, t_{j+1}]$, each of which is sufficiently short so that energy estimates hold even on $[t_{j-2}, t_{j+3}]$; this can be done in view of the uniform estimates on the length of such intervals over compact subsets. Using a partition of unity, we may assume that f is supported in $[t_{k-1}, t_{k+2}]$, and we need to construct a global solution of $Pu = f$ with u supported in $[t_{k-1}, \infty)$. First we obtain u_k as above solving the PDE on $(-\infty, t_{k+2}]$ (that is, $Pu_k - f$ is supported in (t_{k+2}, ∞)) and supported in $[t_{k-1}, t_{k+3}]$. Let $f_{k+1} = Pu_k - f$; this is thus supported in $[t_{k+2}, t_{k+3}]$. We next solve $Pu_{k+1} = -f_{k+1}$ on $(-\infty, t_{k+3}]$ with a result supported in $[t_{k+1}, t_{k+4}]$. Then $P(u_k + u_{k+1}) - f$ is supported in $[t_{k+3}, t_{k+4}]$, etc. Proceeding inductively and noting that the resulting sum is locally finite, we obtain the solution on all of X. □

Well-posedness of the solution will follow once we show that for solutions $u \in H^{1,s'}_{0,\text{b},\text{loc}}(X)$ of $Pu = f$, with $f \in H^{-1,s}_{0,\text{b},\text{loc}}(X)$ supported in $t > t_0$, we in fact have $u \in H^{1,s-1}_{0,\text{b},\text{loc}}(X)$; indeed, this is a consequence of the propagation of singularities. We state this as a theorem now, recalling the standing assumptions as well:

Theorem 4.16. *Assume that* (*TF*) *and* (*PT*) *hold. Suppose $\lambda < (n-1)^2/4$. For $f \in H^{-1,1}_{0,\text{b},\text{loc}}(X)$ supported in $t > t_0$, there exists a unique $u \in H^1_{0,\text{loc}}(X)$ such that* $\operatorname{supp} u \subset \{p : t(p) \geq t_0\}$ *and $Pu = f$. Moreover, for $K \subset X$ compact there is $K' \subset X$ compact, depending on K and t_0 only, such that*

$$\|u|_K\|_{H^1_0(X)} \leq \|f|_{K'}\|_{H^{-1,1}_{0,\text{b}}(X)}. \tag{4-18}$$

Remark 4.17. While we used τ of Lemma 4.10 instead of t throughout, the conclusion of this theorem is invariant under this change (since $\delta_0 > 0$ is arbitrary in Lemma 4.10), and thus is actually valid for the original t as well.

Proof. Uniqueness and (4-18) follow from Corollary 4.13 and the estimate (4-12). By Proposition 4.15, this problem has a solution $u \in H^{1,-1}_{0,\text{b},\text{loc}}(X)$ with the desired support property. By the propagation of singularities, Theorem 8.8, we know $u \in H^1_{0,\text{loc}}(X)$ since u vanishes for $t < t_0$. □

5. Zero-differential operators and b-pseudodifferential operators

To microlocalize, we need to replace $\mathrm{Diff}_b(X)$ by $\Psi_b(X)$ and $\Psi_{bc}(X)$. We refer to [Melrose 1993] for a thorough discussion and [Vasy 2008c, Section 2] for a concise introduction to these operator algebras including all the facts that are required here. In particular, the distinction between $\Psi_b(X)$ and $\Psi_{bc}(X)$ is the same as between $\Psi_{cl}(\mathbb{R}^n)$ and $\Psi(\mathbb{R}^n)$ of *classical*, or *one step polyhomogeneneous*, respectively standard, pseudodifferential operators, that is, elements of the former ($\Psi_b(X)$, respectively $\Psi_{cl}(\mathbb{R}^n)$) are (locally) quantizations of symbols with a full one-step polyhomogeneous asymptotic expansion (also called *classical* symbols), while those of the latter ($\Psi_{bc}(X)$, respectively $\Psi(\mathbb{R}^n)$) are (locally) quantizations of symbols that merely satisfy symbolic estimates. While the former are convenient since they have homogeneous principal symbols, the latter are more useful when one must use approximations (for example, by smoothing operators), as is often the case below. Before proceeding, we recall that points in the b-cotangent bundle ${}^bT^*X$ of X are of the form

$$\underline{\xi}\,\frac{dx}{x}+\sum_{j=1}^{n-1}\underline{\zeta}_j\,dy_j.$$

Thus, $(x, y, \underline{\xi}, \underline{\zeta})$ give coordinates on ${}^bT^*X$. If (x, y, ξ, ζ) are the standard coordinates on T^*X induced by local coordinates on X, that is, if one-forms are written as $\xi\,dx+\zeta\,dy$, then the map $\pi : T^*X \to {}^bT^*X$ is given by $\pi(x, y, \xi, \zeta) = (x, y, x\xi, \zeta)$.

To be a bit more concrete (but again we refer to [Melrose 1993] and [Vasy 2008c, Section 2] for more detail), we can define a large subspace (which in fact is sufficient for our purposes here) of $\Psi^m_{bc}(X)$ and $\Psi^m_b(X)$ locally by explicit *quantization maps*; these can be combined to a global quantization map by a partition of unity as usual. Thus, we have $q = q_m : S^m({}^bT^*X) \to \Psi^m_{bc}(X)$, which restrict to $q : S^m_{cl}({}^bT^*X) \to \Psi^m_b(X)$, with cl denoting classical symbols. Namely, over a local coordinate chart U with coordinates (x, y), where $y = (y_1, \dots, y_{n-1})$, and with a supported in ${}^bT^*_KX$ with $K \subset U$ compact, we may take

$$q(a)u(x,y)=(2\pi)^{-n}\int e^{\imath((x-x')\xi+(y-y')\cdot\zeta)}\phi\Big(\frac{x-x'}{x}\Big)a(x,y,x\xi,\zeta)u(x',y')\,dx'\,dy'\,d\xi\,d\zeta,$$

understood as an oscillatory integral, where $\phi \in \mathscr{C}^\infty_{comp}((-1/2, 1/2))$ is identically 1 near 0, and the integral in x' is over $[0, \infty)$. Note that ϕ is irrelevant as far as the behavior of Schwartz kernels near the diagonal is concerned (it is identically 1 there); it simply localizes to a neighborhood of the diagonal. Somewhat inaccurately, one may write $q(a)$ as $a(x, y, xD_x, D_y)$, so a is symbolic in b-vector fields; a more accurate way of reflecting this is to change variables, writing $\underline{\xi} = x\xi$ and $\underline{\zeta} = \zeta$, so

$$q(a)u(x,y)=(2\pi)^{-n}\int e^{\imath\left(\frac{x-x'}{x}\underline{\xi}+(y-y')\cdot\underline{\zeta}\right)}\phi\Big(\frac{x-x'}{x}\Big)a(x,y,\underline{\xi},\underline{\zeta})u(x',y')\,\frac{dx'}{x}\,dy'\,d\underline{\xi}\,d\underline{\zeta}. \tag{5-1}$$

With this explicit quantization, the *principal symbol* $\sigma_{b,m}(A)$ of $A = q(a)$ is the class $[a]$ of a in $S^m({}^bT^*X)/S^{m-1}({}^bT^*X)$. If a is classical, this class can be further identified with a homogeneous symbol of degree m, that is, an element of $S^m_{hom}({}^bT^*X\setminus o)$. On the other hand, the *operator wave front set* $\mathrm{WF}'_b(A)$

of $A = q(a)$ can be defined by saying that $p \in {}^{\mathrm{b}}T^*X \setminus o$ is *not* in $\mathrm{WF}'_{\mathrm{b}}(A)$ if p has a conic neighborhood Γ in ${}^{\mathrm{b}}T^*X \setminus o$ such that $a = a(x, y, \underline{\xi}, \underline{\zeta})$ is rapidly decreasing (that is, is an order $-\infty$ symbol) in Γ. Thus, A is microlocally order $-\infty$ on the complement of $\mathrm{WF}'_{\mathrm{b}}(A)$.

A somewhat better definition of $\Psi_{\mathrm{bc}}(X)$ and $\Psi_{\mathrm{b}}(X)$ is directly in terms of the Schwartz kernels. The Schwartz kernels are well behaved on the b-double space $X^2_{\mathrm{b}} = [X^2; (\partial X)^2]$ created by blowing up the corner $(\partial X)^2$ in the product space $X^2 = X \times X$; in particular they are smooth away from the diagonal and vanish to infinite order off the front face. In these terms ϕ above localizes to a neighborhood of the diagonal that only intersects the boundary of X^2_{b} in the front face of the blow-up. The equivalence of the two descriptions can be read off directly from (5-1), which shows that the Schwartz kernel is a right b-density valued (this is the factor $(dx'/x)\,dy'$ in (5-1)) distribution conormal to $(x-x')/x = 0$ and $y - y' = 0$, that is, the lift of the diagonal to X^2_{b}.

The space $\Psi_{\mathrm{bc}}(X)$ forms a filtered algebra, so $AB \in \Psi^{m+m'}_{\mathrm{bc}}(X)$ for $A \in \Psi^m_{\mathrm{bc}}(X)$ and $B \in \Psi^{m'}_{\mathrm{bc}}(X)$. In addition, the commutator satisfies $[A, B] \in \Psi^{m+m'-1}_{\mathrm{bc}}(X)$, that is, it is one order lower than the product, but there is no gain of decay at ∂X. We also recall a crucial lemma from [Vasy 2008c, Section 2]:

Lemma 5.1. *For $A \in \Psi^m_{\mathrm{bc}}(X)$ and $A \in \Psi^m_{\mathrm{b}}(X)$, one has $[xD_x, A] \in x\Psi^m_{\mathrm{bc}}(X)$ and $[xD_x, A] \in x\Psi^m_{\mathrm{b}}(X)$, respectively.*

Proof. The lemma is an immediate consequence of xD_x having a commutative normal operator; see [Melrose 1993] for a detailed discussion and [Vasy 2008c, Section 2] for a brief explanation. □

For simplicity of notation we state the results from here through Lemma 5.5 for $\Psi_{\mathrm{b}}(X)$; they work equally well if one replaces $\Psi_{\mathrm{b}}(X)$ by $\Psi_{\mathrm{bc}}(X)$ throughout.

Lemma 4.4 still holds with $\mathrm{Diff}_{\mathrm{b}}(X)$ replaced by $\Psi_{\mathrm{b}}(X)$, but without the awkward restriction on positivity of b-orders (which is simply due to the lack of nontrivial negative order differential operators).

Definition 5.2. Let $\mathrm{Diff}^k_0 \Psi^m_{\mathrm{b}}(X)$ be the (complex) vector space of operators on $\dot{\mathcal{C}}^\infty(X)$ of the form

$$\sum P_j Q_j, \quad \text{with } P_j \in \mathrm{Diff}^k_0(X) \text{ and } Q_j \in \Psi^m_{\mathrm{b}}(X),$$

where the sum is locally finite, and let

$$\mathrm{Diff}_0 \Psi_{\mathrm{b}}(X) = \bigcup_{k=0}^{\infty} \bigcup_{m\in\mathbb{R}}^{\infty} \mathrm{Diff}^k_0 \Psi^m_{\mathrm{b}}(X).$$

We define $\mathrm{Diff}^k_0 \Psi^m_{\mathrm{bc}}(X)$ similarly, by replacing $\Psi_{\mathrm{b}}(X)$ by $\Psi_{\mathrm{bc}}(X)$ throughout the definition.

The ring structure (even with a weight x^r) of $\mathrm{Diff}_0 \Psi_{\mathrm{b}}(X)$ was proved in [Vasy 2010b, Corollary 4.4 and Lemma 4.5], which we recall here. We add to the statements of these results that $\mathrm{Diff}_0 \Psi_{\mathrm{b}}(X)$ is also closed under adjoints with respect to any weighted nondegenerate b-density, and in particular with respect to a nondegenerate 0-density such as $|dg|$, since both $\mathrm{Diff}_0(X)$ and $\Psi_{\mathrm{b}}(X)$ are closed under these adjoints and $(AB)^* = B^*A^*$.

Lemma 5.3. $\mathrm{Diff}_0 \Psi_{\mathrm{b}}(X)$ *is a filtered *-ring under composition (and adjoints) with*

$$AB \in \mathrm{Diff}^{k+k'}_0 \Psi^{m+m'}_{\mathrm{b}}(X) \quad \textit{if } A \in \mathrm{Diff}^k_0 \Psi^m_{\mathrm{b}}(X) \textit{ and } B \in \mathrm{Diff}^{k'}_0 \Psi^{m'}_{\mathrm{b}}(X)$$

and

$$A^* \in \mathrm{Diff}_0^k \Psi_{\mathrm{b}}^m(X) \quad \text{if } A \in \mathrm{Diff}_0^k \Psi_{\mathrm{b}}^m(X)$$

where the adjoint is taken with respect to a (*that is*, any fixed) *nondegenerate* 0-*density. Moreover*, *composition is commutative to leading order in* Diff_0, *that is*, *for* A *and* B *as above and* $k+k' \geq 1$,

$$[A, B] \in \mathrm{Diff}_0^{k+k'-1} \Psi_{\mathrm{b}}^{m+m'}(X).$$

Just like for differential operators, we again have a lemma that improves the b-order (rather than merely the 0-order) of the commutator provided one of the commutants is in $\Psi_{\mathrm{b}}(X)$. Again, it is crucial here that there are no weights on $\Psi_{\mathrm{b}}(X)$.

Lemma 5.4. $[A, B] \in \mathrm{Diff}_0^k \Psi_{\mathrm{b}}^{s+m-1}(X)$ *if* $A \in \Psi_{\mathrm{b}}^s(X)$ *and* $B \in \mathrm{Diff}_0^k \Psi_{\mathrm{b}}^m(X)$,

Proof. Expanding elements of $\mathrm{Diff}_0^k(X)$ as finite sums of products of vector fields and functions, and using that $\Psi_{\mathrm{b}}(X)$ is commutative to leading order, we need to consider commutators $[f, A]$ for $f \in \mathscr{C}^\infty(X)$ and $A \in \Psi_{\mathrm{b}}^s(X)$ and show that this is in $\Psi_{\mathrm{b}}^{s-1}(X)$, which is automatic as $\mathscr{C}^\infty(X) \subset \Psi_{\mathrm{b}}^0(X)$. We also need to consider $[V, A]$ for $V \in \mathcal{V}_0(X)$ and $A \in \Psi_{\mathrm{b}}^s(X)$ and show that this is in $\mathrm{Diff}_0^1 \Psi_{\mathrm{b}}^{s-1}(X)$, that is,

$$[V, A] = \sum_j W_j B_j + C_j \quad \text{for some } B_j, C_j \in \Psi_{\mathrm{b}}^{s-1}(X) \text{ and } W_j \in \mathcal{V}_0(X).$$

But $V = xV'$, where $V' \in \mathcal{V}(X)$, and

$$[V', A] = \sum_j W_j' B_j' + C_j' \quad \text{for some } W_j' \in \mathcal{V}(X) \text{ and } B_j', C_j' \in \Psi_{\mathrm{b}}^{s-1}(X);$$

see [Vasy 2008c, Lemma 2.2]. Meanwhile $B'' = [x, A]x^{-1} \in \Psi_{\mathrm{b}}^{s-1}(X)$, so

$$[V, A] = [x, A]V' + x[V', A] = B''(xV') + \sum_j (xW_j')B_j' + xC_j',$$

which is of the desired form once the first term is rearranged using Lemma 5.3. That is, explicitly $B''(xV') = (xV')B'' + [B'', xV']$, with the last term being an element of $\Psi_{\mathrm{b}}^{s-1}(X)$. ☐

We also have an analogue of Lemma 4.5.

Lemma 5.5. *For any integer* $l \geq 0$,

$$x^l \,\mathrm{Diff}_0^k \Psi_{\mathrm{b}}^m(X) \subset \mathrm{Diff}_0^{k+l} \Psi_{\mathrm{b}}^{m-l}(X).$$

Proof. It suffices to show that $x\Psi_{\mathrm{b}}^m(X) \subset \mathrm{Diff}_0^1 \Psi_{\mathrm{b}}^{m-1}(X)$; the rest follows by induction. Also, we may localize and assume that A is supported in a coordinate patch; note that

$$\Psi_{\mathrm{b}}^{-\infty}(X) \subset \mathrm{Diff}_0^1 \Psi_{\mathrm{b}}^{-\infty}(X)$$

since $\mathscr{C}^\infty(X) \subset \mathrm{Diff}_0^1(X)$. Thus, let $A \in \Psi_{\mathrm{b}}^m(X)$. Then there exist $A_j \in \Psi_{\mathrm{b}}^{m-1}(X)$ for $j = 0, \dots, n-1$, and $R \in \Psi_{\mathrm{b}}^{-\infty}(X)$ such that

$$A = (xD_x)A_0 + \sum_j D_{y_j} A_j + R;$$

to achieve this, one simply needs to use the ellipticity of $L=(xD_x)^2+\sum D_{y_j}^2$ by constructing a parametrix $G\in\Psi_{\mathrm{b}}^{-2}(X)$ to it, and writing $A=LGA+EA$, with $E\in\Psi_{\mathrm{b}}^{-\infty}(X)$. As $x(xD_x), xD_{y_j}\in\mathcal{V}_0(X)$, the conclusion follows. □

As a consequence of our results thus far, we deduce that $\Psi_{\mathrm{b}}^0(X)$ is bounded on $H_0^m(X)$, as stated already in [Vasy 2010b, Lemma 4.7].

Proposition 5.6. *Suppose $m\in\mathbb{Z}$. Any $A\in\Psi_{\mathrm{bc}}^0(X)$ with compact support defines a bounded operator on $H_0^m(X)$, with operator norm bounded by a seminorm of A in $\Psi_{\mathrm{bc}}^0(X)$.*

Proof. For $m\geq 0$ this is a special case of [Vasy 2010b, Lemma 4.7]. The fact that the operator norm is bounded by a seminorm of A in $\Psi_{\mathrm{bc}}^0(X)$ was not explicitly stated there, though follows from the proof. The case $m<0$ follows by duality.

For the convenience of the reader we recall the proof in the case we actually use in this paper, namely $m=1$ (then $m=-1$ follows by duality). Any A as in the statement of the proposition is bounded on $L^2(X)$ with the stated properties. Thus, we need to show that if $V\in\mathcal{V}_0(X)$, then $VA: H_0^1(X)\to L^2(X)$. But $VA=AV+[V,A]$ and $[V,A]\in\mathrm{Diff}_0^1\Psi_{\mathrm{b}}^{-1}(X)\subset\Psi_{\mathrm{b}}^0(X)$. Hence $AV: H_0^1(X)\to L^2(X)$ and $[V,A]: L^2(X)\to L^2(X)$, with the claimed norm behavior. □

If q is a homogeneous function on ${}^{\mathrm{b}}T^*X\setminus o$, then we again consider the Hamilton vector field H_q associated to it on $T^*X^\circ\setminus o$. A calculation with change of coordinates shows that in the b-canonical coordinates given above

$$\mathsf{H}_q=(\partial_{\underline{\xi}}q)x\partial_x-(x\partial_x q)\partial_{\underline{\xi}}+(\partial\underline{\zeta}q)\partial_y-(\partial_y q)\partial\underline{\zeta},$$

so H_q extends to a $\mathscr{C}^\infty$ vector field on ${}^{\mathrm{b}}T^*X\setminus o$ that is tangent to ${}^{\mathrm{b}}T^*_{\partial X}X$. If $Q\in\Psi_{\mathrm{b}}^{m'}(X)$ and $P\in\Psi_{\mathrm{b}}^{m}(X)$, then $[Q,P]\in\Psi_{\mathrm{b}}^{m+m'-1}(X)$ has principal symbol

$$\sigma_{\mathrm{b},m+m'-1}([Q,P])=\frac{1}{\imath}\mathsf{H}_q p.$$

Using Proposition 5.6 we can define a meaningful WF_{b} relative to $H_0^1(X)$. First we recall the definition of the corresponding global function space from [Vasy 2010b, Section 4]:

For $k\geq 0$ the b-Sobolev spaces relative to $H_0^r(X)$ are given by[3]

$$H_{0,\mathrm{b,comp}}^{r,k}(X)=\{u\in H_{0,\mathrm{comp}}^r(X): Au\in H_{0,\mathrm{comp}}^r(X)\text{ for all }A\in\Psi_{\mathrm{b}}^k(X)\}.$$

These can be normed by taking any properly supported elliptic $A\in\Psi_{\mathrm{b}}^k(X)$ and letting

$$\|u\|^2_{H_{0,\mathrm{b,comp}}^{r,k}(X)}=\|u\|^2_{H_0^r(X)}+\|Au\|^2_{H_0^r(X)}.$$

Although the norm depends on the choice of A, for u supported in a fixed compact set, different choices give equivalent norms; see [Vasy 2010b, Section 4] for details in the 0-setting (where supports are not

[3] We do not need weighted spaces, unlike in [Vasy 2010b], so we only state the definition in the special case when the weight is identically 1. On the other hand, we are working on a noncompact space, so we must consider local spaces and spaces of compactly supported functions as in [Vasy 2008c, Section 3]. Note also that we reversed the index convention (which index comes first) relative to [Vasy 2010b], to match the notation for the wave front sets.

an issue), and [Vasy 2008c, Section 3] for an analysis involving supports. We also let $H^{r,k}_{0,\mathrm{b},\mathrm{loc}}(X)$ be the subspace of $H^r_{0,\mathrm{loc}}(X)$ consisting of $u \in H^r_{0,\mathrm{loc}}(X)$ such that $\phi u \in H^{r,k}_{0,\mathrm{b},\mathrm{comp}}(X)$ for any $\phi \in \mathscr{C}^\infty_{\mathrm{comp}}(X)$.

Here it is also useful to have Sobolev spaces with a negative amount of b-regularity, in a manner completely analogous to [Vasy 2008c, Definition 3.15]:

Definition 5.7. Let r be an integer, $k < 0$, and $A \in \Psi_{\mathrm{b}}^{-k}(X)$ be elliptic on ${}^{\mathrm{b}}S^*X$ with proper support. Let $H^{r,k}_{0,\mathrm{b},\mathrm{comp}}(X)$ be the space of all $u \in \mathscr{C}^{-\infty}(X)$ of the form $u = u_1 + Au_2$ with $u_1, u_2 \in H^r_{0,\mathrm{comp}}(X)$. Let

$$\|u\|_{H^{r,k}_{0,\mathrm{b},\mathrm{comp}}(X)} = \inf\{\|u_1\|_{H^r_0(X)} + \|u_2\|_{H^r_0(X)} : u = u_1 + Au_2\}.$$

We also let $H^{r,k}_{0,\mathrm{b},\mathrm{loc}}(X)$ be the space of all $u \in \mathscr{C}^{-\infty}(X)$ such that $\phi u \in H^{r,k}_{0,\mathrm{b},\mathrm{comp}}(X)$ for all $\phi \in \mathscr{C}^\infty_{\mathrm{comp}}(X)$.

As discussed for analogous spaces following [Vasy 2008c, Definition 3.15], this definition is independent of the particular A chosen, and different A give equivalent norms for distributions u supported in a fixed compact set K. Moreover:

Lemma 5.8. *Suppose $r \in \mathbb{Z}$ and $k \in \mathbb{R}$. Any $B \in \Psi^0_{\mathrm{bc}}(X)$ with compact support defines a bounded operator on $H^{r,k}_{0,\mathrm{b}}(X)$, with operator norm bounded by a seminorm of B in $\Psi^0_{\mathrm{bc}}(X)$.*

Proof. Suppose $k \geq 0$ first. Then for an $A \in \Psi^k_{\mathrm{b}}(X)$ as in the definition above,

$$\|Bu\|^2_{H^{r,k}_{0,\mathrm{b},\mathrm{comp}}(X)} = \|Bu\|^2_{H^r_0(X)} + \|ABu\|^2_{H^r_0(X)}.$$

The first term on the right side is bounded in the desired way due to Proposition 5.6. Letting $G \in \Psi_{\mathrm{b}}^{-k}(X)$ be a properly supported parametrix for A such that $GA = \mathrm{Id} + E$ for $E \in \Psi_{\mathrm{b}}^{-\infty}(X)$, we have $ABu = AB(GA - E)u = (ABG)Au - (ABE)u$, with $ABG \in \Psi^0_{\mathrm{bc}}(X)$ and $ABE \in \Psi^{-\infty}_{\mathrm{bc}}(X) \subset \Psi^0_{\mathrm{bc}}(X)$. Thus

$$\|ABu\|_{H^r_0(X)} \leq C\|Au\|_{H^r_0(X)} + C\|u\|_{H^r_0(X)}$$

by Proposition 5.6, with C bounded by a seminorm of B. This completes the proof if $k \geq 0$.

For $k < 0$, let $A \in \Psi_{\mathrm{b}}^{-k}(X)$ be as in the definition. If $u = u_1 + Au_2$, and $G \in \Psi^k_{\mathrm{b}}(X)$ is a parametrix for A such that $AG = \mathrm{Id} + F$ for $F \in \Psi_{\mathrm{b}}^{-\infty}(X)$, then

$$Bu = Bu_1 + BAu_2 = Bu_1 + (AG - F)BAu_2 = Bu_1 + A(GBA)u_2 - (FBA)u_2.$$

Now, $B, FBA, GBA \in \Psi^0_{\mathrm{b}}(X)$ so $Bu \in H^{r,k}_{0,\mathrm{b},\mathrm{comp}}(X)$. Choosing u_1 and u_2 so that

$$\|u_1\|_{H^r_0(X)} + \|u_2\|_{H^r_0(X)} \leq 2\|u\|_{H^{r,k}_{0,\mathrm{b},\mathrm{comp}}(X)}$$

shows the desired continuity, and that the operator norm of B is bounded by a $\Psi^0_{\mathrm{bc}}(X)$-seminorm. □

Now we define the wave front set relative to $H^r_{0,\mathrm{loc}}(X)$. We also allow negative a priori b-regularity relative to this space.

Definition 5.9. Suppose $u \in H^{r,k}_{0,\mathrm{loc}}(X)$, $r \in \mathbb{Z}$ and $k \in \mathbb{R}$. Then $q \in {}^{\mathrm{b}}T^*X \setminus o$ is *not* in $\mathrm{WF}^{r,\infty}_{\mathrm{b}}(u)$ if there is an $A \in \Psi^0_{\mathrm{b}}(X)$ such that $\sigma_{\mathrm{b},0}(A)(q)$ is invertible and $QAu \in H^r_{0,\mathrm{loc}}(X)$ for all $Q \in \mathrm{Diff}_{\mathrm{b}}(X)$, that is, if $Au \in H^{r,\infty}_{0,\mathrm{b},\mathrm{loc}}(X)$.

Moreover, $q \in {}^{\mathrm{b}}T^*X \setminus o$ is *not* in $\mathrm{WF}_{\mathrm{b}}^{r,m}(u)$ if there is an $A \in \Psi_{\mathrm{b}}^m(X)$ such that $\sigma_{\mathrm{b},0}(A)(q)$ is invertible and $Au \in H^r_{0,\mathrm{loc}}(X)$.

Proposition 5.6 implies that $\Psi_{\mathrm{bc}}(X)$ acts microlocally, that is, it preserves WF_{b}; see [Vasy 2008c, Section 3] for a similar argument. In particular, the proofs for both the qualitative and quantitative version of microlocality go through without any significant changes; one simply replaces the use of [Vasy 2008c, Lemma 3.2] by Proposition 5.6.

Lemma 5.10 (see [Vasy 2008c, Lemma 3.9]). *Suppose that* $u \in H^{r,k'}_{0,\mathrm{b},\mathrm{loc}}(X)$ *and* $B \in \Psi^k_{\mathrm{bc}}(X)$. *Then* $\mathrm{WF}_{\mathrm{b}}^{r,m-k}(Bu) \subset \mathrm{WF}_{\mathrm{b}}^{r,m}(u) \cap \mathrm{WF}'_{\mathrm{b}}(B)$.

As in [Vasy 2008c, Section 3], the wave front set microlocalizes the "b-singular support relative to $H^r_{0,\mathrm{loc}}(X)$", meaning this:

Lemma 5.11 (see [Vasy 2008c, Lemma 3.10]). *Suppose* $u \in H^{r,k}_{0,\mathrm{b},\mathrm{loc}}(X)$, $p \in X$. *If* ${}^{\mathrm{b}}S^*_pX \cap \mathrm{WF}_{\mathrm{b}}^{1,m}(u) = \varnothing$, *then in a neighborhood of* p, u *lies in* $H^{1,m}_{0,\mathrm{b}}(X)$, *that is, there is* $\phi \in \mathscr{C}^\infty_{\mathrm{comp}}(X)$ *with* $\phi \equiv 1$ *near* p *such that* $\phi u \in H^{1,m}_{0,\mathrm{b}}(X)$.

Corollary 5.12 (see [Vasy 2008c, Corollary 3.11]). *Suppose* $u \in H^{r,k}_{0,\mathrm{b},\mathrm{loc}}(X)$ *and* $\mathrm{WF}_{\mathrm{b}}^{r,m}(u) = \varnothing$. *Then* $u \in H^{r,m}_{0,\mathrm{b},\mathrm{loc}}(X)$.

In particular, if $u \in H^{r,k}_{0,\mathrm{b},\mathrm{loc}}(X)$ *and* $\mathrm{WF}_{\mathrm{b}}^{r,m}(u) = \varnothing$ *for all* m, *then* $u \in H^{r,\infty}_{0,\mathrm{b},\mathrm{loc}}(X)$, *that is,* u *is conormal in that* $Au \in H^r_{0,\mathrm{loc}}(X)$ *for all* $A \in \mathrm{Diff}_{\mathrm{b}}(X)$ (*or indeed* $A \in \Psi_{\mathrm{b}}(X)$).

Finally, we have the following quantitative bound for which we recall the definition of the wave front set of bounded subsets of $\Psi^k_{\mathrm{bc}}(X)$:

Definition 5.13 (see [Vasy 2008c, Definition 3.12]). Suppose that $\mathscr{B}$ is a bounded subset of $\Psi^k_{\mathrm{bc}}(X)$, and $q \in {}^{\mathrm{b}}S^*X$. We say that $q \notin \mathrm{WF}'_{\mathrm{b}}(\mathscr{B})$ if there is some $A \in \Psi_{\mathrm{b}}(X)$ that is elliptic at q such that $\{AB : B \in \mathscr{B}\}$ is a bounded subset of $\Psi_{\mathrm{b}}^{-\infty}(X)$.

Lemma 5.14 (see [Vasy 2008c, Lemmas 3.13 and 3.18]). *Suppose that* $K \subset {}^{\mathrm{b}}S^*X$ *is compact and* U *is a neighborhood of* K *in* ${}^{\mathrm{b}}S^*X$. *Let* $\tilde{K} \subset X$ *compact, and* $\tilde{U}$ *be a neighborhood of* $\tilde{K}$ *in* X *with compact closure. Let* $Q \in \Psi^k_{\mathrm{b}}(X)$ *be elliptic on* K *with* $\mathrm{WF}'_{\mathrm{b}}(Q) \subset U$, *with Schwartz kernel supported in* $\tilde{K} \times \tilde{K}$. *Let* $\mathscr{B}$ *be a bounded subset of* $\Psi^k_{\mathrm{bc}}(X)$ *with* $\mathrm{WF}'_{\mathrm{b}}(\mathscr{B}) \subset K$ *and Schwartz kernel supported in* $\tilde{K} \times \tilde{K}$. *Then for any* $s \le 0$ *there is a constant* $C > 0$ *such that for* $B \in \mathscr{B}$ *and* $u \in H^{r,s}_{0,\mathrm{b},\mathrm{loc}}(X)$ *with* $\mathrm{WF}_{\mathrm{b}}^{r,k}(u) \cap U = \varnothing$, *we have*

$$\|Bu\|_{H^r_0(X)} \le C(\|u\|_{H^{r,s}_{0,\mathrm{b}}(\tilde{U})} + \|Qu\|_{H^r_0(X)}).$$

We can use this lemma to obtain uniform bounds for pairings. We call a subset $\mathscr{B}$ of $\mathrm{Diff}_0^m \Psi^{2k}_{\mathrm{bc}}(X)$ bounded if its elements are locally finite linear combinations of a fixed, locally finite set of elements of $\mathrm{Diff}_0^m(X)$ with coefficients that lie in a bounded subset of $\Psi^{2k}_{\mathrm{bc}}(X)$.

Corollary 5.15. *Suppose that* $K \subset {}^{\mathrm{b}}S^*X$ *is compact and* U *is a neighborhood of* K *in* ${}^{\mathrm{b}}S^*X$. *Let* $\tilde{K} \subset X$ *be compact, and* $\tilde{U}$ *be a neighborhood of* $\tilde{K}$ *in* X *with compact closure. Let* $Q \in \Psi^k_{\mathrm{b}}(X)$ *be elliptic on* K *with* $\mathrm{WF}'_{\mathrm{b}}(Q) \subset U$, *with Schwartz kernel supported in* $\tilde{K} \times \tilde{K}$. *Let* $\mathscr{B}$ *be a bounded subset of* $\mathrm{Diff}_0^2 \Psi^{2k}_{\mathrm{bc}}(X)$

with $\mathrm{WF}'_{\mathrm{b}}(\mathcal{B}) \subset K$ *and Schwartz kernel supported in* $\tilde{K} \times \tilde{K}$. *Then there is a constant* $C > 0$ *such that for* $B \in \mathcal{B}$ *and* $u \in H^{1,s}_{0,\mathrm{b},\mathrm{loc}}(X)$ *with* $\mathrm{WF}^{1,k}_{\mathrm{b}}(u) \cap U = \varnothing$, *we have*

$$|\langle Bu, u\rangle| \le C(\|u\|_{H^1_0(\tilde{U})} + \|Qu\|_{H^1_0(X)})^2.$$

Proof. Using Lemma 5.3 we can write B as $\sum B'_{ij} P_i^* R_j \Lambda$, where $P_i,\ R_j \in \mathrm{Diff}^1_0(X)$, $\Lambda \in \Psi^k_{\mathrm{b}}(X)$ (which we take to be elliptic on K, but such that Q is elliptic on $\mathrm{WF}'_{\mathrm{b}}(\Lambda)$), B'_{ij} lies in a bounded subset $\mathcal{B}'$ of $\Psi^k_{\mathrm{b}}(X)$ and the sum is finite. Then

$$\begin{aligned}|\langle Bu, u\rangle| &\le \sum_{ij} |\langle R_j \Lambda u, P_i (B'_{ij})^* u\rangle| \le \sum_{ij} \|R_j \Lambda u\|_{L^2(X)} \|P_i (B'_{ij})^* u\|_{L^2(X)} \\ &\le \sum_{ij} \|\Lambda u\|_{H^1_0(X)} \|P_i (B'_{ij})^* u\|_{H^1_0(X)} \le \sum C(\|u\|_{H^{1,s}_{0,\mathrm{b}}(\tilde{U})} + \|Qu\|_{H^1_0(X)})^2,\end{aligned}$$

where in the last step we used Lemma 5.14. □

It is useful to note that infinite order b-regularity relative to $L^2_0(X)$ and $H^1_0(X)$ are the same.

Lemma 5.16. $\mathrm{WF}^{1,\infty}_{\mathrm{b}}(u) = \mathrm{WF}^{0,\infty}_{\mathrm{b}}(u)$ *for* $u \in H^1_{0,\mathrm{loc}}(X)$.

Proof. The complements of the two sides are the set of points $q \in {}^{\mathrm{b}}S^*X$ for which there exist $A \in \Psi^0_{\mathrm{b}}(X)$ (with compactly supported Schwartz kernel, as one may assume) such that $\sigma_{\mathrm{b},0}(A)(q)$ is invertible and $LAu \in H^1_0(X)$, respectively $LAu \in L^2_0(X)$. Since $H^1_0(X) \subset L^2_0(X)$, that $\mathrm{WF}^{0,\infty}_{\mathrm{b}}(u) \subset \mathrm{WF}^{1,\infty}_{\mathrm{b}}(u)$ follows immediately. For the converse, if $LAu \in L^2_0(X)$ for all $L \in \mathrm{Diff}_{\mathrm{b}}(X)$, then $\mathrm{Diff}_0(X) \subset \mathrm{Diff}_{\mathrm{b}}(X)$ shows that $QLAu \in L^2_0(X)$ for $Q \in \mathrm{Diff}^1_0(X)$ and $L \in \mathrm{Diff}_{\mathrm{b}}(X)$, so $LAu \in H^1_0(X)$, that is, $\mathrm{WF}^{1,\infty}_{\mathrm{b}}(u) \subset \mathrm{WF}^{0,\infty}_{\mathrm{b}}(u)$, completing the proof. □

We finally recall that $u \in \mathcal{A}^k(X)$, that is, that u is conormal relative to $x^k L^2_{\mathrm{b}}(X)$, which means that $Lu \in x^k L^2_{\mathrm{b}}(X)$ for all $L \in \mathrm{Diff}_{\mathrm{b}}(X)$, so in particular $u \in x^k L^2_{\mathrm{b}}(X)$. Thus,

$$\mathrm{WF}^{0,\infty}_{\mathrm{b}}(u) = \varnothing \quad \text{if and only if} \quad u \in \mathcal{A}^{(n-1)/2}(X),$$

in view of $L^2_0(X) = x^{(n-1)/2} L^2_{\mathrm{b}}(X)$.

6. Generalized broken bicharacteristics

We recall the structure of the compressed characteristic set and GBB from [Vasy 2010a, Sections 1 and 2]. In that paper X is a manifold with corners and k is the codimension of the highest codimension corner in the local coordinate chart. Thus, for application to this paper, the reader should take $k = 1$ when referring to [Vasy 2010a, Sections 1 and 2]. It is often convenient to work on the cosphere bundle, here ${}^{\mathrm{b}}S^*X$, which is equivalent to working on conic subsets of ${}^{\mathrm{b}}T^*X \setminus o$. In a region where, say,

$$|\underline{\xi}| < C|\underline{\zeta}_{n-1}| \quad \text{and} \quad |\underline{\zeta}_j| < C|\underline{\zeta}_{n-1}| \quad \text{for } j = 1, \dots, n-2, \tag{6-1}$$

with $C > 0$ fixed, we can take

$$x, y_1, \dots, y_{n-1}, \hat{\underline{\xi}}, \hat{\underline{\zeta}}_1, \dots, \hat{\underline{\zeta}}_{n-2}, |\underline{\zeta}_{n-1}| \quad \text{where } \hat{\underline{\xi}} = \underline{\xi}/|\underline{\zeta}_{n-1}| \quad \text{and} \quad \hat{\underline{\zeta}}_j = \underline{\zeta}_j/|\underline{\zeta}_{n-1}|,$$

as (projective) local coordinates on ${}^{\mathrm{b}}T^*X \setminus o$, and hence take

$$x, y_1, \ldots, y_{n-1}, \hat{\underline{\xi}}, \hat{\underline{\zeta}}_1, \ldots, \hat{\underline{\zeta}}_{n-2}$$

as local coordinates on the image of this region under the quotient map in ${}^{\mathrm{b}}S^*X$; see [Vasy 2010a, Equation (1.4)].

First, we choose local coordinates more carefully. In arbitrary local coordinates $(x, y_1, \ldots, y_{n-1})$ on a neighborhood U_0 of a point on $Y = \partial X$, so that Y is given by $x = 0$ inside $x \geq 0$, any symmetric bilinear form on T^*X can be written as

$$\hat{G}(x, y) = A(x, y)\,\partial_x\,\partial_x + \sum_j 2C_j(x, y)\,\partial_x\,\partial_{y_j} + \sum_{i,j} B_{ij}(x, y)\,\partial_{y_i}\,\partial_{y_j} \tag{6-2}$$

with A, B, C smooth. In view of (1-1), using x given there and coordinates y_j on Y pulled by to a collar neighborhood of Y by the product structure, we have in addition

$$A(0, y) = -1 \quad \text{and} \quad C_j(0, y) = 0 \quad \text{for all } y,$$

and $B(0, y) = (B_{ij}(0, y))$ is Lorentzian for all y. Below we write covectors as

$$\alpha = \xi\,dx + \sum_{i=1}^{n-1} \zeta_i\,dy_i. \tag{6-3}$$

Thus,

$$\hat{G}|_{x=0} = -\partial_x^2 + \sum_{i,j=1}^{n-1} B_{ij}(0, y)\,\partial_{y_i}\,\partial_{y_j}, \tag{6-4}$$

and hence the metric function, $p(q) = \hat{G}(q, q)$ for $q \in T^*X$, is

$$p|_{x=0} = -\xi^2 + \zeta \cdot B(y)\zeta. \tag{6-5}$$

Since $A(0, y) = -1 < 0$, we see Y is indeed timelike in that the restriction of the dual metric $\hat{G}$ to N^*Y is negative definite, for locally the conormal bundle N^*Y is given by

$$\{(x, y, \xi, \zeta) : x = 0,\ \zeta = 0\}.$$

We write $h = \zeta \cdot B(y)\zeta$ for the metric function on the boundary. Also, from (6-5),

$$\mathsf{H}_p = -2\xi \cdot \partial_x + \mathsf{H}_h + \beta\partial_\xi + xV, \tag{6-6}$$

where V is a $\mathscr{C}^\infty$ vector field in $\mathscr{U}_0 = T^*U_0$ and β is a $\mathscr{C}^\infty$ function on $\mathscr{U}_0$.

It is sometimes convenient to improve the form of B near a particular point p_0, around which the coordinate system is centered. Namely, since B is Lorentzian, we can further arrange it by adjusting the y_j coordinates so that

$$\sum B_{ij}(0, 0)\partial_{y_i}\partial_{y_j} = \partial^2_{y_{n-1}} - \sum_{i<n-1} \partial^2_{y_i}. \tag{6-7}$$

We now recall from the introduction that $\pi : T^*X \to {}^{\mathrm{b}}T^*X$ is the natural map corresponding to the identification of a section of T^*X as a section of ${}^{\mathrm{b}}T^*X$, and in local coordinates π is given by

$$\pi(x, y, \xi, \zeta) = (x, y, x\xi, \zeta).$$

Moreover, the image under π of the characteristic set $\Sigma \subset T^*X \setminus o$, given by

$$\Sigma = \{q \in T^*X : p(q) = 0\},$$

is the compressed characteristic set $\dot{\Sigma} = \pi(\Sigma)$. Note that (6-5) gives that

$$\dot{\Sigma} \cap \mathcal{U}_0 \cap {}^{\mathrm{b}}T^*_Y X = \{(0, y, 0, \underline{\zeta}) : 0 \leq \underline{\zeta} \cdot B(y)\underline{\zeta},\ \underline{\zeta} \neq 0\}. \tag{6-8}$$

In particular, in view of (6-7), $\dot{\Sigma} \cap \mathcal{U}_0$ lies in the region (6-1), at least after we possibly shrink U_0 (recall that $\mathcal{U}_0 = T^*U_0$), as we assume from now. We also remark that, using (6-6),

$$\pi_*|_{(x,y,\xi,\zeta)}\mathsf{H}_p = -2\xi \cdot (\partial_x + \underline{\xi}\,\partial_{\underline{\xi}}) + \mathsf{H}_h + x\beta\partial_{\underline{\xi}} + x\pi_* V, \tag{6-9}$$

and correspondingly

$$\mathsf{H}_p\pi^*\underline{\xi}\big|_{x=0} = -2\xi^2 = 2(p - \zeta \cdot B(y)\zeta) = -2\zeta \cdot B(y)\zeta, \quad \text{where } (0, y, \xi, \zeta) \in \Sigma. \tag{6-10}$$

As we already noted, $\underline{\zeta}_{n-1}$ cannot vanish on $\dot{\Sigma} \cap \mathcal{U}_0$, so

$$\begin{aligned} \mathsf{H}_p\pi^*(\underline{\xi}/|\underline{\zeta}_{n-1}|)\big|_{x=0} &= -2|\zeta_{n-1}|^{-1}\xi^2 - x\xi|\zeta_{n-1}|^{-2}(\mathsf{H}_h|\zeta_{n-1}|)\big|_{x=0} \\ &= -2|\zeta_{n-1}|^{-1}\zeta \cdot B(y)\zeta,\ (0, y, \xi, \zeta) \in \Sigma. \end{aligned} \tag{6-11}$$

To better understand the generalized broken bicharacteristics for $\Box$, we divide $\dot{\Sigma}$ into two subsets. We thus define the *glancing set* $\mathcal{G}$ as the set of points in $\dot{\Sigma}$ whose preimage under $\hat{\pi} = \pi|_\Sigma$ consists of a single point, and define the *hyperbolic set* $\mathcal{H}$ as its complement in $\dot{\Sigma}$. Thus, ${}^{\mathrm{b}}T^*_{X^\circ}X \cap \dot{\Sigma} \subset \mathcal{G}$ since π is a diffeomorphism on $T^*_{X^\circ}X$, while $q \in \dot{\Sigma} \cap {}^{\mathrm{b}}T^*_Y X$ lies in $\mathcal{G}$ if and only if on $\hat{\pi}^{-1}(\{q\})$, $\xi = 0$. More explicitly, with the notation of (6-8),

$$\begin{aligned} \mathcal{G} \cap \mathcal{U}_0 \cap {}^{\mathrm{b}}T^*_Y X &= \{(0, y, 0, \underline{\zeta}) : \underline{\zeta} \cdot B(y)\underline{\zeta} = 0,\ \underline{\zeta} \neq 0\}, \\ \mathcal{H} \cap \mathcal{U}_0 \cap {}^{\mathrm{b}}T^*_Y X &= \{(0, y, 0, \underline{\zeta}) : \underline{\zeta} \cdot B(y)\underline{\zeta} > 0,\ \underline{\zeta} \neq 0\}. \end{aligned} \tag{6-12}$$

Thus, $\mathcal{G}$ corresponds to generalized broken bicharacteristics that are tangent to Y in view of the vanishing of ξ at $\hat{\pi}^{-1}(\mathcal{G})$ (recall that the ∂_x component of H_p is -2ξ), while $\mathcal{H}$ corresponds to generalized broken bicharacteristics that are normal to Y. Note that if Y is one-dimensional (hence X is 2-dimensional), then $\underline{\zeta} \cdot B(y)\underline{\zeta} = 0$ necessarily implies $\underline{\zeta} = 0$, so in fact $\mathcal{G} \cap {}^{\mathrm{b}}T^*_Y X = \varnothing$, and hence there are no glancing rays.

We next make the role of $\mathcal{G}$ and $\mathcal{H}$ more explicit, which explains the relevant phenomena better. A characterization of GBB, which is equivalent to Definition 1.1, is this:

Lemma 6.1 (see the discussion in [Vasy 2005, Section 1] after the statement of Definition 1.1). *A continuous map $\gamma : I \to \dot{\Sigma}$, where $I \subset \mathbb{R}$ is an interval, is a GBB (in the analytic sense that we use here) if and only if it satisfies the following requirements*:

(i) *If* $q_0 = \gamma(s_0) \in \mathcal{G}$, *then for all* $f \in \mathcal{C}^\infty({}^{\mathrm{b}}T^*X)$,

$$\frac{d}{ds}(f \circ \gamma)(s_0) = \mathsf{H}_p(\pi^* f)(\tilde{q}_0) \quad \text{where } \tilde{q}_0 = \hat{\pi}^{-1}(q_0). \tag{6-13}$$

(ii) *If* $q_0 = \gamma(s_0) \in \mathcal{H}$, *then there exists* $\epsilon > 0$ *such that*

$$\gamma(t) \notin {}^{\mathrm{b}}T^*_Y X \quad \text{if} \quad 0 < |s - s_0| < \epsilon \text{ for } s \in I. \tag{6-14}$$

The idea of the proof of this lemma is that at $\mathcal{G}$, the requirement in (i) is equivalent to Definition 1.1 since $\hat{\pi}^{-1}(q_0)$ contains a single point. On the other hand, at $\mathcal{H}$, the requirement in (ii) follows from Definition 1.1 applied to the functions $f = \pm\underline{\xi}$, using (6-10), to conclude that $\underline{\xi}$ is strictly decreasing at $\mathcal{H}$ along GBB. Since one has $\underline{\xi} = 0$ on $\dot{\Sigma} \cap \{x = 0\}$, we have for a GBB γ through $\gamma(s_0) = q_0 \in \mathcal{H}$, on a punctured neighborhood of s_0, that $\underline{\xi}(\gamma(s)) \neq 0$, so $\gamma(s) \notin {}^{\mathrm{b}}T^*_Y X$ (since $\gamma(s) \in \dot{\Sigma}$). For the converse direction at $\mathcal{H}$ we refer to [Lebeau 1997]; see [Vasy 2005, Section 1] for details.

7. Microlocal elliptic regularity

We first note the form of $\Box$ with commutator calculations in mind. Rather than thinking of the tangential terms xD_y as "too degenerate", we think of xD_x as "too singular" in that it causes the failure of $\Box$ to lie in $x^2 \operatorname{Diff}^2_{\mathrm{b}}(X)$. This makes the calculations rather analogous to the conformal case, and also it facilitates the use of the symbolic machinery for b-pseudodifferential operators (b-PsDOs).

Proposition 7.1. *On a collar neighborhood of* Y, *the form of* $\Box$ *is*

$$-(xD_x)^*\alpha(xD_x) + (xD_x)^*M' + M''(xD_x) + \tilde{P}, \tag{7-1}$$

with

$$\alpha - 1 \in x\mathcal{C}^\infty(X), \qquad M', M'' \in x^2 \operatorname{Diff}^1_{\mathrm{b}}(X) \subset x \operatorname{Diff}^1_0(X),$$
$$\tilde{P} \in x^2 \operatorname{Diff}^2_{\mathrm{b}}(X), \quad \tilde{P} - x^2\Box_h \in x^3 \operatorname{Diff}^2_{\mathrm{b}}(X) \subset x \operatorname{Diff}^2_0(X),$$

where $\Box_h$ *is the d'Alembertian of the conformal metric on the boundary* (*extended to a neighborhood of* Y *using the collar structure*).

Proof. Writing the coordinates as $(z_1, \ldots, z_n)$, the operator $\Box_g$ is given by

$$\Box_g = \sum_{ij} D^*_{z_i} G_{ij} D_{z_j},$$

with adjoints taken with respect to $dg = |\det g|^{1/2} |dz_1 \cdots dz_n|$. With $z_j = y_j$ for $j = 1, \ldots, n-1$ and $z_n = x$, this can be rewritten as

$$\begin{aligned}\Box_g &= \sum_{ij} (xD_{z_i})^* \hat{G}_{ij} (xD_{z_j}) \\ &= (xD_x)^* \hat{G}_{nn}(xD_x) + \sum_{j=1}^{n-1} (xD_x)^* \hat{G}_{nj}(xD_{y_j}) + \sum_{j=1}^{n-1} (xD_{y_j})^* \hat{G}_{jn}(xD_{y_j}) + \sum_{i,j=1}^{n-1} (xD_{y_i})^* \hat{G}_{ij}(xD_{y_j}).\end{aligned}$$

Since $\hat{G}_{nn}+1 \in x\mathscr{C}^\infty(X)$, we find that $\alpha-1 \in x\mathscr{C}^\infty(X)$ by taking $\alpha = -\hat{G}_{nn}$. Since $\hat{G}_{jn}$, $\hat{G}_{nj} \in x\mathscr{C}^\infty(X)$, we find M', $M'' \in x^2 \operatorname{Diff}^1_{\mathrm{b}}(X)$ by taking $M' = \sum_{j=1}^{n-1} \hat{G}_{nj}(xD_{y_j})$ and $M'' = \sum_{j=1}^{n-1}(xD_{y_j})^* \hat{G}_{jn}$. Finally,

$$\tilde{P} = \sum_{ij=1}^{n-1} (xD_{y_i})^* \hat{G}_{ij}(xD_{y_j}) \in x^2 \operatorname{Diff}^2_{\mathrm{b}}(X).$$

Modulo $x^3 \operatorname{Diff}^2_{\mathrm{b}}(X)$, we can pull out the factors of x and restrict $\hat{G}_{ij}$ to Y. Therefore $\tilde{P}$ differs from $x^2\Box_h = x^2 \sum D^*_{y_i} h_{ij} D_{y_j}$ by an element of $x^3 \operatorname{Diff}^2_{\mathrm{b}}(X)$, completing the proof. □

We next state the lemma regarding Dirichlet form that is of fundamental use in both the elliptic and hyperbolic/glancing estimates. Below the main assumption is that $P = \Box_g + \lambda$, with $\Box_g$ as in (7-1). We first recall the notation for local norms:

Remark 7.2. Since X is noncompact and our results are microlocal, we may always fix a compact set $\tilde{K} \subset X$ and assume that all PsDOs have Schwartz kernel supported in $\tilde{K} \times \tilde{K}$. We also let $\tilde{U}$ be a neighborhood of $\tilde{K}$ in X such that $\tilde{U}$ has compact closure, and use the $H^1_0(\tilde{U})$ norm in place of the $H^1_0(X)$ norm to accommodate $u \in H^1_{0,\mathrm{loc}}(X)$. (We may instead take $\phi \in \mathscr{C}^\infty_{\mathrm{comp}}(\tilde{U})$ identically 1 in a neighborhood of $\tilde{K}$, and use $\|\phi u\|_{H^1_0(X)}$.) Below we use the notation $\|\cdot\|_{H^1_{0,\mathrm{loc}}(X)}$ for $\|\cdot\|_{H^1_0(\tilde{U})}$ to avoid having to specify $\tilde{U}$. We also use $\|v\|_{H^{-1}_{0,\mathrm{loc}}(X)}$ for $\|\phi v\|_{H^{-1}_0(X)}$.

Lemma 7.3 (see [Vasy 2008c, Lemma 4.2]). *Suppose that $K \subset {}^{\mathrm{b}}S^*X$ is compact, $U \subset {}^{\mathrm{b}}S^*X$ is open, and $K \subset U$. Suppose that $\mathscr{A} = \{A_r : r \in (0, 1]\}$ is a bounded family of PsDOs in $\Psi^s_{\mathrm{bc}}(X)$ with $\mathrm{WF}'_{\mathrm{b}}(\mathscr{A}) \subset K$, and with $A_r \in \Psi^{s-1}_{\mathrm{b}}(X)$ for $r \in (0, 1]$. Then there are $G \in \Psi^{s-1/2}_{\mathrm{b}}(X)$ and $\tilde{G} \in \Psi^{s+1/2}_{\mathrm{b}}(X)$ with $\mathrm{WF}'_{\mathrm{b}}(G)$, $\mathrm{WF}'_{\mathrm{b}}(\tilde{G}) \subset U$ and $C_0 > 0$ such that for $r \in (0, 1]$ and $u \in H^{1,k}_{0,\mathrm{b},\mathrm{loc}}(X)$ (here $k \le 0$) with neither $\mathrm{WF}^{1,s-1/2}_{\mathrm{b}}(u)$ nor $\mathrm{WF}^{-1,s+1/2}_{\mathrm{b}}(Pu)$ intersecting U, we have*

$$|\langle dA_r u, dA_r u\rangle_G + \lambda\|A_r u\|^2| \le C_0(\|u\|^2_{H^{1,k}_{0,\mathrm{b},\mathrm{loc}}(X)} + \|Gu\|^2_{H^1_0(X)} + \|Pu\|^2_{H^{-1,k}_{0,\mathrm{b},\mathrm{loc}}(X)} + \|\tilde{G}Pu\|^2_{H^{-1}_0(X)}).$$

Remark 7.4. The point of this lemma is G is $1/2$ order lower ($s - 1/2$ versus s) than the *family* $\mathscr{A}$. We will later take the limit $r \to 0$ to gain control of the Dirichlet form evaluated on $A_0 u$, where $A_0 \in \Psi^s_{\mathrm{bc}}(X)$, in terms of lower order information.

The role of A_r for $r > 0$ is to regularize such an argument, that is, to make sure various terms in a formal computation, in which one uses A_0 directly, actually make sense.

The main difference with [Vasy 2008c, Lemma 4.2] is that λ is *not negligible*.

Proof. We have $A_r u \in H^1_0(X)$ for $r \in (0, 1]$, so

$$\langle dA_r u, dA_r u\rangle + \lambda\|A_r u\|^2 = \langle PA_r u, A_r u\rangle.$$

Here the right side is the pairing of $H^{-1}_0(X)$ with $H^1_0(X)$, so by writing $PA_r = A_r P + [P, A_r]$, it can be estimated by

$$|\langle A_r Pu, A_r u\rangle| + |\langle [P, A_r]u, A_r u\rangle|. \tag{7-2}$$

The lemma is thus proved if we show that the first term of (7-2) is bounded by

$$C_0'\big(\|u\|^2_{H^{1,k}_{0,\mathrm{b,loc}}(X)} + \|Gu\|^2_{H^1_0(X)} + \|Pu\|^2_{H^{-1,k}_{0,\mathrm{b,loc}}(X)} + \|\tilde{G}Pu\|^2_{H^{-1}_0(X)}\big), \tag{7-3}$$

the second term is bounded by $C_0''(\|u\|^2_{H^{1,k}_{0,\mathrm{b,loc}}(X)} + \|Gu\|^2_{H^1_0(X)})$. (Recall that the "local" norms were defined in Remark 7.2.)

