ANALYSIS & PDE

Volume 6 No. 8 2013

msp

Analysis & PDE

msp.org/apde

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

The subscription price for 2013 is US $160/year for the electronic version, and $310/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o 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

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 6, No. 8, 2013
dx.doi.org/10.2140/apde.2013.6.1793

L^p AND SCHAUDER ESTIMATES FOR NONVARIATIONAL OPERATORS STRUCTURED ON HÖRMANDER VECTOR FIELDS WITH DRIFT

MARCO BRAMANTI AND MAOCHUN ZHU

Let

$$\mathcal{L} = \sum_{i,j=1}^{q} a_{ij}(x) X_i X_j + a_0(x) X_0,$$

where $X_0, X_1, \dots, X_q$ are real smooth vector fields satisfying Hörmander's condition in some bounded domain $\Omega \subset \mathbb{R}^n$ $(n > q+1)$, and the coefficients $a_{ij} = a_{ji}$, a_0 are real valued, bounded measurable functions defined in Ω, satisfying the uniform positivity conditions

$$\mu|\xi|^2 \le \sum_{i,j=1}^{q} a_{ij}(x)\xi_i\xi_j \le \mu^{-1}|\xi|^2, \quad \mu \le a_0(x) \le \mu^{-1},$$

for a.e. $x \in \Omega$, every $\xi \in \mathbb{R}^q$, and some constant $\mu > 0$.

We prove that if the coefficients a_{ij}, a_0 belong to the Hölder space $C_X^\alpha(\Omega)$ with respect to the distance induced by the vector fields, local Schauder estimates of the following kind hold:

$$\|X_i X_j u\|_{C_X^\alpha(\Omega')} + \|X_0 u\|_{C_X^\alpha(\Omega')} \le c\{\|Lu\|_{C_X^\alpha(\Omega)} + \|u\|_{L^\infty(\Omega)}\}$$

for any $\Omega' \Subset \Omega$.

If the coefficients a_{ij}, a_0 belong to the space $\mathrm{VMO}_{X,\mathrm{loc}}(\Omega)$ with respect to the distance induced by the vector fields, local L^p estimates of the following kind hold, for every $p \in (1, \infty)$:

$$\|X_i X_j u\|_{L^p(\Omega')} + \|X_0 u\|_{L^p(\Omega')} \le c\{\|Lu\|_{L^p(\Omega)} + \|u\|_{L^p(\Omega)}\}.$$

This project was supported by the National Natural Science Foundation of China (grant numbers 11271299 and 11001221).
MSC2010: primary 35H20; secondary 42B20, 35B45, 53C17.
Keywords: Hörmander's vector fields, Schauder estimates, L^p estimates, drift.

1. Introduction

Let us consider a family of real smooth vector fields

$$X_i = \sum_{j=1}^{n} b_{ij}(x)\partial_{x_j}, \quad i = 0, 1, 2, \ldots, q$$

(here $q+1 < n$), defined in some bounded domain Ω of $\mathbb{R}^n$ and satisfying Hörmander's condition: the Lie algebra generated by the X_i at any point of Ω spans $\mathbb{R}^n$. Under these assumptions, Hörmander's operators

$$\mathscr{L} = \sum_{i=1}^{q} X_i^2 + X_0$$

have been studied since the late 1960s. Hörmander [1967] proved that $\mathscr{L}$ is hypoelliptic, while Rothschild and Stein [1976] proved that, for these operators, a priori estimates of L^p type for second order derivatives with respect to the vector fields hold, namely,

$$\sum_{i,j=1}^{q} \|X_i X_j u\|_{L^p(\Omega')} + \|X_0 u\|_{L^p(\Omega')} \le c\left\{\|\mathscr{L}u\|_{L^p(\Omega)} + \|u\|_{L^p(\Omega)} + \sum_{i=1}^{q} \|X_i u\|_{L^p(\Omega)}\right\} \tag{1-1}$$

for any $p \in (1, \infty)$, $\Omega' \Subset \Omega$.

Note that the "drift" vector field X_0 has weight two, compared with the vector fields

$$X_i \quad \text{for } i = 1, 2, \ldots, q.$$

Many more results have been proved in the literature for operators without the drift term ("sum of squares" of Hörmander type) than for complete Hörmander's operators. On the other hand, complete operators owe their interest, for instance, to the class of Kolmogorov–Fokker–Planck operators, which arise naturally in many fields of physics, natural sciences, and finance as the transport-diffusion equations satisfied by the transition probability density of stochastic systems of ODEs which describe some real system governed by a basically deterministic law perturbed by some kind of white noise. The study of Kolmogorov–Fokker–Planck operators in the framework of Hörmander's operators received a strong impulse from [Lanconelli and Polidoro 1994], which started a lively line of research. We refer to [Lanconelli et al. 2002] for a good survey of this field, with further motivations for the study of these equations and related references.

Let us also note that the study of Hörmander's operators is considerably easier when $\mathscr{L}$ is left invariant with respect to a suitable Lie group of translations and homogeneous of degree two with respect to a suitable family of dilations (which are group automorphisms of the corresponding group of translations). In this case we say that $\mathscr{L}$ has an underlying structure of homogeneous group and, by a famous result due to Folland [1975], $\mathscr{L}$ possesses a homogeneous left invariant global fundamental solution, which turns out to be a precious tool in proving a priori estimates.

In the last ten years, more general classes of nonvariational operators structured on Hörmander's vector fields have been studied, namely,

$$\mathcal{L} = \sum_{i,j=1}^{q} a_{ij}(x) X_i X_j, \tag{1-2}$$

$$\mathcal{L} = \sum_{i,j=1}^{q} a_{ij}(x,t) X_i X_j - \partial_t, \tag{1-3}$$

$$\mathcal{L} = \sum_{i,j=1}^{q} a_{ij}(x) X_i X_j + a_0(x) X_0, \tag{1-4}$$

where the matrix $\{a_{ij}(\cdot)\}_{i,j=1}^{q}$ is symmetric positive definite and the coefficients are bounded (a_0 is bounded away from zero) and satisfy suitable mild regularity assumptions; for instance, they belong to Hölder or VMO spaces defined with respect to the distance induced by the vector fields. Since the a_{ij}'s are not C^∞, these operators are no longer hypoelliptic. Nevertheless, a priori estimates on second order derivatives with respect to the vector fields are a natural result which does not in principle require smoothness of the coefficients. Namely, a priori estimates in L^p (with coefficients a_{ij} in $\mathrm{VMO}_X \cap L^\infty$) have been proved for operators (1-2) [Bramanti and Brandolini 2000a] and for operators (1-4) [Bramanti and Brandolini 2000b] but in homogeneous groups; a priori estimates in C_X^α spaces (with coefficients a_{ij} in C_X^α) have been proved for operators (1-3) [Bramanti and Brandolini 2007] and for operators (1-4) [Gutiérrez and Lanconelli 2009] but in homogeneous groups. Here the Hölder space C_X^α and the VMO_X space are defined with respect to the distance induced by the vector fields (see Section 3D for precise definitions).

In the particular case of Kolmogorov–Fokker–Planck operators, which can be written as

$$\mathcal{L} = \sum_{i,j=1}^{q} a_{ij}(x) \partial^2_{x_i x_j} + X_0$$

for a suitable drift X_0, L^p estimates (when a_{ij} are VMO) have been proved [Bramanti et al. 1996] in homogeneous groups, while Schauder estimates (when a_{ij} are Hölder continuous) have been proved [Di Francesco and Polidoro 2006] under more general assumptions (namely, assuming the existence of translations but not necessarily dilations, adapted to the operator). We recall that the idea of proving L^p estimates for nonvariational operators with leading coefficients in $\mathrm{VMO} \cap L^\infty$ (instead of assuming their uniform continuity) appeared for the first time in [Chiarenza et al. 1991; Chiarenza et al. 1993] by Chiarenza, Frasca, and Longo, in the uniformly elliptic case.

The aim of the present paper is to prove both L^p and C^α local estimates for general operators (1-4) structured on Hörmander's vector fields "with drift", without assuming the existence of any group structure, under the appropriate assumptions on the coefficients a_{ij}, a_0. Namely, our basic estimates read as follows:

$$\|u\|_{S_X^{2,p}(\Omega')} \le c\{\|\mathcal{L}u\|_{L^p(\Omega)} + \|u\|_{L^p(\Omega)}\} \tag{1-5}$$

for $p \in (1, \infty)$ and any $\Omega' \Subset \Omega$ if the coefficients are $\mathrm{VMO}_{X,\mathrm{loc}}(\Omega)$, and

$$\|u\|_{C_X^{2,\alpha}(\Omega')} \le c\{\|\mathscr{L}u\|_{C_X^{\alpha}(\Omega)} + \|u\|_{L^\infty(\Omega)}\} \tag{1-6}$$

for $\alpha \in (0, 1)$ and $\Omega' \Subset \Omega$ if the coefficients are $C_X^\alpha(\Omega)$. The related Sobolev and Hölder spaces $S_X^{2,p}$, $C_X^{2,\alpha}$ are those induced by the vector fields X_i, and will be precisely defined in Section 3D. Clearly, these estimates are more general than those contained in all the aforementioned papers.

At first sight, this kind of result could seem a straightforward generalization of existing theories. However, several difficulties exist, some hidden in subtle details. We are going to describe some of them. First of all, we have to remark that in [Rothschild and Stein 1976], although $S_X^{2,p}$ estimates are stated for both sum of squares and complete Hörmander's operators, proofs are given only in the first case. While some adaptations are quite straightforward, this is not always the case. Therefore, some results proved in the present paper can be seen also as a detailed proof of results stated in [Rothschild and Stein 1976], in the drift case. One of the new difficulties in the drift case is related to the proof of suitable representation formulas for second order derivatives $X_i X_j u$ of a test function, in terms of u and $\mathscr{L}u$, via singular integrals and commutators of singular integrals. In turn, the reason why these representation formulas are harder to prove in the presence of a drift relies on the fact that a technical result which allows us to exchange, in a suitable sense, the action of X_i-derivatives with that of suitable integral operators assumes a more involved form when the drift is present.

Once the suitable representation formulas are established, a real variable machinery similar to that used in [Bramanti and Brandolini 2000a; 2007] can be applied, and this is the reason why we have chosen to give in a single paper a unified treatment of L^p and C_X^α estimates. More specifically, one considers a bounded domain Ω endowed with the control distance induced by the vector fields X_i, which has been defined, in the drift case, by Nagel, Stein, and Wainger [Nagel et al. 1985], and the Lebesgue measure, which is locally doubling with respect to these metric balls, as proved in [Nagel et al. 1985]. However, a problem arises when trying to apply to this context known results about singular integrals in metric doubling spaces (or "spaces of homogeneous type", after [Coifman and Weiss 1971]). Namely, what we should know to apply this theory on some domain $\Omega' \Subset \Omega$ is a doubling property such as

$$\mu(B(x, 2r) \cap \Omega') \le c\mu(B(x, r) \cap \Omega') \quad \text{for any } x \in \Omega' \Subset \Omega,\ r > 0 \tag{1-7}$$

while what we actually know, in view of [Nagel et al. 1985], is

$$\mu(B(x, 2r)) \le c\mu(B(x, r)) \quad \text{for any } x \in \Omega' \Subset \Omega,\ 0 < r < r_0. \tag{1-8}$$

It has been known since [Franchi and Lanconelli 1983] that, when Ω' is for instance a metric ball, condition (1-7) follows from (1-8) as soon as the distance satisfies a kind of *segment property* which reads as follows: for any couple of points x_1, x_2 at distance r and for any number $\delta < r$ and $\varepsilon > 0$, there exists a point x_0 having distance $\le \delta$ from x_1 and $\le r - \delta + \varepsilon$ from x_2 (this fact explicitly appears, for instance, from the proof given in [Bramanti and Brandolini 2005, Lemma 4.2]). However, while when the drift term is lacking, the distance induced by the X_i is easily seen to satisfy this property, this is no longer the case when the field X_0 with weight two enters the definition of distance, and, as far as we know, a condition of kind (1-7) has never been proved in this context for a metric ball Ω', or for any other special kind of bounded domain Ω. Thus we are forced to apply a theory of singular integrals which does not require the full strength of the global doubling condition (1-7). A first possibility is to consider the context of *nondoubling spaces*, as studied by Tolsa, Nazarov, Treil, and Volberg, and other authors (see, for instance, [Tolsa 2001; Nazarov et al. 2003] and the references therein). Results of L^p and C^α continuity for singular integrals of this kind, applicable to our context, have been proved in [Bramanti 2010]. However, to prove our L^p estimates (1-5), we also need some *commutator estimates*, of the kind of the well-known result proved by [Coifman et al. 1976], which, as far as we know, are not presently available in the framework of general nondoubling quasimetric (or metric) spaces. For this reason, we have recently developed [Bramanti and Zhu 2012] a theory of *locally homogeneous spaces* which is quite a natural framework where all the results we need about singular integrals and their commutators with BMO functions can be proved. To give a unified treatment of both L^p and C^α estimates, here we have decided to prove both by exploiting the results in [Bramanti and Zhu 2012]. We note that our Schauder estimates could also be obtained by applying the results in [Bramanti 2010], while L^p estimates could not.

Once the basic estimates on second order derivatives are established, a natural, but nontrivial, extension consists in proving similar estimates for derivatives of (weighted) order $k+2$, in terms of k derivatives of $\mathscr{L}u$ (assuming, of course, that the coefficients of the operator possess the corresponding further regularity). In the presence of a drift, it is reasonable to restrict this study to the case of k even, as already appears from the analog result proved in homogeneous groups [Bramanti and Brandolini 2000b]. Even in this case, a proof of this extension seems to be a difficult task, and we have decided not to lengthen the paper to address this problem.

2. Assumptions and main results

We now state precisely our assumptions and main results. All the function spaces involved in the statements below will be precisely defined in Section 3D. Our basic assumption is as follows.

Assumption (H). Let

$$\mathscr{L} = \sum_{i,j=1}^{q} a_{ij}(x) X_i X_j + a_0(x) X_0,$$

where the $X_0, X_1, \dots, X_q$ are real smooth vector fields satisfying Hörmander's condition (see Section 3A) in some bounded domain $\Omega \subset \mathbb{R}^n$ and the coefficients $a_{ij} = a_{ji}$, a_0 are real valued, bounded measurable functions defined in Ω, satisfying the uniform positivity conditions

$$\mu|\xi|^2 \le \sum_{i,j=1}^{q} a_{ij}(x)\xi_i\xi_j \le \mu^{-1}|\xi|^2, \quad \mu \le a_0(x) \le \mu^{-1},$$

for a.e. $x \in \Omega$, every $\xi \in \mathbb{R}^q$, and some constant $\mu > 0$.

Our main results are contained in the next two theorems.

Theorem 2.1. *In addition to* (H), *assume that the coefficients* a_{ij}, a_0 *belong to* $C_X^\alpha(\Omega)$ *for some* $\alpha \in (0,1)$. *Then, for every domain* $\Omega' \Subset \Omega$, *there exists a constant* $c > 0$ *depending on* Ω', Ω, X_i, α, μ, $\|a_{ij}\|_{C_X^\alpha(\Omega)}$, *and* $\|a_0\|_{C_X^\alpha(\Omega)}$ *such that, for every* $u \in C_X^{2,\alpha}(\Omega)$, *one has*

$$\|u\|_{C_X^{2,\alpha}(\Omega')} \le c\{\|\mathcal{L}u\|_{C_X^\alpha(\Omega)} + \|u\|_{L^\infty(\Omega)}\}.$$

Theorem 2.2. *In addition to* (H), *assume that the coefficients* a_{ij}, a_0 *belong to the space* $\mathrm{VMO}_{X,\mathrm{loc}}(\Omega)$. *Then, for every* $p \in (1,\infty)$, *any* $\Omega' \Subset \Omega$, *there exists a constant* c *depending on* X_i, n, q, p, μ, Ω', Ω, *and the* VMO *moduli of* a_{ij} *and* a_0 *such that, for every* $u \in S_X^{2,p}(\Omega)$,

$$\|u\|_{S_X^{2,p}(\Omega')} \le c\{\|\mathcal{L}u\|_{L^p(\Omega)} + \|u\|_{L^p(\Omega)}\}.$$

Remark 2.3. Under the assumptions of the previous theorems, it is not restrictive to assume $a_0(x)$ to be equal to 1, for we can always rewrite (1-4) in the form

$$\sum_{i,j=1}^{q} \frac{a_{ij}}{a_0} X_i X_j + X_0 = \frac{f}{a_0}$$

and apply the a priori estimates to this equation, controlling C_X^α or VMO moduli of the new coefficients a_{ij}/a_0 in terms of the analogous moduli of a_{ij}, a_0, and the constant μ. Therefore, throughout the following we will always take $a_0 \equiv 1$.

3. Known results and preparatory results from real analysis and geometry of vector fields

3A. ***Hörmander's vector fields, lifting, and approximation.*** Let $X_0, X_1, \dots, X_q$ be a system of real smooth vector fields

$$X_i = \sum_{j=1}^{n} b_{ij}(x)\partial_{x_j}, \quad i = 0, 1, 2, \dots, q$$

$(q+1 < n)$ defined in some bounded, open and connected subset Ω of $\mathbb{R}^n$. Let us assign to each X_i a weight p_i, saying that

$$p_0 = 2 \quad \text{and} \quad p_i = 1 \quad \text{for } i = 1, 2, \dots, q.$$

For any multiindex

$$I = (i_1, i_2, \dots, i_k), \quad 0 \le i_j \le q,$$

we define the weight of I as

$$|I| = \sum_{j=1}^{k} p_{i_j}$$

and we set

$$X_I = X_{i_1} X_{i_2} \cdots X_{i_k},$$
$$X_{[I]} = [X_{i_1}, [X_{i_2}, \dots [X_{i_{k-1}}, X_{i_k}] \dots]],$$

where $[X, Y] = XY - YX$ for any couple of vector fields X, Y.

We will say that $X_{[I]}$ is a *commutator of weight* $|I|$. As usual, $X_{[I]}$ can be seen either as a differential operator or as a vector field. We will write

$$X_{[I]} f$$

to denote the differential operator $X_{[I]}$ acting on a function f, and

$$(X_{[I]})_x$$

to denote the vector field $X_{[I]}$ evaluated at the point $x \in \Omega$.

We shall say that $X = \{X_0, X_1, \dots, X_q\}$ satisfies *Hörmander's condition of weight* s if these vector fields, together with their commutators of weight $\leq s$, span the tangent space at every point $x \in \Omega$.

Let ℓ be the free Lie algebra of weight s on $q+1$ generators, that is, the quotient of the free Lie algebra with $q+1$ generators by the ideal generated by the commutators of weight at least $s+1$. We say that the vector fields $X_0, \dots, X_q$, which satisfy Hörmander's condition of weight s at some point $x_0 \in \mathbb{R}^n$, are *free up to order* s *at* x_0 if $n = \dim \ell$, as a vector space (note that inequality $\leq$ always holds). The famous lifting theorem proved by Rothschild and Stein [1976, p. 272] reads as follows.

Theorem 3.1. *Let* $X = (X_0, X_1, \dots, X_q)$ *be* C^∞ *real vector fields on a domain* $\Omega \subset \mathbb{R}^n$ *satisfying Hörmander's condition of weight* s *in* Ω*. Then, for any* $\bar{x} \in \Omega$*, in terms of new variables,* $h_{n+1}, \dots, h_N$*, there exist smooth functions* $\lambda_{il}(x, h)$ $(0 \leq i \leq q,\ n+1 \leq l \leq N)$ *defined in a neighborhood* $\widetilde{U}$ *of* $\bar{\xi} = (\bar{x}, 0) \in \mathbb{R}^N$ *such that the vector fields* $\widetilde{X}_i$ *given by*

$$\widetilde{X}_i = X_i + \sum_{l=n+1}^{N} \lambda_{il}(x, h) \frac{\partial}{\partial h_l}, \quad i = 0, \dots, q,$$

satisfy Hörmander's condition of weight s *and are free up to weight* s *at every point in* $\widetilde{U}$*.*

Let $\widetilde{X} = (\widetilde{X}_0, \widetilde{X}_1, \dots, \widetilde{X}_q)$ be the lifted vector fields which are free up to weight s at some point $\xi \in \mathbb{R}^N$ and let ℓ be the free Lie algebra generated by $\widetilde{X}$. For each j, $1 \leq j \leq s$, we can select a family $\{\widetilde{X}_{j,k}\}_k$ of commutators of weight j, with $\widetilde{X}_{1,k} = \widetilde{X}_k$, $\widetilde{X}_{2,1} = \widetilde{X}_0$, $k = 1, 2, \dots, q$, such that $\{\widetilde{X}_{j,k}\}_{jk}$ is a basis of ℓ, that is to say, there exists a set A of double-indices α such that $\{\widetilde{X}_\alpha\}_{\alpha \in A}$ is a basis of ℓ. Note that $\operatorname{Card} A = N$, which allows us to identify ℓ with $\mathbb{R}^N$.

Now, in $\mathbb{R}^N$ we can consider the group structure of $N(q+1, s)$, which is the simply connected Lie group associated to ℓ. We will write $\circ$ for the Lie group operation (which we think of as a *translation*) and assume that the group identity is the origin. It is also possible to assume that $u^{-1} = -u$ (the group inverse is the Euclidean opposite). We can naturally define *dilations* in $N(q+1, s)$ by

$$D(\lambda)((u_\alpha)_{\alpha \in A}) = (\lambda^{|\alpha|} u_\alpha)_{\alpha \in A} \tag{3-1}$$

with $|j, k| = j$. These are group automorphisms, hence $N(q+1, s)$ is a *homogeneous group*, in the sense of Stein [1993, pp. 618–622]. We will call this group $\mathbb{G}$, leaving the numbers q, s implicitly understood.

We can define in $\mathbb{G}$ a *homogeneous norm* $\|\cdot\|$ as follows. For any $u \in \mathbb{G}$, $u \neq 0$, set

$$\|u\| = r \Leftrightarrow \left| D\left(\frac{1}{r}\right) u \right| = 1,$$

where $|\cdot|$ denotes the Euclidean norm.

The function

$$d_{\mathbb{G}}(u, v) = \|v^{-1} \circ u\|$$

is a *quasidistance*, that is

$$\begin{aligned} &d_{\mathbb{G}}(u, v) \geq 0 \quad \text{and} \quad d_{\mathbb{G}}(u, v) = 0 \quad \text{if and only if } u = v, \\ &d_{\mathbb{G}}(u, v) = d_{\mathbb{G}}(v, u), \\ &d_{\mathbb{G}}(u, v) \leq c(d_{\mathbb{G}}(u, z) + d_{\mathbb{G}}(z, v)) \end{aligned} \tag{3-2}$$

for every $u, v, z \in \mathbb{G}$ and some positive constant $c(\mathbb{G}) \geq 1$. We define the balls with respect to $d_{\mathbb{G}}$ as

$$B(u, r) := \{v \in \mathbb{R}^N : d_{\mathbb{G}}(u, v) < r\}.$$

It can be proved [Stein 1993, p. 619] that the Lebesgue measure in $\mathbb{R}^N$ is the Haar measure of $\mathbb{G}$. Therefore, by (3-1),

$$|B(u, r)| = |B(u, 1)| r^Q$$

for every $u \in \mathbb{G}$ and $r > 0$, where $Q = \sum_{\alpha \in A} |\alpha|$. We will call Q the *homogeneous dimension* of $\mathbb{G}$.

Let τ_u be the left translation operator acting on functions: $(\tau_u f)(v) = f(u \circ v)$. We say that a differential operator P on $\mathbb{G}$ is *left invariant* if $P(\tau_u f) = \tau_u(Pf)$ for every smooth function f.

We say that a differential operator P on $\mathbb{G}$ is *homogeneous of degree* $\delta > 0$ if

$$P(f(D(\lambda)u)) = \lambda^\delta (Pf)(D(\lambda)u)$$

for every test function f and every $\lambda > 0$, $u \in \mathbb{G}$. We also say that a function f is *homogeneous of degree* $\delta \in \mathbb{R}$ if

$$f(D(\lambda)u) = \lambda^\delta f(u) \quad \text{for every } \lambda > 0,\ u \in \mathbb{G}.$$

Clearly, if P is a differential operator homogeneous of degree δ_1 and f is a homogeneous function of degree δ_2, then Pf is a homogeneous function of degree $\delta_2 - \delta_1$, while fP is a differential operator, homogeneous of degree $\delta_1 - \delta_2$.

Let Y_α be the left invariant vector field which agrees with $\partial/(\partial u_\alpha)$ at 0 and set $Y_{1,k} = Y_k$, $k = 1, \dots, q$, $Y_{2,1} = Y_0$. The differential operator $Y_{i,k}$ is homogeneous of degree i, and $\{Y_\alpha\}_{\alpha \in A}$ is a basis of the free Lie algebra ℓ.

A differential operator on $\mathbb{G}$ is said to have *local degree less than or equal to* λ if, after taking the Taylor expansion at 0 of its coefficients, each term obtained is a differential operator homogeneous of degree $\leq \lambda$.

Also, a function on $\mathbb{G}$ is said to have *local degree greater than or equal to* λ if, after taking the Taylor expansion at 0 of its coefficients, each term obtained is a homogeneous function of degree $\geq \lambda$. For $\xi, \eta \in \widetilde{U}$, define the map

$$\Theta_\eta(\xi) = (u_\alpha)_{\alpha \in A}$$

with $\xi = \exp\bigl(\sum_{\alpha \in A} u_\alpha \widetilde{X}_\alpha\bigr)\eta$. We will also write $\Theta(\eta, \xi) = \Theta_\eta(\xi)$.

We can now state Rothschild and Stein's approximation theorem [1976, p. 273].

Theorem 3.2. *In the coordinates given by $\Theta(\eta,\cdot)$ we can write $\widetilde{X}_i = Y_i + R_i^\eta$ on an open neighborhood of* 0, *where R_i^η is a vector field of local degree at most* 0 *for $i = 1, \dots, q$ (and at most* 1 *for $i = 0$) depending smoothly on η. Explicitly, this means that, for every $f \in C_0^\infty(\mathbb{G})$,*

$$\widetilde{X}_i[f(\Theta(\eta,\cdot))](\xi) = (Y_i f + R_i^\eta f)(\Theta(\eta,\xi)). \tag{3-3}$$

More generally, for every double-index $(i,k) \in A$, we can write

$$\widetilde{X}_{i,k}[f(\Theta(\eta,\cdot))](\xi) = (Y_{i,k} f + R_{i,k}^\eta f)(\Theta(\eta,\xi)), \tag{3-4}$$

where $R_{i,k}^\eta$ is a vector field of local degree $\le i-1$ depending smoothly on η.