The first term is straightforward to estimate. Let $\Lambda \in \Psi_{\mathrm{b}}^{-1/2}(X)$ be elliptic with $\Lambda^- \in \Psi_{\mathrm{b}}^{1/2}(X)$ a parametrix, so that

$$E = \Lambda\Lambda^- - \mathrm{Id} \quad \text{and} \quad E' = \Lambda^-\Lambda - \mathrm{Id} \in \Psi_{\mathrm{b}}^{-\infty}(X).$$

Then

$$\langle A_r Pu, A_r u\rangle = \langle (\Lambda\Lambda^- - E)A_r Pu, A_r u\rangle = \langle \Lambda^- A_r Pu, \Lambda^* A_r u\rangle - \langle A_r Pu, E^* A_r u\rangle.$$

Since $\Lambda^- A_r$ is uniformly bounded in $\Psi_{\mathrm{bc}}^{s+1/2}(X)$ and $\Lambda^* A_r$ is uniformly bounded in $\Psi_{\mathrm{bc}}^{s-1/2}(X)$, we have $\langle \Lambda^- A_r Pu, \Lambda^* A_r\rangle$ is uniformly bounded, with a bound like (7-3) using Cauchy–Schwartz and Lemma 5.14. Indeed, by Lemma 5.14, if we choose any $G \in \Psi_{\mathrm{b}}^{s-1/2}(X)$ that is elliptic on K, there is a constant $C_1 > 0$ such that

$$\|\Lambda^* A_r u\|^2_{H^1_0(X)} \le C_1\big(\|u\|^2_{H^{1,k}_{0,\mathrm{b,loc}}(X)} + \|Gu\|^2_{H^1_0(X)}\big).$$

Similarly, by Lemma 5.14 and its analogue for $\mathrm{WF}_{\mathrm{b}}^{-1,s}$, if we choose any $\tilde{G} \in \Psi_{\mathrm{b}}^{s+1/2}(X)$ that is elliptic on K, there is a constant $C_1' > 0$ such that

$$\|\Lambda^- A_r Pu\|^2_{H^{-1}_0(X)} \le C_1'\big(\|Pu\|^2_{H^{-1,k}_{0,\mathrm{b,loc}}(X)} + \|\tilde{G}Pu\|^2_{H^{-1}_0(X)}\big).$$

Combining these gives, with $C_0' = C_1 + C_1'$, the desired result:

$$\begin{aligned}|\langle \Lambda^- A_r Pu, \Lambda^* A_r u\rangle| &\le \|\Lambda^- A_r Pu\|\,\|\Lambda^* A_r u\| \le \|\Lambda^- A_r Pu\|^2 + \|\Lambda^* A_r u\|^2 \\ &\le C_0'(\|u\|^2_{H^{1,k}_{0,\mathrm{b,loc}}(X)} + \|Gu\|^2_{H^1_0(X)} + \|Pu\|^2_{H^{-1,k}_{0,\mathrm{b,loc}}(X)} + \|\tilde{G}Pu\|^2_{H^{-1}_0(X)}).\end{aligned}$$

A similar argument, using that A_r is uniformly bounded in $\Psi_{\mathrm{bc}}^{s+1/2}(X)$ (in fact in $\Psi_{\mathrm{bc}}^{s}(X)$), and $E^* A_r$ is uniformly bounded in $\Psi_{\mathrm{bc}}^{s-1/2}(X)$ (in fact in $\Psi_{\mathrm{bc}}^{-\infty}(X)$), shows that $\langle A_r Pu, E^* A_r u\rangle$ is uniformly bounded.

Now we turn to the second term in (7-2), whose uniform boundedness is a direct consequence of Lemma 5.4 and Corollary 5.15. Indeed, by Lemma 5.4, $[P, A_r]$ is a bounded family in $\mathrm{Diff}_0^2\,\Psi_{\mathrm{bc}}^{s-1}(X)$; hence $A_r^*[P, A_r]$ is a bounded family in $\mathrm{Diff}_0^2\,\Psi_{\mathrm{bc}}^{2s-1}(X)$. Then one can apply Corollary 5.15 to conclude that

$$\langle A_r^*[P, A_r]u, u\rangle \le C'\big(\|u\|^2_{H^{1,k}_{0,\mathrm{b,loc}}(X)} + \|Gu\|^2_{H^1(X)}\big). \qquad \square$$

A more precise version, in terms of requirements on Pu, is the following. Here, as in Section 2, we fix a positive definite inner product on the fibers of ${}^0T^*X$ (that is, a Riemannian 0-metric) to compute $\|dv\|^2_{L^2(X;{}^0T^*X)}$; since v has support in a compact set below, the choice of the inner product is irrelevant.

Lemma 7.5 (see [Vasy 2008c, Lemma 4.4]). *Suppose that $K \subset {}^{\mathrm{b}}S^*X$ is compact, $U \subset {}^{\mathrm{b}}S^*X$ is open, and $K \subset U$. Suppose that $\mathcal{A} = \{A_r : r \in (0,1]\}$ is a bounded family of PsDOs in $\Psi^s_{\mathrm{bc}}(X)$ with $\mathrm{WF}'_{\mathrm{b}}(\mathcal{A}) \subset K$ and with $A_r \in \Psi^{s-1}_{\mathrm{b}}(X)$ for $r \in (0,1]$. Then there are $G \in \Psi^{s-1/2}_{\mathrm{b}}(X)$ and $\tilde{G} \in \Psi^s_{\mathrm{b}}(X)$ with $\mathrm{WF}'_{\mathrm{b}}(G), \mathrm{WF}'_{\mathrm{b}}(\tilde{G}) \subset U$ and $C_0 > 0$ such that for $\epsilon > 0$, $r \in (0,1]$, $u \in H^{1,k}_{0,\mathrm{b,loc}}(X)$ (where $k \le 0$) with neither $\mathrm{WF}^{1,s-1/2}_{\mathrm{b}}(u)$ nor $\mathrm{WF}^{-1,s}_{\mathrm{b}}(Pu)$ intersecting U, we have*

$$\begin{aligned}&\big|\langle dA_ru, dA_ru\rangle_G + \lambda\|A_ru\|^2\big|\\ &\qquad \le \epsilon\|dA_ru\|^2_{L^2(X;{}^0T^*X)} + C_0\big(\|u\|^2_{H^{1,k}_{0,\mathrm{b,loc}}(X)} + \|Gu\|^2_{H^1_0(X)} + \epsilon^{-1}\|Pu\|^2_{H^{-1,k}_{0,\mathrm{b,loc}}(X)} + \epsilon^{-1}\|\tilde{G}Pu\|^2_{H^{-1}_0(X)}\big).\end{aligned}$$

Remark 7.6. The point of this lemma is that on the one hand the new term $\epsilon\|dA_ru\|^2$ can be absorbed in the left hand side in the elliptic region and hence is negligible; on the other hand, there is a gain in the order of $\tilde{G}$ (s versus $s+1/2$ in the previous lemma).

Proof. We need only modify the previous proof slightly, by estimating the term $|\langle A_rPu, A_ru\rangle|$ in (7-2) differently, namely

$$|\langle A_rPu, A_ru\rangle| \le \|A_rPu\|_{H^{-1}_0(X)}\|A_ru\|_{H^1_0(X)} \le \tilde{\epsilon}\|A_ru\|^2_{H^1_0(X)} + \tilde{\epsilon}^{-1}\|A_rPu\|^2_{H^{-1}_0(X)}.$$

Now the lemma follows by using Lemma 5.14 and the remark following it: Choosing any $\tilde{G} \in \Psi^s_{\mathrm{b}}(X)$ that is elliptic on K gives a constant $C'_1 > 0$ such that

$$\|A_rPu\|^2_{H^{-1}_0(X)} \le C'_1\big(\|Pu\|^2_{H^{-1,k}_{0,\mathrm{b,loc}}(X)} + \|\tilde{G}Pu\|^2_{H^{-1}_0(X)}\big).$$

We then use the Poincaré inequality to estimate $\|A_ru\|_{H^1_0(X)}$ by $C_2\|dA_ru\|_{L^2(X)}$, and finish the proof exactly as for Lemma 7.3. □

We next state microlocal elliptic regularity. For this result the restrictions on $\lambda \in \mathbb{C}$ are weak (only a half-line is disallowed), but on the other hand, a solution u satisfying our hypotheses may not exist for values of λ when $\lambda \notin (-\infty, (n-1)^2/4)$.

Proposition 7.7 (microlocal elliptic regularity). *Suppose that $P = \Box + \lambda$, $\lambda \in \mathbb{C} \setminus [(n-1)^2/4, \infty)$ and $m \in \mathbb{R}$ or $m = \infty$. Suppose $u \in H^{1,k}_{0,\mathrm{b,loc}}(X)$ for some $k \le 0$. Then*

$$\mathrm{WF}^{1,m}_{\mathrm{b}}(u) \setminus \dot{\Sigma} \subset \mathrm{WF}^{-1,m}_{\mathrm{b}}(Pu).$$

Proof. We first prove a slightly weaker result in which $\mathrm{WF}^{-1,m}_{\mathrm{b}}(Pu)$ is replaced by $\mathrm{WF}^{-1,m+1/2}_{\mathrm{b}}(Pu)$ — we rely on Lemma 7.3. We then prove the original statement using Lemma 7.5.

Suppose that $q \in {}^{\mathrm{b}}T^*_YX \setminus \dot{\Sigma}$. We may assume iteratively that $q \notin \mathrm{WF}^{1,s-1/2}_{\mathrm{b}}(u)$; we need to prove then that $q \notin \mathrm{WF}^{1,s}_{\mathrm{b}}(u)$ provided $s \le m+1/2$ (note that the inductive hypothesis holds for $s = k+1/2$ since $u \in H^{1,k}_{0,\mathrm{b,loc}}(X)$). We use local coordinates (x, y) as in Section 6, centered so that $q \in {}^{\mathrm{b}}T^*_{(0,0)}X$ and arranging that (6-7) holds. We further group the variables as $y = (y', y_{n-1})$, with corresponding b-dual variables $(\underline{\zeta}', \underline{\zeta}_{n-1})$. We denote the Euclidean norm by $|\underline{\zeta}'|$.

Let $A \in \Psi^s_{\mathrm{b}}(X)$ be such that

$$\mathrm{WF}'_{\mathrm{b}}(A) \cap \mathrm{WF}^{1,s-1/2}_{\mathrm{b}}(u) = \varnothing \quad \text{and} \quad \mathrm{WF}'_{\mathrm{b}}(A) \cap \mathrm{WF}^{1,s+1/2}_{\mathrm{b}}(Pu) = \varnothing$$

and that $\mathrm{WF}'_{\mathrm{b}}(A)$ in a small conic neighborhood U of q, with U such that for a suitable $C > 0$ or $\epsilon > 0$,

(i) $\underline{\zeta}_{n-1}^2 < C\underline{\xi}^2$ if $\underline{\xi}(q) \neq 0$,

(ii) $|\underline{\xi}| < \epsilon|\underline{\zeta}|$ for all j, and $|\underline{\zeta}'|/|\underline{\zeta}_{n-1}| > 1+\epsilon$ if $\underline{\xi}(q) = 0$ and $\underline{\zeta}(q) \cdot B(y(q))\underline{\zeta}(q) < 0$.

Let $\Lambda_r \in \Psi_{\mathrm{b}}^{-2}(X)$ for $r > 0$, such that $\mathscr{L} = \{\Lambda_r : r \in (0, 1]\}$ is a bounded family in $\Psi_{\mathrm{b}}^0(X)$, and $\Lambda_r \to \mathrm{Id}$ as $r \to 0$ in $\Psi_{\mathrm{b}}^{\tilde\epsilon}(X)$ for $\tilde\epsilon > 0$. For example, the symbol of Λ_r could be taken as $(1 + r(|\underline{\zeta}|^2 + |\underline{\xi}|^2))^{-1}$. Let $A_r = \Lambda_r A$. Let a be the symbol of A, and let A_r have symbol $(1 + r(|\underline{\zeta}|^2 + |\underline{\xi}|^2))^{-1}a$ for $r > 0$, so $A_r \in \Psi_{\mathrm{b}}^{s-2}(X)$ for $r > 0$, and A_r is uniformly bounded in $\Psi_{\mathrm{bc}}^s(X)$, and $A_r \to A$ in $\Psi_{\mathrm{bc}}^{s+\tilde\epsilon}(X)$.

By Lemma 7.3,

$$\langle dA_r u, dA_r u\rangle_G + \lambda\|A_r u\|^2$$

is uniformly bounded for $r \in (0, 1]$, so

$$\langle dA_r u, dA_r u\rangle_G + \operatorname{Re}\lambda\|A_r u\|^2 \quad \text{and} \quad \operatorname{Im}\lambda\|A_r u\|^2$$

are uniformly bounded. If $\operatorname{Im}\lambda \neq 0$, then taking the imaginary part at once shows that $\|A_r u\|$ is in fact uniformly bounded. On the other hand, whether $\operatorname{Im}\lambda = 0$ or not,

$$\begin{aligned}\langle dA_r u, dA_r u\rangle_G = \int_X A(x, y) x D_x A_r u\, \overline{x D_x A_r u}\, dg + \int_X \sum B_{ij}(x, y) x D_{y_i} A_r u\, \overline{x D_{y_j} A_r u}\, dg \\ + \int_X \sum C_j(x, y) x D_x A_r u\, \overline{x D_{y_j} A_r u}\, dg + \int_X \sum C_j(x, y) x D_{y_j} A_r u\, \overline{x D_x A_r u}\, dg.\end{aligned}$$

Using that $A(x, y) = -1 + xA'(x, y) + \sum(y_j - y_j(q))A_j(x, y)$, we see that if A_r is supported where $x < \delta$ and $|y_j - y_j(q)| < \delta$ for all j, then for some $C > 0$ (independent of A_r),

$$\left|\int_X A(x, y)\, x D_x A_r u\, \overline{x D_x A_r u}\, dg - \int_X A(0, y(q))\, x D_x A_r u\, \overline{x D_x A_r u}\, dg\right| \leq C\delta\|x D_x A_r u\|^2, \quad (7\text{-}4)$$

with analogous estimates[4] for $B_{ij}(x, y) - B_{ij}(0, y(q))$ and for $C_j(x, y)$. Thus, there exists $\tilde C > 0$ and $\delta_0 > 0$ such that if $\delta < \delta_0$ and A is supported where $|x| < \delta$ and $|y - y(q)| < \delta$, then

$$\begin{aligned}\int_X \big((1 - \tilde C\delta)|x D_x A_r u|^2 - \operatorname{Re}\lambda|A_r u|^2\big)\, dg + \sum_{j=1}^{n-2}\int_X \Big((1 - \tilde C\delta)\sum_j x D_{y_j} A_r u\, \overline{x D_{y_j} A_r u}\Big)\, dg \\ - \int_X \Big((1 + \tilde C\delta)\sum_j x D_{y_{n-1}} A_r u\, \overline{x D_{y_{n-1}} A_r u}\Big)\, dg \\ \leq |\langle dA_r u, dA_r u\rangle_G + \operatorname{Re}\lambda\|A_r u\|^2|. \quad (7\text{-}5)\end{aligned}$$

Now we distinguish the cases $\underline{\xi}(q) = 0$ and $\underline{\xi}(q) \neq 0$. If $\underline{\xi}(q) = 0$, we choose $\delta \in (0, 1/(2\tilde C))$ with $\delta < \delta_0$, so that

$$(1 - \tilde C\delta)(|\underline{\zeta}'|^2/\underline{\zeta}_{n-1}^2) > 1 + 2\tilde C\delta$$

[4]Recall that $C_j(0, y) = 0$ and $B_{ij}(0, y(q)) = 0$ if $i \neq j$ and $B_{ij}(0, y(q)) = 1$ if $i = j = n-1$ and $B_{ij}(0, y(q)) = -1$ if $i = j \neq n-1$.

on a neighborhood of $\mathrm{WF}'_{\mathrm{b}}(A)$, which is possible in view of (ii) at the beginning of the proof. Then the second integral on the left side of (7-5) can be written as $\|BxA_ru\|^2$, with the symbol of B given by

$$\big((1-\tilde{C}\delta)|\underline{\zeta}'|^2-(1+\tilde{C}\delta)\underline{\zeta}_{n-1}^2\big)^{1/2}$$

(which is $\geq \delta|\underline{\zeta}_{n-1}|$), modulo a term

$$\int_X FxA_ru\,\overline{xA_ru}\,dg \quad \text{for } F\in\Psi_{\mathrm{b}}^1(X).$$

But $A_r^*xFxA_r$ is uniformly bounded in $x^2\Psi_{\mathrm{bc}}^{2s+1}(X)\subset \mathrm{Diff}_0^2\,\Psi_{\mathrm{bc}}^{2s-1}(X)$, so this expression is uniformly bounded as $r\to 0$ by Corollary 5.15. We thus deduce that

$$\int_X\big((1-\tilde{C}\delta)|xD_xA_ru|^2-\mathrm{Re}\,\lambda|A_ru|^2\big)\,dg+\|BxA_ru\|^2$$

is uniformly bounded as $r\to 0$.

If $\underline{\xi}(q)\neq 0$, and A is supported in $|x|<\delta$, then

$$\tilde{C}\delta\int_X\delta^{-2}|x^2D_xA_ru|^2\,dg\leq \tilde{C}\delta\int_X|xD_xA_ru|^2\,dg.$$

On the other hand, near $\{q':\underline{\xi}(q')=0\}$, for $\delta>0$ sufficiently small,

$$\int_X\Big(\frac{\tilde{C}\delta}{\delta^2}|x^2D_xA_ru|^2-|xD_{y_{n-1}}A_ru|^2\Big)\,dg=\|BxA_ru\|^2+\int_X FxA_ru\,\overline{xA_ru}\,dg,$$

with the symbol of B given by $((\tilde{C}/\delta)\underline{\xi}^2-\underline{\zeta}_{n-1}^2)^{1/2}$ (which does not vanish on U for $\delta>0$ small), while $F\in\Psi_{\mathrm{b}}^1(X)$, so the second term on the right side is uniformly bounded as $r\to 0$ just as above. We thus deduce in this case that

$$\int_X((1-2\tilde{C}\delta)|xD_xA_ru|^2\,dg-\mathrm{Re}\,\lambda|A_ru|^2)+\|BxA_ru\|^2$$

is uniformly bounded as $r\to 0$.

If $\mathrm{Im}\,\lambda\neq 0$ then we already saw that $\|A_ru\|_{L^2}$ is uniformly bounded, so we deduce that

$$A_ru,\ \ xD_xA_ru \text{ and } BxA_ru \text{ are uniformly bounded in } L^2(X). \tag{7-6}$$

If $\mathrm{Im}\,\lambda=0$ but $\lambda<(n-1)^2/4$, then the Poincaré inequality allows us to reach the same conclusion, since on the one hand in case (ii)

$$(1-\tilde{C}\delta)\|xD_xA_ru\|^2-\mathrm{Re}\,\lambda\|A_ru\|^2,$$

and in case (i)

$$(1-2\tilde{C}\delta)\|xD_xA_ru\|^2-\mathrm{Re}\,\lambda\|A_ru\|^2,$$

are uniformly bounded; on the other hand by Proposition 2.3, for $\delta > 0$ sufficiently small there exists $c > 0$ such that

$$(1 - 2\tilde{C}\delta)\|xD_x A_r u\|^2 - \operatorname{Re}\lambda \|A_r u\|^2 \geq c(\|xD_x A_r u\|^2 + \|A_r u\|^2).$$

Correspondingly there are sequences $A_{r_k}u$, $xD_x A_{r_k}u$ and $BxA_{r_k}u$, weakly convergent in $L^2(X)$, and such that $r_k \to 0$, as $k \to \infty$. Since they respectively converge to Au, $xD_x Au$ and $BxAu$ in $\mathscr{C}^{-\infty}(X)$, we deduce that the weak limits are Au, $xD_x Au$ and $BxAu$, which therefore lie in $L^2(X)$. Consequently, $q \notin \mathrm{WF}_{\mathrm{b}}^{1,s}(u)$, hence proving the proposition with $\mathrm{WF}_{\mathrm{b}}^{-1,m}(Pu)$ replaced by $\mathrm{WF}_{\mathrm{b}}^{-1,m+1/2}(Pu)$.

To obtain the optimal result, we note that due to Lemma 7.5 we still have, for any $\epsilon > 0$, that

$$\langle dA_r u, dA_r u\rangle_G - \epsilon \|dA_r u\|^2$$

is uniformly bounded above for $r \in (0, 1]$. By arguing just as above, with B as above, for sufficiently small $\epsilon > 0$, the right side gives an upper bound for

$$\int_X \big((1 - 2\tilde{C}\delta - \epsilon)|xD_x A_r u|^2 - \operatorname{Re}\lambda |A_r u|^2\big)\, dg + \|BxA_r u\|^2,$$

which is thus uniformly bounded as $r \to 0$. The proof is then finished exactly as above. □

The analogous argument works for the conformally compact elliptic problem, that is, on asymptotically hyperbolic spaces, to give that for $\lambda \in \mathbb{C} \setminus [(n-1)^2/4, \infty)$, local solutions of $(\Delta_g - \lambda)u$ are actually conormal to Y provided they lie in $H_0^1(X)$ locally, or indeed in $H_{0,\mathrm{b}}^{1,-\infty}(X)$.

8. Propagation of singularities

In this section we prove propagation of singularities for P by positive commutator estimates. We do so by first performing a general commutator calculation in Proposition 8.1, then using it to prove rough propagation estimates first at hyperbolic, then at glancing points, in Propositions 8.2 and 8.6, respectively. An argument originally due to Melrose and Sjöstrand [1978] then proves the main theorems, Theorems 8.8 and 8.9. Finally we discuss consequences of these results.

We first describe the form of commutators of P with $\Psi_{\mathrm{b}}(X)$. We state this as an analogue of [Vasy 2010a, Proposition 3.10], and later in the section we follow the structure of [Vasy 2010a] as well. Given Proposition 8.1 below, the proof of propagation of singularities proceeds with the same commutant construction as in [Vasy 2008c]; see also [Vasy 2008a]. Although it is in a setting that is more complicated in some ways, since it deals with the equation on differentials forms, we follow the structure of [Vasy 2010a] since it was written in a more systematic way than [Vasy 2008c]. Recall from the introduction that $\underline{\xi}$ is the variable b-dual to x, and $\hat{\underline{\xi}} = \underline{\xi}/|\zeta_{n-1}|$.

Proposition 8.1. *Suppose $\mathscr{A} = \{A_r : r \in (0, 1]\}$ is a family of operators $A_r \in \Psi_{\mathrm{b}}^0(X)$ uniformly bounded in $\Psi_{\mathrm{bc}}^{s+1/2}(X)$, of the form $A_r = A\Lambda_r$, with $A \in \Psi_{\mathrm{b}}^0(X)$, $a = \sigma_{\mathrm{b},0}(A)$ and $w_r = \sigma_{\mathrm{b},s+1/2}(\Lambda_r)$. Then*

$$\imath[A_r^* A_r, \Box] = (xD_x)^* C_r^\sharp (xD_x) + (xD_x)^* x C_r' + x C_r''(xD_x) + x^2 C_r^\flat, \tag{8-1}$$

where

$$C_r^\sharp \in L^\infty((0,1];\Psi_{\mathrm{bc}}^{2s}(X)),\quad C_r', C_r'' \in L^\infty((0,1];\Psi_{\mathrm{bc}}^{2s+1}(X)),\quad C_r^\flat \in \Psi_{\mathrm{bc}}^{2s+2}(X),$$

and

$$\begin{aligned}\sigma_{\mathrm{b},2s}(C_r^\sharp) &= 2w_r^2 a(V^\sharp a + a\tilde{c}_r^\sharp),\\ \sigma_{\mathrm{b},2s+1}(C_r') = \sigma_{\mathrm{b},2s+1}(C_r'') &= 2w_r^2 a(V'a + a\tilde{c}_r'),\\ \sigma_{\mathrm{b},2s+2}(C_r^\flat) &= 2w_r^2 a(V^\flat a + a\tilde{c}_r^\flat),\end{aligned}$$

with $\tilde{c}_r^\sharp$, $\tilde{c}_r'$, $\tilde{c}_r^\flat$ *uniformly bounded in* S^{-1}, S^0, S^1, *respectively,* $V^\sharp$, V', $V^\flat$ *smooth and homogeneous of degree* $-1, 0, 1$ *respectively on* ${}^{\mathrm{b}}T^*X \setminus o$, *and where* $V^\sharp|_Y$ *and* $V'|_Y$ *annihilate* $\underline{\xi}$ *and*

$$V^\flat|_Y = 2h\partial_{\underline{\xi}} - \mathsf{H}_h. \tag{8-2}$$

Proof. In Proposition 7.1, $\Box$ is decomposed into a sum of products of weighted b-operators, so analogously expanding the commutator, all calculations can be done in $x^l\Psi_{\mathrm{b}}(X)$ for various values of l. In particular, keeping in mind Lemma 5.1 (which gives the additional order of decay),

$$\imath[A_r^*A_r, xD_x], \imath[A_r^*A_r, (xD_x)^*] \in L^\infty((0,1]_r, x\Psi_{\mathrm{b}}^{2s+1}(X)),$$

with principal symbol $-2w_r^2 a x\partial_x a - 2a^2 w_r(x\partial_x w_r)$. By this observation, all commutators with factors of xD_x or $(xD_x)^*$ in (7-1) can be absorbed into the "next term" of (8-1), so $[A_r^*A_r, (xD_x)^*]\alpha(xD_x)$ is absorbed into $xC_r''(xD_x)$, $(xD_x)\alpha[A_r^*A_r, xD_x]$ is absorbed into $(xD_x)^*xC_r'$, and $[A_r^*A_r, (xD_x)^*]M'$ and $M''[A_r^*A_r, (xD_x)]$ are absorbed into $x^2C_r^\flat$. The principal symbols of these terms are of the desired form, that is, after factoring out $2w_r^2 a$, they are the result of a vector field applied to a plus a multiple of a, and this vector field is $-\alpha\partial_x$ in the case of the first two terms (thus annihilating $\underline{\xi}$), and is $-mx^{-1}\partial_x$ in the case of the last two terms, which in view of $m = \sigma_{\mathrm{b},1}(M') = \sigma_{\mathrm{b},1}(M'') \in x^2S^1$, shows that it actually does not affect $V^\flat|_Y$.

Next, $\imath(xD_x)^*[A_r^*A_r, \alpha](xD_x)$ can be absorbed into (and can be taken equal to) $(xD_x)^*C_r^\sharp(xD_x)$ with principal symbol of $C_r^\sharp$ given by

$$-(\partial_y\alpha)\partial_{\underline{\zeta}}(a^2w_r^2) - (x\partial_x\alpha)\partial_{\underline{\xi}}(a^2w_r^2)$$

in local coordinates; thus again is of the desired form since the $\partial_{\underline{\xi}}$ term has a vanishing factor of x preceding it.

Since $[A_r^*A_r, M']$ and $[A_r^*A_r, M'']$ are uniformly bounded in $x^2\Psi_{\mathrm{b}}^{2s+1}(X)$, the corresponding commutators can be absorbed into $(xD_x)^*xC_r'$ and $xC_r''(xD_x)$, respectively, without affecting the principal symbols of C_r' and C_r'' at Y, and possessing the desired form.

Next, $\tilde{P} = x^2\Box_h + R$, with $R \in x^3\,\mathrm{Diff}_{\mathrm{b}}^2(X)$, so $[A_r^*A_r, R]$ is uniformly bounded in $x^3\Psi_{\mathrm{b}}^{2s+2}(X)$, and thus can be absorbed into $C_r^\flat$ without affecting its principal symbol at Y, and it has the desired form. Finally, $\imath[A_r^*A_r, x^2\Box_h] \in x^2\Psi_{\mathrm{b}}^{2s+2}(X)$ has principal symbol $\partial_{\underline{\xi}}(a^2w_r^2)2x^2h - x^2\mathsf{H}_h(a^2w_r^2)$, and can thus be absorbed into $C_r^\flat$, yielding the stated principal symbol at Y. $\Box$

We start our propagation results with the propagation estimate at hyperbolic points.

Proposition 8.2 (normal, or hyperbolic, propagation). *Suppose* $P = \Box_g + \lambda$, *with* $\lambda \in \mathbb{C} \setminus [(n-1)^2/4, \infty)$. *Let* $q_0 = (0, y_0, 0, \underline{\zeta}_0) \in \mathcal{H} \cap {}^{\mathrm{b}}T^*_Y X$, *and let*

$$\eta = -\hat{\underline{\xi}}$$

be the function defined in the local coordinates discussed above, and suppose that $u \in H^{1,k}_{0,\mathrm{b},\mathrm{loc}}(X)$ *for some* $k \le 0$, $q_0 \notin \mathrm{WF}_{\mathrm{b}}^{-1,\infty}(f)$ *and* $f = Pu$. *If* $\operatorname{Im} \lambda \le 0$ *and there exists a conic neighborhood* U *of* q_0 *in* ${}^{\mathrm{b}}T^*X \setminus o$ *such that*

$$q \notin \mathrm{WF}_{\mathrm{b}}^{1,\infty}(u) \quad \text{if } q \in U \text{ and } \eta(q) < 0, \tag{8-3}$$

then $q_0 \notin \mathrm{WF}_{\mathrm{b}}^{1,\infty}(u)$.

In fact, if the wave front set assumptions are relaxed to $q_0 \notin \mathrm{WF}_{\mathrm{b}}^{-1,s+1}(f)$ (*with* $f = Pu$) *and the existence of a conic neighborhood* U *of* q_0 *in* ${}^{\mathrm{b}}T^*X \setminus o$ *such that*

$$q \notin \mathrm{WF}_{\mathrm{b}}^{1,s}(u) \quad \text{if } q \in U \text{ and } \eta(q) < 0, \tag{8-4}$$

then we can still conclude that $q_0 \notin \mathrm{WF}_{\mathrm{b}}^{1,s}(u)$.

Remark 8.3. As follows immediately from the proof given below, in (8-3) and (8-4), one can replace $\eta(q) < 0$ by $\eta(q) > 0$, that is, one has the conclusion for either direction (backward or forward) of propagation, *provided one also switches the sign of* $\operatorname{Im}\lambda$ *when it is nonzero that is, the assumption should be* $\operatorname{Im}\lambda \ge 0$. In particular, if $\operatorname{Im}\lambda = 0$, one obtains propagation estimates both along increasing and along decreasing η.

Note that η is *increasing* along the GBB of $\Box_{\hat{g}}$ by (6-11). Thus, the hypothesis region $\{q \in U : \eta(q) < 0\}$ on the left side of (8-3) is *backwards* from q_0, so this proposition, roughly speaking, propagates regularity *forwards*.

Moreover, every neighborhood U of $q_0 = (y_0, \underline{\zeta}_0) \in \mathcal{H} \cap {}^{\mathrm{b}}T^*_Y X$ in $\dot{\Sigma}$ contains an open set of the form

$$\{q : |x(q)|^2 + |y(q) - y_0|^2 + |\hat{\underline{\zeta}}(q) - \hat{\underline{\zeta}}_0|^2 < \delta\}, \tag{8-5}$$

see [Vasy 2008c, Equation (5.1)]. Note also that (8-3) implies the same statement with U replaced by any smaller neighborhood of q_0 and in particular for the set (8-5), provided that δ is sufficiently small. We can also assume by the same observation that $\mathrm{WF}_{\mathrm{b}}^{-1,s+1}(Pu) \cap U = \varnothing$. Furthermore, we can also arrange that $h(x, y, \underline{\xi}, \underline{\zeta}) > |(\underline{\xi}, \underline{\zeta})|^2 |\underline{\zeta}_0|^{-2} h(q_0)/2$ on U since $\underline{\zeta}_0 \cdot B(y_0)\underline{\zeta}_0 = h(0, y_0, 0, \underline{\zeta}_0) > 0$. We write

$$\hat{h} = |\underline{\zeta}_{n-1}|^{-2} h = |\underline{\zeta}_{n-1}|^{-2} \underline{\zeta} \cdot B(y)\underline{\zeta}$$

for the rehomogenized version of h, which is thus homogeneous of degree zero and bounded below by a positive constant on U.

Proof. This proposition is the analogue of [Vasy 2008c, Proposition 6.2], and since the argument is similar, we mainly emphasize the differences. These enter by virtue of λ not being negligible and the use of the Poincaré inequality. In [Vasy 2008c], one uses a commutant $A \in \Psi^0_{\mathrm{b}}(X)$ and weights $\Lambda_r \in \Psi^0_{\mathrm{b}}(X)$ for $r \in (0, 1)$, which are uniformly bounded in $\Psi^{s+1/2}_{\mathrm{bc}}(X)$, with $A_r = A\Lambda_r$, in order to obtain the propagation of $\mathrm{WF}_{\mathrm{b}}^{1,s}(u)$ with the notation of that paper, whose analogue is $\mathrm{WF}_{\mathrm{b}}^{1,s}(u)$ here (the difference is the space relative to which one obtains b-regularity: $H^1(X)$ in the previous paper, the zero-Sobolev space $H_0^1(X)$ here). One can use *exactly the same* commutant as in [Vasy 2008c]. Then Proposition 8.1 lets one calculate $\imath[A_r^* A_r, P]$ to obtain an expression completely analogous to [Vasy 2008c, Equation (6.18)] in the hyperbolic case. We also refer to [Vasy 2010a] because, although it studies a more delicate problem, namely natural boundary conditions (which are not scalar), the main ingredient of the proof, the commutator calculation, is written up exactly as above in Proposition 8.1; see [Vasy 2010a, Proposition 3.10] and the way it is used subsequently in Proposition 5.1 there.

As in the proof of [Vasy 2010a, Proposition 5.1], we first construct a commutant by defining its scalar principal symbol a. This completely follows the scalar case; see the proof [Vasy 2008c, Proposition 6.2]. Next we show how to obtain the desired estimate.

So, as in the proof [Vasy 2008c, Proposition 6.2], let

$$\omega(q) = |x(q)|^2 + |y(q) - y_0|^2 + |\hat{\underline{\xi}}(q) - \hat{\underline{\xi}}_0|^2, \tag{8-6}$$

with $|\cdot|$ denoting the Euclidean norm. For $\epsilon > 0$ and $\delta > 0$, with other restrictions to be imposed later on, let

$$\phi = \eta + \frac{1}{\epsilon^2 \delta}\omega, \tag{8-7}$$

Let $\chi_0 \in \mathscr{C}^\infty(\mathbb{R})$ be equal to 0 on $(-\infty, 0]$ and $\chi_0(t) = \exp(-1/t)$ for $t > 0$. Thus, $t^2\chi_0'(t) = \chi_0(t)$ for $t \in \mathbb{R}$. Let $\chi_1 \in \mathscr{C}^\infty(\mathbb{R})$ be 0 on $(-\infty, 0]$ and 1 on $[1, \infty)$, with $\chi_1' \geq 0$ satisfying $\chi_1' \in \mathscr{C}^\infty_{\mathrm{comp}}((0, 1))$. Finally, let $\chi_2 \in \mathscr{C}^\infty_{\mathrm{comp}}(\mathbb{R})$ be supported in $[-2c_1, 2c_1]$ and identically 1 on $[-c_1, c_1]$, where c_1 satisfies $|\hat{\underline{\xi}}|^2 < c_1/2$ in $\dot{\Sigma} \cap U$. Thus, $\chi_2(|\hat{\underline{\xi}}|^2)$ is a cutoff in $|\hat{\underline{\xi}}|$, with its support properties ensuring that $d\chi_2(|\hat{\underline{\xi}}|^2)$ is supported in $|\hat{\underline{\xi}}|^2 \in [c_1, 2c_1]$ and hence outside $\dot{\Sigma}$ — it should be thought of as a factor that microlocalizes near the characteristic set but effectively commutes with P (since we already have the microlocal elliptic result). Then, for $\digamma > 0$ large, to be determined, let

$$a = \chi_0(\digamma^{-1}(2 - \phi/\delta))\chi_1(\eta/\delta + 2)\chi_2(|\hat{\underline{\xi}}|^2); \tag{8-8}$$

so a is a homogeneous degree zero $\mathscr{C}^\infty$ function on a conic neighborhood of q_0 in ${}^{\mathrm{b}}T^*X \setminus o$. Indeed as we will see momentarily, a has for any $\epsilon > 0$ compact support inside this neighborhood (regarded as a subset of ${}^{\mathrm{b}}S^*X$, that is, quotienting out by the $\mathbb{R}^+$-action) for δ sufficiently small, so in fact it is globally well defined. In fact, on $\operatorname{supp} a$ we have $\phi \leq 2\delta$ and $\eta \geq -2\delta$. Since $\omega \geq 0$, the first of these inequalities implies that $\eta \leq 2\delta$, so on $\operatorname{supp} a$

$$|\eta| \leq 2\delta. \tag{8-9}$$

Hence,

$$\omega \leq \epsilon^2\delta(2\delta - \eta) \leq 4\delta^2\epsilon^2. \tag{8-10}$$

In view of (8-6) and (8-5), this shows that given any $\epsilon_0 > 0$ there exists $\delta_0 > 0$ such that a is supported in U for any $\epsilon \in (0, \epsilon_0)$ and $\delta \in (0, \delta_0)$. The role that $\digamma$ large plays (in the definition of a) is that it increases the size of the first derivatives of a relative to the size of a; hence it allows us to give a bound for a in terms of a small multiple of its derivative along the Hamilton vector field, much like the stress energy tensor was used to bound other terms by making χ' large relative to χ in the (nonmicrolocal) energy estimate.

Now let $A_0 \in \Psi_b^0(X)$ with $\sigma_{b,0}(A_0) = a$, supported in the coordinate chart. Also let Λ_r be scalar and have symbol

$$|\underline{\zeta}_{n-1}|^{s+1/2}(1 + r|\underline{\zeta}_{n-1}|^2)^{-s}\,\mathrm{Id} \quad \text{for } r \in [0, 1), \tag{8-11}$$

so $A_r = A\Lambda_r \in \Psi_b^0(X)$ for $r > 0$ and it is uniformly bounded in $\Psi_{bc}^{s+1/2}(X)$. Then, for $r > 0$,

$$\begin{aligned}\langle \imath A_r^* A_r P u, u\rangle - \langle \imath A_r^* A_r u, P u\rangle &= \langle \imath [A_r^* A_r, P] u, u\rangle + \langle \imath (P - P^*) A_r^* A_r u, u\rangle \\ &= \langle \imath [A_r^* A_r, P] u, u\rangle - 2\,\mathrm{Im}\,\lambda \|A_r u\|^2.\end{aligned} \tag{8-12}$$

We can compute this using Proposition 8.1. We arrange the terms of the proposition so that the terms in which a vector field differentiates χ_1 and χ_2 are included in E_r and E'_r, respectively. Thus, we have

$$\imath A_r^* A_r P - \imath P A_r^* A_r = (x D_x)^* C_r^\sharp (x D_x) + (x D_x)^* x C'_r + x C''_r (x D_x) + x^2 C_r^\flat + E_r + E'_r + F_r, \tag{8-13}$$

with

$$\begin{aligned}\sigma_{b,2s}(C_r^\sharp) &= w_r^2\big(\digamma^{-1}\delta^{-1} a |\underline{\zeta}_{n-1}|^{-1}(\hat f^\sharp + \epsilon^{-2}\delta^{-1} f^\sharp)\chi'_0\chi_1\chi_2 + a^2\tilde c_r^\sharp\big),\\ \sigma_{b,2s+1}(C'_r) &= w_r^2\big(\digamma^{-1}\delta^{-1} a(\hat f' + \delta^{-1}\epsilon^{-2} f')\chi'_0\chi_1\chi_2 + a^2 \tilde c'_r\big),\\ \sigma_{b,2s+1}(C''_r) &= w_r^2\big(\digamma^{-1}\delta^{-1} a(\hat f'' + \delta^{-1}\epsilon^{-2} f'')\chi'_0\chi_1\chi_2 + a^2 \tilde c''_r\big),\\ \sigma_{b,2s+2}(C_r) &= w_r^2\big(\digamma^{-1}\delta^{-1}|\underline{\zeta}_{n-1}| a(4\hat h + \hat f^\flat + \delta^{-1}\epsilon^{-2} f^\flat)\chi'_0\chi_1\chi_2 + a^2\tilde c_r^\flat\big),\end{aligned} \tag{8-14}$$

where $f^\sharp$, f', f'' and $f^\flat$ as well as $\hat f^\sharp$, $\hat f'$, $\hat f''$ and $\hat f^\flat$ are all smooth functions on ${}^bT^*X\setminus o$, homogeneous of degree 0 (independent of ϵ and δ), and $\hat h = |\underline{\zeta}_{n-1}|^{-2} h$ is the rehomogenized version of h. Moreover, $f^\sharp$, f', f'' and $f^\flat$ arise from when ω is differentiated in $\chi(\digamma^{-1}(2-\phi/\delta))$, and thus vanish when $\omega = 0$, while $\hat f^\sharp$, $\hat f'$, $\hat f''$ and $\hat f^\flat$ arise when η is differentiated in $\chi(\digamma^{-1}(2-\phi/\delta))$, and comprise all such terms with the exception of those arising from the $\partial_{\underline{\xi}}$ component of $V^\flat|_Y$ (which gives $4\hat h = 4|\zeta_{n-1}|^{-2} h$ on the last line above) and hence are the sums of functions vanishing at $x = 0$ (corresponding to us only specifying the restrictions of the vector fields in (8-2) at Y) and functions vanishing at $\hat{\underline{\xi}} = 0$ (when $|\underline{\zeta}_{n-1}|^{-1}$ in $\eta = -\underline{\xi}|\underline{\zeta}_{n-1}|^{-1}$ is differentiated).[5]

In this formula we think of

$$4\digamma^{-1}\delta^{-1} w_r^2 a |\underline{\zeta}_{n-1}| \hat h \chi'_0\chi_1\chi_2 \tag{8-15}$$

as the main term; note that $\hat h$ is positive near q_0. Compared to this, the terms with a^2 are negligible, for they can all be bounded by

$$c\digamma^{-1}(\digamma^{-1}\delta^{-1} w_r^2 a |\underline{\zeta}_{n-1}|^{-1}\chi'_0\chi_1\chi_2)$$

(see (8-15)), that is, by a small multiple of $\digamma^{-1}\delta^{-1} w_r^2 a |\underline{\zeta}_{n-1}|^{-1}\chi'_0\chi_1\chi_2$ when $\digamma$ is taken large, using that $2 - \phi/\delta \le 4$ on $\operatorname{supp} a$ and

$$\chi_0(\digamma^{-1}t) = (\digamma^{-1}t)^2\chi'_0(\digamma^{-1}t) \le 16\digamma^{-2}\chi'_0(\digamma^{-1}t) \quad \text{for } t \le 4; \tag{8-16}$$

[5] Terms of the latter kind did not occur in [Vasy 2008c] since time-translation invariance was assumed, but it does occur in [Vasy 2008b] and [Vasy 2010a], where the Lorentzian scalar setting is considered.

see the discussion in [Vasy 2008b, Section 6] and following [Vasy 2008c, Equation (6.19)].

The vanishing condition on the $f^\sharp$, f', f'', $f^\flat$ ensures that, on $\operatorname{supp} a$,

$$|f^\sharp|, |f'|, |f''|, |f^\flat| \le C\omega^{1/2} \le 2C\epsilon\delta, \tag{8-17}$$

so the corresponding terms can thus be estimated using $w_r^2 F^{-1}\delta^{-1} a|\underline{\zeta}_{n-1}|^{-1}\chi_0'\chi_1\chi_2$ provided ϵ^{-1} is not too large; that is, there exists $\tilde\epsilon_0 > 0$ such that if $\epsilon > \tilde\epsilon_0$, the terms with $f^\sharp$, f', f'', $f^\flat$ can be treated as error terms.

On the other hand, we have

$$|\hat f^\sharp|, |\hat f'|, |\hat f''|, |\hat f^\flat| \le C|x| + C|\hat{\underline{\xi}}| \le C\omega^{1/2} + C|\hat{\underline{\xi}}| \le 2C\epsilon\delta + C|\hat{\underline{\xi}}|. \tag{8-18}$$

Now, $|\hat{\underline{\xi}}| \le 2|x|$ on $\dot\Sigma$ (for $|\underline{\xi}| = x|\xi| \le 2|x||\underline{\zeta}_{n-1}|$ with U sufficiently small). Therefore we can write $\hat f^\sharp = \hat f^\sharp_\sharp + \hat f^\sharp_\flat$ with $\hat f^\sharp_\flat$ supported away from $\dot\Sigma$ and $\hat f^\sharp_\sharp$ satisfying

$$|\hat f^\sharp_\sharp| \le C|x| + C|\hat{\underline{\xi}}| \le C'|x| \le C'\omega^{1/2} \le 2C'\epsilon\delta; \tag{8-19}$$

we can also obtain a similar decomposition for $\hat f'$, $\hat f''$ and $\hat f^\flat$.

Indeed, using (8-16) it is useful to rewrite (8-14) as

$$\begin{aligned}
\sigma_{\mathrm{b},2s}(C_r^\sharp) &= w_r^2 F^{-1}\delta^{-1} a|\underline{\zeta}_{n-1}|^{-1}(\hat f^\sharp + \epsilon^{-2}\delta^{-1} f^\sharp + F^{-1}\delta\hat c_r^\sharp)\chi_0'\chi_1\chi_2,\\
\sigma_{\mathrm{b},2s+1}(C_r') &= w_r^2\delta^{-1}F^{-1} a(\hat f' + \delta^{-1}\epsilon^{-2} f' + F^{-1}\delta\hat c_r')\chi_0'\chi_1\chi_2,\\
\sigma_{\mathrm{b},2s+1}(C_r'') &= w_r^2\delta^{-1}F^{-1} a(\hat f'' + \delta^{-1}\epsilon^{-2} f'' + F^{-1}\delta\hat c_r'')\chi_0'\chi_1\chi_2,\\
\sigma_{\mathrm{b},2s+2}(C_r^\flat) &= w_r^2\delta^{-1}F^{-1} a|\underline{\zeta}_{n-1}|(4\hat h + \hat f^\flat + \delta^{-1}\epsilon^{-2} f^\flat + F^{-1}\hat c_r^\flat)\chi_0'\chi_1\chi_2,
\end{aligned} \tag{8-20}$$

where

- $f^\sharp$, f', f'' and $f^\flat$ are smooth functions on ${}^{\mathrm{b}}T^*X \setminus o$ that are homogeneous of degree 0 and satisfy (8-17) (and are independent of F, ϵ, δ, r);
- $\hat f^\sharp$, $\hat f'$, $\hat f''$ and $\hat f^\flat$ are smooth functions on ${}^{\mathrm{b}}T^*X \setminus o$, homogeneous of degree 0, with $\hat f^\sharp = \hat f^\sharp_\sharp + \hat f^\sharp_\flat$, where $\hat f^\sharp_\sharp$, $\hat f'_\sharp$, $\hat f''_\sharp$, $\hat f^\flat_\sharp$ satisfy (8-19) (and are independent of F, ϵ, δ, r), while $\hat f^\sharp_\flat$, $\hat f'_\flat$, $\hat f''_\flat$, $\hat f^\flat_\flat$ are supported away from $\dot\Sigma$; and
- $\hat c_r^\sharp$, $\hat c_r'$, $\hat c_r''$ and $\hat c_r^\flat$ are smooth functions on ${}^{\mathrm{b}}T^*X \setminus o$ that are homogeneous of degree 0 and uniformly bounded in ϵ, δ, r, F.

Let

$$b_r = 2w_r|\underline{\zeta}_{n-1}|^{1/2}(F\delta)^{-1/2}(\chi_0\chi_0')^{1/2}\chi_1\chi_2,$$

and let $\tilde B_r \in \Psi_{\mathrm{b}}^{s+1}(X)$ with principal symbol b_r. Then let

$$C \in \Psi_{\mathrm{b}}^0(X) \quad \text{and} \quad \sigma_{\mathrm{b},0}(C) = |\underline{\zeta}_{n-1}|^{-1}h^{1/2}\psi = \hat h^{1/2}\psi,$$

where $\psi \in S^0_{\mathrm{hom}}({}^{\mathrm{b}}T^*X \setminus o)$ is identically 1 on U considered as a subset of ${}^{\mathrm{b}}S^*X$; recall from Remark 8.3 that $\hat h$ is bounded below by a positive quantity here.

If $\tilde{C}_r \in \Psi_b^{2s}(X)$ with principal symbol

$$\sigma_{b,2s}(\tilde{C}_r) = -4w_r^2 F^{-1}\delta^{-1} a|\underline{\zeta}_{n-1}|^{-1}\chi_0'\chi_1\chi_2 = -|\underline{\zeta}_{n-1}|^{-2} b_r^2,$$

then we deduce from (8-13)–(8-20) that[6]

$$\begin{aligned}\imath A_r^* A_r P - \imath P A_r^* A_r \\ = \tilde{B}_r^*\big(C^* x^2 C + x R^\flat x + (xD_x)^*\tilde{R}' x + x\tilde{R}''(xD_x) + (xD_x)^* R^\sharp (xD_x)\big)\tilde{B}_r + R_r'' + E_r + E_r' \qquad (8\text{-}21)\end{aligned}$$

with

$$\begin{gathered} R^\flat \in \Psi_b^0(X), \quad \tilde{R}', \tilde{R}'' \in \Psi_b^{-1}(X), \quad R^\sharp \in \Psi_b^{-2}(X), \\ R_r'' \in L^\infty((0,1); \mathrm{Diff}_0^2\,\Psi_b^{2s-1}(X)), \quad E_r, E_r' \in L^\infty((0,1); \mathrm{Diff}_0^2\,\Psi_b^{2s}(X)), \end{gathered}$$

with $\mathrm{WF}_b'(E) \subset \eta^{-1}((-\infty,-\delta]) \cap U$ and $\mathrm{WF}_b'(E') \cap \dot{\Sigma} = \varnothing$, and with $r^\flat = \sigma_{b,0}(R^\flat)$, $\tilde{r}' = \sigma_{b,-1}(\tilde{R}')$, $\tilde{r}'' = \sigma_{b,-1}(\tilde{R}'')$, $r^\sharp \in \sigma_{b,-2}(R^\sharp)$, and

$$\begin{gathered} |r^\flat| \le C_2(\delta\epsilon + \epsilon^{-1} + \delta F^{-1}), \quad |\underline{\zeta}_{n-1}\tilde{r}'| \le C_2(\delta\epsilon + \epsilon^{-1} + \delta F^{-1}), \\ |\underline{\zeta}_{n-1}\tilde{r}''| \le C_2(\delta\epsilon + \epsilon^{-1} + \delta F^{-1}), \quad |\underline{\zeta}_{n-1}^2 r^\sharp| \le C_2(\delta\epsilon + \epsilon^{-1} + \delta F^{-1}). \end{gathered}$$

This is almost completely analogous to [Vasy 2008c, Equation (6.18)] with the understanding that each term therein inside the parentheses attains an additional factor of x^2 (corresponding to $\Box$ being in $\mathrm{Diff}_0^2(X)$ rather than $\mathrm{Diff}^2(X)$), which we partially include in xD_x (vs. D_x). The only difference is the presence of the δF^{-1} term, which however is treated like the $\epsilon\delta$ term for F sufficiently large; hence the rest of the proof proceeds very similarly to that paper. We go through this argument to show the role that λ and the Poincaré inequality play, and in particular how the restrictions on λ arise.

Having calculated the commutator, we proceed to estimate the "error terms" $R^\flat$, $\tilde{R}'$, $\tilde{R}''$ and $R^\sharp$ as operators. We start with $R^\flat$. By the standard square root construction to prove the boundedness of PsDOs on L^2, see e.g. the discussion after [Vasy 2008c, Remark 2.1], there exists $R_\flat^\flat \in \Psi_b^{-1}(X)$ such that

$$\|R^\flat v\| \le 2\sup|r^\flat|\,\|v\| + \|R_\flat^\flat v\| \quad \text{for all } v \in L^2(X).$$

Here $\|\cdot\|$ is the $L^2(X)$ norm, as usual. Thus, we can estimate, for any $\gamma > 0$,

$$\begin{aligned} |\langle R^\flat v, v\rangle| &\le \|R^\flat v\|\,\|v\| \le 2\sup|r^\flat|\,\|v\|^2 + \|R_\flat^\flat v\|\,\|v\| \\ &\le 2C_2(\delta\epsilon + \epsilon^{-1} + \delta F^{-1})\|v\|^2 + \gamma^{-1}\|R_\flat^\flat v\|^2 + \gamma\|v\|^2. \end{aligned}$$

Now we turn to $\tilde{R}'$. Let $T \in \Psi_b^{-1}(X)$ be elliptic (which we use to shift the orders of PsDOs at our convenience), with symbol $|\underline{\zeta}_{n-1}|^{-1}$ on $\operatorname{supp} a$, and with $T^- \in \Psi_b^1(X)$ a parametrix, so $T^-T = \mathrm{Id} + F$

[6] The $f_\sharp^\sharp$ terms are included in $R^\sharp$, while the $f_\flat^\sharp$ terms are included in E', and similarly for the other analogous terms in f', f'', $f^\flat$. Moreover, in view of Lemma 5.4, we can freely rearrange factors, e.g., writing $C^* x^2 C$ as xC^*Cx if we wish, with the exception of commuting powers of x with xD_x or $(xD_x)^*$ since we need to regard the latter as elements of $\mathrm{Diff}_0^1(X)$ rather than $\mathrm{Diff}_b^1(X)$. Indeed, the difference between rearrangements has lower b-order than the product, in this case being in $x^2\Psi_b^{-1}(X)$, which in view of Lemma 5.5, at the cost of dropping powers of x, can be translated into a gain in 0-order, that is, $x^2\Psi_b^{-1}(X) \subset \mathrm{Diff}_0^2\,\Psi_b^{-3}(X)$, with the result that these terms can be moved to the "error term" $R'' \in L^\infty((0,1); \mathrm{Diff}_0^2\,\Psi_b^{2s-1}(X))$.

with $F \in \Psi_b^{-\infty}(X)$. Then there exists $\tilde{R}'_\flat \in \Psi_b^{-1}(X)$ such that

$$\begin{aligned}\|(\tilde{R}')^* w\| = \|(\tilde{R}')^*(T^- T - F)w\| &\le \|((\tilde{R}')^* T^-)(Tw)\| + \|(\tilde{R}')^* F w\| \\ &\le 2C_2(\delta\epsilon + \epsilon^{-1} + \delta F^{-1})\|Tw\| + \|\tilde{R}'_\flat T w\| + \|(\tilde{R}')^* F w\|\end{aligned}$$

for all w with $Tw \in L^2(X)$, and similarly, there exists $\tilde{R}''_\flat \in \Psi_b^{-1}(X)$ such that

$$\|\tilde{R}'' w\| \le 2C_2(\delta\epsilon + \epsilon^{-1} + \delta F^{-1})\|Tw\| + \|\tilde{R}''_\flat T w\| + \|\tilde{R}'' F w\|.$$

Finally, there exists $R^\sharp_\flat \in \Psi_b^{-1}(X)$ such that

$$\|(T^-)^* R^\sharp w\| \le 2C_2(\delta\epsilon + \epsilon^{-1} + \delta F^{-1})\|Tw\| + \|R^\sharp_\flat T w\| + \|(T^-)^* R^\sharp F w\|$$

for all w with $Tw \in L^2(X)$. Thus,

$$\begin{aligned}|\langle xv, (\tilde{R}')^*(xD_x)v\rangle| &\le 2C_2(\delta\epsilon + \epsilon^{-1} + \delta F^{-1})\|TxD_x v\|\,\|xv\| \\ &\qquad + 2\gamma\|xv\|^2 + \gamma^{-1}\|\tilde{R}'_\flat TxD_x v\|^2 + \gamma^{-1}\|F' xD_x v\|^2, \\ |\langle \tilde{R}'' xD_x v, xv\rangle| &\le 2C_2(\delta\epsilon + \epsilon^{-1} + \delta F^{-1})\|TxD_x v\|\,\|xv\| \\ &\qquad + 2\gamma\|xv\|^2 + \gamma^{-1}\|\tilde{R}''_\flat TxD_x v\|^2 + \gamma^{-1}\|F'' xD_x v\|^2,\end{aligned}$$

and, writing $xD_x v = T^- T(xD_x v) - F(xD_x v)$ in the right factor, and taking the adjoint of T^-,

$$\begin{aligned}|\langle R^\sharp xD_x v, xD_x v\rangle| &\le 2C_2(\delta\epsilon + \epsilon^{-1} + \delta F^{-1})\|T(xD_x)v\|\,\|T(xD_x)v\| + 2\gamma\|T(xD_x)v\|^2 \\ &\quad + \gamma^{-1}\|R^\sharp_\flat T(xD_x)v\|^2 + \gamma^{-1}\|F(xD_x)v\|^2 + \|R^\sharp(xD_x)v\|\,\|F^\sharp(xD_x v)\|,\end{aligned}$$

with $F', F'', F^\sharp \in \Psi_b^{-\infty}(X)$.

Now, by (8-21),

$$\begin{aligned}\langle \imath[A_r^* A_r, P]u, u\rangle = \|Cx\tilde{B}_r u\|^2 &+ \langle R^\flat x\tilde{B}_r u, x\tilde{B}_r u\rangle + \langle \tilde{R}'' xD_x \tilde{B}_r u, x\tilde{B}_r u\rangle \\ &+ \langle x\tilde{B}_r u, (\tilde{R}')^* xD_x \tilde{B}_r u\rangle + \langle R^\sharp xD_x \tilde{B}_r u, xD_x \tilde{B}_r u\rangle \\ &+ \langle R''_r u, u\rangle + \langle (E_r + E'_r)u, u\rangle \qquad (8\text{-}22)\end{aligned}$$

On the other hand, this commutator can be expressed as in (8-12), so

$$\begin{aligned}&\langle \imath A_r^* A_r P u, u\rangle - \langle \imath A_r^* A_r u, Pu\rangle \\ &\quad = -2\,\mathrm{Im}\,\lambda\|A_r u\|^2 + \|Cx\tilde{B}_r u\|^2 + \langle R^\flat x\tilde{B}_r u, x\tilde{B}_r u\rangle + \langle \tilde{R}'' xD_x \tilde{B}_r u, x\tilde{B}_r u\rangle \\ &\qquad + \langle x\tilde{B}_r u, (\tilde{R}')^* xD_x \tilde{B}_r u\rangle + \langle R^\sharp xD_x \tilde{B}_r u, xD_x \tilde{B}_r u\rangle + \langle R''_r u, u\rangle + \langle (E_r + E'_r)u, u\rangle, \qquad (8\text{-}23)\end{aligned}$$

so the signs of the first two terms agree if $\mathrm{Im}\,\lambda < 0$, and the $\mathrm{Im}\,\lambda$ term vanishes if λ is real.

Assume for the moment that $\mathrm{WF}_b^{-1,s+3/2}(Pu) \cap U = \varnothing$ — this is certainly the case in our setup if $q_0 \notin \mathrm{WF}_b^{-1,\infty}(Pu)$, but this assumption is a little stronger than $q_0 \notin \mathrm{WF}_b^{-1,s+1}(Pu)$, which is what we need to assume for the second paragraph in the statement of the proposition. We deal with the weakened hypothesis $q_0 \notin \mathrm{WF}_b^{-1,s+1}(Pu)$ at the end of the proof. Returning to (8-23), the utility of the commutator calculation is that we have good information about Pu (this is where we use that we have a microlocal solution of the PDE!). Namely, we estimate the left hand side as

$$\begin{aligned}|\langle A_r Pu, A_r u\rangle| &\le |\langle (T^-)^* A_r Pu, TA_r u\rangle| + |\langle A_r Pu, FA_r u\rangle| \\ &\le \|(T^-)^* A_r Pu\|_{H_0^{-1}(X)}\|TA_r u\|_{H_0^1(X)} + \|A_r Pu\|_{H_0^{-1}(X)}\|FA_r u\|_{H_0^1(X)}.\end{aligned} \tag{8-24}$$

Since $(T^-)^* A_r$ is uniformly bounded in $\Psi_{\mathrm{bc}}^{s+3/2}(X)$ and TA_r is uniformly bounded in $\Psi_{\mathrm{bc}}^{s-1/2}(X)$, both with $\mathrm{WF}_{\mathrm{b}}'$ in U, with $\mathrm{WF}_{\mathrm{b}}^{-1,s+3/2}(Pu)$ and $\mathrm{WF}_{\mathrm{b}}^{1,s-1/2}(u)$, respectively, disjoint from them, we deduce (using Lemma 5.14 and its H_0^{-1} analogue) that $|\langle (T^-)^* A_r Pu, TA_r u\rangle|$ is uniformly bounded. Similarly, taking into account that FA_r is uniformly bounded in $\Psi_{\mathrm{b}}^{-\infty}(X)$, we see that $|\langle A_r Pu, FA_r u\rangle|$ is also uniformly bounded, so $|\langle A_r Pu, A_r u\rangle|$ is uniformly bounded for $r \in (0, 1]$.