Some other important properties of the map Θ are stated in the next theorem (see [Rothschild and Stein 1976, pp. 284–287]).

Theorem 3.3. *Let $\bar{\xi} \in \mathbb{R}^N$ and $\widetilde{U}$ be a neighborhood of $\bar{\xi}$ such that for any $\eta \in \widetilde{U}$ the map $\Theta(\eta,\cdot)$ is well defined in $\widetilde{U}$. For $\xi, \eta \in \widetilde{U}$, define*

$$\rho(\eta,\xi) = \|\Theta(\eta,\xi)\|, \tag{3-5}$$

where $\|\cdot\|$ is the homogeneous norm defined above. Then

(a) $\Theta(\eta,\xi) = \Theta(\xi,\eta)^{-1} = -\Theta(\xi,\eta)$ *for every $\xi, \eta \in \widetilde{U}$;*

(b) *ρ is a quasidistance in $\widetilde{U}$ (that is satisfies the three properties* (3-2)*);*

(c) *under the change of coordinates $u = \Theta_\xi(\eta)$, the measure element becomes*

$$d\eta = c(\xi)\cdot(1+\omega(\xi,u))\,du, \tag{3-6}$$

where $c(\xi)$ is a smooth function, bounded and bounded away from zero in $\widetilde{U}$, $\omega(\xi,u)$ is a smooth function in both variables with

$$|\omega(\xi,u)| \le c\|u\|,$$

and an analogous statement is true for the change of coordinates $u = \Theta_\eta(\xi)$.

Remark 3.4. As we recalled in the introduction, in [Rothschild and Stein 1976] detailed proofs are given only when the drift term X_0 is lacking. A proof of the lifting and approximation results explicitly covering the drift case can be found in [Bramanti et al. 2010], where the theory is also extended to the case of nonsmooth Hörmander's vector fields. We refer to the introduction of [Bramanti et al. 2010] for further bibliographic remarks about existing alternative proofs of the lifting and approximation theorems.

3B. ***Metric induced by vector fields.*** Let us start by recalling the definition of control distance given by Nagel, Stein, Wainger [Nagel et al. 1985] for Hörmander's vector fields with drift.

Definition 3.5. For any $\delta > 0$, let $C(\delta)$ be the class of absolutely continuous mappings $\varphi\colon [0,1] \to \Omega$ which satisfy

$$\varphi'(t) = \sum_{|I|\le s} \lambda_I(t)(X_{[I]})_{\varphi(t)} \quad \text{for a.e. } t \in (0,1) \tag{3-7}$$

with $|\lambda_I(t)| \le \delta^{|I|}$. We define

$$d(x, y) = \inf\{\delta : \text{there exists } \varphi \in C(\delta) \text{ with } \varphi(0) = x, \varphi(1) = y\}.$$

The finiteness of d immediately follows by Hörmander's condition: since the vector fields $\{X_{[I]}\}_{|I|\le s}$ span $\mathbb{R}^n$, we can always join any two points x, y with a curve φ of the kind (3-7); moreover, d turns out to be a distance. Analogously to what Nagel, Stein, and Wainger [Nagel et al. 1985] do when X_0 is lacking, in [Bramanti et al. 2013] the following notion is introduced.

Definition 3.6. For any $\delta > 0$, let $C_1(\delta)$ be the class of absolutely continuous mappings $\varphi : [0, 1] \to \Omega$ which satisfy

$$\varphi'(t) = \sum_{i=0}^{q} \lambda_i(t)(X_i)_{\varphi(t)} \quad \text{for a.e. } t \in (0, 1)$$

with $|\lambda_0(t)| \le \delta^2$ and $|\lambda_j(t)| \le \delta$ for $j = 1, \dots, q$. We define

$$d_X(x, y) = \inf\{\delta : \text{there exists } \varphi \in C_1(\delta) \text{ with } \varphi(0) = x, \varphi(1) = y\}.$$

Note that the finiteness of $d_X(x, y)$ for any two points $x, y \in \Omega$ is not a trivial fact, but depends on a connectivity result ("Chow's theorem"); moreover, it can be proved that d and d_X are locally equivalent, and that d_X is still a distance (see [Bramanti et al. 2013], where these results are proved in the more general setting of nonsmooth vector fields). From now on we will always refer to d_X as the *control distance* induced by the system of Hörmander's vector fields X. It is well-known that this distance is topologically equivalent to the Euclidean one. For any $x \in \Omega$, we set

$$B(x, r) = \{y \in \Omega : d_X(x, y) < r\}.$$

The basic result about the measure of metric balls is the famous local doubling condition.

Theorem 3.7 [Nagel et al. 1985]. *For every $\Omega' \Subset \Omega$ there exist positive constants c, r_0 such that, for any $x \in \Omega'$, $r \le r_0$,*

$$|B(x, 2r)| \le c|B(x, r)|.$$

As already pointed out in the introduction, the distance d_X does *not* satisfy the segment property: given two points at distance r, it is generally impossible to find a third point at distance $r/2$ from both. A weaker property which this distance actually satisfies is contained in the next lemma, and will be useful when dealing with the properties of Hölder spaces C^α_X.

Lemma 3.8. *For any $x, y \in \Omega$, positive integer n, $\varepsilon > 0$, we can join x to y with a curve γ and find $n + 1$ points $p_0 = x, p_1, p_2, \dots, p_n = y$ on γ, such that*

$$d_X(p_j, p_{j+1}) \le \frac{1+\varepsilon}{\sqrt{n}} d_X(x, y) \quad \textit{for } j = 0, 2, \dots, n-1.$$

Proof. For any $x, y \in \Omega$ with $d_X(x, y) = R$, any $\varepsilon > 0$, by Definition 3.6 we can join x and y with a curve $\gamma(t)$ satisfying

$$\gamma(0) = x, \quad \gamma(1) = y, \quad \gamma'(t) = \sum_{i=0}^{q} \lambda_i(t)(X_i)_{\gamma(t)},$$

with $|\lambda_i(t)| \le R(1+\varepsilon)$, for $i = 1, \dots, q$ and $|\lambda_0(t)| \le (R(1+\varepsilon))^2$.

Let $\gamma_j(t) = \gamma((t+j)/n)$ for $j = 0, 1, 2, \dots, n-1$. Then $\gamma_j(t)$ satisfies

$$\gamma_j(0) = \gamma\left(\frac{j}{n}\right) =: p_j, \quad \gamma_j(1) = \gamma\left(\frac{j+1}{n}\right) = p_{j+1}.$$

In particular, $p_0 = x$ and $p_n = y$. Moreover,

$$\gamma_j'(t) = \frac{1}{n}\sum_{i=0}^{q} \lambda_i\left(\frac{t+j}{n}\right)(X_i)_{\gamma_j(t)}$$

with

$$\left|\frac{1}{n}\lambda_0\left(\frac{t+j}{n}\right)\right| \le \left(\frac{R(1+\varepsilon)}{\sqrt{n}}\right)^2, \quad \left|\frac{1}{n}\lambda_i\left(\frac{t+j}{n}\right)\right| < \frac{R(1+\varepsilon)}{\sqrt{n}}$$

for $i = 1, \dots, q$, $j = 0, 2, \dots, n-1$. Thus

$$d_X(p_j, p_{j+1}) \le \frac{R(1+\varepsilon)}{\sqrt{n}}$$

for $j = 0, 2, \dots, n-1$, so we are done. □

The free lifted vector fields $\widetilde{X}_i$ induce, in the neighborhood where they are defined, a control distance $d_{\widetilde{X}}$; we will denote by $\widetilde{B}(\xi, r)$ the corresponding metric balls. In this lifted setting we can also consider the quasidistance ρ defined in (3-5). The two functions turn out to be equivalent.

Lemma 3.9. *Let $\bar{\xi}, \widetilde{U}$ be as in Theorem 3.3. There exists $\widetilde{B}(\bar{\xi}, R) \subset \widetilde{U}$ such that the distance $d_{\widetilde{X}}$ is equivalent to the quasidistance ρ in* (3-5) *in $\widetilde{B}(\bar{\xi}, R)$, and both are greater than the Euclidean distance*; *namely*, *there exist positive constants c_1, c_2, c_3 such that*

$$c_1|\xi - \eta| \le c_2\rho(\eta, \xi) \le d_{\widetilde{X}}(\eta, \xi) \le c_3\rho(\eta, \xi) \quad \textit{for every } \xi, \eta \in \widetilde{B}(\bar{\xi}, R).$$

This fact is proved in [Nagel et al. 1985]; see also [Bramanti et al. 2010, Proposition 22].

3C. ***Locally homogeneous spaces.*** We are now going to recall the notion of *locally homogeneous space*, introduced in [Bramanti and Zhu 2012]. Roughly speaking, a locally homogeneous space is a set Ω endowed with a function d which is a quasidistance on any compact subset, and a measure μ which is locally doubling, in a sense which will be made precise below. In our concrete situation, our set is endowed with a function d which is a *distance* in Ω, and a locally doubling measure. We can therefore give the following definition, which is simpler than that given in [Bramanti and Zhu 2012].

Definition 3.10. Let (Ω, d) be a metric space, and let μ be a positive regular Borel measure in Ω.

Assume there exists an increasing sequence $\{\Omega_n\}_{n=1}^{\infty}$ of bounded measurable subsets of Ω such that

$$\bigcup_{n=1}^{\infty} \Omega_n = \Omega \tag{3-8}$$

and, for any $n = 1, 2, 3, \ldots,$

(i) the closure of Ω_n in Ω is compact,

(ii) there exists $\varepsilon_n > 0$ such that

$$\{x \in \Omega : d(x, y) < 2\varepsilon_n \quad \text{for some } y \in \Omega_n\} \subset \Omega_{n+1}, \tag{3-9}$$

(iii) there exists $C_n > 1$ such that, for any $x \in \Omega_n$, $0 < r \leq \varepsilon_n$, we have

$$0 < \mu(B(x, 2r)) \leq C_n \mu(B(x, r)) < \infty. \tag{3-10}$$

(Note that for $x \in \Omega_n$ and $r \leq \varepsilon_n$ we also have $B(x, 2r) \subset \Omega_{n+1}$.)

We say that $(\Omega, \{\Omega_n\}_{n=1}^{\infty}, d, \mu)$ is a *(metric) locally homogeneous space* if the above assumptions hold.

Any space satisfying the above definition a fortiori satisfies the definition of locally homogeneous space given in [Bramanti and Zhu 2012].

Next, we discuss some facts about local singular kernels. For fixed Ω_n, Ω_{n+1}, and a fixed ball $B(\bar{x}, R_0)$, with $\bar{x} \in \Omega_n$ and $R_0 < 2\varepsilon_n$ (hence $B(\bar{x}, R_0) \subset \Omega_{n+1}$), let $K(x, y)$ be a measurable function defined for $x, y \in B(\bar{x}, R_0)$, $x \neq y$. We now list a series of possible assumptions on the kernel K which are involved in the theorems that we will apply in the following.

(i) We say that K satisfies the *standard estimates* for some $\nu \in [0, 1)$ if the following hold:

$$|K(x, y)| \leq \frac{A d(x, y)^{\nu}}{\mu(B(x, d(x, y)))} \tag{3-11}$$

for $x, y \in B(\bar{x}, R_0)$ with $x \neq y$, and

$$|K(x_0, y) - K(x, y)| + |K(y, x_0) - K(y, x)| \leq \frac{B d(x_0, y)^{\nu}}{\mu(B(x_0, d(x_0, y)))} \left(\frac{d(x_0, x)}{d(x_0, y)}\right)^{\beta} \tag{3-12}$$

for any $x_0, x, y \in B(\bar{x}, R_0)$ with $d(x_0, y) > 2d(x_0, x)$, and some $\beta > 0$.

(ii) We say that K satisfies the *cancellation property* if the following holds: there exists $C > 0$ such that, for a.e. $x \in B(\bar{x}, R_0)$ and every $\varepsilon_1, \varepsilon_2$ such that $0 < \varepsilon_1 < \varepsilon_2$ and $B_\rho(x, \varepsilon_2) \subset \Omega_{n+1}$,

$$\left| \int_{\Omega_{n+1}, \varepsilon_1 < \rho(x, y) < \varepsilon_2} K(x, y)\, d\mu(y) \right| + \left| \int_{\Omega_{n+1}, \varepsilon_1 < \rho(x, z) < \varepsilon_2} K(z, x)\, d\mu(z) \right| \leq C, \tag{3-13}$$

where ρ is any *quasidistance* (see (3-2)) equivalent to d in Ω_{n+1} and B_ρ denotes ρ-balls.

(iii) We say that K satisfies the *convergence condition* if the following holds: for a.e. $x \in B(\bar{x}, R_0)$ such that $B_\rho(x, R) \subset \Omega_{n+1}$, there exists

$$h_R(x) \equiv \lim_{\varepsilon \to 0} \int_{\Omega_{n+1}, \varepsilon < \rho(x,y) < R} K(x, y)\, d\mu(y), \tag{3-14}$$

where ρ is any quasidistance equivalent to d in Ω_{n+1}.

Application of the abstract theory to our setting. Let's now explain how this abstract setting will be used to describe our concrete situation. The a priori estimates we will prove in Theorems 2.1 and 2.2 involve a fixed subdomain $\Omega' \Subset \Omega$. Let us fix this Ω' once and for all. For any $\bar{x} \in \Omega'$ we can perform in a suitable neighborhood of $\bar{x}$ the lifting and approximation procedure as explained in Section 3A. Let $\bar{\xi} = (\bar{x}, 0) \in \mathbb{R}^N$ and $\widetilde{B}(\bar{\xi}, R)$ be as in Lemma 3.9. Then we can choose

$$\widetilde{\Omega} = \widetilde{B}(\bar{\xi}, R);\ \widetilde{\Omega}_k = \widetilde{B}\left(\bar{\xi}, \frac{kR}{k+1}\right) \quad \text{for } k = 1, 2, 3, \dots.$$

By the properties of $d_{\widetilde{X}}$ that we have listed in Section 3B, and particularly Theorem 3.7, we see that

$$(\widetilde{\Omega}, \{\widetilde{\Omega}_k\}_{k=1}^{\infty}, d_{\widetilde{X}}, d\xi)$$

is a metric locally homogeneous space. The function $\rho(\xi, \eta) = \|\Theta(\eta, \xi)\|$ will play the role of the quasidistance appearing in conditions (3-13) and (3-14), in view of Lemma 3.9. This is the basic setting where we will apply several results about singular integrals in locally homogeneous spaces, which have been proved in [Bramanti and Zhu 2012]. Here we do not repeat the statements of all those theorems. Instead, we will give a precise reference to [Bramanti and Zhu 2012] for each one. We just note that, since in our situation we are dealing with a *metric* locally homogeneous space, the constants which are called B_n in [Bramanti and Zhu 2012], here are equal to 1.

In the space of the original variables (Ω, d_X, dx), instead, we will not apply singular integral estimates, but we will again use the local doubling condition when we establish some important properties of function spaces C^α and VMO (see Section 3D). Note that if Ω_k is an increasing sequence of domains with $\Omega_k \Subset \Omega_{k+1} \Subset \Omega$, we can say that

$$(\Omega, \{\Omega_k\}_k, d_X, dx)$$

is a metric locally homogeneous space.

3D. *Function spaces.* The aim of this section is twofold. First, we want to define the basic function spaces we will need and point out their main properties; second, we want to find a relation between function spaces defined over a ball $B(\bar{x}, r) \subset \Omega \subset \mathbb{R}^n$ and those over the corresponding lifted ball $\widetilde{B}(\bar{\xi}, r) \subset \mathbb{R}^N$. More precisely, we need to know that $f(x)$ belongs to some function space on B if and only if $\tilde{f}(x, h) = f(x)$ belongs to the analogous function space on $\widetilde{B}$. This last fact relies on the following known result; see [Nagel et al. 1985, Lemmas 3.1 and 3.2, p. 139].

Theorem 3.11. *Let us denote by B and $\widetilde{B}$ the balls defined with respect to d_X and $d_{\widetilde{X}}$, respectively. There exist constants $\delta_0 \in (0,1)$, $r_0, c_1, c_2 > 0$ such that*

$$c_1 \operatorname{vol}(\widetilde{B}_r(x,h)) \leq \operatorname{vol}(B_r(x)) \cdot \operatorname{vol}\{h' \in \mathbb{R}^{N-n} : (z,h') \in \widetilde{B}_r(x,h)\} \leq c_2 \operatorname{vol}(\widetilde{B}_r(x,h)) \tag{3-15}$$

*for every $x \in \Omega$, $z \in B_{\delta_0 r}(x)$, and $r \leq r_0$. (Here "*vol*" stands for the Lebesgue measure in the appropriate dimension, x denotes a point in $\mathbb{R}^n$, and h a point in $\mathbb{R}^{N-n}$). More precisely, the condition $z \in B_{\delta_0 r}(x)$ is needed only for the validity of the first inequality in* (3-15)*. Moreover,*

$$d_{\widetilde{X}}((x,h),(x',h')) \geq d_X(x,x'). \tag{3-16}$$

Finally, the projection of the lifted ball $\widetilde{B}_r(x,h)$ on $\mathbb{R}^n$ is just the ball $B(x,r)$, and this projection is onto.

A consequence of the above theorem is the following.

Corollary 3.12. *For any positive function g defined in $B_r(x) \subset \Omega$, $r \leq r_0$, one has*

$$\frac{c_1}{|B_{\delta_0 r}(x)|}\int_{B_{\delta_0 r}(x)} g(y)\,dy \leq \frac{1}{|\widetilde{B}_r(x,h)|}\int_{\widetilde{B}_r(x,h)} g(y)\,dy\,dh' \leq \frac{c_2}{|B_r(x)|}\int_{B_r(x)} g(y)\,dy, \tag{3-17}$$

where δ_0 is the constant in Theorem 3.11.

Proof. By (3-15) and the locally doubling condition, we have, for some fixed $\delta_0 < 1$ as in Theorem 3.11,

$$\begin{aligned}\frac{1}{|\widetilde{B}_r(x,h)|}\int_{\widetilde{B}_r(x,h)} g(y)\,dy\,dh' &= \frac{1}{|\widetilde{B}_r(x,h)|}\int_{B_r(x)} g(y)\,dy \int_{\{h'\in\mathbb{R}^{N-n}:(y,h')\in\widetilde{B}_r(x,h)\}} dh' \\ &\geq \frac{c_1}{|\widetilde{B}_r(x,h)|}\int_{B_{\delta_0 r}(x)} \frac{|\widetilde{B}_r(x,h)|}{|B_r(x)|} g(y)\,dy \geq \frac{c}{|B_{\delta_0 r}(x)|}\int_{B_{\delta_0 r}(x)} g(y)\,dy,\end{aligned}$$

where in the last inequality we exploited the doubling condition $|B_r(x)| \leq c|B_{\delta_0 r}(x)|$, which holds because $B_r(x) \subset \Omega$ and $r \leq r_0$. The proof of the second inequality in (3-17) is analogous but easier, since it involves the second inequality in (3-15), which does not require the condition $y \in B_{\delta_0 r}(x)$. □

3D.1. *Hölder spaces.*

Definition 3.13. For any $0 < \alpha < 1$, $u : \Omega \to \mathbb{R}$, let

$$\begin{aligned} |u|_{C_X^\alpha(\Omega)} &= \sup\left\{\frac{|u(x)-u(y)|}{d_X(x,y)^\alpha} : x,y \in \Omega, x \neq y\right\}, \\ \|u\|_{C_X^\alpha(\Omega)} &= |u|_{C^\alpha(\Omega)} + \|u\|_{L^\infty(\Omega)}, \\ C_X^\alpha(\Omega) &= \{u : \Omega \to \mathbb{R} : \|u\|_{C^\alpha(\Omega)} < \infty\}. \end{aligned}$$

Also, for any positive integer k, let

$$C_X^{k,\alpha}(\Omega) = \{u : \Omega \to \mathbb{R} : \|u\|_{C^{k,\alpha}(\Omega)} < \infty\},$$

with

$$\|u\|_{C_X^{k,\alpha}(\Omega)} = \sum_{|I|=1}^{k} \sum_{j_i=0}^{q} \|X_{j_1}\cdots X_{j_l} u\|_{C^\alpha(\Omega)} + \|u\|_{C^\alpha(\Omega)},$$

where $I = (j_1, j_2, \dots, j_l)$.

We will set $C^{\alpha}_{X,0}(\Omega)$ and $C^{k,\alpha}_{X,0}(\Omega)$ for the subspaces of $C^{\alpha}_{X}(\Omega)$ and $C^{k,\alpha}_{X}(\Omega)$ of functions which are compactly supported in Ω, and set $C^{\alpha}_{\widetilde{X}}(\widetilde{B})$, $C^{k,\alpha}_{\widetilde{X}}(\widetilde{B})$, $C^{\alpha}_{\widetilde{X},0}(\widetilde{B})$, and $C^{k,\alpha}_{\widetilde{X},0}(\widetilde{B})$ for the analogous function spaces over $\widetilde{B}$ defined by the $\widetilde{X}_i$.

We will also write $C^{k,0}_{X}(\Omega)$ to denote the space of functions with continuous X-derivatives up to weight k.

Let us note that we will sometimes also need to use the classical spaces of (possibly compactly supported) continuously differentiable functions, denoted as usual by C^1 (or C^1_0).

The next proposition, adapted from [Bramanti and Brandolini 2007, Proposition 4.2], collects some properties of C^{α} functions which will be useful later. We will apply these properties mainly in the context of lifted variables, that is, for the vector fields $\widetilde{X}_i$ on a ball $\widetilde{B}(\bar{\xi}, R)$.

Proposition 3.14. *Let $B(\bar{x}, 2R)$ be a fixed ball where the vector fields X_i and the control distance d are well defined.*

(i) *For any $\delta \in (0, 1)$ and any $f \in C^1(B(\bar{x}, (1+\delta)R))$, one has*

$$|f(x) - f(y)| \le \frac{c}{\delta} d_X(x, y) \left(\sum_{i=1}^{q} \sup_{B(\bar{x},(1+\delta)R)} |X_i f| + d_X(x, y) \sup_{B(\bar{x},(1+\delta)R)} |X_0 f| \right) \tag{3-18}$$

for any $x, y \in B(\bar{x}, R)$.

If $f \in C^1_0(B(\bar{x}, R))$, one can simply write, for any $x, y \in B(\bar{x}, R)$,

$$|f(x) - f(y)| \le c d_X(x, y) \left(\sum_{i=1}^{q} \sup_{B(\bar{x},R)} |X_i f| + d_X(x, y) \sup_{B(\bar{x},R)} |X_0 f| \right). \tag{3-19}$$

In particular, for $f \in C^1_0(B(\bar{x}, R))$,

$$|f|_{C^{\alpha}(B(\bar{x},R))} \le c R^{1-\alpha} \cdot \left(\sum_{i=1}^{q} \sup_{B(\bar{x},R)} |X_i f| + R \sup_{B(\bar{x},R)} |X_0 f| \right). \tag{3-20}$$

The assumption $f \in C^1$ (or C^1_0) can be replaced by $f \in C^2_X$ (or $C^2_{X,0}$, respectively).

(ii) *For any couple of functions $f, g \in C^{\alpha}_X(B(\bar{x}, R))$, one has*

$$|fg|_{C^{\alpha}_X(B(\bar{x},R))} \le |f|_{C^{\alpha}_X(B(\bar{x},R))} \|g\|_{L^{\infty}(B(\bar{x},R))} + |g|_{C^{\alpha}_X(B(\bar{x},R))} \|f\|_{L^{\infty}(B(\bar{x},R))}$$

and

$$\|fg\|_{C^{\alpha}_X(B(\bar{x},R))} \le 2 \|f\|_{C^{\alpha}_X(B(\bar{x},R))} \|g\|_{C^{\alpha}_X(B(\bar{x},R))}. \tag{3-21}$$

Moreover, if both f and g vanish at least at a point of $B(\bar{x}, R)$, then

$$|fg|_{C^{\alpha}_X(B(\bar{x},R))} \le c R^{\alpha} |f|_{C^{\alpha}_X(B(\bar{x},R))} |g|_{C^{\alpha}_X(B(\bar{x},R))}. \tag{3-22}$$

(iii) *Let $B(x_i, r)$ $(i = 1, 2, \dots, k)$ be a finite family of balls of the same radius r such that $\bigcup_{i=1}^k B(x_i, 2r) \subset \Omega$. Then, for any $f \in C_X^\alpha(\Omega)$,*

$$\|f\|_{C_X^\alpha(\bigcup_{i=1}^k B(x_i,r))} \le c \sum_{i=1}^k \|f\|_{C_X^\alpha(B(x_i,2r))} \tag{3-23}$$

with c depending on the family of balls, but not on f.

(iv) *There exists $r_0 > 0$ such that, for any $f \in C_{X,0}^{2,\alpha}(B(\bar{x}, R))$ and $0 < r \le r_0$, we have the interpolation inequality*

$$\|X_0 f\|_{L^\infty(B(\bar{x},R))} \le r^{\alpha/2} |X_0 f|_{C_X^\alpha(B(\bar{x},R))} + \frac{2}{r} \|f\|_{L^\infty(B(\bar{x},R))}. \tag{3-24}$$

Proof. The proof of (ii)–(iii) is similar to that in [Bramanti and Brandolini 2007, Proposition 4.2], hence we will only prove (i) and (iv).

Throughout this proof we will write d for d_X. (Actually, we will apply this proposition both to d_X and to $d_{\widetilde{X}}$).

(i) Fix $\delta \in (0, 1)$ and let $R' = (1 + \delta)R$. Let us distinguish two cases.