Thus,

$$\begin{aligned}&\|Cx\tilde B_r u\|^2 - \operatorname{Im}\lambda\|A_r u\|^2 \\ &\le 2|\langle A_r Pu, A_r u\rangle| + |\langle (E_r+E_r')u, u\rangle| + \big(2C_2(\delta\epsilon+\epsilon^{-1}+\delta\digamma^{-1})+\gamma\big)\|x\tilde B_r u\|^2 + \gamma^{-1}\|R_\flat^\flat x\tilde B_r u\|^2 \\ &\quad + 4C_2(\delta\epsilon+\epsilon^{-1}+\delta\digamma^{-1})\|x\tilde B_r u\|\|T(xD_x)\tilde B_r u\| + \gamma^{-1}\|\tilde R_\flat' T(xD_x)\tilde B_r u\|^2 + \gamma^{-1}\|\tilde R_\flat'' T(xD_x)\tilde B_r u\|^2 \\ &\quad + 4\gamma\|x\tilde B_r u\|^2 + \big(2C_2(\delta\epsilon+\epsilon^{-1}+\delta\digamma^{-1})+2\gamma\big)\|T(xD_x)\tilde B_r u\|^2 \\ &\qquad + \gamma^{-1}\|R_\flat^\sharp T(xD_x)\tilde B_r u\|^2 + \|R^\sharp(xD_x)\tilde B_r u\|\,\|F(xD_x)\tilde B_r u\| \\ &\qquad\quad + \gamma^{-1}\|F(xD_x)\tilde B_r u\|^2 + \gamma^{-1}\|F'(xD_x)\tilde B_r u\|^2 + \gamma^{-1}\|F''(xD_x)\tilde B_r u\|^2.\end{aligned} \tag{8-25}$$

All terms but the ones involving C_2 or γ (not γ^{-1}) remain bounded as $r \to 0$. The C_2 and γ terms can be estimated by writing $T(xD_x) = (xD_x)T' + T''$ for some $T', T'' \in \Psi_{\mathrm{b}}^{-1}(X)$, and using Lemma 7.3 and the Poincaré lemma where necessary. Namely, we use either $\operatorname{Im}\lambda \ne 0$ or $\lambda < (n-1)^2/4$ to control $xD_x L\tilde B_r u$ and $L\tilde B_r u$ in $L^2(X)$ in terms of $\|x\tilde B_r u\|_{L^2}$ where $L \in \Psi_{\mathrm{b}}^{-1}(X)$; this is possible by factoring $D_{y_{n-1}}$ (which is elliptic on $\mathrm{WF}'(\tilde B_r)$) out of $\tilde B_r$ modulo an error $\tilde F_r$ bounded in $\Psi_{\mathrm{bc}}^{s}(X)$, which in turn can be incorporated into the "error" given by the right hand side of Lemma 7.3. Thus, there exists $C_3 > 0$, $G \in \Psi_{\mathrm{b}}^{s-1/2}(X)$ and $\tilde G \in \Psi_{\mathrm{b}}^{s+1/2}(X)$ as in Lemma 7.3 such that

$$\|xD_x L\tilde B_r u\|^2 + \|L\tilde B_r u\| \le C_3\big(\|x\tilde B_r u\|^2 + \|u\|^2_{H^{1,k}_{0,\mathrm{b,loc}}(X)} + \|Gu\|^2_{H^1_0(X)} + \|Pu\|^2_{H^{-1,k}_{0,\mathrm{b,loc}}(X)} + \|\tilde G Pu\|^2_{H_0^{-1}(X)}\big).$$

We further estimate $\|x\tilde B_r u\|$ in terms of $\|Cx\tilde B_r u\|$ and $\|u\|_{H^1_{0,\mathrm{loc}}(X)}$ using that C is elliptic on $\mathrm{WF}_{\mathrm{b}}'(B)$ and Lemma 5.14. We conclude, using $\operatorname{Im}\lambda \le 0$, taking ϵ sufficiently large, then γ and δ_0 sufficiently small, and finally $\digamma$ sufficiently large, that there exist $\gamma > 0$, $\epsilon > 0$, $\delta_0 > 0$ and $C_4 > 0$ and $C_5 > 0$ such that for $\delta \in (0, \delta_0)$,

$$\begin{aligned}C_4\|x\tilde B_r u\|^2 &\le 2|\langle A_r Pu, A_r u\rangle| + |\langle (E_r+E_r')u, u\rangle| \\ &\quad + C_5\big(\|Gu\|^2_{H^1_0(X)} + \|\tilde G Pu\|^2_{H_0^{-1}(X)}\big) + C_5\big(\|u\|_{H^{1,k}_{0,\mathrm{b,loc}}(X)} + \|Pu\|_{H^{-1,k}_{0,\mathrm{b,loc}}(X)}\big).\end{aligned}$$

Letting $r \to 0$ now keeps the right hand side bounded, proving that $\|x\tilde B_r u\|$ is uniformly bounded as $r \to 0$; hence $x\tilde B_0 u \in L^2(X)$ (see the proof of Proposition 7.7). In view of Lemma 7.3 and the Poincaré inequality (as in the proof of Proposition 7.7), this proves that $q_0 \notin \mathrm{WF}_{\mathrm{b}}^{1,s}(u)$, and hence proves the first statement of the proposition.

In fact, recalling that we needed $q_0 \notin \mathrm{WF}_{\mathrm{b}}^{-1,s+3/2}(Pu)$ for the uniform boundedness in (8-24), this proves a slightly weaker version of the second statement of the proposition with $\mathrm{WF}_{\mathrm{b}}^{-1,s+1}(Pu)$ replaced by $\mathrm{WF}_{\mathrm{b}}^{-1,s+3/2}(Pu)$. For the more precise statement we modify (8-24) — this is the only term in (8-25) that needs modification to prove the optimal statement. Let $\tilde{T} \in \Psi_{\mathrm{b}}^{-1/2}(X)$ be elliptic, $\tilde{T}^- \in \Psi_{\mathrm{b}}^{1/2}(X)$ a parametrix, with $\tilde{F} = \tilde{T}^-\tilde{T} - \mathrm{Id} \in \Psi_{\mathrm{b}}^{-\infty}(X)$. Then, similarly to (8-24), we have for any $\gamma > 0$,

$$\begin{aligned}|\langle A_r Pu, A_r u\rangle| &\le |\langle (\tilde{T}^-)^* A_r Pu, \tilde{T} A_r u\rangle| + |\langle A_r Pu, \tilde{F} A_r u\rangle| \\ &\le \gamma^{-1}\|(\tilde{T}^-)^* A_r Pu\|^2_{H_0^{-1}(X)} + \gamma\|\tilde{T} A_r u\|^2_{H^1(X)} + \|A_r Pu\|_{H^{-1}(X)}\|\tilde{F} A_r u\|_{H_0^1(X)}.\end{aligned} \tag{8-26}$$

The last term on the right hand side can be estimated as before. As $(\tilde{T}^-)^* A_r$ is bounded in $\Psi_{\mathrm{bc}}^{s+1}(X)$ with $\mathrm{WF}_{\mathrm{b}}'$ disjoint from U, we see that $\|(\tilde{T}^-)^* A_r Pu\|_{H_0^{-1}(X)}$ is uniformly bounded. Moreover, $\|\tilde{T} A \Lambda_r u\|^2_{H_0^1(X)}$ can be estimated, using Lemma 7.3 and the Poincaré inequality, by $\|x D_{y_{n-1}} \tilde{T} A \Lambda_r u\|^2_{L^2(X)}$ modulo terms that are uniformly bounded as $r \to 0$. The principal symbol of $D_{y_{n-1}} \tilde{T} A$ is $\underline{\zeta}_{n-1}\sigma_{\mathrm{b},-1/2}(\tilde{T})a$, with $a = \chi_0\chi_1\chi_2$, where χ_0 stands for $\chi_0(A_0^{-1}(2-\phi/\delta))$, etc., so we can write

$$|\underline{\zeta}_{n-1}|^{1/2} a = |\underline{\zeta}_{n-1}|^{1/2}\chi_0\chi_1\chi_2 = A_0^{-1}(2-\phi/\delta)|\underline{\zeta}_{n-1}|^{1/2}(\chi_0\chi_0')^{1/2}\chi_1\chi_2 = F^{-1/2}\delta^{1/2}(2-\phi/\delta)\tilde{b},$$

where we used that

$$\chi_0'(F^{-1}(2-\phi/\delta)) = F^2(2-\phi/\delta)^{-2}\chi_0(F^{-1}(2-\phi/\delta))$$

when $2-\phi/\delta > 0$, while a and $\tilde{b}$ vanish otherwise. Correspondingly, using that $|\underline{\zeta}_{n-1}|^{1/2}\sigma_{\mathrm{b},-1/2}(\tilde{T})$ is $\mathscr{C}^\infty$ and homogeneous degree zero near the support of a in ${}^{\mathrm{b}}T^*X \setminus o$, we can write $D_{y_{n-1}}\tilde{T} A = G\tilde{B} + F$, with $G \in \Psi_{\mathrm{b}}^0(X)$ and $F \in \Psi_{\mathrm{b}}^{-1/2}(X)$. Thus, modulo terms that are bounded as $r \to 0$, we can estimate $\|x D_{y_{n-1}}\tilde{T} A\Lambda_r u\|^2$ (and hence $\|\tilde{T} A \Lambda_r u\|^2_{H_0^1(X)}$) from above by $C_6\|x\tilde{B}_r u\|^2$. Therefore, modulo terms that are bounded as $r \to 0$, for $\gamma > 0$ sufficiently small, $\gamma\|\tilde{T} A_r u\|^2_{H_0^1(X)}$ can be absorbed into $\|Cx\tilde{B}_r u\|^2$. As the treatment of the other terms on the right hand side of (8-25) requires no change, we deduce as above that $x\tilde{B}_0 u \in L^2(X)$, which (in view of Lemma 7.3) proves that $q_0 \notin \mathrm{WF}_{\mathrm{b}}^{1,s}(u)$, completing the proof of the iterative step.

We need to make one more remark to prove the proposition for $\mathrm{WF}_{\mathrm{b}}^{1,\infty}(u)$; namely we need to show that the neighborhoods of q_0 that are disjoint from $\mathrm{WF}_{\mathrm{b}}^{1,s}(u)$ do not shrink uncontrollably to $\{q_0\}$ as $s \to \infty$. This argument parallels the last paragraph of the proof of [Hörmander 1985, Proposition 24.5.1]. In fact, note that above we have proved that the elliptic set of $\tilde{B} = \tilde{B}_s$ is disjoint from $\mathrm{WF}_{\mathrm{b}}^{1,s}(u)$. In the next step, when we are proving $q_0 \notin \mathrm{WF}_{\mathrm{b}}^{1,s+1/2}(u)$, we decrease $\delta > 0$ slightly (by an arbitrary small amount), thus decreasing the support of $a = a_{s+1/2}$ in (8-8), to make sure that $\operatorname{supp} a_{s+1/2}$ is a subset of the elliptic set of the union of $\tilde{B}_s$ with the region $\eta < 0$, and hence that $\mathrm{WF}_{\mathrm{b}}^{1,s}(u) \cap \operatorname{supp} a_{s+1/2} = \varnothing$. Each iterative step thus shrinks the elliptic set of $\tilde{B}_s$ by an arbitrarily small amount, which allows us to conclude that q_0 has a neighborhood U' such that $\mathrm{WF}_{\mathrm{b}}^{1,s}(u) \cap U' = \varnothing$ for all s. This proves that

$q_0 \notin \mathrm{WF}_b^{1,\infty}(u)$, and indeed that $\mathrm{WF}_b^{1,\infty}(u) \cap U' = \varnothing$, for if $A \in \Psi_b^m(X)$ with $\mathrm{WF}_b'(A) \subset U'$, then $Au \in H_0^1(X)$ by Lemma 5.10 and Corollary 5.12. □

Before turning to tangential propagation we need a technical lemma, which roughly states that when applied to solutions of $Pu = 0$ with $u \in H_0^1(X)$, the operators xD_x and Id are not merely bounded by $xD_{y_{n-1}}$ microlocally, but are small compared to it, provided that $\lambda \in \mathbb{C} \setminus [(n-1)^2/4, \infty)$. This result is the analogue of [Vasy 2008c, Lemma 7.1], and is proved as there, with the only difference being that the term $\langle \lambda A_r u, A_r u\rangle$ cannot be dropped; instead it is treated just as in Proposition 7.7 above. Below a δ-neighborhood refers to a δ-neighborhood with respect to the metric associated to any Riemannian metric on the manifold ${}^bT^*X$, and we identify ${}^bS^*X$ as the unit ball bundle with respect to some fiber metric on ${}^bT^*X$.

Lemma 8.4 (see [Vasy 2008c, Lemma 7.1]). *Suppose that $P = \Box_g + \lambda$, with $\lambda \in \mathbb{C} \setminus [(n-1)^2/4, \infty)$. Suppose $u \in H^{1,k}_{0,b,\mathrm{loc}}(X)$, and suppose that we are given $K \subset {}^bS^*X$ compact satisfying*

$$K \subset \mathcal{G} \cap T^*Y \setminus \mathrm{WF}_b^{-1,s+1/2}(Pu).$$

Then there exist $\delta_0 > 0$ and $C_0 > 0$ with the following property. Let $\delta < \delta_0$. Let $U \subset {}^bS^*X$ be open in a δ-neighborhood of K, and let $\mathcal{A} = \{A_r : r \in (0,1]\}$ be a bounded family of PsDOs in $\Psi^s_{bc}(X)$ with $\mathrm{WF}_b'(\mathcal{A}) \subset U$, and with $A_r \in \Psi_b^{s-1}(X)$ for $r \in (0,1]$.

Then there exist $G \in \Psi_b^{s-1/2}(X)$ and $\tilde{G} \in \Psi_b^{s+1/2}(X)$ with $\mathrm{WF}_b'(G), \mathrm{WF}_b'(\tilde{G}) \subset U$ and $\tilde{C}_0 = \tilde{C}_0(\delta) > 0$ such that for all $r > 0$,

$$\begin{aligned} &\|xD_xA_ru\|^2 + \|A_ru\|^2 \\ &\quad \le C_0\delta\|xD_{y_{n-1}}A_ru\|^2 + \tilde{C}_0\big(\|u\|^2_{H^{1,k}_{0,b,\mathrm{loc}}(X)} + \|Gu\|^2_{H_0^1(X)} + \|Pu\|^2_{H^{-1,k}_{0,b,\mathrm{loc}}(X)} + \|\tilde{G}Pu\|^2_{H_0^{-1}(X)}\big). \end{aligned} \tag{8-27}$$

The meaning of $\|u\|_{H^{1,k}_{0,b,\mathrm{loc}}(X)}$ and $\|Pu\|^2_{H^{-1,k}_{0,b,\mathrm{loc}}(X)}$ is stated in Remark 7.2.

Remark 8.5. Since K is compact, this is essentially a local result. In particular, we may assume that K is a subset of ${}^bT^*X$ over a suitable local coordinate patch. Moreover, we may assume that $\delta_0 > 0$ is sufficiently small so that $D_{y_{n-1}}$ is elliptic on U.

Proof. By Lemma 7.3 applied with K replaced by $\mathrm{WF}_b'(\mathcal{A})$ in the hypothesis (note that the latter is compact), we already know that

$$|\langle dA_ru, dA_ru\rangle_G + \lambda\|A_ru\|^2| \le C_0'\big(\|u\|^2_{H^{1,k}_{0,b,\mathrm{loc}}(X)} + \|Gu\|^2_{H_0^1(X)} + \|Pu\|^2_{H^{-1,k}_{0,b,\mathrm{loc}}(X)} + \|\tilde{G}Pu\|^2_{H_0^{-1}(X)}\big). \tag{8-28}$$

for some $C_0' > 0$ and for some G and $\tilde{G}$ as in the statement of the lemma. Freezing the coefficients at Y, as in the proof of Proposition 7.7 — see [Vasy 2008c, Lemma 7.1] for details — we deduce that

$$\begin{aligned} &\big|\|xD_xA_ru\|^2 - \lambda\|A_ru\|^2\big| \\ &\qquad \le \int_X \big(B_{ij}(0,y)(xD_{y_i})A_ru\,\overline{(xD_{y_j})A_ru}\big)|dg| + C_1\delta\|xD_{y_{n-1}}A_ru\|^2 \\ &\qquad\qquad + C_0''\big(\|u\|^2_{H^{1,k}_{0,b,\mathrm{loc}}(X)} + \|Gu\|^2_{H_0^1(X)} + \|Pu\|^2_{H^{-1,k}_{0,b,\mathrm{loc}}(X)} + \|\tilde{G}Pu\|^2_{H^{-1}(X)}\big). \end{aligned} \tag{8-29}$$

Now, one can show that

$$\Big|\int_X \Big(\sum (D_{y_i}^* B_{ij}(0,y) D_{y_j}) x A_r u \, \overline{x A_r u}\Big) \, |dg|\Big| \\ \le C_2 \delta \|D_{y_{n-1}} A_r u\|^2 + \tilde{C}_2(\delta)\big(\|u\|^2_{H^{1,k}_{0,\mathrm{b,loc}}(X)} + \|Gu\|^2_{H^1_0(X)}\big) \quad (8\text{-}30)$$

precisely as in the proof of [Vasy 2008c, Lemma 7.1]. Equations (8-29)–(8-30) imply (8-27) with the left side replaced by $|\|xD_x A_r u\|^2 - \lambda \|A_r u\|^2|$. If $\operatorname{Im}\lambda \neq 0$, we get the desired bound for $\|A_r u\|^2$ by taking the imaginary part of $\|xD_x A_r u\|^2 - \lambda \|A_r u\|^2$; hence taking the real part gives the desired bound for $\|xD_x A_r u\|^2$ as well. If $\operatorname{Im}\lambda = 0$ but $\lambda < (n-1)^2/4$, we finish the proof using the Poincaré inequality; see the proof of Proposition 7.7. □

We finally state the tangential, or glancing, propagation result.

Proposition 8.6 (tangential, or glancing, propagation). *Suppose $P = \Box_g + \lambda$ with $\lambda \in \mathbb{C} \setminus [(n-1)^2/4, \infty)$. Let U_0 be a coordinate chart in X, and let U be open with $\overline{U} \subset U_0$. Let $u \in H^{1,k}_{0,\mathrm{b,loc}}(X)$ for some $k \le 0$, and let $\tilde{\pi} : T^*X \to T^*Y$ be the coordinate projection*

$$\tilde{\pi} : (x, y, \xi, \zeta) \mapsto (y, \zeta).$$

Given $K \subset {}^{\mathrm{b}}S^*_U X$ compact with

$$K \subset (\mathcal{G} \cap {}^{\mathrm{b}}T^*_Y X) \setminus \mathrm{WF}_{\mathrm{b}}^{-1,\infty}(f), \quad \textit{where } f = Pu, \quad (8\text{-}31)$$

there exist constants $C_0 > 0$ and $\delta_0 > 0$ such that the following holds. If $\operatorname{Im}\lambda \le 0$, $q_0 = (y_0, \underline{\zeta}_0) \in K$, $\alpha_0 = \hat{\pi}^{-1}(q_0)$ and $W_0 = \tilde{\pi}_*|_{\alpha_0} \mathsf{H}_p$ considered as a constant vector field in local coordinates, and for some $0 < \delta < \delta_0$, $C_0\delta \le \epsilon < 1$ and for all $\alpha = (x, y, \xi, \zeta) \in \Sigma$, there holds

$$\pi(\alpha) \notin \mathrm{WF}_{\mathrm{b}}^{1,\infty}(u) \quad \textit{if} \quad \alpha \in T^*X \;\; \textit{and} \;\; |\tilde{\pi}(\alpha - (\alpha_0 - \delta W_0))| \le \epsilon\delta \;\; \textit{and} \;\; |x(\alpha)| \le \epsilon\delta, \quad (8\text{-}32)$$

then $q_0 \notin \mathrm{WF}_{\mathrm{b}}^{1,\infty}(u)$.

In addition, $\mathrm{WF}_{\mathrm{b}}^{-1,\infty}(f)$ may be replaced by $\mathrm{WF}_{\mathrm{b}}^{-1,s+1}(f)$, and $\mathrm{WF}_{\mathrm{b}}^{1,\infty}(u)$ may be replaced by $\mathrm{WF}_{\mathrm{b}}^{1,s}(u)$, $s \in \mathbb{R}$.

Remark 8.7. Just like Proposition 8.2, this result gives regularity propagation in the *forward* direction along W_0, that is, to conclude regularity at q_0, one needs to know regularity in the *backward* W_0-direction from q_0.

One can again change the direction of propagation, that is, replace δ by $-\delta$ in $\alpha - (\alpha_0 - \delta W_0)$, provided one also changes the sign of $\operatorname{Im}\lambda$ to $\operatorname{Im}\lambda \ge 0$. In particular, if $\operatorname{Im}\lambda = 0$, one obtains propagation estimates in both the forward and backward directions.

Proof. The proof follows closely that of [Vasy 2008c, Proposition 7.3], which is corrected at a point in [Vasy 2008a], so we merely point out the main steps. Again, one uses a commutant $A \in \Psi^0_{\mathrm{b}}(X)$ and weights $\Lambda_r \in \Psi^0_{\mathrm{b}}(X)$ for $r \in (0,1)$, uniformly bounded in $\Psi^{s+1/2}_{\mathrm{bc}}(X)$, with $A_r = A\Lambda_r$, in order to obtain the propagation of $\mathrm{WF}_{\mathrm{b}}^{1,s}(u)$ with the notation of that paper, whose analogue is $\mathrm{WF}_{\mathrm{b}}^{1,s}(u)$ here (the difference is the space relative to which one obtains b-regularity: it is $H^1(X)$ in the previous paper, but the zero-Sobolev space $H^1_0(X)$ here). One can use *exactly the same* commutants as in [Vasy 2008c], with a small correction given in [Vasy 2008a]. Then Proposition 8.1 lets one calculate $\imath[A_r^* A_r, P]$ to obtain a completely analogous expression to the formulas below [Vasy 2008c, Equation (7.16)], as corrected. The rest of the argument is completely analogous as well. Again, we refer the reader to [Vasy 2010a] because the commutator calculation is written up exactly as above in Proposition 8.1 (see [Vasy 2010a, Proposition 3.10]) and it is used subsequently in 6.1 there the same way it needs to be used here — any modifications are analogous to those in Proposition 8.2 and arise due to the nonnegligible nature of λ.

Again, we first construct the symbol a of our commutator following the (corrected) proof [Vasy 2008c, Proposition 7.3]. Note that (with $\tilde{p} = x^{-2}\sigma_{\mathrm{b},2}(\tilde{P}) = h$)

$$W_0(q_0) = \mathsf{H}_{\tilde{p}}(q_0),$$

and let

$$W = |\underline{\zeta}_{n-1}|^{-1} W_0,$$

so W is homogeneous of degree zero (with respect to the $\mathbb{R}^+$-action on the fibers of $T^*Y \setminus o$). We use

$$\tilde{\eta} = (\operatorname{sgn}(\underline{\zeta}_{n-1})_0)(y_{n-1} - (y_{n-1})_0)$$

now to measure propagation, since $\underline{\zeta}_{n-1}^{-1}\mathsf{H}_{\tilde{p}}(y_{n-1}) = 2 > 0$ at q_0 by (6-7), so $\mathsf{H}_{\tilde{p}}\tilde{\eta}$ is $2|\underline{\zeta}_{n-1}| > 0$ at q_0. Note that $\tilde{\eta}$ is thus increasing along GBB of $\hat{g}$.

First, we require

$$\rho_1 = \tilde{p}(y, \hat{\underline{\zeta}}) = |\underline{\zeta}_{n-1}|^{-2}\tilde{p}(y, \underline{\zeta});$$

note that $d\rho_1 \neq 0$ at q_0 for $\underline{\zeta} \neq 0$ there, but $\mathsf{H}_{\tilde{p}}\tilde{p} \equiv 0$, so

$$W\rho_1(q_0) = 0.$$

Next, $\dim T^*Y = 2n-2$ since $\dim Y = n-1$; hence $\dim S^*Y = 2n-3$. With a slight abuse of notation, we also regard q_0 as a point in S^*Y — recall that $S^*Y = (T^*Y \setminus o)/\mathbb{R}^+$. We can also regard W as a vector field on S^*Y in view of its homogeneity. Since W does not vanish as a vector in $T_{q_0}S^*Y$ in view of $W\tilde{\eta}(q_0) \neq 0$ since $\tilde{\eta}$ is homogeneous degree zero and hence a function on S^*Y, the kernel of W in $T^*_{q_0}S^*Y$ has dimension $2n-4$. Thus there exist homogeneous degree zero functions ρ_j for $j = 2, \dots, 2n-4$ on T^*Y (and hence functions on S^*Y) such that

$$\begin{aligned} \rho_j(q_0) &= 0 \quad \text{for } j = 2, \dots, 2n-4,\\ W\rho_j(q_0) &= 0 \quad \text{for } j = 2, \dots, 2n-4,\\ d\rho_j(q_0) &\quad\;\;\, \text{for } j = 1, \dots, 2n-4 \text{ are linearly independent at } q_0. \end{aligned} \tag{8-33}$$

By dimensional considerations, the $d\rho_j(q_0)$ for $j = 1, \dots, 2n-4$, together with $d\tilde{\eta}$, span the cotangent space of S^*Y at q_0, that is, of the quotient of T^*Y by the $\mathbb{R}^+$-action, so the ρ_j, together with $\tilde{\eta}$, can be used as local coordinates on a chart $\tilde{\mathcal{U}}_0 \subset S^*Y$ near q_0. We also let $\tilde{\mathcal{U}}$ be a neighborhood of q_0 in ${}^{\mathrm{b}}S^*X$ such that ρ_j, together with $\tilde{\eta}$, x and $\hat{\underline{\xi}}$, are local coordinates on $\tilde{\mathcal{U}}$; this holds if $\tilde{\mathcal{U}}_0$ is identified with a subset of $\mathcal{G} \cap {}^{\mathrm{b}}S^*_Y X$ and $\tilde{\mathcal{U}}$ is a product neighborhood of this in ${}^{\mathrm{b}}S^*X$ in terms of the coordinates (6-1).

Note that since $\hat{\underline{\xi}} = 0$ on $\dot\Sigma \cap {}^{\mathrm{b}}S^*_Y X$, for points q in $\dot\Sigma$ one can ensure that $\hat{\underline{\xi}}$ is small by ensuring that $\tilde\pi(q)$ is close to q_0 and $x(q)$ is small; see the discussion around (8-5) and after (8-7). By reducing $\tilde{\mathcal{U}}$ if needed (this keeps all previously discussed properties), we may also assume that it is disjoint from $\mathrm{WF}_{\mathrm{b}}^{-1,\infty}(f)$.

Hence,

$$|\underline{\zeta}_{n-1}|^{-1} W_0 \rho_j = \sum_{i=1}^{2n-4} \tilde F_{ji}\rho_i + \tilde F_{j,2n-3}\tilde\eta \quad \text{for } j = 2, \dots, 2n-4,$$

with $\tilde F_{ji}$ smooth for $i = 1, \dots, 2n-3$ and $j = 2, \dots, 2n-4$. Then we extend ρ_j to a function on ${}^{\mathrm{b}}T^*X \setminus o$ (using the coordinates $(x, y, \underline{\xi}, \underline{\zeta})$), and conclude that

$$|\underline{\zeta}_{n-1}|^{-1} \mathsf{H}_{\tilde p}\rho_j = \sum_{l=1}^{2n-4} \tilde F_{jl}\rho_l + \tilde F_{j,2n-3}\tilde\eta + \tilde F_{j0}x \quad \text{for } j = 2, \dots, 2n-4, \tag{8-34}$$

with $\tilde F_{jl}$ smooth. Similarly, with $\check F_l$ smooth,

$$|\underline{\zeta}_{n-1}|^{-1} \mathsf{H}_{\tilde p}\tilde\eta = 2 + \sum_{l=1}^{2n-4} \check F_l\rho_l + \check F_{2n-3}\tilde\eta + \check F_0 x. \tag{8-35}$$

Let

$$\omega = |x|^2 + \sum_{j=1}^{2n-4} \rho_j^2. \tag{8-36}$$

Finally, we let

$$\phi = \tilde\eta + \omega/(\epsilon^2\delta), \tag{8-37}$$

and define a by

$$a = \chi_0(F^{-1}(2 - \phi/\delta))\chi_1((\tilde\eta\delta)/\epsilon\delta + 1)\chi_2(|\underline{\xi}|^2/\underline{\zeta}_{n-1}^2), \tag{8-38}$$

with χ_0, χ_1 and χ_2 as in the case of the normal propagation estimate, stated after (8-7). We always assume $\epsilon < 1$, so we have

$$\phi \le 2\delta \quad \text{and} \quad \tilde\eta \ge -\epsilon\delta - \delta \ge -2\delta \quad \text{on } \operatorname{supp} a.$$

Since $\omega \ge 0$, the first of these inequalities implies that $\tilde\eta \le 2\delta$, so

$$|\tilde\eta| \le 2\delta \quad \text{on } \operatorname{supp} a. \tag{8-39}$$

Hence,

$$\omega \le \epsilon^2\delta(2\delta - \tilde\eta) \le 4\delta^2\epsilon^2. \tag{8-40}$$

Thus, $\operatorname{supp} a$ lies in $\tilde{\mathcal{U}}$ for $\delta > 0$ sufficiently small. Moreover,

$$\tilde\eta \in [-\delta - \epsilon\delta, -\delta] \quad \text{and} \quad \omega^{1/2} \le 2\epsilon\delta \quad \text{on } \operatorname{supp} d\chi_1, \tag{8-41}$$

so this region lies in (8-32) after ϵ and δ are both replaced by appropriate constant multiples, namely the present δ should be replaced by $\delta/(2|(\underline{\zeta}_{n-1})_0|)$.

We proceed as in the case of hyperbolic points, letting $A_0 \in \Psi_{\mathrm{b}}^0(X)$ with $\sigma_{\mathrm{b},0}(A_0) = a$, supported in the coordinate chart. Also let Λ_r be scalar, with symbol

$$|\underline{\zeta}_{n-1}|^{s+1/2}(1 + r|\underline{\zeta}_{n-1}|^2)^{-s}\,\mathrm{Id} \quad \text{for } r \in [0, 1), \tag{8-42}$$

so $A_r = A\Lambda_r \in \Psi_{\mathrm{b}}^0(X)$ for $r > 0$ and it is uniformly bounded in $\Psi_{\mathrm{bc}}^{s+1/2}(X)$. Then, for $r > 0$,

$$\begin{aligned}\langle \imath A_r^* A_r P u, u\rangle - \langle \imath A_r^* A_r u, P u\rangle &= \langle \imath [A_r^* A_r, P] u, u\rangle + \langle \imath (P - P^*) A_r^* A_r u, u\rangle \\ &= \langle \imath [A_r^* A_r, P] u, u\rangle - 2\,\mathrm{Im}\,\lambda \|A_r u\|^2.\end{aligned} \tag{8-43}$$

and we compute the commutator here using Proposition 8.1. We arrange the terms of the proposition so that the terms in which a vector field differentiates χ_1 are included in E_r and the terms in which a vector fields differentiates χ_2 are included in E_r'. Thus, we have

$$\imath A_r^* A_r P - \imath P A_r^* A_r = (x D_x)^* C_r^\sharp (x D_x) + (x D_x)^* x C_r' + x C_r''(x D_x) + x^2 C_r^\flat + E_r + E_r' + F_r, \tag{8-44}$$

with

$$\begin{aligned}\sigma_{\mathrm{b},2s}(C_r^\sharp) &= w_r^2\big(\digamma^{-1}\delta^{-1} a |\underline{\zeta}_{n-1}|^{-1}(\hat{f}^\sharp + \epsilon^{-2}\delta^{-1} f^\sharp)\chi_0'\chi_1\chi_2 + a^2\tilde{c}_r^\sharp\big),\\ \sigma_{\mathrm{b},2s+1}(C_r') &= w_r^2\big(\digamma^{-1}\delta^{-1} a (\hat{f}' + \delta^{-1}\epsilon^{-2} f')\chi_0'\chi_1\chi_2 + a^2\tilde{c}_r'\big),\\ \sigma_{\mathrm{b},2s+1}(C_r'') &= w_r^2\big(\digamma^{-1}\delta^{-1} a (\hat{f}'' + \delta^{-1}\epsilon^{-2} f'')\chi_0'\chi_1\chi_2 + a^2\tilde{c}_r''\big),\\ \sigma_{\mathrm{b},2s+2}(C_r^\flat) &= w_r^2\big(\digamma^{-1}\delta^{-1} |\underline{\zeta}_{n-1}| a (4 + \hat{f}^\flat + \delta^{-1}\epsilon^{-2} f^\flat)\chi_0'\chi_1\chi_2 + a^2\tilde{c}_r^\flat\big),\end{aligned} \tag{8-45}$$

where $f^\sharp$, f', f'' and $f^\flat$ as well as $\hat{f}^\sharp$, $\hat{f}'$, $\hat{f}''$ and $\hat{f}^\flat$ are all smooth functions on ${}^{\mathrm{b}}T^*X \setminus o$, homogeneous of degree 0 (and independent of ϵ and δ). Moreover, $f^\sharp$, f', f'', $f^\flat$ arise when ω is differentiated in $\chi_0(\digamma^{-1}(2 - \phi/\delta))$, while $\hat{f}^\sharp$, $\hat{f}'$, $\hat{f}''$ and $\hat{f}^\flat$ arise when $\tilde{\eta}$ is differentiated in $\chi_0(\digamma^{-1}(2 - \phi/\delta))$, and comprise all such terms with the exception of part of that arising from the $-\mathsf{H}_h$ component of $V^\flat|_Y$ (which gives the 4 on the last line above, modulo a term included in $\hat{f}^\flat$ and vanishing $\omega = 0$). In addition, since $V^\bullet\rho^2 = 2\rho V^\bullet\rho$ for any function ρ, the terms $f^\bullet$ for $\bullet = \sharp, ', '', \flat$ have vanishing factors of ρ_l and x, with the structure of the remaining factor dictated by the form of $V^\bullet\rho_l$ and $V^\bullet x$, respectively. Thus, using (8-34) to compute $f^\flat$, (8-35) to compute $\hat{f}^\flat$, we have

$$\begin{aligned} f^\sharp &= \sum_k \rho_k f_k^\sharp + x f_0^\sharp, & f^\flat &= \sum_{kl} \rho_k\rho_l f_{kl}^\flat + \sum_k \rho_k x f_k^\flat + x^2 f_0 + \sum_k \rho_k \tilde{\eta} f_{k+}^\flat,\\ f^\bullet &= \sum_k \rho_k f_k^\bullet + x f_0^\bullet \quad \text{for } \bullet = ', '', & \hat{f}^\flat &= x\hat{f}_0^\flat + \sum_k \rho_k \hat{f}_k^\flat + \tilde{\eta}\hat{f}_+^\flat,\end{aligned}$$

with $f_k^\sharp$ etc. smooth. We deduce that

$$\epsilon^{-2}\delta^{-1}|f^\sharp| \le C\epsilon^{-1}, \qquad |\hat{f}^\sharp| \le C, \tag{8-46}$$

$$\epsilon^{-2}\delta^{-1}|f^\bullet| \le C\epsilon^{-1}, \qquad |\hat{f}^\bullet| \le C \quad \text{for } \bullet = ', '', \tag{8-47}$$

$$\epsilon^{-2}\delta^{-1}|f^\flat| \le C\epsilon^{-1}\delta, \qquad |\hat{f}^\flat| \le C\delta. \tag{8-48}$$

We remark that although thus far we worked with a single $q_0 \in K$, the same construction works with q_0 in a neighborhood $\mathcal{U}_{q_0'}$ of a fixed $q_0' \in K$, with a *uniform* constant C. In view of the compactness of K, this suffices (by the rest of the argument we present below) to give the uniform estimate of the proposition.

Since (8-46)–(8-48) are exactly the same (with slightly different notation) as (6.16)–(6.18) of [Vasy 2010a], the rest of the proof is analogous, except that [Vasy 2010a, Lemma 4.6] is replaced by Lemma 8.4 here. Thus, for a small constant $c_0 > 0$ to be determined, which we may assume to be less than C, we demand below that the expressions on the right sides of (8-46) are bounded by $c_0(\epsilon\delta)^{-1}$, those on the right sides of (8-47) are bounded by $c_0(\epsilon\delta)^{-1/2}$, and those on the right sides of (8-48) are bounded by c_0. This demand is due to the appearance of two, one, and zero, respectively, factors of xD_x in (8-44) for the terms whose principal symbols are affected by these, taking into account that in view of Lemma 8.4 we can estimate $\|Q_i v\|$ by $C_{\mathcal{G},K}(\epsilon\delta)^{1/2}\|D_{y_{n-1}}v\|$ if v is microlocalized to a $\epsilon\delta$-neighborhood of $\mathcal{G}$, which is the case for us with $v = A_r u$ in terms of support properties of a.

Thus, recalling that $c_0 > 0$ is to be determined, we require that

$$(C/c_0)^2\delta \leq \epsilon \leq 1, \tag{8-49}$$

and

$$\delta < (c_0/C)^2; \tag{8-50}$$

see [Vasy 2010a, Proposition 6.1] for motivation. Then with ϵ, δ satisfying (8-49) and (8-50) and hence $\delta^{-1} > (C/c_0)^2 > C/c_0$, the bounds (8-46)–(8-48) give that

$$\epsilon^{-2}\delta^{-1}|f^\sharp| \leq c_0\delta^{-1}\epsilon^{-1}, \qquad |\hat{f}^\sharp| \leq c_0\delta^{-1}\epsilon^{-1}, \tag{8-51}$$

$$\epsilon^{-2}\delta^{-1}|f^\bullet| \leq c_0\delta^{-1/2}\epsilon^{-1/2}, \qquad |\hat{f}^\bullet| \leq c_0\delta^{-1/2}\epsilon^{-1/2} \quad \text{for } \bullet = {}', {}'' \tag{8-52}$$

$$\epsilon^{-2}\delta^{-1}|f^\flat| \leq c_0, \qquad |\hat{f}^\flat| \leq c_0, \tag{8-53}$$

as desired. One deduces that

$$\imath A_r^* A_r P - \imath P A_r^* A_r = \tilde{B}_r^*\big(C^* x^2 C + x R^\flat x + (xD_x)^*\tilde{R}' x + x\tilde{R}''(xD_x) + (xD_x)^* R^\sharp (xD_x)\big)\tilde{B}_r + R_r'' + E_r + E_r' \tag{8-54}$$

with

$$R^\flat \in \Psi_{\mathrm{b}}^0(X), \quad \tilde{R}', \tilde{R}'' \in \Psi_{\mathrm{b}}^{-1}(X), \quad R^\sharp \in \Psi_{\mathrm{b}}^{-2}(X),$$
$$R_r'' \in L^\infty((0,1); \mathrm{Diff}_0^2\,\Psi_{\mathrm{b}}^{2s-1}(X)), \quad E_r, E_r' \in L^\infty((0,1); \mathrm{Diff}_0^2\,\Psi_{\mathrm{b}}^{2s}(X)),$$

with

$$\mathrm{WF}_{\mathrm{b}}'(E) \subset \tilde{\eta}^{-1}((-\delta - \epsilon\delta, -\delta]) \cap \omega^{-1}([0, 4\delta^2\epsilon^2)) \subset \tilde{\mathcal{U}}$$

(see (8-41)), $\mathrm{WF}_{\mathrm{b}}'(E') \cap \dot{\Sigma} = \varnothing$, and with $r^\flat = \sigma_{\mathrm{b},0}(R^\flat)$, $\tilde{r}' = \sigma_{\mathrm{b},-1}(\tilde{R}')$, $\tilde{r}'' = \sigma_{\mathrm{b},-1}(\tilde{R}'')$, $r^\sharp \in \sigma_{\mathrm{b},-2}(R^\sharp)$,

$$|r^\flat| \leq 2c_0 + C_2\delta F^{-1}, \qquad |\underline{\zeta}_{n-1}\tilde{r}'| \leq 2c_0\delta^{-1/2}\epsilon^{-1/2} + C_2\delta F^{-1},$$
$$|\underline{\zeta}_{n-1}\tilde{r}''| \leq 2c_0\delta^{-1/2}\epsilon^{-1/2} + C_2\delta F^{-1}, \quad |\underline{\zeta}_{n-1}^2 r^\sharp| \leq 2c_0\delta^{-1}\epsilon^{-1} + C_2\delta F^{-1}.$$

These are analogues of the result of the second displayed equation after [Vasy 2008c, Equation (7.16)], as corrected in [Vasy 2008a], with the small (at this point arbitrary) constant c_0 replacing some constants given there in terms of ϵ and δ; see [Vasy 2010a, Equation (6.25)] for estimates stated in exactly the same form in the form-valued setting. The rest of the argument proceeds as in the proof of [Vasy 2008c, Proposition 7.3], taking into account [Vasy 2008a], and using Lemma 8.4 in place of [Vasy 2008c, Lemma 7.1]. □

Since for λ real, $\lambda < (n-1)^2/4$, both forward and backward propagation are covered by these two results (see Remarks 8.3 and 8.7), we deduce our main result on the propagation of singularities:

Theorem 8.8. *Suppose that $P = \Box + \lambda$, with $\lambda < (n-1)^2/4$, for $m \in \mathbb{R}$ or $m = \infty$. Suppose $u \in H^{1,k}_{0,\mathrm{b},\mathrm{loc}}(X)$ for some $k \leq 0$. Then*

$$(\mathrm{WF}^{1,m}_{\mathrm{b}}(u) \cap \dot{\Sigma}) \setminus \mathrm{WF}^{-1,m+1}_{\mathrm{b}}(Pu)$$

is a union of maximally extended generalized broken bicharacteristics of the conformal metric $\hat{g}$ in

$$\dot{\Sigma} \setminus \mathrm{WF}^{-1,m+1}_{\mathrm{b}}(Pu).$$

In particular, if $Pu = 0$, then $\mathrm{WF}^{1,\infty}_{\mathrm{b}}(u) \subset \dot{\Sigma}$ is a union of maximally extended generalized broken bicharacteristics of $\hat{g}$.

Proof. The proof proceeds as that of [Vasy 2008c, Theorem 8.1], since Propositions 8.2 and 8.6 are complete analogues of [Vasy 2008c, Propositions 6.2 and 7.3]. Given the results of the preceding sections of [Vasy 2008c], the argument proving [Vasy 2008c, Theorem 8.1] is itself only a slight modification of an argument originally due to Melrose and Sjöstrand [1978], as presented by Lebeau [1997] (although we do not need Lebeau's treatment of corners here).

For the convenience of the reader we give a very sketchy version of the proof. To start with, propagation of singularities has already been proved in X°; this is the theorem of Duistermaat and Hörmander [Hörmander 1971]. Now, the theorem can easily be localized — the global version follows by a Zorn's lemma argument; see [Vasy 2008c, proof of Theorem 8.1] for details. Indeed, in view of the Duistermaat and Hörmander's result, it suffices to show that if

$$q_0 \in \mathrm{WF}^{1,m}_{\mathrm{b}}(u) \setminus \mathrm{WF}^{-1,m+1}_{\mathrm{b}}(Pu) \quad \text{and} \quad q_0 \in {}^{\mathrm{b}}T^*_Y X, \tag{8-55}$$

then

$$\begin{aligned}&\text{there exists a generalized broken bicharacteristic } \gamma : [-\epsilon_0, 0] \to \dot{\Sigma},\\ &\text{with } \epsilon_0 > 0, \quad \gamma(0) = q_0, \quad \gamma(s) \in \mathrm{WF}^{1,m}_{\mathrm{b}}(u) \setminus \mathrm{WF}^{-1,m+1}_{\mathrm{b}}(Pu), \quad s \in [-\epsilon_0, 0],\end{aligned} \tag{8-56}$$

for the existence of a GBB on $[0, \epsilon_0]$ can be demonstrated similarly by replacing the forward propagation estimates by backward ones, and, directly from Definition 1.1, piecing together the two GBBs gives one defined on $[-\epsilon_0, \epsilon_0]$. Note that (8-55) implies that $q_0 \in \mathcal{G} \cup \mathcal{H}$ by microlocal elliptic regularity, Proposition 7.7.

Now suppose $q_0 \in (\mathrm{WF}^{1,m}_{\mathrm{b}}(u) \setminus \mathrm{WF}^{-1,m+1}_{\mathrm{b}}(Pu)) \cap {}^{\mathrm{b}}T^*_Y X \cap \mathcal{H}$. We use the notation of Proposition 8.2. Then γ in (8-55) is constructed by taking a sequence $q_n \to q_0$, where $q_n \in T^*X^\circ$ and $\eta(q_n) = -\underline{\hat{\xi}}(q_n) < 0$

and GBB $\gamma_n : [-\epsilon_0, 0] \to \dot\Sigma$ with $\gamma_n(0) = q_n$ and with $\gamma_n(s) \in (\mathrm{WF}_{\mathrm{b}}^{1,m}(u) \setminus \mathrm{WF}_{\mathrm{b}}^{-1,m+1}(Pu)) \cap T^*X^\circ$ for $s \in [-\epsilon_0, 0]$. Once this is done, by compactness of GBB with image in a compact set (see [Vasy 2008c, Proposition 5.5] and [Lebeau 1997, Proposition 6]), one can extract a uniformly convergent subsequence, converging to some γ, giving (8-56). Now, the q_n arise directly from Proposition 8.2, by shrinking U (via shrinking δ in (8-5)); namely under our assumption on q_0, for each such U there must exist a $q \in \mathrm{WF}_{\mathrm{b}}^{1,m}(u)$ in $U \cap \{\eta < 0\}$. The γ_n then arise from the theorem of Duistermaat and Hörmander, using that $\eta(q_n) < 0$ implies that the backward GBB from q_n cannot meet Y for some time ϵ_0, uniform in n — this is essentially due to η being strictly increasing along GBB microlocally, and η vanishing at $\dot\Sigma \cap {}^{\mathrm{b}}T^*_Y X$: So as long as η is negative, the GBB cannot hit the boundary. For more details, see the proof of [Vasy 2008c, Theorem 8.1].

Finally, suppose $q_0 \in (\mathrm{WF}_{\mathrm{b}}^{1,m}(u) \setminus \mathrm{WF}_{\mathrm{b}}^{-1,m+1}(Pu)) \cap {}^{\mathrm{b}}T^*_Y X \cap \mathcal{G}$, which is the more technical case. This part of the argument is present in essentially the same form in [Melrose and Sjöstrand 1978]. Lebeau [1997, Proposition VII.1] gives a very nice presentation; see the proof of [Vasy 2008c, Theorem 8.1] for an overview with more details. The rough idea for constructing the GBB γ for (8-56) is to define approximations to it using Proposition 8.6. First, recall that in Proposition 8.6, applied at q_0, W_0 is the coordinate projection (push forward) of H_p, evaluated at $\hat\pi^{-1}(q_0)$, to T^*Y. Thus, one should think of the point $\tilde\pi(q_0) - \delta W_0$ in T^*Y as an $\mathcal{O}(\delta^2)$ approximation of where a backward GBB should be after "time" (that is, parameter value) δ. This is used as follows: Given $\delta > 0$, Proposition 8.6 gives the existence of a point q_1 in $\mathrm{WF}_{\mathrm{b}}^{1,m}(u)$ that is, roughly speaking, $\mathcal{O}(\delta^2)$ from $\tilde\pi(q_1) - (\tilde\pi(q_0) - \delta W_0)$, with $x(q_1)$ being $\mathcal{O}(\delta^2)$ as well. Then, from q_1, one can repeat this procedure (replacing q_0 by q_1 in Proposition 8.6) — there are some technical issues corresponding to q_1 being in the boundary or not, and also whether in the former case the backward GBB hits the boundary in time δ. Taking $\delta = 2^{-N}\epsilon_0$, this gives $2^N + 1$ points q_j corresponding to the dyadic points on the parameter interval $[-\epsilon_0, 0]$. It is helpful to consider this as analogous to a discrete approximation of solving an ODE without the presence of the boundary by taking steps of size $2^{-N}\epsilon_0$. Defining $\gamma_N(s)$ for only these dyadic values, one can then get a subsequence γ_{N_k} that converges, as $k \to \infty$, at $s = 2^{-n} j\epsilon_0$ for all $n \geq 1$ and $0 \leq j \leq 2^n$ integers. (Note that $\gamma_{N_k}(s)$ is defined for these values of s for k sufficiently large!) One then checks as in Lebeau's proof that the result is the restriction of a GBB to dyadic parameter values. Again, we refer to [Lebeau 1997, Proposition VII.1] and the proof of [Vasy 2008c, Theorem 8.1] for more details. □

In fact, even if $\operatorname{Im}\lambda \neq 0$, we get one-sided statements:

Theorem 8.9. *Suppose that $P = \Box + \lambda$ and $\operatorname{Im}\lambda > 0$, and $m \in \mathbb{R}$ or $m = \infty$. Suppose $u \in H^{1,k}_{0,\mathrm{b},\mathrm{loc}}(X)$ for some $k \leq 0$. Then*

$$(\mathrm{WF}_{\mathrm{b}}^{1,m}(u) \cap \dot\Sigma) \setminus \mathrm{WF}_{\mathrm{b}}^{-1,m+1}(Pu)$$

is a union of maximally forward extended (*and in the case* $\operatorname{Im}\lambda < 0$ backward extended) *generalized broken bicharacteristics of the conformal metric $\hat g$ in*

$$\dot\Sigma \setminus \mathrm{WF}_{\mathrm{b}}^{-1,m+1}(Pu).$$

In particular, if $Pu = 0$, *then* $\mathrm{WF}_{\mathrm{b}}^{1,\infty}(u) \subset \dot{\Sigma}$ *is a union of maximally extended generalized broken bicharacteristics of* $\hat{g}$.

Proof. The proof proceeds again as for Theorem 8.8, but now Propositions 8.2 and 8.6 only allow propagation in one direction. Thus, if $\operatorname{Im}\lambda < 0$, they allow one to conclude that if a point in $\dot{\Sigma} \setminus \mathrm{WF}_{\mathrm{b}}^{-1,m+1}(Pu)$ is in $\mathrm{WF}_{\mathrm{b}}^{1,m}(u)$, then there is another point in $\mathrm{WF}_{\mathrm{b}}^{1,m}(u)$ that is roughly along a *backward* GBB segment emanating from it. Then an actual backward GBB can be constructed as in [Melrose and Sjöstrand 1978; Lebeau 1997]. □

In the absence of b-wave front set we can easily read off the actual expansion at the boundary as well.

Proposition 8.10. *Suppose that* $P = \Box + \lambda$, *where* $\lambda \in \mathbb{C}$. *Let*

$$s_\pm(\lambda) = \tfrac{1}{2}(n-1) \pm \sqrt{\tfrac{1}{4}(n-1)^2 - \lambda}.$$

Suppose $u \in H^1_{0,\mathrm{loc}}(X)$, $\mathrm{WF}_{\mathrm{b}}^{1,\infty}(u) = \varnothing$ *and* $Pu \in \dot{\mathscr{C}}^\infty(X)$. *Then*

$$u = x^{s_+(\lambda)} v_+ \quad \textit{and} \quad v_+ \in \mathscr{C}^\infty(X). \tag{8-57}$$

Conversely, if $\lambda < (n-1)^2/4$, *given any* $g_+ \in \mathscr{C}^\infty(Y)$, *there exists* $v_+ \in \mathscr{C}^\infty(X)$ *and* $v_+|_Y = g_+$ *such that* $u = x^{s_+(\lambda)} v_+$ *satisfies* $Pu \in \dot{\mathscr{C}}^\infty(X)$; *in particular* $u \in H^1_{0,\mathrm{loc}}(X)$ *and* $\mathrm{WF}_{\mathrm{b}}^{1,\infty}(u) = \varnothing$.

This proposition reiterates the importance of the constraint on λ in that

$$x^{(n-1)/2+i\alpha} \notin H^1_{0,\mathrm{loc}}(X) \quad \text{for } \alpha \in \mathbb{R};$$

for $\lambda \geq (n-1)^2/4$, the growth or decay relative to $H^1_{0,\mathrm{loc}}(X)$ does not distinguish between the two approximate solutions $x^{s_\pm(\lambda)} v_\pm$ having $v_\pm \in \mathscr{C}^\infty(X)$.

Proof. For the first part of the lemma, by Lemma 5.16 and the remark after, we have $u \in \mathscr{A}^{(n-1)/2}(X)$ under our assumptions. By (7-1),

$$P + \big((xD_x + \imath(n-1))(xD_x) - \lambda\big) \in x\,\mathrm{Diff}^2_{\mathrm{b}}(X). \tag{8-58}$$

This is, up to a change in overall the sign of the second summand,

$$(xD_x + \imath(n-1))(xD_x) - \lambda,$$

the same as the analogous expression in the de Sitter setting; see the first line of the proof of [Vasy 2010b, Lemma 4.13]. Thus, the proof of that lemma goes through without changes — the reader needs to keep in mind that $u \in \mathscr{A}^{(n-1)/2}(X)$ excludes one of the indicial roots from appearing in the argument of that lemma. (In the de Sitter setting, in [Vasy 2010b, Lemma 4.13] there was no a priori weight, relative to which one has conormality, specified.)

The converse again works as in [Vasy 2010b, Lemma 4.13] using (8-58). □

We can now state the "inhomogeneous Dirichlet problem":

Theorem 8.11. *Assume (TF) and (PT). Suppose $\lambda < (n-1)^2/4$, and $s_+(\lambda) - s_-(\lambda) = 2\sqrt{(n-1)^2/4 - \lambda}$ is not an integer, and $P = P(\lambda) = \Box_g + \lambda$.*

Given $v_0 \in \mathscr{C}^\infty(Y)$ and $f \in \dot{\mathscr{C}}^\infty(X)$, both supported in $\{t \geq t_0\}$, the problem

$$Pu = f, \quad u|_{t<t_0} = 0, \quad u = x^{s_-(\lambda)} v_- + x^{s_+(\lambda)} v_+, \quad v_\pm \in \mathscr{C}^\infty(X), \quad v_-|_Y = v_0,$$

has a unique solution

If $s_+(\lambda) - s_-(\lambda)$ is an integer, the same conclusion holds if we replace the condition $v_- \in \mathscr{C}^\infty(X)$ by $v_- \in \mathscr{C}^\infty(X) + x^{s_+(\lambda)-s_-(\lambda)} \log x\, \mathscr{C}^\infty(X)$.

Proof. The proof of [Vasy 2010b, Lemma 4.13] shows that there exists $\tilde{u}$, supported in $t \geq t_0$, such that $\tilde{u} = x^{s_-(\lambda)} v_-$, v_- is as in the statement of the theorem, and $P\tilde{u} \in \dot{\mathscr{C}}^\infty(X)$. Now let u' be the solution of $Pu' = f - P\tilde{u}$ supported in $\{t \geq t_0\}$, whose existence follows from Theorem 4.16, and which is of the form $x^{s_+(\lambda)} v_+$ by Theorem 8.8 and Proposition 8.10. Then $u = \tilde{u} + u'$ solves the PDE as stated. Uniqueness follows from the basic well-posedness theorem, Theorem 4.16. □

Finally we add well-posedness of possibly rough initial data:

Theorem 8.12. *Assume (TF) and (PT). Suppose $f \in H^{-1,m+1}_{0,\mathrm{b},\mathrm{loc}}(X)$ for some $m \in \mathbb{R}$, and let $m' \leq m$. Then* (1-6) *has a unique solution in $H^{1,m'}_{0,\mathrm{b},\mathrm{loc}}(X)$, which in fact lies in $H^{1,m}_{0,\mathrm{b},\mathrm{loc}}(X)$, and for all compact $K \subset X$ there exists a compact $K' \subset X$ and a constant $C > 0$ such that*

$$\|u\|_{H^{1,m}_0(K)} \leq C \|f\|_{H^{-1,m+1}_{0,\mathrm{b}}(K')}.$$

Remark 8.13. It should be emphasized that if one only wants to prove this result, without microlocal propagation, one could use more elementary energy estimates.

Proof. If $m \geq 0$, then by Theorem 4.16, (1-6) has a unique solution in $H^1_{0,\mathrm{loc}}(X)$, and by propagation of singularities it lies in $H^{1,m}_{0,\mathrm{b},\mathrm{loc}}(X)$, with the desired estimate. Moreover, again by the propagation of singularities, any solution of (1-6) in $H^{1,m'}_{0,\mathrm{b},\mathrm{loc}}(X)$ lies in $H^{1,m}_{0,\mathrm{b},\mathrm{loc}}(X)$, so the solution is indeed unique even in $H^{1,m'}_{0,\mathrm{b},\mathrm{loc}}(X)$.

If $m < 0$, uniqueness and the stability estimate follow as above. To see existence, let $T_0 < t_0$, and let $f_j \to f$ such that $f_j \in H^{-1,1}_{0,\mathrm{b},\mathrm{loc}}$ and $\operatorname{supp} f_j \subset \{t > T_0\}$. This can be achieved by taking $A_r \in \Psi^{-\infty}_{\mathrm{bc}}(X)$ with properly supported Schwartz kernel (of sufficiently small support) such that $\{A_r : r \in (0,1]\}$ is a bounded family in $\Psi^0_{\mathrm{bc}}(X)$, converging to Id in $\Psi^\epsilon_{\mathrm{bc}}(X)$ for $\epsilon > 0$; then with $f_j = A_{r_j} f$, $r_j \to 0$, we have the desired properties. By Theorem 4.16, (1-6) with f replaced by f_j has a unique solution $u_j \in H^1_{0,\mathrm{loc}}(X)$. Moreover, by the propagation of singularities, one has a uniform estimate

$$\|u_k - u_j\|_{H^{1,m}_0(K)} \leq C \|f_k - f_j\|_{H^{-1,m+1}_{0,\mathrm{b}}(K')},$$

with C independent of j and k. In view of the convergence of the f_j in $H^{-1,m+1}_{0,\mathrm{b}}(K')$, we deduce the convergence of the u_j in $H^{1,m}_{0,\mathrm{b}}(K)$ to some $u \in H^{1,m}_{0,\mathrm{b}}(K)$; hence (by uniqueness) we deduce the existence of $u \in H^{1,m}_{0,\mathrm{b},\mathrm{loc}}(X)$ solving $Pu = f$ with support in $\{t \geq T_0\}$. However, as $\operatorname{supp} f \subset \{t \geq t_0\}$, uniqueness shows the vanishing of u on $\{t < t_0\}$, proving the theorem. □

Acknowledgments

I am very grateful to Dean Baskin, Rafe Mazzeo and Richard Melrose for helpful discussions. I would also like to thank the careful referee whose comments helped to improve the exposition significantly and also led to the removal of some very confusing typos.