Case 1: $d(x, y) < R' - \max(d(\bar{x}, x), d(\bar{x}, y))$. Let $\varepsilon > 0$ be such that

$$d(x, y) + \varepsilon < R' - \max(d(\bar{x}, x), d(\bar{x}, y)), \tag{3-25}$$

hence, by Definition 3.6, there exists a curve $\varphi(t)$ such that $\varphi(0) = x$, $\varphi(1) = y$, and

$$\varphi'(t) = \sum_{i=0}^q \lambda_i(t) (X_i)_{\varphi(t)}$$

with $|\lambda_i(t)| \le (d(x, y) + \varepsilon)$, $|\lambda_0(t)| \le (d(x, y) + \varepsilon)^2$ for $i = 1, \dots, q$. By (3-25),

$$B(x, d(x, y) + \varepsilon) \subset B(\bar{x}, R'),$$

hence every point $\gamma(t)$ for $t \in (0, 1)$ belongs to $B(\bar{x}, R')$. Then we can write

$$\begin{aligned} |f(x) - f(y)| &= \left| \int_0^1 \frac{d}{dt} f(\varphi(t))\, dt \right| = \left| \int_0^1 \sum_{i=0}^q \lambda_i(t) (X_i f)_{\varphi(t)}\, dt \right| \\ &\le (d(x, y) + \varepsilon) \sum_{i=1}^q \sup_{B(\bar{x},R')} |X_i f| + (d(x, y) + \varepsilon)^2 \sup_{B(\bar{x},R')} |X_0 f|, \end{aligned}$$

and since ε is arbitrary, this implies (3-19) and, in particular, (3-18). We note that the above argument relies on the differentiability of f along the curve φ, which holds under either the assumption $f \in C^1(B(\bar{x}, (1 + \delta)R))$ or $f \in C_X^2(B(\bar{x}, (1 + \delta)R))$ (since X_0 has weight two).

Case 2: $d(x, y) \ge R' - \max(d(\bar{x}, x), d(\bar{x}, y))$. Let us write

$$|f(x) - f(y)| \le |f(x) - f(\bar{x})| + |f(\bar{x}) - f(y)| = A + B.$$

Each of the terms A, B can be bounded by an argument similar to that in *Case* 1 (since both x and y can be joined to $\bar{x}$ by curves contained in $B(\bar{x}, R)$), giving

$$|f(x) - f(y)| \le [d(x, \bar{x}) + d(y, \bar{x})] \cdot \left\{ \sum_{i=1}^{q} \sup_{B(\bar{x},R)} |X_i f| + [d(x, \bar{x}) + d(y, \bar{x})] \sup_{B(\bar{x},R)} |X_0 f| \right\}.$$

Now it is enough to show that

$$d(x, \bar{x}) + d(y, \bar{x}) \le \frac{c}{\delta} d(x, y).$$

To show this, let $r := \max(d(\bar{x}, x), d(\bar{x}, y))$. Then

$$d(x, \bar{x}) + d(y, \bar{x}) \le 2r \le \frac{2}{\delta}(R' - r) \le \frac{2}{\delta} d(x, y),$$

where the second inequality holds since $r < R$ and $R' = (1+\delta)R$, and the last inequality is the assumption $d(x, y) \ge R' - \max(d(\bar{x}, x), d(\bar{x}, y))$. This completes the proof of (3-18), which immediately implies (3-19) and (3-20).

(iv) Let $f \in C^{2,\alpha}_{X,0}(B(\bar{x}, R))$. For any $x \in B(\bar{x}, R)$, let $\gamma(t)$ be the curve such that

$$\gamma'(t) = (X_0)_{\gamma(t)}, \quad \gamma(0) = x.$$

This $\gamma(t)$ will be defined at least for $t \in [0, r_0]$ where $r_0 > 0$ is a number only depending on $B(\bar{x}, R)$ and X_0. Then, for any $r \in (0, r_0)$, we can write, for some $\theta \in (0, 1)$,

$$f(\gamma(r)) - f(\gamma(0)) = r \frac{d}{dt}[f(\gamma(t))]_{t=\theta r} = r(X_0 f)(\gamma(\theta r)),$$

hence

$$(X_0 f)(x) = (X_0 f)(\gamma(0)) - (X_0 f)(\gamma(\theta r)) + \frac{1}{r}[f(\gamma(r)) - f(\gamma(0))]$$

and since, by definition of γ and d, $d(\gamma(0), \gamma(\theta r)) \le (\theta r)^{1/2}$, we get

$$\begin{aligned} |(X_0 f)(x)| &\le |(X_0 f)(\gamma(0)) - (X_0 f)(\gamma(\theta r))| + \frac{2}{r} \|f\|_{L^\infty} \\ &\le (\theta r)^{\alpha/2} |X_0 f|_{C^\alpha_X(B(\bar{x},R))} + \frac{2}{r} \|f\|_{L^\infty(B(\bar{x},R))} \\ &\le r^{\alpha/2} |X_0 f|_{C^\alpha_X(B(\bar{x},R))} + \frac{2}{r} \|f\|_{L^\infty(B(\bar{x},R))}, \end{aligned}$$

so we are done. □

Next, we are going to study the relation between the spaces $C^\alpha_X(B_R)$ and $C^\alpha_{\widetilde{X}}(\widetilde{B}_R)$.

Proposition 3.15. *Let $\widetilde{B}(\bar{\xi}, R)$ be a lifted ball* (*see the end of Section 3C*), *with $\bar{\xi} = (\bar{x}, 0)$. If f is a function defined in $B(\bar{x}, R)$ and $\tilde{f}(x, h) = f(x)$ is regarded as a function defined on $\widetilde{B}_R(\bar{\xi}, R)$, the following inequalities hold* (*whenever the right-hand side is finite*):

$$\begin{aligned} &|\tilde{f}|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R))} \le |f|_{C^\alpha_X(B(\bar{x},R))}, \\ &|f|_{C^\alpha_X(B(\bar{x},s))} \le \frac{c}{(t-s)^2} |\tilde{f}|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},t))} \quad \textit{for } 0 < s < t < R, \end{aligned} \tag{3-26}$$

where c also depends on R. Moreover,

$$|\widetilde{X}_{i_1}\widetilde{X}_{i_2}\cdots\widetilde{X}_{i_k}\tilde{f}|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R))} \le |X_{i_1}X_{i_2}\cdots X_{i_k}f|_{C^\alpha_X(B(\bar{x},R))}, \tag{3-27}$$

$$|X_{i_1}X_{i_2}\cdots X_{i_k}f|_{C^\alpha_X(B(\bar{x},s))} \le \frac{c}{(t-s)^2}|\widetilde{X}_{i_1}\widetilde{X}_{i_2}\cdots\widetilde{X}_{i_k}\tilde{f}|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},t))} \tag{3-28}$$

for $0 < s < t < R$ and $i_j = 0, 1, 2, \dots, q$.

As already done in [Bramanti and Brandolini 2007, Proposition 8.3], to prove the above relation between Hölder spaces over B and $\widetilde{B}$ we have to exploit an equivalent integral characterization of Hölder continuous functions, analogous to the one established in the classical case by Campanato [1963]. However, to avoid integration over sets of the kind $\Omega \cap B(x, r)$ (with the related problem of assuring a suitable doubling condition), we need to apply the local version of this result which has been established in [Bramanti and Zhu 2012].

Definition 3.16. For $\bar{x} \in \Omega'$, $B(\bar{x}, R) \subset \Omega$, $f \in L^1(B(\bar{x}, R))$, $\alpha \in (0, 1)$, and $0 < s < t \le 1$, let

$$M_{\alpha,B_{sR},B_{tR}}(f) = \sup_{x\in B(\bar{x},sR),\, r\le(t-s)R}\ \inf_{c\in\mathbb{R}} \frac{1}{r^\alpha|B_r(x)|}\int_{B_r(x)} |f(y)-c|\,dy.$$

If $f \in C^\alpha_X(B(\bar{x}, R))$, then

$$M_{\alpha,B_{sR},B_{tR}}(f) \le |f|_{C^\alpha(B_R(x_0))}.$$

Moreover, we get the following.

Lemma 3.17. *For $\bar{x} \in \Omega'$, $B(\bar{x}, 2R_0) \subset \Omega$, $R < R_0$, $\alpha \in (0, 1)$, and $0 < s < t \le 1$, if $f \in L^1(B(\bar{x}, tR))$ is a function such that $M_{\alpha,B_{sR},B_{tR}}(f) < \infty$, then there exists a function f^*, a.e. equal to f, such that $f^* \in C^\alpha_X(B(\bar{x}, sR))$ and*

$$|f^*|_{C^\alpha_X(B(\bar{x},sR))} \le \frac{c}{(t-s)^2} M_{\alpha,B_{sR},B_{tR}}(f)$$

for some c independent of f, s, t.

Proof. We can apply [Bramanti and Zhu 2012, Theorem 9.2] choosing $\Omega_k = B(\bar{x}, sR)$, $\Omega_{k+1} = B(\bar{x}, tR)$, $\varepsilon_n = R(t - s)$. The locally doubling constant can be chosen independently of R, since $B(\bar{x}, 2R_0) \subset \Omega$, $R < R_0$. We conclude that there exists a function f^*, a.e. equal to f, such that

$$|f^*(x) - f^*(y)| \le cM_{\alpha,B_{sR},B_{tR}}(f)d_X(x, y)^\alpha$$

for any $x, y \in B(\bar{x}, sR)$ with $d_X(x, y) \le R(t - s)/2$.

Now if x, y are any two points in $B_{sR}(x_0)$, and $r = d_X(x, y)$, by Lemma 3.8 we can find $n + 1$ points $x_0 = x, x_1, x_2, \dots, x_n = y$ in $B_{sR}(x_0)$ such that

$$d_X(x_i, x_{i-1}) \le \frac{2r}{\sqrt{n}}.$$

Let n be the least integer such that $2r/\sqrt{n} \le R(t-s)/2$. Then

$$|f^*(x) - f^*(y)| \le \sum_{i=1}^{n} |f^*(x_i) - f^*(x_{i-1})| \le \sum_{i=1}^{n} c M_{\alpha, B_{sR}, B_{tR}}(f) d_X(x_i, x_{i-1})^\alpha$$
$$\le n c M_{\alpha, B_{sR}, B_{tR}}(f) d_X(x, y)^\alpha.$$

Let us find an upper bound on n. We know that

$$\sqrt{n} \le c \frac{d_X(x, y)}{R(t-s)} \le \frac{c}{t-s},$$

since $d_X(x, y) \le 2R$ for $x, y \in B_{tR}(x_0)$. Hence $n \le c/(t-s)^2$ and the lemma is proved. □

Proof of Proposition 3.15. The first inequality immediately follows by (3-16). To prove the second one, let $0 < s < t < 1$ and $x \in B(\bar{x}, \delta_0 s R)$, where δ_0 is the number in Theorem 3.11, $r \le R(t-s)$, $\bar{\xi} = (\bar{x}, 0)$. Since the projection $\pi : \widetilde{B}((x, s), \delta) \to B(x, \delta)$ is onto (see Theorem 3.11), there exists $h \in \mathbb{R}^{N-n}$ such that $\xi = (x, h) \in \widetilde{B}(\bar{\xi}, \delta_0 s R)$. Then, by Corollary 3.12, we have

$$\frac{1}{r^\alpha} \frac{c}{|B_{\delta_0 r}(x)|} \int_{B_{\delta_0 r}(x)} |f(y) - k| \, dy \le \frac{c}{r^\alpha} \frac{1}{|\widetilde{B}(\xi, r)|} \int_{\widetilde{B}(\xi, r)} |\tilde{f}(\eta) - k| \, d\eta; \tag{3-29}$$

choosing $k = f(x) = \tilde{f}(\xi)$, the latter quantity is

$$\le \frac{c}{r^\alpha} |\tilde{f}|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\xi, r))} r^\alpha = c |\tilde{f}|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\xi, r))}.$$

Since $r \le R(t-s)$ and $d(\xi, \bar{\xi}) < \delta_0 s R$, we have the inclusion

$$\widetilde{B}(\xi, r) \subset \widetilde{B}(\bar{\xi}, \delta_0 s R + R(t-s)) =: \widetilde{B}(\bar{\xi}, R')$$

so that (3-29) implies

$$M_{\alpha, B(\bar{x}, \delta_0 s R), B(\bar{x}, \delta_0 t R)}(f) \le c |\tilde{f}|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi}, R'))},$$

and, by Lemma 3.17, we conclude

$$|f^*|_{C^\alpha_X(B(\bar{x}, \delta_0 s R))} \le \frac{c}{(t-s)^2} |\tilde{f}|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi}, R'))}.$$

Note that $R' - \delta_0 s R = R(t-s)$, hence, changing our notation to

$$\delta_0 s R = s', \quad R' = t',$$

we get

$$|f^*|_{C^\alpha_X(B(\bar{x}, s'))} \le \frac{c}{(t'-s')^2} |\tilde{f}|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi}, t'))}$$

for $0 < s' < t' < R$, with c also depending on R. This is (3-26).

Now inequalities (3-27) and (3-28) also follow, because $\widetilde{X}_i \tilde{f} = \widetilde{X_i f}$, hence the same reasoning can be iterated to higher order derivatives. □

3D.2. *Sobolev spaces.*

Definition 3.18. If $X = (X_0, X_1, \dots, X_q)$ is any system of smooth vector fields satisfying Hörmander's condition in a domain $\Omega \subset \mathbb{R}^n$, the Sobolev space $S_X^{2,p}(\Omega)$ $(1 < p < \infty)$ consists of L^p-functions with 2 (weighted) derivatives with respect to the vector fields X_i, in L^p. Explicitly,

$$\|u\|_{S_X^{2,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \sum_{i=1}^{2} \|D^i u\|_{L^p(\Omega)},$$

$$\text{where } \|D^1 u\|_{L^p(\Omega)} = \sum_{i=1}^{q} \|X_i u\|_{L^p(\Omega)};\ \|D^2 u\|_{L^p(\Omega)} = \|X_0 u\|_{L^p(\Omega)} + \sum_{i,j=1}^{q} \|X_i X_j u\|_{L^p(\Omega)}.$$

Also, we can define the spaces of functions vanishing at the boundary saying that $u \in S_{0,X}^{2,p}(\Omega)$ if there exists a sequence $\{u_k\}$ of $C_0^\infty(\Omega)$ functions converging to u in $S_X^{2,p}(\Omega)$. Similarly, we can define the Sobolev spaces $S_{\widetilde{X}}^{2,p}(\widetilde{B})$, $S_{\widetilde{X},0}^{2,p}(\widetilde{B})$ over a lifted ball $\widetilde{B}$, induced by the $\widetilde{X}$.

The following has been proved [Bramanti and Brandolini 2000a, Proposition 3.5].

Proposition 3.19. *If $u \in S_X^{2,p}(\Omega)$ and $\varphi \in C_0^\infty(\Omega)$, then $u\varphi \in S_{0,X}^{2,p}(\Omega)$, and an analogous property holds for the space $S_{0,\widetilde{X}}^{2,p}(\widetilde{B})$.*

Moreover, we have the following.

Theorem 3.20. *Let $f \in L^p(B(x,r)$, $\tilde{f}(x,h) = f(x)$, and $\widetilde{B}(\xi, r)$ be the lifted ball of $B(x,r)$, with $\xi = (x, 0) \in \mathbb{R}^N$. Then*

$$c_1 \|f\|_{L^p(B(x,\delta_0 r))} \le \|\tilde{f}\|_{L^p(\widetilde{B}(\xi,r))} \le c_2 \|f\|_{L^p(B(x,r))},$$
$$c_1 \|f\|_{S_X^{2,p}(B(x,\delta_0 r))} \le \|\tilde{f}\|_{S_{\widetilde{X}}^{2,p}(\widetilde{B}(\xi,r))} \le c_2 \|f\|_{S_X^{2,p}(B(x,r))},$$

where $\delta_0 < 1$ is the number appearing in Theorem 3.11.

Proof. The first inequality follows by Theorem 3.11; the second follows by the first, since

$$\widetilde{X}_i \tilde{f} = X_i \tilde{f} = \widetilde{(X_i f)}. \qquad \square$$

3D.3. *Vanishing mean oscillation.* Let us recall the following abstract definition.

Definition 3.21 [Bramanti and Zhu 2012, Definition 6.1]. Let $(\Omega, \{\Omega_n\}_{n=1}^\infty, d, \mu)$ be a metric locally homogeneous space (see Section 3C). For any function $u \in L^1(\Omega_{n+1})$ and $r > 0$ with $r \le \varepsilon_n$, set

$$\eta^*_{u,\Omega_n,\Omega_{n+1}}(r) = \sup_{t \le r} \sup_{x_0 \in \Omega_n} \frac{1}{\mu(B(x_0,t))} \int_{B(x_0,t)} |u(x) - u_B| \, d\mu(x),$$

where $u_B = \mu(B(x_0,t))^{-1} \int_{B(x_0,t)} u$. We say that $u \in \mathrm{BMO}_{\mathrm{loc}}(\Omega_n, \Omega_{n+1})$ if

$$\|u\|_{\mathrm{BMO}_{\mathrm{loc}}(\Omega_n,\Omega_{n+1})} = \sup_{r \le \varepsilon_n} \eta^*_{u,\Omega_n,\Omega_{n+1}}(r) < \infty.$$

We say that $u \in \mathrm{VMO}_{\mathrm{loc}}(\Omega_n, \Omega_{n+1})$ if $u \in \mathrm{BMO}_{\mathrm{loc}}(\Omega_n, \Omega_{n+1})$ and

$$\eta^*_{u,\Omega_n,\Omega_{n+1}}(r) \to 0 \quad \text{as } r \to 0.$$

The function $\eta^*_{u,\Omega_n,\Omega_{n+1}}$ will be called the VMO local modulus of u in (Ω_n, Ω_{n+1}).

We need to specialize this definition to our concrete situation. First, let us endow our domain Ω with the structure

$$(\Omega, \{\Omega_k\}_k, d_X, dx)$$

of locally homogeneous space described at the end of Section 3C. Then:

Definition 3.22 (local VMO). We say that $a \in \mathrm{VMO}_{X,\mathrm{loc}}(\Omega)$ if

$$a \in \mathrm{VMO}_{\mathrm{loc}}(\Omega_k, \Omega_{k+1}) \quad \text{for every } k.$$

More explicitly, this means that, for any fixed $\Omega' \Subset \Omega$, the function

$$\eta^*_{u,\Omega',\Omega}(r) = \sup_{t \le r} \sup_{x_0 \in \Omega'} \frac{1}{|B_t(x_0)|} \int_{B_t(x_0)} |u(x) - u_{B_t(x_0)}|\, dx,$$

is finite for $r \le r_0$ and vanishes for $r \to 0$, where r_0 is the number such that the local doubling condition of Theorem 3.7 holds:

$$|B(x, 2r)| \le c|B(x, r)| \quad \text{for any } x \in \Omega',\ r \le r_0.$$

As for Hölder continuous and Sobolev functions, we need a comparison result for VMO functions in the original variables and the lifted ones. By Corollary 3.12 we immediately have the following.

Proposition 3.23. *Let $a \in \mathrm{VMO}_{X,\mathrm{loc}}(\Omega)$. Then, for any $\Omega' \Subset \Omega$, $x_0 \in \Omega'$, $B(x_0, R)$, and $\widetilde{\Omega}_k = \widetilde{B}(\xi_0, kR/(k+1))$ as before, we have that $\tilde{a}(x, h) = a(x)$ belongs to the class $\mathrm{VMO}_{\mathrm{loc}}(\widetilde{\Omega}_k, \widetilde{\Omega}_k)$ for every k, with*

$$\eta^*_{\tilde{a},\widetilde{\Omega}_k,\widetilde{\Omega}_{k+1}}(r) \le c\eta^*_{a,\Omega',\Omega}(r).$$

In other words, the $\mathrm{VMO}_{\mathrm{loc}}$ modulus of the original function a controls the $\mathrm{VMO}_{\mathrm{loc}}$ modulus of its lifted version.

4. Operators of type λ and representation formulas

4A. ***Differential operators and fundamental solutions.*** We now define various differential operators that we will handle in the following. Our main interest is to study the operator

$$\mathcal{L} = \sum_{i,j=1}^{q} a_{ij}(x) X_i X_j + X_0,$$

under the assumption (H) in Section 2. Recall that, in view of Remark 2.3, we have set $a_0(x) \equiv 1$.

For any $\bar{x} \in \Omega$, we can apply the "lifting theorem" to the vector fields X_i (see Section 3A for the statement and notation), obtaining new vector fields $\widetilde{X}_i$ which are free up to weight s and satisfy

Hörmander's condition of weight s in a neighborhood of $\bar{\xi} = (\bar{x}, 0) \in \mathbb{R}^N$. For $\xi = (x, t) \in \widetilde{B}(\bar{\xi}, R)$, with $\widetilde{B}(\bar{\xi}, R)$ as in Lemma 3.9, set

$$\tilde{a}_{ij}(x, t) = a_{ij}(x),$$

and let

$$\widetilde{\mathscr{L}} = \sum_{i,j=1}^{q} \tilde{a}_{ij}(\xi) \widetilde{X}_i \widetilde{X}_j + \widetilde{X}_0 \tag{4-1}$$

be the lifted operator, defined in $\widetilde{B}(\bar{\xi}, R)$. Next, we freeze $\widetilde{\mathscr{L}}$ at some point $\xi_0 \in \widetilde{B}(\bar{\xi}, R)$, and consider the frozen lifted operator

$$\widetilde{\mathscr{L}}_0 = \sum_{i,j=1}^{q} \tilde{a}_{ij}(\xi_0) \widetilde{X}_i \widetilde{X}_j + \widetilde{X}_0. \tag{4-2}$$

To study $\widetilde{\mathscr{L}}_0$, in view of the "approximation theorem" (Theorem 3.2), we will consider the approximating operator, defined on the homogeneous group $\mathbb{G}$,

$$\mathscr{L}_0^* = \sum_{i,j=1}^{q} \tilde{a}_{ij}(\xi_0) Y_i Y_j + Y_0,$$

and its transpose,

$$\mathscr{L}_0^{*T} = \sum_{i,j=1}^{q} \tilde{a}_{ij}(\xi_0) Y_i Y_j - Y_0,$$

where $\{Y_i\}$ are the left invariant vector fields on the group $\mathbb{G}$ defined in Section 3A.

We will apply to $\mathscr{L}_0^*$ and $\mathscr{L}_0^{*T}$ several results proved in [Bramanti and Brandolini 2000b], which in turn are based on [Folland 1975, Theorem 2.1 and Corollary 2.8; Folland and Stein 1974, Proposition 8.5]. They are collected in the following theorem.

Theorem 4.1. *Assume that the homogeneous dimension of* $\mathbb{G}$ *is* $Q \geq 3$. *For every* $\xi_0 \in \widetilde{B}(\bar{\xi}, R)$, *the operator* $\mathscr{L}_0^*$ *has a unique fundamental solution* $\Gamma(\xi_0; \cdot)$ *such that*

(a) $\Gamma(\xi_0; \cdot) \in C^\infty(\mathbb{R}^N \setminus \{0\})$;

(b) $\Gamma(\xi_0; \cdot)$ *is homogeneous of degree* $(2 - Q)$;

(c) *for every test function* f *and every* $v \in \mathbb{R}^N$,

$$f(v) = \int_{\mathbb{R}^N} \Gamma(\xi_0; u^{-1} \circ v) \mathscr{L}_0^* f(u)\, du;$$

moreover, for every $i, j = 1, \ldots, q$, *there exist constants* $\alpha_{ij}(\xi_0)$ *such that*

$$Y_i Y_j f(v) = \mathrm{PV} \int_{\mathbb{R}^N} Y_i Y_j \Gamma(\xi_0; u^{-1} \circ v) \mathscr{L}_0^* f(u)\, du + \alpha_{ij}(\xi_0) \cdot \mathscr{L}_0^* f(v); \tag{4-3}$$

(d) $Y_i Y_j \Gamma(\xi_0; \cdot)$ *is homogeneous of degree* $-Q$;

(e) *for every* $R > r > 0$,

$$\int_{r<\|u\|<R} Y_i Y_j \Gamma(\xi_0; u)\, du = \int_{\|u\|=1} Y_i Y_j \Gamma(\xi_0; u)\, d\sigma(u) = 0.$$

In (4-3) the notation PV $\int_{\mathbb{R}^N} \cdots du$ stands for $\lim_{\varepsilon\to 0} \int_{\|u^{-1}\circ v\|>\varepsilon} \cdots du$.

Remark 4.2. By [Folland 1975, remark on p. 174], we know that the fundamental solution of the transposed operator $\mathscr{L}_0^{*T}$ is

$$\Gamma^T(\xi_0; u) = \Gamma(\xi_0; u^{-1}) = \Gamma(\xi_0; -u).$$

(However, beware that $Y_i \Gamma^T(\xi_0; u) \neq \pm Y_i \Gamma(\xi_0; -u)$.)

Throughout the following, we will set, for $i, j = 1, \dots, q$,

$$\Gamma_{ij}(\xi_0; u) = Y_i Y_j [\Gamma(\xi_0; \cdot)](u),$$
$$\Gamma_{ij}^T(\xi_0; u) = Y_i Y_j [\Gamma^T(\xi_0; \cdot)](u).$$

A second fundamental result we need contains a bound on the derivatives of Γ, uniform with respect to ξ_0.

Theorem 4.3 [Bramanti and Brandolini 2000b, Theorem 12]. *For every multi-index β, there exists a constant $c = c(\beta, \mathbb{G}, \mu)$ such that, for any $i, j = 1, \dots, q$,*

$$\sup_{\substack{\xi\in\widetilde{B}(\bar{\xi},R)\\ \|u\|=1}} \left| \left(\frac{\partial}{\partial u}\right)^{\beta} \Gamma_{ij}(\xi; u) \right| \leq c;$$

moreover, for the α_{ij} appearing in (4-3), *the uniform bound*

$$\sup_{\xi\in\widetilde{B}(\bar{\xi},R)} |\alpha_{ij}(\xi)| \leq c_2$$

holds for some constant $c_2 = c_2(\mathbb{G}, \mu)$.

Remark 4.4. Theorems 4.1 and 4.3 still hold replacing Γ by Γ^T and Γ_{ij} by Γ_{ij}^T.

4B. ***Operators of type λ.*** As in [Rothschild and Stein 1976; Bramanti and Brandolini 2000a], we are going to build a parametrix for $\widetilde{\mathscr{L}}$ shaped on the homogeneous fundamental solution of $\mathscr{L}_0^*$. More generally, we need to define a class of integral operators with different degrees of singularity. The next definition is adapted from [Bramanti and Brandolini 2000a], the difference being the necessity, in the present case, to consider integral kernels shaped on the fundamental solutions of both $\mathscr{L}_0^*$ and $\mathscr{L}_0^{*T}$.

Definition 4.5. For any $\xi_0 \in \widetilde{B}(\bar{\xi}, R)$, we say that $k(\xi_0; \xi, \eta)$ is a *frozen kernel of type* λ (over the ball $\widetilde{B}(\bar{\xi}, R)$) for some nonnegative integer λ (we will use $\lambda = 0, 1, 2$) if, for every positive integer m, we can write, for $\xi, \eta \in \widetilde{B}(\bar{\xi}, R)$,

$$
\begin{aligned}
k(\xi_0; \xi, \eta) &= k'(\xi_0; \xi, \eta) + k''(\xi_0; \xi, \eta) \\
&= \left\{ \sum_{i=1}^{H_m} a_i(\xi) b_i(\eta) D_i \Gamma(\xi_0; \cdot) + a_0(\xi) b_0(\eta) D_0 \Gamma(\xi_0; \cdot) \right\} (\Theta(\eta, \xi)) \\
&\qquad + \left\{ \sum_{i=1}^{H_m} a_i'(\xi) b_i'(\eta) D_i' \Gamma^T(\xi_0; \cdot) + a_0'(\xi) b_0'(\eta) D_0' \Gamma^T(\xi_0; \cdot) \right\} (\Theta(\eta, \xi)),
\end{aligned}
$$

where $a_i, b_i, a_i', b_i' \in C_0^\infty(\widetilde{B}(\bar{\xi}, R))$ $(i = 0, 1, \dots, H_m)$, and D_i and D_i' are differential operators such that, for $i = 1, \dots, H_m$, D_i and D_i' are homogeneous of degree $\leq 2 - \lambda$ (so that $D_i \Gamma(\xi_0; \cdot)$, and $D_i' \Gamma^T(\xi_0; \cdot)$ are homogeneous functions of degree $\geq \lambda - Q$); D_0 and D_0' are differential operators such that $D_0 \Gamma(\xi_0; \cdot)$ and $D_0' \Gamma^T(\xi_0; \cdot)$ have m (weighted) derivatives with respect to the vector fields Y_i $(i = 0, 1, \dots, q)$. Moreover, the coefficients of the differential operators D_i, D_i' for $i = 0, 1, \dots, H_m$ possibly depend also on the variables ξ, η, in such a way that the joint dependence on (ξ, η, u) is smooth.

In order to simplify notation, we will not always express explicitly this dependence of the coefficients of D_i on ξ, η. Only if necessary will we write, for instance, $a_i(\xi) b_i(\eta) D_i^{\xi,\eta} \Gamma(\xi_0; \Theta(\eta, \xi))$ to recall this dependence.

Remark 4.6. Note that if a smooth function $c(\xi, \eta, u)$ is $D(\lambda)$-homogeneous of some degree β with respect to u, any ξ or η derivative of c has the same homogeneity with respect to u, since

$$
c(\xi, \eta, D(\lambda)u) = \lambda^\beta c(\xi, \eta, u) \quad \text{implies } \frac{\partial c}{\partial \xi_i}(\xi, \eta, D(\lambda)u) = \lambda^\beta \frac{\partial c}{\partial \xi_i}(\xi, \eta, u).
$$

Hence any derivative

$$
\left(\frac{\partial}{\partial \xi_i} D_i^{\xi,\eta} \right) \Gamma(\xi_0; \cdot), \qquad \left(\frac{\partial}{\partial \eta_i} D_i^{\xi,\eta} \right) \Gamma(\xi_0; \cdot)
$$

has the same homogeneity as

$$
D_i^{\xi,\eta} \Gamma(\xi_0; \cdot).
$$

Here and in the following, the symbol $((\partial/\partial \xi_i) D_i^{\xi,\eta}) f$ means that we have taken the ξ_i-derivative of the coefficients of the differential operator $D_i^{\xi,\eta}$, which acts on the u variables but contains ξ, η as parameters; the resulting differential operator acts on the function $f(u)$.

Definition 4.7. For any $\xi_0 \in \widetilde{B}(\bar{\xi}, R)$, we say that $T(\xi_0)$ is a *frozen operator of type* $\lambda \geq 1$ (over the ball $\widetilde{B}(\bar{\xi}, R)$) if $k(\xi_0; \xi, \eta)$ is a frozen kernel of type λ and

$$
T(\xi_0) f(\xi) = \int_{\widetilde{B}} k(\xi_0; \xi, \eta) f(\eta) \, d\eta
$$

for $f \in C_0^\infty(\widetilde{B}(\bar{\xi}, R))$. We say that $T(\xi_0)$ is a *frozen operator of type* 0 if $k(\xi_0; \xi, \eta)$ is a frozen kernel of type 0 and

$$
T(\xi_0) f(\xi) = \mathrm{PV} \int_{\widetilde{B}} k(\xi_0; \xi, \eta) f(\eta) \, d\eta + \alpha(\xi_0, \xi) f(\xi),
$$

where α is a bounded measurable function, smooth in ξ, and the principal value integral exists. Explicitly, this principal value is defined by

$$\mathrm{PV}\int_{\widetilde{B}} k(\xi_0;\xi,\eta)f(\eta),d\eta = \lim_{\varepsilon\to 0}\int_{\|\Theta(\eta,\xi)\|>\varepsilon} k(\xi_0;\xi,\eta)f(\eta)\,d\eta.$$

Definition 4.8. If $k(\xi_0;\xi,\eta)$ is a frozen kernel of type $\lambda \geq 0$, we say that $k(\xi;\xi,\eta)$ is a *variable kernel of type* λ (over the ball $\widetilde{B}(\bar{\xi},R)$), and

$$Tf(\xi) = \int_{\widetilde{B}} k(\xi;\xi,\eta)f(\eta)\,d\eta$$

is a *variable operator of type* λ. If $\lambda = 0$, the integral must be taken in principal value sense and a term $\alpha(\xi,\xi)f(\xi)$ must be added.

In reference to Definition 4.5, we will call the k' and k'' parts of k “the frozen kernels of type λ modeled on Γ and Γ^T”, respectively. Analogously we will sometimes speak of frozen operators of type λ modeled on Γ or Γ^T, to denote that the kernel has this special form.

A common operation on frozen operators is *transposition*.

Definition 4.9. If $T(\xi_0)$ is a frozen operator of type $\lambda \geq 0$ over $\widetilde{B}(\bar{\xi},R)$, we will denote by $T(\xi_0)^T$ the transposed operator, formally defined by

$$\int_{\widetilde{B}} f(\xi)T(\xi_0)^T g(\xi)\,d\xi = \int_{\widetilde{B}} g(\xi)T(\xi_0)f(\xi)\,d\xi$$

for any $f, g \in C_0^\infty(\widetilde{B}(\bar{\xi},R))$.

Clearly, if $k(\xi_0,\xi,\eta)$ is the kernel of $T(\xi_0)$, then $k(\xi_0,\eta,\xi)$ is the kernel of $T(\xi_0)^T$. It is useful to note the following.

Proposition 4.10. *If $T(\xi_0)$ is a frozen operator of type $\lambda \geq 0$ over $\widetilde{B}(\bar{\xi},R)$, modeled on Γ or Γ^T, then $T(\xi_0)^T$ is a frozen operator of type λ, modeled on Γ^T or Γ, respectively. In particular, the transpose of a frozen operator of type λ is still a frozen operator of type λ.*

Proof. Let D be any differential operator on the group $\mathbb{G}$. For any $f \in C_0^\infty(\widetilde{B}(\bar{\xi},R))$, let $f'(u) = f(-u)$. Let D' be the differential operator defined by the identity

$$D'f = (D(f'))'.$$

Clearly, if D is homogeneous of some degree β, the same is true for D'; if $D\Gamma(\xi_0;\cdot)$ or $D\Gamma^T(\xi_0;\cdot)$ has m (weighted) derivatives with respect to the vector fields Y_i $(i = 0, 1, \ldots, q)$, the same is true for $D'\Gamma(\xi_0;\cdot)$ or $D'\Gamma^T(\xi_0;\cdot)$. Also, recalling that $\Gamma^T(\xi_0;u) = \Gamma(\xi_0;-u)$, we have

$$(D'\Gamma)(u) = (D\Gamma^T)(-u) \quad \text{and} \quad (D'\Gamma^T)(u) = (D\Gamma)(-u).$$

Moreover, these identities can be iterated, for instance,

$$(D_1D_2\Gamma)(-u) = (D_1(D_2\Gamma))(-u) = (D_1'(D_2\Gamma)')(u) = (D_1'D_2'\Gamma^T)(u).$$

Therefore, if

$$k'(\xi_0, \xi, \eta) = \left\{ \sum_{i=1}^{H_m} a_i(\xi) b_i(\eta) D_i \Gamma(\xi_0; \cdot) + a_0(\xi) b_0(\eta) D_0 \Gamma(\xi_0; \cdot) \right\} (\Theta(\eta, \xi))$$

is a frozen kernel of type λ modeled on Γ, then

$$\begin{aligned} k'(\xi_0, \eta, \xi) &= \left\{ \sum_{i=1}^{H_m} a_i(\eta) b_i(\xi) D_i \Gamma(\xi_0; \cdot) + a_0(\xi) b_0(\eta) D_0 \Gamma(\xi_0; \cdot) \right\} (-\Theta(\eta, \xi)) \\ &= \left\{ \sum_{i=1}^{H_m} a_i(\eta) b_i(\xi) D_i' \Gamma^T(\xi_0; \cdot) + a_0(\xi) b_0(\eta) D_0' \Gamma^T(\xi_0; \cdot) \right\} (\Theta(\eta, \xi)) \end{aligned}$$

is a frozen kernel of type λ modeled on Γ^T. Analogously one can prove the converse. □

We now have to deal with the relations between operators of type λ and the differential operators represented by the vector fields $\widetilde{X}_i$. This is a study which was carried out in [Rothschild and Stein 1976, Section 14] and adapted to nonvariational operators in [Bramanti and Brandolini 2000a]. We are interested in two main results. Roughly speaking, the first says that the composition, in any order, of an operator of type λ with the $\widetilde{X}_i$ or $\widetilde{X}_0$ derivative is an operator of type $\lambda - 1$ or $\lambda - 2$, respectively. The second says that the $\widetilde{X}_i$ derivative of an operator of type λ can be rewritten as the sum of other operators of type λ, each acting on a different $\widetilde{X}_j$ derivative, plus a suitable remainder. In [Rothschild and Stein 1976] these results are proved only for a system of Hörmander vector fields of weight one (that is, without the drift), and several arguments are very condensed. Hence we need to extend and modify some arguments in [Rothschild and Stein 1976, Section 14] to cover the present situation. Moreover, as in [Bramanti and Brandolini 2000a], we need to keep under careful control the dependence of any quantity on the frozen point ξ_0 appearing in $\Gamma(\xi_0, \cdot)$. For these and other technical reasons, we prefer to write complete proofs of these properties. The first result is the following.

Theorem 4.11 [Rothschild and Stein 1976, Theorem 8]. *Suppose $T(\xi_0)$ is a frozen operator of type $\lambda \geq 1$. Then $\widetilde{X}_k T(\xi_0)$ and $T(\xi_0)\widetilde{X}_k$ ($k = 1, 2, \dots, q$) are operators of type $\lambda - 1$. If $\lambda \geq 2$, then $\widetilde{X}_0 T(\xi_0)$ and $T(\xi_0)\widetilde{X}_0$ are operators of type $\lambda - 2$.*

To prove this, we begin by stating the following two lemmas.

Lemma 4.12. *If $k(\xi_0; \xi, \eta)$ is a frozen kernel of type $\lambda \geq 1$ over $\widetilde{B}(\bar{\xi}, R)$, then $(\widetilde{X}_j k)(\xi_0; \cdot, \eta)(\xi)$ ($j = 1, 2, \dots, q$) is a frozen kernel of type $\lambda - 1$. If $\lambda \geq 2$, then $(\widetilde{X}_0 k)(\xi_0; \cdot, \eta)(\xi)$ is a frozen kernel of type $\lambda - 2$.*

Proof. This basically follows by the definition of kernel of type λ and Theorem 3.2. When the $\widetilde{X}_j$ derivative acts on the ξ variable of a kernel $D_i^{\xi} \Gamma(\xi_0, \cdot)$, one also has to take into account Remark 4.6.

Here we just want to point out the following fact. The prototype of a frozen kernel of type 2 is the function

$$a(\xi) \Gamma(\xi_0; \Theta(\eta, \xi)) b(\eta).$$

Note that the computation

$$\widetilde{X}_i[a(\cdot)\Gamma(\xi_0;\Theta(\eta,\cdot))b(\eta)](\xi) = a(\xi)[(Y_i + R_i^\eta)\Gamma(\xi_0;\cdot)](\Theta(\eta,\xi))b(\eta) + (\widetilde{X}_i a)(\xi)\Gamma(\xi_0;\Theta(\eta,\xi))b(\eta)$$

in particular generates the term

$$a(\xi)(R_i^\eta\Gamma)(\xi_0;\cdot)(\Theta(\eta,\xi))b(\eta),$$

where the differential operator R_i^η has coefficients depending on η. In the proof of Theorem 4.11 we will see another basic computation on frozen kernels which generates differential operators with coefficients also depending on ξ. This is the reason why Definition 4.5 allows for this kind of dependence. □

Lemma 4.13. *If $T(\xi_0)$ is a frozen operator of type $\lambda \geq 1$ over $\widetilde{B}(\bar{\xi}, R)$, then $\widetilde{X}_i T(\xi_0)$ $(i = 1, 2, \ldots, q)$ is a frozen operator of type $\lambda - 1$. If $\lambda \geq 2$, then $\widetilde{X}_0 T(\xi_0)$ is a frozen operator of type $\lambda - 2$.*

Proof. With reference to Definition 4.5, it is enough to consider the part k' of the kernel of T, the proof for k'' being completely analogous. So, let us consider the operator

$$\widetilde{X}_i T(\xi_0) \quad (i = 1, 2, \ldots, q),$$

where $T(\xi_0)$ has kernel k'.

If $\lambda > 1$, the result immediately follows by the previous lemma. If $\lambda = 1$, then

$$T(\xi_0)f(\xi) = \int_{\widetilde{B}(\bar{\xi},R)} a(\xi)b(\eta)D_1\Gamma(\xi_0;\Theta(\eta,\xi))f(\eta)\,d\eta + T'(\xi_0)f(\xi),$$

where $T'(\xi_0)$ is a frozen operator of type 2 and D_1 is a 1-homogeneous differential operator. We already know that $\widetilde{X}_i T'(\xi_0)$ is a frozen operator of type 1, so it remains to show that

$$\widetilde{X}_i \int_{\widetilde{B}(\bar{\xi},R)} a(\xi)b(\eta)D_1\Gamma(\xi_0;(\Theta(\eta,\xi)))f(\eta)\,d\eta$$

is a frozen operator of type 0. To do this, we have to apply a distributional argument, which will be used several times in the following. Let us compute, for any $\omega \in C_0^\infty(\widetilde{B}(\bar{\xi}, R))$,

$$\begin{aligned}
&\int_{\widetilde{B}(\bar{\xi},R)} \widetilde{X}_i^T\omega(\xi) \int_{\widetilde{B}(\bar{\xi},R)} a(\xi)b(\eta)D_1^\xi\Gamma(\xi_0;(\Theta(\eta,\xi)))f(\eta)\,d\eta\,d\xi \\
&\qquad = \lim_{\varepsilon\to 0}\int_{\widetilde{B}(\bar{\xi},R)} \widetilde{X}_i^T\omega(\xi) \int_{\widetilde{B}(\bar{\xi},R)} a(\xi)b(\eta)\varphi_\varepsilon(\Theta(\eta,\xi))D_1^\xi\Gamma(\xi_0;(\Theta(\eta,\xi)))f(\eta)\,d\eta\,d\xi,
\end{aligned}$$

where $\varphi_\varepsilon(u) = \varphi(D(\varepsilon^{-1})u)$ and $\varphi \in C_0^\infty(\mathbb{R}^N)$, $\varphi(u) = 0$ for $\|u\| < 1$, $\varphi(u) = 1$ for $\|u\| > 2$. Here we have written D_1^ξ to recall that the coefficients of the differential operator D_1 also depend (smoothly) on ξ

as a parameter. By Theorem 3.2,

$$\begin{aligned}
&\int_{\widetilde{B}(\bar{\xi},R)} \widetilde{X}_i^T \omega(\xi) \int_{\widetilde{B}(\bar{\xi},R)} a(\xi) b(\eta) \varphi_\varepsilon(\Theta(\eta,\xi)) D_1^{\xi}\Gamma(\xi_0; (\Theta(\eta,\xi))) f(\eta)\, d\eta\, d\xi \\
&\qquad = \int_{\widetilde{B}(\bar{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\bar{\xi},R)} (\widetilde{X}_i^T \omega)(\xi) a(\xi) \varphi_\varepsilon(\Theta(\eta,\xi)) D_1^{\xi}\Gamma(\xi_0; (\Theta(\eta,\xi)))\, d\xi\, d\eta \\
&\qquad = \int_{\widetilde{B}(\bar{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) (\widetilde{X}_i a)(\xi) \varphi_\varepsilon(\Theta(\eta,\xi)) D_1^{\xi}\Gamma(\xi_0; (\Theta(\eta,\xi)))\, d\xi\, d\eta \\
&\qquad\quad + \int_{\widetilde{B}(\bar{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) a(\xi) \varphi_\varepsilon(\Theta(\eta,\xi)) (\widetilde{X}_i D_1^{\xi})\Gamma(\xi_0; (\Theta(\eta,\xi)))\, d\xi\, d\eta \\
&\qquad\quad + \int_{\widetilde{B}(\bar{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) a(\xi) [(Y_i + R_i^{\eta})(\varphi_\varepsilon D_1^{\xi}\Gamma(\xi_0; \cdot))](\Theta(\eta,\xi))\, d\xi\, d\eta \\
&\qquad =: A_\varepsilon + B_\varepsilon + C_\varepsilon. \qquad (4\text{-}4)
\end{aligned}$$

(For the meaning of the symbol $\widetilde{X}_i D_1^{\xi}$ appearing in the term B_ε, see Remark 4.6.) Now, for $\varepsilon \to 0$,

$$\begin{aligned}
A_\varepsilon \to &\int_{\widetilde{B}(\bar{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) (\widetilde{X}_i a)(\xi) D_1\Gamma(\xi_0; (\Theta(\eta,\xi)))\, d\xi\, d\eta \\
&= \int_{\widetilde{B}(\bar{\xi},R)} f(\eta) S_1(\xi_0) \omega(\eta)\, d\eta = \int_{\widetilde{B}(\bar{\xi},R)} \omega(\eta) S_1(\xi_0)^T f(\eta)\, d\eta, \qquad (4\text{-}5)
\end{aligned}$$

where $S_1(\xi_0)$ is a frozen operator of type 1, and $S_1(\xi_0)^T$ is still a frozen operator of type 1, by Proposition 4.10. Next,

$$\begin{aligned}
B_\varepsilon \to &\int_{\widetilde{B}(\bar{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) a(\xi) (\widetilde{X}_i D_1^{\xi})\Gamma(\xi_0; (\Theta(\eta,\xi)))\, d\xi\, d\eta \\
&= \int_{\widetilde{B}(\bar{\xi},R)} f(\eta) S_1'(\xi_0) \omega(\eta)\, d\eta = \int_{\widetilde{B}(\bar{\xi},R)} \omega(\eta) S_1'(\xi_0)^T f(\eta)\, d\eta, \qquad (4\text{-}6)
\end{aligned}$$

where, by Remark 4.6, $S_1'(\xi_0)$ is a frozen operator of type 1, and the same is still true for $S_1'(\xi_0)^T$. Finally,

$$\begin{aligned}
C_\varepsilon = &\int_{\widetilde{B}(\bar{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) a(\xi) [\varphi_\varepsilon Y_i D_1\Gamma(\xi_0; \cdot)](\Theta(\eta,\xi))\, d\xi\, d\eta \\
&+ \int_{\widetilde{B}(\bar{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) a(\xi) [\varphi_\varepsilon R_i^{\eta} D_1\Gamma(\xi_0; \cdot)](\Theta(\eta,\xi))\, d\xi\, d\eta \\
&+ \int_{\widetilde{B}(\bar{\xi},R)} b(\eta) f(\eta) \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) a(\xi) [(Y_i + R_i^{\eta}) \varphi_\varepsilon D_1\Gamma(\xi_0; \cdot)](\Theta(\eta,\xi))\, d\xi\, d\eta \\
=: &\, C_\varepsilon^1 + C_\varepsilon^2 + C_\varepsilon^3. \qquad (4\text{-}7)
\end{aligned}$$

Now

$$C_\varepsilon^1 \to \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) \left\{ \mathrm{PV} \int_{\widetilde{B}(\bar{\xi},R)} a(\xi) Y_i D_1\Gamma(\xi_0; \Theta(\eta,\xi)) b(\eta) f(\eta) d\eta \right\} d\xi = \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) T(\xi_0) f(\xi) d\xi, \quad (4\text{-}8)$$

where $T(\xi_0)$ is a frozen operator of type 0. Note that the principal value exists because the kernel $Y_i D_1 \Gamma(\xi_0; u)$ has a vanishing integral over spherical shells $\{u \in \mathbb{G} : r_1 < \|u\| < r_2\}$ (see Theorem 4.1).

$$C_\varepsilon^2 \to \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) \left\{ \int_{\widetilde{B}(\bar{\xi},R)} a(\xi) R_i^\eta D_1 \Gamma(\xi_0; \Theta(\eta,\xi)) b(\eta) f(\eta)\, d\eta \right\} d\xi = \int_{\widetilde{B}(\bar{\xi},R)} \omega(\xi) S(\xi_0) f(\xi)\, d\xi, \quad (4\text{-}9)$$

where $S(\xi_0)$ is a frozen operator of type 1. To handle C_ε^3, let us perform the change of variables $u = \Theta(\eta,\xi)$, which, by Theorem 3.3, gives

$$C_\varepsilon^3 = \int_{\widetilde{B}(\bar{\xi},R)} (bf)(\eta) \int_{\|u\|<R} (\omega a)(\Theta(\eta,\cdot)^{-1}(u)) [(Y_i + R_i^\eta) \varphi_\varepsilon D_1 \Gamma(\xi_0; \cdot)](u) \cdot c(\eta)(1 + O(\|u\|))\, du\, d\eta.$$

On the other hand, $Y_i \varphi_\varepsilon(u) = (1/\varepsilon) Y_i \varphi(D(1/\varepsilon)u)$, while $R_i^\eta \varphi_\varepsilon(u)$ is uniformly bounded in ε. Hence the change of variables $D(1/\varepsilon)u = v$ gives

$$\begin{aligned}
C_\varepsilon^3 &= \int_{\widetilde{B}(\bar{\xi},R)} (bf)(\eta) \int_{\|v\|<R/\varepsilon} (\omega a)(\Theta(\eta,\cdot)^{-1}(D(\varepsilon)v)) \left[\frac{1}{\varepsilon} Y_i \varphi(v) + O(1) \right] \\
&\qquad\qquad\qquad\qquad \cdot c(\eta) \varepsilon^{1-Q} D_1^\eta \Gamma(\xi_0; v)(1 + O(\varepsilon\|v\|)) \varepsilon^Q\, dv\, d\eta \\
&\to \int_{\widetilde{B}(\bar{\xi},R)} (bcf)(\eta) \int_{\|v\|\le 2} (\omega a)(\Theta(\eta,\cdot)^{-1}(0)) Y_i \varphi(v) D_1^\eta \Gamma(\xi_0; v)\, dv\, d\eta \\
&= \int_{\widetilde{B}(\bar{\xi},R)} (\omega abcf)(\eta) \int_{\|v\|\le 2} Y_i \varphi(v) D_1^\eta \Gamma(\xi_0; v)\, dv\, d\eta \\
&= \int_{\widetilde{B}(\bar{\xi},R)} (\omega abcf)(\eta) \alpha(\xi_0,\eta)\, d\eta, \qquad (4\text{-}10)
\end{aligned}$$

which is the integral of ω times the multiplicative part of a frozen operator of type 0. It is worthwhile (although not logically necessary to prove the theorem) to realize that the quantity $\alpha(\xi_0,\eta)$ appearing in (4-10) actually does not depend on the function φ. Namely, recalling that $Y_i \varphi(v)$ is supported in the spherical shell $1 \le \|v\| \le 2$ with $\varphi(u) = 1$ for $\|u\| = 2$ and $\varphi(u) = 0$ for $\|u\| = 1$, an integration by parts gives

$$\int_{1\le\|v\|\le 2} Y_i \varphi(v) D_1^\eta \Gamma(\xi_0; v)\, dv = -\int_{1\le\|v\|\le 2} \varphi(v) Y_i D_1^\eta \Gamma(\xi_0; v)\, dv + \int_{\|v\|=2} D_1^\eta \Gamma(\xi_0; v) n_i\, d\sigma(v)$$

with $n_i = \sum_{j=1}^N b_{ij}(u) \nu_j$, where $Y_i = \sum_{j=1}^N b_{ij}(u) \partial_{u_j}$ and ν is the outer normal on $\|v\| = 2$. The vanishing property of the kernel $Y_i D_1^\xi \Gamma(\xi_0; \cdot)$ implies that if φ is a radial function, the first integral vanishes. Therefore,

$$\alpha(\xi_0,\eta) = \int_{\|v\|=2} D_1^\eta \Gamma(\xi_0; v) n_i\, d\sigma(v),$$

which also shows that $\alpha(\xi_0,\eta)$ smoothly depends on η and is bounded in ξ_0 (by Theorem 4.3). By (4-4)–(4-6) and (4-8)–(4-10) we have therefore proved that

$$\widetilde{X}_i T(\xi_0) f(\xi) = S_1(\xi_0)^T f(\xi) + S_1'(\xi_0)^T f(\xi) + T(\xi_0) f(\xi) + \alpha(\xi_0,\xi)(abcf)(\xi),$$

which is a frozen operator of type 0. This completes the proof of the first statement. The proof of the fact that if $\lambda \geq 2$, then $\widetilde{X}_0 T(\xi_0)$ is a frozen operator of type $\lambda - 2$ is completely analogous. □

The above two lemmas imply the assertion on $\widetilde{X}_k T(\xi_0)$ and $\widetilde{X}_0 T(\xi_0)$ in Theorem 4.11. To prove the assertions about $T(\xi_0)\widetilde{X}_k$ and $T(\xi_0)\widetilde{X}_0$ we need a way to express ξ-derivatives of the integral kernel in terms of η-derivatives of the kernel, in order to integrate by parts. This will involve the use of *right invariant vector fields* on the group $\mathbb{G}$: throughout the following, we will denote by

$$Y_{i,k}^R$$

the right invariant vector field on $\mathbb{G}$ satisfying $Y_{i,k}^R f(0) = Y_{i,k} f(0)$.