References

[Anderson 2004] M. T. Anderson, "On the structure of asymptotically de Sitter and anti-de Sitter spaces", *Adv. Theor. Math. Phys.* **8**:5 (2004), 861–894. MR 2006i:53102 Zbl 1096.83018

[Anderson 2005] M. T. Anderson, "Existence and stability of even-dimensional asymptotically de Sitter spaces", *Ann. Henri Poincaré* **6**:5 (2005), 801–820. MR 2007c:53090 Zbl 1100.83004

[Anderson 2008] M. T. Anderson, "Einstein metrics with prescribed conformal infinity on 4-manifolds", *Geom. Funct. Anal.* **18**:2 (2008), 305–366. MR 2009f:53061 Zbl 1148.53033

[Anderson and Chruściel 2005] M. T. Anderson and P. T. Chruściel, "Asymptotically simple solutions of the vacuum Einstein equations in even dimensions", *Comm. Math. Phys.* **260**:3 (2005), 557–577. MR 2007b:58041 Zbl 1094.83002

[Bachelot 2008] A. Bachelot, "The Dirac system on the anti-de Sitter universe", *Comm. Math. Phys.* **283**:1 (2008), 127–167. MR 2009f:83029 Zbl 1153.83013

[Baskin 2010] D. Baskin, "A parametrix for the fundamental solution of the Klein–Gordon equation on asymptotically de Sitter spaces", *J. Funct. Anal.* **259**:7 (2010), 1673–1719. MR 2011h:58044 Zbl 1200.35247

[Bony and Häfner 2008] J.-F. Bony and D. Häfner, "Decay and non-decay of the local energy for the wave equation on the de Sitter–Schwarzschild metric", *Comm. Math. Phys.* **282**:3 (2008), 697–719. MR 2010h:58041 Zbl 1159.35007

[Breitenlohner and Freedman 1982a] P. Breitenlohner and D. Z. Freedman, "Positive energy in anti-de Sitter backgrounds and gauged extended supergravity", *Phys. Lett. B* **115**:3 (1982), 197–201. MR 83i:83055 Zbl 0606.53044

[Breitenlohner and Freedman 1982b] P. Breitenlohner and D. Z. Freedman, "Stability in gauged extended supergravity", *Ann. Physics* **144**:2 (1982), 249–281. MR 84g:83044 Zbl 0606.53044

[Dafermos and Rodnianski 2005] M. Dafermos and I. Rodnianski, "A proof of Price's law for the collapse of a self-gravitating scalar field", *Invent. Math.* **162**:2 (2005), 381–457. MR 2006i:83016 Zbl 1088.83008

[Dafermos and Rodnianski 2007] M. Dafermos and I. Rodnianski, "The wave equation on Schwarzschild–de Sitter space times", preprint, 2007. arXiv 07092766

[Dafermos and Rodnianski 2009] M. Dafermos and I. Rodnianski, "The red-shift effect and radiation decay on black hole spacetimes", *Comm. Pure Appl. Math.* **62**:7 (2009), 859–919. MR 2011b:83059 Zbl 1169.83008

[Fefferman and Graham 1985] C. Fefferman and C. R. Graham, *Conformal invariants: The mathematical heritage of Élie Cartan* (Lyon, 1984), Astérisque, Société Mathématique de France, Paris, 1985. MR 87g:53060 Zbl 0602.53007

[Graham and Lee 1991] C. R. Graham and J. M. Lee, "Einstein metrics with prescribed conformal infinity on the ball", *Adv. Math.* **87**:2 (1991), 186–225. MR 92i:53041 Zbl 0765.53034

[Graham and Witten 1999] C. R. Graham and E. Witten, "Conformal anomaly of submanifold observables in AdS/CFT correspondence", *Nuclear Phys. B* **546**:1-2 (1999), 52–64. MR 2000h:81286 Zbl 0944.81046

[Graham and Zworski 2003] C. R. Graham and M. Zworski, "Scattering matrix in conformal geometry", *Invent. Math.* **152**:1 (2003), 89–118. MR 2004c:58064 Zbl 1030.58022

[Holzegel 2010] G. Holzegel, "On the massive wave equation on slowly rotating Kerr-AdS spacetimes", *Comm. Math. Phys.* **294**:1 (2010), 169–197. MR 2011d:58070 Zbl 1210.58023

[Hörmander 1971] L. Hörmander, "On the existence and the regularity of solutions of linear pseudo-differential equations", *Enseignement Math.* (2) **17** (1971), 99–163. MR 48 #9458 Zbl 0224.35084

[Hörmander 1985] L. Hörmander, *The analysis of linear partial differential operators, III: Pseudodifferential operators*, Grundlehren der Mathematischen Wissenschaften **274**, Springer, Berlin, 1985. MR 87d:35002a Zbl 0601.35001

[Lebeau 1997] G. Lebeau, “Propagation des ondes dans les variétés à coins”, *Ann. Sci. École Norm. Sup.* (4) **30**:4 (1997), 429–497. MR 98d:58183 Zbl 0891.35072

[Mazzeo and Melrose 1987] R. R. Mazzeo and R. B. Melrose, “Meromorphic extension of the resolvent on complete spaces with asymptotically constant negative curvature”, *J. Funct. Anal.* **75**:2 (1987), 260–310. MR 89c:58133 Zbl 0636.58034

[Melrose 1993] R. B. Melrose, *The Atiyah–Patodi–Singer index theorem*, Research Notes in Mathematics **4**, A K Peters, Wellesley, MA, 1993. MR 96g:58180 Zbl 0796.58050

[Melrose and Sjöstrand 1978] R. B. Melrose and J. Sjöstrand, “Singularities of boundary value problems, I”, *Comm. Pure Appl. Math.* **31**:5 (1978), 593–617. MR 58 #11859 Zbl 0368.35020

[Melrose and Sjöstrand 1982] R. B. Melrose and J. Sjöstrand, “Singularities of boundary value problems, II”, *Comm. Pure Appl. Math.* **35**:2 (1982), 129–168. MR 83h:35120 Zbl 0546.35083

[Melrose and Taylor 1985] R. B. Melrose and M. E. Taylor, “Near peak scattering and the corrected Kirchhoff approximation for a convex obstacle”, *Adv. in Math.* **55**:3 (1985), 242–315. MR 86m:35095 Zbl 0591.58034

[Melrose et al. 2008] R. B. Melrose, A. S. Barreto, and A. Vasy, “Asymptotics of solutions of the wave equation on de Sitter–Schwarzschild space”, preprint, 2008. arXiv 0811.2229

[Melrose et al. 2011] R. B. Melrose, A. S. Barreto, and A. Vasy, “Analytic continuation and semiclassical resolvent estimates on asymptotically hyperbolic spaces”, preprint, 2011. arXiv 1103.3507

[Sá Barreto and Zworski 1997] A. Sá Barreto and M. Zworski, “Distribution of resonances for spherical black holes”, *Math. Res. Lett.* **4**:1 (1997), 103–121. MR 97m:83063 Zbl 0883.35120

[Sjöstrand 1980] J. Sjöstrand, “Propagation of analytic singularities for second order Dirichlet problems”, *Comm. Partial Differential Equations* **5**:1 (1980), 41–93. MR 81e:35031a Zbl 0458.35026

[Taylor 1976] M. E. Taylor, “Grazing rays and reflection of singularities of solutions to wave equations”, *Comm. Pure Appl. Math.* **29**:1 (1976), 1–38. MR 53 #1035 Zbl 0318.35009

[Taylor 1996] M. E. Taylor, *Partial differential equations: Basic theory*, Texts in Applied Mathematics **23**, Springer, New York, 1996. MR 98b:35002a Zbl 0869.35001

[Vasy 2005] A. Vasy, “Propagation of singularities for the wave equation on manifolds with corners”, in *Séminaire: Équations aux Dérivées Partielles,* 2004–2005 (Exposé 20), École Polytech., Palaiseau, 2005. MR 2006j:58045

[Vasy 2008a] A. Vasy, “A correction to ‘Propagation of singularities for the wave equation on manifolds with corners”’, preprint, 2008, Available at math.stanford.edu/∼andras/psmc-corr.pdf.

[Vasy 2008b] A. Vasy, “Diffraction by edges”, *Modern Phys. Lett. B* **22**:23 (2008), 2287–2328. MR 2010b:58039 Zbl 1159.78302

[Vasy 2008c] A. Vasy, “Propagation of singularities for the wave equation on manifolds with corners”, *Ann. of Math.* (2) **168**:3 (2008), 749–812. MR 2009i:58037 Zbl 1171.58007

[Vasy 2010a] A. Vasy, “Diffraction at corners for the wave equation on differential forms”, *Comm. Partial Differential Equations* **35**:7 (2010), 1236–1275. MR 2012a:58051 Zbl 1208.58026

[Vasy 2010b] A. Vasy, “The wave equation on asymptotically de Sitter-like spaces”, *Adv. Math.* **223**:1 (2010), 49–97. MR 2011i:58046 Zbl 1191.35064

[Witten 1998] E. Witten, “Anti de Sitter space and holography”, *Adv. Theor. Math. Phys.* **2**:2 (1998), 253–291. MR 99e:81204c Zbl 0914.53048

[Yagdjian and Galstian 2009] K. Yagdjian and A. Galstian, “The Klein–Gordon equation in anti-de Sitter spacetime”, *Rend. Semin. Mat. Univ. Politec. Torino* **67**:2 (2009), 271–292. MR 2011d:35282 Zbl 1184.35109

Received 23 Dec 2009. Revised 11 Oct 2010. Accepted 22 Dec 2010.

ANDRÁS VASY: andras@math.stanford.edu
Department of Mathematics, Stanford University, Stanford, CA 94305-2125, United States
http://math.stanford.edu/∼andras/

ANALYSIS AND PDE
Vol. 5, No. 1, 2012
dx.doi.org/10.2140/apde.2012.5.145

SMALL DATA SCATTERING AND SOLITON STABILITY IN $\dot{H}^{-1/6}$ FOR THE QUARTIC KDV EQUATION

HERBERT KOCH AND JEREMY L. MARZUOLA

We prove scattering for perturbations of solitons in the scaling space appropriate for the quartic nonlinearity, namely $\dot{H}^{-1/6}$. The article relies strongly on refined estimates for a KdV equation linearized at the soliton. In contrast to the work of Tao, we are able to work purely in the scaling space without additional regularity assumptions, allowing us to construct wave operators and a weak version of inverse wave operators.

1. Introduction and statement of results

The generalized Korteweg–de Vries (KdV) equation

$$\begin{cases} \partial_t \psi + \partial_x(\partial_x^2 \psi + \psi^p) = 0 & \text{for } t, x \in \mathbb{R}, \\ \psi(0, x) = \psi_0(x) \end{cases} \tag{1-1}$$

has an explicit soliton solution

$$\psi_c(x,t) = Q_{p,c,c^2t+x_0}(x) := c^{2/(p-1)} Q_p(c(x-(x_0+c^2t)))$$

with $c > 0$, $x_0 \in \mathbb{R}$ and

$$Q_p = \left(\frac{p+1}{2}\right)^{1/(p-1)} \operatorname{sech}^{2/(p-1)}\left(\frac{p-1}{2}x\right). \tag{1-2}$$

Well-posedness of the generalized KdV equation was established by Kenig, Ponce and Vega [Kenig et al. 1993] in H^s for some s depending on p. The case $p = 4$ (quartic KdV) is particularly interesting as it is the only subcritical power nonlinearity that does not lead to a completely integrable system. The critical space for the quartic KdV equation is $H^{-1/6}$. Grünrock [2005] obtained local wellposedness in H^s for $s > -1/6$ and the endpoint $\dot{H}^{-1/6}$ was reached by Tao [2007]. Though wellposedness is not the main focus of this note, we will return to this question in Section 7 and use spaces of bounded p variation and their predual (see the appendix and [Hadac et al. 2009]) to simplify and strengthen Tao's wellposedness result in the critical space.

The solutions $Q_{c,y}$ are called traveling waves or solitons. These are minimizers of the constrained variational problem

$$\min\{E(w) : w \in H^1,\ \|w\|_{L^2} = \mu > 0\}, \tag{1-3}$$

Marzuola was funded by a Hausdorff Center postdoc at the University of Bonn and by a National Science Foundation postdoctoral fellowship. Koch was partially supported by the DFG through Sonderforschungsbereich 611.
MSC2000: 35K40, 35Q51, 35Q53.
Keywords: Korteweg–de-Vries, solitons, scattering.

where

$$E(u) = \int \Big(\tfrac{1}{2} u_x^2 - \frac{1}{p+1} u^{p+1} \Big)\, dx.$$

Minimizers also are extremals of the Lagrangian

$$S(u) = E(u) + \frac{\lambda}{2} \int u^2\, dx, \tag{1-4}$$

where λ is a Lagrangian multiplier. Existence of the minimizer has been shown by Berestycki and Lions [1983] using the constrained minimization problem

$$\min\{T(w) : w \in H^1,\ V(w) = \tilde{\mu}\},$$

where

$$T(w) = \int w_x^2\, dx \quad \text{and} \quad V(w) = \frac{\lambda}{2} \int w^2\, dx - \frac{1}{p+1} \int w^{p+1}\, dx.$$

The function Q in (1-2) is the unique positive even solution to the Euler–Lagrange equation

$$-Q_{xx} - Q^p + Q = 0 \tag{1-5}$$

to (1-4) with $\lambda = 1$. It is a critical point of $S(u)$ again with $\lambda = 1$, a minimizer of E with constraint $\|u\|_{L^2} = \mu$, where

$$\mu^2 = \|Q_p(x)\|_{L^2}^2 = \Big(\frac{p+1}{2} \Big)^{2/(p-1)} \frac{\Gamma\big(\frac{p+1}{p-1}\big)\sqrt{\pi}}{\Gamma\big(\frac{p+3}{2(p-1)}\big)} \tag{1-6}$$

and hence the quadratic form

$$K(\psi) := \int \tfrac{1}{2} w'^2 + \tfrac{1}{2} w^2 - \tfrac{1}{2} p Q^{p-1} w^2\, dx \geq 0 \quad \text{for } \langle w, Q \rangle = 0 \tag{1-7}$$

is nonnegative on the tangent space that is, the functions orthogonal to Q.

The stability of solitons for generic KdV equations has been studied in several seminal works. Orbital stability was first effectively established in the work of Weinstein [1985]. Then asymptotic stability of solitons for KdV was first observed by Pego and Weinstein [1994], who proved that solitons for KdV are stable under perturbations in exponentially weighted spaces. Later, Martel and Merle [2001a; 2005; 2001b] and Martel [2006] refined this result to observe that solitons for generalized KdV equations are indeed stable under perturbations in the energy space, but measured within a moving reference frame. As mentioned above, for the case $p = 4$, building on the multilinear estimates of Grünrock [2005] and the work of Martel and Merle, Tao [2007] assumes smallness in $H^1 \cap \dot{H}^{-1/6}$ and obtains scattering in $\dot{H}^{-1/6}$. We will give a more thorough introduction to previous stability results including rigorous definitions of stability in Section 2.

In the sequel we will focus on the case $p = 4$ and omit p in the notation. It seems that any further progress is tied to an understanding of the linearization, or more precisely of the linear equation

$$u_t + \partial_x \mathcal{L} u = 0 \tag{1-8}$$

and its adjoint

$$v_t + \mathcal{L}\partial_x v = 0, \tag{1-9}$$

which have the explicit solutions (with $\widetilde{Q} = c\partial_c Q_c|_{c=1}$)

$$u = a(\widetilde{Q} + 2tQ') + bQ' \quad \text{and} \quad v = cQ,$$

where

$$\widetilde{Q} := c\frac{d}{dc}c^{2/(p-1)}Q_p(cx)\Big|_{c=1} = \frac{2}{p-1}Q_p + xQ_p', \tag{1-10}$$

usually evaluated at $c = 1$.

Thus both equations (1-8) and (1-9) have linearly growing solutions. It is one of the first contributions of this paper that both equations are uniformly L^2 bounded once we take into account these modes, and, moreover, there are local energy estimates global in time once we remove these modes. In particular the assumption of Pego and Weinstein on the absence of embedded eigenvalues holds.

Our goal is to build on the arguments of Weinstein [1985] and Martel and Merle [2001a; 2005] to establish some type of asymptotic soliton stability for generalized KdV equations by a direct analysis of the equation itself. We apply a variant of Weinstein's and Martel and Merle's arguments to the linear equations (1-8) and (1-9) and their relatives with variable scale and velocity, and control nonlinear terms through estimates for linear equations.

Specifically, we define projection operators related to the spectrum of $\mathcal{L}$:

$$P_{Q'}^{\perp}\psi = \psi - \frac{\langle \psi, Q'\rangle}{\langle Q', Q'\rangle}Q', \quad \tilde{P}\psi = \psi - \frac{\langle \psi, Q\rangle}{\langle Q, \widetilde{Q}\rangle}\widetilde{Q}. \tag{1-11}$$

We obtain the main linear estimates, which in their simplest form can be written as follows.

Theorem 1. *Let S be the solution operator for* (1-8) *and S^* the solution operator for* (1-9). *Then, we have*

$$\sup_t \|S(t)\tilde{P}^* u_0\|_{L^2} + \|\operatorname{sech}(x)\partial_x P_{Q'}^{\perp} S(t)\tilde{P}^* u_0\|_{L^2(\mathbb{R}^2)} \lesssim \|u_0\|_{L^2}, \tag{1-12}$$

$$\sup_t \|S^*(t)P_{Q'}^{\perp} v(t)\|_{L^2} + \|\operatorname{sech}(x)\partial_x \tilde{P} S^*(t)P^{\perp}\|_{L^2(\mathbb{R}^2)} \lesssim \|v_0\|_{L^2}. \tag{1-13}$$

The linear estimates presented in the sequel may be generalized to any subcritical power $p < 5$. We provide variants of Theorem 1 for linearization at solitons with variable scale and velocity as well as estimates in scales of Banach spaces similar to estimates for the Airy equation.

Even near the trivial solution dominating the nonlinear part globally by the linear parts requires to work in a scale invariant space similar to $\dot{H}^{-1/6}$. On the positive side it will lead to scattering for perturbations of a soliton in $\dot{H}^{-1/6}$, without the smallness condition of Tao in the energy space (2-4). The study of the linear equation will lead to a fairly precise understanding of its properties, which seems to be new — we hope that it will provide a model for many other questions on the stability of solitons.

As is standard in the study of stability, we take

$$\psi(x, t) = Q_{c(t)}(x - y(t)) + w(x, t).$$

Then, we have

$$\partial_t w + \partial_x(\partial_x^2 w + 4Q_c^3 w) = -\dot{c}(\partial_c Q_c)(x-y) + \dot{y}(Q_c')(x-y) - \partial_x(\partial_x^2 Q_c - c^2 Q_c + Q_c^4) - c^2(Q_c'(x-y)) - \partial_x(6Q_c^2(x-y)w^2 + 4Q_c(x-y)w^3 + w^4). \tag{1-14}$$

The standard choice of $\dot{c}$ and $\dot{y}$ ensures orthogonality conditions for w. Due to low time regularity we are forced to relax the orthogonality conditions to

$$\frac{\dot{c}}{c}\langle Q_c, \widetilde{Q}_c\rangle = \langle w, Q_c\rangle, \tag{1-15}$$

$$(\dot{y}-c^2)\langle Q_c', Q_c'\rangle = -\kappa\langle w, Q_c'\rangle, \tag{1-16}$$

where $\kappa \gg 1$.

From an implicit function theorem argument similar to that in the proof of [Martel and Merle 2001b, Proposition 1], there exist unique $c(0)$ and $y(0)$ so that $w(\cdot\,,0)$ is orthogonal to $Q_{c(0)}(\cdot - y(0))$ and $Q'_{c(0)}(\cdot - y(0))$ provided the distance of ψ to the set of solitons is small in a suitable norm.

We consider the equations above as ordinary differential equations for c and y, coupled with the partial differential equation.

Using the decomposition and linear estimates, in Sections 8.2 and 8.3 we can prove (referring to later sections for the definition of the function spaces, with $\dot{B}_\infty^{-1/6,2}$ slightly larger than $\dot{H}^{-1/6}$) the following global result:

Theorem 2. *There exists $\epsilon > 0$ and $c > 0$ such that given* (1-1) *with initial data of the form*

$$\min_{c_0, y_0} \|\psi_0 - Q_{c_0}(x-y_0)\|_{\dot{B}_\infty^{-1/6,2}} \le \epsilon,$$

there exist unique functions c and y with

$$\langle w(0), Q_{c(0)}\rangle = \langle w(0), Q'_{c(0)}\rangle = 0, \quad \dot{c} \in L^1 \cap C^0, \quad \dot{y} - c^2 \in L^2 \cap C^0,$$

and a function $w(x,t) \in \dot{X}_\infty^{-1/6}$ such that

$$\psi(x,t) = Q_{c(t),y(t)}(x) + w(x,t)$$

satisfies the quartic KdV equation, and w, c and y satisfy (1-15), (1-16) *and* (1-14). *Moreover,*

$$\|\dot{c}\|_{L^1\cap C^0} + \|\dot{y}-c^2\|_{L^2\cap C^0} + \|w\|_{\dot{X}_\infty^{-1/6}} \le c\|w_0\|_{\dot{B}_\infty^{-1/6,2}}.$$

In addition, there exists a function $z_0 \in \dot{B}_\infty^{-1/6,2}$ such that

$$\|w(t) - e^{-t\partial_x^3} z_0\|_{\dot{B}_\infty^{-1/6,2}} \to 0$$

and

$$\|w(\cdot) - e^{-\cdot\,\partial_x^3} z_0\|_{X_\infty^{-1/6}((t,\infty))} \to 0 \quad \text{as } t \to \infty$$

if $w(0)$ is in the closure of C_0^∞.

In fact, we prove a far stronger result than this, though Theorem 2 captures the main ideas. Finally, in Section 8.4 we show for a function v, there exists a quantity $J(v)$ defined in (8-8) such that we have the following:

Theorem 3. *Let v_0 be in the closure of C_0^∞ in $\dot{B}_\infty^{-1/6,2}$, let $c_\infty > 0$, and let $y_0 \in \mathbb{R}$. Let v be the solution to the linear homogeneous KdV equation. Assume that*

$$J(v) \le \delta \quad \text{for some } \delta = \delta(\|v_0\|_{\dot{B}_\infty^{-1/6,2}}).$$

Then there exists a solution Ψ to the quartic KdV equation, a function $y \in C^1([0,\infty))$, a function $c \in C^1([0,\infty),(0,\infty))$ such that $w = \Psi - Q_{c,y}$, where c and y satisfy equations (1-15), (1-16), (1-14), *and*

$$\langle w(0), Q_{c(0)}(\cdot - y(0))\rangle = \langle w(0), Q'_{c(0)}(\cdot - y(0))\rangle = 0,$$
$$c(t) \to c_\infty, \quad y(0) = y_0, \quad w(t) - v(t) \to 0 \quad \text{in } \dot{B}_\infty^{-1/6,2} \text{ as } t \to \infty.$$

Moreover, if in addition $v_0 \in L^2$, then $\Psi \in C(\mathbb{R}, L^2(\mathbb{R}))$ and

$$\|v_0\|_{L^2}^2 + \|Q_{c_\infty,0}\|_{L^2}^2 = \|\Psi(t)\|_{L^2}.$$

There exists $\varepsilon > 0$ such that the assumptions are satisfied if $\|v_0\|_{\dot{B}_\infty^{-1/6,2}} \le \varepsilon$.

Remark 1.1. The conclusions in Theorems 2 and 3 hold as well in the spaces $\dot{B}_\infty^{-1/6,2} \cap \dot{H}^s \cap H^\sigma$ for any $-1 < s \le 0$ and $\sigma \ge 0$, allowing one to prove uniform bounds in higher Sobolev norms; see Section 7.1. In particular, given initial data in $\dot{B}_\infty^{-1/6,2} \cap \dot{H}^s \cap H^\sigma$, J small will imply stability and scattering in $\dot{B}_\infty^{-1/6,2} \cap \dot{H}^s \cap H^\sigma$. Specifically, we note one can prove boundedness and scattering in the energy space H^1 intersected with $\dot{B}_\infty^{-1/6,2}$.

To motivate the construction of our nonlinear iteration spaces, in Section 3 we first derive some refined estimates for the linear KdV equation

$$\begin{cases} \partial_t u + \partial_x^3 u = f, \\ u(0,x) = u_0(x). \end{cases} \tag{1-17}$$

Then, in Section 4 we discuss the spectral and mapping properties of the operator $\mathcal{L}$ and derive linear estimates for the systems (1-8) and (1-9) and their relatives

$$u_t + u_{xxx} + (Q_{c(t)}(x - x(t))u)_x = f.$$

In Section 5, we combine local smoothing estimates as for (1-17), where we treat the Q terms as error terms with the virial identity and energy conservation for (1-8) to prove uniform bounds for a projection of the solution v assuming orthogonality of the initial data to Q'.

With this first result at hand we pursue a standard though nontrivial path and employ pseudodifferential techniques and duality to derive similar estimates in a full scale of function spaces. The Littlewood–Paley decomposition at low frequencies is severely affected by the term containing Q. This is done in Section 6 with main result Proposition 6.7.

Theorems 2 and 3 are proven in the final two sections by combining the wellposedness arguments and the linear estimates.

2. Review of previous soliton stability results

To begin, we consider the linearized operator

$$\mathscr{L}\psi = -\psi'' - pQ^{p-1}\psi + \psi$$

associated to the Euler–Lagrange equation (1-5) of (1-4) with $\lambda = 1$, respectively the constraint variational problem (1-3) with Lagrange multiplier 1. It is one of the remarkable operators for which almost everything is known about the spectrum and scattering; see [Lamb 1980, Section 2.4 and 2.5], and [Titchmarsh 1962, Section 4.19]. The operator

$$\mathscr{L}_M\psi = -\psi_{xx} - M\operatorname{sech}^2(x)\psi$$

has the continuous spectrum $[0, \infty)$ and the ground state $\psi_0(x) = \operatorname{sech}^\alpha(x)$ with eigenvalue α^2 provided $M = \alpha(\alpha+1)$, with $\alpha > 0$. The other eigenvalues are $(\alpha - j)^2$ for $1 \le j < \alpha$ together with the eigenfunctions can be obtained as follows: Let $\psi_{0,M}$ be the ground state with the constant M. Then,

$$\psi_{j,(\alpha+j)(\alpha+j+1)}(x) = \prod_{l=1}^{j}\Big(\frac{d}{dx} - (\alpha+l)\tanh(x)\Big)\operatorname{sech}^\alpha(x)$$

is the j eigenfunction to the potential with $M = (\alpha+j)(\alpha+j+1)$. We consider this information useful, and we will use these results, even if the arguments could easily be adapted to a much larger class of nonlinearities.

Clearly $\mathscr{L}Q' = 0$ and a short calculation or a comparison with the results above shows that $Q^{(p+1)/2}$ is the ground state with eigenvalue $1 - (p+1)^2/4$. There is no other eigenvalue if $p \ge 3$, but there are other eigenvalues in $(0, 1)$ if $p < 3$. As an immediate consequence $K(\psi) \ge \|\psi\|_{L^2}^2$ if $\langle\psi, Q'\rangle = \langle\psi, Q^{(p+1)/2}\rangle = 0$.

We recall that K is positive definite on the orthogonal complement of Q. We follow [Weinstein 1985] and use this bound to establish a lower bound on a different codimension 2 subspace if $p < 5$. There exists $\delta > 0$ such that

$$K(\psi) \ge \delta\|\psi\|_{H^1}^2 \quad \text{for all } \psi \text{ with } \langle\psi, Q^{p-1}Q'\rangle = \langle\psi, Q\rangle = 0. \tag{2-1}$$

It suffices to verify this statement independently for odd and even functions. For odd functions the quadratic form is nonnegative, with a null space spanned by Q'. Positivity follows from $\langle Q', Q^{p-1}Q'\rangle \ne 0$. The argument for even functions is harder, but again the quadratic form is nonnegative since Q is a local minimizer of the constraint variational problem.

Let ψ_j be a minimizing sequence with $\|\psi_j\|_{H^1} = 1$. Suppose that the left hand side of (2-1) converges to 0. The sequence maximizes $\int Q^{p-1}\psi_j^2\,dx$. There exists a weakly converging subsequence which convergences against a nontrivial even limit ψ since $\psi \to \int Q^{p-1}\psi^2\,dx > 0$ is weakly lower semicontinuous. Moreover $\langle\psi, Q\rangle = 0$ and $\|\psi\|_{H^1} \le 1$. Rescaling if necessary we see that $\|\psi\|_{H^1} = 1$.

We want to show that $K(\psi) > 0$ and argue by contradiction. Suppose that $K(\psi) = 0$. Then by (1-7) ψ is a minimizer of K under the sole constraint $\langle Q, \psi \rangle = 0$ and hence it satisfies the Euler–Lagrange equations

$$\mathcal{L}\psi = \lambda Q.$$

But then ψ is a multiple of $\tilde{Q}$ since

$$\mathcal{L}\tilde{Q} = -2Q$$

is the unique symmetric function with this property. However $\langle Q, \tilde{Q} \rangle \neq 0$ if $p \neq 5$, and hence $\psi = 0$, which contradicts our construction and thus implies the existence of $\delta > 0$ with

$$K(\psi) \geq \delta \|\psi\|_{H^1}.$$

Observe that here the subcriticality condition $p < 5$ enters crucially.

Given ψ we define the parameters c_0 and x_0 by the variational problem

$$\|\psi - Q_{c_0,x_0}\|_{H^1}^2 = \inf_{c,x} \|\psi - Q_{c,x}\|_{H^1}^2.$$

Following Weinstein [1985] we claim

$$\|\psi - Q_{c_0,x_0}\|_{H^1}^2 \leq c(E(\psi) - E(Q_c)), \tag{2-2}$$

provided the left hand side is sufficiently small. This is a consequence of the lower bound for the quadratic form (2-1).

Lyapunov stability of solitons has been shown in the seminal work of Weinstein.

Theorem [Weinstein 1985, Theorem 4]. *Let $\varepsilon > 0$. There exists $\delta > 0$ such that*

$$\inf_{x_0} \|\psi(t) - Q_1(x - x_0)\|_{H^1} \leq \varepsilon \quad \textit{if } \|\psi_0 - Q_1\|_{H^1} \leq \delta.$$

This is a direct consequence of the conservation of the L^2 norm and the energy, plus (2-2).

The study of asymptotic stability began with Pego and Weinstein [1994] in spaces with growing exponential weights. The effect of the weight is twofold. First, there is not much the soliton could interact with on its path to the right. Secondly, small solitons that are slow and prevent asymptotic stability in L^2 carry a weight that makes them exponentially decreasing in time. A key assumption is the absence of embedded eigenvalues of $\partial_x \mathcal{L}$, other than 0 with eigenfunction Q' and the generalized eigenfunction $\tilde{Q}$. Pego and Weinstein verify this assumption for $p = 2$ and $p = 3$ and show that it fails at at most a finite number of values for p between 2 and 5. It is a consequence of the virial identity below that there are no nonzero purely imaginary eigenvalues of $\partial_x \mathcal{L}$.

The exponential weight pushes the continuous spectrum of $\partial_x \mathcal{L}$ to the left, makes the problem more parabolic, and allows the use of techniques from smooth dynamical systems, in particular of a center manifold reduction that is a restriction of the flow to a two dimensional manifold.

Martel and Merle [2001a; 2005] and Martel [2006] introduced a virial identity or monotonicity formula for the adjoint problem (1-9) as well as for nonlinear problems. Let

$$\eta(x) = -\frac{p+1}{p-1}\frac{Q'}{Q} = \frac{p+1}{2}\tanh\frac{p-1}{2}x$$

and suppose that v satisfies the Equation (1-9). By direct computation we have

$$-\frac{d}{dt}\int \eta v^2\,dx = \langle (3(\mathscr{L}+\tfrac{1}{4}(p+1)^2-1)Q^{(p-1)/2}v,\, Q^{(p-1)/2}v\rangle, \tag{2-3}$$

where the quadratic form is nonnegative and it has by the spectral theory of Schrödinger operators with $\operatorname{sech}^2(x)$ potentials a one-dimensional null space spanned by Q. There are two consequences: the quantity on the left hand side is monotonically decreasing, and the right hand side controls the H^1 norm of $Q^{(p-1)/2}v$ provided v is orthogonal to a vector $\overline{Q}$ with $\langle \overline{Q}, Q\rangle \neq 0$. Hence, if $v(0)$ is orthogonal to Q' and $\widetilde{Q}$, which is preserved under the evolution,

$$\|Q^{(p-1)/2}v\|_{H^1} \le c\sup_t \|v(t)\|_{L^2}.$$

The left hand side is controlled provided we obtain a bound on $\sup_t\|v(t)\|_{L^2}$. Martel and Merle [2001a; 2005] use this and related observations together with the a priori control on the deviation of the solution to the set of solitons in ingenious ways for indirect arguments: The existence of a solution H^1 close to solitons, but not asymptotically converging to the soliton "on the right" leads to the existence of impossible objects.

Later, Côte [2006] constructed solutions with specific asymptotic conditions including many soliton solutions for positive time. This shows that L^2 convergence to a soliton will not be true without restricting the set where convergence is studied.

Already L^2 conservation precludes asymptotic stability of the trivial solution. The relevant notion instead of asymptotic stability is for unitary problems the notion of scattering. Suppose that $\psi(0)$ is close to a soliton. We seek a function w satisfying the Airy equation as well as $c(t)$ and $y(t)$ and a Banach space X such that $\|\psi - Q_{c(t)}(x-y(t)) - w(t)\|_X \to 0$ as $t\to\infty$. Tao [2007] verifies scattering in the following sense: Suppose that

$$\|\psi(0)-Q(0)\|_{H^1} + \|\psi(0)-Q\|_{\dot H^{-1/6}} \ll 1. \tag{2-4}$$

Then scattering holds with $X = \dot H^{-1/6}$. Tao relies on the work of Martel and Merle, and in particular on Weinstein's a priori estimate of the difference to the soliton.

3. The Airy equation

For purposes of understanding and motivating dispersive estimates for the linearized KdV equation, here we study and collect results for the Airy equation

$$\begin{cases} v_t + v_{xxx} = 0, \\ v(x,0) = v_0(x). \end{cases} \tag{3-1}$$

The solution operator defines a unitary group $S(t)$ with the kernel

$$K(t,x) = t^{-1/3} Ai(xt^{-1/3}),$$

where as $x \to \infty$ the Airy function is roughly $x^{-1/4} e^{-x^{3/2}}$, and as $x \to -\infty$ the Airy function is roughly $\mathrm{Re}(x^{-1/4} e^{-ix^{3/2}})$. Strichartz estimates for solutions,

$$\|u\|_{L^p L^q} \le c \| |D|^{-1/p} u_0 \|_{L^2} \tag{3-2}$$

where $L^p L^q$ is the standard space time norm such that the L^p norm in time of the L^q norm in space and

$$\frac{2}{p} + \frac{1}{q} = \frac{1}{2},$$

follow as an immediate consequence. Of particular interest for this work are the homogeneous Strichartz pair $(p,q) = (6,6)$ as well as the endpoint Strichartz pair $(p,q) = (4,\infty)$. For an overview of Airy function asymptotics, see [Fedoryuk 1993].

Local smoothing estimates for (3-1) go back to [Kato 1983]. Here we are interested in a more general version of them. Let $\gamma(t,x) \ge 1$ be a smooth bounded increasing function. We calculate

$$\frac{d}{dt} \int \gamma u^2 \, dx = \int (\gamma_t + \gamma^{(3)}) u^2 - 3\gamma' u_x^2 \, dx \tag{3-3}$$

and search for conditions ensuring that the right hand side is nonpositive. We assume

$$\partial_x^3 \gamma \le -\tfrac{2}{3} \partial_t \gamma \tag{3-4}$$

with the easiest case being $\gamma(t,x) = \gamma_0(x-t)$, for which we assume

$$\gamma_0^{(3)} \le \tfrac{2}{3} \gamma_0'. \tag{3-5}$$

We get

$$\frac{d}{dt} \int \gamma u^2 \, dx + \int \gamma' (u_x^2 + \tfrac{1}{3} u^2) \, dx \le 0. \tag{3-6}$$

Let us fix a particular example,

$$\gamma_0(x) = 1 + \int_{-\infty}^{x} (1 + |y|^2)^{-(1+\varepsilon)/2} dy. \tag{3-7}$$

It satisfies the criteria and, provided ε is sufficiently small, a straightforward calculation gives (3-5). Next, it is instructive to consider a scaling. For $\mu > 0$ and γ_0 as above we define

$$\gamma_\mu(t,x) = \gamma_0(\mu^{-1}(x - \mu^{-2} t)).$$

Then,

$$\frac{d}{dt} \int \gamma_\mu u^2 \, dx + \int \gamma_\mu' (u_x^2 + \frac{1}{3\mu^2} u^2) \, dx \le 0. \tag{3-8}$$

One may easily generalize this inequality by choosing $t \to y(t)$ with $\dot{y} \ge \frac{1}{8} \mu^{-2}$, and setting $\gamma(t,x) = \gamma_0(\mu^{-1}(x - y(t)))$. In the sequel we will always restrict ourselves to $\mu = 1$.

The virial identity clearly generalizes to functions spaces with different regularity. To see this, we first define the space H^s_ρ (and similarly L^2_ρ) by the norm

$$\|u\|^2_{H^s_\rho} = \int |\langle D\rangle^s u|^2 \rho^2(x)\, dx < \infty,$$

where $\rho > 0$ with uniformly bounded derivatives of order up to k for some $k \ge |s|$ and $\langle D\rangle^s$ is defined through the Fourier multiplication $(1+|\xi|^2)^{s/2}$. Similarly we define ρH^s where $u \in \rho H^s$ if and only if

$$u = \rho f \quad \text{for } f \in H^s \quad \text{and} \quad \|u\|_{\rho H^s} = \inf_{u=\rho f} \|f\|_{H^s}.$$

The function ρ will often depend on t. Given a Banach space X, we denote the space of X-valued L^2 functions by $L^2 X$, and give the obvious meaning to $L^2 \rho H^s$ and $L^2 H^s_\rho$. Such spaces will be explored further in Section 4.

Remark 3.1. We note that $\rho H^s = H^s_{\rho^{-1}}$, if ρ is nonnegative, up to equivalent norms. However as we wish to highlight the use of duality throughout the linear analysis and construction of iteration spaces, we adopt the ρH^s convention.

If γ satisfies the assumptions above and

$$\begin{cases} u_t - u_{xxx} = f, & \text{where } f \in L^2\sqrt{\gamma'}H^{-1}, \\ u(0,x) = u_0(x), & \text{where } u_0 \in L^2, \end{cases} \tag{3-9}$$

we obtain by an obvious modification of the argument above

$$\|u\|_{L^\infty L^2} + \|u\|_{L^2 H^1_{\sqrt{\gamma'}}} \le c\left(\|u(0)\|_{L^2} + \|f\|_{L^2\sqrt{\gamma'}H^{-1}}\right). \tag{3-10}$$

We turn to a useful technical result.

Lemma 3.1. *Let $m \in C^\infty(\mathbb{R})$ satisfy $|m^{(j)}(\xi)| \le c_j \langle\xi\rangle^{s-j}$ for $j \ge 1$ and let $m(D)$ be the Fourier multiplier defined by m. Suppose that $\gamma \in C^\infty$,*

$$|\gamma^{(j)}(x)| \lesssim \gamma(x) \quad \text{for } j \ge 0, \text{ and}$$
$$|1 - \gamma(x)/\gamma(y)| \lesssim c(|x-y| + |x-y|^N) \quad \text{for some } N.$$

For any $a \in \mathbb{R}$ we have

$$\|\gamma^{-a}[m(D), \gamma^a]\langle D\rangle^{1-s} f\|_{L^2} + \|[m(D), \gamma^a]\gamma^{-a}\langle D\rangle^{1-s} f\|_{L^2} \le c_{s,a}\|f\|_{L^2}$$

and

$$\|\langle D\rangle^{1-s}[m(D), \gamma^a]\gamma^{-a} f\|_{L^2} + \|\langle D\rangle^{1-s}\gamma^{-a}[m(D), \gamma^a] f\|_{L^2} \le c_{s,a}\|f\|_{L^2}.$$

The most important example of m is the Fourier multiplier $\langle D\rangle^s$ defined by the function $(1+|\xi|^2)^{s/2}$.

Proof. We begin with the estimate of the first term in the first inequality, the second term being similar. We decompose $m(D) = m_0(D) + m_1(D)$, where the convolution kernel $m_0(x)$ of $m_0(D)$ is supported in $|x| \le 2$, and the one for $m_1(D)$ is supported in $|x| \ge 1$. The convolution kernel $m_1(x)$ together with its derivatives decays exponentially.

The integral kernel of $\gamma^{-a}[m_1(D), \gamma^a]$ is

$$K_1(x, y) = m_1(x-y)\Big(1 - \Big(\frac{\gamma(x)}{\gamma(y)}\Big)^a\Big).$$

The kernel and its derivatives decay like $\langle x-y\rangle^{-N}$, which implies

$$\|\gamma^{-a}[m_1(D), \gamma^a] f\|_{H^N} \le c_N \|f\|_{H^{-N}}$$

for all $N > 0$ by Schur's lemma. It remains to prove

$$\|\gamma^{-a}[m_0(D), \gamma^a]\langle D\rangle^{1-s} f\|_{L^2} \le c_{s,a}\|f\|_{L^2}.$$

We decompose $\langle D\rangle^s = D_0 + D_1$. The bound for $\gamma^{-a}[m_0(D), \gamma^a]D_0$ follows from standard pseudodifferential calculus. The bound for the term with D_1 follows from

$$\|\gamma^{-a}[m_0(D), \gamma^a] f\|_{L^2} \le c_N \|f\|_{H^N},$$

which again follows easily by standard pseudodifferential calculus. □

Lemma 3.2. *Suppose that*

$$\begin{cases} u_t + u_{xxx} = f, & \text{where } f \in \sqrt{\gamma'} H^{s-1}, \\ u(0, x) = u_0(x), & \text{where } u_0 \in H^s. \end{cases} \tag{3-11}$$

Then

$$\|u\|_{L^\infty H^s} + \|u\|_{L^2 H^{s+1}_{\sqrt{\gamma'}}} \le c\left(\|u(0)\|_{H^s} + \|f\|_{L^2 \sqrt{\gamma'} H^{s-1}}\right). \tag{3-12}$$

Moreover, if

$$\begin{cases} u_t + u_{xxx} = (\operatorname{sech}^2(x - x(t)) f)_x + \partial_x g, \\ u(0, x) = u_0(x), \end{cases} \tag{3-13}$$

with $\dot{x} \ge \delta$, then

$$\|u\|_{L^\infty \dot{H}^{-1}} + \|u\|_{L^2 H^0_{\sqrt{\gamma'}}} \lesssim \|u(0)\|_{\dot{H}^{-1}} + \|f\|_{L^2 H^{-1}} + \|g\|_{L^1 L^2}. \tag{3-14}$$

Proof. We set $v = \langle D\rangle^s u$, where u satisfies (3-11); hence

$$v_t + v_{xxx} = \langle D\rangle^s f,$$

and

$$\|u\|_{L^\infty H^s} + \|u\|_{L^2 H^{s+1}_{\sqrt{\gamma'}}} = \|v\|_{L^\infty L^2} + \|v\|_{L^2 H^1_{\sqrt{\gamma'}}} \le c(\|v(0)\|_{L^2} + \|\langle D\rangle^s f\|_{L^2 \gamma H^{-1}}),$$

where the first term is equal to $\|u(0)\|_{H^s}$ and

$$\begin{aligned} \|\langle D\rangle^s f\|_{L^2\sqrt{\gamma'} H^{-1}} &= \|\langle D\rangle^{-1}(\gamma')^{-1/2}\langle D\rangle^s f\|_{L^2 L^2} \\ &\le \|f\|_{L^2\sqrt{\gamma'} H^{s-1}} + \|[\langle D\rangle^{-1}, (\gamma')^{-1/2}]\langle D\rangle^s f\|_{L^2 L^2} \\ &\le \|f\|_{L^2\sqrt{\gamma'} H^{s-1}} + \|(\gamma')^{-1/2}\langle D\rangle^{s-2} f\|_{L^2 L^2}. \end{aligned}$$

The last inequality follows from Lemma 3.1 applied with γ' for γ, $a = 1/2$, and -1 for s. This implies the desired estimate (3-12). Now suppose that u satisfies (3-13) and let v be the antiderivative of u with respect to x. It satisfies

$$\begin{cases} v_t + v_{xxx} = \operatorname{sech}^2(x - x(t)) f + g, \\ v(0, x) = v_0(x); \end{cases} \tag{3-15}$$

hence

$$\|u\|_{L^\infty \dot{H}^{-1}} + \|u\|_{L^2 H^0_{\sqrt{\gamma'}}} \le c\|g\|_{L^1 L^2} + \|f\|_{L^2 H^{-1}}.$$ □

4. Properties of the Schrödinger operator

We briefly recall notions from the introduction. Given $p > 1$, solitons of the form $Q_p(x-t)$ satisfy (1-5) and it is not hard to verify that all bounded solutions are translates of $\pm Q_p$ in Equation (1-2). Similarly $Q_{p,c} = c^{2/(p-1)} Q_p(cx)$ satisfies

$$\partial_x^2 (Q_p)_c - c^2 (Q_p)_c + (Q_p)_c^p = 0. \tag{4-1}$$

We will focus on $p = 4$ and omit again p from the notation. Let $'$ denote differentiation with respect to x and $\dot{}$ differentiation with respect to time. We recall the definition of $\widetilde{Q}$ from (1-10) and $\widetilde{Q}_c = c\partial_c Q_c$ respectively $\widetilde{\widetilde{Q}}_c = c\partial_c c\partial_c Q_c$, the corresponding differentiation at c. There are many explicit calculations, and we collect some of them here. Using the properties of $Q_c(x) = c^{2/3} Q(cx)$, it follows that

$$\|Q_c\|_{L^2} = c^{1/6} \|Q_1\|_{L^2}, \quad \langle \widetilde{Q}_c, Q_c \rangle = \tfrac{1}{2} c\partial_c \|Q_c\|_{L^2}^2 = \tfrac{1}{6} \|Q_c\|_{L^2}^2, \tag{4-2}$$

where the L^2 norm is given by (1-6), and

$$\|Q_c'\|_{L^2} = c^{7/6} \|Q_1'\|_{L^2}. \tag{4-3}$$

In addition,

$$\partial_x Q_c = c^{5/3} Q'(cx) \quad \text{and} \quad c\partial_c Q_c = \left(\tfrac{2}{3} Q_c + x Q_c'\right) = \widetilde{Q}_c = c^{2/3} \widetilde{Q}(cx).$$

The operator $\mathcal{L}_c$ is defined by

$$\mathcal{L}_c u = -u_{xx} + c^2 u - 4 Q_c^3 u, \tag{4-4}$$

where we mostly omit y and c if $c = 1$. We recall that virtually everything is known about the spectrum of $\mathcal{L}$; see [Andrews et al. 1999; Lamb 1980; Titchmarsh 1962]. We summarize the findings below. We also refer to [Martel 2006; Weinstein 1985] and the references therein for extensive discussions of these properties for more general operators of type similar to $\mathcal{L}$.

By direct differentiation in x of (1-5), we see $\mathcal{L}Q' = 0$. Hence, the null space of $\mathcal{L}$ consists at least of the space $\alpha Q'$ for all $\alpha \in \mathbb{R}$. Similarly, by differentiation in c of (4-1), we see $\mathcal{L}(\widetilde{Q}) = -2Q$, so $\partial_x \mathcal{L}$ has at least a 2-dimensional generalized null space. Also, since $Q' = 0$ only at $x = 0$, we know from the Sturm oscillation theorem that there exists some $\lambda_0 > 0$ and $\mathcal{Q}_0 > 0$ such that $\mathcal{L}\mathcal{Q}_0 = -\lambda_0 \mathcal{Q}_0$, the unique negative eigenstate of $\mathcal{L}$. Note, because $\mathcal{L}$ is a sech^2 potential perturbation of the Laplacian, it is possible to exactly construct $\mathcal{Q}_0 = Q^{5/2}$ and $\lambda_0 = 21/4$ using standard techniques. The analysis above summarizes the entire discrete spectral decomposition for $\mathcal{L}$.

Following the analysis in [Weinstein 1985, Propositions 2.7 and 2.9], if

$$\langle \tilde{u}, Q\rangle = 0 \quad \text{and} \quad \langle \tilde{u}, Q'\rangle = 0,$$

then there exists $k_0 > 0$ such that

$$\langle \tilde{u}, \mathcal{L}\tilde{u}\rangle \geq k_0 \|\tilde{u}\|_{L^2}^2. \tag{4-5}$$

Here k_0 depends only on the power $p = 4$ in (1-1).

We will consider $\rho = e^\nu$ with $\nu \in C^{|s|+1}$ with

$$|\nu^{(j)}(x)| \leq \varepsilon \tag{4-6}$$

for $0 \leq j \leq |s| + 1$ and a small constant ε to be chosen later. Clearly we may regularize ν and hence $\rho = e^\nu$ without changing the spaces. Then

$$u \in H^s_\rho \iff \rho u \in H^s \iff u \in \rho^{-1} H^s.$$

It is quite obvious that the dual space of H^s_ρ is ρH^{-s} with isometric norms, and this statement does not depend on the regularity of ρ. We recall the definition of the projectors (1-11).

Lemma 4.1. *For all $s \in \mathbb{R}$, there exists $C > 0$ such that*

$$\|P^\perp_{Q'} u\|_{H^{s+2}} \leq C\|\mathcal{L}u\|_{H^s}, \quad \|P^\perp_{Q'} u\|_{\rho H^{s+2}} \leq C\|\mathcal{L}u\|_{\rho H^s}, \quad \|P^\perp_{Q'} u\|_{H^{s+2}_\rho} \leq C\|\mathcal{L}u\|_{H^s_\rho}.$$

Proof. The first inequality is an immediate consequence of the nature of the spectrum described above along with ellipticity. The second and the third statement are equivalent because $H^s_\rho = \rho^{-1} H^s$, with equivalent norms.

Fix $\mu = 1 - (p+1)^2/4$, where $p = 4$. For $\lambda = \lambda_0 + i\lambda_1$ in the complex half plane left of μ, we obtain the resolvent estimate

$$|\lambda - \mu| \|u\|_{L^2} \leq \|(\mathcal{L} - \lambda)u\|_{L^2}$$

and also for some $1 > \kappa > 0$, we have

$$\begin{aligned}\operatorname{Re} \int u\overline{(\mathcal{L} - \lambda)u}\, dx &\geq |\lambda - \mu| \|u\|_{L^2}^2 + \langle (\mathcal{L} - \mu)u, u\rangle \\ &\geq \tfrac{1}{2}|\lambda - mu| \|u\|_{L^2}^2 + \kappa \|u_x\|_{L^2}^2 + (1-\kappa)\langle (\mathcal{L} - \mu)u, u\rangle \\ &\qquad + \left(\tfrac{1}{2}|\lambda - \mu| + \kappa(1-\mu) - 4\kappa \|Q\|_{L^\infty}^3\right) \|u\|_{L^2}^2 \\ &\geq \tfrac{1}{2}|\mu - \lambda_0| \|u\|_{L^2}^2 + \min\left\{ \frac{|\mu - \lambda_0|}{8\|Q\|_{L^\infty}^3}, \frac{1}{2} \right\} \|u_x\|_{L^2}^2\end{aligned}$$

by the obvious choice of κ.

We obtain the estimate for λ with real part at most μ:

$$|\mu - \lambda| \|u\|_{L^2} + \min\{|\lambda - \mu|, 1\} \|u_x\|_{L^2} \leq C \operatorname{Re}\langle (\mathcal{L} - \lambda)u, u\rangle \leq C \|\langle (\mathcal{L} - \lambda)u\|_{H^{-1}} \|u\|_{H^1}.$$

These estimates imply that the resolvent $(\mathcal{L}-\lambda)^{-1}$ defines a continuous uniformly bounded map (for $\operatorname{Re}\lambda \le \lambda_0 < \mu$) from H^{-1} to H^1. Moreover,

$$\|u\|_{L^2} \le |\lambda-\mu|^{-1}\|\mathcal{L}u\|_{L^2} \quad \text{and} \quad \|u_x\|_{L^2}^2 \le |\lambda-\mu|^{-1}\max\left\{\frac{1}{2}, \frac{8\|Q\|_{L^\infty}}{|\lambda-\mu|}\right\}\|(\mathcal{L}-\lambda)u\|_{L^2}^2.$$

We turn to the weighted estimates and calculate formally

$$e^{\nu}(\mathcal{L}-\lambda)e^{-\nu} = \mathcal{L}-\lambda-|\nu'|^2+\partial_x\nu'+\nu'\partial_x,$$

and hence, since $\partial_x\nu'+\nu'\partial_x$ is antisymmetric,

$$\operatorname{Re}\int u\overline{e^{\nu}(\mathcal{L}-\lambda)e^{-\nu}u)}\,dx = \operatorname{Re}\int u\overline{(\mathcal{L}-\lambda)u}\,dx - \|\nu'u\|_{L^2}^2 \ge \tfrac{1}{2}|\lambda_0-\mu|\|u\|_{L^2}^2$$

if $\varepsilon \le \sqrt{|\lambda_0-\mu|/2}$, which we assume in the sequel. As above we obtain with an explicit constant C

$$\|u\|_{H^1} \le C\|e^{\nu}(\mathcal{L}-\lambda)e^{-\nu}u\|_{H^{-1}}. \tag{4-7}$$

It follows from these estimates that given $\delta > 0$ there is a single resolvent family (for $\operatorname{Re}\lambda < \mu-\delta$) mapping $\rho H^{-1} \to \rho H^1$ and from $H^{-1}_\rho \to H^1_\rho$, provided ε is sufficiently small.

Recall that $\mathcal{L}$ has a zero eigenvalue with eigenfunction Q' and a single negative eigenvalue $-\lambda_0$ with a ground state $\mathcal{Q}_0$. Let P be the orthogonal projection to the orthogonal complement of these two eigenfunctions. The remaining spectrum is contained in $[\rho,\infty)$, where $\rho>0$ is either 1 (if $p\ge 3$), or the next positive eigenvalue, which can be easily be calculated. Moreover, $\mathcal{L}$ is selfadjoint. The resolvent $R(\lambda) = (\mathcal{L}-\lambda)^{-1}$ is a holomorphic map in $\mathbb{C}\setminus(1,\infty)$ with simple poles in μ, 0, and possibly some other eigenvalues in (0, 1). In addition, $R_0(\lambda) = R(\lambda)P$ has a continuous and hence holomorphic extension to $\lambda = 0$ and $\lambda = -\lambda_0$, which is uniformly bounded in each half plane strictly left of ρ.

By Equation (4-7) the resolvent is uniformly bounded on the weighted spaces if λ is in the half plane left of $-\mu$. Decreasing ε if necessary (so that the orthogonal projection $P^{\perp}_{Q'}$ along Q' is bounded in the weighted space), we obtain the same statement for $R_0(\lambda)$. Now complex interpolation implies

$$\|\mathcal{L}^{-1}Pf\|_{H^1_\rho} \le C\|Pf\|_{H^{-1}_\rho}.$$

This implies the desired estimates for $s=-1$.

Standard elliptic theory extends this estimate to

$$\|u\|_{H^{s+2}_\rho} \le C\|(\mathcal{L}-\lambda)u\|_{H^s_\rho}, \tag{4-8}$$

$$\|u\|_{\rho H^{s+2}} \le C\|(\mathcal{L}-\lambda)u\|_{\rho H^s} \tag{4-9}$$

first to all $s \ge -1$, and then, by duality, to all $s \in \mathbb{R}$. The first estimate is the special situation when ν is constant.

We conclude with the trivial observation that we may replace (4-6) by $\lim_{x\to\in\infty}\nu^j = 0$, which holds for $\rho(x) = (1+|x|^2)^a$ for all real numbers a, since in that case we may choose an equivalent norm that satisfies (4-6). □

5. Energy methods for the linearized equation

We turn to a study of what we call the linear u-problem

$$\begin{cases} u_t = \partial_x(\mathcal{L}u), \\ u(0,x) = u_0, \end{cases} \tag{5-1}$$

where

$$\mathcal{L}u = (-\partial_x^2 + 1 - 4Q^3)u = \left(-\partial_x^2 + 1 - 10\operatorname{sech}^2\left(\tfrac{3}{2}x\right)\right)u.$$

We note here that $\mathcal{L}$ is the operator that results from linearization of the KdV equation about Q when we work in a moving reference frame or in other words make the change of variables $x \to x - t$. Indeed, setting $\psi(x,t) = Q(x-t) + u(x-t,t)$ and plugging into (1-1), we get

$$\partial_t u = -\partial_x(\partial_x^2 u - u + (Q+u)^4 - Q^4 + \partial_x^2 Q - Q + Q^4) = \partial_x(\mathcal{L}u) - \partial_x(6Q^2u^2 + 4Qu^3 + u^4).$$

For reasons that will become clear in the sequel, we also consider the linear v-problem

$$\begin{cases} v_t = \mathcal{L}(\partial_x v), \\ v(0,x) = v_0. \end{cases} \tag{5-2}$$

The two equations (5-2) and (5-1) are related in many ways.

(1) They are dual to each other.

(2) If u satisfies the u equation, then $v = \partial_x u$ satisfies the v equation.

(3) If v satisfies the v equation, then $u = \mathcal{L}v$ satisfies the u equation.