Lemma 4.14. *For any $f \in C_0^\infty(\mathbb{G})$ and η, ξ in a neighborhood of ξ_0, we can write, for any $i = 1, 2, \dots, s$, $k = 1, 2, \dots, k_i$ (recall s is the step of the Lie algebra),*

$$\widetilde{X}_{i,k}[f(\Theta(\cdot\,, \xi))](\eta) = -(Y_{i,k}^R f)(\Theta(\eta, \xi)) + ((R_{i,k}^\xi)' f)(\Theta(\eta, \xi)), \tag{4-11}$$

where $(R_{i,k}^\xi)'$ is a vector field of local degree $\leq i - 1$ smoothly depending on ξ.

Proof. We start with the following.

Claim. For any function f defined on $\mathbb{G}$, let

$$f'(u) = f(-u)$$

(recall that $-u = u^{-1}$); then the following identities hold:

$$Y_{i,k}(f') = -(Y_{i,k}^R f)'. \tag{4-12}$$

Proof. Let us define the vector fields $\widehat{Y}_{i,k}$ by

$$Y_{i,k}(f') = -(\widehat{Y}_{i,k} f)'. \tag{4-13}$$

Then, for any $a \in \mathbb{G}$, denoting by L_a, R_a the corresponding operators of left and right translation, respectively (acting on functions), we have

$$(\widehat{Y}_{i,k} R_a f)' = -Y_{i,k}((R_a f)') = -Y_{i,k}(L_{-a} f') = -L_{-a} Y_{i,k} f' = L_{-a}(-Y_{i,k} f') = L_{-a}(\widehat{Y}_{i,k} f)' = (R_a \widehat{Y}_{i,k} f)',$$

hence $\widehat{Y}_{i,k}$ are right invariant vector fields. Also, note that, for any vector field $Y = \sum a_j(u)\partial_{u_j}$, we have

$$Y(f')(0) = -(Yf)(0),$$

because

$$\begin{aligned} Y(f')(u) &= \sum a_j(u)\partial_{u_j}[f(-u)] = -\sum a_j(u)(\partial_{u_j} f)(-u) \text{ implies} \\ Y(f')(0) &= -\sum a_j(0)(\partial_{u_j} f)(0) = -(Yf)(0). \end{aligned}$$

Hence, by (4-13), we know that $\widehat{Y}_k f(0) = Y_k f(0)$. Therefore $\widehat{Y}_k$ is the right invariant vector field which coincides with Y_k at the origin, that is, $\widehat{Y}_k = Y_k^R$. □

By (3-4) and (4-12),

$$
\begin{aligned}
&\widetilde{X}_{i,k}[f(\Theta(\,\cdot\,,\xi))](\eta)\\
&\qquad = \widetilde{X}_{i,k}[f'(\Theta(\xi,\,\cdot\,))](\eta) = (Y_{i,k}f' + R^{\xi}_{i,k}f')(\Theta(\xi,\eta))\\
&\qquad = -(Y^R_{i,k}f)'(\Theta(\xi,\eta)) + R^{\xi}_{i,k}f'(\Theta(\xi,\eta)) = -(Y^R_{i,k}f)(\Theta(\eta,\xi)) + ((R^{\xi}_{i,k})'f)(\Theta(\eta,\xi)),
\end{aligned}
\tag{4-14}
$$

where

$$((R^{\xi}_{i,k})'f)(u) = (R^{\xi}_{i,k}f')(-u)$$

is a differential operator of degree $\le i-1$. This proves (4-11). □

Proof of Theorem 4.11. As we noted after Lemma 4.13, we are left to prove the assertion about $T(\xi_0)\widetilde{X}_i$ and $T(\xi_0)\widetilde{X}_0$. We only give the proof for the case $\lambda \ge 1$, $i = 1, \ldots, q$. The proof for $\lambda \ge 2$, $i = 0$ being very similar. Like in the proof of Lemma 4.13, it is enough to consider the part k' of the kernel of T, the proof for k'' being completely analogous (see Definition 4.5). Let us expand

$$k'(\xi_0;\xi,\eta) = \left\{\sum_{j=1}^{H_m} a_j(\xi)b_j(\eta)D_j\Gamma(\xi_0;\,\cdot\,) + a_0(\xi)b_0(\eta)D_0\Gamma(\xi_0;\,\cdot\,)\right\}(\Theta(\eta,\xi)),$$

where $D_0\Gamma(\xi_0;\,\cdot\,)$ has bounded Y_i-derivatives ($i = 1, 2, \ldots, q$). We can consider each of the terms

$$T_j(\xi_0)\widetilde{X}_i f(\xi) \equiv \int a_j(\xi)b_j(\eta)D^{\eta}_j\Gamma(\xi_0;\Theta(\eta,\xi))\widetilde{X}_i f(\eta)\,d\eta$$

(this time it is important to recall the η-dependence of the coefficients of D_j) and distinguish 2 cases:

(i) $D_j\Gamma$ is homogeneous of degree $\ge 2-Q$ or it is regular (that is, $D_j\Gamma$ has bounded Y_i-derivatives);

(ii) $T_j(\xi_0)$ is a frozen operator of type 1 and $D_j\Gamma$ is homogeneous of degree $1-Q$.

Case (i). We can integrate by parts, recalling that the transpose of $\widetilde{X}_i$ is

$$(\widetilde{X}_i)^T g(\eta) = -\widetilde{X}_i g(\eta) + c_i(\eta)g(\eta)$$

with c_i smooth functions:

$$
\begin{aligned}
&T_j(\xi_0)\widetilde{X}_i f(\xi)\\
&\quad= \int c_i(\eta)a_j(\xi)b_j(\eta)D^{\eta}_j\Gamma(\xi_0;\Theta(\eta,\xi))f(\eta)\,d\eta - \int a_j(\xi)(\widetilde{X}_i b_j)(\eta)D^{\eta}_j\Gamma(\xi_0;\Theta(\eta,\xi))f(\eta)\,d\eta\\
&\qquad - \int a_j(\xi)b_j(\eta)\widetilde{X}_i[D^{\eta}_j\Gamma(\xi_0;\Theta(\,\cdot\,,\xi))](\eta)f(\eta)\,d\eta - \int a_j(\xi)b_j(\eta)(\widetilde{X}^{\eta}_i D^{\eta}_j)\Gamma(\xi_0;\Theta(\eta,\xi))f(\eta)\,d\eta\\
&\quad= A(\xi) + B(\xi) + C(\xi) + D(\xi).
\end{aligned}
$$

Now, $A(\xi)+B(\xi)$ is still an operator of type λ, applied to f; in particular, it can be seen as operator of type $\lambda-1$; the same is true for $D(\xi)$, by Remark 4.6. To study $C(\xi)$, we apply Lemma 4.14, which gives

$$\widetilde{X}_i[D^{\eta}_j\Gamma(\xi_0;\Theta(\,\cdot\,,\xi))](\eta) = -(Y^R_i D^{\eta}_j\Gamma)(\xi_0,\Theta(\eta,\xi)) + ((R^{\xi}_i)'D^{\eta}_j\Gamma)(\xi_0,\Theta(\eta,\xi)).$$

Since Y_i^R is homogeneous of degree 1, $a_j(\xi)b_j(\eta)Y_i^R D_j^\eta\Gamma(\xi_0,\Theta(\eta,\xi))$ is a kernel of type $\lambda-1$. Since $(R_i^\xi)'$ is a differential operator of degree ≤ 0, the kernel $a_j(\xi)b_j(\eta)((R_i^\xi)' D_j^\eta\Gamma)(\xi_0,\Theta(\eta,\xi))$ is of type λ.

Note that, even when the coefficients of the differential operator D_j (in the expression $D_j\Gamma(\xi_0;\Theta(\eta,\xi))$) do not depend on ξ and η, this procedure introduces, with the operator $(R_i^\xi)'$, a new ξ-dependence of the coefficients. Compare this with our remark in the proof of Lemma 4.12.

Case (ii). In this case the kernel $(Y_i^R D_j\Gamma)$ is singular, so that the computation must be handled with more care. We can write

$$T_j(\xi_0)\widetilde{X}_i f(\xi)=\lim_{\varepsilon\to 0}\int a_j(\xi)b_j(\eta)\varphi_\varepsilon(\Theta(\xi,\eta))D_j\Gamma(\xi_0;\Theta(\eta,\xi))\widetilde{X}_i f(\eta)\,d\eta\equiv\lim_{\varepsilon\to 0}T_\varepsilon(\xi)$$

with φ_ε as in the proof of Lemma 4.13. Note that, choosing a radial φ, we have $\varphi_\varepsilon(\Theta(\xi,\eta))=\varphi_\varepsilon(\Theta(\eta,\xi))$. Then

$$\begin{aligned}
T_\varepsilon(\xi)= &\int c_i(\eta)a_j(\xi)b_j(\eta)\varphi_\varepsilon(\Theta(\xi,\eta))D_j\Gamma(\xi_0;\Theta(\eta,\xi))f(\eta)\,d\eta\\
&-\int a_j(\xi)(\widetilde{X}_i b_j)(\eta)\varphi_\varepsilon(\Theta(\xi,\eta))D_j\Gamma(\xi_0;\Theta(\eta,\xi))f(\eta)\,d\eta\\
&-\int a_j(\xi)b_j(\eta)\widetilde{X}_i[\varphi_\varepsilon(\Theta(\cdot,\xi))D_j\Gamma(\xi_0;\Theta(\cdot,\xi))](\eta)f(\eta)\,d\eta\\
&-\int a_j(\xi)b_j(\eta)\varphi_\varepsilon(\Theta(\xi,\eta))(\widetilde{X}_i^\eta D_j^\eta)\Gamma(\xi_0;\Theta(\eta,\xi))f(\eta)\,d\eta\\
=&:A_\varepsilon(\xi)+B_\varepsilon(\xi)+C_\varepsilon(\xi)+D_\varepsilon(\xi).
\end{aligned}$$

Now $A_\varepsilon(\xi)+B_\varepsilon(\xi)+D_\varepsilon(\xi)$ converge to an operator of type λ, as $A(\xi)$, $B(\xi)$, $D(\xi)$ are in Case (i), while, by Theorem 3.2 and Lemma 4.14,

$$\begin{aligned}
C_\varepsilon(\xi)= &-\int a_j(\xi)b_j(\eta)f(\eta)(Y_i\varphi_\varepsilon)(\Theta(\eta,\xi))D_j\Gamma(\xi_0;\Theta(\eta,\xi))\,d\eta\\
&-\int a_j(\xi)b_j(\eta)f(\eta)(R_i^\xi\varphi_\varepsilon)(\Theta(\eta,\xi))D_j\Gamma(\xi_0;\Theta(\eta,\xi))\,d\eta\\
&+\int a_j(\xi)b_j(\eta)f(\eta)\varphi_\varepsilon(\Theta(\eta,\xi))(Y_i^R D_j\Gamma)(\xi_0,\Theta(\eta,\xi))\,d\eta\\
&-\int a_j(\xi)b_j(\eta)f(\eta)\varphi_\varepsilon(\Theta(\eta,\xi))((R_i^\xi)' D_j\Gamma)(\xi_0,\Theta(\eta,\xi))\,d\eta\\
=&:E_\varepsilon(\xi)+F_\varepsilon(\xi)+G_\varepsilon(\xi)+H_\varepsilon(\xi).
\end{aligned}$$

Now $H_\varepsilon(\xi)$ tends to an operator of type 1 and $G_\varepsilon(\xi)$ tends to

$$\mathrm{PV}\int a_j(\xi)b_j(\eta)f(\eta)(Y_i^R D_j\Gamma)(\xi_0,\Theta(\eta,\xi))\,d\eta,$$

which is an operator of type 0. As to $E_\varepsilon(\xi)$, the same computation as in the proof of Lemma 4.13 gives

$$E_\varepsilon(\xi)\to\alpha(\xi_0,\xi)(abcf)(\xi)$$

with

$$\alpha(\xi_0,\xi)=\int Y_i\varphi(v)D_1^{\xi}\Gamma(\xi_0;v)\,dv,$$

which is the multiplicative part of an operator of type 0. A similar computation shows that $F_\varepsilon(\xi)\to 0$. □

Let us come to the second main result of this section. In [Rothschild and Stein 1976, corollary on p. 296], the following fact is proved for a family of Hörmander's vector fields without the drift $\widetilde{X}_0$: for any frozen operator $T(\xi_0)$ of type 1, $i=1,2,\dots,q$, there exist operators $T_{ij}(\xi_0)$, $T_i(\xi_0)$ of type 1 such that

$$\widetilde{X}_iT(\xi_0)=\sum_{j=1}^{q}T_{ij}(\xi_0)\widetilde{X}_j+T_i(\xi_0).$$

This possibility of exchanging the order of integral and differential operators will be crucial in the proof of representation formulas. However, such an identity cannot be proved in this form when the drift $\widetilde{X}_0$ is present. Instead, we are going to prove the following, which will be enough for our purposes.

Theorem 4.15. *If $T(\xi_0)$ is a frozen operator of type $\lambda\ge 1$, $i=1,2,\dots,q$, then*

$$\widetilde{X}_iT(\xi_0)=\sum_{k=1}^{q}T_k^i(\xi_0)\widetilde{X}_k+\sum_{h,j=1}^{q}\tilde{a}_{hj}(\xi_0)T^{hi}(\xi_0)\widetilde{X}_j+T_0^i(\xi_0)+T^i(\xi_0)\widetilde{\mathscr{L}}_0, \tag{4-15}$$

where $T_k^i(\xi_0)$ $(k=0,1,\dots,q)$ and $T^{hi}(\xi_0)$ are frozen operators of type λ, $T^i(\xi_0)$ are frozen operators of type $\lambda+1$, and $\tilde{a}_{hj}(\xi_0)$ are the frozen coefficients of $\widetilde{\mathscr{L}}_0$.

If $T(\xi_0)$ is a frozen operator of type $\lambda\ge 2$, then

$$\widetilde{X}_0T(\xi_0)=\sum_{k=1}^{q}T_k(\xi_0)\widetilde{X}_k+\sum_{h,j=1}^{q}\tilde{a}_{hj}(\xi_0)T^{h}(\xi_0)\widetilde{X}_j+T_0(\xi_0)+T(\xi_0)\widetilde{\mathscr{L}}_0, \tag{4-16}$$

where $T_k(\xi_0)$ $(k=0,1,\dots,q)$ and $T^h(\xi_0)$ are frozen operators of type $\lambda-1$, $T(\xi_0)$ is a frozen operator of type λ.

We start with the following lemma, similar to that proved in [Rothschild and Stein 1976, p. 296].

Lemma 4.16. *For any vector field $\widetilde{X}_{j_0,k_0}$ $(j_0=1,2,\dots,s,\ k_0=1,2,\dots,k_{j_0})$, there exist smooth functions*

$$\{a_{jk}^{j_0k_0\eta}\}_{\substack{j=1,2,\dots,s\\ k=1,2,\dots,h_j}}$$

having local degree $\ge\max\{j-j_0,0\}$ and smoothly depending on η, such that, for any $f\in C_0^\infty(\mathbb{G})$, one can write

$$\widetilde{X}_{j_0,k_0}[f(\Theta(\eta,\cdot))](\xi)=\sum_{\substack{j=1,2,\dots,s\\ k=1,2,\dots,k_j}}a_{jk}^{j_0k_0\eta}(\Theta(\eta,\xi))\widetilde{X}_{j,k}[f(\Theta(\cdot,\xi))](\eta)+(R_{j_0,k_0}^{\xi,\eta}f)(\Theta(\eta,\xi)), \tag{4-17}$$

where $R_{j_0}^{\xi,\eta}$ is a vector field of local degree $\le j_0-1$, smoothly depending on ξ, η.

Proof. By Theorem 3.2 we know that

$$\widetilde{X}_{j_0,k_0}[f(\Theta(\eta,\cdot))](\xi)=(Y_{j_0,k_0}f+R^{\eta}_{j_0,k_0}f)(\Theta(\eta,\xi))\equiv(Z^{\eta}_{j_0,k_0}f)(\Theta(\eta,\xi)), \tag{4-18}$$

where $Z^{\eta}_{j_0,k_0}$ is a vector field of local degree $\le j_0$, smoothly depending on η. To rewrite $(Z^{\eta}_{j_0,k_0}f)$ in a suitable form, we start from the following identities:

$$Y_{i,k}=\frac{\partial}{\partial u_{ik}}+\sum_{i<l\le s}\sum_{r=1}^{k_l}g^{ik}_{lr}(u)\frac{\partial}{\partial u_{lr}} \tag{4-19}$$

for any $i=1,2,\dots,s$ and $k=1,2,\dots,k_i$;

$$Y_{i,k}=\sum g^{ik}_{lr}(u)Y^{R}_{l,r}, \tag{4-20}$$

where $g^{ik}_{lr}(u)$ are homogeneous of degree $l-i$; see [Rothschild and Stein 1976, p. 295]. Hence we can write

$$Z^{\eta}_{j_0,k_0}=\sum a^{\eta}_{jk}(u)\frac{\partial}{\partial u_{jk}},$$

where a_{jk} has local degree $\ge j-j_0$ and smoothly depends on η. By inverting (for any i,k) the triangular system (4-19), we obtain

$$\frac{\partial}{\partial u_{jk}}=Y_{j,k}+\sum_{j<l\le s}\sum_{r=1}^{k_l}f^{jk}_{lr}(u)Y_{l,r},$$

where each $f^{jk}_{lr}(u)$ is homogeneous of degree $l-j$. Also using (4-20), we have

$$(Z^{\eta}_{j_0,k_0}f)(u)=\sum a^{\eta}_{jk}(u)[(Y_{j,k}f)(u)+\sum_{j<l\le s}f^{jk}_{lr}(u)(Y_{l,r}f)(u)]=\sum b^{\eta}_{lr}(u)(Y^{R}_{l,r}f)(u), \tag{4-21}$$

where

$$b^{\eta}_{lr}\text{ has local degree}\ge\max\{l-j_0,0\} \tag{4-22}$$

and smoothly depends on η. Then, by Lemma 4.14,

$$(Z^{\eta}_{j_0,k_0}f)(\Theta(\eta,\xi))=\sum_{l,r}-b^{\eta}_{lr}(\Theta(\eta,\xi))\widetilde{X}_{l,r}[f(\Theta(\cdot,\xi))](\eta)+\sum_{l,r}(b^{\eta}_{lr}(R^{\xi}_{l,r})'f)(\Theta(\eta,\xi)), \tag{4-23}$$

where $(R^{\xi}_{l,r})'$ is a differential operator of local degree $\le l-1$, hence the differential operator on $\mathbb{G}$

$$R^{\xi,\eta}_{j_0,k_0}\equiv\sum_{l,r}b^{\eta}_{lr}(R^{\xi}_{l,r})'\text{ has local degree}\le j_0-1 \tag{4-24}$$

and depends smoothly on ξ,η. Collecting (4-18), (4-22), (4-23), (4-24), the lemma is proved, with $a^{j_0k_0\eta}_{jk}=-b^{\eta}_{jk}$. □

Thanks to this lemma, we can prove the following, which is similar to [Rothschild and Stein 1976, Theorem 9].

Theorem 4.17. (i) *Suppose $T(\xi_0)$ is a frozen operator of type $\lambda \geq 1$. Given a vector field $\widetilde{X}_i$ for $i = 1, 2, \dots, q$, there exist frozen operators $T^i(\xi_0)$ of type λ, and $T^i_{jk}(\xi_0)$, frozen operators of type $\lambda + j - 1$, such that*

$$\widetilde{X}_i T(\xi_0) = \sum_{j,k} T^i_{jk}(\xi_0) \widetilde{X}_{j,k} + T^i(\xi_0). \tag{4-25}$$

(ii) *Suppose $T(\xi_0)$ is a frozen operator of type $\lambda \geq 2$. There exist $T^0(\xi_0)$ and $T^0_{jk}(\xi_0)$, frozen operators of type $\lambda - 1$ and $\lambda + \max\{j-2, 0\}$, respectively, such that*

$$\widetilde{X}_0 T(\xi_0) = \sum_{j,k} T^0_{jk}(\xi_0) \widetilde{X}_{j,k} + T^0(\xi_0). \tag{4-26}$$

Proof. First of all, it is enough to consider the part k' of the kernel of $T(\xi_0)$, the proof for k'' being completely analogous (see Definition 4.5).

(i) If $T(\xi_0)$ is a frozen operator of type $\lambda \geq 1$ with kernel k', we can write it as

$$T(\xi_0) f(\xi) = \int a(\xi) D\Gamma(\xi_0; \Theta(\eta, \xi)) b(\eta) f(\eta)\, d\eta + T'(\xi_0) f(\xi),$$

where $D\Gamma(\xi_0, \cdot)$ is homogeneous of degree $\lambda - Q$ and $T'(\xi_0)$ is a frozen operator of degree $\lambda + 1$. Since $\widetilde{X}_i T'(\xi_0)$ is a frozen operator of type λ, it already has the form $T^i(\xi_0)$ required by the theorem, hence it is enough to prove that

$$\widetilde{X}_i \int a(\xi) D\Gamma(\xi_0; \Theta(\eta, \xi)) b(\eta) f(\eta)\, d\eta$$

can be rewritten in the form

$$\sum_{j,k} T^i_{jk}(\xi_0) \widetilde{X}_{j,k} f(\xi) + T^i(\xi_0) f(\xi)$$

with $T^i_{jk}(\xi_0)$ and $T^i(\xi_0)$ frozen operators of type $\lambda + j - 1$ and λ, respectively. Next, we have to distinguish two cases.

Case 1: $\lambda \geq 2$. In this case the $\widetilde{X}_i$ derivative can be taken under the integral sign, writing

$$\begin{aligned}
&\widetilde{X}_i \int a(\xi) D\Gamma(\xi_0; \Theta(\eta, \xi)) b(\eta) f(\eta)\, d\eta \\
&\qquad = \int (\widetilde{X}_i a)(\xi) D\Gamma(\xi_0; \Theta(\eta, \xi)) b(\eta) f(\eta)\, d\eta + \int a(\xi) \widetilde{X}_i [D\Gamma(\Theta(\eta, \cdot))](\xi) b(\eta) f(\eta)\, d\eta \\
&\qquad =: A(\xi) + B(\xi).
\end{aligned}$$

Now $A(\xi)$ is a frozen operator of type λ, while applying Lemma 4.16 with $j_0 = 1$ we get

$$\begin{aligned}
B(\xi) = &\int a(\xi) \sum_{l,r} a^i_{lr}(\Theta(\eta, \xi)) \widetilde{X}_{l,r} [D\Gamma(\xi_0; \Theta(\,\cdot\,, \xi))](\eta) b(\eta) f(\eta)\, d\eta \\
&+ \int a(\xi) (R^\xi_i D\Gamma)(\xi_0; \Theta(\eta, \xi)) b(\eta) f(\eta)\, d\eta \\
=:&\ C(\xi) + D(\xi),
\end{aligned}$$

where R_i^{ξ} are differential operators of local degree ≤ 0, and the a_{lr}^{i} have local degree $\geq l-1$. Hence D is a frozen operator of type λ, while, since the transposed vector field of $\widetilde{X}_{l,r}$ is

$$\widetilde{X}_{l,r}^{T} = -\widetilde{X}_{l,r} + c_{l,r}$$

with $c_{l,r}$ smooth functions,

$$\begin{aligned} C(\xi) = & -a(\xi)\sum_{l,r}\int \widetilde{X}_{l,r}[a_{lr}^{i}(\Theta(\,\cdot\,,\xi))b(\,\cdot\,)](\eta)D\Gamma(\xi_0;\Theta(\eta,\xi))f(\eta)\,d\eta \\ & + a(\xi)\sum_{l,r}\int a_{lr}^{i}(\Theta(\eta,\xi))D\Gamma(\xi_0;\Theta(\eta,\xi))c_{l,r}(\eta)b(\eta)f(\eta)\,d\eta \\ & - a(\xi)\sum_{l,r}\int a_{lr}^{i}(\Theta(\eta,\xi))D\Gamma(\xi_0;\Theta(\eta,\xi))b(\eta)\widetilde{X}_{l,r}f(\eta)\,d\eta. \end{aligned}$$

The first two terms in the last expression are still frozen operators of type λ applied to f, while the third is a sum of operators of type $\lambda+l-1$ applied to $\widetilde{X}_{l,r}f$, as required by the theorem.

Case 2: $\lambda = 1$. In this case we have to compute the derivative of the integral in a distributional sense, as was already done in the proof of Lemma 4.13. With the same meaning of φ_ε, let us compute

$$\lim_{\varepsilon\to 0}\widetilde{X}_i\int a(\xi)\varphi_\varepsilon(\Theta(\eta,\xi))D\Gamma(\xi_0;\Theta(\eta,\xi))b(\eta)f(\eta)\,d\eta.$$

Actually, this gives exactly the same result as in case 1:

$$\begin{aligned} & \widetilde{X}_i\int a(\xi)\varphi_\varepsilon(\Theta(\eta,\xi))D\Gamma(\xi_0;\Theta(\eta,\xi))b(\eta)f(\eta)\,d\eta \\ & = \int(\widetilde{X}_i a)(\xi)\varphi_\varepsilon(\Theta(\eta,\xi))D\Gamma(\xi_0;\Theta(\eta,\xi))b(\eta)f(\eta)\,d\eta + \int a(\xi)\widetilde{X}_i[(\varphi_\varepsilon D\Gamma)(\Theta(\eta,\cdot))](\xi)b(\eta)f(\eta)\,d\eta \\ & = A_\varepsilon(\xi)+B_\varepsilon(\xi), \end{aligned}$$

where $A_\varepsilon(\xi)\to\int(\widetilde{X}_i a)(\xi)D\Gamma(\xi_0;\Theta(\eta,\xi))b(\eta)f(\eta)\,d\eta$ and

$$\begin{aligned} B_\varepsilon(\xi) = & \int a(\xi)\sum_{l,r}a_{lr}^{i}(\Theta(\eta,\xi))\widetilde{X}_{l,r}[\varphi_\varepsilon(\Theta(\,\cdot\,,\xi))D\Gamma(\xi_0;\Theta(\,\cdot\,,\xi))](\eta)b(\eta)f(\eta)\,d\eta \\ & + \int a(\xi)(R_i^{\xi}(\varphi_\varepsilon D\Gamma))(\xi_0;\Theta(\eta,\xi))b(\eta)f(\eta)\,d\eta \\ =: & \; C_\varepsilon(\xi)+D_\varepsilon(\xi), \end{aligned}$$

where $C_\varepsilon(\xi)$ converges to the expression called $C(\xi)$ in the computation of case 1; as for $D_\varepsilon(\xi)$,

$$R_i^{\xi}(\varphi_\varepsilon D\Gamma) = (R_i^{\xi}\varphi_\varepsilon)D\Gamma + \varphi_\varepsilon R_i^{\xi}D\Gamma.$$

Now, $\varphi_\varepsilon R_i^{\xi}D\Gamma \to R_i^{\xi}D\Gamma$ while $(R_i^{\xi}\varphi_\varepsilon)D\Gamma\to 0$, R_i^{ξ} being a vector field of local degree ≤ 0. Hence $D_\varepsilon(\xi)$ also converges to the expression called $D(\xi)$ in the computation of case 1, and we are done.

(ii) Now let $T(\xi_0)$ be a frozen operator of type $\lambda \geq 2$ with kernel k'. As in (i), it is enough to prove that

$$\widetilde{X}_0 \int a(\xi) D\Gamma(\xi_0; \Theta(\eta,\xi)) b(\eta) f(\eta)\, d\eta,$$

where $D\Gamma$ is homogeneous of degree $\lambda - Q$ can be rewritten in the form

$$\sum_{j,k} T^0_{jk}(\xi_0) \widetilde{X}_{j,k} f(\xi) + T^0(\xi_0) f(\xi)$$

with $T^0_{jk}(\xi_0)$ and $T^0(\xi_0)$ frozen operators of type $\lambda + j - 2$ and $\lambda - 1$, respectively. Let us consider only the case $\lambda \geq 3$, the case $\lambda = 2$ being handled with the modification seen in (i), Case 2. By Lemma 4.16,

$$\begin{aligned}
\widetilde{X}_0 \int a(\xi) D\Gamma(\xi_0; \Theta(\eta,\xi)) b(\eta) f(\eta)\, d\eta \\
&= \int (\widetilde{X}_0 a)(\xi) D\Gamma(\xi_0; \Theta(\eta,\xi)) b(\eta) f(\eta)\, d\eta \\
&\quad + \int a(\xi) \sum_{l,r} a^0_{lr}(\Theta(\eta,\xi)) \widetilde{X}_{l,r}[D\Gamma(\xi_0; \Theta(\cdot\,,\xi))](\eta) b(\eta) f(\eta)\, d\eta \\
&\quad + \int a(\xi)(R^\xi_0 D\Gamma)(\xi_0; \Theta(\eta,\xi)) b(\eta) f(\eta)\, d\eta \\
&=: A(\xi) + C(\xi) + D(\xi),
\end{aligned}$$

where R^ξ_0 are now differential operators of local degree ≤ 1, and the a^0_{lr} have local degree $\geq \max\{j-2, 0\}$. Then $A(\xi)$ is a frozen operator of type λ, applied to f; $D(\xi)$ is a frozen operator of type $\lambda - 1$, applied to f. Moreover,

$$\begin{aligned}
C(\xi) = &- a(\xi) \sum_{l,r} \int \widetilde{X}_{l,r}[a^0_{lr}(\Theta(\cdot\,,\xi)) b(\cdot)](\eta) D\Gamma(\xi_0; \Theta(\eta,\xi)) f(\eta)\, d\eta \\
&+ a(\xi) \sum_{l,r} \int a^0_{lr}(\Theta(\eta,\xi)) D\Gamma(\xi_0; \Theta(\eta,\xi)) c_{l,r}(\eta) b(\eta) f(\eta)\, d\eta \\
&- a(\xi) \sum_{l,r} \int a^0_{lr}(\Theta(\eta,\xi)) D\Gamma(\xi_0; \Theta(\eta,\xi)) b(\eta) \widetilde{X}_{l,r} f(\eta)\, d\eta,
\end{aligned}$$

where the first two terms are still frozen operators of type λ, applied to f, while the third is the sum of frozen operators of type $\lambda + \max\{j-2, 0\}$ applied to $\widetilde{X}_{l,r} f$. □

Proof of Theorem 4.15. It suffices to prove (4-15), since the proof of (4-16) is similar. So, if $\widetilde{X}_i T(\xi_0)$ is like in (4-15), let us apply Theorem 4.17 and rewrite $\widetilde{X}_i T(\xi_0)$ like in (4-25). Now, let us consider one of the terms $T^i_{jk}(\xi_0)\widetilde{X}_{j,k}$ appearing in (4-25).

If $j = 1$, the term is already in the form required by the theorem we are proving.

If $j = 2$, then $\widetilde{X}_{2,k}$ can be written as a combination of commutators of the vector fields $\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_q$, plus (possibly) the field $\widetilde{X}_0$. Then $T^i_{2k}(\xi_0)\widetilde{X}_{2,k}$ contains terms $T^i_{2k}(\xi_0)\widetilde{X}_h\widetilde{X}_j$ and possibly a term $T^i_{2k}(\xi_0)\widetilde{X}_0$.

By Theorem 4.17, we know T^i_{2k} is a frozen operator of type $\lambda+1$. Now

$$T^i_{2k}(\xi_0)\widetilde{X}_h\widetilde{X}_j = (T^i_{2k}(\xi_0)\widetilde{X}_h)\widetilde{X}_j = T^i_k(\xi_0)\widetilde{X}_j,$$

where, by Theorem 4.11, $T^i_k(\xi_0)$ is a frozen operator of type λ; on the other hand, by (4-2),

$$\begin{aligned} T^i_{2k}(\xi_0)\widetilde{X}_0 &= T^i_{2k}(\xi_0)\biggl(\widetilde{\mathscr{L}}_0 - \sum_{h,j=1}^{q} \tilde{a}_{hj}(\xi_0)\widetilde{X}_h\widetilde{X}_j\biggr) \\ &= T^i_{2k}(\xi_0)\widetilde{\mathscr{L}}_0 - \sum_{h,j=1}^{q} \tilde{a}_{hj}(\xi_0)(T^i_{2k}(\xi_0)\widetilde{X}_h)\widetilde{X}_j \;=\; T^i_{2k}(\xi_0)\widetilde{\mathscr{L}}_0 - \sum_{h,j=1}^{q} \tilde{a}_{hj}(\xi_0)T^i_{h,k}(\xi_0)\widetilde{X}_j, \end{aligned}$$

with $T^i_{2k}(\xi_0)$ and $T^i_{h,k}(\xi_0)$ frozen operators of type $\lambda+1$ and λ, respectively, which is in the form allowed by the thesis of the theorem we are proving.

Finally, if $j > 2$, it is enough to look at the final part of the differential operator $\widetilde{X}_{j,k}$. It is always possible to rewrite $\widetilde{X}_{j,k}$ either as $\widetilde{X}_{j-1,k}\widetilde{X}_{1,k}$ or as $\widetilde{X}_{j-2,k}\widetilde{X}_{2,k}$. In the first case, we have

$$T^i_{jk}(\xi_0)\widetilde{X}_{j,k} = (T^i_{jk}(\xi_0)\widetilde{X}_{j-1,k})\widetilde{X}_{1,k} = T'^i_{jk}(\xi_0)\widetilde{X}_{1,k},$$

with $T'^i_{jk}(\xi_0)$ frozen operator of type λ, which is already in the proper form; in the second case, we have

$$T^i_{jk}(\xi_0)\widetilde{X}_{j,k} = (T^i_{jk}(\xi_0)\widetilde{X}_{j-2,k})\widetilde{X}_{2,k} = T'^i_j(\xi_0)\widetilde{X}_{2,k}$$

with $T'^i_{jk}(\xi_0)$ frozen operator of type $\lambda+1$, and then we can proceed as in the case $j = 2$. □

4C. *Parametrix and representation formulas.* Throughout this subsection we will make extensive use of computations on frozen operators of type λ. To make our formulas more readable, we will use the symbols

$$T(\xi_0), \quad S(\xi_0), \quad P(\xi_0)$$

(possibly with some indices) to denote frozen operators of type 0, 1, 2, respectively.

In order to prove representation formulas for second order derivatives, we start with the following parametrix identities, analogous to [Rothschild and Stein 1976, Theorem 10; Bramanti and Brandolini 2000a, Theorem 3.1].

Theorem 4.18. *Given $a \in C_0^\infty(\widetilde{B}(\bar{\xi}, R))$, there exist $S_{ij}(\xi_0)$, $S_0(\xi_0)$, $S^*_{ij}(\xi_0)$, $S^*_0(\xi_0)$, frozen operators of type* 1 *and $P(\xi_0)$, $P^*(\xi_0)$, frozen operators of type* 2 *(over the ball $\widetilde{B}(\bar{\xi}, R)$) such that*

$$aI = \widetilde{\mathscr{L}}_0^T P^*(\xi_0) + \sum_{i,j=1}^{q} \tilde{a}_{ij}(\xi_0)S^*_{ij}(\xi_0) + S^*_0(\xi_0), \tag{4-27}$$

$$aI = P(\xi_0)\widetilde{\mathscr{L}}_0 + \sum_{i,j=1}^{q} \tilde{a}_{ij}(\xi_0)S_{ij}(\xi_0) + S_0(\xi_0), \tag{4-28}$$

where I denotes the identity. Moreover, $S_{ij}^*(\xi_0)$, $S_0^*(\xi_0)$, $P^*(\xi_0)$ are modeled on Γ^T, while $S_{ij}(\xi_0)$, $S_0(\xi_0)$, $P(\xi_0)$ are modeled on Γ. Explicitly,

$$P^*(\xi_0)f(\xi) = -\frac{a(\xi)}{c(\xi)}\int_{\widetilde{B}}\Gamma^T(\xi_0;\Theta(\eta,\xi))\,b(\eta)\,f(\eta)\,d\eta,$$
$$P(\xi_0)f(\xi) = -b(\xi)\int_{\widetilde{B}}\frac{a(\eta)}{c(\eta)}\Gamma(\xi_0;\Theta(\eta,\xi))\,f(\eta)\,d\eta,$$

where c is the function appearing in Theorem 3.3(c).

Sketch of the proof. Let us define

$$P^*(\xi_0)f(\xi) = -\frac{a(\xi)}{c(\xi)}\int_{\widetilde{B}}\Gamma^T(\xi_0;\Theta(\eta,\xi))b(\eta)f(\eta)\,d\eta,$$

where $a, b \in C_0^\infty(\widetilde{B}(\bar{\xi}, R))$ such that $ab = a$ and $c(\xi)$ is the function appearing in the formula of change of variables (3-6). Let us compute $\widetilde{\mathscr{L}}_0^T P^*(\xi_0)f$ for $f \in C_0^\infty(\widetilde{B}(\bar{\xi}, R))$. We can apply a distributional argument like in the proof of Lemma 4.13. For $\omega \in C_0^\infty(\widetilde{B}(\bar{\xi}, R))$, let us evaluate

$$\int_{\widetilde{B}}\widetilde{\mathscr{L}}_0\omega(\xi)P^*(\xi_0)f(\xi)\,d\xi = \lim_{\varepsilon\to 0}\int_{\widetilde{B}}\widetilde{\mathscr{L}}_0\omega(\xi)P_\varepsilon^*(\xi_0)f(\xi)\,d\xi,$$

where

$$P_\varepsilon^*(\xi_0)f(\xi) = -\frac{a(\xi)}{c(\xi)}\int_{\widetilde{B}}\varphi_\varepsilon(\Theta(\eta,\xi))\Gamma^T(\xi_0;\Theta(\eta,\xi))b(\eta)f(\eta)\,d\eta$$

with φ_ε as in the proof of Lemma 4.13. Now, computing the integral

$$\int_{\widetilde{B}}\widetilde{\mathscr{L}}_0\omega(\xi)P_\varepsilon^*(\xi_0)f(\xi)\,d\xi$$

and taking its limit for $\varepsilon \to 0$, by the same techniques used in Section 4B, we can prove (4-27). Transposing this identity, one finds (4-28). □

Now, starting from (4-28) and reasoning as in the proof of [Bramanti and Brandolini 2000a, Theorem 3.2], applying Theorem 4.11 and Theorem 4.15, one can easily prove the next two theorems.

Theorem 4.19 (representation of $\widetilde{X}_m\widetilde{X}_l u$ by frozen operators). *Let $a \in C_0^\infty(\widetilde{B}(\bar{\xi}, R))$, $\xi_0 \in \widetilde{B}(\bar{\xi}, R)$. Then, for any $m, l = 1, 2, \ldots, q$, there exist frozen operators over the ball $\widetilde{B}(\bar{\xi}, R)$ such that, for any $u \in C_0^\infty(\widetilde{B}(\bar{\xi}, R))$,*

$$\widetilde{X}_m\widetilde{X}_l(au) = T_{lm}(\xi_0)\widetilde{\mathscr{L}}_0 u + \sum_{k=1}^q T_{lm,k}(\xi_0)\widetilde{X}_k u + T_{lm}^0(\xi_0)u$$
$$+\sum_{i,j=1}^q \tilde{a}_{ij}(\xi_0)\left\{\sum_{k=1}^q T_{lm,k}^{ij}(\xi_0)\widetilde{X}_k u + \sum_{h,k=1}^q \tilde{a}_{hk}(\xi_0)T_{lm,h}^{\prime ij}(\xi_0)\widetilde{X}_k u + S_{lm}^{ij}(\xi_0)\widetilde{\mathscr{L}}_0 u + T_{lm}^{ij}(\xi_0)u\right\}. \quad (4\text{-}29)$$

(All the $T_{\dots}(\xi_0)$ are frozen operators of type 0 *and $S^{ij}_{lm}(\xi_0)$ are of type* 1.) *Also,*

$$\begin{aligned}\widetilde{X}_m\widetilde{X}_l(au) &= T_{lm}(\xi_0)\widetilde{\mathscr{L}}u + T_{lm}(\xi_0)\biggl(\sum_{i,j=1}^{q}[\tilde{a}_{ij}(\xi_0)-\tilde{a}_{ij}(\,\cdot\,)]\widetilde{X}_i\widetilde{X}_j u\biggr) + \sum_{k=1}^{q} T_{lm,k}(\xi_0)\widetilde{X}_k u + T^0_{lm}(\xi_0)u \\ &\quad + \sum_{i,j=1}^{q}\tilde{a}_{ij}(\xi_0)\biggl\{\sum_{k=1}^{q} T^{ij}_{lm,k}(\xi_0)\widetilde{X}_k u + \sum_{h,k=1}^{q}\tilde{a}_{hk}(\xi_0)T'^{ij}_{lm,h}(\xi_0)\widetilde{X}_k u + S^{ij}_{lm}(\xi_0)\widetilde{\mathscr{L}}u \\ &\quad + S^{ij}_{lm}(\xi_0)\biggl(\sum_{i,j=1}^{q}[\tilde{a}_{ij}(\xi_0)-\tilde{a}_{ij}(\,\cdot\,)]\widetilde{X}_i\widetilde{X}_j u\biggr) + T^{ij}_{lm}(\xi_0)u\biggr\}. \qquad (4\text{-}30)\end{aligned}$$

Remark 4.20. The representation formulas of the above theorem have a cumbersome aspect, due to the presence of the coefficients $\tilde{a}_{ij}(\xi_0)$ which appear several times as multiplicative factors. Anyway, if we agree to leave implicitly understood in the symbol of frozen operators the possible multiplication by the coefficients $\tilde{a}_{ij}$, our formulas assume the following more compact form

$$\widetilde{X}_m\widetilde{X}_l(au) = T_{lm}(\xi_0)\widetilde{\mathscr{L}}_0 u + \sum_{k=1}^{q} T^{lm}_k(\xi_0)\widetilde{X}_k u + T^0_{lm}(\xi_0)u$$

and

$$\widetilde{X}_m\widetilde{X}_l(au) = T_{lm}(\xi_0)\widetilde{\mathscr{L}}u + T_{lm}(\xi_0)\biggl(\sum_{i,j=1}^{q}[\tilde{a}_{ij}(\xi_0)-\tilde{a}_{ij}(\,\cdot\,)]\widetilde{X}_i\widetilde{X}_j u\biggr) + \sum_{k=1}^{q} T^{lm}_k(\xi_0)\widetilde{X}_k u + T^0_{lm}(\xi_0)u.$$

In the proof of a priori estimates, when we take $C^{\alpha}_{\widetilde{X}}$ or L^p norms of both sides of these identities, the multiplicative factors $\tilde{a}_{hj}$ will be simply bounded by taking, respectively, the $C^{\alpha}_{\widetilde{X}}$ or the L^∞ norms of the $\tilde{a}_{hj}$; hence leaving these factors implicitly understood is harmless.

The above theorem is suited to the proof of $C^{\alpha}_{\widetilde{X}}$ estimates for $\widetilde{X}_i\widetilde{X}_j u$. In order to prove L^p estimate for $\widetilde{X}_i\widetilde{X}_j u$ we need the following variation.

Theorem 4.21 (representation of $\widetilde{X}_m\widetilde{X}_l u$ by variable operators). *Given $a \in C^\infty_0(\widetilde{B}(\bar{\xi}, R))$, for any $m, l = 1, 2, \dots, q$, there exist variable operators over the ball $\widetilde{B}(\bar{\xi}, R)$ such that, for any $u \in C^\infty_0(\widetilde{B}(\bar{\xi}, R))$,*

$$\begin{aligned}&\widetilde{X}_m\widetilde{X}_l(au) \\ &\quad = T_{lm}\widetilde{\mathscr{L}}u + \sum_{i,j=1}^{q}[\tilde{a}_{ij}, T_{lm}]\widetilde{X}_i\widetilde{X}_j u + \sum_{k=1}^{q} T_{lm,k}\widetilde{X}_k u + T^0_{lm}u \\ &\qquad + \sum_{i,j=1}^{q}\tilde{a}_{ij}\biggl\{\sum_{k=1}^{q} T^{ij}_{lm,k}\widetilde{X}_k u + \sum_{h,k=1}^{q}\tilde{a}_{hk}T'^{ij}_{lm,h}\widetilde{X}_k u + S^{ij}_{lm}\widetilde{\mathscr{L}}u + \sum_{i,j=1}^{q}[\tilde{a}_{ij}, S^{ij}_{lm}]\widetilde{X}_i\widetilde{X}_j u + T^{ij}_{lm}u\biggr\}. \qquad (4\text{-}31)\end{aligned}$$

Here all the $T_{\dots}$ are variable operators of type 0, *S^{ij}_{lm} is of type* 1, *$[a, T]$ denotes the commutator of the multiplication for a with the operator T, and $\tilde{a}_{ij}$ are the coefficients of the operator $\widetilde{\mathscr{L}}$ (which are no longer frozen at ξ_0).*

Remark 4.22. The above representation formula can be written in a shorter way as

$$\widetilde{X}_m \widetilde{X}_l(au) = T_{lm}\widetilde{\mathcal{L}}u + \sum_{i,j=1}^{q} [\tilde{a}_{ij}, T_{lm}]\, \widetilde{X}_i \widetilde{X}_j u + \sum_{k=1}^{q} T_{lm,k} \widetilde{X}_k u + T_{lm}^0 u$$

if we leave understood in the symbol of variable operators the possible multiplication by the coefficients $\tilde{a}_{ij}$; see the previous remark.

5. Singular integral estimates for operators of type zero

The proof of a priori estimates on the derivatives $\widetilde{X}_i \widetilde{X}_j u$ will follow, as will be explained in Section 6 and Section 7, combining the representation formulas proved in Section 4C with suitable C^α or L^p estimates for "operators of type zero". To be more precise, the results we need are the $C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi}, R))$ continuity of a *frozen operator of type zero* and the $L^p(\widetilde{B}(\bar{\xi}, R))$ continuity of a *variable operator of type zero*, together with the $L^p(\widetilde{B}(\bar{\xi}, r))$ estimate for the commutator of a variable operator of type zero with the multiplication with a VMO function, implying that the $L^p(\widetilde{B}(\bar{\xi}, r))$ norm of the commutator vanishes as $r \to 0$. All these results will be derived in the present section, as an application of abstract results proved in [Bramanti and Zhu 2012] in the context of locally homogeneous spaces (see Section 3C). To apply them, we need to check that our kernels of type zero satisfy suitable properties. Moreover, to study *variable* operators of type zero, we also have to resort to the classical technique of expansion in series of spherical harmonics, dating back to Calderón and Zygmund [1957], and already applied in the framework of vector fields in [Bramanti and Brandolini 2000b; 2000a]. This study will be split into two subsections, the first devoted to frozen operators on C^α and the second to variable operators on L^p.

5A. $C^\alpha_{\widetilde{X}}$ *continuity of frozen operators of type* 0.

The goal of this section is the proof of the following.

Theorem 5.1. *Let $\widetilde{B}(\bar{\xi}, R)$ be as before, $\xi_0 \in \widetilde{B}(\bar{\xi}, R)$, and let $T(\xi_0)$ be a frozen operator of type $\lambda \geq 0$ over $\widetilde{B}(\bar{\xi}, R)$. Then there exists $c > 0$ depending on R, $\{\widetilde{X}_i\}$, α, and μ, such that, for any $r < R$ and $u \in C^\alpha_{\widetilde{X},0}(\widetilde{B}(\bar{\xi}, r))$,*

$$\|T(\xi_0)u\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},r))} \leq c\|u\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},r))}. \tag{5-1}$$

To prove this, we will apply theorems proved in [Bramanti and Zhu 2012] about the C^α continuity of singular and fractional integrals in spaces of locally homogeneous type, taking

$$\Omega_k = \widetilde{B}\left(\bar{\xi}, \frac{kR}{k+1}\right) \quad \text{for } k = 1, 2, 3, \ldots. \tag{5-2}$$

Following notation and assumptions in Definition 4.5, our frozen kernel of type zero can be written as

$$k(\xi_0; \xi, \eta) = k'(\xi_0; \xi, \eta) + k''(\xi_0; \xi, \eta).$$

We will prove Theorem 5.1 for the operator with kernel k', the proof for k'' being completely analogous. Let us split k' as

$$\begin{aligned}k'(\xi_0;\xi,\eta) &= a_1(\xi)b_1(\eta)D_1\Gamma(\xi_0;\Theta(\eta,\xi))+\left\{\sum_{i=2}^{H_m}a_i(\xi)b_i(\eta)D_i\Gamma(\xi_0;\cdot)+a_0(\xi)b_0(\eta)D_0\Gamma(\xi_0;\cdot)\right\}(\Theta(\eta,\xi))\\ &=: k_S(\xi,\eta)+k_F(\xi,\eta),\end{aligned}$$

where $D_1\Gamma(\xi_0; u)$ is homogeneous of degree $-Q$ while all the kernels $D_i\Gamma(\xi_0; u)$ are homogeneous of some degree $\geq 1-Q$ and $D_0\Gamma(\xi_0; u)$ is regular. Recall that all these kernels may also have an explicit (smooth) dependence on ξ, η; we will write $D_i^{\xi,\eta}\Gamma(\xi_0; \Theta(\eta, \xi))$ to point out this fact when it is important.

We want to apply [Bramanti and Zhu 2012, Theorem 5.4] (about singular integrals) to the kernel k_S and [Bramanti and Zhu 2012, Theorem 5.8] (about fractional integrals) to each term of the kernel k_F.

We start with the following result, very similar to [Bramanti and Brandolini 2000a, Proposition 2.17].

Proposition 5.2. *Let $W^{\xi,\eta}(\cdot)$ be a function defined on the homogeneous group $\mathbb{G}$, smooth outside the origin and homogeneous of degree $\ell-Q$ for some nonnegative integer ℓ, smoothly depending on the parameters ξ, $\eta\in\widetilde{B}(\bar{\xi}, R)$, and let*

$$K(\xi,\eta)=W^{\xi,\eta}(\Theta(\eta,\xi))$$

be defined for ξ, $\eta\in\widetilde{B}(\bar{\xi}, R)$. Then K satisfies the following.

(i) *The growth condition: there exists a constant c such that*

$$|K(\xi,\eta)|\leq c\cdot\sup_{\|u\|=1}|W^{\xi,\eta}(u)|\cdot d_{\widetilde{X}}(\xi,\eta)^{\ell-Q}.$$

(ii) *The mean value inequality: there exists a constant $c>0$ such that, for every ξ_0, ξ, η with $d_{\widetilde{X}}(\xi_0,\eta)\geq 2d_{\widetilde{X}}(\xi_0,\xi)$,*

$$|K(\xi_0,\eta)-K(\xi,\eta)|+|K(\eta,\xi_0)-K(\eta,\xi)|\leq C\frac{d_{\widetilde{X}}(\xi_0,\xi)}{d_{\widetilde{X}}(\xi_0,\eta)^{Q+1-\ell}}, \tag{5-3}$$

where the constant C has the form

$$c\sup_{\substack{\|u\|=1\\ \xi,\eta\in\widetilde{B}(\bar{\xi},R)}}\{|\nabla_u W^{\xi,\eta}(u)|+|\nabla_\xi W^{\xi,\eta}(u)|+|\nabla_\eta W^{\xi,\eta}(u)|\}.$$

(iii) *The cancellation property: if $\ell=0$ and W satisfies the vanishing property*

$$\int_{r<\|u\|<R}W^{\xi,\eta}(u)\,du=0\quad \text{for every } R>r>0 \text{ and } \xi,\eta\in\widetilde{B}(\bar{\xi},R), \tag{5-4}$$

then, for any positive integer k, for every $\varepsilon_2>\varepsilon_1>0$ and $\xi\in\Omega_k$ (see (5-2)) such that $\widetilde{B}(\xi,\varepsilon_2)\subset\Omega_{k+1}$,

$$\left|\int_{\Omega_{k+1},\varepsilon_1<\rho(\xi,\eta)<\varepsilon_2}K(\xi,\eta)\,d\eta\right|+\left|\int_{\Omega_{k+1},\varepsilon_1<\rho(\xi,\eta)<\varepsilon_2}K(\eta,\xi)\,d\eta\right|\leq C\cdot(\varepsilon_2-\varepsilon_1), \tag{5-5}$$

where the constant C has the form

$$c \sup_{\substack{\|u\|=1 \\ \xi,\eta\in\widetilde{B}(\bar{\xi},R)}} \{|W^{\xi,\eta}(u)| + |\nabla_\xi W^{\xi,\eta}(u)| + |\nabla_\eta W^{\xi,\eta}(u)|\}.$$

Proof. (i) is trivial, by the homogeneity of W, and the equivalence between $d_{\widetilde{X}}$ and ρ (see Lemma 3.9).