We observe that $u = Q'$ is a solution to the u equation, and hence $\langle v, Q'\rangle$ is preserved by the flow for v. In particular orthogonality is preserved by the evolution. Similarly, $v = Q$ is a solution to the v equation and $\langle u, Q\rangle$ is preserved by the u flow. Moreover, $u = aQ' + b(\widetilde{Q} + 2tQ')$ satisfies the u equation for all coefficients a and b. As a consequence both equations admit solutions that grow linearly with time. Moreover, if v satisfies the v equation, then

$$\frac{d}{dt}\langle v, \widetilde{Q}\rangle + 2t\langle v, Q'\rangle = 0$$

and v is orthogonal to $\widetilde{Q}$ and Q' provided it is initially.

Inspired by a set of ideas collected from [Martel and Merle 2008] and the references therein, let us look at a virial identity for (5-2), namely

$$I_\eta(v) = -\int \eta(x) v^2\, dx,$$

where $\eta(x)$ will be defined in the sequel. We have

$$
\begin{aligned}
-\frac{d}{dt} I_\eta(v) &= -2 \int \eta(x) v(\mathcal{L} v_x)\, dx \\
&= -2 \int \eta(x) v((-\partial_x^2 + 1 - 4Q^3) v_x)\, dx \\
&= 2 \int \eta v \partial_x^3 v\, dx - 2 \int \eta v \partial_x v\, dx + 8 \int \eta Q^3 v \partial_x v\, dx \\
&= -2 \int \eta' v \partial_x^2 v\, dx - 2 \int \eta \partial_x v \partial_x^2 v\, dx + \int \eta_x v^2\, dx \\
&\qquad\qquad - 4 \int \eta' Q^3 v^2\, dx - 4 \int \eta \partial_x \big(\operatorname{sech}^2\big(\tfrac{3}{2}x\big)\big) v^2\, dx \\
&= 3 \int \eta' v_x^2\, dx + 2 \int \eta'' v \partial_x v\, dx + \int \eta' v^2\, dx - 4 \int \eta' Q^3 v^2\, dx - 12 \int \eta Q^2 Q' v^2\, dx.
\end{aligned}
$$

As in [Martel 2006], we take

$$
\eta(x) = -\frac{5}{3} \frac{Q'}{Q} = \tfrac{5}{3} \tanh\big(\tfrac{3}{2}x\big), \tag{5-3}
$$

which is similar to x near 0 and bounded at ∞. Note, the sign convention here is chosen to match that of [Martel and Merle 2008]. By direct computation we have

$$
\begin{aligned}
\eta'(x) &= Q^3(x), & (Q^3\eta)' &= -5Q^3 + 3Q^6, \\
\frac{\eta'''(x)}{\eta'(x)} &= 9\big(1 - \tfrac{3}{5} Q^3(x)\big), & \eta^2(x) &= \big(\tfrac{5}{3}\big)^2 \big(1 - \tfrac{2}{5} Q^3(x)\big), \\
\left(\frac{\eta''(x)}{\eta'(x)}\right)^2 &= 9\big(1 - \tfrac{2}{3} Q^3(x)\big), & |\eta| &\le \tfrac{5}{3}.
\end{aligned}
$$

Proposition 5.1. *If v satisfies the v-KdV equation and v is orthogonal to $\tilde{Q}$ and Q', then there exists some $C > 0$ such that given η as in* (5-3), *we have*

$$
\frac{d}{dt} I_\eta(v) + C \|\operatorname{sech}\big(\tfrac{3}{2}x\big) v\|_{H^1}^2 \le 0.
$$

Proof of Proposition 5.1. Following the formalism presented above, we see

$$
-\frac{d}{dt} I_\eta(v) = -2 \int \mathcal{L}(\partial_x v) v \eta\, dx = 3 \int (\partial_x v)^2 \eta'\, dx + \int v^2 (-\eta''' + \eta' - 4(Q^3\eta)')\, dx.
$$

Selecting

$$
\tilde{w}(t,x) = v(t,x) \sqrt{\eta'(x)}
$$

we see

$$
-\frac{d}{dt} I_\eta(v) = 3 \int (\partial_x \tilde{w})^2\, dx + \int A(x) \tilde{w}^2\, dx,
$$

where

$$A(x) = 1 + \frac{1}{2}\frac{\eta'''}{\eta'} - \frac{3}{4}\left(\frac{\eta''}{\eta'}\right)^2 - 4\frac{(Q^3\eta)'}{\eta'} = \frac{75}{4} - 12Q^3.$$

Hence,

$$-\frac{d}{dt}I_\eta(v) = 3\Big(\langle \mathscr{L}\tilde{w}, \tilde{w}\rangle + \frac{21}{4}\int \tilde{w}^2\,dx\Big).$$

Since $\mathscr{L}\partial_x Q = 0$, we know that given $v = Q$, we have

$$-\frac{d}{dt}I_\eta(v) = 0.$$

However, $v = Q$ corresponds directly to $\tilde{w} = Q^{5/2}$, which is the ground state or $\mathscr{L}$, which has exact eigenvalue $-21/4$. Then, since $\langle Q, \widetilde{Q}\rangle \neq 0$, our orthogonality condition $v \perp \widetilde{Q}$ is enough to guarantee that there exists $C > 0$ such that

$$\tilde{B}(\tilde{w}, \tilde{w}) \geq C\|\tilde{w}\|^2_{H^1} = C\|\sqrt{\eta'}v\|^2_{H^1},$$

which is the desired result. □

We note in the case of more general weight functions η, virial identity methods are still applicable even if perhaps analytic proofs of the virial identities are more challenging.

By choosing the multiplier $\gamma(v - v_{xx})$ with $\gamma = \gamma_0(x - t)$ for γ_0 as in (3-7), we see

$$\begin{aligned}\frac{d}{dt}\int \gamma(v^2 + v_x^2)\,dx = -3\int \gamma' v_x^2\,dx + \int \gamma^{(3)} v^2\,dx - \int \gamma' v^2\,dx \\ + 4\int \gamma' Q^3 v^2\,dx + 12\int \gamma Q^2 Q' v^2\,dx \\ - \int (3\gamma' v_{xx}^2 + \gamma' v_x^2 - \gamma^{(3)} v_x^2)\,dx + \int 4\gamma' Q^3 v_x^2\,dx, \quad (5\text{-}4)\end{aligned}$$

which consists of a number of negative semidefinite terms. All nonnegative semidefinite terms are easily dominated by a multiple of $\|v\|^2_{H^1_{\mathrm{sech}(3x/2)}}$, the term in Proposition 5.1.

Finally, note that by direct computation

$$\partial_t\langle \mathscr{L}^{-1}v, v\rangle = 0. \quad (5\text{-}5)$$

Now, let us define an energy for the solution v of (5-2) to be

$$E(v) = \int \gamma(x)(v^2 + v_x^2)\,dx + \lambda_E\int \eta(x)v^2\,dx + \Lambda_E\langle \mathscr{L}^{-1}v, v\rangle, \quad (5\text{-}6)$$

where $\eta(x)$ is chosen as in (5-3).

Proposition 5.2. *Let us assume v satisfies the v-KdV equation and v is orthogonal to $\widetilde{Q}$ and Q'. There exist λ_E, Λ_E, $\delta > 0$ such that*

$$E(v) \sim \|v\|^2_{H^1} \tag{5-7}$$

$$\frac{d}{dt} E(v) + \delta \|v\|^2_{H^2_{\sqrt{\gamma'}}} \leq 0. \tag{5-8}$$

Proof. From (5-4) and the proof of Proposition 5.1, we see easily one may choose a λ_E that depends only on δ and C so that (5-8) holds for all $\Lambda_2 > 0$. We choose Λ_E large to achieve $E(v) \geq C'' \|v\|^2_{H^1}$. There exists some constant C' such that $E(v) \leq C' \|v\|^2_{H^1}$. Thus the estimate follows given the orthogonality conditions on v. □

The assertions of Proposition 5.2 are robust under suitable perturbations. We turn to the analysis of the time dependent problem

$$v_t - (-\partial_x^2 - 4Q^3_{c(t),y(t)})\partial_x v = \alpha(t) Q_{c(t),y(t)} + \beta(t) Q'_{c(t),y(t)}, \tag{5-9}$$

where

$$\alpha(t) = -\frac{(\dot{c}/c)\langle v, \tilde{\widetilde{Q}}_{c(t),y(t)}\rangle + (\dot{y} - c^2)\langle v, \widetilde{Q}'_{c(t),y(t)}\rangle}{\langle Q_{c(t),y(t)}, \widetilde{Q}_{c(t),y(t)}\rangle}, \tag{5-10}$$

$$\beta(t) = -\frac{(\dot{c}/c)\langle v, \widetilde{Q}'_{c(t),y(t)}\rangle + (\dot{y} - c^2)\langle v, \widetilde{Q}''_{c(t),y(t)}\rangle}{\langle Q'_{c(t),y(t)}, Q'_{c(t),y(t)}\rangle}. \tag{5-11}$$

Here,

$$\tilde{\widetilde{Q}}_{c(t),y(t)} = \tfrac{2}{3} \widetilde{Q}_{c(t),y(t)} + x \widetilde{Q}'_{c(t),y(t)} = c(t)\partial_c \widetilde{Q}_{c(t),y(t)}. \tag{5-12}$$

For simplicity of exposition, in the sequel we suppress the t and y dependence and instead write simply $Q_{c(t),y(t)} = Q_c$ unless we want to stress the dependence on $y(t)$ and t. Similarly we recall

$$\mathcal{L}_c v = \mathcal{L}_{c,y} v = -v_{xx} + c^2 v - 4Q^3_{c,y} v. \tag{5-13}$$

The terms on the right hand side ensure that $\langle v(0), Q'_{c(0),y(0)}\rangle = 0$ implies $\langle v(t), Q'_{c(t),y(t)}\rangle = 0$, and, in addition, $\langle v(0), \widetilde{Q}_{c(0),y(0)}\rangle = 0$ implies $\langle v(t), \widetilde{Q}_{c(t),y(t)}\rangle = 0$. We choose $\gamma(x,t) = \gamma_0(x - y(t))$ and we prove the following:

Proposition 5.3. *There exists a δ, λ, $\Lambda > 0$ such that the following is true: Suppose that*

$$|c(t) - 1| + |\dot{c}(t)| + |\dot{y}(t) - c^2(t)| < \delta \tag{5-14}$$

for all $t \geq 0$ and define

$$E(v) = \int \gamma(x,t)(v^2 + v_x^2)\,dx + \lambda \int \eta_{1,y(t)}(x) v^2\,dx + \Lambda \langle \mathcal{L}^{-1}_{1,y(t)} v, v\rangle, \tag{5-15}$$

where we suppress the dependence of E and v on t. Then

$$E(v) \sim \|v\|^2_{H^1}, \tag{5-16}$$

for all $t > 0$ provided

$$\langle v, Q_c' \rangle = \langle v, \widetilde{Q}_c \rangle = 0. \tag{5-17}$$

Moreover, if v satisfies the system consisting of (5-9), (5-10) *and* (5-11) *and $v(\,\cdot\,,0)$ is orthogonal to $\widetilde{Q}_{c(0),y(0)}$ and $Q'_{c(0),y(0)}$ (which implies the orthogonality for all t) we have*

$$\frac{d}{dt}E(v) + \delta\|v\|^2_{H^2_{\sqrt{\gamma'}}} \le 0. \tag{5-18}$$

Proof. Since $\langle v(t), Q'_{c(t),y(t)} \rangle = 0$, we have

$$\langle \mathcal{L}^{-1}_{c(t),y(t)} v, v \rangle \ge C\|v\|^2_{H^{-1}}$$

for some $C > 0$, as seen in (4-5). Here and in the remaining part of this section we use the Moore–Penrose inverse, which is by an abuse of notation given by the orthogonal projection to the complement of Q', followed by an inversion of $\mathcal{L}$ on this orthogonal subspace. Let us look at a slightly different quantity (where we replace c by 1) given by $\langle \mathcal{L}^{-1}_{1,y(t)} v, v \rangle$. Then, since $\langle v, Q_c' \rangle = \langle v, \widetilde{Q}_c \rangle = 0$, for $|c-1|$ small enough we have

$$\langle \mathcal{L}^{-1}_{1,y(t)} v, v \rangle \ge 2C\|P^{\perp}_{Q'_{1,y(t)}} v\|^2_{H^{-1}} \ge 2C\|v\|^2_{H^{-1}} - C'|c-1|\|v\|^2_{H^{-1}} \ge C\|v\|^2_{H^{-1}}$$

for some constants $C, C' > 0$ and $\delta \le C/C'$. The properties are similar to the previous proposition, but the calculations are more tedious. We consider them to be important for the understanding of the linearization. We recall that we suppress the dependence of Q and $\mathcal{L}$ on y in the notation below. Then we have

$$\begin{aligned}\frac{d}{dt}\langle \mathcal{L}_1^{-1} v, v \rangle &= 2\langle v_t, \mathcal{L}_1^{-1} v \rangle - 12\dot{y}\langle Q_1^2 Q_1' \mathcal{L}_1^{-1} v, \mathcal{L}_1^{-1} v \rangle + 2\dot{y}\frac{\langle v, Q_1' \rangle}{\langle Q_1', Q_1' \rangle}\langle Q''_{1,y}, \mathcal{L}_1^{-1} v \rangle \\ &= 2I_1 - 12I_2 + 2I_3,\end{aligned}$$

where I_2 originates from the differentiation of the inverse and I_3 from the dependence of the implicit projection on time. We have

$$\begin{aligned} I_1 &= \langle \mathcal{L}_c \partial_x v, \mathcal{L}_1^{-1} v \rangle - c^2 \langle \partial_x v, \mathcal{L}_1^{-1} v \rangle + \alpha \langle Q_c', \mathcal{L}_1^{-1} v \rangle + \beta \langle Q_c, \mathcal{L}_1^{-1} v \rangle, \\ \langle \mathcal{L}_c \partial_x v, \mathcal{L}_1^{-1} v \rangle &= \langle (\mathcal{L}_c - \mathcal{L}_1)\partial_x v, \mathcal{L}_1^{-1} v \rangle = (c^2-1)\langle \partial_x v, \mathcal{L}_1^{-1} v \rangle + 4\langle (Q_c^3 - Q_1^3)\partial_x v, \mathcal{L}^{-1} v \rangle, \\ [\mathcal{L}_1^{-1}, \partial_x] &= -\mathcal{L}_1^{-1}[\mathcal{L}_1, \partial_x]\mathcal{L}_1^{-1} = 12\mathcal{L}_1^{-1} Q_1^2 Q_1' \mathcal{L}_1^{-1} + \mathcal{L}_1^{-1}[\partial_x, P_{Q_1'}] + [\partial_x, P_{Q_1'}]\mathcal{L}_1^{-1}, \\ [\partial_x, P_{Q_1'}]v &= -\frac{\langle v, Q_1' \rangle}{\langle Q_1', Q_1' \rangle} Q_1'' + \frac{\langle v, Q_1'' \rangle}{\langle Q_1', Q_1' \rangle} Q_1', \\ \langle \partial_x v, \mathcal{L}_1^{-1} v \rangle &= \tfrac{1}{2}\langle v, [\mathcal{L}_1^{-1}, \partial_x] v \rangle = 6\langle Q_1^2 Q_1' \mathcal{L}_1^{-1} v, \mathcal{L}_1^{-1} v \rangle, \\ \langle v, Q_1' \rangle &= \langle v, Q_1' - Q_c' \rangle \end{aligned}$$

by the orthogonality conditions and

$$\langle \mathcal{L}_1^{-1} v, Q_c' \rangle = \langle \mathcal{L}_1^{-1} v, Q_c' - Q_1' \rangle \quad \text{and} \quad \langle \mathcal{L}_1^{-1} v, Q_c \rangle = \langle \mathcal{L}_1^{-1} v, (\mathcal{L}_1^{-1} - \mathcal{L}_c^{-1}) Q_c \rangle$$

because of the orthogonality conditions and since $\mathcal{L}_c^{-1} Q_c = \widetilde{Q}_c$.

Altogether, and applying Lemma 4.1, we have

$$\left| \frac{d}{dt} \langle \mathcal{L}_1^{-1} v, v \rangle \right| \leq \mathbb{O}(|c^2 - 1| + |\dot{y} - 1| + |\dot{c}|) \|v\|^2_{H^{-1/2}_{\sqrt{\gamma'}}}, \tag{5-19}$$

which we will control by the virial identity below.

We now look at virial weights of the form

$$\eta(x, t) = -\frac{5}{3} \frac{Q_1'(x - y(t))}{Q_1(x - y(t))} = \tfrac{5}{3} \tanh\big(\tfrac{3}{2}(x - y(t))\big),$$

which has properties similar to those of $\eta(x)$ with appropriate changes for the unit scaling.

We have defined v such that

$$v(x, t) \perp \widetilde{Q}_{c(t), y(t)} \quad \text{and} \quad v(x, t) \perp Q'_{c(t), y(t)} \quad \text{for all } t \geq 0.$$

Following the formalism presented above and in [Martel 2006], select

$$\tilde{w}(t, x) = v(t, x) \sqrt{\eta'(x)}.$$

Then,

$$-\frac{d}{dt} I_\eta(v) = 3 \int (\partial_x \tilde{w})^2 \, dx + \int A(x) \tilde{w}^2 \, dx - 2 \int \eta_{1,y} \Big(\beta(t) Q'_{c,y} + \alpha(t) Q_{c,y}\Big) v \, dx$$
$$+ \tfrac{3}{2} c^2 (\dot{y} - c^2) \int \operatorname{sech}^2\big(\tfrac{3}{2} c(x - y(t))\big) v^2 \, dx + \tfrac{3}{2} c^4 \int \operatorname{sech}^2\big(\tfrac{3}{2} c(x - y(t))\big) v^2 \, dx,$$

where

$$A(x, t) = 1 + \frac{1}{2} \frac{\eta'''}{\eta'} - \frac{3}{4} \Big(\frac{\eta''}{\eta'}\Big)^2 - 4 \frac{(Q^3 \eta)'}{\eta'} = \frac{75}{4} - 12 Q^3.$$

Hence,

$$-\frac{d}{dt} I_\eta(v) > 3\Big(\langle \mathcal{L} \tilde{w}, \tilde{w} \rangle + \tfrac{21}{4} c^2 \int \tilde{w}^2 \, dx\Big) + \mathbb{O}(|1^2 - c^2| + |\dot{y} - c^2|) \|v\|^2_{L^2_{\sqrt{\gamma'}}}.$$

From above, we know that $3\big(\langle \mathcal{L} \tilde{w}, \tilde{w} \rangle + \frac{21}{4} \int \tilde{w}^2 \, dx\big) = 0$ for $v = Q_{1,y(t)}$. This corresponds to $\tilde{w} = Q^{5/2}_{1,y(t)}$, which is the ground state or $\mathcal{L}_{1,y(t)}$. Hence, $v = Q$ is the ground state of the quadratic form

$$3\Big(\langle \mathcal{L} \tilde{w}, \tilde{w} \rangle + \tfrac{21}{4} \int w^2 \, dx\Big).$$

From Lemma 4.1, our orthogonality condition $v \perp \widetilde{Q}_{c(t), y(t)}$ is enough to guarantee there exists $\delta > 0$ such that

$$\frac{d}{dt} I_\eta(v) + \|v\|^2_{H^1_{\sqrt{\gamma'}}} \leq 0$$

provided $|c^2 - 1^2| + |\dot{y} - c^2|$ is small for all $t \geq 0$, which follows from our assumptions on the initial perturbation.

The time-dependent version of

$$\frac{d}{dt}\int \gamma(x,t)(v^2+v_x^2)\,dx$$

is done in full generality in the analysis of (3-13) in Section 3 for the Airy equation. The terms that we have to control are the same as for constant c and $\dot{y}$, plus the terms coming from the right hand side. Those are easy to control. Namely,

$$2\left|\int \big[(\alpha Q+\beta Q')\gamma-\partial_x(\gamma(\alpha Q'+\beta Q''))\big]\,v\,dx\right| \lesssim (|\dot{c}|+|\dot{y}-c^2|)\||v|^2\|_{L^2_{\sqrt{\gamma'}}}$$

for γ as in Section 3. □

Note, above we have always assumed the proper orthogonality conditions, but without them we easily obtain the following estimate for solutions of the v equation:

$$\|v\|_{L^\infty H^1\cap L^2 H^2_{\sqrt{\gamma'}}} \le C\big(\|v(0)\|_{H^1}+\sup_t|\langle v(\,\cdot\,,t),\, Q'_{1,y(t)}\rangle|+\|\langle v(\,\cdot\,,t),\, \widetilde{Q}_{1,y(t)}\rangle\|_{L^2([0,\infty)}\big). \tag{5-20}$$

6. Function spaces and projection operators

In this section we construct the function spaces for our nonlinear analysis using properties of the linear evolution we studied in Sections 3–5. Based on the energy functional (5-6) for the v-equation, it seems natural to look at

$$v\in X^1=L^\infty H^1\cap L^2 H^2_{\sqrt{\gamma'}},$$

where $\gamma=\gamma_0(x)$ is as in (3-5), and again by convention we set L^pX to be the L^p norm in time of the X norm in space.

Then, as follows naturally from the equation, we define $Y^1=L^1H^1+L^2\sqrt{\gamma'}L^2$.

Generically, we define

$$X^s=L^\infty H^s\cap L^2 H^{s+1}_{\sqrt{\gamma'}} \quad\text{and}\quad Y^s=L^1H^s+L^2\sqrt{\gamma'}H^{s-1},$$

where we note $(Y^s)^*=X^{-s}$.

6.1. ***The scale of energy spaces.*** Let us study the v-equation

$$\begin{cases}(\partial_t-\mathcal{L}\partial_x)v=f_0+\sqrt{\gamma'}f_1=f,\\ v(0,x)=v_0,\end{cases} \tag{6-1}$$

where $f_0\in L^1H^s$, $f_1\in L^2L^{s-1}$ and $v_0\in H^s$. We assume that the orthogonality conditions

$$v_0\perp Q',\quad v_0\perp \widetilde{Q} \tag{6-2}$$

and

$$(f_0+\sqrt{\gamma'}f_1)\perp\widetilde{Q},\quad (f_0+\sqrt{\gamma'}f_1)\perp Q' \quad\text{for all } t \tag{6-3}$$

hold.

Proposition 6.1. *There exists a unique solution $v \in X^s$ that satisfies*

$$\|v\|_{X^s} \le c(\|v_0\|_{H^s} + \|f_0 + \sqrt{\gamma'} f_1\|_{Y^s}).$$

Moreover, $v(t)$ is orthogonal to Q' and $\widetilde{Q}$.

Note, Theorem 1 is an immediate consequence.

Proof. We begin by considering the case $s = 1$. The previous section implies the estimate

$$\|v\|_{X^s} \le c(\|v_0\|_{H^s} + \|f_0\|_{L^1(H^s)}).$$

if $f_1 = 0$ by a variation of constants argument. We retrace the steps and its modifications needed for f_1. Using the multipliers from the energy inequalities, we need the obvious estimates

$$\left|\iint f_0 \gamma (v - v_{xx})\, dx\, dt\right| + \left|\iint f_0 \eta v dx\, dt\right| + \left|\iint f_0 \mathscr{L}^{-1} v\, dx\, dt\right| \le c\|v\|_{L^\infty H^1} \|f_0\|_{L^1 H^1}$$

and, using Lemma 4.1,

$$\left|\iint \sqrt{\gamma'} f_1 (\gamma (v - v_{xx})\, dx\, dt\right| + \left|\iint \sqrt{\gamma'} f_1 \eta v\, dx\, dt\right| + \left|\iint \sqrt{\gamma'} f_1 \mathscr{L}^{-1} v\, dx\, dt\right| \le c\|f_1\|_{L^2} \|v\|_{L^2 H^2_{\sqrt{\gamma'}}}.$$

It is not hard to see that $v(t)$ remains orthogonal to Q' and $\widetilde{Q}$ so that we can close the argument as in the previous section. We obtain the desired estimate for $s = 1$:

$$\|v\|_{X^1} \le c(\|v_0\|_{H^1} + \|f_0 + \sqrt{\gamma'} f_1\|_{Y^1}).$$

We denote the solution operator for the inhomogeneous v-problem (u-problem) by S_v (S_u) and we write

$$\|S_v f\|_{X^1} \le c\|f\|_{Y^1}. \tag{6-4}$$

The role of the two orthogonality conditions are different: The equation is invariant under the addition of a multiple of Q to v, and orthogonality to Q' is conserved. Orthogonality to $\widetilde{Q}$ was needed for the virial identity of Martel and Merle, whereas orthogonality of v and Q' entered the control of the H^{-1} norm by the Moore–Penrose inverse of $\mathscr{L}$. Without orthogonality one still obtains (5-20).

Suppose now that v satisfies

$$\begin{cases} v_t - \mathscr{L}\partial_x v = f, \\ v(x, 0) = v_0. \end{cases} \tag{6-5}$$

Let ε be a small constant. We apply $(1 + \varepsilon^2 D^2)^{(s-1)/2}$ to both sides of the equation and denote $v^s = (1 + \varepsilon^2 D^2)^{(s-1)/2} v$. It satisfies

$$v^s_t - \mathscr{L}\partial_x v^s = (1 + \varepsilon^2 D^2)^{(s-1)/2} f + [(1 + \varepsilon^2 D^2)^{(s-1)/2}, 4Q^3]\partial_x v.$$

Hence, applying (5-20)

$$\begin{aligned} \|v\|_{X^s} &\le c_1 \|v^s\|_{X^1} \\ &\le c_2\big(\|(1+\varepsilon^2 D^2)^{(s-1)/2} f\|_{Y^1} + \|[(1+\varepsilon^2 D^2)^{(s-1)/2}, 4Q^3]\partial_x v\|_{Y^1} + \sup_t |\langle v^s, Q'\rangle| + \|\langle v^s, \widetilde{Q}\rangle\|_{L^2}\big). \end{aligned}$$

and we turn to the commutator term.

Lemma 6.2. *Let $\phi \in C^\infty(\mathbb{R})$ satisfy $|\phi| + |\phi'| \le Ce^{-|x|}$. Let $k(x, y)$ be the kernel of the operator*

$$[(1+\varepsilon^2 D^2)^{s/2}, \phi](1+\varepsilon^2 D^2)^{-s/2}.$$

Then,

$$|k(x, y)| \le c\varepsilon |s| e^{-(|x|+|y|)/4-|x-y|/(4\varepsilon)}.$$

We postpone its proof. By Lemma 6.2 (with $\phi = 4Q^3$ and $s-1$) and Schur's Lemma

$$\begin{aligned}\|[(1+\varepsilon^2 D^2)^{(s-1)/2}, 4Q^3]\partial_x v\|_{Y^1} &\le \|(\gamma')^{-1/2}[(1+\varepsilon^2 D^2)^{(s-1)/2}, 4Q^3](1+\varepsilon^2 D^2)^{(1-s)/2}\partial_x v_s\|_{L^2} \\ &\le c\varepsilon \|(\gamma')^{1/2}\partial_x v^s\|_{L^2},\end{aligned} \tag{6-6}$$

and by Lemma 3.1, after rescaling, as for the constant coefficient equation, we have

$$\|(1+\varepsilon^2 D^2)^{(s-1)/2} f\|_{Y^1} \le c\|f\|_{Y^s}.$$

For all Schwartz functions,

$$\big\|(1+|x|^2)^N\big((1+\varepsilon^2 D^2)^{s/2}\phi - \phi\big)\big\|_{L^2} \le C\varepsilon.$$

If $\langle v, \widetilde{Q}\rangle = \langle v, Q'\rangle = 0$, then

$$|\langle v^s, \widetilde{Q}\rangle| = \big|\langle v, \widetilde{Q}\rangle - \langle v^s, \widetilde{Q} - (1+\varepsilon^2 D^2)^{-\frac{s}{2}}\widetilde{Q}\rangle\big| \le C\varepsilon\|\gamma'^{1/2} v^s\|_{L^2}, \tag{6-7}$$

$$|\langle v^s, Q'\rangle| = \big|\langle v, Q'\rangle - \langle v^s, Q' - (1+\varepsilon^2 D^2)^{-\frac{s}{2}} Q'\rangle\big| \le C\varepsilon\|v^s\|_{L^2}. \tag{6-8}$$

Suppose that $\langle f, Q'\rangle = \langle f, \widetilde{Q}\rangle = 0$. Then we obtain for all $s \in \mathbb{R}$ from (6-6), (6-7) and (6-8)

$$\|v^s\|_{X^s} \le c(\|f\|_{Y^s} + \varepsilon\|v^s\|_{X^s})$$

and hence

$$\|v\|_{X^s} \lesssim \|f\|_{Y^s}, \tag{6-9}$$

which again implies for solutions v to $v_t - \mathscr{L}\partial_x v = P_{Q'} f$, given by the variation of constants formula, the bound $\|\tilde{P}^* v\|_{X^s} \le C\|f\|_{Y^s}$ or equivalently (recall (1-11))

$$\|\tilde{P}^* S_v P_{Q'}^{\perp}\|_{Y^s \to X^s} \lesssim 1. \tag{6-10}$$

Using spacetime duality, we consider

$$\begin{cases}(\partial_t - \partial_x \mathscr{L})u = g,\\ u(0, x) = 0.\end{cases}$$

The estimate adjoint to (6-10) is

$$\|P_{Q'}^{\perp} S_u \tilde{P}\|_{Y^s \to X^s} \lesssim 1, \tag{6-11}$$

which completes the proof. □

Proof. We turn to the proof of Lemma 6.2.

Let $\hat{\phi}$ be the Fourier transform of ϕ, which, because of the exponential decay extends to a holomorphic function $\hat{\phi}$ in the strip $\{z : |\operatorname{Im} z| < 1\}$. Moreover there exists C such that

$$\int |\hat{\phi}(\xi + i\sigma)|\, d\xi \le C \quad \text{if } |\sigma| \le \tfrac{1}{2}.$$

This estimate in turn implies exponential decay. Let $k(x, y)$ be the integral kernel of

$$[(1+\varepsilon^2 D^2)^{s/2}, \phi](1+\varepsilon^2 D^2)^{-s/2}.$$

We claim

$$|k(x, y)| \le c_N \varepsilon |s| e^{-\delta(|x|+|y|)} e^{-\delta|x-y|/\varepsilon}, \tag{6-12}$$

which implies Lemma 6.2.

The symplectic Fourier transform

$$\hat{k}(\xi, \eta) = \frac{1}{2\pi} \int e^{-i\xi x + iy\eta} k(x, y)\, dx\, dy$$

satisfies

$$\hat{k}(\xi, \eta) = \left(\left(\frac{1+\varepsilon^2\xi^2}{1+\varepsilon^2\eta^2} \right)^{s/2} - 1 \right) \hat{\phi}(\xi - \eta).$$

We set $a = \varepsilon(\xi+\eta)/2$ and $b = (\xi-\eta)/2$. Then $\hat{k}(\xi, \eta) = \hat{g}(a, b)$, where

$$\hat{g}(a, b) = \left(\left(\frac{1+(\varepsilon b + a)^2}{1+(\varepsilon b - a)^2} \right)^{s/2} - 1 \right) \hat{\phi}(2b)$$

and

$$\begin{aligned} k(x, y) &= (2\pi)^{-1} \int e^{i(x\xi - y\eta)} \hat{g}(\varepsilon(\xi+\eta)/2, (\xi-\eta)/2))\, d\xi\, d\eta \\ &= 2\varepsilon(2\pi)^{-1} \int e^{i(\frac{x-y}{\varepsilon} a + b(x+y))} \hat{g}(a, b)\, da\, db =: 2\varepsilon g((x-y)/\varepsilon, x+y). \end{aligned}$$

The function $\hat{g}$ expands to a holomorphic function in a to the strip $\{z : |\operatorname{Im} z| < 1/2\}$ if $\varepsilon |\operatorname{Im} b| < 1/2$. Clearly,

$$\frac{1+(a+\varepsilon b)^2}{1+(a-\varepsilon b)^2} = 1 + \frac{4(\varepsilon b)^2}{1+(a-\varepsilon b)^2} + 4\varepsilon b \frac{a - \varepsilon b}{1+(a-\varepsilon b)^2},$$

and hence we define the error term h by the right hand side of

$$\left(\frac{1+(\varepsilon b + a)^2}{1+(\varepsilon b - a)^2} \right)^{s/2} - 1 = 2s\varepsilon b \frac{a - \varepsilon b}{1+(\varepsilon b - a)^2} + h(\varepsilon b, a).$$

It satisfies

$$|h(\varepsilon b, a)| \le c s^2 \varepsilon^2 |b|^2 (1 + |\varepsilon b - a|)^{-2} \quad \text{if } |\varepsilon \operatorname{Im} b + a| \le 1/2.$$

Hence,

$$\left| \int e^{i(av+bw)} h(\varepsilon b, a) \hat{\phi}(2b)\, da\, db \right| \le c s^2 \varepsilon^2 e^{-(|v|+|w|)/4}$$

by the extension of a and b to a suitable complex strip. The leading term contributing to g can be calculated:

$$\begin{aligned} g_0(v,w) &= (2\pi)^{-1}\int e^{i(av+bw)}\frac{a-\varepsilon b}{1+(a-\varepsilon b)^2} b\hat{\phi}(2b)\,da\,db \\ &= i\frac{v}{|v|}e^{-|v|}\int e^{i(b(w+\varepsilon v))} b\hat{\phi}(2b)\,db = \sqrt{\pi}2i\frac{v}{|v|}e^{-|v|}\phi'((w+\varepsilon v)/2)). \end{aligned}$$

The leading term for k is

$$k_0(x,y) = \sqrt{\pi}2i\varepsilon\frac{x-y}{|x-y|}e^{-|x-y|/\varepsilon}\phi'(x). \qquad \square$$

6.2. ***U and V space estimates.*** In this section, we generalize and improve Theorem 1 using the U^p and V^p spaces as defined in [Hadac et al. 2009] and in the appendix. For notational simplicity, let us define

$$U^p = U^p_{KdV} \quad \text{and} \quad V^p = V^p_{KdV}.$$

We begin with a number of estimates that we will use often in the sequel.

Let $c, y \in C^1$ satisfy (5-14) and let $\gamma(x,t) = \gamma_0(x - y(t))$. Then,

$$\|aQ'_{c(t),y(t)} + b\widetilde{Q}_{c(t),y(t)}\|_{Y^0} \lesssim \|a\|_{L^2+L^1} + \|b\|_{L^2+L^1};$$

hence

$$\|P^{\perp}_{Q'}\tilde{P}f\|_{DU^2+L^2\sqrt{\gamma'}H^{-1}} \lesssim \|f\|_{DU^2} + \|\langle f, Q\rangle\|_{L^2+L^1} + \|\langle f, Q'\rangle\|_{L^2+L^1}.$$

We consider

$$w_t + w_{xxx} = f, \quad \text{with } w(0) = u_0.$$

Then,

$$\|w\|_{U^2} \lesssim \|u_0\|_{L^2} + \|f\|_{DU^2}$$

and, since $U^2 \subset L^2H^1_{\sqrt{\gamma'}}$,

$$\|\langle w, Q\rangle\|_{L^2} + \|\langle w, \widetilde{Q}\rangle\|_{L^2} \lesssim \|f\|_{DU^2} + \|u_0\|_{L^2}.$$

Hence, with $v = \tilde{P}P^{\perp}w$, we have

$$\|v\|_{L^2H^1_{\sqrt{\gamma'}}} \lesssim \|f\|_{DU^2} + \|u_0\|_{L^2}.$$

We calculate

$$\begin{aligned} (\partial_t + c^2 - \partial_x\mathcal{L})\Big(\frac{\langle w,Q\rangle}{\langle Q,\widetilde{Q}\rangle}\widetilde{Q} + \frac{\langle w,Q'\rangle}{\langle Q',\widetilde{Q}'\rangle}Q'\Big) &= \frac{\dot{c}}{c}\frac{\langle w,Q\rangle}{\langle Q,\widetilde{Q}\rangle}\widetilde{\widetilde{Q}} + (\dot{y}-c^2)\frac{\langle w,Q'\rangle}{\langle Q',\widetilde{Q}'\rangle}Q'' \\ &\quad + \Big((\dot{y}-c^2)\frac{\langle w,Q\rangle}{\langle Q,\widetilde{Q}\rangle} + \frac{\dot{c}}{c}\frac{\langle w,Q'\rangle}{\langle Q',\widetilde{Q}'\rangle}\Big)\widetilde{Q}' - \tilde{\alpha}\widetilde{Q} - \tilde{\beta}Q', \end{aligned}$$

where $\tilde{\alpha}$ and $\tilde{\beta}$ are the time derivatives of the coefficients of $\widetilde{Q}$ and Q' and

$$\widetilde{\widetilde{Q}} = (x-y)\widetilde{Q}' + \tfrac{2}{3}\widetilde{Q}.$$

Hence, assuming that u_0 satisfies the orthogonality conditions, that w and v are as above, and with g defined through the previous calculations,

$$\partial_t v + c^2 v_x - \partial_x \mathcal{L}_{c,y} v = \tilde{\alpha}\widetilde{Q} + \tilde{\beta} Q' + g + f, \quad \text{with } v(0) = u_0, \tag{6-13}$$

where we collect the properties of v and g in the following.

Lemma 6.3. *Assuming* (5-14), *we have* $\langle v(t), Q\rangle = \langle v(t), Q'\rangle = 0$ *and*

$$\|v\|_{V^2 \cap L^2 H^1_{\sqrt{\gamma'}}} + \|g\|_{L^2\sqrt{\gamma'}H^{-1}} \lesssim \|u_0\|_{L^2} + \|f\|_{DU^2} + \|\langle f, Q\rangle\|_{L^2+L^1} + \|\langle f, Q'\rangle\|_{L^2+L^1}. \tag{6-14}$$

Proof. We claim that

$$\|\tilde{\alpha}\|_{L^1+L^2} + \|\tilde{\beta}\|_{L^1+L^2} \le c(\|w_0\|_{L^2} + \|f\|_{DU^2}),$$

the proof of which we postpone. Assuming its validity we put the term $4\partial_x Q v$ in (6-13) on the right hand side. We bound $\|v\|_{V^2}$ in terms of $\|w_0\|_{L^2}$ and the right hand side in DV^2. Since $DU^2 \subset DV^2$ and $L^2\sqrt{\gamma'}H^{-1} \subset DV^2$, we can control all terms on the right hand side.

The only missing piece is the $L^2 + L^1$ bound for α and β. There are two different arguments: Either we can follow the calculation above and calculate $\tilde{\alpha}$ and $\tilde{\beta}$ above, or we can test by Q and Q' and use orthogonality to obtain the standard equations for α and β. We use the first approach and recall the calculations after (1-10). Then

$$\begin{aligned} \frac{d}{dt}\langle w, Q\rangle &= \langle f, Q\rangle + \dot{y}\langle w, Q'\rangle + \frac{\dot{c}}{c}\langle w, \widetilde{Q}\rangle, \\ \frac{d}{dt}\langle w, Q'\rangle &= \langle f, Q'\rangle + \dot{y}\langle w, Q''\rangle + \frac{\dot{c}}{c}\langle w, \widetilde{Q}'\rangle. \end{aligned} \tag{6-15}$$

There is one more term entering the coefficient of Q' coming from applying the linear operator to $\widetilde{Q}$, which gives

$$-2\frac{\langle w, Q\rangle}{\langle Q, \widetilde{Q}\rangle} c^2 Q'.$$

All these terms are easily controlled. □

We return to the analysis of the time dependent v-problem

$$v_t + c^2 c_x - \mathcal{L}\partial_x v = \alpha(t) Q_{c(t),y(t)} + \beta(t) Q'_{c(t),y(t)} + f, \tag{6-16}$$

where

$$\alpha(t) = -\frac{(\dot{c}/c)\langle v, \tilde{\widetilde{Q}}_{c(t),y(t)}\rangle + (\dot{y} - c^2)\langle v, \widetilde{Q}'_{c(t),y(t)}\rangle}{\langle Q_{c(t),y(t)}, \widetilde{Q}_{c(t),y(t)}\rangle} - \frac{\langle \widetilde{Q}, f\rangle}{\langle Q, \widetilde{Q}\rangle}, \tag{6-17}$$

$$\beta(t) = -\frac{(\dot{c}/c)\langle v, \widetilde{Q}'_{c(t),y(t)}\rangle + (\dot{y} - c^2)\langle v, \widetilde{Q}''_{c(t),y(t)}\rangle}{\langle Q'_{c(t),y(t)}, Q'_{c(t),y(t)}\rangle} - \frac{\langle Q', f\rangle}{\langle Q', Q'\rangle}, \tag{6-18}$$

with the initial data $v(x, 0) = v_0(x)$ orthogonal to $\widetilde{Q}$ and Q'. Then also $v(t)$ satisfies these orthogonality conditions. We combine the arguments of the previous subsection with those of Proposition 5.3:

Lemma 6.4. *Suppose that* (5-14) *holds. There exists a unique solution v to* (6-16) *and* (6-17) *and* (6-18) *that satisfies*

$$\|\langle D\rangle^s v\|_{X^0\cap V^2} \le c(\|v_0\|_{H^s} + \|\langle D\rangle^s f\|_{Y^0+DV^2}).$$

Moreover $v(t)$ is orthogonal to Q' and $\widetilde{Q}$.

Proof. We begin with $s=0$. We write $f = f_U + f_Y$ with $f_U \in DU^2$ and $f_Y \in Y^0$. Let $\tilde{v}$ be defined with $f = f_U$ as in Lemma 6.3. It satisfies

$$\|\tilde{v}\|_{V^2\cap X^0} \le C(\|f_U\|_{DU^2} + \|u_0\|_{L^2}).$$

Let us take $v = \tilde{v} + w$, where w satisfies

$$w_t + w_x - \partial_x \mathscr{L} w = \alpha \widetilde{Q} + \beta Q' + f_Y + g, \quad \text{with } w(0) = 0,$$

with g as in Lemma 6.3 and by Lemma 6.4

$$\|w\|_{X^0} \lesssim \|f_Y\|_{Y^0} + \|g\|_{Y^0} \lesssim \|f\|_{DU^2+Y^0} + \|u_0\|_{L^2}.$$

We put the term $4\partial_x(Q^3 w)$ to the right hand side, which we easily control in Y^0 as well as α and β and we arrive at

$$\|v\|_{V^2\cap X^0} \le C\big(\|v_0\|_{L^2} + \|f\|_{Y^0+DU^2} + \|\langle f, Q\rangle\|_{L^2+L^1} + \|\langle f, Q'\rangle\|_{L^2+L^1}\big). \tag{6-19}$$

The case of general s follows by the same arguments as above. □

Our main interest will be in similar estimates for the u problem below.

We consider the u equations

$$u_t + c^2 u_x - \partial_x(\mathscr{L}_{c,y} u) = \alpha \widetilde{Q} + \beta Q' + f, \tag{6-20}$$

with initial data $u(0) = u_0$ that satisfies $\langle u_0, Q\rangle = \langle u_0, Q'\rangle = 0$, together with the modal equations

$$\alpha(t) = -\frac{(\dot{c}/c)\langle u, \widetilde{Q}\rangle + \langle f, Q\rangle}{\langle Q, \widetilde{Q}\rangle}, \tag{6-21}$$

$$\beta(t) = -\frac{(\dot{y} - c^2)\langle u, Q''\rangle + (\dot{c}/c)\langle u, \widetilde{Q}'\rangle + \langle u, \mathscr{L} Q_{xx}\rangle + \langle f, Q'\rangle}{\langle Q', Q'\rangle}, \tag{6-22}$$

which again ensures the orthogonality of $u(t)$ with Q and Q'.

We obtain first the analog of Lemma 6.4.

Lemma 6.5. *Suppose that* (5-14) *holds. There exists a unique solution u to* (6-20), (6-21) *and* (6-22) *that satisfies*

$$\|u\|_{X^0\cap U^2} \le c(\|u_0\|_{L^2} + \|f\|_{Y^0+DV^2}).$$

Moreover, *$u(t)$ is orthogonal to Q' and Q.*

It is not difficult to construct solutions; however we are interested in global estimates. Moreover we may restrict to a finite time interval and assume that all the data as well as u are smooth and decay at infinity.

We set $v = \mathscr{L}u$. It satisfies the orthogonality conditions

$$\langle v, Q' \rangle = 0 = \langle u, Q \rangle = \langle \mathscr{L}^{-1} v, Q \rangle = \langle v, \widetilde{Q} \rangle.$$

Moreover, v satisfies

$$v_t + c^2 v_x - \mathscr{L}\partial_x v = -2c^2 \alpha Q + 12 Q^2 ((\dot{c}/c)\widetilde{Q} + (\dot{y} - c^2) Q') u + \mathscr{L} f$$

and we may apply Lemma 6.4 with $s = -2$:

$$\|v\|_{X^{-2}} \lesssim \|\mathscr{L}u(0)\|_{H^{-2}} + \|\mathscr{L}f\|_{Y^{-2}} + (|\dot{c}/c| + |\dot{y} - c^2|) \|u\|_{L^2(H^{-3}_{\sqrt{\gamma'}})}.$$

We apply Lemma 4.1 several times to get

$$\|u\|_{X^0} \lesssim \|\mathscr{L}v\|_{X^{-2}} \lesssim \|u_0\|_{L^2} + \|f\|_{Y^0} + \sup_t (|\dot{c}/c| + |\dot{y} - c^2|) \|u\|_{X^{-1}}.$$

To complete the proof we observe that by (5-14) we may subtract the last term on the right hand side from both sides to arrive at the desired estimate. The inclusion of V^2 and DU^2 works now exactly as for the v equation.

We collect the results for the case $s = 0$, which is the only estimate we will need later on.

Proposition 6.6. *Suppose that* (5-14) *holds. There exists a unique solution v to* (6-16), (6-17) *and* (6-18) *that satisfies*

$$\|v\|_{V^2 \cap X^0} \le c(\|v_0\|_{L^2} + \|f\|_{DU^2 + Y^0}).$$

Moreover, $v(t)$ is orthogonal to Q' and $\widetilde{Q}$. Similarly there is a unique solution u to (6-20), (6-21) *and* (6-22) *that satisfies*

$$\|u\|_{V^2 \cap X^0} \le c(\|u_0\|_{L^2} + \|f\|_{DU^2 + Y^0}).$$

Moreover, $u(t)$ is orthogonal to Q' and Q.

6.3. *Littlewood–Paley decomposition.* We consider functions c and y satisfying (5-14) We set $\lambda \in \Lambda_0 = 1.01^{\mathbb{N}}$ and let P_λ be the Littlewood–Paley decomposition with Fourier multipliers supported in the set $\{\xi : 1.01^{-1}\lambda \le |\xi| \le 1.01\lambda\}$ if $\lambda > 1$ and $\{\xi : |\xi| \le 1\}$ if $\lambda = 0$. Then, we denote

$$u_\lambda = P_\lambda u.$$

The Besov spaces are defined as the set of all tempered distributions for which the norm

$$\|v\|_{B^{s,p}_q} = \|\lambda^s \|v_\lambda\|_{L^p}\|_{l^q(\Lambda_0)}$$

is finite. Here $s \in \mathbb{R}$ and $1 \le p, q \le \infty$. Similarly we define the homogeneous spaces $\dot{B}^{s,p}_q$ with the summation over $\Lambda = 1.01^{\mathbb{Z}}$, where the frequency $\lambda = 1$ plays no special role. There is an ambiguity about the meaning of v_0, which differs depending on whether we consider $B^{s,p}_q$ or the homogeneous space $\dot B^{s,p}_q$.

We define the spaces X^s_∞ and Y^s_∞ using the norms

$$\|u\|_{X^s_\infty} = \sup_{\lambda\in\Lambda_0} \lambda^s \|u_\lambda\|_{V^2\cap X^0} \quad \text{and} \quad \|f\|_{Y^s_\infty} = \sup_{\lambda\in\Lambda_0} \lambda^s \|f_\lambda\|_{DU^2+Y^0}.$$

The homogeneous spaces $\dot X^s_\infty$ and $\dot Y^s_\infty$ are defined in the same way as the homogeneous Besov space $\dot B^{s,p}_q$ with $\Lambda = 1.01^{\mathbb{Z}}$, though with a slight modification for $s<0$ in the Y spaces due to the ρ multiplier. Namely, we take

$$\begin{aligned} \|u\|_{\dot X^s_\infty} &= \sup_{\lambda\in\Lambda} \left(\lambda^s \|u_\lambda\|_{V^2\cap X^0}\right), \\ \|F\|_{\dot Y^s_\infty} &= \inf_{F=f+g} \Big(\sup_{\lambda\in\Lambda} \lambda^s \|f_\lambda\|_{DU^2} + \sup_{\lambda\in\Lambda_0} \lambda^s \|g_\lambda\|_{Y^0}\Big), \end{aligned} \tag{6-23}$$

where there is a slight abuse of notation since the operators in f_0 and g_0 are taking on two different meanings, the homogeneous projection for f_0 and the inhomogeneous projection for g_0.

We study

$$\begin{cases} u_t + u_x + \partial_x \mathscr{L} u = \alpha \widetilde{Q} + \beta Q_x + f + \partial_x(\rho g), \\ u(x,0) = 0, \end{cases} \tag{6-24}$$

where α is given by (6-21) and β by (6-22). As a first step we obtain a weighted L^2 bound for u in (6-25) below.

Let $f = f^+ + f^-$ and $g = g^+ + g^-$ be a decomposition into high ($|\xi|>1$) and low ($|\xi|\le 1$) frequencies. We define

$$\begin{cases} v_t + c^2 v_x - \mathscr{L} v_x = \alpha_+ Q + \beta_+ Q' + (\partial_x^{-1} f^+ + \rho g^+), \\ v(x,0) = 0, \end{cases}$$

where

$$\begin{aligned} \langle Q, \widetilde{Q}\rangle \alpha_+ &= (c^2 - \dot y)\langle v, \tilde{Q}'\rangle - (\dot c/c)\langle v, \widetilde{\widetilde{Q}}\rangle - \langle \partial_x^{-1} f^+ + \rho g^+, \widetilde{Q}\rangle, \\ \langle Q', Q'\rangle \beta_+ &= (c^2 - \dot y)\langle v, Q''\rangle - (\dot c/c)\langle v, \widetilde{Q}'\rangle + \langle f^+ + \partial_x(\rho g^+), Q\rangle \end{aligned}$$

ensure $\tilde P^* P^\perp_{Q'} v = 0$. Then by Proposition 6.6

$$\|v\|_{X^0} \lesssim \|\partial_x^{-1} f^+ + \rho g^+\|_{DU^2+Y^0} \lesssim \|F^+\|_{\dot Y^s_\infty} \lesssim \|F\|_{\dot Y^s_\infty},$$

where the second inequality holds for all $s > -1$.

As a simple consequence, we obtain

$$\|P_{Q'} \partial_x v\|_{L^2 L^2_\rho} \lesssim \|F\|_{\dot Y^s}$$

and compute similar to arguments above

$$\begin{aligned} &(\partial_t - c^2\partial_x + \partial_x \mathscr{L})(P_{Q'}\partial_x v) \\ &\qquad = \left(\alpha_+ - \frac{d}{dt}\frac{\langle v, Q''\rangle}{\langle Q', Q'\rangle}\right) Q' + \left(\beta_+ + \frac{\langle v, Q''\rangle}{\langle Q', Q'\rangle}(\dot y - c^2)\right) Q'' + \frac{\langle v, Q''\rangle}{\langle Q', Q'\rangle}\frac{\dot c}{c}\widetilde{Q}' + f_+ + \partial_x(\rho g_+). \end{aligned}$$

We make the ansatz $u = P_{Q'}\partial_x v + u_-$ and observe that $\langle \partial_x v, Q\rangle = 0$ by construction. Then,

$$\begin{aligned}\partial_t u_- + c^2\partial_x u_- - \partial_x \mathscr{L} u_- \\ = \alpha \widetilde{Q} + \beta Q' + f_- + \partial_x(\rho g_-) + \Big(\frac{\dot c}{c}\frac{\langle v, \widetilde{Q}'\rangle}{\langle Q', Q'\rangle} - \frac{\langle f^+ + \partial_x(\rho g^+), Q\rangle}{\langle Q', Q'\rangle}\Big) Q'' - \frac{\langle v, Q''\rangle}{\langle Q', Q'\rangle}\frac{\dot c}{c}\widetilde{Q}',\end{aligned}$$

where α and β ensure orthogonality. Later we will need the obvious identity (integrate by parts in the second term)

$$\Big\langle f_- + \partial_x(\rho g_-) - \Big(\frac{\langle f^+ + \partial_x(\rho g^+), Q\rangle}{\langle Q', Q'\rangle}\Big) Q'', Q\Big\rangle = \langle F, Q\rangle.$$

Then, $u = \partial_x P^{\perp}_{Q'} v + u_-$ and hence with $F^+ = f^+ + \partial_x \rho g^+$ we have

$$\|u\|_{L^2 L^2_\rho} \lesssim \|F\|_{Y^s} + \|\langle F, Q\rangle\|_{L^2+L^1} + \|\langle F^+, Q\rangle\|_{L^2+L^1}. \tag{6-25}$$

By (6-21) we see

$$\|\alpha\|_{L^1} \lesssim \|\dot c\|_{L^2\cap L^\infty}(\|F\|_{Y^s} + \|\langle F, Q\rangle\|_{L^2+L^1}) + \|\langle F, Q\rangle\|_{L^1} \tag{6-26}$$

and, using (5-14)

$$\|\alpha\|_{L^2} \lesssim \|F\|_{Y^s} + \|\langle F, Q\rangle\|_{L^2} \tag{6-27}$$

and by (6-22)

$$\|\beta\|_{L^2} \lesssim \|F\|_{Y^s} + \|\langle F, Q'\rangle\|_{L^2}. \tag{6-28}$$

We turn to the frequency localized equation

$$\begin{cases}(u_\lambda)_t + (u_\lambda)_{xxx} = -P_\lambda \partial_x(4Q^3 u) + \alpha P_\lambda \widetilde{Q} + \beta P_\lambda Q_x + P_\lambda f + P_\lambda \partial_x(\rho g),\\ u^\lambda(x, 0) = 0.\end{cases}$$

Observe that by using first the boundedness of Fourier multipliers on U^2, DV^2 and the dual of the embedding $U^2 \subset L^2 H^1_\rho$, we have

$$\|P_\lambda \partial_x(4Q^3 u)\|_{DV^2} \lesssim \lambda \|Q^3 u\|_{DV^2} \lesssim \lambda \|Q^2 u\|_{L^2 L^2} \lesssim \lambda\big(\|F\|_{\dot Y^s_\infty} + \|\langle F, Q\rangle\|_{L^2+L^1}\big).$$

If $\lambda > 1$, then by Lemma 3.1

$$\|[P_\lambda \partial_x, Q^3]u\|_{L^2 L^2_\rho} \lesssim \|u\|_{L^2 L^2_\rho}.$$

Repeating these estimates for the term containing g and using the estimates of the previous section we obtain for $\lambda \le 1$,

$$\|u_\lambda\|_{V^2\cap L^2 H^1_\rho} \lesssim \|P_\lambda f\|_{DU^2} + \lambda\big(\|F\|_{Y^s} + \|\langle F, Q\rangle\|_{L^2+L^1} + \|\langle F^+, Q\rangle\|_{L^2+L^1}\big) + \lambda^{1/2}\|\alpha\|_{L^1},$$

since

$$\|\alpha \widetilde{Q}\|_{L^1 \dot B^{-1/2,2}_\infty} \lesssim \|\alpha\|_{L^1}$$

and, for $\lambda > 1$,

$$\|u_\lambda\|_{V^2\cap L^2 H^1_\rho} \lesssim \|f_\lambda\|_{DU^2} + \|g_\lambda\|_{L^2} + \|F\|_{Y^s} + \|\langle F, Q\rangle\|_{L^2+L^1} + \|\langle F^+, Q\rangle\|_{L^2+L^1} + \|\langle F, Q'\rangle\|_{L^2+L^1}.$$

As a result, we arrive at the following key fact.

Proposition 6.7. *Suppose* (5-14) *holds for some small* δ, *that* $-1/2 < s < 0$, $F \in Y^s$ *and*

$$\begin{cases} u_t + u_{xxx} + 4\partial_x(Q^3 u) = \alpha \widetilde{Q} + \beta Q_x + F, \\ u(x,0) = 0, \end{cases}$$

where α *and* β *are defined in* (6-21) *and* (6-22). *Then*,

$$\|u\|_{\dot{X}^s_\infty} \lesssim \|F\|_{\dot{Y}^s_\infty} + \|\langle F, Q\rangle\|_{L^1} + \|\langle F, Q'\rangle\|_{L^2+L^1} + \|\langle F, Q^+\rangle\|_{L^2+L^1}.$$

This result will play a large role in the nonlinear analysis required to prove asymptotic stability.

For future use, we denote by L^p_I, $\dot{X}^{-1/6}_{\infty,I}$, etc. the function spaces on the space time set $I \times \mathbb{R}$, and specifically we set L^p_T, $\dot{X}^{-1/6}_{\infty,T}$ for $I = (0,T)$. All previous constructions carry over to finite time intervals.

7. Local wellposedness for the quartic KdV equation

In this section we study local wellposedness for the quartic generalized KdV equation

$$\begin{cases} \partial_t \psi - \partial_{xxx}\psi - (\psi^4)_x = 0, \\ \psi(0,x) = \psi_0(x). \end{cases} \tag{7-1}$$

Let v be the solution to the Airy equation with the same initial data, that is,

$$\begin{cases} v_t + v_{xxx} = 0, \\ v(0,x) = \psi_0(x). \end{cases} \tag{7-2}$$

The main local wellposedness is the next result.

Theorem 4. *Let* $r_0 > 0$. *There exist* $\epsilon_0, \delta_0 > 0$ *such that*, *if* $0 < T \leq \infty$,

$$\|\psi_0\|_{\dot{B}^{-1/6,2}_\infty} \leq r_0 \tag{7-3}$$

and

$$\sup_\lambda \|v_\lambda\|_{L^6([0,T],\mathbb{R})} \leq \delta_0, \tag{7-4}$$

then there is a unique solution $\psi = v + w$ *with* $\|w\|_{\dot{X}^{-1/6}_{\infty,T}} \leq \epsilon_0$. *Moreover*, *the function* w (*and hence* ψ) *depends analytically on the initial data.*

By the Strichartz estimates for linear KdV (see also (7-5) and (7-6) below), given v as in (7-2) we have

$$\sup_\lambda \|v_\lambda\|_{L^6} \leq \kappa_0 \|v\|_{\dot{X}^{-1/6}_{\infty,T}},$$

and by the definition of the spaces

$$\|v\|_{\dot{X}^{-1/6}_{\infty,T}} \leq \kappa_1 \big(\|\psi(0)\|_{\dot{B}^{-1/6,2}_\infty} + \|\partial_t v + \partial_{xxx} v\|_{\dot{Y}^{-1/6}_{\infty,T}} \big).$$

Hence, we obtain global existence from Theorem 4 for (7-1) if

$$\|\psi\|_{\dot{B}^{-1/6,2}_\infty} \leq \min\Big\{1, \frac{\delta(1)}{(\kappa_0\kappa_1)}\Big\},$$

where $\delta(1)$ is the δ (which depends on r_0) evaluated at $r_0 = 1$.