In order to prove (ii), fix ξ_0, η and let $r = \frac{1}{2}\rho(\eta,\xi_0)$. Condition $\rho(\eta,\xi_0) > 2\rho(\xi,\xi_0)$ means that ξ is a point ranging in $\widetilde{B}_r(\xi_0)$. Applying (3-18) to the function

$$f(\xi) = K(\xi,\eta),$$

we can write

$$|f(\xi) - f(\xi_0)| \le c d_{\widetilde{X}}(\xi,\xi_0)\Bigg(\sum_{i=1}^{q} \sup_{\zeta\in\widetilde{B}(\xi_0,\frac{3}{4}d_{\widetilde{X}}(\xi_0,\eta))} |\widetilde{X}_i f(\zeta)| + d_{\widetilde{X}}(\xi,\xi_0) \sup_{\zeta\in\widetilde{B}(\xi_0,\frac{3}{4}d_{\widetilde{X}}(\xi_0,\eta))} |\widetilde{X}_0 f(\zeta)|\Bigg).$$

Noting that, for $\zeta \in \widetilde{B}(\xi_0, \frac{3}{4}d_{\widetilde{X}}(\xi_0,\eta))$,

$$\begin{aligned}|\widetilde{X}_i K(\,\cdot\,,\eta)(\zeta)| &= |\widetilde{X}_i(W^{\zeta,\eta}(\Theta(\,\cdot\,,\eta)))(\zeta) + (\widetilde{X}_i W^{\cdot,\eta}(\Theta(\zeta,\eta)))(\zeta)| \\ &\le |(Y_i W + R_i^\eta W)(\Theta(\eta,\zeta))| + |(\widetilde{X}_i W^{\cdot,\eta}(\Theta(\zeta,\eta)))(\zeta)|\end{aligned}$$

and recalling that, by Remark 4.6, $\nabla_\zeta W^{\zeta,\eta}(u)$ has the same u homogeneity as $W^{\zeta,\eta}(u)$, we get

$$\begin{aligned}|\widetilde{X}_i K(\,\cdot\,,\eta)(\zeta)| &\le \sup_{\substack{\|u\|=1 \\ \zeta,\eta\in\widetilde{B}(\bar{\xi},R)}} |\nabla_u W^{\xi,\eta}(u)| \frac{c}{\rho(\zeta,\eta)^{Q-\ell+1}} + \sup_{\substack{\|u\|=1 \\ \zeta,\eta\in\widetilde{B}(\bar{\xi},R)}} |\nabla_\zeta W^{\zeta,\eta}(u)| \frac{c}{\rho(\zeta,\eta)^{Q-\ell}} \\ &\le \sup_{\substack{\|u\|=1 \\ \zeta,\eta\in\widetilde{B}(\bar{\xi},R)}} \{|\nabla_u W^{\zeta,\eta}(u)| + |\nabla_\zeta W^{\zeta,\eta}(u)|\} \frac{c}{d_{\widetilde{X}}(\xi_0,\eta)^{Q-\ell+1}}.\end{aligned}$$

Analogously,

$$|\widetilde{X}_0 K(\,\cdot\,,\eta)(\zeta)| \le \sup_{\substack{\|u\|=1 \\ \zeta,\eta\in\widetilde{B}(\bar{\xi},R)}} \{|\nabla_u W^{\zeta,\eta}(u)| + |\nabla_\zeta W^{\zeta,\eta}(u)|\} \frac{c}{d_{\widetilde{X}}(\xi_0,\eta)^{Q-\ell+2}},$$

hence

$$|K(\xi,\eta) - K(\xi_0,\eta)| \le C \frac{d_{\widetilde{X}}(\xi,\xi_0)}{d_{\widetilde{X}}(\xi_0,\eta)^{Q-\ell+1}}$$

with

$$C = c \sup_{\substack{\|u\|=1 \\ \zeta,\eta\in\widetilde{B}(\bar{\xi},R)}} \{|\nabla_u W^{\zeta,\eta}(u)| + |\nabla_\zeta W^{\zeta,\eta}(u)|\}.$$

To get the analogous bound for $|K(\eta,\xi_0) - K(\eta,\xi)|$, it is enough to apply the previous estimate to the function

$$\widetilde{K}(\xi,\eta) = \widetilde{W}^{\xi,\eta}(\Theta(\eta,\xi)) \quad \text{with } \widetilde{W}^{\xi,\eta}(u) = W^{\eta,\xi}(u^{-1}).$$

This completes the proof of (ii).

To prove (iii), we first ignore the dependence on the parameters ξ, η, and then we will show how to modify our argument to take them into account. By the change of variables $u = \Theta(\eta, \xi)$, Theorem 3.3(c) gives

$$\int_{\Omega_{k+1}, \varepsilon_1 < \rho(\xi,\eta) < \varepsilon_2} W(\Theta(\eta, \xi))\, d\eta = c(\xi) \int_{\varepsilon_1 < \|u\| < \varepsilon_2} W(u)(1 + \omega(\xi, u)))\, du,$$

which, by the vanishing property of W, equals

$$c(\xi) \int_{\varepsilon_1 < \|u\| < \varepsilon_2} W(u)\omega(\xi, u)\, du.$$

Then

$$\begin{aligned}\left| \int_{\Omega_{k+1}, \varepsilon_1 < \rho(\xi,\eta) < \varepsilon_2} W(\Theta(\eta, \xi))\, d\eta \right| &\le c \cdot \int_{\varepsilon_1 < \|u\| < \varepsilon_2} |W(u)|\, \|u\|\, du \\ &\le c \cdot \sup_{\|u\|=1} |W| \cdot \int_{\varepsilon_1 < \|u\| < \varepsilon_2} \|u\|^{1-Q}\, du \le c \cdot \sup_{\|u\|=1} |W| \cdot (\varepsilon_2 - \varepsilon_1).\end{aligned}$$

Analogously, one can prove the bound on $W(\Theta(\xi, \eta))$. Now, to keep track of the possible dependence of W on the parameters ξ, η, let us write

$$\begin{aligned}&\int_{\Omega_{k+1}, \varepsilon_1 < \rho(\xi,\eta) < \varepsilon_2} W^{\xi,\eta}(\Theta(\eta, \xi))\, d\eta \\ &\qquad = \int_{\Omega_{k+1}, \varepsilon_1 < \rho(\xi,\eta) < \varepsilon_2} W^{\xi,\xi}(\Theta(\eta, \xi))\, d\eta + \int_{\Omega_{k+1}, \varepsilon_1 < \rho(\xi,\eta) < \varepsilon_2} [W^{\xi,\eta}(\Theta(\eta, \xi)) - W^{\xi,\xi}(\Theta(\eta, \xi))\,]\, d\eta \\ &\qquad =: I + II.\end{aligned}$$

The term I can be bounded as above, while

$$|W^{\xi,\eta}(u) - W^{\xi,\xi}(u)| \le |\xi - \eta||\nabla_\eta W^{\xi,\eta'}(u)|$$

for some point η' near ξ and η. Recalling again that the function $\nabla_\eta W^{\xi,\eta'}(\,\cdot\,)$ has the same homogeneity as $W^{\xi,\eta'}(\,\cdot\,)$, while

$$|\xi - \eta| \le c d_{\widetilde{X}}(\xi, \eta) \le c\rho(\xi, \eta),$$

we have

$$|II| \le c \sup_{\substack{\|u\|=1 \\ \xi,\eta \in \widetilde{B}(\bar{\xi}, R)}} |\nabla_\eta W^{\xi,\eta}(u)| \int_{\Omega_{k+1}, \varepsilon_1 < \|u\| < \varepsilon_2} \|u\|^{1-Q}\, du$$

and the same reasoning as above applies. This proves the bound on $|\int K(\xi, \eta)\, d\eta|$ in (5-5). The proof of the bound on $|\int K(\eta, \xi)\, d\eta|$ is analogous, since the vanishing property (5-4) also implies the same bound for $\int_{r < \|u\| < R} W^{\xi,\eta}(u^{-1})\, du$. □

Proposition 5.2 implies that $D_1\Gamma(\xi_0; \Theta(\eta, \xi))$ satisfies the standard estimates, cancellation property, and convergence condition stated in Section 3C. Note that (5-5) implies both the cancellation property and the convergence condition.

In order to apply [Bramanti and Zhu 2012, Theorem 5.4] to the kernel $k_S(\xi, \eta)$, we still need to prove that the singular integral T with kernel $k_S(\xi, \eta)$ satisfies a condition $T(1) \in C^{\gamma}_{\widetilde{X}}$. This result is more delicate than the previous conditions, and is contained in the following.

Proposition 5.3. *Let*

$$\widetilde{h}(\xi) := \lim_{\varepsilon \to 0} \int_{\rho(\xi,\eta)>\varepsilon} \widetilde{K}(\xi, \eta)\, d\eta$$

with

$$\widetilde{K}(\xi, \eta) = a_1(\xi) b_1(\eta) D_1^{\xi,\eta} \Gamma(\xi_0; \Theta(\eta, \xi)),$$

$D_1^{\xi,\eta}\Gamma(\xi_0; \cdot\,)$ *homogeneous of degree $-Q$ and satisfying the vanishing property*

$$\int_{r<\|u\|<R} D_1^{\xi,\eta} \Gamma(\xi_0; u)\, du = 0 \quad \textit{for every } R > r > 0, \textit{ any } \xi, \eta \in \widetilde{B}(\bar{\xi}, R).$$

Then

$$\widetilde{h} \in C^{\gamma}_{\widetilde{X}}(\widetilde{B}(\bar{\xi}, R)) \quad \textit{for any } \gamma \in (0, 1). \tag{5-6}$$

Proof. Since a_1, b_1 are compactly supported in $\widetilde{B}(\bar{\xi}, R)$, we can choose a radial cutoff function

$$\phi(\xi, \eta) = f(\rho(\xi, \eta))$$

with

$$f(\|u\|) = 1 \quad \text{for } \|u\| \le R, \qquad f(\|u\|) = 0 \quad \text{for } \|u\| \ge 2R,$$

so that $\widetilde{K}(\xi, \eta) = \widetilde{K}(\xi, \eta)\phi(\xi, \eta)$. To begin with, let us prove the assertion without taking into consideration the dependence of $D_1^{\xi,\eta}\Gamma(\xi_0; u)$ on ξ, η. Then

$$\begin{aligned} \widetilde{h}(\xi) = a_1(\xi) b_1(\xi) \lim_{\varepsilon \to 0} \int_{\rho(\xi,\eta)>\varepsilon} & \phi(\xi, \eta) D_1 \Gamma(\xi_0; \Theta(\eta, \xi))\, d\eta \\ & + a_1(\xi) \int \phi(\xi, \eta) D_1 \Gamma(\xi_0; \Theta(\eta, \xi))[b_1(\eta) - b_1(\xi)]\, d\eta \\ =: I(\xi) + II(\xi). & \end{aligned}$$

Now,

$$\begin{aligned} I(\xi) &= a_1(\xi) b_1(\xi) c(\xi) \lim_{\varepsilon \to 0} \int_{\|u\|>\varepsilon} f(\|u\|) D_1 \Gamma(\xi_0; u)(1 + \omega(\xi, u))\, du \\ &= a_1(\xi) b_1(\xi) c(\xi) \int f(\|u\|) D_1 \Gamma(\xi_0; u) \omega(\xi, u)\, du, \end{aligned}$$

by the vanishing property, with ω smoothly depending on ξ and uniformly bounded by $c\|u\|$. Hence $I(\xi)$ is Lipschitz continuous and, in particular, Hölder continuous of any exponent $\gamma \in (0, 1)$. Moreover,

$$II(\xi) = a_1(\xi) \int_{\widetilde{B}(\bar{\xi}, R)} \kappa(\xi, \eta)\, d\eta \quad \text{with } \kappa(\xi, \eta) = \phi(\xi, \eta) D_1 \Gamma(\xi_0; \Theta(\eta, \xi))[b_1(\eta) - b_1(\xi)].$$

It is not difficult to check that the kernel $\kappa(\xi, \eta)$ satisfies the standard estimates of *fractional* integrals (3-11) and (3-12) for any $\nu \in (0, 1)$ (actually, for $\nu = 1$). Hence, by [Bramanti and Zhu 2012, Theorem 5.8], the operator with kernel κ is continuous on $C^\gamma(\widetilde{B}(\bar{\xi}, R))$; in particular, it maps the function 1 into $C^\gamma(\widetilde{B}(\bar{\xi}, R))$, which proves that $II(\xi)$ is Hölder continuous.

To conclude the proof, we have to show how to take into account the possible dependence of $D_1^{\xi,\eta}\Gamma(\xi_0; u)$ on ξ, η. Let us start with the η dependence.

$$\begin{aligned}\widetilde{h}(\xi) = a_1(\xi)b_1(\xi) \lim_{\varepsilon\to 0}\int_{\rho(\xi,\eta)>\varepsilon} &\phi(\xi,\eta) D_1^\eta \Gamma(\xi_0; \Theta(\eta,\xi))\, d\eta \\ &+ a_1(\xi)\int \phi(\xi,\eta) D_1^\eta\Gamma(\xi_0;\Theta(\eta,\xi))[b_1(\eta)-b_1(\xi)]\, d\eta \\ =: I'(\xi) + II'(\xi).&\end{aligned}$$

The term $II'(\xi)$ can be handled as the term $II(\xi)$ above. As to $I'(\xi)$,

$$\begin{aligned}I'(\xi) = a_1(\xi)b_1(\xi)c(\xi)\lim_{\varepsilon\to 0}\int_{\|u\|>\varepsilon} & f(\|u\|) D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0; u)\, du \\ &+ a_1(\xi)b_1(\xi)c(\xi)\int f(\|u\|) D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0; u)\omega(\xi, u)\, du.\end{aligned}$$

The second term can be handled as above, while the first one requires some care. By the vanishing property of $D_1^\zeta\Gamma(\xi_0; u)$ for any fixed ζ, we can write

$$\lim_{\varepsilon\to 0}\int_{\|u\|>\varepsilon} f(\|u\|) D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0; u)\, du = \lim_{\varepsilon\to 0}\int_{\|u\|>\varepsilon} f(\|u\|)[D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0; u) - D_1^\xi\Gamma(\xi_0; u)]\, du.$$

On the other hand,

$$D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0; u) = D_1^\xi\Gamma(\xi_0; u) + D_0^\xi\Gamma(\xi_0; u),$$

where D_0^ξ is a vector field of local weight ≤ 0, smoothly depending on ξ. Hence

$$\lim_{\varepsilon\to 0}\int_{\|u\|>\varepsilon} f(\|u\|) D_1^{\Theta(\cdot,\xi)^{-1}(u)}\Gamma(\xi_0; u)\, du = \int f(\|u\|) D_0^\xi\Gamma(\xi_0; u)\, du,$$

which can be handled as the term $I(\xi)$ above.

Dependence on the variable ξ can be taken into account as follows. If

$$\begin{aligned}\widetilde{h}(\xi) &= a_1(\xi)b_1(\xi)\lim_{\varepsilon\to 0}\int_{\rho(\xi,\eta)>\varepsilon}\phi(\xi,\eta) D_1^{\xi,\eta}\Gamma(\xi_0; \Theta(\eta,\xi))\, d\eta \\ &\equiv \lim_{\varepsilon\to 0}\int F_\varepsilon(\xi,\xi,\eta) \quad \text{with } F_\varepsilon(\zeta,\xi,\eta) = a_1(\xi)b_1(\xi)\chi_{\rho(\xi,\eta)>\varepsilon}(\eta)\phi(\xi,\eta) D_1^{\zeta,\eta}\Gamma(\xi_0;\Theta(\eta,\xi))\, d\eta,\end{aligned}$$

then

$$\begin{aligned}\widetilde{h}(\xi_1) - \widetilde{h}(\xi_2) &= \lim_{\varepsilon\to 0}\int [F_\varepsilon(\xi_1,\xi_1,\eta) - F_\varepsilon(\xi_2,\xi_1,\eta)]\, d\eta + \lim_{\varepsilon\to 0}\int [F_\varepsilon(\xi_2,\xi_1,\eta) - F_\varepsilon(\xi_2,\xi_2,\eta)]\, d\eta \\ &=: A(\xi_1,\xi_2) + B(\xi_1,\xi_2).\end{aligned}$$

Now,

$$|A(\xi_1,\xi_2)| \le c\rho(\xi_1,\xi_2)$$

by the smoothness of $\xi \mapsto D_1^{\xi,\eta}\Gamma(\xi_0; u)$. As to $B(\xi_1, \xi_2)$, it is enough to apply the previous reasoning to $D_1^{\zeta,\eta}\Gamma(\xi_0; \Theta(\eta, \xi))$, for any fixed ζ, to conclude that

$$\left|\lim_{\varepsilon\to 0}\int [F(\zeta, \xi_1, \eta) - F(\zeta, \xi_2, \eta)]\, d\eta\right| \le c\rho(\xi_1, \xi_2)^\gamma$$

for some constant uniformly bounded in ζ, and then apply this inequality taking $\zeta = \xi_2$. □

Conclusion of the proof of Theorem 5.1. Recall that a frozen operator of type zero is written as

$$T(\xi_0) f(\xi) = \mathrm{PV}\int_{\widetilde{B}} k(\xi_0; \xi, \eta)\, f(\eta)\, d\eta + \alpha(\xi_0, \xi) f(\xi),$$

where α is a bounded measurable function, smooth in ξ. The multiplicative part

$$f(\xi) \longmapsto \alpha(\xi_0, \xi) f(\xi)$$

clearly maps $C^\alpha_{\widetilde{X}}$ in $C^\alpha_{\widetilde{X}}$, since $\alpha(\xi_0, \cdot)$ is smooth, with operator norm bounded by some constant depending on the vector fields and the ellipticity constant μ, by Theorem 4.3.

Let us now consider the integral part. With the notation introduced at the beginning of this section, let us consider first

$$k_S(\xi, \eta) = a_1(\xi) b_1(\eta) D_1^{\xi,\eta}\Gamma(\xi_0; \Theta(\eta, \xi)),$$

where $D_1^{\xi,\eta}\Gamma(\xi_0; u)$ is homogeneous of degree $-Q$ and satisfies the vanishing property (5-4). By Proposition 5.2, $k_S(\xi, \eta)$ satisfies conditions (i), (ii), and (iii) in Section 3C, with constants bounded by

$$c \sup_{\|u\|=1} \{|D^2\Gamma(\xi_0, u)| + |D^3\Gamma(\xi_0, u)|\}, \tag{5-7}$$

where the symbols D^2, D^3 denote standard derivatives of orders 2, 3, respectively, with respect to u, and the constant c depends on the vector fields but not on the point ξ_0. By Proposition 5.3, condition (5-6) is also satisfied by $k_S(\xi, \eta)$, with the $C^\gamma_{\widetilde{X}}$ norm bounded by a quantity of the kind (5-7). Hence, by [Bramanti and Zhu 2012, Theorem 5.4], the operator with kernel $k_S(\xi, \eta)$ satisfies the assertion of the theorem we are proving, with a constant bounded by a quantity like (5-7). In turn, by Theorem 4.3, this quantity can be bounded by a constant depending on the vector fields and the ellipticity constant μ of the matrix $a_{ij}(x)$.

Let us now come to the kernel

$$k_F(\xi, \eta) = \left\{\sum_{i=2}^{H} a_i(\xi) b_i(\eta) D_i^{\xi,\eta}\Gamma(\xi_0; \cdot) + a_0(\xi) b_0(\eta) D_0^{\xi,\eta}\Gamma(\xi_0; \cdot)\right\}(\Theta(\eta, \xi)),$$

where each function $D_i^{\xi,\eta}\Gamma(\xi_0; u)$ $(i = 2, 3, \ldots, H)$ is homogeneous of some degree $\ge 1 - Q$, while $D_0^{\xi,\eta}\Gamma(\xi_0; u)$ is bounded and smooth. By Proposition 5.2, each kernel

$$a_i(\xi) b_i(\eta) D_i^{\xi,\eta}\Gamma(\xi_0; \Theta(\eta, \xi))$$

satisfies the standard estimates (i) in Section 3C for some $\nu > 0$, hence we can apply [Bramanti and Zhu 2012, Theorem 5.8] to the integral operators defined by these kernels, and conclude as above. Finally, the integral operator with regular kernel is clearly C^{γ} continuous. □

5B. L^p *continuity of variable operators of type* 0 *and their commutators.* In this subsection we are going to prove the following.

Theorem 5.4. *Let T be a variable operator of type* 0 (*see Section 4B*) *over the ball* $\widetilde{B}(\bar{\xi}, R)$, *and* $p \in (1, \infty)$. *Then*

(i) *there exists $c > 0$, depending on p, R, $\{\widetilde{X}_i\}_{i=0}^q$, and μ such that*

$$\|Tu\|_{L^p(\widetilde{B}(\bar{\xi},r))} \leq c\|u\|_{L^p(\widetilde{B}(\bar{\xi},r))}$$

for every $u \in L^p(\widetilde{B}(\bar{\xi}, r))$ and $r \leq R$;

(ii) *for every $a \in \mathrm{VMO}_{X,\mathrm{loc}}(\Omega)$, any $\varepsilon > 0$, there exists $r \leq R$ such that, for every $u \in L^p(\widetilde{B}(\bar{\xi}, r))$,*

$$\|T(\tilde{a}u) - \tilde{a} \cdot Tu\|_{L^p(\widetilde{B}(\bar{\xi},r))} \leq \varepsilon\|u\|_{L^p(\widetilde{B}(\bar{\xi},r))}, \tag{5-8}$$

where $\tilde{a}(x, h) = a(x)$. The number r depends on p, R, $\{\widetilde{X}_i\}_{i=0}^q$, μ, $\eta^*_{a,\Omega',\Omega}$, and ε (see Section 3D.3 for the definition of $\mathrm{VMO}_{X,\mathrm{loc}}(\Omega)$ *and* $\eta^*_{a,\Omega',\Omega}$).

A basic difference between the context here and that of the previous section is that here we are considering *variable* kernels and operators of type zero. To reduce the study of these operators to that of constant operators of type zero we will make use of the classical technique of expansion in series of spherical harmonics, as already done in [Bramanti and Brandolini 2000a].

Proof. This proof is similar to that of [Bramanti and Brandolini 2000a, Theorem 2.11]. Recall that a variable operator of type zero is written as

$$Tf(\xi) = \mathrm{PV}\int_{\widetilde{B}} k(\xi; \xi, \eta) f(\eta)\, d\eta + \alpha(\xi, \xi) f(\xi),$$

where $\alpha(\xi_0, \xi)$ is a bounded measurable function in ξ_0, smooth in ξ. The multiplicative part

$$f(\xi) \longmapsto \alpha(\xi, \xi) f(\xi)$$

clearly maps L^p into L^p, with operator norm bounded by some constant depending on the vector fields and the ellipticity constant μ, by Theorem 4.3. Moreover, this part does not affect the commutator of T.

As to the integral part of T, let us split the variable kernel as

$$k(\xi; \xi, \eta) = k'(\xi; \xi, \eta) + k''(\xi; \xi, \eta).$$

Like in the previous section, it is enough to prove our result for the kernel k'. Let us expand it as

$$\begin{aligned} k'(\xi; \xi, \eta) &= \sum_{i=1}^{H} a_i(\xi) b_i(\eta) D_i^{\xi,\eta}\Gamma(\xi; \Theta(\eta, \xi)) + a_0(\xi) b_0(\eta) D_0^{\xi,\eta}\Gamma(\xi; \Theta(\eta, \xi)) \\ &=: k_U(\xi; \xi, \eta) + k_B(\xi; \xi, \eta), \end{aligned}$$

where the kernels $D_i^{\xi,\eta}\Gamma(\xi;u)$ (for $i=1,2,3,\dots,H$) are homogeneous of some degree $\geq -Q$, $D_i^{\xi,\eta}\Gamma(\xi;u)$ satisfies the cancellation property, and $D_0^{\xi,\eta}\Gamma(\xi;u)$ is bounded in u and smooth in ξ,η. The kernels k_U and k_B are "unbounded" and "bounded", respectively.

The operator with kernel k_B is obviously L^p continuous. Moreover, it satisfies the commutator estimate (5-8) by [Bramanti and Zhu 2012, Theorem 7.3], since

$$|k_B(\xi;\xi,\eta)| \leq ca_0(\xi)b_0(\eta)$$

and the constant function 1 obviously satisfies the standard estimates (3-11), (3-12) with $\nu=1$.

To handle the kernel k_U, we expand each of its terms in series of spherical harmonics, exactly like in [Bramanti and Brandolini 2000a, Section 2.4]:

$$D_i^{\xi,\eta}\Gamma(\xi;u)=\sum_{m=0}^{\infty}\sum_{k=1}^{g_m} c_{i,km}^{\xi,\eta}(\xi)K_{i,km}(u),$$

where $K_{i,km}(u)$ are homogeneous kernels which, on the sphere $\|u\|=1$, coincide with the spherical harmonics, and $c_{i,km}^{\xi,\eta}(\cdot)$ are the corresponding Fourier coefficients.