In any case, if condition (7-4) is satisfied for $T = \infty$, then, since $\psi_\lambda \in V^2$, the function $e^{t\partial_{xxx}}\psi_\lambda$ is of bounded 2-variation with values in L^2 (see the appendix), and hence it has a limit in L^2 as $t \to \infty$. This implies that $\sum_\lambda \lim_{t\to\infty} e^{t\partial_{xxx}}\psi_\lambda =: S(\psi_0)$ exists and is the scattering state. If in addition ψ_0 is in the closure of C_0^∞ in $\dot{B}_\infty^{-1/6,2}$, then we may exchange the summation and limit.

Under the same assumptions we can solve the initial value problem with initial data $\psi_0(T) = e^{-T\partial_{xxx}}\psi_0$, which, by an easy limit as $T \to \infty$, gives the inverse of the map S. We will later see similar constructions for perturbation of the soliton.

It is not hard to see that if ψ_0 is in the closure of C_0^∞ in $\dot{B}_\infty^{-1/6,2}$, then we can achieve condition (7-4) by choosing T small. This implies local existence with smooth dependence on initial data. Moreover, since we obtain smooth dependence on the initial data, if we have any global solution $\psi(t)$ in the closure of C_0^∞ and perturb the initial data by an amount ε, we obtain a solution at least with a life span $T = -c\ln\varepsilon$ by easy perturbation arguments. In particular, if the initial datum lies in an ε neighborhood of a soliton, then the solution exists at least until time $\sim |\ln\varepsilon|$ and remains in a small neighborhood until that time.

Before turning to the proof we remark that in this section we work with the weaker norms

$$\|u\|_{\dot{X}^{-1/6}} = \sup_\lambda \lambda^{-1/6}\|u_\lambda\|_{V^2} \quad \text{and} \quad \|f\|_{\dot{Y}^{-1/6}} = \sup_\lambda \lambda^{-1/6}\|f_\lambda\|_{DU^2}.$$

On the other hand, since the results remain trivially true for the original definition of the spaces we keep the notation.

Proof. First, we recall some estimates for $u \in U^2_{KdV}$. Let $m(\xi,\xi_1) = m(\xi,\xi-\xi_1)$. Then

$$\|u\|_{L^6_t L^6_x} \lesssim \||D|^{-1/6}u(0)\|_{L^2} \qquad (L^6 \text{ Strichartz estimate}), \tag{7-5}$$

$$\begin{aligned}\Big\|\int_{\mathbb{R}} m(\xi,\xi_1)|\xi_1^2-(\xi-\xi_1)^2|^{1/2}\hat{u}_1(\xi_1)\hat{u}_2(\xi-\xi_1)\,d\xi_1\Big\|_{L^2(\mathbb{R}^2)}& \\ \lesssim \sup\frac{|m(\xi,\xi_1)|^2}{|\xi_1^2-(\xi-\xi_1)^2|^{\frac12}}\|u_1(0)\|_{L^2}\|u_2(0)\|_{L^2}& \qquad (\text{bilinear estimate}).\end{aligned} \tag{7-6}$$

The bilinear estimate is a variant of standard estimates as in [Grünrock 2005]. The most important choice is $m = |\xi_1^2 - (\xi-\xi_1)^2|^{1/2}$.

Let $m(\xi,\xi_1)$ be a function that satisfies $m(\xi,\xi-\xi_1) = m(\xi,\xi_1)$. Then,

$$\begin{aligned}\Big\|\int m(\xi,\xi_1)e^{it(\xi_1^3+(\xi-\xi_1)^3)}\hat{u}_1(0,\xi_1)\hat{u}_2(0,\xi-\xi_1)\,d\xi_1\Big\|^2_{L^2}&\\ =\int m(\xi,\xi_1)m(\xi,\eta_1)e^{3it(\xi_1^2-\xi\xi_1-\eta_1^2+\xi\eta_1)}\hat{u}_1(\xi_1)\hat{u}_2(\xi-\xi_1)\overline{\hat{u}_2(\xi-\eta_1)\hat{u}_1(\eta_1)}\,dt\,d\xi_1\,d\eta_1\,d\xi&\\ =\int\frac{|m(\xi,\eta_1)|^2}{|\eta_1^2-(\xi-\eta_1)^2|}|u_1(\eta_1)|^2|u_2(\xi-\eta_1)|^2\,d\xi\,d\eta_1&\\ \le\sup\frac{|m(\xi,\xi_1)|^2}{|\xi_1^2-(\xi-\xi_1)^2|}\|u_1\|L^{2^2}\|u_2\|L^{2^2}&\end{aligned}$$

since

$$\phi(\xi_1) = \xi\xi_1^2 - \xi^2\xi_1 - \xi\eta_1^2 + \xi^2\eta_1 = \xi(\xi_1 - \eta_1)(\xi_1 + \eta_1 - \xi)$$

vanishes if $\xi_1 = \eta_1$ or $\xi_1 = \xi - \eta_1$ and

$$\phi'(\eta_1) = \xi(2\eta_1 - \xi) = \xi(\eta_1 - (\xi - \eta_1)) \quad \text{and} \quad \phi'(\xi - \eta_1) = \xi((\xi - \eta_1) - \xi_1).$$

These results immediately imply (see the appendix for more information) for $\lambda \geq 1.1\mu$ the estimates

$$\|u_\lambda\|_{L_T^6} \lesssim \lambda^{-1/6}\|u_\lambda\|_{U_T^2}, \tag{7-7}$$

$$\|u_\lambda u_\mu\|_{L_T^2} \lesssim \lambda^{-1}\|u_\lambda\|_{U_T^2}\|u_\mu\|_{U_T^2}. \tag{7-8}$$

By interpolating the bilinear estimate and the Strichartz estimate, if $2 < p \leq 3$,

$$\|u_\lambda u_\mu\|_{L_T^p} \lesssim \lambda^{-1}(\mu^{-1/6}\lambda^{5/6})^{(3p-6)/p}\|u_\lambda\|_{U_T^2}\|u_\mu\|_{U_T^2} \tag{7-9}$$

and, if $\rho \ll \mu \sim \lambda$,

$$\|(u_\lambda u_\mu)_\rho\|_{L_T^2} \lesssim \lambda^{-1/2}\rho^{-1/2}\|u_\lambda\|_{U_T^2}\|u_\mu\|_{U_T^2}. \tag{7-10}$$

Interpolating once again, we have

$$\|(u_\lambda u_\mu)_\rho\|_{L_T^p} \lesssim \lambda^{-1/2}\rho^{-1/2}(\lambda^{1/6}\rho^{1/2})^{(3p-6)/p}\|u_\lambda\|_{U_T^p}\|u_\mu\|_{U_T^p}. \tag{7-11}$$

We proceed with a standard fixed point argument, which requires bounds on the nonlinearity. The solution $\psi = v + w$ is constructed by studying

$$\begin{cases} w_t + w_{xxx} + (v+w)_x^4 = 0, \\ w(0) = 0, \end{cases} \tag{7-12}$$

where again

$$\begin{cases} v_t + v_{xxx} = 0, \\ v(0) = \psi_0. \end{cases}$$

Then, the key estimate is contained in the following.

Lemma 7.1. *There exists $r > 0$ independent of T such that given $v_k \in \dot{X}_{\infty,T}^{-1/6}$ for $k = 1, 2, 3, 4$ we have*

$$\|\partial_x(v_1v_2v_3v_4)\|_{\dot{Y}_{\infty,T}^{-1/6}} \leq r\prod_{k=1}^{4}\|v_k\|_{\dot{X}_{\infty,T}^{-1/6}}, \tag{7-13}$$

and, with v and w defined by (7-2) *and* (7-12), *respectively*,

$$\|\partial_x(v^3w)\|_{\dot{Y}_{\infty,T}^{-1/6}} \leq r\sup_\lambda\|v_\lambda\|_{L^6}\|\psi_0\|_{B_\infty^{-1/6,2}}^2\|w\|_{X_{\infty,T}^{-1/6}}. \tag{7-14}$$

We apply these estimates to $v^4 + 4v^3w + 6v^2w^2 + 4vw^3 + w^4$. Either we may choose to estimate one factor v in L^6 or the dependence on w is at least quadratic. Suppose that $\|w\|_{\dot{X}_{\infty,T}^{-1/6}} \leq \mu$. We obtain

$$\|\partial_x(v+w)^4\|_{\dot{Y}_{\infty,T}^{-1/6}} \leq 6r(\kappa_1^3\delta r_0^3 + \kappa_1^2\delta\mu r_0^2 + \kappa_1^2 r_0^2\mu^2 + \kappa_1 r_0\mu^3 + \mu^4).$$

If $\mu \le \kappa_1 r_0$, then the right hand side is bounded by

$$20r(\kappa_1^3 \delta r_0^3 + \kappa_1^2 \mu^2 r_0^2).$$

Suppose that

$$\mu \le \min\Big\{\kappa_1 r_0, \frac{1}{40r\kappa_1^2 r_0^2}\Big\} \quad \text{and} \quad \delta \le \frac{\mu}{40r\kappa_1^4 r_0^3}.$$

If w solves

$$\begin{cases} w_t + w_{xxx} + (v+W)_x^4 = 0, \\ w(0) = 0 \end{cases}$$

and $\|W\|_{\dot{X}^{-1/6}_{\infty,T}} \le \mu$, then w exists and satisfies $\|w\|_{\dot{X}^{-1/6}_{\infty,T}} \le \mu$.

Standard arguments then allow one to construct a unique solution satisfying the contraction assumption, possibly after decreasing μ by an absolute multiplicative factor. □

It remains to prove Lemma 7.1. By duality, it suffices to verify that

$$\lambda \Big| \int v_1 v_2 v_3 v_4 u_\lambda \, dx\, dt \Big| \le C\lambda^{1/6} \|u_\lambda\|_{V^2} \prod_{k=1}^{4} \|v_k\|_{\dot{X}^{-1/6}_{\infty,T}}$$

and

$$\lambda \Big| \int v^3 w u_\lambda \, dx\, dt \Big| \le C\lambda^{1/6} \|u_\lambda\|_{V^2} \sup_\mu \|v_\mu\|_{L^6} \Big(\sup_\mu \mu^{-1/6} \|v_\mu\|_{U^2} \Big)^2 \|w\|_{\dot{X}^{-1/6}_{\infty,T}},$$

where $u_\lambda \in V^2$ is frequency localized at frequency λ.

By summation, the statement of the lemma holds provided we can prove the following.

Lemma 7.2. *We have for $\lambda_1 \le \lambda_2 \sim \lambda_3 \sim \lambda_4 \sim \lambda_5$ and $\varepsilon > 0$*

$$\lambda_5 \int v_{1,\lambda_1} v_{2,\lambda_2} v_{3,\lambda_3} v_{4,\lambda_4} v_{5,\lambda_5} \, dx\, dt \lesssim \lambda_{\min}^{-1/3} \lambda_5^{-\frac{1}{6}} \Big(\frac{\lambda_{\max}}{\lambda_{\min}}\Big)^{\varepsilon} \prod_{k=1}^{5} \|v_{k,\lambda_k}\|_{V^2} \tag{7-15}$$

and

$$\lambda_{\max} \int v_{\lambda_1} v_{\lambda_2} v_{\lambda_3} u_{\lambda_4} w_{\lambda_5} \, dx\, dt \lesssim \Big(\frac{\lambda_{\min}}{\lambda_{\max}}\Big)^{1/6} \Big(\frac{\lambda_{\max}}{\lambda_{\min}}\Big)^{\varepsilon} \times \sup_\mu \|v_\mu\|_{L^6} \Big(\sup_\mu \mu^{-1/6} \|v_\mu\|_{U^2} \Big)^2 \|u_{\lambda_4}\|_{V^2} \|w_{\lambda_5}\|_{V^2}, \tag{7-16}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ respectively are the maximal and minimal λ_j.

Proof. We claim that

$$\Big| \int v_{1,\lambda_1} v_{2,\lambda_2} v_{3,\lambda_3} v_{4,\lambda_4} v_{5,\lambda_5} \, dx\, dt \Big| \le C\lambda_{\max}^{-1} \|v_{1,\lambda_1}\|_{U^2} \|v_{2,\lambda_2}\|_{U^2} \|v_{3,\lambda_3}\|_{L^6} \|v_{4,\lambda_4}\|_{L^6} \|v_{5,\lambda_5}\|_{L^6} \tag{7-17}$$

provided

$$|\lambda_1 - \lambda_2| \ge \tfrac{1}{10} \lambda_{\max}. \tag{7-18}$$

This estimate is a consequence of Hölder's inequality and the bilinear estimate (7-6). We recall that

$$\|v_\lambda\|_{L^6} \lesssim \lambda^{-1/6}\|v_\lambda\|_{U^6} \lesssim \lambda^{-1/6}\|v_\lambda\|_{V^2}.$$

To obtain a nontrivial integral there have to be elements in the support of the Fourier transforms that add up to zero. Unless there is at least one pair of (λ_j, λ_k) satisfying (7-18), the integral is zero. Hence, we would obtain (7-15) if we were allowed to replace the V^2 norms there by U^2 norms for the first two factors. Observe that we may reorganize the factors as we wish.

Let us assume $\lambda_1 \lesssim \lambda_2 \lesssim \lambda_3 \lesssim \lambda_4 \lesssim \lambda_5$. We consider first the case when $\lambda_4 \le 1.05\lambda_1$. Then, if there are elements in the support of the truncations on the Fourier side adding up to zero — otherwise the integral vanishes — either

$$0.8\frac{\lambda_5}{4} \le \lambda_1 \le \lambda_4 \le 1.1\lambda_1 \le 1.2\frac{\lambda_5}{4} \quad \text{or} \quad 0.6\frac{\lambda_5}{2} \le \lambda_1 \le \lambda_4 \le 1.1\lambda_1 \le 1.4\frac{\lambda_5}{2}.$$

In this case we can replace the U^2 norms by V^2 norms as follows. We decompose into low and high modulation as

$$v_{j,\lambda_j} = v^l_{j,\lambda_j} + v^h_{j,\lambda_j},$$

where v^l_j is defined by the Fourier multiplier projecting to $|\tau - \xi^3| \le \lambda_5^3/1000$. Then, we have

$$\|v^l_{j,\lambda_j}\|_{V^2} + \|v^h_{j,\lambda_j}\|_{V^2} \le \|v_{j,\lambda_j}\|_{V^2} \quad \text{and} \quad \|v^h_{j,\lambda_j}\|_{L^2} \lesssim \lambda_5^{-\frac{3}{2}}\|v_{j,\lambda_j}\|_{V^2}.$$

We refer to the appendix and [Hadac et al. 2009] for more information.

We expand the product. The integral over the product of the five $v^l_{j,\lambda}$ vanishes because of the support of the Fourier transforms. Hence at least one term has high modulation. We estimate it in L^2, put another term into L^∞ and the others into L^6 using Hölder's inequality. We estimate the L^∞ norm through energy and Bernstein's inequality.

Hence

$$\left|\int v_{1,\lambda_1} v_{2,\lambda_2} v_{3,\lambda_3} v_{4,\lambda_4} v_{5,\lambda_5}\, dx\, dt\right| \lesssim \lambda_5^{-3/2} \prod_{j=1}^{5} \|v_{j,\lambda_j}\|_{V^2}, \tag{7-19}$$

which implies the desired estimate.

It remains to study $\lambda_1 \lesssim \lambda_2 \lesssim \lambda_3 \lesssim \lambda_4 \lesssim \lambda_5$, $\lambda_4 \ge 1.05\lambda_1$. The most difficult case is $\lambda_5 \le 1.02\lambda_2$ since otherwise we apply two stronger bilinear estimates. For simplicity we consider $\lambda_1 \ll \lambda$ where $\lambda_2 = \lambda_5 = \lambda$. We have to bound

$$\int_{\sum \xi_j = 0} \prod \hat{u}_{j,\lambda_j}(\xi_j)\, d\xi_2\, d\xi_5\, dt$$

with $\xi_1 = -\sum_{j=2}^5 \xi_j$. We may restrict the integration to $\sum_{j=2}^5 \xi_j \sim \lambda_1$ and $\xi_j \sim \lambda$. By symmetry it suffices to consider

$$\int_{\sum \xi_j = 0} \chi_{||\xi_3| - |\xi_2|| \sim \lambda_1} \prod \hat{v}_{j,\lambda_j}(\xi_j)\, d\xi_2\, d\xi_5\, dt.$$

We choose $\varepsilon > 0$ small, p, q so that $1/p = (1-\varepsilon)/2 + \varepsilon/3$, $1/q = \varepsilon/2 + (1-\varepsilon)/3$. By Hölder's inequality

$$\left|\int v_{1,\lambda_1}(v_{2,\lambda}v_{3,\lambda})_{\lambda_1} v_{4,\lambda}v_{5,\lambda}\,dx\,dt\right| \lesssim \lambda^{-1}(\lambda^{5/6}\lambda_1^{-1/6})^{\varepsilon}\lambda^{-1/2}\lambda_1^{-1/2}(\lambda^{1/6}\lambda_1^{1/2})^{1-\varepsilon}\lambda_3^{-1/6}\prod_j \|v_{j,\lambda_j}\|_{U^{12/5}}.$$
$$\lesssim \lambda^{-3/2}(\lambda^{2/3}\lambda_1^{-2/3})^{\varepsilon}\prod_j \|v_{j,\lambda_j}\|_{U^{12/5}}.$$

For the second part we would like to put one v term into L^6, and up to two into U^2. This can be easily be done if there are two frequencies of v that differ by a small constant times $\lambda_{\max}$. If not it is not hard to see that in the argument above we can put one term into L^6. □

7.1. *Variants and extensions of wellposedness for the quartic KdV equation.* The arguments of the last sections have implications for wellposedness questions in other function spaces. Given $1 \le p \le \infty$, $\omega \in C^1((0,\infty),(0,\infty))$ and $T \in (0,\infty]$, we define the function space $X^{\omega}_{p,T}$ as the set of all distributions for which the norm

$$\|u\|^p_{X^{\omega}_{p,T}} = \sum_\lambda (\omega(\lambda)\|u_\lambda\|_{V^2})^p, \tag{7-20}$$

with obvious modifications if $p = \infty$ is finite. We will always assume that

$$\sup|\omega'|/\omega < \infty, \tag{7-21}$$
$$\inf \omega'/\omega > -1. \tag{7-22}$$

This is a Banach space provided for some $C > 0$ we have

$$\liminf_{\lambda\to 0}\omega(\lambda)\lambda^{1/2} > C; \tag{7-23}$$

otherwise we obtain a Banach space of equivalence classes of functions. Similarly, we define the Banach space

$$\|f\|^p_{Y^{\omega}_{p,T}} = \sum_\lambda (\omega(\lambda)\|f_\lambda\|_{DU^2})^p. \tag{7-24}$$

The definition of $B^{\omega,p}_q$ follows the same pattern. It is not hard to see that

$$\int uf\,dx\,dt \lesssim \|u\|_{X^{\omega}_{p,T}}\|f\|_{Y^{\omega^{-1}}_{p',T}} \quad\text{and}\quad \|f\|_{Y^{\omega^{-1}}_{p',T}} \lesssim \sup_{\|u\|_{X^{\omega}_{p,T}}\le 1}\int uf\,dx\,dt.$$

Moreover, we may expand the inner product into dyadic pieces and apply uniformly elliptic pseudo-differential operators to the pieces. In particular, we may replace differentiation by multiplication on the dyadic pieces and vice versa.

Proposition 7.3. *The following estimate holds*:

$$\|\partial_x(u^4)\|_{Y^{\omega}_{p,T}} \le C\sup_\lambda\|u_\lambda\|_{L^6}\|u\|^2_{\dot X^{-1/6}_{\infty,T}}\|u\|_{X^{\omega}_{p,T}}.$$

Proof. Given $v \in X^{\omega^{-1}}_{p',T}$, we expand $\int \partial_x(u^4)v\,dx\,dt$ into dyadic pieces, to which we apply the arguments and estimate (7-16) from the previous section. By symmetry

$$\sum_{\lambda_j} \lambda_5 \Big| \int u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} u_{\lambda_4} v_{\lambda_5}\,dx\,dt \Big| \lesssim \sum_{\lambda_1 \le \lambda_2 \le \lambda_3 \le \lambda_4, \lambda_5} \lambda_5 \Big| \int u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} u_{\lambda_4} v_{\lambda_5}\,dx\,dt \Big|.$$

If $\lambda_5 \sim \lambda_4$ we obtain

$$\sum_{\lambda_1 \le \lambda_2 \le \lambda_3 \le \lambda_4 \sim \lambda_5} \lambda_5 \Big| \int u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} u_{\lambda_4} v_{\lambda_5}\,dx\,dt \Big|$$
$$\lesssim \sup_\mu \|u_\mu\|_{L^6} \Big(\sup_\mu \mu^{-1/6} \|u_\mu\|_{U^2} \Big)^2 \times \sum_{\lambda_1 \le \lambda_2 \le \lambda_3 \le \lambda_4 \sim \lambda_5} \lambda_1^{1/6} \lambda_5^{-1/6} \Big(\frac{\lambda_5}{\lambda_1} \Big)^{\varepsilon} \|u_{\lambda_4}\|_{V^2} \|v_{\lambda_5}\|_{V^2},$$

which is bounded by

$$\sup_\mu \|u_\mu\|_{L^6} \Big(\sup_\mu \mu^{-1/6} \|u_\mu\|_{U^2} \Big)^2 \|u\|_{X^{\omega}_{p,T}} \|v\|_{X^{\omega^{-1}}_{p',T}}.$$

The other extreme is

$$\sum_{\lambda_5 \le \lambda_1 \le \lambda_2 \le \lambda_3 \sim \lambda_4} \lambda_5 \Big| \int u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} u_{\lambda_4} v_{\lambda_5}\,dx\,dt \Big|$$
$$\le \sup_\mu \|u_\mu\|_{L^6} \Big(\sup_\mu \mu^{-1/6} \|u_\mu\|_{U^2} \Big)^2 \times \sum_{\lambda_5 \le \lambda_1 \le \lambda_2 \le \lambda_3 \sim \lambda_4} \lambda_5 \lambda_4^{-1} \Big(\frac{\lambda_4}{\lambda_5} \Big)^{\varepsilon} \|u_{\lambda_4}\|_{V^2} \|v_{\lambda_5}\|_{V^2},$$

which satisfies the same estimate provided $\sum_{\lambda \le \mu} \lambda \omega(\lambda) \lesssim \mu \omega_\mu$. However, this is ensured by (7-22). The remaining cases are similar and the result follows. □

From Proposition 7.3, we can prove the following corollary to Theorem 4.

Corollary 7.4. *Suppose that ω satisfies* (7-21), (7-22) *and* (7-23). *If $\psi_0 \in \dot{B}^{-1/6,2}_{\infty} \cap B^{\omega,2}_p$ is the initial data for a solution of* (7-1) *and v satisfies* (7-4), *then the solution ψ of Theorem 4 is in $X^{\omega}_{p,T}$ and satisfies*

$$\|\psi\|_{X^{\omega}_{p,T}} \le C \|\psi_0\|_{B^{\omega,2}_p}.$$

In addition:

Corollary 7.5. *Suppose that ψ_0 lies in the closure of C_0^∞ in $\dot{B}^{-1/6}_{\infty,T}$. Then, it follows that*

$$(t \to \psi(t)) \in C([0,T], \dot{B}^{-1/6,2}_{\infty}).$$

If $T = \infty$, then $e^{t\partial^3_{xxx}}\psi$ converges to the scattering data as $t \to \infty$ in $\dot{B}^{-1/6}_{\infty,T}$. If in addition $\psi_0 \in L^2$, then

$$(t \to \psi(t)) \in C([0,T], L^2)$$

and $e^{t\partial^3_{xxx}}\psi$ converges also in L^2.

There exists ω satisfying the assumptions above, with $\omega(\lambda)\lambda^{-1/6} \to \infty$ as $\lambda \to \infty$ and $\lambda \to 0$ and $\|\psi_0\|_{B^{\omega,2}_{\infty}} < \infty$. By Corollary 7.4 the $X^{\omega}_{\infty,T}$ is controlled by the initial data. Hence

$$\lambda^{-1/6} \|v_\lambda\|_{DV^2 \cap X^0} \to 0 \text{as } \lambda \to \infty \text{ or } \lambda \to 0.$$

By the previous argument the deviation of the solution to the linear solution tends to zero as the considered interval shrinks to zero. This implies continuity. Continuity at infinity always holds in V^p.

The second part requires an obvious specialization of Corollary 7.4 to the case $\omega = 1$, plus a repetition of the argument for scattering.

Particular examples for ω are $\langle\lambda\rangle^s$ for $s \ge -1/6$ and $\lambda^s + \lambda^\sigma$ for $-1/2 \le s \le -1/6 \le \sigma$. It is not hard to see that we can replace the homogeneous spaces by inhomogeneous ones if we restrict to finite T and allow the constants to depend on T.

8. Stability and scattering for perturbations of the soliton

8.1. ***Setup and main result.*** We return now to the full nonlinear problem (7-1). Let us take

$$\psi(x,t) = Q_{c(t)}(x - y(t)) + w(x,t).$$

Then, we have

$$\begin{aligned}\partial_t w + \partial_x(\partial_x^2 w + 4Q_c^3 w) = -\dot c(\partial_c Q_c)(x-y) + \dot y(Q_c')(x-y) - \partial_x(\partial_x^2 Q_c - c^2 Q_c + Q_c^4) - c^2(Q_c'(x-y)) \\ - \partial_x(6Q_c^2(x-y)w^2 + 4Q_c(x-y)w^3 + w^4).\end{aligned}$$

Hence,

$$\begin{aligned}\partial_t w + \partial_x(\partial_{xx}^2 w + 4Q_c^3 w) = -(\dot c/c)\widetilde{Q}_c(x-y) + (\dot y - c^2)Q_c'(x-y) \\ - \partial_x(6Q_c^2(x-y)w^2 + 4Q_c(x-y)w^3 + w^4). \quad (8\text{-}1)\end{aligned}$$

To use the dispersive estimates proved in Section 6, we wish to have

$$w \perp Q_c(x-y) \quad \text{and} \quad w \perp Q_c'(x-y). \quad (8\text{-}2)$$

To get more regularity for y and c, we ask for (8-2) only asymptotically and hence take as in (1-15) and (1-16) the modal equations

$$(\dot c/c)\langle Q_c, \widetilde{Q}_c\rangle = \langle w, Q_c\rangle, \quad (8\text{-}3)$$

$$(\dot y - c^2)\langle Q_c', Q_c'\rangle = -\kappa\langle w, Q_c'\rangle, \quad (8\text{-}4)$$

where $\kappa > 0$ is taken to be large.

We calculate

$$\begin{aligned}\frac{d}{dt}\langle w, Q\rangle &= \langle w_t, Q\rangle + \dot y\langle w, Q'\rangle + (\dot c/c)\langle w, \widetilde{Q}\rangle \\ &= \langle w, \mathcal{L}Q'\rangle - (\dot c/c)\langle Q, \widetilde{Q}\rangle + \langle 6Q^2w^2 + 4Qw^3 + w^4, Q'\rangle + (\dot y - c^2)\langle w, Q'\rangle + (\dot c/c)\langle w, \widetilde{Q}\rangle\end{aligned}$$

and

$$\begin{aligned}\frac{d}{dt}\langle w, Q'\rangle &= \langle w_t, Q'\rangle + \dot y\langle w, Q''\rangle + (\dot c/c)\langle w, \widetilde{Q}'\rangle \\ &= \langle w, \mathcal{L}Q''\rangle + (\dot y - c^2)\langle Q', Q'\rangle + \langle 6Q^2w^2 + 4Qw^3 + w^4, Q''\rangle \\ &\qquad + (\dot y - c^2)\langle w, Q''\rangle + (\dot c/c)\langle w, \widetilde{Q}'\rangle.\end{aligned}$$

Hence,

$$\frac{d}{dt}\langle w, Q\rangle + \langle w, Q\rangle = -\kappa \frac{\langle w, Q'\rangle^2}{\langle Q', Q'\rangle} - \frac{\langle w, Q\rangle\langle w, \tilde{Q}\rangle}{\langle Q, \tilde{Q}\rangle} + \langle 6Q^2w^2 + 4Qw^3 + w^4, Q'\rangle \tag{8-5}$$

and

$$\frac{d}{dt}\langle w, Q'\rangle + \kappa\langle w, Q'\rangle - \langle w, \mathscr{L}Q''\rangle = -\kappa \frac{\langle w, Q'\rangle\langle w, Q''\rangle}{\langle Q', Q'\rangle} + \frac{\langle w, Q\rangle\langle w, \tilde{Q}'\rangle}{\langle Q, \tilde{Q}\rangle} + \langle 6Q^2w^2 + 4Qw^3 + w^4, Q''\rangle. \tag{8-6}$$

The right hand sides are at least quadratic in w, and, as we shall see, small compared to $\|w_0\|$ in a suitable sense. As a consequence the orthogonality conditions are approximately satisfied for large t. In addition, $\dot{c}$ and $\dot{y} - c^2$ are small and continuous.

We study the initial value problem $w(0) = w_0$. Let again v be the solution to the linear problem. We will prove scattering for small perturbations of the soliton in $\dot{B}^{-1/6,2}_\infty$. It will be important for the reverse problem that we will achieve something slightly stronger.

Using the notation

$$\Gamma = \left\{ y \in C([0,\infty)) : y(0) = 0, |\dot{y} - 1| \le \tfrac{1}{10} \right\}, \tag{8-7}$$

we define for any interval I the quantity

$$J_I(v) = \sup_\lambda \Big(\|v_\lambda\|_{L^6_I} + \lambda^{1/4-1/6}\|v_\lambda\|_{L^4_I L^\infty} + \sup_{y\in\Gamma} \lambda^{-1/6} \int_{\mathbb{R}\times I} \gamma_0'(x - y(t))(v_\lambda^2 + (\partial_x v_\lambda)^2)\, dx\, dt \Big). \tag{8-8}$$

Proposition 8.1. *Let v be a solution of* (3-1) *with initial data $v_0 \in \dot{B}^{-1/6,2}_\infty$. Then,*

$$J_{[0,\infty)}(v) \lesssim \|v_0\|_{\dot{B}^{-1/6,2}_\infty}.$$

Moreover, if v_0 is in the closure of C_0^∞ in $\dot{B}^{-1/6,2}_\infty$, then

$$\lim_{t\to\infty} J_{[t,\infty)}(v) = 0.$$

Proof. The first statement is an immediate consequence of the Strichartz estimate and local smoothing. For the second statement we fix $\varepsilon > 0$. There are at most finitely many $v_{0,\lambda}$ of norm larger than ε/c. Hence it suffices to verify the statement for a single λ. Since $v_\lambda \in L^6L^6$ and L^4L^∞, we have

$$\lim_{t\to\infty} \|v_\lambda\|_{L^6(\mathbb{R}\times(t,\infty))} = \lim_{t\to\infty} \|v_\lambda\|_{L^4_{[t,\infty)}L^\infty} = 0.$$

Let I be a bounded interval. Then the map

$$\Gamma \to \lambda^{-1/6} \int_{\mathbb{R}\times I} \gamma_0'(x - y(t))(v_\lambda^2 + (\partial_x v_\lambda)^2)\, dx\, dt$$

is continuous with respect to uniform convergence, hence it assumes its maximum. Given $j \ge 1$, let $y_j : [2^j, 2^{j+1}] \to \mathbb{R}$ be the path for which this quantity is maximal. We choose two paths y_o and y_e with $\gamma(0) = 0$ and the difference between 1 and the derivative at most 0.2, one which coincides with y_j for j

even on the corresponding intervals, and one which does so for j odd. For both paths we have the local smoothing estimate. But this implies the claim. □

Let $y \in \Gamma$. The function spaces $\dot X^s_{\infty;T}$ and $\dot Y^s_{\infty;T}$ depend on y but not on c. This dependence is not reflected in the notation. In addition, let $c \in C^1([0,\infty))$. We assume (5-14), $\dot c \in L^2$ and $\dot y - c^2 \in L^2$ in this section, which we have to verify for the solutions we study, and turn to a study of a priori estimates for solutions to (8-1), (8-3) and (8-4), and we recall (8-5) and (8-6). Because of translation and scaling invariance we may restrict ourselves to a study for $y(0) = 0$ and $c(0) = 1$. Moreover, we may and do assume that the orthogonality conditions hold at time 0, that is,

$$\langle w_0, Q\rangle = \langle w_0, Q'\rangle = 0.$$

The main result is the following sharpened version of Theorem 2.

Proposition 8.2. *Let $C > 0$. There exist $\varepsilon > 0$ and $K > 0$ such that for $\|w_0\|_{\dot B^{-1/6,2}_\infty} < C$ and $J_{[0,T)}(v) \le \varepsilon$ and for v a solution of* (3-1) *with initial data w_0, the solution w in the system of equations* (1-15)–(1-14) *satisfies* (5-14),

$$\|w\|_{\dot X^{-1/6}_{\infty,T}} \le K J^{1/2}_{[0,T)}(v),$$

with K depending on C but not on time. Moreover, if $J_{(0,\infty)}(v) \le \varepsilon$, then there exists a unique $\eta \in \dot B^{-1/6,2}_\infty$ such that

$$\lim_{t\to\infty} e^{t\partial^3_{xxx}} w_\lambda(t) = \eta_\lambda,$$

with convergence in L^2. In addition

$$\lim_{t\to\infty} \|w(t)\|_{\dot B^{-1/6}_{\infty,T}} = \|\Psi\|_{\dot B^{-1/6,2}_\infty}.$$

Remark 8.1. Variants in the spirit of Corollary 7.4 can be easily obtained by including the arguments there, which will establish Theorems 2 and 3 with higher Sobolev regularity as stated in Remark 1.1.

The proof consists of three step,s a preliminary part consisting of an important initialization, multilinear estimates that are less critical variants of those of the last section, and a priori estimates for the nonlinear equation using the multilinear estimates and the linearized equation.

We recall that v satisfies $v_t + v_{xxx} = 0$ with initial data $v(0) = w_0$. We want to control the difference between v and the solution ν to $\nu_t + c^2\partial_x \nu - \partial_x \mathscr{L}\nu = \alpha \widetilde Q + \beta Q'$ with initial data $\nu(0) = w_0$ with α and β ensuring $\langle \nu, Q\rangle = \langle \nu, Q'\rangle = 0$, which we assume to hold initially. We recall that (5-14) is a standing assumption.

For simplicity, let us define $J = J_{[0,\infty)}(v)$. The following result is the first step of the proof.

Lemma 8.3. *Suppose that $w_0 \in B^{-1/6,2}_\infty$ satisfies the orthogonality conditions. Then*

$$\|\nu\|_{\dot X^{-1/6}_\infty} \lesssim \|w_0\|_{\dot B^{-1/6,2}_\infty}$$

and

$$\|\nu - \tilde P P^\perp_{Q'} v\|_{\dot X^{-1/6}_\infty} + \|\alpha\|_{L^1\cap L^2} + \|\beta\|_{L^2} \lesssim J. \tag{8-9}$$

Proof. The first bound on v is an immediate consequence of Proposition 6.7. The second statement is more delicate. As a first step we consider $u = \tilde{P} P^{\perp} v$. It satisfies

$$\sup_{\lambda}\big(\|u_\lambda\|_{L^6} + \lambda^{-1/6+1/4}\|u_\lambda\|_{L^4 L^\infty} + \lambda^{-1/6}\|u_\lambda\|_{L^2 H^1_\rho}\big) \lesssim J, \tag{8-10}$$

since $\|\langle v, Q\rangle\|_{L^2\cap L^6} + \|\langle v, Q'\rangle\|_{L^2\cap L^6} \lesssim J$. We calculate

$$\partial_t u + c^2 u_x - \partial_x \mathscr{L}_{c,y} u = G, \quad \text{with } u(0) = w_0, \tag{8-11}$$

where

$$G = -4\partial_x(Q^3 u) - \Big(\frac{d}{dt}\frac{\langle v, Q\rangle}{\langle Q, \widetilde{Q}\rangle}\Big)\widetilde{Q} - \Big(\frac{d}{dt}\frac{\langle v, Q'\rangle}{\langle Q', Q'\rangle}\Big)Q' \\ - \frac{\langle v, Q\rangle}{\langle Q, \widetilde{Q}\rangle}(\partial_t + c^2\partial_x - \partial_x\mathscr{L})\tilde{Q} - \frac{\langle v, Q'\rangle}{\langle Q', Q'\rangle}(\partial_t + c^2\partial_x - \partial_x\mathscr{L})\tilde{Q}'.$$

We consider the terms separately. Any derivative falling on $\langle Q, \widetilde{Q}\rangle$ or $\langle Q', Q'\rangle$ can be computed using (4-2) and (4-3), yielding a factor $\dot{c}/c$. Next,

$$\frac{d}{dt}\langle v, Q\rangle = \langle \partial_t v + c^2 v', Q\rangle + (c^2 - \dot{y})\langle v, Q'\rangle + (\dot{c}/c)\langle v, \widetilde{Q}\rangle = -\langle Q_x^4, v\rangle + (c^2 - \dot{y})\langle v, Q'\rangle + (\dot{c}/c)\langle v, \widetilde{Q}\rangle$$

and

$$\frac{d}{dt}\langle v, Q'\rangle = \langle \dot{v} + c^2 v', Q'\rangle + (c^2 - \dot{y})\langle v, Q''\rangle + (\dot{c}/c)\langle v, \tilde{Q}'\rangle \\ = -\langle v, \mathscr{L}Q''\rangle - 4\langle Q^3 Q'', v\rangle + (c^2 - \dot{y})\langle v, Q''\rangle + (\dot{c}/c)\langle v, \widetilde{Q}'\rangle.$$

Moreover,

$$(\partial_t + c^2\partial_x - \partial_x\mathscr{L})\widetilde{Q} = (\dot{c}/c)\widetilde{\widetilde{Q}} + (c^2 - \dot{y})\widetilde{Q}' + 2c^2 Q', \\ (\partial_t + c^2\partial_x - \partial_x\mathscr{L})Q' = (\dot{c}/c)\widetilde{Q}' + (c^2 - \dot{y})Q''.$$

We write $G = \alpha\widetilde{Q} + \beta Q' + g$, where, using again (4-2) and (4-3),

$$\langle Q, \widetilde{Q}\rangle\alpha = -(\dot{c}/c)\langle v, \widetilde{Q}\rangle - (c^2 - \dot{y})\langle v, Q'\rangle + \tfrac{1}{3}(\dot{c}/c)\langle v, Q\rangle + \langle v, (Q^4)_x\rangle,$$

$$\langle Q', Q'\rangle\beta = -(\dot{c}/c)\langle v, \widetilde{Q}'\rangle - (c^2 - \dot{y})\langle v, Q''\rangle + \tfrac{10}{3}(\dot{c}/c)\langle v, Q'\rangle + 4\langle v, Q^3 Q''\rangle + \langle v, \mathscr{L}_c Q_c''\rangle - 2c^2\langle v, Q'\rangle,$$

$$g = -4\partial_x(Q^3 v) - \frac{\langle v, Q\rangle}{\langle Q, \widetilde{Q}\rangle}((\dot{c}/c)\widetilde{\widetilde{Q}} + (c^2 - \dot{y})\widetilde{Q}') - (\langle v, Q'\rangle/\langle Q', Q'\rangle)(\frac{\dot{c}}{c}\widetilde{Q}' + (c^2 - \dot{y})Q'').$$

By Lemma 6.3, we have $\|g\|_{Y^0} \lesssim J$. The difference $w = v - u$ satisfies (abusing the notation slightly by denoting by α and β new quantities)

$$w_t + c^2 w_x - \partial_x\mathscr{L}w = \alpha\widetilde{Q} + \beta Q' - g$$

with initial data $w(0) = 0$, and again by Lemma 6.3

$$\|w\|_{X^0\cap V^2} \lesssim \|g\|_{Y^0} \lesssim J.$$

We rewrite the equation for v as

$$v_t + v_{xxx} = -\partial_x(4Qv) + \alpha\widetilde{Q} + \beta Q' =: F.$$

Decompose $v = u + w$. We recall that

$$\langle Q, \widetilde{Q}\rangle\alpha = -(\dot{c}/c)\langle v, \widetilde{Q}\rangle; \tag{8-12}$$

hence

$$\|\alpha\|_{L^1} + \|F\|_{\dot{Y}_\infty^{-1/6}} \lesssim (\|\dot{c}/c\|_{L^2} + \|\dot{y} - c^2\|_{L^2})J.$$

The L^2 bound for β is simpler. The estimates for the linear equation imply now (8-9). □

As it will be used in the sequel, we note the following simple consequence of Lemma 8.3. Namely, we have

$$J_{c,y}(v) \lesssim J(v), \tag{8-13}$$

where we denote by $J_{c,y}$ the quantity analogous to J, but for the given path dictated by the c and y modulation parameters. After this nontrivial preliminary step, we continue with the proof of Proposition 8.2. The strategy is to write the equation in terms of

$$u = \Psi - Q_{c(t),y(t)} - v$$

and expand the nonlinearity. In the next step we study multilinear estimates, which in the last step are combined with Proposition 6.7 to obtain the a priori estimates.

8.2. *Multilinear estimates.* We proceed as for the initial value problem and bound multilinear expressions. In this section we collect nonlinear estimates in terms of the V^2 spaces to prove Proposition 8.2.

Lemma 8.4. *Let u be a tempered distribution and u_λ its frequency localization. Let ϕ be a Schwartz function. Then,*

$$\|\phi u_\lambda\|_{L^2} \lesssim \min\{\lambda^{1/2-\varepsilon}, \lambda^{-1}\}\big(\|u_\lambda\|_{L^2(\gamma')} + \|\partial_x u_\lambda\|_{L^2(\gamma')}\big).$$

Here ε is the constant of (3-7).

Proof. We begin with the case $\lambda \geq 1$, in which case we prove the stronger estimate where we replace ϕ by γ' as defined in Section 3. Let $\chi \in C_0^\infty$ be supported in $\{\xi : \frac{1}{2} \leq |\xi| \leq 2\}$. Then,

$$\begin{aligned}\sqrt{\gamma'}u_\lambda &= \sqrt{\gamma'}\partial_x^{-1}\chi\Big(\frac{\partial_x}{\lambda}\Big)\partial_x u_\lambda = \lambda^{-1}\sqrt{\gamma'}\Big(\frac{\partial_x}{\lambda}\Big)^{-1}\chi\Big(\frac{\partial_x}{\lambda}\Big)\partial_x u_\lambda \\ &= \lambda^{-1}\Big(\frac{\partial_x}{\lambda}\Big)^{-1}\chi\Big(\frac{\partial_x}{\lambda}\Big)\sqrt{\gamma'}\partial_x u_\lambda + \lambda^{-1}\Big(\sqrt{\gamma'}, \Big(\frac{\partial_x}{\lambda}\Big)^{-1}\chi\Big(\frac{\partial_x}{\lambda}\Big)\Big)\partial_x u_\lambda,\end{aligned}$$

where $(\partial_x/\lambda)^{-1}\chi(\partial_x/\lambda)$ is an L^2 bounded Fourier multiplier. As a result,

$$\Big\|\lambda^{-1}\Big(\frac{\partial_x}{\lambda}\Big)^{-1}\chi\Big(\frac{\partial_x}{\lambda}\Big)\sqrt{\gamma'}\partial_x u_\lambda\Big\|_{L^2} \lesssim \lambda^{-1}\|u_\lambda\|_{L^2(\gamma')}.$$

We estimate the second term on the right hand side using the adjoint of Lemma 3.1 with $a=1$ and $s=0$:

$$\left\|\lambda^{-1}\left(\sqrt{\gamma'},\left(\frac{\partial_x}{\lambda}\right)^{-1}\chi\left(\frac{\partial_x}{\lambda}\right)\right)(\gamma')^{-1/2}\sqrt{\gamma'}\partial_x u_\lambda\right\|_{L^2}\lesssim\lambda^{-1}\|u_\lambda\|_{L^2(\gamma')}.$$

We turn to $\lambda<1$. Clearly,

$$\|\phi u_\lambda\|_{L^2}\le\|\phi(\gamma')^{-1/2}\|_{L^2}\|\sqrt{\gamma'}u_\lambda\|_{L^\infty}.$$

Let $\tilde{\chi}=\sin(x)/x$, which is the inverse Fourier transform (up to a constant factor) of the characteristic function of the interval $[-1,1]$. Let $x_0\in\mathbb{R}$. We define

$$g_\lambda(x)=u_\lambda(x)\tilde{\chi}(\lambda(x-x_0)/100).$$

Then, g_λ satisfies roughly the same frequency localization as v_λ, and it coincides with u_λ at x_0. Thus, by Bernstein's inequalities,

$$|\sqrt{\gamma'(x_0)}u_\lambda(x_0)|\le c\lambda^{1/2}\sqrt{\gamma'(x_0)}\|g_\lambda\|_{L^2}\le c\lambda^{1/2}\sup_{x,x_0}\frac{\sqrt{\gamma'(x_0)}\tilde{\chi}(\lambda(x-x_0)/100)}{\sqrt{\gamma'(x)}}\|\sqrt{\gamma'}u_\lambda\|_{L^2}.$$

Now the elementary estimate

$$\sup_{x,x_0}\sqrt{\gamma'(x_0)}\frac{|\tilde{\chi}(\lambda(x-x_0)/100)|}{\sqrt{\gamma'(x)}}\le c\lambda^{-\varepsilon}$$

completes the proof. □

We proceed to prove the necessary multilinear estimates.

Lemma 8.5. *Let* c, y *satisfy* (5-14), $u\in\dot{X}_\infty^{-1/6}$ *and let* v *and* Q *be as in Proposition 8.2. Then, the following estimates hold*:

$$\|\partial_x(u_1u_2u_3Q)\|_{\dot{Y}_\infty^{-1/6}}\lesssim\prod_{j=1}^{3}\|u_j\|_{\dot{X}_\infty^{-1/6}},\tag{8-14}$$

$$\|\partial_x(u_1u_2Q^2)\|_{\dot{Y}_\infty^{-1/6}}\lesssim\prod_{j=1}^{2}\|u_j\|_{\dot{X}_\infty^{-1/6}},\tag{8-15}$$

$$\|\partial_x(v^2uQ)\|_{\dot{Y}_\infty^{-1/6}}\lesssim J^{1/2}\|v\|^{3/2}_{\dot{X}_\infty^{-1/6}}\|u\|_{\dot{X}_\infty^{-1/6}}\tag{8-16}$$

$$\|\partial_x(vuQ^2)\|_{\dot{Y}_\infty^{-1/6}}\lesssim J^{1/2}\|v\|^{1/2}_{\dot{X}_\infty^{-1/6}}\|u\|_{\dot{X}_\infty^{-1/6}}.\tag{8-17}$$

Proof. We begin with the dual Strichartz estimate

$$\|f_\lambda\|_{DV^2}\lesssim\lambda^{-1/4}\|f_\lambda\|_{L^{4/3}L^1}.$$

By construction, spatial Fourier multipliers in V^p, U^p, DU^p and DV^p are bounded by the supremum of the multiplier; hence

$$\|P_\lambda\partial_x(Q^2u_{1,\lambda_1}u_{2,\lambda_2})\|_{DU^2}\lesssim\lambda^{3/4}\|Q^2u_{1,\lambda_1}u_{2,\lambda_2}\|_{L^{\frac{4}{3}}L^1}$$

and

$$\|Q^2 u_{1,\lambda_1} u_{2,\lambda_2}\|_{L^{\frac{4}{3}}L^1} \le \|Q u_{1,\lambda_1}\|_{L^2L^2} \|Q u_{2,\lambda_2}\|^{1}_{L^2L^2}/2 \|u_{2,\lambda_2}\|^{1/2}_{L^\infty L^2}$$
$$\lesssim \min\{1, \lambda_1^{-1}\} \min\{1, \lambda_2^{-1/2}\} \lambda_1^{1/6} \lambda_2^{1/6} \|u_1\|_{\dot X^{-1/6}_\infty} \|u_2\|_{\dot X^{-1/6}_\infty}.$$

This is summable for $\lambda_j \in \Gamma$ and we obtain the desired estimate for $\lambda \le 1$. Assume now that $\lambda \ge 1$. Then, using Hölder and Bernstein and $|Q'| \lesssim Q$

$$\lambda^{-1/6} \|P_\lambda \partial_x Q^2 u_{1,\lambda_1} u_{2,\lambda_2}\|_{L^2\gamma H^{-1}} \lesssim \lambda^{-1/6} \|Q u_{1,\lambda_1}\|^{1/2}_{L^\infty} \|Q u_{2,\lambda_2}\|^{1/2}_{L^\infty} \|Q u_{1,\lambda_1}\|^{1/2}_{L^2} \|Q u_{2,\lambda_2}\|^{1/2}_{L^2}$$
$$\lesssim \lambda^{-1/6} \lambda_1^{5/12} \min\{1, \lambda_1^{-1/2}\} \lambda_2^{5/12} \min\{1, \lambda_2^{-1/2}\} \|u_1\|_{\dot X^{-1/6}_\infty} \|u_2\|_{\dot X^{-1/6}_\infty},$$

which can easily be summed over λ_1 and λ_2 if $\lambda \ge 1$. This implies (8-15) and also (8-17).

We approach estimate (8-14) similarly: We expand u_j and observe that the expressions are symmetric; hence it suffices to sum over $\lambda_1 \le \lambda_2 \le \lambda_3$. If $\lambda_1 \lesssim 1$ we argue as above and estimate u_{1,λ_1} in L^∞, followed by Bernstein's inequality. So we restrict to the case $\lambda_1 \gg 1$.

Then, using that Q is integrable,

$$\lambda \|P_\lambda Q u_{1,\lambda_1} u_{2,\lambda_2} u_{3,\lambda_3}\|_{DU^2} \lesssim \lambda^{3/4} \prod_{j=1}^{3} \|u_{j,\lambda_j}\|_{L^4L^\infty} \lesssim \lambda^{3/4} (\lambda_1\lambda_2\lambda_3)^{-1/12} \prod_{j=1}^{3} \|u_{j,\lambda_j}\|_{\dot X^{-1/6}_{\infty,T}},$$

which is easily summable if $\lambda \lesssim 1 \lesssim \lambda_1, \lambda_2, \lambda_3$. If $\lambda > 1$, we argue differently. To simplify the argument we assume that the Fourier transform of Q is supported in $[-1, 1]$ — handling the tail is straightforward but technical. Instead of bounding $\lambda \|P_\lambda Q u_{1,\lambda_1} u_{2,\lambda_2} u_{3,\lambda_3}\|_{DU^2}$, we employ duality and study

$$I = \left| \int Q u_{1,\lambda_1} u_{2,\lambda_2} u_{3,\lambda_3} u_{4,\lambda_4} \, dx \, dt \right|$$

assuming that $1 \ll \lambda_1 \le \lambda_2 \le \lambda_3$. Then, we have

$$I \le \|Q u_{3,\lambda_3}\|_{L^2} \|u_{1,\lambda_1}\|_{L^6} \|u_{2,\lambda_2}\|_{L^6} \|u_{4,\lambda_4}\|_{L^6} \lesssim \lambda_3^{-5/6} \lambda_4^{-1/6} \prod_{j=1}^{3} \|u_{j,\lambda_j}\|_{\dot X^{-1/6}_{\infty,T}} \|u_{4,\lambda_4}\|_{V^2}.$$

The factor $\lambda_3^{-5/6} \lambda_4^{5/6}$ is summable for fixed λ_4 over $1 \le \lambda_1 \le \lambda_2 \le \lambda_3$, $1 \le \lambda_4 \lesssim \lambda_3$ — this suffices since $I = 0$ if λ_4 is much larger than λ_3. As a result, we have proven estimate (8-14) and, after checking the proof, (8-16). □

We turn to bounds for inner products occurring as inner products of the right hand side of (8-1) with Q and Q', and at the right hand side of (8-5) and (8-6).

Lemma 8.6. *Let $u \in \dot X^{-1/6}_\infty$, and let v and Q be as in Proposition 8.2. In addition, let $\psi_0(t)$ be a one parameter family of Schwartz functions parametrized by t with uniformly bounded seminorms and $\psi(x, t) = \psi_0(t, x - y(t))$. Then for all $1 \le p < 3/2$*

$$\|\langle \partial_x (u_1 u_2 u_3 u_4), \psi \rangle\|_{L^p} \lesssim \prod_{j=1}^{4} \|u_j\|_{\dot X^{-1/6}_\infty}, \tag{8-18}$$

where we consider the L^p norm with respect to time and

$$\|\langle\partial_x(v^3u),\psi\rangle\|_{L^p}\lesssim J\|v\|^2_{\dot X^{-1/6}_\infty}\|u\|_{\dot X^{-1/6}_\infty}. \tag{8-19}$$

For all $1\le p<2$, *we have*

$$\|\langle\partial_x(u_1u_2u_3Q),\psi\rangle\|_{L^p}\lesssim\prod_{j=1}^3\|u_j\|_{\dot X^{-1/6}_\infty}, \tag{8-20}$$

$$\|\langle\partial_x(v^2uQ),\psi\rangle\|_{L^p}\lesssim J\|v\|_{\dot X^{-1/6}_\infty}\|u\|_{\dot X^{-1/6}_\infty}. \tag{8-21}$$

For all $1\le p<3$, *we have*

$$\|\langle\partial_x(u_1u_2Q^2),\psi\rangle\|_{L^p}\lesssim\prod_{j=1}^2\|u_j\|_{\dot X^{-1/6}_\infty}, \tag{8-22}$$

$$\|\langle\partial_x(vuQ^2),\psi\rangle\|_{L^p}\lesssim J\|u\|_{\dot X^{-1/6}_\infty}. \tag{8-23}$$

Proof. We expand the terms in (8-18) and we consider

$$I_p:=\|\langle u_{1,\lambda_1}u_{2,\lambda_2}u_{3,\lambda_3}u_{4,\lambda_4},\psi\rangle\|_{L^p}.$$

By symmetry it suffices to look at the case $\lambda_1\le\lambda_2\le\lambda_3\le\lambda_4$. If $p=1$ we bound the terms using Hölder's and Bernstein's inequalities as above:

$$\begin{aligned}I_1&\lesssim\|u_{1,\lambda_1}\|_{L^\infty}\|u_{2,\lambda_2}\|_{L^\infty}\||\psi|^{1/2}u_{3,\lambda_3}\|_{L^2}\||\psi|^{1/2}u_{4,\lambda_4}\|_{L^2}\\&\lesssim\lambda_1^{2/3}\lambda_2^{2/3}\min\{\lambda_3^{1/6},\lambda_3^{-5/6}\}\min\{\lambda_4^{1/6},\lambda_4^{-5/6}\}\prod_{j=1}^4\|u_j\|_{\dot X^{-1/6}_\infty},\end{aligned}$$

which is easily summable. We obtain by Hölder's inequality

$$I_{3/2}\lesssim\prod_{j=1}^4\|u_{j,\lambda_j}\|_{L^6}\lesssim\prod_{j=1}^4\|u_j\|_{\dot X^{-1/6}_\infty},$$

which we use if $1\le\lambda_1\le\lambda_4$. If $\lambda_1\le 1$, we estimate the corresponding term in L^∞, apply Bernstein's inequality, and argue as in the next case. Interpolation with the L^1 estimate yields a summable expression as long as $p<3/2$.

We turn to estimate (8-20), denote again the p-norms by I_p and expand again

$$I_1\lesssim\|u_{1,\lambda_1}\|_{L^\infty}\|Qu_{2,\lambda_2}\|_{L^2}\|(\partial_x\psi)u_{3,\lambda_3}\|_{L^2}\lesssim\lambda_1^{2/3}\min\{\lambda_2^{1/6},\lambda_2^{-5/6}\}\min\{\lambda_3^{1/6},\lambda_3^{-5/6}\}\prod\|u_j\|_{\dot X^{-1/6}_\infty},$$

which again is easily summable over $\lambda_1\le\lambda_2\le\lambda_3$. Also

$$I_2\lesssim\prod_{j=1}^3\|u_{j,\lambda_j}\|_{L^6}\lesssim\prod_{j=1}^3\|u_{j,\lambda_j}\|_{\dot X^{-1/6}_\infty},$$

which is almost summable, and by interpolation we obtain the bounds for any $p<2$.

The estimate (8-22) with $p = 1$ follows by the same arguments as above. It is even simpler. Again we may restrict ourselves to $\lambda_1 \le \lambda_2$. For $p = 3$ we put estimate u_{j,λ_j} into L^6 and again the full statement follows by interpolation. A simple check of the proof reveals that the arguments above imply (8-23), (8-21) and (8-19). □

The right hand sides of (8-5) and (8-6) are functions of t, for which we have bounds in L^p for $1 \le p < 3/2$ in terms of $\|w\|_{\dot X^{-1/6}_{\infty,T}}$. In the second equation, (8-6), the term $\langle w, \mathcal{L} Q_{xx}\rangle$ plays a special role: It is in L^q for $2 \le q \le \infty$, but not in L^p for any $p < 2$ in general. In particular we cannot control the deviation of y from the linear movement.

Equation (8-5) and (8-6) can be considered as scalar linear ordinary differential equations for $\langle w, Q\rangle$ and $\langle w, Q'\rangle$. The kernel for the fundamental solution is uniformly bounded in L^p in the first case for all p, and in the second case it is bounded in L^1 by $1/\kappa$, whereas the L^∞ norm is 1.

We collect the consequences as follows.

Lemma 8.7. *Suppose that w solves* (8-1) *with $\langle w(0), Q\rangle = \langle w(0), Q'\rangle = 0$ and $w = v + u$, where v solves* (7-2) *with initial data $w(0)$. Then,*

$$\sup_t |\langle w(t), Q\rangle| \lesssim (J + \|u\|_{\dot X^{-1/6}_\infty})^2 (1 + \|w\|_{\dot X^{-1/6}_\infty})^2, \tag{8-24}$$

$$\sup_t |\langle w(t), Q'\rangle| \lesssim (J + \|u\|_{\dot X^{-1/6}_\infty})^2 (1 + \|w\|_{\dot X^{-1/6}_\infty})^2 + \kappa^{-1/2} \|w\|_{\dot X^{-1/6}_{\infty,T}}. \tag{8-25}$$

Moreover, if $1 \le p < 3/2$, then

$$\Big\| \frac{d}{dt} \langle w(t), Q\rangle \Big\|_{L^p(0,\infty)} \lesssim (J + \|u\|_{\dot X^{-1/6}_\infty})^2 (1 + \|w\|_{\dot X^{-1/6}_\infty})^2.$$

We may write $\frac{d}{dt}\langle w(t), Q'\rangle = \gamma_1 + \gamma_2$ such that

$$\|\gamma_1\|_{L^p(0,\infty)} \lesssim (J + \|u\|_{\dot X^{-1/6}_\infty})^2 (1 + \|w\|_{\dot X^{-1/6}_\infty})^2 \quad \text{and} \quad \|\gamma_2\|_{L^2(0,\infty)} \lesssim (J + \|u\|_{\dot X^{-1/6}_\infty}).$$

Finally, it follows that

$$\begin{aligned} \sup_t (|c(t) - 1| + |\dot c|) + \|\dot c\|_{L^1} &\lesssim (J + \|u\|_{\dot X^{-1/6}_\infty})^2 (1 + \|w\|_{\dot X^{-1/6}_\infty})^2, \\ \sup_t |\dot y - 1| &\lesssim (J + \|u\|_{\dot X^{-1/6}_\infty})^2 (1 + \|w\|_{\dot X^{-1/6}_\infty})^2 + \kappa^{-1/2} (J + \|u\|_{\dot X^{1/6}_\infty}), \\ \|\dot y - c^2\|_{L^2} &\lesssim (J + \|u\|_{\dot X^{-1/6}_\infty})(1 + \|w\|_{\dot X^{-1/6}_\infty})^3. \end{aligned} \tag{8-26}$$

Proof. This is an immediate consequence of Lemma 8.5 and basic properties of the simple ordinary differential equations. □

The estimates of this subsection remain true if we consider a time integral instead of $(0, \infty)$.

8.3. *Global bounds and scattering near the soliton.* We now complete the proof of Proposition 8.2.

Proof. By the local existence result there exists a local solution in a neighborhood of the soliton.

The decomposition $\psi = Q_{c(t),y(t)} + w$ together with the modal equations (8-3) and (8-4) implies existence of C^1 functions $c(t)$ and $y(t)$ that satisfy (8-1), (8-3) and (8-4) up to fixed time. We recall that after rescaling and shifting, $\langle w_0, Q\rangle = \langle w_0, Q'\rangle = 0$, $c(0) = 1$ and $y(0) = 0$.

As in the first step we denote the solution to the linear equation with initial data $w(0)$ by v. It satisfies the estimates of Lemma 8.3 and (8-13) provided (5-14) is satisfied.

We suppose that (ψ, c, y) is a solution up to time T, such that $u = \psi - Q_{c(t),y(t)} - v$ satisfies for some k_1, k_2 to be chosen later the conditions

$$\|P^{\perp}_{Q'} u\|_{\dot X^{-1/6}_{\infty,T}} \le 2k_1 J^{1/2} \quad \text{and} \quad \|u\|_{\dot X^{-1/6}_{\infty,T}} \le 2k_2 J^{1/2}. \tag{8-27}$$

We shall see that there exist δ, k_1 and k_2 such that if in addition $J \le \delta$, then

$$\|P^{\perp}_{Q'} u\|_{\dot X^{-1/6}_{\infty,T}} \le k_1 J^{1/2} \quad \text{and} \quad \|u\|_{\dot X^{-1/6}_{\infty,T}} \le k_2 J^{1/2}. \tag{8-28}$$

This implies the estimate conditionally depending on (8-27). Observe that by Lemma 8.7 control of the norms implies validity of (5-14) if δ is sufficiently small. In particular, the estimates on the linear equations hold.

On the other hand, if we fix C and δ we can apply a continuity argument with the initial data τw_0. The estimate clearly holds for small τ and the norms depend (for finite time) continuously on τ. This implies the a priori estimate uniformly for all T. The scattering statement is an immediate consequence: Combine the fact that functions in V^2 are left-continuous at infinity with a frequency envelope argument as above. It remains to derive (8-28) from (8-27) for suitably chosen k_1, k_2 and δ.

We formulate the crucial estimate in the following.

Lemma 8.8. *Let C be given and let v and Q be as in Proposition 8.2. There exist k_1, k_2 and δ such that, if* (8-27) *holds, and $\|w_0\|_{\dot B^{-1/6,2}_{\infty}} \le C$ and $J_{(0,T)}(v) \le \delta$ hold, then*

$$\|P^{\perp}_{Q'} u\|_{\dot X^{-1/6}_{\infty,T}} \le c_3\big(\|u\|^2_{\dot X^{-1/6}_{\infty,T}} + J^{1/2}_{(0,T)}(v)\|w_0\|^{3/2}_{\dot B^{-1/6,2}_{\infty}} + J_{(0,T)}(v)\big).$$

We postpone the proof of Lemma 8.8. Clearly $\langle u, Q'\rangle = \langle w, Q'\rangle$ and the same is true for its derivatives. By Lemma 8.7 and simple properties of ODEs, we have with implicit constants depending on the size of the initial data that

$$\Big\|\langle u, Q'\rangle\Big\|_{L^1+L^2} + \Big\|\frac{d}{dt}\langle u, Q'\rangle\Big\|_{L^1+L^2} \lesssim \|u\|^2_{\dot X^{-1/6}_{\infty}} + J + \|\langle w, Q''\rangle\|_{L^2}$$

and

$$\|\langle w, Q''\rangle\|_{L^2} \le \|\langle v, Q''\rangle\|_{L^2} + \|\langle P^{\perp}_{Q'} u, Q''\rangle\|_{L^2}.$$

Hence,

$$\|\langle u, Q'\rangle\|_{L^1+L^2} + \Big\|\frac{d}{dt}\langle u, Q'\rangle\Big\|_{L^1+L^2} \lesssim \|u\|^2_{\dot X^{-1/6}_{\infty}} + J + \|P^{\perp}_{Q'} u\|_{\dot X^{-1/6}_{\infty}}.$$

The crucial point is that the right hand side only contains the projection of u, not u itself. We obtain easily

$$\|(\partial_t + \partial_x^3)\gamma(t) Q'\|_{\dot Y^{-1/6}_{\infty,T}} \lesssim \|\gamma\|_{L^2} + \|\gamma'\|_{L^1+L^2}.$$

As a result, using estimates similar to those in Lemma 8.7 we have

$$\Big\| \frac{\langle u, Q'\rangle}{\langle Q', Q'\rangle} Q' \Big\|_{X^{-1/6}_{\infty,T}} \le k_4\big(\|u\|^2_{\dot X^{-1/6}_{\infty,T}} + J + \|P^\perp u\|_{\dot X^{-1/6}_{\infty,T}}\big). \tag{8-29}$$

By Lemma 8.8 and (8-27),

$$\|P^\perp_{Q'} u\|_{\dot X^{-1/6}_{\infty,T}} \le c_3(\|u\|^2_{\dot X^{-1/6}_{\infty,T}} + J^{1/2}_{(0,T)}(v) + J) \le c_3(4k_2^2+1)J + c_3 J^{1/2}$$

using, as we may, $\|u\|_{\dot X^{-1/6}_{\infty,T}} \le 1$ and, by the estimate (8-29) and (8-27), we have

$$\|u\|_{\dot X^{-1/6}_{\infty,T}} \le k_4\big(J + c_3(8k_2^2+1)J(v) + c_3 J^{1/2}\big).$$

We choose first k_1, then k_2 and finally δ small to complete the proof. □

Proof of Lemma 8.8. We write the equation for $u = w - \nu$, with $u_t + c^2 u_x - \partial_x \mathcal{L}_{c,y} u =: G$. We have

$$G = (\dot c/c + \alpha_1)\widetilde Q + (\dot y - c^2 + \beta_1)Q' - \partial_x\big(6Q_c^2(u+\nu)^2 + 4Q_c'(u+\nu)^3 + (u+\nu)^4\big),$$

where α_1 and β_1 ensure the orthogonality conditions for ν, that is, (8-12). We recall that they satisfy

$$\|\alpha_1\|_{L^1} \lesssim \|\dot c\|_{L^2} J \quad \text{and} \quad \|\alpha_1\|_{L^2} + \|\beta_1\|_{L^2} \lesssim J.$$

To apply Proposition 6.7 we have to project u. This leads to a calculation similar to Lemma 8.3. Let $\mu = P^\perp_{Q'} \tilde P u$ and $\mu_t + c^2\partial_x \mu - \partial_x \mathcal{L}\mu =: H$. Then, using $\langle u, Q\rangle = \langle w, Q\rangle$ and $\langle u, Q'\rangle = \langle w, Q'\rangle$,

$$\begin{aligned} H = G &- \Big(\frac{d}{dt}\frac{\langle w, Q\rangle}{\langle Q, \widetilde Q\rangle}\Big)\widetilde Q - \Big(\frac{d}{dt}\frac{\langle w, Q'\rangle}{\langle Q', Q'\rangle}\Big)Q' \\ &\qquad - \frac{\langle w, Q\rangle}{\langle Q, \widetilde Q\rangle}\Big(\frac{\dot c}{c}\tilde{\widetilde Q} + (c^2 - \dot y)\widetilde Q' + 2Q'\Big) - \frac{\langle w, Q'\rangle}{\langle Q', Q'\rangle}\big((\dot c/c)\widetilde Q' + (c^2-\dot y)Q''\big) \\ &= \alpha \widetilde Q + \beta Q' + g, \end{aligned}$$

where

$$\begin{aligned} -g = \partial_x\big(6Q_c^2(u+\nu)^2 + 4Q_c'(u+\nu)^3 + (u+\nu)^4\big) + \frac{\langle w, Q\rangle}{\langle Q, \widetilde Q\rangle}\frac{\dot c}{c}\tilde{\widetilde Q} \\ + \Big(\frac{\langle w, Q\rangle}{\langle Q, \widetilde Q\rangle}(c^2 - \dot y) + \frac{\langle w, Q'\rangle}{\langle Q', Q'\rangle}\frac{\dot c}{c}\Big)\widetilde Q' + \frac{\langle w, Q'\rangle}{\langle Q', Q'\rangle}(c^2-\dot y)Q''. \end{aligned}$$

By construction $u(0) = 0$. We apply Proposition 6.7 to get

$$\|u\|_{\dot X^s_{\infty,T}} \lesssim \|g\|_{\dot Y^{-1/6}_{\infty,T}} + \|\langle g, Q\rangle\|_{L^1} + \|\langle g^+, Q\rangle\|_{L^2} + \|\langle g, Q'\rangle\|_{L^2}.$$

By Lemma 7.1, Lemma 8.5 and Lemma 8.6, we get

$$\|g\|_{\dot Y^{-1/6}_\infty} \lesssim \|u\|^2_{\dot X^{-1/6}_\infty} + J\|w_0\|_{\dot B^{-1/6,2}_\infty},$$

and by Lemma 8.7

$$\|\langle g, Q\rangle\|_{L^1} + \|\langle g^+, Q\rangle\|_{L^2} + \|\langle g, Q\rangle\|_{L^2} \lesssim \|u\|^2_{\dot X^{-1/6}_\infty} + J^{1/2}\|w_0\|^{3/2}_{\dot B^{-1/6,2}_\infty}.$$

Together, we have

$$\|\tilde{P}^* P^\perp u\|_{\dot{X}^{-1/6}_{\infty,T}} \lesssim \|u\|^2_{\dot{X}^{-1/6}_{\infty,T}} + J^{1/2}\|w_0\|^{3/2}_{\dot{B}^{-1/6,2}_\infty}. \qquad \square$$

Proposition 8.2 generalizes straightforwardly to smaller function spaces in the style of Section 7.1.

8.4. ***An almost inverse wave operator result.*** In this section we will construct solutions with given asymptotic behavior, proving Theorem 3. This is a partial converse statement to Proposition 8.2.

Remark 8.2. Theorem 3 is quite satisfactory in several respects. It shows which asymptotic properties may characterize a solution. The main missing piece is uniqueness of the solution Ψ. It implies existence of a solution for small scattering data, and, for arbitrary scattering states, existence of a solution with given scattering data for large t.

Proof. We turn to the time-reversed equation

$$\partial_t w + \partial_x(\partial_x^2 w + 4Q_c^3 w) = (\dot{y} - c^2)\langle w, Q_{xx}\rangle + (\dot{c}/c)\langle w, \tilde{Q}'\rangle + \langle 6Q^2w^2 + 4Qw^3 + w^4, Q_{xx}\rangle \quad \text{(8-30)}$$

with

$$(\dot{c}/c)\langle Q_c, \tilde{Q}_c\rangle = -\langle w, Q_c\rangle \quad \text{and} \quad (\dot{y} - c^2)\langle Q_c', Q_c'\rangle = \kappa\langle w, Q_c'\rangle.$$

Let v be the solution to the Airy equation with initial data v_0. We may and do assume that $y_0 = 0$. By Proposition 8.1 we know that $\lim_{t\to\infty} J_{[t,\infty)}(v) = 0$. Given $S > 0$ and y^S satisfying $|y^S(S) - c_\infty^2 S| < \hat{\delta} S$, we solve the backwards initial value problem

$$\Psi(S) = v(S) + Q_{c_\infty, y^S}.$$

We choose $1 \gg \hat{\delta} \gg \delta$ to ensure that $|\dot{y} - c^\infty| \le \hat{\delta}$ for the solutions under consideration. The arguments of the previous section allow one to do that down to a largest time t^{S,y^S} for which

$$|y(t^{S,y^S}) - c_\infty^2 t^{S,y^S}| = \hat{\delta} t^{S,y^S}.$$

We want to show that the infimum of the t^{S,y^S} as a function of y^S is attained for some y^S and it is equal to zero if $\hat{\delta}$ is sufficiently small. Suppose not, and denote the infimum by $\tau > 0$. By continuous dependence on y^S, given $\varepsilon > 0$, there exists an interval $[a, b]$ such that the solution exists down to a time smaller than $(1+\varepsilon)\tau$, and $y^{S,a}((1+\varepsilon)\tau) = (c_\infty^2 - \delta)(1+\varepsilon)\tau$ and $y^{S,b}((1+\varepsilon)\tau) = (c_\infty^2 + \delta)(1+\varepsilon)\tau$. Hence, there exists $y^{S,\varepsilon}$ with

$$y^{S,a}((1+\varepsilon)\tau) = c_\infty^2(1+\varepsilon)\tau.$$

But then, if $\hat{\delta}$ is sufficiently small, we see that a positive infimum is not possible, and moreover this construction gives a limit that is a solution denoted again by (Ψ^S, y^S) with $y^S(0) = 0$.

We consider the limit $S \to \infty$. Since $\dot{y}^S - c_\infty^2$ and $\dot{c}^S$ are small there exists a converging subsequence y^{S_j}, c^{S_j}, $S_j \to \infty$ that converges to c and y. There are corresponding solutions Ψ_j, u_j and w_j of the corresponding equation. We extend w_j beyond S_j by v. By the stability result, given $\delta > 0$ we find $T > 0$ such that

$$\|w_j - v\|_{\dot{X}^{-1/6}_{\infty,[T,\infty)}} \le \delta.$$

Using a frequency envelope there exists Λ such that

$$\lambda^{-1/6}\|(w_j)_\lambda\|_{V^2(T,\infty)} \lesssim \delta$$

whenever $\lambda > \Lambda$ or $\lambda^{-1} > \Lambda$.

In particular,

$$\|(w_j - w_l)(t)\|_{\dot{B}^{-1/6,2}_\infty} \le \delta$$

for $t \ge T(\delta)$ and j and l sufficiently big. Again, using J small we are able to deduce that (w_j) is a Cauchy sequence in $\dot{X}_T^{-1/6}$ and the limit is the desired solution. □

Acknowledgments

We wish to thank Axel Grünrock and Yvan Martel for helpful comments on an early version of the result and the anonymous referee for helping to improve the paper in several places.

Appendix: Setup and properties of the U^p, V^p spaces for the linear KdV equation

To define the function spaces U^2, V^2, we summarize [Hadac et al. 2009, Section 2], where we suggest the reader look for further details. Let $\mathfrak{Z}$ be the set of finite partitions $-\infty < t_0 < t_1 < \ldots < t_K = \infty$. In the following, we consider functions taking values in $L^2 := L^2(\mathbb{R}^d;\mathbb{C})$, but in the general part of this section L^2 may be replaced by an arbitrary Hilbert space.

Definition A.1. Let $1 \le p < \infty$. For $\{t_k\}_{k=0}^K \in \mathfrak{Z}$ and $\{\phi_k\}_{k=0}^{K-1} \subset L^2$ with $\sum_{k=0}^{K-1}\|\phi_k\|_{L^2}^p = 1$ we call the function $a : \mathbb{R} \to L^2$ given by

$$a = \sum_{k=1}^{K} \chi_{[t_{k-1},t_k)}\phi_{k-1}$$

a U^p-atom, where χ_I is the standard cutoff function to interval I. Furthermore, we define the atomic space

$$U^p := \Big\{ u = \sum_{j=1}^{\infty} \lambda_j a_j \;\Big|\; a_j \text{ a } U^p\text{-atom,, } \lambda_j \in \mathbb{C} \text{ such that } \sum_{j=1}^{\infty} |\lambda_j| < \infty \Big\}$$

with norm

$$\|u\|_{U^p} := \inf\Big\{ \sum_{j=1}^{\infty} |\lambda_j| \;\Big|\; u = \sum_{j=1}^{\infty} \lambda_j a_j,\ \lambda_j \in \mathbb{C},\ a_j \text{ a } U^p\text{-atom} \Big\}. \tag{A-1}$$

Atoms are bounded in the supremum norm, and hence every convergence here implies uniform convergence.

Proposition A.2. *Let $1 \le p < q < \infty$.*

(1) *The expression $\|\cdot\|_{U^p}$ is a norm. The space U^p is complete and hence a Banach space.*

(2) *The embeddings $U^p \subset U^q$ have norm* 1.

(3) *For* $u \in U^p$, *all one sided limits exist, including at* $\pm\infty$, u *is continuous from the right, and the limit at* $-\infty$ *is zero.*

(4) *The subspace of continuous functions* U^p_c *is closed.*

Definition A.3. Let $1 \le p < \infty$. We define V^p as the normed space of all functions $v : \mathbb{R} \to L^2$ for which the norm

$$\|v\|_{V^p} := \sup_{\{t_k\}_{k=0}^K \in \mathcal{Z}} \left(\sum_{k=1}^{K} \|v(t_k) - v(t_{k-1})\|_{L^2}^p \right)^{1/p} \tag{A-2}$$

is finite. Here we understand $v(\infty)$ as zero. Let V^p_- denote the subspace of all right-continuous functions with limit 0 at $-\infty$.

Taking the partition $\{t, \infty\}$, one sees that the supremum norm is not larger than the V^p norm.

Proposition A.4. *Let* $1 \le p < q < \infty$.

(1) *The expression* $\|\cdot\|_{V^p}$ *is a norm and* V^p *is complete.*

(2) *For* $v \in V^p$, *all one sided limits including at* $\pm\infty$ *exist.*

(3) *The subspace* V^p_- *is closed.*

(4) *The embedding* $U^p \subset V^p_-$ *is continuous and* $\|u\|_{V^p} \le 2^{1/p}\|u\|_{U^p}$.

(5) *The embeddings* $V^p \subset V^q$ *are continuous and* $\|v\|_{V^q} \le \|v\|_{V^p}$.

From the proof of [Hadac et al. 2009, Proposition 2.17], we have the following:

Lemma A.5. *Let* $f \in V^p_-$, *with* $q > p$. *Then, given* $\delta > 0$ *and* $m > 1$, *there exist* $f_1 \in U^p$ *and* $f_2 \in U^q$ *such that* $f = f_1 + f_2$ *and*

$$m^{-1}\|f_1\|_{U^p} + e^{\delta m}\|f_2\|_{U^q} \lesssim \|f\|_{V^p}.$$

The following corollary is obvious.

Corollary A.6. *The space* V^p_- *is continuously embedded in* U^q *for* $q > p$.

There is a bilinear map, B, which for $1/p + 1/q = 1$ and $1 < p, q < \infty$ can formally be written as

$$B(f, g) = -\int f_t g \, dt, \quad \text{for } f \in V^p, \, g \in U^q.$$

It satisfies $|B(f, g)| \le \|f\|_{V^p}\|g\|_{U^q}$, which is natural if we replace g by an atom. The map

$$V^p \ni f \to (g \to B(f, g)) \in (U^q)^*$$

is an isometric bijection. Moreover,

$$\|u\|_{U^p} = \sup\{B(u, v) : v \in C(\mathbb{R}), \, \|v\|_{V^p} = 1\}.$$

If $v \in V^p_-$, then

$$\|v\|_{V^q} = \sup\{B(u, v) : u \in C(\mathbb{R}), \, \|u\|_{U^p} = 1\}.$$

If the distributional derivative of u is in L^1 and $v \in V^p$, then

$$B(u,v) = -\int u_t v \, dt.$$

If $f \in L^1$, then $F(t) = \int_{-\infty}^t f ds \in V^p$ for all $p \geq 1$, and hence $F \in U^p$. Moreover, $\|f\|_{DU^p} := \|F\|_{U^p} \leq \|f\|_{L^1}$. We denote by DU^p the metric completion of L^1 in the norm given by the duality pairing. Similarly we define DV^q.

There is a close relation to Besov spaces, namely

$$B_1^{1/p,p} \subset U^p \subset V^p \subset B_\infty^{1/p,p} \tag{A-3}$$

with continuous embeddings. These embeddings clarify the relation to $X^{s,b}$ spaces below.

We claim that the convolution with an L^1 function η defines a bounded operator on U^p and V^p with norm $\leq \|\eta\|_{L^1}$. Because of the duality statement it suffices to verify boundedness on U^p. We approximate the characteristic function by a sum of Dirac measures. The convolution with an atom clearly has norm at most 1. Convergence in U^1 to the convolution with the characteristic function is immediate. The full statement is an immediate consequence, as well as the boundedness of the convolution by a Schwarz function on U^p and V^p. In particular smooth projections on high and low frequencies are bounded.

Following Bourgain's strategy for the Fourier restriction spaces, we define the adapted function spaces

$$U^p_{KdV} = S(-t)U^p \quad \text{and} \quad V^p_{KdV} = S(-t)V^p$$

and similarly DU^p and DV^p.

Again, we define a bilinear map B_{KdV} such that for $u \in V^p_{KdV}$ and $v \in U^q_{KdV}$, we have for a function u with $(\partial_t + \partial_x^3)u \in L^1L^2$

$$B_{KdV}(u,v) = -\int \langle (\partial_t + \partial_x^3)u, v\rangle \, dt.$$

Note, this bilinear map is well defined and gives a duality relation. Hence,

$$\|u\|_{DV^p_{KdV}} = \sup_{\|f\|_{U^q_{KdV}} \leq 1} \int uf \, dx \, dt \quad \text{and} \quad \|u\|_{DU^p_{KdV}} = \sup_{\|f\|_{V^q_{KdV}} \leq 1} \int uf \, dx \, dt.$$

Moreover, we may restrict f to suitable subspaces. More details on how the construction of such atomic spaces allows us to put u_t in the dual space are included in [Hadac et al. 2009].

By the construction of our spaces, we obtain for a solution u of the linear KdV equation

$$\begin{cases} u_t + u_{xxx} = f, \\ u(0,x) = u_0(x), \end{cases} \tag{A-4}$$

the estimates

$$\|u\|_{V^2_{KdV}} \lesssim \|u_0\|_{L^2} + \|f\|_{DV^2_{KdV}}, \tag{A-5}$$

$$\|u\|_{U^2_{KdV}} \lesssim \|u_0\|_{L^2} + \|f\|_{DU^2_{KdV}}, \tag{A-6}$$

which follow trivially from the construction of the V^2_{KdV}, DV^2_{KdV} and U^2_{KdV}, DU^2_{KdV} spaces.

Spatial Fourier multipliers act on U^p, V^q, DU^p, DV^q in the obvious way and their operator norm is bounded by the supremum of the multiplier.

Let (p, q) be a Strichartz pair. Then,

$$\|u\|_{L^p L^q} \le c \| |D|^{-1/p} u\|_{U^p}$$

and the dual estimate

$$\|f\|_{DV^{p'}} \le c \| |D|^{-1/p} f\|_{L^{p'} L^{q'}}$$

hold. The first estimate is not hard to check on atoms. Since convergence in U^p and in $L^p L^q$ both imply pointwise convergence for subsequences we obtain the full estimate. The second estimate follows by duality.

Similarly the local smoothing estimates carry over to U^p spaces and to DV^q. Let $c(t)$ and $y(t)$ satisfy (5-14). Then

$$\|u\|_{L^2 H^1_{\sqrt{\gamma'}})} \le c\|u\|_{U^2} \quad \text{and} \quad \|f\|_{DV^2} \le c\|f\|_{L^2 \sqrt{\gamma'} H^{-1}}.$$

In the same fashion the bilinear estimates for solutions to the free equation imply bilinear estimates for functions in U^2.

The smooth decomposition into high and low modulation (that is, the smooth projection of the frequencies to $\tau - \xi^3$ large and respectively small) is bounded in U^2 and V^2, and the L^2 norm of the high modulation part gains the inverse of square root of the truncation as a factor by the embeddings (A-3).

References

[Andrews et al. 1999] G. E. Andrews, R. Askey, and R. Roy, *Special functions*, Encyclopedia of Mathematics and its Applications **71**, Cambridge University Press, 1999. MR 2000g:33001 Zbl 0920.33001

[Berestycki and Lions 1983] H. Berestycki and P.-L. Lions, "Nonlinear scalar field equations, I: Existence of a ground state", *Arch. Rational Mech. Anal.* **82**:4 (1983), 313–345. MR 84h:35054a Zbl 0533.35029

[Côte 2006] R. Côte, "Construction of solutions to the subcritical gKdV equations with a given asymptotical behavior", *J. Funct. Anal.* **241**:1 (2006), 143–211. MR 2007h:35279 Zbl 1157.35091

[Fedoryuk 1993] M. V. Fedoryuk, *Asymptotic analysis*, Springer, Berlin, 1993. MR 95m:34091 Zbl 0782.34001

[Grünrock 2005] A. Grünrock, "A bilinear Airy-estimate with application to gKdV-3", *Differential Integral Equations* **18**:12 (2005), 1333–1339. MR 2007b:35282 Zbl 1212.35412

[Hadac et al. 2009] M. Hadac, S. Herr, and H. Koch, "Well-posedness and scattering for the KP-II equation in a critical space", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26**:3 (2009), 917–941. MR 2010d:35301 Zbl 1169.35372

[Kato 1983] T. Kato, "On the Cauchy problem for the (generalized) Korteweg-de Vries equation", pp. 93–128 in *Studies in applied mathematics*, Adv. Math. Suppl. Stud. **8**, Academic Press, New York, 1983. MR 86f:35160 Zbl 0549.34001

[Kenig et al. 1993] C. E. Kenig, G. Ponce, and L. Vega, "Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle", *Comm. Pure Appl. Math.* **46**:4 (1993), 527–620. MR 94h:35229 Zbl 0808.35128

[Lamb 1980] G. L. Lamb, Jr., *Elements of soliton theory*, Wiley, New York, 1980. MR 82f:35165 Zbl 0445.35001

[Martel 2006] Y. Martel, "Linear problems related to asymptotic stability of solitons of the generalized KdV equations", *SIAM J. Math. Anal.* **38**:3 (2006), 759–781. MR 2007i:35204 Zbl 1126.35055

[Martel and Merle 2001a] Y. Martel and F. Merle, "Asymptotic stability of solitons for subcritical generalized KdV equations", *Arch. Ration. Mech. Anal.* **157**:3 (2001), 219–254. MR 2002b:35182 Zbl 0981.35073

[Martel and Merle 2001b] Y. Martel and F. Merle, "Instability of solitons for the critical generalized Korteweg-de Vries equation", *Geom. Funct. Anal.* **11**:1 (2001), 74–123. MR 2002g:35182 Zbl 0985.35071

[Martel and Merle 2005] Y. Martel and F. Merle, "Asymptotic stability of solitons of the subcritical gKdV equations revisited", *Nonlinearity* **18**:1 (2005), 55–80. MR 2006i:35319 Zbl 1064.35171

[Martel and Merle 2008] Y. Martel and F. Merle, "Asymptotic stability of solitons of the gKdV equations with general nonlinearity", *Math. Ann.* **341**:2 (2008), 391–427. MR 2008k:35416 Zbl 1153.35068

[Pego and Weinstein 1994] R. L. Pego and M. I. Weinstein, "Asymptotic stability of solitary waves", *Comm. Math. Phys.* **164**:2 (1994), 305–349. MR 95h:35209 Zbl 0805.35117

[Tao 2007] T. Tao, "Scattering for the quartic generalised Korteweg-de Vries equation", *J. Differential Equations* **232**:2 (2007), 623–651. MR 2008i:35178 Zbl 1171.35107

[Titchmarsh 1962] E. C. Titchmarsh, *Eigenfunction expansions associated with second-order differential equations, Part I*, 2nd ed., Clarendon Press, Oxford, 1962. MR 31 #426 Zbl 0099.05201

[Weinstein 1985] M. I. Weinstein, "Modulational stability of ground states of nonlinear Schrödinger equations", *SIAM J. Math. Anal.* **16**:3 (1985), 472–491. MR 86i:35130 Zbl 0583.35028

Received 1 Feb 2010. Revised 11 Aug 2010. Accepted 1 Oct 2010.

HERBERT KOCH: koch@math.uni-bonn.de
Mathematisches Institut, Universität Bonn, Endenicher Allee 60, D-53115 Bonn, Germany

JEREMY L. MARZUOLA: marzuola@math.unc.edu
Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599, United States

ANALYSIS AND PDE
Vol. 5, No. 1, 2012
dx.doi.org/10.2140/apde.2012.5.199

A REMARK ON BARELY $\dot{H}^{s_p}$-SUPERCRITICAL WAVE EQUATIONS

TRISTAN ROY

We prove that a good $\dot{H}^{s_p}$ critical theory for the $3D$ wave equation $\partial_{tt}u - \triangle u = -|u|^{p-1}u$ can be extended to prove global well-posedness of smooth solutions of at least one $3D$ barely $\dot{H}^{s_p}$-supercritical wave equation $\partial_{tt}u - \triangle u = -|u|^{p-1}ug(|u|)$, with g growing slowly to infinity, provided that a Kenig-Merle type condition is satisfied. This result is related to those obtained by Tao and the author for the particular case $s_p = 1$, showing global regularity for g growing logarithmically with radial data and for g growing doubly logarithmically with general data.

1. Introduction

For fixed $p > 3$, let $\tilde{H}^2 := \dot{H}^2(\mathbb{R}^3) \cap \dot{H}^{s_p}(\mathbb{R}^3)$ and $\tilde{H}^1 := \dot{H}^1(\mathbb{R}^3) \cap \dot{H}^{s_p-1}(\mathbb{R}^3)$, where $s_p := \frac{3}{2} - \frac{2}{p-1}$. We consider the wave equation

$$\begin{cases} \partial_{tt}u - \triangle u = -|u|^{p-1}ug(|u|), \\ u(0) := u_0 \in \tilde{H}^2, \\ \partial_t u(0) := u_1 \in \tilde{H}^1, \end{cases} \tag{1-1}$$

where $u : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{C}$ is a complex-valued scalar field and g is a smooth, real-valued positive function defined on the set of nonnegative numbers and satisfying

$$0 \leq g'(x) \lesssim \frac{1}{x}. \tag{1-2}$$

This condition says that g grows more slowly than any positive power of u.

We shall see that (1-1) has many connections with the defocusing power-type wave equation

$$\begin{cases} \partial_{tt}u - \triangle u = -|u|^{p-1}u, \\ u(0) := u_0 \in \dot{H}^{s_p}(\mathbb{R}^3), \\ \partial_t u(0) := u_1 \in \dot{H}^{s_p-1}(\mathbb{R}^3). \end{cases} \tag{1-3}$$

It is known that if u satisfies (1-3), then u_λ defined by

$$u_\lambda(t,x) := \frac{1}{\lambda^{2/(p-1)}} u\left(\frac{t}{\lambda}, \frac{x}{\lambda}\right), \tag{1-4}$$

satisfies the same equation, but with data

$$u_\lambda(0,x) = \frac{1}{\lambda^{2/(p-1)}} u_0\left(\frac{x}{\lambda}\right) \quad \text{and} \quad \partial_t u_\lambda(0,x) = \frac{1}{\lambda^{2/(p-1)+1}} u_1\left(\frac{x}{\lambda}\right).$$

MSC2000: 35Q55.
Keywords: wave equation, global existence, barely supercritical.

Notice that (1-3) is $\dot{H}^{s_p}(\mathbb{R}^3)$ critical, which means that the $\dot{H}^{s_p}(\mathbb{R}^3) \times \dot{H}^{s_p-1}(\mathbb{R}^3)$-norm of $(u(0), \partial_t u(0))$ is invariant under the scaling defined above.

We recall the local existence theory. From [Ginibre and Velo 1989; Lindblad and Sogge 1995], we know that there exists a positive constant $\delta := \delta\big(\|(u_0, u_1)\|_{\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)}\big) > 0$ and a time of local existence $T_l > 0$ such that if

$$\left\| \cos(tD)u_0 + \frac{\sin(tD)}{D} \right\|_{L_t^{2(p-1)} L_x^{2(p-1)}([0,T_l]\times\mathbb{R}^3)} \leq \delta \tag{1-5}$$

then there exists a unique solution $(u, \partial_t u)$ in

$$\mathscr{C}\left([0, T_l],\ \dot{H}^{s_p}(\mathbb{R}^3)\right) \cap L_t^{2(p-1)} L_x^{2(p-1)}([0, T_l]\times\mathbb{R}^3) \cap D^{\frac{1}{2}-s_p} L_t^4 L_x^4([0, T_l]\times\mathbb{R}^3) \times \mathscr{C}\left([0, T_l],\ \dot{H}^{s_p-1}(\mathbb{R}^3)\right)$$

of (1-3)[1] in the integral equation sense, i.e., u satisfies the Duhamel formula

$$u(t) := \cos(tD)u_0 + \frac{\sin(tD)}{D}u_1 - \int_0^t \frac{\sin(t-t')D}{D}\left(|u|^{p-1}u\right)(t')\,dt'. \tag{1-6}$$

It follows that we can define a maximal time interval of existence $I_{\max} = (-T_-, T_+)$. Moreover,

$$\|u\|_{L_t^{2(p-1)} L_x^{2(p-1)}(J)} < \infty, \quad \|D^{s_p-\frac{1}{2}}u\|_{L_t^4 L_x^4(J)} < \infty, \quad \text{and} \quad \|(u, \partial_t u)\|_{L_t^\infty \dot{H}^{s_p}\times L_t^\infty \dot{H}^{s_p-1}(J)} < \infty$$

for any compact subinterval $J \subset I_{\max}$. See [Kenig and Merle 2006] or [Tao 2006a] for more explanations.

Now we turn to the global well-posedness theory of "(1-3)". In view of the local well-posedness theory, one can prove (see [Kenig and Merle 2011] and references), after some effort, that it is enough to find a finite upper bound of $\|u\|_{L_t^{2(p-1)} L_x^{2(p-1)}(I\times\mathbb{R}^3)}$ on arbitrary long time intervals I, and, if this is the case, then the solution scatters to a solution of the linear wave equation. No blow-up has been observed for (1-3). Therefore it is believed that the following scattering conjecture is true:

Conjecture 1.1 (scattering conjecture). *Assume that u is the solution of* (1-3) *with data* $(u_0, u_1) \in \dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)$. *Then u exists for all time t and there exists* $C_1 := C_1\big(\|(u_0, u_1)\|_{\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)}\big)$ *such that*

$$\|u\|_{L_t^{2(p-1)} L_x^{2(p-1)}(\mathbb{R}\times\mathbb{R}^3)} \leq C_1. \tag{1-7}$$

The case $s_p = 1$ (equivalently, $p = 5$) is particular. Indeed the solution

$$(u, \partial_t u) \in \mathscr{C}\big([0, T_l],\ \dot{H}^1(\mathbb{R}^3)\big) \times \mathscr{C}\big([0, T_l],\ L^2(\mathbb{R}^3)\big)$$

satisfies the conservation of the energy $E(t)$ defined by

$$E(t) := \frac{1}{2}\int_{\mathbb{R}^3} |\partial_t u(t,x)|^2\,dx + \frac{1}{2}\int_{\mathbb{R}^3} |\nabla u(t,x)|^2\,dx + \frac{1}{6}\int_{\mathbb{R}^3} |u|^6(t,x)\,dx. \tag{1-8}$$

[1]The $L_t^{2(p-1)} L_x^{2(p-1)}(\mathbb{R}\times\mathbb{R}^3)$-norm of u is invariant under the scaling (1-4). The choice of the space $L_t^{2(p-1)} L_x^{2(p-1)}$ in which we place the solution u is not unique. There exists an infinite number of spaces of the form $L_t^q L_x^r$ scale invariant in which we can establish a local well-posedness theory.

In other words, $E(t) = E(0)$. This is why this equation is often called energy-critical: the exponent $s_p = 1$ corresponds precisely to the minimal regularity required for (1-8) to be defined. The global well-posedness of (1-8) in the energy class and in higher regularity spaces is now understood. Rauch [1981] proved the global existence of smooth solutions of this equation with small data. Struwe [1988] showed that the result still holds for large data but with the additional assumption of spherical symmetry of the data. The general case (large data, no symmetry assumption) was finally settled by Grillakis [1990; 1992]. Shatah and Struwe [1994] and independently Kapitanski [1994] proved global existence of solutions in the energy class. Bahouri and Gérard [1999] reproved this result by using a compactness method and results from Bahouri and Shatah [1998]. In particular, they showed that the $L_t^{2(5-1)} L_x^{2(5-1)}(\mathbb{R} \times \mathbb{R}^3)$-norm of the solution is bounded by an unspecified finite quantity. Lately Tao [2006b] found an exponential tower type bound of this norm. All these proofs of global existence of solutions of the energy-critical wave equation have as a common key point the conservation of energy, which leads, in particular, to the control of the $\dot{H}^1 \times L^2$-norm of the solution $(\partial_t u(t), u(t))$.

If $s_p < 1$, or equivalently, $p < 5$, we are in the energy-subcritical equation. The scattering conjecture is an open problem. Nevertheless, some partial results are known if we consider the same problem (1-3), but with data $(u_0, u_1) \in H^s \times H^{s-1}$, $s_p < s$. More precisely, it is proved in [Kenig et al. 2000; Gallagher and Planchon 2003; Bahouri and Chemin 2006; Roy 2007; Roy 2009a] that there exists $s_0 := s_0(p)$ such that $s_p < s_0 < 1$ and such that (1-3) is globally well-posed in $H^s \times H^{s-1}$, for $s > s_0$.

If $s_p > 1$, or, equivalently, $p > 5$, we are in the energy-supercritical regime. The global behavior of the solution is, in this regime, very poorly understood. Indeed, following the theory of the energy-critical wave equation, the first step would be to prove that the $\dot{H}^{s_p} \times \dot{H}^{s_p-1}$-norm of the solution is bounded for all time by a finite quantity depending only on the $\dot{H}^{s_p} \times \dot{H}^{s_p-1}$-norm of the initial data. Unfortunately, the control of this norm is a very challenging problem, since there are no known conservation laws in high regularity Sobolev spaces. Kenig and Merle [2011] recently proved, at least for radial data, that this step would be the last, by using their concentration compactness/rigidity theorem method [Kenig and Merle 2006]. More precisely, they showed that if $\sup_{t \in I_{\max}} \|(u(t), \partial_t u(t))\|_{\dot{H}^{s_p}(\mathbb{R}^3) \times \dot{H}^{s_p-1}(\mathbb{R}^3)} < \infty$, then Conjecture 1.1 is true.

As mentioned before, the energy supercritical regime is almost *terra incognita*. Nevertheless, Tao [2007] observed that the technology used to prove global well-posedness of smooth solutions of (1-3) can be extended, after some effort, to some equations of the type (1-1), with $p = 5$ and radial data. More precisely, he proved global regularity of (1-1) with $g(x) := \log(2 + x^2)$. This phenomenon, in fact, does not depend on the symmetry of the data: it was proved in [Roy 2009b] that there exists a unique global smooth solution of (1-1) with $g(x) := \log^c \log(10 + x^2)$ and $0 < c < \frac{8}{225}$.

Equations of the type (1-1) are called *barely $\dot{H}^{s_p}$-supercritical* wave equations. Indeed, the condition (1-2) basically says that for every $\epsilon > 0$, there exist two constants $c_1 := c_1(p)$ and $c_2 := c_2(p, \epsilon)$ such that

$$c_1(p) \leq g(|u|) \leq c_2(p, \epsilon)|u|^{\epsilon} \quad \text{for } |u| \text{ large.} \tag{1-9}$$

Since the critical exponent of the equation $\partial_{tt} u - \triangle u = -|u|^{p-1+\epsilon} u$ is $s_{p+\epsilon} = s_p + O(\epsilon)$, the nonlinearity of (1-1) is barely $\dot{H}^{s_p}$-supercritical.

The goal of this paper is to check that this phenomenon, observed for $s_p = 1$, still holds for other values of s_p. The standard local well-posedness theory shows us that it is enough to control the pointwise-in-time $\tilde{H}^2 \times \tilde{H}^1$-norm of the solution. In this paper, we will use an alternative local well-posedness theory. We shall prove:

Proposition 1.2 (local existence for barely $\dot{H}^{s_p}$-supercritical wave equation). *Assume that g satisfies* (1-2) *and*

$$g''(x) = O\left(\frac{1}{x^2}\right). \tag{1-10}$$

Let M be such that $\|(u_0,\ u_1)\|_{\tilde{H}^2\times\tilde{H}^1} \le M$. Then there exists $\delta := \delta(M) > 0$ small such that, if T_l satisfies

$$\left\| \cos(tD)u_0 + \frac{\sin tD}{D}u_1 \right\|_{L_t^{2(p-1)} L_x^{2(p-1)}([0,T_l]\times\mathbb{R}^3)} \le \delta, \tag{1-11}$$

then there exists a unique $(u, \partial_t u)$ in

$$\mathscr{C}\big([0,\ T_l],\ \tilde{H}^2\big) \cap L_t^{2(p-1)} L_x^{2(p-1)}\big([0,\ T_l]\big) \cap D^{\frac{1}{2}-s_p} L_t^4 L_x^4\big([0,\ T_l]\big) \cap D^{\frac{1}{2}-2} L_t^4 L_x^4\big([0,\ T_l]\big) \times \mathscr{C}\big([0,\ T_l],\ \tilde{H}^1\big)$$

that solves (1-1) *in the integral equation sense*; *i.e., u satisfies the Duhamel formula*

$$u(t) := \cos(tD)u_0 + \frac{\sin tD}{D}u_1 - \int_0^t \frac{\sin(t-t')D}{D}\left(|u(t')|^{p-1}u(t')g(|u(t')|)\right) dt'. \tag{1-12}$$

Notice the many similarities between Proposition 1.2 and the local well-posedness theory for (1-3).

This allows us to define a maximum time interval of existence $I_{\max,g} = [-T_{-,g},\ T_{+,g}]$ such that, for any compact subinterval $J \subset I_{\max,g}$, the quantities

$$\|u\|_{L_t^{2(p-1)} L_x^{2(p-1)}(J)}, \quad \|D^{s_p-\frac{1}{2}}u\|_{L_t^4 L_x^4(J)}, \quad \|D^{2-\frac{1}{2}}u\|_{L_t^4 L_x^4(J)}, \quad \|(u,\ \partial_t u)\|_{L_t^\infty \tilde{H}^2(J)\times L_t^\infty \tilde{H}^1(J)}$$

are all finite. Again, see [Kenig and Merle 2006] or [Tao 2006a] for more explanations.

Now we set up the problem. In view of the comments above for $s_p = 1$, we need to make two assumptions. First we will work with a "good" $\dot{H}^{s_p}(\mathbb{R}^3)$ theory: therefore we will assume that Conjecture 1.1 is true. Then, we also would like to work with $\dot{H}^{s_p}(\mathbb{R}^3) \times \dot{H}^{s_p-1}(\mathbb{R}^3)$ bounded solutions $(u(t), \partial_t u(t))$; more precisely, we will assume this:

Condition 1.3 (of Kenig–Merle type). *Let g be a function that satisfies* (1-2) *and that is constant for x large. Then there exists $C_2 := C_2\big(\|(u_0,\ u_1)\|_{\tilde{H}^2\times\tilde{H}^1},\ g\big)$ such that*

$$\sup_{t\in I_{\max,g}} \|\,(u(t), \partial_t u(t))\,\|_{\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)} \le C_2. \tag{1-13}$$

Remark 1.4. In the particular case $s_p = 1$, it is not difficult to see that Condition 1.3 is satisfied. Indeed, u satisfies the energy conservation law

$$E_b(t) := \frac{1}{2}\int_{\mathbb{R}^3} \big(\partial_t u(t,x)\big)^2\, dx + \frac{1}{2}\int_{\mathbb{R}^3} |\nabla u(t,x)|^2\, dx + \int_{\mathbb{R}^3} F(u(t,x), \bar{u}(t,x))\, dx, \tag{1-14}$$

with

$$F(z,\bar{z}) = |z|^{5+1}\int_0^1 t^5\,\mathrm{Re}\,(g(t|z|))\,dt = |z|^{5+1}\int_0^1 t^5\,g(t|z|)\,dt. \tag{1-15}$$

Since g is bounded, we have $|F(z,\bar{z})| \lesssim |z|^6$. By using the Sobolev embeddings $\|u_0\|_{L_x^6} \lesssim \|u_0\|_{\tilde{H}^2}$ and $\|u(t)\|_{L_x^6} \lesssim \|u(t)\|_{\tilde{H}^2}$, we easily conclude that Condition 1.3 holds. The energy conservation law was often in [Tao 2007; Roy 2009b].

Here is the main result of this paper:

Theorem 1.5. *Let p be fixed.*

(1) *There exists a function $\tilde{g}$ satisfying* (1-2) *and*

$$\lim_{x\to\infty} \tilde{g}(x) = \infty \tag{1-16}$$

and such that the solution of (1-1) (*with $g := \tilde{g}$*) *exists for all time*, *provided that the scattering conjecture and Condition 1.3 are satisfied.*

(2) *There exists a function f depending on T and $\|(u_0,u_1)\|_{\tilde{H}^2\times\tilde{H}^1}$ such that*

$$\|u\|_{L_t^\infty \tilde{H}^2([-T,T])} + \|\partial_t u\|_{L_t^\infty \tilde{H}^1([-T,T])} \leq f\big(T, \|(u_0,u_1\|)_{\tilde{H}^2\times\tilde{H}^1}\big). \tag{1-17}$$

Theorem 1.5 shows that a "good" $\dot{H}^{s_p}(\mathbb{R}^3)$ theory for (1-3) can be extended, at least, to one barely $\dot{H}^{s_p}(\mathbb{R}^3)$-supercritical equation, with $\tilde{g}$ going to infinity.

Remark 1.6. Apart from its dependence on p, the function $\tilde{g}$ is universal: it does not depend on an upper bound of the initial data. Moreover, $\tilde{g}$ is unbounded: it goes to infinity with as x.

Remark 1.7. In fact, Theorem 1.5 holds for a weaker version of Condition 1.3: there exists a function C_2 such that for all subinterval $I \subset I_{\max,g}$

$$\sup_{t\in I}\big\|(u(t),\partial_t u(t))\big\|_{\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)} \leq C_2, \tag{1-18}$$

with $C_2 := C_2\big(\|(u_0,u_1)\|_{\tilde{H}^2\times\tilde{H}^1},\ g,\ |I|\big)$. See the proof of Theorem 1.5 and, in particular, (5-21), (5-33) and (5-48).

We recall some basic properties and estimates. If $t_0 \in [t_1,t_2]$, if $F \in L_t^{\tilde{q}} L_x^{\tilde{r}}([t_1,\ t_2])$ and if $(u,\partial_t u) \in C\big([t_1,t_2],\ \dot{H}^m(\mathbb{R}^3)\big)\times C\big([t_1,t_2],\ \dot{H}^{m-1}(\mathbb{R}^3)\big)$ satisfy

$$u(t):\quad \cos(tD)u_0 + \frac{\sin tD}{D}u_1 - \int_{t_0}^t \frac{\sin(t-t')D}{D}F(t')\,dt', \tag{1-19}$$

with data $(u(t_0),\partial_t u(t_0)) \in \dot{H}^m(\mathbb{R}^3)\times\dot{H}^{m-1}(\mathbb{R}^3)$, then we have the Strichartz estimates [Ginibre and Velo 1995; Lindblad and Sogge 1995]

$$\|u\|_{L_t^q L_x^r([t_1,\,t_2])} + \|u\|_{L_t^\infty \dot{H}^m(\mathbb{R}^3)([t_1,\,t_2])} + \|\partial_t u\|_{L_t^\infty \dot{H}^{m-1}(\mathbb{R}^3)([t_1,\,t_2])} \lesssim \|(u(t_0),\ \partial_t u(t_0))\|_{\dot{H}^m(\mathbb{R}^3)\times\dot{H}^{m-1}(\mathbb{R}^3)} + \|F\|_{L_t^{\tilde{q}} L_x^{\tilde{r}}([t_1,\,t_2])}. \tag{1-20}$$

Here (q, r) is m-wave admissible, i.e.,

$$(q, r) \in (2, \infty] \times [2, \infty] \quad \text{and} \quad \frac{1}{q} + \frac{3}{r} = \frac{3}{2} - m; \tag{1-21}$$

moreover,

$$\frac{1}{q} + \frac{3}{r} = \frac{1}{\tilde{q}} + \frac{3}{\tilde{r}} - 2. \tag{1-22}$$

We set some notation that will appear throughout the paper.

We write $A \lesssim B$ if there exists a universal nonnegative constant $C' > 0$ such that $A \leq C'B$. The notation $A = O(B)$ means $A \lesssim B$. More generally, we write $A \lesssim_{a_1,\dots,a_n} B$ if there exists a nonnegative constant $C' = C(a_1, \dots, a_n)$ such that $A \leq C'B$. We say that C'' is the constant determined by $\lesssim$ in $A \lesssim_{a_1,\dots,a_n} B$ if C'' is the smallest possible C' such that $A \leq C'B$. We write $A \ll_{a_1,\dots,a_n} B$ if there exists a universal small nonnegative constant $c = c(a_1, \dots, a_n)$ such that $A \leq cB$. Following [Kenig and Merle 2011], we define, on an interval I,

$$\|u\|_{S(I)} := \|u\|_{L_t^{2(p-1)} L_x^{2(p-1)}(I)}, \quad \|u\|_{W(I)} := \|u\|_{L_t^4 L_x^4(I)}, \quad \|u\|_{\tilde{W}(I)} := \|u\|_{L_t^{\frac{4}{3}} L_x^{\frac{4}{3}}(I)}. \tag{1-23}$$

We also define the quantity

$$Q(I, u) := \|D^{s_p - \frac{1}{2}} u\|_{W(I)} + \|D^{2-\frac{1}{2}} u\|_{W(I)} + \|u\|_{L_t^\infty \tilde{H}^2(I)} + \|\partial_t u\|_{L_t^\infty \tilde{H}^1(I)} \tag{1-24}$$

Let X be a Banach space and $r \geq 0$. Then

$$\boldsymbol{B}(X, r) := \{f \in X : \|f\|_X \leq r\} \tag{1-25}$$

We recall also the well-known Sobolev embeddings. We have

$$\|h\|_{L^\infty(\mathbb{R}^3)} \lesssim \|h\|_{\tilde{H}^2}, \tag{1-26}$$

$$\|h\|_{S(I)} \lesssim \|D^{s_p - \frac{1}{2}} h\|_{L_t^{2(p-1)} L_x^{\frac{6(p-1)}{2p-3}}(I)}. \tag{1-27}$$

We shall combine (1-27) with the Strichartz estimates, since $\left(2(p-1), \frac{6(p-1)}{2p-3}\right)$ is $\frac{1}{2}$- wave admissible.

We also recall some Leibnitz rules [Christ and Weinstein 1991; Kenig et al. 1993]. We have

$$\|D^\alpha F(u)\|_{L_t^q L_x^r(I)} \lesssim \|F'(u)\|_{L_t^{q_1} L_x^{r_1}(I)} \|D^\alpha u\|_{L_t^{q_2} L_x^{r_2}(I)}, \tag{1-28}$$

with $\alpha > 0$, r, r_1, r_2 lying in $[1, \infty]$, $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2}$, and $\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2}$.

The Leibnitz rule for products is

$$\|D^\alpha (uv)\|_{L_t^q L_x^r(I)} \lesssim \|D^\alpha u\|_{L_t^{q_1} L_x^{r_1}(I)} \|v\|_{L_t^{q_2} L_x^{r_2}(I)} + \|D^\alpha u\|_{L_t^{q_3} L_x^{r_3}(I)} \|v\|_{L_t^{q_4} L_x^{r_4}(I)}, \tag{1-29}$$

with $\alpha > 0$, r, r_1, r_2 lying in $[1, \infty]$, $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2}$, $\frac{1}{q} = \frac{1}{q_3} + \frac{1}{q_4}$, $\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2}$, and $\frac{1}{r} = \frac{1}{r_3} + \frac{1}{r_4}$.

If $F \in C^2$, we can write

$$F(x) - F(y) = \int_0^1 F'\big(tx + (1-t)y\big)(x - y)\, dt. \tag{1-30}$$

By using (1-28) and (1-29) the Leibnitz rule for differences can be formulated as

$$\|D^{\alpha}(F(u)-F(v))\|_{L_t^q L_x^r(I)} \lesssim \sup_{t\in[0,1]} \|F'(tu+(1-t)v)\|_{L_t^{q_1} L_x^{r_1}(I)} \|D^{\alpha}(u-v)\|_{L_t^{q_2} L_x^{r_2}(I)} + \sup_{t\in[0,1]} \|F''(tu+(1-t)v)\|_{L_t^{q_1'} L_x^{r_1'}(I)} \big(\|D^{\alpha}u\|_{L_t^{q_2'} L_x^{r_2'}(I)} + \|D^{\alpha}v\|_{L_t^{q_2'} L_x^{r_2'}(I)}\big) \|u-v\|_{L_t^{q_3'} L_x^{r_3'}(I)}, \tag{1-31}$$

with $\alpha>0$, $r_1, r_2, r_1', r_2', r_3'$ lying in $[1,\ \infty]$, $\frac{1}{q}=\frac{1}{q_1}+\frac{1}{q_2}$, $\frac{1}{r}=\frac{1}{r_1}+\frac{1}{r_2}$, $\frac{1}{q}=\frac{1}{q_1'}+\frac{1}{q_2'}+\frac{1}{q_3'}$, and $\frac{1}{r}=\frac{1}{r_1'}+\frac{1}{r_2'}+\frac{1}{r_3'}$.

We shall apply these formulas to several formulas of $F(u)$, and, in particular, to $F(u) := |u|^{p-1}ug(|u|)$. Notice that, by (1-2) and (1-10), we have $F'(x) \sim |x|^{p-1}g(|x|)$ and $F''(x) \sim |x|^{p-2}g(|x|)$. Notice also that, by (1-2) again, we have, for $t \in [0,1]$,

$$g\,(|tx+(1-t)y|) \le g\,(2\max(|x|,|y|)) \le g\big(\max(|x|,|y|)+\log 2\big) \lesssim g(|x|)+g(|y|). \tag{1-32}$$

This will allow us to estimate easily

$$\sup_{t\in[0,1]} \|F'(tu+(1-t)v)\|_{L_t^{q_1} L_x^{r_1}(I)} \quad \text{and} \quad \sup_{t\in[0,1]} \|F''(tu+(1-t)v)\|_{L_t^{q_1} L_x^{r_1}(I)}.$$

Now we explain the main ideas of this paper. We shall prove, in Section 3, that very many values functions g, a special property for the solution of (1-1) holds.

Proposition 1.8 (control of $S(I)$-norm and of norm of initial data imply control of $L_t^{\infty}\tilde{H}^2(I)\times L_t^{\infty}\tilde{H}^1(I)$ norm). *Let I be a compact subinterval of $I_{\max,g}$ (so $\|u\|_{S(I)}<\infty$) and assume that $0\in I$. Assume that g satisfies* (1-2), (1-10) *and*[2]

$$\int_1^{\infty} \frac{1}{yg^2(y)}\,dy = \infty. \tag{1-33}$$

Let $A\ge 0$ such that $\|(u_0,u_1)\|_{\tilde{H}^2\times\tilde{H}^1} \le A$. Let u be the solution of (1-1). *There exists a constant $C>0$ such that*

$$\|(u,\partial_t u)\|_{L_t^{\infty}\tilde{H}^2(I)\times L_t^{\infty}\tilde{H}^1(I)} \le (2C)^N A, \tag{1-34}$$

with $N := N(I)$, such that

$$\int_{2CA}^{(2C)^N A} \frac{1}{yg^2(y)}\,dy \gg \|u\|_{S(I)}^{2(p-1)}. \tag{1-35}$$

Moreover we shall give a criterion of global well-posedness (proved in Section 4):

Proposition 1.9 (criterion of global well-posedness). *Assume that $|I_{\max,g}|<\infty$. Assume that g satisfies* (1-2), (1-10) *and* (1-33). *Then*

$$\|u\|_{S(I_{\max,g})} = \infty. \tag{1-36}$$

The first step is to prove global well-posedness of (1-1), with $g := g_1$ a nondecreasing function that is constant for x large (say $x \ge C_1'$, with C_1' to be determined). By Proposition 1.9, it is enough to find an upper bound of the $S([-T,T])$-norm of the solution $u_{[1]}$ for T arbitrarily large. This can indeed be done, by proving that g_1 can be considered as a subcritical perturbation of the nonlinearity. In other words, $g_1(|u|)|u|^{p-1}u$ will play the same role as that of $|u|^{p-1}u(1-|u|^{-\alpha})$ for some $\alpha>0$. Once we have noticed

[2] Condition (1-33) basically says that g grows slowly on average.

that this comparison is possible, we shall estimate the relevant norms (in particular, $\|u_{[1]}\|_{S([-T,T])}$)) using perturbation theory, Conjecture 1.1 and Condition 1.3, in the spirit of [Zhang 2006]. We expect to find a bound of the form

$$\|u_{[1]}\|_{S([-T,T])} \leq C_3\big(\|(u_0,u_1)\|_{\tilde{H}^2\times\tilde{H}^1},\ T\big), \tag{1-37}$$

with C_3 increasing as T or $\|(u_0,u_1)\|_{\tilde{H}^2\times\tilde{H}^1}$ grows. Notice that if we restrict $[-T,T]$ to the interval $[-1,1]$ and if the $\tilde{H}^2\times\tilde{H}^1$-norm of the initial data (u_0,u_1) is bounded by 1, then we can prove, using (1-37), (1-26) and Proposition 1.8, that the $L_t^\infty L_x^\infty([-T,T])$-norm of the solution $u_{[1]}$ is bounded by a constant (denoted by C_1) on $[-1,1]$. Therefore, if h is a smooth extension of g_1 outside $[0,C_1]$, and if u is the solution of (1-1) (with $g:=h$), we expect to prove that $u=u_{[1]}$ on $[-1,1]$ and for data $\|(u_0,u_1)\|_{\tilde{H}^2\times\tilde{H}^1}\leq 1$. This implies in particular, by (1-37), that we have a finite upper bound $\|u\|_{S([-1,1])}$.

We are not done yet. There are two problems. First, g_1 does not go to infinity. Second, we only control $\|u\|_{S([-1,1])}$ for data $\|(u_0,u_1)\|_{\tilde{H}^2\times\tilde{H}^1}\leq 1$: we would like to control $\|u\|_{S(\mathbb{R})}$ for arbitrary data. In order to overcome these difficulties we iterate the procedure described above. More precisely, given a function g_{i-1} that is constant for $x\geq C_{i-1}$ and such that $u_{[i-1]}$, a solution of (1-1) with $g=g_{i-1}$, satisfies $\|u_{[i-1]}\|_{S([-(i-1),i-1])}<\infty$, we construct a function g_i that

- is an extension of g_{i-1} outside $[0,C_{i-1}]$, and
- is increasing and constant (say equal to $i+1$) for $x\geq C'_i$, with C'_i to be determined.

Again, we shall prove that the g_i may be regarded as a subcritical perturbation of the nonlinearity $(i+1)|u|^{p-1}u$. This allow us to control $\|u_{[i]}\|_{S([-i,i])}$, by using perturbation theory, Conjecture 1.1, and Condition 1.3. Using Proposition 1.8 and (1-26), we can find a finite upper bound for $\|u_{[i]}\|_{L_t^\infty L_x^\infty([-i,i])}$. We assign the value of this upper bound to C_i. To conclude the argument we let $\tilde{g}=\lim_{i\to\infty} g_i$. Given $T>0$, we can find a j such that $[-T,T]\subset[-j,j]$ and $\|(u_0,u_1)\|_{\tilde{H}^2\times\tilde{H}^1}\leq j$. We prove that $u=u_{[j]}$ on $[-j,j]$, where u is a solution of (1-1) with $g:=\tilde{g}$. Since we have a finite upper bound of $\|u_{[j]}\|_{S([-j,j])}$, we also control $\|u\|_{S([-j,j])}$ and $\|u\|_{S([-T,T])}$. Theorem 1.5 follows from Proposition 1.9.

2. Proof of Proposition 1.2

In this section we prove Proposition 1.2 for barely $\dot{H}^{s_p}(\mathbb{R}^3)$-supercritical wave equations (1-1). The proof is based upon standard arguments. Here we have chosen to modify an argument in [Kenig and Merle 2011].

For δ, T_l, C, M to be chosen and such that (1-11) holds we define

$$\begin{aligned} B_1 &:= \boldsymbol{B}\big(\mathscr{C}([0,T_l],\ \tilde{H}^2)\cap D^{\frac12-s_p}W([0,T_l])\cap D^{\frac12-2}W([0,T_l]),\,2CM\big),\\ B_2 &:= \boldsymbol{B}\big(S([0,T_l]),\ 2\delta\big),\\ B' &:= \boldsymbol{B}\big(\mathscr{C}([0,T_l],\ \tilde{H}^1),\,2CM\big), \end{aligned} \tag{2-1}$$

and

$$X := \big\{(u,\partial_t u) : u\in B_1\cap B_2,\ \partial_t u\in B'\big\}. \tag{2-2}$$

Let

$$\Psi(u,\ \partial_t u) := \begin{pmatrix} \cos(tD)u_0 + \dfrac{\sin(tD)}{D}u_1 - \displaystyle\int_0^t \dfrac{\sin(t-t')D}{D}\big(|u(t')|^{p-1}u(t')g(|u(t')|)\big)\,dt' \\ -D\sin(tD)u_0 + \cos(tD)u_1 - \displaystyle\int_0^t \cos(t-t')D\,\big(|u(t')|^{p-1}u(t')g(|u(t')|)\big)\,dt' \end{pmatrix}. \tag{2-3}$$

<u>*Ψ maps X to X.*</u> Indeed, in view of (1-11), (1-20), and the fractional Leibnitz rule (1-28) applied to $\alpha \in \{s_p - \frac{1}{2},\ 2-\frac{1}{2}\}$ and

$$F(u) := |u|^{p-1}ug(|u|)$$

and by applying the multipliers $D^{2-\frac{1}{2}}$ and $D^{s_p-\frac{1}{2}}$ to the Strichartz estimates with $m=\frac{1}{2}$, we have

$$\begin{aligned} &Q([0,\ T_l]) \\ &\quad\lesssim \|(u_0,\ u_1)\|_{\tilde{H}^2(\mathbb{R}^3)\times\tilde{H}^1(\mathbb{R}^3)} + \big\|D^{s_p-\frac{1}{2}}\big(|u|^{p-1}ug(|u|)\big)\big\|_{\tilde{W}([0,T_l])} + \big\|D^{2-\frac{1}{2}}\big(|u|^{p-1}ug(|u|)\big)\big\|_{\tilde{W}([0,T_l])} \\ &\quad\le CM + C\big(\|D^{s_p-\frac{1}{2}}u\|_{W([0,T_l])} + \|D^{2-\frac{1}{2}}u\|_{W([0,T_l])}\big)\|u\|^{p-1}_{S([0,T_l])}g(\|u\|_{L^\infty_t L^\infty_x([0,T_l])}) \\ &\quad\le CM + (2\delta)^{p-1}C(2CM)g(2CM) \end{aligned} \tag{2-4}$$

for some $C>0$ and

$$\begin{aligned} \|u\|_{S([0,T_l])} - \delta &\lesssim \big\|D^{s_p-\frac{1}{2}}\big(|u|^{p-1}ug(|u|)\big)\big\|_{\tilde{W}([0,T_l])} \\ &\lesssim \|u\|^{p-1}_{S([0,T_l])}\|D^{s_p-\frac{1}{2}}u\|_{W([0,T_l])}\, g(\|u\|_{L^\infty_t L^\infty_x([0,T_l])}) \lesssim (2\delta)^{p-1}(2CM)g(2CM). \end{aligned} \tag{2-5}$$

Choosing $\delta=\delta(M)>0$ small enough we see that $\Psi(X)\subset X$.

<u>*Ψ is a contraction.*</u> Indeed we have

$$\begin{aligned} &\|\Psi(u)-\Psi(v)\|_X \\ &\lesssim \big\|D^{s_p-\frac{1}{2}}(|u|^{p-1}ug(|u|)-|v|^{p-1}vg(|v|))\big\|_{\tilde{W}([0,T_l])} + \big\|D^{2-\frac{1}{2}}(|u|^{p-1}ug(|u|)-|v|^{p-1}vg(|v|))\big\|_{\tilde{W}([0,T_l])} \\ &\lesssim \big(g(\|u\|_{L^\infty_t L^\infty_x([0,T_l])}) + g(\|v\|_{L^\infty_t L^\infty_x([0,T_l])})\big) \\ &\quad\times\Big(\big(\|u\|^{p-1}_{S([0,T_l])} + \|v\|^{p-1}_{S([0,T_l])}\big)\big(\|D^{s_p-\frac{1}{2}}(u-v)\|_{W([0,T_l])} + \|D^{2-\frac{1}{2}}(u-v)\|_{W([0,T_l])}\big) \\ &\qquad + \big(\|u\|^{p-2}_{S([0,T_l])} + \|v\|^{p-2}_{S([0,T_l])}\big)\|u-v\|_{S([0,T_l])} \\ &\qquad\quad \times\big(\|D^{s_p-\frac{1}{2}}u\|_{W([0,T_l])} + \|D^{2-\frac{1}{2}}u\|_{W([0,T_l])} + \|D^{s_p-\frac{1}{2}}v\|_{W([0,T_l])} + \|D^{2-\frac{1}{2}}v\|_{W([0,T_l])}\big)\Big) \\ &\lesssim \big(g(2CM)(2\delta)^{p-1} + (2\delta)^{p-2}(2CM)\big)\|u-v\|_X. \end{aligned} \tag{2-6}$$

In these computations, we applied the Leibnitz rule for differences to $\alpha\in\{s_p-\frac{1}{2}, 2-\frac{1}{2}\}$ and

$$F(u) := |u|^{p-1}ug(|u|).$$

Therefore, if $\delta=\delta(M)>0$ is small enough, Ψ is a contraction.

3. Proof of Proposition 1.8

To show Proposition 1.8, it is enough to prove that $Q(I) < \infty$. Without loss of generality we can assume that $A \gg 1$. Then we divide I into subintervals $(I_i)_{1 \le i \le N}$ such that

$$\|u\|_{S(I_i)} = \frac{\eta}{g^{1/(p-1)}((2C)^i A)} \tag{3-1}$$

for some $C \gtrsim 1$ and $\eta > 0$ constants to be chosen later, except maybe the last one. Notice that such a partition always exists since by (1-33) we get, for $N := N(I)$ large enough,

$$\sum_{i=1}^{N} \frac{1}{g^2((2C)^i A)} \ge \int_1^N \frac{1}{g^2((2C)^x A)}\, dx \gtrsim \int_{2CA}^{(2C)^N A} \frac{1}{y g^2(y)}\, dy \gg \|u\|_{S(I)}^{2(p-1)}. \tag{3-2}$$

We get, by a similar reasoning as used in Section 2

$$\begin{aligned} Q(I_1, u) &\lesssim \|(u_0, u_1)\|_{\tilde{H}^2(\mathbb{R}^3) \times \tilde{H}^1(\mathbb{R}^3)} + \left\| D^{s_p - \frac{1}{2}} \left(|u|^{p-1} u g(|u|) \right) \right\|_{\tilde{W}(I_1)} + \left\| D^{2-\frac{1}{2}} \left(|u|^{p-1} u g(|u|) \right) \right\|_{\tilde{W}(I_1)} \\ &\lesssim A + \left(\|D^{s_p - \frac{1}{2}} u\|_{W(I_1)} + \|D^{2-\frac{1}{2}} u\|_{W(I_1)} \right) \|u\|_{S(I_1)}^{p-1} g(\|u\|_{L_t^\infty L_x^\infty(I_1)}) \\ &\lesssim A + \|u\|_{S(I_1)}^{p-1} Q(I_1, u) g(Q(I_1, u)). \end{aligned} \tag{3-3}$$

We choose C to be equal to the constant determined by $\lesssim$ in (3-3). Without loss of generality we can assume that $C > 1$. By a continuity argument, iteration on i, we get, for $\eta \ll 1$, (1-34).

4. Proof of Proposition 1.9

To prove Proposition 1.9, we argue as follows: by time reversal symmetry it is enough to prove that $T_{+,g} < \infty$. If $\|u\|_{S(I_{\max,g})} < \infty$ then we have $Q([0, T_{+,g}], u) < \infty$: this follows by slightly adapting the proof of Proposition 1.8. Consequently, by the dominated convergence theorem, there would exist a sequence $t_n \to T_{+,g}$ such that $\|u\|_{S([t_n, T_{+,g}])} \ll \delta$ and $\|D^{s_p - \frac{1}{2}} u\|_{W([t_n, T_{+,g}])} \ll \delta$ if n is large enough, with δ defined in Proposition 1.2. But, by (1-19) and (1-20),

$$\begin{aligned} &\left\| \cos((t - t_n)D) u(t_n) + \frac{\sin(t - t_n)D}{D} u_1 \right\|_{S([t_n, T_{+,g}])} \\ &\qquad \lesssim \|u\|_{S([t_n, T_{+,g}])} + \|u\|_{S([t_n, T_{+,g}])}^{p-1} \|D^{s_p - \frac{1}{2}} u\|_{W([t_n, T_{+,g}])} g(Q([0, T_{+,g}, u])) \ll \delta, \end{aligned} \tag{4-1}$$

and consequently, by continuity, there would exist $\tilde{T} > T_{+,g}$ such that

$$\left\| \cos((t - t_n)D) u(t_n) + \frac{\sin(t - t_n)D}{D} \partial_t u(t_n) \right\|_{S([t_n, \tilde{T}])} \le \delta, \tag{4-2}$$

which would contradict the definition of $T_{+,g}$.

Remark 4.1. Notice that if we have the stronger bound $\|u\|_{S(I_{\max,g})} \le C$ with $C := C\left(\|(u_0, u_1)\|_{\tilde{H}^2 \times \tilde{H}^1} \right) < \infty$, then not only $I_{\max,g} = (-\infty, +\infty)$ but also u scatters as $t \to \pm\infty$. Indeed, by Proposition 1.9, $I_{\max,g} = \mathbb{R}$. Then by time reversal symmetry it is enough to assume that $t \to \infty$. Let $v(t) := (u(t), \partial_t u(t))$.

We are looking for $v_{+,0} := \big(u_{+,0},\ u_{+,1}\big)$ such that

$$\big\| v(t) - K(t) v_{+,0} \big\|_{\tilde{H}^2 \times \tilde{H}^1} \to 0 \tag{4-3}$$

as $t \to \infty$. Here

$$K(t) := \begin{pmatrix} \cos tD & (\sin tD)/D \\ -D \sin tD & \cos tD \end{pmatrix} \tag{4-4}$$

We have

$$K^{-1}(t) = \begin{pmatrix} \cos tD & -(\sin tD)/D \\ D \sin tD & \cos tD \end{pmatrix}. \tag{4-5}$$

Notice that $K^{-1}(t)$ and $K(t)$ are bounded in $\tilde{H}^2 \times \tilde{H}^1$. Therefore it is enough to prove that $K^{-1}(t)v(t)$ has a limit as $t \to \infty$. But since $K^{-1}(t)v(t) = (u_0, u_1) - K^{-1}(t)\,(u_{\mathrm{nl}}(t),\ \partial_t u_{\mathrm{nl}}(t))$ — where

$$u_{\mathrm{nl}}(t) := -\int_0^t \frac{\sin\,(t-t')D}{D} \big(|u(t')|^{p-1} u(t') g(|u(t')|)\big)\, dt'$$

denotes the nonlinear part of the solution (1-12) — it suffices to prove that $K^{-1}(t)\,(u_{\mathrm{nl}}(t), \partial_t u_{\mathrm{nl}}(t))$ has a limit. But

$$\begin{aligned} &\big\| K^{-1}(t_1) u_{\mathrm{nl}}(t_1) - K^{-1}(t_2) u_{\mathrm{nl}}(t_2) \big\|_{\tilde{H}^2 \times \tilde{H}^1} \\ &\qquad \lesssim \big\| (u_{\mathrm{nl}}, \partial_t u_{\mathrm{nl}}) \big\|_{L_t^\infty \tilde{H}^2([t_1,t_2]) \times L_t^\infty \tilde{H}^1([t_1,t_2])} \\ &\qquad \lesssim \Big(\big\| D^{s_p - \frac{1}{2}} (|u|^{p-1} u g(|u|)) \big\|_{\tilde{W}([t_1,t_2])} + \big\| D^{2-\frac{1}{2}} (|u|^{p-1} u g(|u|)) \big\|_{\tilde{W}([t_1,t_2])} \Big) \qquad (4\text{-}6) \\ &\qquad \lesssim \Big(\big\| D^{s_p - \frac{1}{2}} u \big\|_{W([t_1,t_2])} + \big\| D^{2-\frac{1}{2}} u \big\|_{W([t_1,t_2])} \Big) \|u\|^{p-1}_{S([t_1,t_2])} g\big(\|u\|_{L_t^\infty L_x^\infty(\mathbb{R})} \big). \end{aligned}$$

It remains to prove that $Q(\mathbb{R}) < \infty$ in order to conclude that the Cauchy criterion is satisfied, which would imply scattering. This follows from $\|u\|_{S(\mathbb{R})} < \infty$ and a slight modification of the proof of Proposition 1.8.

5. Construction of the function *g*

In this section we prove Theorem 1.5. Let

$$\mathrm{Up}(i) := \big\{ (T,\ (u_0, u_1)) : 0 \le T \le i,\ \|(u_0, u_1)\|_{\tilde{H}^2 \times \tilde{H}^1} \le i \big\} \tag{5-1}$$

As i ranges over $\{1, 2, \dots\}$ we construct, for each set $\mathrm{Up}(i)$, a function g_i satisfying (1-2) and (1-10). Moreover it is constant for large values of $|x|$. The function g_{i+1} depends on g_i; the construction of g_i is made by induction on i. More precisely:

Lemma 5.1. *Let $A \gg 1$. There exist two sequences of numbers $\{C_i\}_{i \ge 0}$, $\{C_i'\}_{i \ge 0}$ and a sequence of functions $\{g_i\}_{i \ge 0}$ such that, for all $(T,\ (u_0, u_1)) \in \mathrm{Up}(i)$, we have*

- $g_0 := 1$, $C_0 := 0$, $C_0' = 0$;
- $\{C_i\}_{i \ge 0}$ *and* $\{C_i'\}_{i \ge 0}$ *are positive, nondecreasing, and satisfy*

$$A C_{i-1} < C_i' < A C_i \tag{5-2}$$

for $i \geq 1$ *and*

$$C_i \geq i; \tag{5-3}$$

- g_i *is smooth, nondecreasing, and satisfies* (1-2), (1-10),

$$\int_1^{C_i'} \frac{1}{t g_i^2(t)}\, dt \to \infty \quad \text{as } i \to \infty, \tag{5-4}$$

and

$$g_i(|x|) = \begin{cases} g_{i-1}(|x|) & \text{if } |x| \leq AC_{i-1}, \\ i+1 & \text{if } |x| \geq C_i'; \end{cases} \tag{5-5}$$

- *the solution* $u_{[i]}$ *of the wave equation*

$$\begin{cases} \partial_{tt} u_{[i]} - \triangle u_{[i]} = -|u_{[i]}|^{p-1} u_{[i]} g_i(|u_{[i]}|), \\ u_{[i]}(0) = u_0 \in \tilde{H}^2, \\ \partial_t u_{[i]}(0) = u_1 \in \tilde{H}^1 \end{cases} \tag{5-6}$$

satisfies

$$\max\left(\|u_{[i]}\|_{S([-i,i])}, \left\|(u_{[i]}, \partial_t u_{[i]})\right\|_{L_t^\infty \tilde{H}^2([-T,T]) \times L_t^\infty \tilde{H}^1([-T,T])} \right) \leq C_i. \tag{5-7}$$

We postpone the proof until page 212. Assume the lemma is true and let $\tilde{g} = \lim_{i \to \infty} g_i$. Clearly $\tilde{g}$ is smooth; it satisfies (1-2) and (1-10). It also goes to infinity. Moreover let u be the solution of (1-1) with $g := \tilde{g}$. We want to prove that the solution u exists for all time. Let $T_0 \geq 0$ be a fixed time. Let $j := j(T_0, \|u_0\|_{\tilde{H}^2}, \|u_1\|_{\tilde{H}^1}) > 0$ be the smallest positive integer such that $[-T_0, T_0] \subset [-j, j]$ and $\|(u_0, u_1)\|_{\tilde{H}^2 \times \tilde{H}^1} \leq j$. We claim that

$$\|(u, \partial_t u)\|_{L_t^\infty \tilde{H}^2([-T_0,T_0]) \times L_t^\infty \tilde{H}^1([-T_0,T_0])} \leq C_j \quad \text{and} \quad \|u\|_{S([-T_0,T_0])} \leq C_j.$$

Indeed, let

$$F_j := \left\{ t \in [0, j] : \|(u, \partial_t u)\|_{L_t^\infty \tilde{H}^2([-t,t]) \times L_t^\infty \tilde{H}^1([-t,t])} \leq C_j \text{ and } \|u\|_{S([-t,t])} \leq C_j \right\}. \tag{5-8}$$

We must show that F_j coincides with $[0, j]$. Certainly F_j is nonempty, since it contains 0; see (5-3).

<u>F_j *is closed.*</u> Indeed, let $\tilde{t} \in \bar{F}_j$. There exists a sequence $(t_n)_{n \geq 1}$ in $[0, j]$ such that $t_n \to \tilde{t}$, $\|u\|_{S([-t_n,t_n])} \leq C_j$, and $\|(u, \partial_t u)\|_{L_t^\infty \tilde{H}^2([-t_n,t_n]) \times L_t^\infty \tilde{H}^1([-t_n,t_n])} \leq C_j$. It is enough to prove that $\|u\|_{S([-\tilde{t},\tilde{t}])}$ is finite and then apply dominated convergence. There are two cases:

- If $\operatorname{card}\{t_n : t_n \leq \tilde{t}\} < \infty$, there exists n_0 large enough such that $t_n \geq \tilde{t}$ for $n \geq n_0$ and

$$\|u\|_{S([-\tilde{t},\tilde{t}])} \leq \|u\|_{S([-t_n,t_n])} < \infty. \tag{5-9}$$

- If $\operatorname{card}\{t_n : t_n \leq \tilde{t}\} = \infty$, we can assume by passing to a subsequence that $t_n \leq \tilde{t}$. Let $n_0 \geq 1$ be fixed. Since

$$\left\| \cos(t - t_{n_0}) D u(t_{n_0}) + \frac{\sin(t - t_{n_0}) D}{D} \partial_t u(t_{n_0}) \right\|_{S([t_{n_0}, \tilde{t}])} \lesssim \left\| (u(t_{n_0}), \partial_t u(t_{n_0})) \right\|_{\tilde{H}^2 \times \tilde{H}^1} \lesssim C_j, \tag{5-10}$$

we conclude from the dominated convergence theorem that there is $n_1 := n_1(n_0)$ large enough that

$$\| \cos(t - t_{n_0}) D u(t_{n_0}) + \frac{\sin(t - t_{n_0}) D}{D} \partial_t u(t_{n_0}) \|_{S([t_{n_1}, \tilde{t}])} \leq \delta, \tag{5-11}$$

with $\delta := \delta(C_j)$ defined in Proposition 1.2. Therefore, by Proposition 1.2, we have $\|u\|_{S([t_{n_1},\tilde{t}])} < \infty$. Similarly, $\|u\|_{S([-\tilde{t},-t_{n_1}])} < \infty$. Combining these inequalities with $\|u\|_{S([-t_{n_1},t_{n_1}])} \le C_j$, we eventually get $\|u\|_{S([-\tilde{t},\tilde{t}])} < \infty$, as desired.

F_j is open. Indeed, let $\bar{t} \in F_j$. By Proposition 1.2 there exists $\alpha > 0$ such that if $t \in (\bar{t}-\alpha, \bar{t}+\alpha) \cap [0, j]$ then $[-t, t] \subset I_{\max,\tilde{g}}$ and $\|u\|_{L_t^\infty L_x^\infty([-t,t])} \lesssim \|u\|_{L_t^\infty \tilde{H}^2([-t,t])} \lesssim C_j$. Also, by (5-7), $[-t, t] \subset I_{\max,g_j}$. In view of these remarks, we conclude, after slightly adapting the proof of Proposition 1.8, that $Q([-t, t], u) \lesssim_j 1$ and $Q([-t, t], u_{[j]}) \lesssim_j 1$. We divide $[-t, t]$ into a finite number of subintervals $(I_i)_{i\le k} = ([a_i, b_i])_{1\le i\le k}$ that satisfy, for $\eta \ll 1$ to be defined later, the following properties:

(1) $1 \le i \le k$: $\|u_{[j]}\|_{S(I_i)} \le \eta$, $\|u\|_{S(I_i)} \le \eta$, $\|D^{s_p-\frac{1}{2}} u_{[j]}\|_{W(I_i)} \le \eta$, $\|D^{s_p-\frac{1}{2}} u\|_{W(I_i)} \le \eta$, $\|D^{2-\frac{1}{2}} u\|_{W(I_i)} \le \eta$, and $\|D^{2-\frac{1}{2}} u_{[j]}\|_{W(I_i)} \le \eta$.

(2) $1 \le i < k$: $\|u_{[j]}\|_{S(I_i)} = \eta$ or $\|u\|_{S(I_i)} = \eta$ or $\|D^{s_p-\frac{1}{2}} u_{[j]}\|_{W(I_i)} = \eta$ or $\|D^{s_p-\frac{1}{2}} u\|_{W(I_i)} = \eta$ or $\|D^{2-\frac{1}{2}} u\|_{W(I_i)} = \eta$, or $\|D^{2-\frac{1}{2}} u_{[j]}\|_{W(I_i)} = \eta$.

Notice that, by (1-2), we have

$$\|g_j(|u|) - g_j(|u_{[j]}|)\|_{L_t^\infty L_x^\infty(I_i)} \lesssim \|u - u_{[j]}\|_{L_t^\infty L_x^\infty(I_i)} \lesssim \|u - u_{[j]}\|_{L_t^\infty \tilde{H}^2(I_i)}. \tag{5-12}$$

Consider $w = u - u_{[j]}$. Applying the Leibnitz rules (1-28), (1-31), and (1-29), together with (5-12), we have

$$\begin{aligned}
&Q(I_1, w)\\
&\lesssim \big\|D^{s_p-\frac{1}{2}}(|u|^{p-1}u(\tilde{g}-g_j)(|u|))\big\|_{\tilde{W}(I_1)} + \big\|D^{2-\frac{1}{2}}(|u|^{p-1}u(\tilde{g}-g_j)(|u|))\big\|_{\tilde{W}(I_1)}\\
&\quad + \big\|D^{s_p-\frac{1}{2}}(|u|^{p-1}u - |u_{[j]}|^{p-1}u_{[j]})g_j(|u|)\big\|_{\tilde{W}(I_1)} + \big\|D^{2-\frac{1}{2}}(|u|^{p-1}u - |u_{[j]}|^{p-1}u_{[j]})g_j(|u|)\big\|_{\tilde{W}(I_1)}\\
&\quad + \big\|D^{s_p-\frac{1}{2}}(|u_{[j]}|^{p-1}u_{[j]}(g_j(|u|) - g_j(|u_{[j]}|)))\big\|_{\tilde{W}(I_1)} + \big\|D^{2-\frac{1}{2}}(|u_{[j]}|^{p-1}u_{[j]}(g_j(|u|) - g_j(|u_{[j]}|)))\big\|_{\tilde{W}(I_1)}\\
&\lesssim (\tilde{g}-g_j)(\|u\|_{L_t^\infty \tilde{H}^2(I_1)})\big(\|D^{s_p-\frac{1}{2}}u\|_{W(I_1)} + \|D^{2-\frac{1}{2}}u\|_{W(I_1)}\big)\|u\|_{S(I_1)}^{p-1}\\
&\quad + g_j\big(\|u\|_{L_t^\infty \tilde{H}^2(I_1)}\big)\Big(\big(\|u_{[j]}\|_{S(I_1)}^{p-1} + \|u\|_{S(I_1)}^{p-1}\big)\big(\|D^{s_p-\frac{1}{2}}w\|_{W(I_1)} + \|D^{2-\frac{1}{2}}w\|_{W(I_1)}\big)\\
&\qquad\qquad + \big(\|u_{[j]}\|_{S(I_1)}^{p-2} + \|u\|_{S(I_1)}^{p-2}\big)\,\|w\|_{S(I_1)}\\
&\qquad\qquad \times\big(\|D^{s_p-\frac{1}{2}}u\|_{W(I_1)} + \|D^{2-\frac{1}{2}}u\|_{W(I_1)} + \|D^{s_p-\frac{1}{2}}u_{[j]}\|_{W(I_1)} + \|D^{2-\frac{1}{2}}u_{[j]}\|_{W(I_1)}\big)\Big)\\
&\quad + \|g_j'(|u|)\|_{L_t^\infty L_x^\infty(I_1)}\big(\|D^{s_p-\frac{1}{2}}u\|_{W(I_1)} + \|D^{2-\frac{1}{2}}u\|_{W(I_1)}\big)\big(\|u\|_{S(I_1)}^{p-2} + \|u_{[j]}\|_{S(I_1)}^{p-2}\big)\|w\|_{S(I_1)}\\
&\quad + \big\|g_j(|u|) - g_j(|u_{[j]}|)\big\|_{L_t^\infty L_x^\infty(I_1)}\,\|D^{s_p-\frac{1}{2}}u_{[j]}\|_{W(I_1)}\,\|u_{[j]}\|_{S(I_1)}^{p-1}\\
&\quad + \|u_{[j]}\|_{S(I_1)}^{p-1}\,\|u_{[j]}\|_{L_t^\infty \tilde{H}^2(I_1)}\big(\|w\|_{L_t^\infty \tilde{H}^2(I_1)}\big(\|D^{s_p-\frac{1}{2}}u\|_{W(I_1)} + \|D^{s_p-\frac{1}{2}}u_{[j]}\|_{W(I_1)}\big) + \|D^{s_p-\frac{1}{2}}w\|_{W(I_1)}\big)\\
&\lesssim g_j(C_j)\eta^{p-1}Q(I_1, w) + \eta^{p-1}Q(I_1, w) + \eta^p Q(I_1, w) + C_j\eta^{p-1}\big(\eta Q(I_1, w) + Q(I_1, w)\big),
\end{aligned} \tag{5-13}$$

since, by choosing A large enough and by the construction of $\tilde{g}$, we have

$$(\tilde{g} - g_j)(\|u\|_{L_t^\infty \tilde{H}^2(I_1)}) = 0. \tag{5-14}$$

We conclude via a continuity argument that $Q(I_1, w) = 0$, so $u = u_{[j]}$ on I_1. In particular, $u(b_1) = u_{[j]}(b_1)$. By iteration on i, it is not difficult to see that $u = u_{[j]}$ on $[-t, t]$. Hence $(\bar{t} - \alpha, \bar{t} + \alpha) \cap [0, j] \subset F_j$, by (5-7). Thus F_j is open.

The upshot is that $F_j = [0, j]$, so $\|u\|_{S([-T_0,T_0])} \leq C_j$. This proves global well-posedness. Moreover, since j depends on T_0 and $\|(u_0, u_1)\|_{\tilde{H}^2 \times \tilde{H}^1}$, we get (1-17).

Proof of Lemma 5.1. The proof extends to the end of the paper. We must establish *a priori* bounds.

Step 1: Construction of g_1.

Basically, g_1 is a nonnegative function that increases and is equal to 2 for x large. Recall that $[-T, T] \subset [-1, 1]$ and $\|(u_0, u_1)\|_{\tilde{H}^2 \times \tilde{H}^1} \leq 1$. Let $I \subset [-T, T]$.

Observe that the point $(\infty-, 3+) := \left(\frac{3+\epsilon}{\epsilon}, 3 + \epsilon\right)$ with $\epsilon \ll 1$ is $\frac{1}{2}$-wave admissible.

We would like to chop I (satisfying $\|\cdot\|_{L_t^\infty L_x^3(I)} < \infty$) into subintervals I_j such that $\|\cdot\|_{L_t^\infty L_x^3(I_j)}$ is as small as wanted. Unfortunately this is impossible because the L_t^∞-norm is pathological. Instead we will apply this process to $\|\cdot\|_{L_t^{\infty-} L_x^{3+}}$. This creates slight variations almost everywhere in the process of the construction of g_i. Details with respect to these slight perturbations have been omitted for the sake of readability: they are left to the reader, who should ignore the $+$ and $-$ signs at the first reading.

We define

$$X(I) := D^{\frac{1}{2}-s_p} L_t^{\infty-} L_x^{3+}(I) \cap D^{\frac{1}{2}-s_p} W(I) \cap S(I) \cap L_t^\infty \dot{H}^{s_p}(I) \times L_t^\infty \dot{H}^{s_p-1}(I). \tag{5-15}$$

Let g_1 be a smooth function, defined on the set of nonnegative real numbers, nondecreasing, and such that $h_1 := g_1 - 2$ satisfies the following properties: $h_1(0) = -1$, h is nondecreasing, and $h_1(x) = 0$ if $|x| \geq 1$. It is not difficult to see that (1-2) and (1-10) are satisfied.

Observe that

$$|h_1(x)| \lesssim \frac{1}{|x|^{\frac{p-1}{2}-}} \tag{5-16}$$

and

$$|h_1'(x)| \lesssim \frac{1}{|x|^{\frac{p+1}{2}-}}. \tag{5-17}$$

Let $u_{[1]}$ and $v_{[1]}$ be solutions to the equations

$$\begin{cases} \partial_{tt} u_{[1]} - \triangle u_{[1]} = -|u_{[1]}|^{p-1} u_{[1]} g_1(|u_{[1]}|), \\ u_{[1]}(0) = u_0 \in \tilde{H}^2, \\ \partial_t u_{[1]}(0) = u_1 \in \tilde{H}^1 \end{cases} \tag{5-18}$$

and

$$\begin{cases} \partial_{tt} v_{[1]} - \triangle v_{[1]} = -2 \left|v_{[1]}\right|^{p-1} v_{[1]}, \\ v_{[1]}(0) = u_0, \\ \partial_t v_{[1]}(0) = u_1. \end{cases} \tag{5-19}$$

Step 1a. We claim that $\|v_{[1]}\|_{X(\mathbb{R})} < \infty$. Indeed, since we assumed that Conjecture 1.1 is true, we can divide $\mathbb{R}$ into subintervals $(I_j = [t_j, t_{j+1}])_{1 \leq j \leq l}$ such that

$$\|v_{[1]}\|_{S(I_j)} = \eta \quad \text{and} \quad \|v_{[1]}\|_{S(I_l)} \leq \eta,$$

with $\eta \ll 1$. Then

$$\begin{aligned}\|v_{[1]}\|_{X(I_{j+1})} &\lesssim \big\|(v_{[1]}(t_j),\partial_t v_{[1]}(t_j))\big\|_{\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)} + \big\|D^{s_p-\frac{1}{2}}(|v_{[1]}|^{p-1}v_{[1]})\big\|_{\tilde{W}(I_{j+1})}\\ &\lesssim \big\|(v_{[1]}(t_j),\partial_t v_{[1]}(t_j))\big\|_{\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)} + \big\|D^{s_p-\frac{1}{2}}v_{[1]}\big\|_{W(I_{j+1})}\|v_{[1]}\|^{p-1}_{S(I_{j+1})}\\ &\lesssim \|v_{[1]}\|_{X(I_j)} + \eta^{p-1}\|v_{[1]}\|_{X(I_{j+1})}.\end{aligned} \tag{5-20}$$

Notice that $l \lesssim 1$: this follows from Conjecture 1.1, Condition 1.3 and the inequality

$$\begin{aligned}\|(u_0,u_1)\|_{\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)} \le \sup_{t\in I_{\max,g_1}}\|(u(t),\partial_t u(t))\|_{\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)} \le C_2\big(\|(u_0,u_1)\|_{\tilde{H}^2\times\tilde{H}^1}\big)\\ \lesssim 1,\end{aligned} \tag{5-21}$$

following from Condition 1.3 and the assumption $\|(u_0,u_1)\|_{\tilde{H}^2\times\tilde{H}^1} \le 1$. (At this stage, we only need to know that $\|(u_0,u_1)\|_{\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)} \le \|(u_0,u_1)\|_{\tilde{H}^2\times\tilde{H}^1} \le 1$ and apply Conjecture 1.1. Therefore the introduction of $\sup_{t\in I_{\max,g_1}}\|(u(t),\partial_t u(t))\|_{\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)}$ in (5-21) is redundant. This is done on purpose. Indeed, we will use Condition 1.3 in other parts of the argument: see (5-33).)

Now by a standard continuity argument and iteration on j we have

$$\|v_{[1]}\|_{X(\mathbb{R})} \lesssim 1 \tag{5-22}$$

Step 1b. We control $\|u_{[1]}-v_{[1]}\|_{X([-\tilde{t},\tilde{t}])}$, for $\tilde{t} \ll 1$ to be chosen later. By time reversal symmetry it is enough to control $\|u_{[1]}-v_{[1]}\|_{X([0,\tilde{t}])}$. To this end we consider $w_{[1]} := u_{[1]}-v_{[1]}$. We get

$$\partial_{tt}w_{[1]} - \triangle w_{[1]} = -|w_{[1]}+v_{[1]}|^{p-1}(v_{[1]}+w_{[1]})g_1(v_{[1]}+w_{[1]}) + 2|v_{[1]}|^{p-1}v_{[1]}.$$

Let $\eta' \ll 1$. By (5-22), we can divide $[0,\tilde{t}]$ into subintervals $(J_k=[t'_k,t'_{k+1}])_{1\le k\le m}$ that satisfy

$$\|D^{s_p-\frac{1}{2}}v_{[1]}\|_{L_t^{\infty-}L_x^{3+}(J_k)} = \eta' \;\text{ or }\; \|D^{s_p-\frac{1}{2}}v_{[1]}\|_{W(J_k)} = \eta' \qquad \text{for } 1\le k<m, \tag{5-23}$$

$$\|D^{s_p-\frac{1}{2}}v_{[1]}\|_{W(J_k)} \le \eta' \text{ and } \|D^{s_p-\frac{1}{2}}v_{[1]}\|_{L_t^{\infty-}L_x^{3+}(J_k)} \le \eta' \qquad \text{for } 1\le k\le m. \tag{5-24}$$

We have

$$\|w_{[1]}\|_{X(J_{k+1})} \lesssim \big\|\big(w_{[1]}(t'_k),\partial_t w_{[1]}(t'_k)\big)\big\|_{\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)} + A_1 + A_2,$$

where

$$\begin{aligned}A_1 &:= \|D^{s_p-\frac{1}{2}}\big(2|v_{[1]}|^{p-1}v_{[1]} - 2|v_{[1]}+w_{[1]}|^{p-1}(v_{[1]}+w_{[1]})\big)\|_{\tilde{W}(J_{k+1})},\\ A_2 &:= \|D^{s_p-\frac{1}{2}}\big(h_1(|v_{[1]}+w_{[1]}|)|v_{[1]}+w_{[1]}|^{p-1}(v_{[1]}+w_{[1]})\big)\|_{L_t^1L_x^{\frac{3}{2}}(J_{k+1})}.\end{aligned} \tag{5-25}$$

By the fractional Leibnitz rule applied to $q(x) := |x|^{p-1} x h(x)$, (5-16), (5-17), Sobolev embedding and Hölder in time we have

$$\begin{aligned} A_2 &\lesssim \left\| |v_{[1]} + w_{[1]}|^{\frac{p-1}{2}+} \right\|_{L_t^{1+} L_x^{3-}(J_{k+1})} \left\| D^{s_p - \frac{1}{2}} (v_{[1]} + w_{[1]}) \right\|_{L_t^{\infty-} L_x^{3+}(J_{k+1})} \\ &\lesssim \| v_{[1]} + w_{[1]} \|^{\frac{p-1}{2}+}_{L_t^{\frac{p-1}{2}+} L_x^{\frac{3(p-1)+}{2}} (J_{k+1})} \left\| D^{s_p - \frac{1}{2}} (v_{[1]} + w_{[1]}) \right\|_{L_t^{\infty-} L_x^{3+}(J_{k+1})} \\ &\lesssim \tilde{t} \left\| D^{s_p - \frac{1}{2}} (v_{[1]} + w_{[1]}) \right\|^{\frac{p+1}{2}+}_{L_t^{\infty-} L_x^{3+}(J_{k+1})} \lesssim \tilde{t} (\eta')^{\frac{p+1}{2}+} + \tilde{t} \left\| D^{s_p - \frac{1}{2}} w_{[1]} \right\|^{\frac{p+1}{2}+}_{L_t^{\infty-} L_x^{3+}(J_{k+1})} \\ &\lesssim \tilde{t} (\eta')^{\frac{p+1}{2}+} + \tilde{t} \| w_{[1]} \|^{\frac{p+1}{2}+}_{X(J_{k+1})}. \end{aligned} \tag{5-26}$$

For A_1 we follow [Kenig and Merle 2011, p. 9]:

$$\begin{aligned} A_1 &\lesssim \left(\| v_{[1]} \|^{p-1}_{S(J_{k+1})} + \| w_{[1]} \|^{p-1}_{S(J_{k+1})} \right) \| D^{s_p - \frac{1}{2}} w_{[1]} \|_{W(J_{k+1})} \\ &\qquad + \left(\| v_{[1]} \|^{p-2}_{S(J_{k+1})} + \| w_{[1]} \|^{p-2}_{S(J_{k+1})} \right) \left(\left\| D^{s_p - \frac{1}{2}} v_{[1]} \right\|_{W(J_{k+1})} + \left\| D^{s_p - \frac{1}{2}} w_{[1]} \right\|_{W(J_{k+1})} \right) \| w_{[1]} \|_{S(J_{k+1})} \\ &\lesssim (\eta')^{p-1} \| w_{[1]} \|_{X(J_{k+1})} + \| w_{[1]} \|^{p}_{X(J_{k+1})} + (\eta')^{p-2} \| w_{[1]} \|^2_{X(J_{k+1})} + \eta' \| w_{[1]} \|^{p-1}_{X(J_{k+1})}. \end{aligned} \tag{5-27}$$

This follows from (1-31) and (1-27). Therefore we have

$$\begin{aligned} \| w_{[1]} \|_{X(J_{k+1})} &\lesssim \| w_{[1]} \|_{X(J_k)} + (\eta')^{\frac{p+1}{2}+} \tilde{t} + \tilde{t} \| w_{[1]} \|^{\frac{p+1}{2}+}_{X(J_{k+1})} \\ &\quad + (\eta')^{p-1} \| w_{[1]} \|_{X(J_{k+1})} + \| w_{[1]} \|^{p}_{X(J_{k+1})} + (\eta')^{p-2} \| w_{[1]} \|^2_{Y(J_{k+1})} + \eta' \| w_{[1]} \|^{p-1}_{X(J_{k+1})}. \end{aligned} \tag{5-28}$$

Let C be the constant determined by (5-28). By induction, we have

$$\| w_{[1]} \|_{X(J_k)} \leq (2C)^k \tilde{t}, \tag{5-29}$$

provided that for $1 \leq k \leq m-1$ we have

$$\begin{aligned} C (\eta')^{\frac{p+1}{2}} \tilde{t} &\ll C (2C)^k \tilde{t}, \\ C \tilde{t} \left((2C)^k \tilde{t} \right)^{\frac{p+1}{2}+} &\ll (2C)^k \tilde{t}, \\ C (\eta')^{p-1} (2C)^{k+1} \tilde{t} &\ll C (2C)^k \tilde{t}, \\ C \left((2C)^k \tilde{t} \right)^p &\ll C (2C)^k \tilde{t}, \\ C (\eta')^{p-2} \left((2C)^{k+1} \tilde{t} \right)^2 &\ll C (2C)^k \tilde{t}, \\ \eta' \left((2C)^{k+1} \tilde{t} \right)^{p-1} &\ll C (2C)^k. \end{aligned} \tag{5-30}$$

These inequalities are satisfied if $\eta' \ll 1$ and

$$\tilde{t} \ll 1 \tag{5-31}$$

since $k \leq m-1$ and, by (5-22), $m \lesssim 1$. We conclude that

$$\| w_{[1]} \|_{X([0, \tilde{t}])} \lesssim 1. \tag{5-32}$$

Step 1c. We control $\|u_{[1]}\|_{X([-T,T])}$. By time reversal symmetry, it is enough to control $\|u_{[1]}\|_{X([0,T])}$. Recall that $T \leq 1$. We chop $T \leq 1$ into subintervals $(J_{k'} = [a_{k'}, b_{k'}])_{1 \leq k' \leq l'}$ such that $|J_{k'}| = \tilde{t}$ for $1 \leq k' < l'$ and $|J_{l'}| \leq \tilde{t}$. Notice that, by Condition 1.3, we have

$$\begin{aligned} \|(u(a_{k'}), \partial_t u(a_{k'}))\|_{\dot{H}^{s_p}(\mathbb{R}^3) \times \dot{H}^{s_p-1}(\mathbb{R}^3)} &\leq \sup_{t \in I_{\max, g_1}} \|(u(t), \partial_t u(t))\|_{\dot{H}^{s_p}(\mathbb{R}^3) \times \dot{H}^{s_p-1}(\mathbb{R}^3)} \\ &\leq C_2 \left(\|(u_0, u_1)\|_{\tilde{H}^2 \times \tilde{H}^1} \right) \lesssim 1, \end{aligned} \tag{5-33}$$

taking advantage of the assumption $\|(u_0, u_1)\|_{\tilde{H}^2 \times \tilde{H}^1} \leq 1$. For each k' let $v_{[1,k']}$ be the solution of

$$\begin{cases} \partial_{tt} v_{[1,k']} - \triangle v_{[1,k']} = -|v_{[1,k']}|^{p-1} v_{[1,k']}, \\ v_{[1,k']}(a_{k'}) = u_{[1]}(a_{k'}), \\ \partial_t v_{[1,k']}(a_{k'}) = \partial_t u_{[1]}(a_{k'}); \end{cases} \tag{5-34}$$

in particular, $v_{[1,k']} = v_{[1]}$. By slightly modifying the proof of Step 1b and letting $v_{[1,k']}$ play the role of $v_{[1]}$, this leads, by (5-33), to

$$\|v_{[1,k']}\|_{X(\mathbb{R})} \lesssim 1 \tag{5-35}$$

and

$$\|w_{[1,k']}\|_{X(J_{k'})} \lesssim 1, \tag{5-36}$$

with $w_{[1,k']} := u_{[1]} - v_{[1,k']}$. Therefore $\|u_{[1]}\|_{X(J_{k'})} \lesssim 1$, and summing over $J_{k'}$ we have

$$\|u_{[1]}\|_{X([0,T])} \lesssim 1. \tag{5-37}$$

Step 1d. We control $\|(u_{[1]}, \partial_t u_{[1]})\|_{L_t^\infty \tilde{H}^2([-1,1]) \times L_t^\infty \tilde{H}^1([-1,1])}$ and $\|u_{[1]}\|_{S([-1,1])}$. We get from (5-37)

$$\|u_{[1]}\|_{S([-1,1])} \lesssim 1. \tag{5-38}$$

To conclude Step 1: By Proposition 1.8 and (5-38) we have

$$\|(u_{[1]}, \partial_t u_{[1]})\|_{L_t^\infty \tilde{H}^2([-1,1]) \times L_t^\infty \tilde{H}^1([-1,1])} \lesssim 1. \tag{5-39}$$

Therefore

$$\max \left(\|u_{[1]}\|_{S([-1,1])},\ \|(u_{[1]}, \partial_t u_{[1]})\|_{L_t^\infty \tilde{H}^2([-1,1]) \times L_t^\infty \tilde{H}^1([-1,1])} \right) \lesssim 1. \tag{5-40}$$

We let C_1' in the statement of Lemma 5.1 be equal to 1. We can assume without the loss of generality that the constant implicit in $\lesssim$ in (5-40) is larger than 1; let C_1 in the statement of Lemma 5.1 be this constant. Then C_1' and C_1 satisfy (5-2) and (5-3).

Step 2: Construction of g_i from g_{i-1}.

Recall that $[-T, T] \subset [-i, i]$ and $\|(u_0, u_1)\|_{\tilde{H}^2 \times \tilde{H}^1} \leq i$. In view of (5-5) it is enough to construct g_i for $|x| > AC_{i-1}$. It is clear that, by choosing C_i' large enough, we can construct find a function $\tilde{g}_i$ defined on $[AC_{i-1},\ C_i']$ such that g_i, defined by

$$g_i(x) := \begin{cases} g_{i-1}(x) & \text{if } |x| \leq AC_{i-1}, \\ \tilde{g}_i(x) & \text{if } C_i' \geq |x| \geq AC_{i-1}, \\ i+1 & \text{if } |x| \geq C_i' \end{cases} \tag{5-41}$$

is smooth and slowly increasing; also it satisfies (1-2), (1-10), and

$$\int_{AC_{i-1}}^{C_i'} \frac{1}{y g_i^2(y)}\, dy \geq i. \tag{5-42}$$

It remains to determine C_i in the statement of Lemma 5.1. To do that we slightly modify the reasoning in Step 1.

We sketch the argument. Let $h_i(x) := g_i(x) - (i+1)$. Then $h_i(x) = 0$ if $|x| > C_i'$. It is not difficult to see that

$$|h_i(x)| \lesssim_i \frac{1}{|x|^{\frac{p-1}{2}+}}, \tag{5-43}$$

$$|h_i'(x)| \lesssim_i \frac{1}{|x|^{\frac{p+1}{2}+}}. \tag{5-44}$$

Let $u_{[i]}$ and $v_{[i]}$ be the solutions of the equations

$$\begin{cases} \partial_{tt} u_{[i]} - \triangle u_{[i]} = -|u_{[i]}|^{p-1} u_{[i]} g_i(|u_{[i]}|), \\ u_{[i]}(0) := u_0, \\ \partial_t u_{[i]}(0) := u_1 \end{cases} \tag{5-45}$$

and

$$\begin{cases} \partial_{tt} v_{[i]} - \triangle v_{[i]} = -(i+1)|v_{[i]}|^{p-1} v_{[i]}, \\ v_{[i]}(0) := u_0, \\ \partial_t v_{[i]}(0) := u_1 \end{cases} \tag{5-46}$$

Step 2a. We have

$$\|v_{[i]}\|_{X(\mathbb{R})} \lesssim_i 1, \tag{5-47}$$

by adapting the proof of Step 1a. Notice, in particular, that we can use Conjecture 1.1 and control $\|v_{[i]}\|_{S(\mathbb{R})}$ since $w_{[i]} := (i+1)^{\frac{1}{p-1}} v_{[i]}$ satisfies $\partial_{tt} w_{[i]} - \triangle w_{[i]} = -|w_{[i]}|^{p-1} w_{[i]}$.

Step 2b. We have $\|u_{[i]} - v_{[i]}\|_{X([0,\tilde{t}])} \lesssim_i 1$ for $\tilde{t} \ll_i 1$, by adapting the proof of Step 1b. The dependance on i basically comes from (5-43), (5-44) and (5-46).

Step 2c. We prove that $\|u_{[i]}\|_{X([-T,T])} \lesssim_{i,p} 1$. By time reversal symmetry, it is enough to control $\|u_{[i]}\|_{X([0,T])}$. Recall that $T \leq i$. We chop $[0, T]$ into subintervals $(J_{k'} = [a_{k'}, b_{k'}])_{1 \leq k' \leq l'}$ such that $|J_{k'}| = \tilde{t}$ for $1 \leq k' < l'$ and $|J_{l'}| \leq \tilde{t}$ (with $\tilde{t}$ defined in Step 2b). By Condition 1.3 and the assumption $\|(u_0, u_1)\|_{\tilde{H}^2 \times \tilde{H}^1} \leq i$, we have

$$\|(u_{[i]}(a_{k'}), \partial_t u_{[i]}(a_{k'}))\|_{\dot{H}^{s_p}(\mathbb{R}^3) \times \dot{H}^{s_p-1}(\mathbb{R}^3)} \leq \sup_{t \in I_{\max, g_i}} \left\| (u_{[i]}(t),\ \partial_t u_{[i]}(t)) \right\|_{\dot{H}^{s_p}(\mathbb{R}^3) \times \dot{H}^{s_p-1}(\mathbb{R}^3)} \lesssim_i 1. \tag{5-48}$$

We introduce

$$\begin{cases} \partial_{tt} v_{[i,k']} - \triangle v_{[i,k']} = -(i+1)|v_{[i,k']}|^{p-1} v_{[i,k']}, \\ v_{[i,k']}(a_{k'}) = u_{[i]}(a_{k'}), \\ \partial_t v_{[i,k']}(a_{k'}) = \partial_t u_{[i]}(a_{k'}) \end{cases} \tag{5-49}$$

and, by using (5-48), we can prove that

$$\|u_{[i]}\|_{S([-i,i])} \lesssim_i 1. \tag{5-50}$$

Step 2d. By using Proposition 1.8 and (5-50) we get

$$\max\left(\|u_{[i]}\|_{S([-i,i])},\ \|(u_{[i]}, \partial_t u_{[i]})\|_{L_t^\infty \tilde{H}^2([-i,i]) \times L_t^\infty \tilde{H}^1([-i,i])}\right) \lesssim_i 1. \tag{5-51}$$

We can assume without loss of generality that the constant implicit in $\lesssim$ is larger than i and C_i'. Let C_i be this constant; (5-2) and (5-3) are satisfied.

This concludes Step 2, and the proof of Lemma 5.1. □

Acknowledgements

The author would like to thank Terence Tao for suggesting this problem and for valuable discussions related to this work.

References

[Bahouri and Chemin 2006] H. Bahouri and J.-Y. Chemin, "On global well-posedness for defocusing cubic wave equation", *Int. Math. Res. Not.* **2006** (2006), Art. ID 54873. MR 2008b:35178

[Bahouri and Gérard 1999] H. Bahouri and P. Gérard, "High frequency approximation of solutions to critical nonlinear wave equations", *Amer. J. Math.* **121**:1 (1999), 131–175. MR 2000i:35123 Zbl 0919.35089

[Bahouri and Shatah 1998] H. Bahouri and J. Shatah, "Decay estimates for the critical semilinear wave equation", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **15**:6 (1998), 783–789. MR 99h:35136 Zbl 0924.35084

[Christ and Weinstein 1991] F. M. Christ and M. I. Weinstein, "Dispersion of small amplitude solutions of the generalized Korteweg-de Vries equation", *J. Funct. Anal.* **100**:1 (1991), 87–109. MR 92h:35203 Zbl 0743.35067

[Gallagher and Planchon 2003] I. Gallagher and F. Planchon, "On global solutions to a defocusing semi-linear wave equation", *Rev. Mat. Iberoamericana* **19**:1 (2003), 161–177. MR 2004k:35265 Zbl 1036.35142

[Ginibre and Velo 1989] J. Ginibre and G. Velo, "Scattering theory in the energy space for a class of nonlinear wave equations", *Comm. Math. Phys.* **123**:4 (1989), 535–573. MR 90i:35172 Zbl 0698.35112

[Ginibre and Velo 1995] J. Ginibre and G. Velo, "Generalized Strichartz inequalities for the wave equation", *J. Funct. Anal.* **133**:1 (1995), 50–68. MR 97a:46047 Zbl 0849.35064

[Grillakis 1990] M. G. Grillakis, "Regularity and asymptotic behaviour of the wave equation with a critical nonlinearity", *Ann. of Math.* (2) **132**:3 (1990), 485–509. MR 92c:35080 Zbl 0736.35067

[Grillakis 1992] M. G. Grillakis, "Regularity for the wave equation with a critical nonlinearity", *Comm. Pure Appl. Math.* **45**:6 (1992), 749–774. MR 93e:35073 Zbl 0785.35065

[Kapitanski 1994] L. Kapitanski, "Global and unique weak solutions of nonlinear wave equations", *Math. Res. Lett.* **1**:2 (1994), 211–223. MR 95f:35158 Zbl 0841.35067

[Kenig and Merle 2006] C. E. Kenig and F. Merle, "Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear Schrödinger equation in the radial case", *Invent. Math.* **166**:3 (2006), 645–675. MR 2007g:35232

[Kenig and Merle 2011] C. E. Kenig and F. Merle, "Nondispersive radial solutions to energy supercritical non-linear wave equations, with applications", *Amer. J. Math.* **133**:4 (2011), 1029–1065. MR 2823870 Zbl 05947418

[Kenig et al. 1993] C. E. Kenig, G. Ponce, and L. Vega, "Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle", *Comm. Pure Appl. Math.* **46**:4 (1993), 527–620. MR 94h:35229 Zbl 0808.35128

[Kenig et al. 2000] C. E. Kenig, G. Ponce, and L. Vega, "Global well-posedness for semi-linear wave equations", *Comm. Partial Differential Equations* **25**:9-10 (2000), 1741–1752. MR 2001h:35128

[Lindblad and Sogge 1995] H. Lindblad and C. D. Sogge, "On existence and scattering with minimal regularity for semilinear wave equations", *J. Funct. Anal.* **130**:2 (1995), 357–426. MR 96i:35087 Zbl 0846.35085

[Rauch 1981] J. Rauch, "I: The u^5 Klein–Gordon equation; II: Anomalous singularities for semilinear wave equations", pp. 335–364 in *Nonlinear partial differential equations and their applications* (Paris, 1978/1979), edited by H. Brezis and J. L. Lions, Res. Notes in Math. **53**, Pitman, Boston, MA, 1981. MR 83a:35066

[Roy 2007] T. Roy, "Global well-posedness for the radial defocusing cubic wave equation on $\mathbb{R}^3$ and for rough data", *Electron. J. Differential Equations* (2007), No. 166, 22 pp. MR 2008m:35247

[Roy 2009a] T. Roy, "Adapted linear-nonlinear decomposition and global well-posedness for solutions to the defocusing cubic wave equation on $\mathbb{R}^3$", *Discrete Contin. Dyn. Syst.* **24**:4 (2009), 1307–1323. MR 2010j:35327 Zbl 1211.47023

[Roy 2009b] T. Roy, "Global existence of smooth solutions of a 3D log-log energy-supercritical wave equation", *Anal. PDE* **2**:3 (2009), 261–280. MR 2011k:35145 Zbl 1195.35222

[Shatah and Struwe 1994] J. Shatah and M. Struwe, "Well-posedness in the energy space for semilinear wave equations with critical growth", *Internat. Math. Res. Notices* 7 (1994), 303–309. MR 95e:35132 Zbl 0830.35086

[Struwe 1988] M. Struwe, "Globally regular solutions to the u^5 Klein–Gordon equation", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **15**:3 (1988), 495–513. MR 90j:35142 Zbl 0728.35072

[Tao 2006a] T. Tao, *Nonlinear dispersive equations: Local and global analysis*, CBMS Regional Conference Series in Mathematics **106**, Amer. Math. Soc., Providence, 2006. MR 2008i:35211 Zbl 1106.35001

[Tao 2006b] T. Tao, "Spacetime bounds for the energy-critical nonlinear wave equation in three spatial dimensions", *Dyn. Partial Differ. Equ.* **3**:2 (2006), 93–110. MR 2007c:35116

[Tao 2007] T. Tao, "Global regularity for a logarithmically supercritical defocusing nonlinear wave equation for spherically symmetric data", *J. Hyperbolic Differ. Equ.* **4**:2 (2007), 259–265. MR 2009b:35294 Zbl 1124.35043

[Zhang 2006] X. Zhang, "On the Cauchy problem of 3-D energy-critical Schrödinger equations with subcritical perturbations", *J. Differential Equations* **230**:2 (2006), 422–445. MR 2007h:35325

Received 26 Apr 2010. Revised 17 Jul 2010. Accepted 16 Aug 2010.

TRISTAN ROY: triroy@ias.edu
School of Mathematics, Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, Institute For Advanced Study

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.berkeley.edu/apde.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in APDE are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Analysis & PDE

Volume 5 No. 1 2012

2157-5045(2012)5:1;1-J