Let us first prove the assertion without taking into account the dependence of the coefficients $c_{i,km}^{\xi,\eta}(\xi)$ on η. Then the operator with kernel k_U can be written as

$$Sf(\xi)=\sum_{m=0}^{\infty}\sum_{k=1}^{g_m} c_{i,km}^{\xi}(\xi)S_{i,km}f(\xi) \tag{5-9}$$

with

$$S_{i,km}f(\xi)=a_i(\xi)\int_{\widetilde{B}} b_i(\eta)K_{i,km}(\Theta(\eta,\xi))f(\eta)\,d\eta.$$

The number g_m in (5-9) is the dimension of the space of spherical harmonics of degree m in $\mathbb{R}^N$; it is known that

$$g_m \leq c(N)\cdot m^{N-2} \quad \text{for every } m=1,2,\dots. \tag{5-10}$$

For every $p\in(1,\infty)$ we can write

$$\|Sf\|_{L^p(\widetilde{B}(\bar{\xi},r))} \leq \sum_{m=0}^{\infty}\sum_{k=1}^{g_m}\|c_{i,km}^{\cdot}(\cdot)\|_{L^\infty(\widetilde{B}(\bar{\xi},r))}\|S_{i,km}f\|_{L^p(\widetilde{B}(\bar{\xi},r))}$$

and

$$\|S(\tilde{a}f)-\tilde{a}\cdot Sf\|_{L^p(\widetilde{B}(\bar{\xi},r))} \leq \sum_{m=0}^{\infty}\sum_{k=1}^{g_m}\|c_{i,km}^{\cdot}(\cdot)\|_{L^\infty(\widetilde{B}(\bar{\xi},r))}\|S_{i,km}(\widetilde{a}f)-\tilde{a}\cdot S_{i,km}f\|_{L^p(\widetilde{B}(\bar{\xi},r))}.$$

Now each $S_{i,km}$ is a frozen operator of type $\lambda\geq 0$, and the same arguments as in the previous section show that the kernel of $S_{i,km}$ satisfies the assumptions (i), (ii), and (iii) in Section 3C with constants bounded by

$$c\cdot\sup_{\|u\|=1}|\nabla_u K_{km}(u)|,$$

(with c depending on the vector fields); in turn, by known properties of spherical harmonics, we have

$$\sup_{\|u\|=1} |\nabla_u K_{km}(u)| \le c(N) m^{N/2},$$

so that, by [Bramanti and Zhu 2012, Theorems 5.3 and 5.7], we conclude as in [Bramanti and Brandolini 2000a, p. 807] that

$$\|S_{i,km} f\|_{L^p(\widetilde{B}(\bar{\xi},r))} \le c \cdot m^{N/2} \|f\|_{L^p(\widetilde{B}(\bar{\xi},r))} \quad \text{for } i = 1, 2, \dots, H,$$

where we have also taken into account Remark 5.5 below.

Analogously, applying [Bramanti and Zhu 2012, Theorems 7.1 and 7.2], we have the commutator estimate

$$\|S_{i,km}(\tilde{a} f) - \tilde{a} \cdot S_{i,km} f\|_{L^p(\widetilde{B}(\bar{\xi},r))} \le \varepsilon \cdot m^{N/2} \|f\|_{L^p(\widetilde{B}(\bar{\xi},r))} \quad \text{for } i = 1, 2, \dots, H,$$

for any $\varepsilon > 0$, provided r is small enough, depending on ε and $\eta^*_{\tilde{a},\Omega_{k+2},\Omega_{k+3}}$ (see (5-2) and Definition 3.21 for the meaning of symbols). Then, by Proposition 3.23, the constant r depends on the function a only through the local VMO modulus $\eta^*_{a,\Omega',\Omega}$.

Next, again by known properties of spherical harmonics, we can say that, for any positive integer h, there exists c_h such that

$$|c^{\zeta}_{i,km}(\xi)| \le c_h \cdot m^{-2h} \sup_{\|u\|=1, |\beta|=2h} \left| \left(\frac{\partial}{\partial u}\right)^{\beta} D^{\zeta}_i \Gamma(\xi; u) \right|.$$

By the uniform estimates contained in Theorem 4.3, the last expression is bounded by Cm^{-2h}, for some constant C depending on h, the vector fields, and the ellipticity constant μ. Also taking into account (5-10) and choosing h large enough, we conclude

$$\|Sf\|_{L^p(\widetilde{B}(\bar{\xi},r))} \le \sum_{m=0}^{\infty} C g_m m^{-2h} m^{N/2} \|f\|_{L^p(\widetilde{B}(\bar{\xi},r))} = c\|f\|_{L^p(\widetilde{B}(\bar{\xi},r))}$$

and

$$\|S(\tilde{a} f) - \tilde{a} \cdot Sf\|_{L^p(\widetilde{B}(\bar{\xi},r))} \le c\varepsilon \|f\|_{L^p(\widetilde{B}(\bar{\xi},r))}$$

for any $\varepsilon > 0$, provided r is small enough.

We are left to show how the previous argument needs to be modified to take into account the possible dependence of $D^{\xi,\eta}_i \Gamma(\xi; u)$ (and then of $c^{\xi,\eta}_{i,km}(\xi)$) on η. Let us expand

$$D^{\zeta,\Theta(\cdot,\zeta)^{-1}(u)}_i \Gamma(\xi; u) = \sum_{m=0}^{\infty} \sum_{k=1}^{g_m} c^{\zeta}_{i,km}(\xi) K_{i,km}(u)$$

so that

$$D^{\zeta,\eta}_i \Gamma(\xi; \Theta(\eta,\zeta)) = \sum_{m=0}^{\infty} \sum_{k=1}^{g_m} c^{\zeta}_{i,km}(\xi) K_{i,km}(\Theta(\eta,\zeta)).$$

The kernels $K_{i,km}$ are the same as above, hence the estimates on the operators $S_{i,km}$ and their commutators remain unchanged. As to the coefficients $c^{\zeta}_{i,km}(\xi)$, we now have to write, for any positive integer h and some constant c_h,

$$|c^{\zeta}_{i,km}(\xi)| \le c_h \cdot m^{-2h} \sup_{\|u\|=1,|\beta|=2h} \left| \left(\frac{\partial}{\partial u}\right)^{\beta} (D_i^{\zeta,\Theta(\cdot,\zeta)^{-1}(u)}\Gamma(\xi;u)) \right|.$$

Now, from the identity

$$\frac{\partial}{\partial u_j}(D_i^{\zeta,\Theta(\cdot,\zeta)^{-1}(u)}\Gamma(\xi;u)) = \frac{\partial}{\partial u_j}(D_i^{\zeta,\eta}\Gamma(\xi;u))_{/\eta=\Theta(\cdot,\zeta)^{-1}(u)} + \sum_l \frac{\partial}{\partial \eta_l}(D_i^{\zeta,\eta}\Gamma(\xi;u)) \frac{\partial}{\partial u_j}(\Theta(\cdot,\zeta)^{-1}(u))_l,$$

it is easy to see that we can still get a uniform bound of the kind

$$|c^{\zeta}_{i,km}(\xi)| \le C \cdot m^{-2h}$$

with C depending on h, the vector fields, and the ellipticity constant μ. □

Remark 5.5. In the statements of all the theorems about singular integrals proved in [Bramanti and Zhu 2012], the constant depends on the kernel only through the constants involved in the assumptions. Actually, we need some additional information about this dependence. A standard sublinearity argument allows us to say that if, for example, our assumptions on the kernel are (3-11), (3-12), and (3-13), the constant in our upper bound will have the form

$$c \cdot (A + B + C),$$

where A, B, and C are the constants appearing in (3-11), (3-12), and (3-13), and c does not depend on the kernel. This fact has been used in the above proof and will be used again.

6. Schauder estimates

We are now in position to apply all the machinery presented in the previous sections to prove our main results, that is, C^{α}_X and L^p estimates on X_iX_ju in terms of u and $\mathcal{L}u$. We will prove C^{α}_X estimates (Theorem 2.1) in this section, and L^p estimates (Theorem 2.2) in Section 7.

Let us recall the setting described at the end of Section 3C. For a fixed subdomain $\Omega' \Subset \Omega \subset \mathbb{R}^n$ and a fixed point $\bar{x} \in \Omega'$, let us consider a lifted ball $\widetilde{B}(\bar{\xi}, R) \subset \mathbb{R}^N$ (with $\bar{\xi} = (\bar{x}, 0)$) where the lifted vector fields $\widetilde{X}_i$ are defined and satisfy Hörmander's condition and the map Θ is defined and satisfies the properties stated in Section 3A.

According to the procedure followed in [Bramanti and Brandolini 2007, Section 5], the proof of C^{α}_X a priori estimates for second order derivatives will proceed in three steps: first, in the space of lifted variables, for test functions supported in a ball $\widetilde{B}(\bar{\xi}, r)$ with r small enough; then for any function in $C^{2,\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi}, r))$ (not necessarily vanishing at the boundary); then for any function in $C^{2,\alpha}_X(B(\bar{x}, r))$, that is in the original space.

The first step in the proof of Schauder estimates is contained in the following.

Theorem 6.1. *Let $\widetilde{B}(\bar{\xi}, R)$ be as before. Then there exist $R_0 < R$ and $c > 0$ such that, for every $u \in C^{2,\alpha}_{\widetilde{X},0}(\widetilde{B}(\bar{\xi}, R_0))$,*

$$\|u\|_{C^{2,\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R_0))} \le c\{\|\widetilde{\mathcal{L}}u\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R_0))} + \|u\|_{L^\infty(\widetilde{B}(\bar{\xi},R_0))}\},$$

where c and R_0 depend on R, $\{\widetilde{X}_i\}$, α, μ, and $\|\tilde{a}_{ij}\|_{C^\alpha(\widetilde{B}(\bar{\xi},R))}$.

The proof is quite similar to that of [Bramanti and Brandolini 2007, Theorem 5.2] and will be omitted. We just point out the facts which it relies upon:

- the representation formula proved in Theorem 4.19;
- Theorem 5.1 about singular integrals on $C^{\alpha}_{\widetilde{X}}$;
- several properties of $C^{2,\alpha}_{\widetilde{X}}$ functions, collected in Proposition 3.14.

The second step in the proof of Schauder estimates consists in establishing a priori estimates for functions not necessarily compactly supported.

Theorem 6.2. *There exist $r_0 < R_0$ and $c, \beta > 0$ (with R_0 as in Theorem 6.1) such that, for every $u \in C^{2,\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi}, r_0))$, $0 < t < s < r_0$,*

$$\|u\|_{C^{2,\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},t))} \le \frac{c}{(s-t)^\beta}\{\|\widetilde{\mathcal{L}}u\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},s))} + \|u\|_{L^\infty(\widetilde{B}(\bar{\xi},s))}\},$$

where r_0, c depend on R, $\{\widetilde{X}_i\}_{i=1}^q$, α, μ, and $\|\tilde{a}_{ij}\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R))}$ and β depends on $\{\widetilde{X}_i\}_{i=0}^q$ and α.

As in [Bramanti and Brandolini 2007], this result relies on interpolation inequalities for $C^{k,\alpha}_{\widetilde{X}}$ norms and the use of suitable cutoff function. The following result can be proved as [Bramanti and Brandolini 2007, Lemma 6.2] by the results in Proposition 3.14.

Lemma 6.3 (cutoff functions). *For any $0 < \rho < r$ and $\xi \in \widetilde{B}(\bar{\xi}, R)$, there exists $\varphi \in C_0^\infty(\mathbb{R}^N)$ with the following properties.*

(i) $0 \leqslant \varphi \leqslant 1$, $\varphi \equiv 1$ *on* $\widetilde{B}(\xi, \rho)$, *and* $\operatorname{sprt}\varphi \subseteq \widetilde{B}(\xi, r)$.

(ii) *For* $i, j = 1, 2, \ldots, q$,

$$|\widetilde{X}_i\varphi| \leqslant \frac{c}{r-\rho};\ |\widetilde{X}_0\varphi|, |\widetilde{X}_i\widetilde{X}_j\varphi| \leqslant \frac{c}{(r-\rho)^2}. \tag{6-1}$$

(iii) *For any* $f \in C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi}, R))$ *and* $r - \rho$ *small enough,*

$$\begin{aligned} \|f\,\widetilde{X}_i\varphi\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R))} &\leqslant \frac{c}{(r-\rho)^2}\|f\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R))},\\ \|f\,\widetilde{X}_0\varphi\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R))}, \|f\widetilde{X}_i\widetilde{X}_j\varphi\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R))} &\leqslant \frac{c}{(r-\rho)^3}\|f\|_{C^{\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R))}. \end{aligned} \tag{6-2}$$

We will write

$$\widetilde{B}_\rho(\xi) \prec \varphi \prec \widetilde{B}_r(\xi)$$

to indicate that φ satisfies all the previous properties.

Proposition 6.4 (interpolation inequality for test functions). *Let*

$$H = \sum_{i=1}^{q} \widetilde{X}_i^2 + \widetilde{X}_0$$

and let $\widetilde{B}(\bar{\xi}, R)$ be as before. Then, for every $\alpha \in (0,1)$, there exist constants $\gamma \geq 1$ and $c > 0$, depending on α, R and $\{\widetilde{X}_i\}$, such that, for every $\varepsilon \in (0,1)$ and every $f \in C_0^\infty(\widetilde{B}(\bar{\xi}, R/2))$,

$$\|\widetilde{X}_l f\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R/2))} \leq \varepsilon \|Hf\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R/2))} + \frac{c}{\varepsilon^\gamma} \|f\|_{L^\infty(\widetilde{B}(\bar{\xi},R/2))} \tag{6-3}$$

for $l = 1, 2, \ldots, q$; moreover, we have

$$\|Df\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R/2))} \leq \varepsilon \|\widetilde{\mathscr{L}} f\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R/2))} + \frac{c}{\varepsilon^\gamma} \|f\|_{L^\infty(\widetilde{B}(\bar{\xi},R/2))}, \tag{6-4}$$

where D is any vector field of local degree ≤ 1.

To prove Proposition 6.4, we need the following.

Lemma 6.5. *Let $P(\xi_0)$ be a frozen operator of type $\lambda \geq 1$ over $\widetilde{B}(\bar{\xi}, R)$ and $\alpha \in (0,1)$. Then there exist positive constants $\gamma > 1$ and c, depending on α, μ, and $\{\widetilde{X}_i\}$, such that, for every $f \in C_0^\infty(\widetilde{B}(\bar{\xi}, R))$ and $\varepsilon \in (0,1)$,*

$$\|PHf\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R))} \leq \varepsilon \|Hf\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R))} + \frac{c}{\varepsilon^\gamma} \|f\|_{L^\infty(\widetilde{B}(\bar{\xi},R))}. \tag{6-5}$$

Moreover, (6-5) still holds if H is replaced by any differential operator of weight two, like $\widetilde{X}_i \widetilde{X}_j$ or $\widetilde{X}_0$.

The proof of this lemma is very similar to that of [Bramanti and Brandolini 2007, Lemma 7.2]. It exploits the properties of cutoff functions (Lemma 6.3), inequality (3-19), and fractional integral estimates, relying on [Bramanti and Zhu 2012, Theorem 5.7] and Remark 5.5.

Proof of Proposition 6.4. By Theorem 4.18, we can write

$$af = PHf(\xi) + Sf,$$

where P and S are frozen operators of type 2 and 1, respectively, over $\widetilde{B}(\bar{\xi}, R)$. More precisely, they should be called "constant kernels of type 2 and 1", since they satisfy the definition of frozen kernels with the matrix $\{\tilde{a}_{ij}(\xi_0)\}$ replaced by the identity matrix.

If we assume $a = 1$ on $\widetilde{B}(\bar{\xi}, R/2)$, then, for $f \in C_0^\infty(\widetilde{B}(\bar{\xi}, R/2))$, we obtain

$$f = PHf(\xi) + Sf, \tag{6-6}$$

and therefore, by Theorem 4.11,

$$\widetilde{X}_i f = S_1 Hf(\xi) + Tf, \tag{6-7}$$

where S_1 and T are frozen operators of type 1 and 0, respectively. Substituting (6-6) in (6-7) yields

$$\widetilde{X}_i f = S_1 Hf(\xi) + TPHf + TSf,$$

and therefore, by Theorem 5.1 and Lemma 6.5,

$$\|\widetilde{X}_i f\|_\alpha \le \|S_1 Hf\|_\alpha + c\{\|PHf\|_\alpha + \|Sf\|_\alpha\} \le c\{\varepsilon\|Hf\|_\alpha + \varepsilon^{-\gamma}\|f\|_\infty + \|Sf\|_\alpha\}, \tag{6-8}$$

where all the norms are taken over $\widetilde{B}(\bar{\xi}, R/2)$. We end the proof by showing that, for an operator S of type 1,

$$\|Sf\|_\alpha \le c\|f\|_{L^\infty},$$

which by (6-8) will complete the proof of the first inequality in the proposition. Indeed, if

$$Sf(\xi) = \int_{\widetilde{B}_R} k(\xi, \eta) f(\eta)\, d\eta,$$

we have

$$|Sf(\xi_1) - Sf(\xi_2)| \le \|f\|_{L^\infty(\widetilde{B}_R)} \int_{\widetilde{B}(\bar{\xi}, R)} |k(\xi_1, \eta) - k(\xi_2, \eta)|\, d\eta. \tag{6-9}$$

Moreover,

$$\begin{aligned}\int_{\widetilde{B}_R} |k(\xi_1, \eta) - k(\xi_2, \eta)|\, d\eta &= \int_{\widetilde{B}(\bar{\xi}, R), \rho(\xi_1, \eta) > M\rho(\xi_1, \xi_2)} |k(\xi_1, \eta) - k(\xi_2, \eta)|\, d\eta \\ &\qquad + \int_{\widetilde{B}(\bar{\xi}, R), \rho(\xi_1, \eta) \le M\rho(\xi_1, \xi_2)} |k(\xi_1, \eta) - k(\xi_2, \eta)|\, d\eta \\ &=: I + II.\end{aligned}$$

Then

$$\begin{aligned} I &\le \int_{\rho(\xi_1, \eta) > M\rho(\xi_1, \xi_2)} \frac{c}{\rho(\xi_1, \eta)^{Q-1}} \frac{\rho(\xi_1, \xi_2)}{\rho(\xi_1, \eta)}\, d\eta \\ &= \rho(\xi_1, \xi_2)^\alpha \int_{\rho(\xi_1, \eta) > M\rho(\xi_1, \xi_2)} \frac{\rho(\xi_1, \eta)^{1-\alpha}}{\rho(\xi_1, \eta)^Q} \frac{\rho(\xi_1, \xi_2)^{1-\alpha}}{\rho(\xi_1, \eta)^{1-\alpha}}\, d\eta \\ &\le c\rho(\xi_1, \xi_2)^\alpha \int_{\widetilde{B}_R} \frac{\rho(\xi_1, \eta)^{1-\alpha}}{\rho(\xi_1, \eta)^Q}\, d\eta \le c\rho(\xi_1, \xi_2)^\alpha R^{1-\alpha}, \end{aligned}$$

where in the last inequality we have used the following standard computation (which will be useful again):

$$\int_{\widetilde{B}(\bar{\xi}, R), \rho(\xi_1, \eta) < r} \frac{d\eta}{\rho(\xi_1, \eta)^{Q-\beta}} \le cr^\beta \quad \text{for any } \xi_1 \in \widetilde{B}(\bar{\xi}, R). \tag{6-10}$$

As to II,

$$II \le \int_{\rho(\xi_1, \eta) \le M\rho(\xi_1, \xi_2)} |k(\xi_1, \eta)|\, d\eta + \int_{\rho(\xi_1, \eta) \le M\rho(\xi_1, \xi_2)} |k(\xi_2, \eta)|\, d\eta.$$

Since there exists $M_1 > 0$ such that if $\rho(\xi_1, \eta) \le M\rho(\xi_1, \xi_2)$, then $\rho(\xi_2, \eta) \le M_1\rho(\xi_1, \xi_2)$,

$$II \le c\left\{\int_{\rho(\xi_1, \eta) \le M\rho(\xi_1, \xi_2)} \frac{1}{\rho(\xi_1, \eta)^{Q-1}}\, d\eta + \int_{\rho(\xi_2, \eta) \le M_1\rho(\xi_1, \xi_2)} \frac{1}{\rho(\xi_2, \eta)^{Q-1}}\, d\eta\right\},$$

which, again by (6-10), is

$$\le c\rho(\xi_1, \xi_2) \le c\rho(\xi_1, \xi_2)^\alpha R^{1-\alpha}.$$

Hence, for every $\alpha \in (0, 1)$,

$$\int_{\widetilde{B}_R} |k(\xi_1, \eta) - k(\xi_2, \eta)| \, d\eta \le c_\alpha \rho(\xi_1, \xi_2)^\alpha R^{1-\alpha},$$

and, by (6-9),

$$|Sf|_\alpha \le c \|f\|_{L^\infty}.$$

Moreover,

$$|Sf(\xi)| \le \int_{\widetilde{B}_R} |k(\xi, \eta) f(\eta)| \, d\eta \le \|f\|_{L^\infty} \int_{\rho(\xi,\eta) \le cR} \frac{c}{\rho(\xi, \eta)^{Q-1}} \, d\eta \le cR \|f\|_{L^\infty},$$

hence

$$\|Sf\|_\alpha \le c \|f\|_{L^\infty}.$$

This completes the proof of (6-3). A similar argument gives (6-4). □

Theorem 6.6 (interpolation inequality). *There exist positive constants c, γ and $r_1 < R$ such that, for any $u \in C^{2,\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi}, r_1))$, $0 < \rho < r_1$, $0 < \delta < 1/3$,*

$$\|\widetilde{D}u\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},\rho))} \le \delta \sum_{i=1}^{q} \|\widetilde{D}^2 u\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}(\bar{\xi},r_1))} + \frac{c}{\delta^\gamma (r_1 - \rho)^{2\gamma}} \|u\|_{L^\infty(\widetilde{B}(\bar{\xi},r_1))},$$

where

$$\|\widetilde{D}u\| \equiv \sum_{i=1}^{q} \|\widetilde{X}_i u\| \quad \text{and} \quad \|\widetilde{D}^2 u\| \equiv \sum_{i,j=1}^{q} \|\widetilde{X}_i \widetilde{X}_i u\| + \|\widetilde{X}_0 u\|.$$

The constants c, r_1, γ depend on α, $\{\widetilde{X}_i\}$; γ is as in Proposition 6.4.

Proof. The proof can be carried out exactly as in [Bramanti and Brandolini 2007, Proposition 7.4], exploiting the properties of cutoff functions (Lemma 6.3), the interpolation inequality for test functions (Proposition 6.4), and (3-20) in Proposition 3.14. □

We are now ready to complete the second step in the proof of Schauder estimates.

Proof of Theorem 6.2. This proof can now be carried out exactly like in [Bramanti and Brandolini 2007, Theorem 5.3], exploiting Schauder estimates for functions with small support (Theorem 6.1), the properties of Hölder continuous functions contained in (3-20), (3-21), and (3-24), the properties of cutoff functions (Lemma 6.3), and the interpolation inequalities contained in Theorem 6.6 and (6-4). □

Proof of Theorem 2.1. We finally come back to our original context, which we are going to recall. We have a bounded domain Ω where our vector fields and coefficients are defined, and a fixed subdomain $\Omega' \Subset \Omega$. Fix $\bar{x} \in \Omega'$ and R such that in $B(\bar{x}, R) \subset \Omega$ all the construction of the previous two subsections (lifting to $\widetilde{B}(\bar{\xi}, R)$ and so on) can be performed. Let r_0 be as in Theorem 6.2. To begin with, we want to prove Schauder estimates for functions $u \in C^{2,\alpha}_X(B(\bar{x}, r_0))$. By Proposition 3.15 we know that the function $\widetilde{u}(x, h) = u(x)$ belongs to $C^{2,\alpha}_{\widetilde{X}}(B(\bar{\xi}, r_0))$, so we can apply to $\widetilde{u}$ Schauder estimates contained in Theorem 6.2. Combining this fact with the two estimates in Proposition 3.15 and choosing t, s such that

$$r_0 > t > s > 0 \quad \text{and} \quad t - s = r_0 - t,$$

we get, for some exponent $\omega > 2$,

$$\begin{aligned}
\|u\|_{C_X^{2,\alpha}(B(\bar{x},s))} &\le \frac{c}{(t-s)^2}\|\widetilde{u}\|_{C_{\widetilde{X}}^{2,\alpha}(\widetilde{B}(\bar{\xi},t))} \\
&\le \frac{c}{(r_0-t)^\omega}(\|\widetilde{\mathcal{L}}\widetilde{u}\|_{C_{\widetilde{X}}^{\alpha}(\widetilde{B}(\bar{\xi},r_0))} + \|\widetilde{u}\|_{L^\infty(\widetilde{B}(\bar{\xi},r_0))}) \\
&\le \frac{c}{(r_0-s)^\omega}(\|\mathcal{L}u\|_{C_X^{\alpha}(B(\bar{x},r_0))} + \|u\|_{L^\infty(B(\bar{x},r_0))}),
\end{aligned} \tag{6-11}$$

since $\widetilde{\mathcal{L}}\widetilde{u} = \widetilde{(\mathcal{L}u)}$. Next, let us choose a family of balls $B(x_i, r_0)$ in Ω such that

$$\Omega' \subset \bigcup_{i=1}^k B(x_i, r_0/2) \subset \bigcup_{i=1}^k B(x_i, r_0) \subset \Omega.$$

Then, by Proposition 3.14(v) and (6-11), with $s = r_0/2$,

$$\begin{aligned}
\|u\|_{C_X^{2,\alpha}(\Omega')} \le \|u\|_{C_X^{2,\alpha}(\cup B(x_i,r_0/2))} &\le c\sum_{i=1}^k \|u\|_{C_X^{2,\alpha}(B(x_i,r_0))} \\
&\le c\sum_{i=1}^k \{\|\mathcal{L}u\|_{C_X^{\alpha}(B(x_i,r_0))} + \|u\|_{L^\infty(B(x_i,r_0))}\} \\
&\le c\{\|\mathcal{L}u\|_{C_X^{\alpha}(\Omega)} + \|u\|_{L^\infty(\Omega)}\}
\end{aligned}$$

with c also depending on r_0. Finally, let us note that the constant c depends on the coefficients a_{ij} through the norms

$$\|\tilde{a}_{ij}\|_{C_{\widetilde{X}}^{\alpha}(\widetilde{B}(\bar{\xi},R_0))},$$

which in turn are bounded by the norms

$$\|a_{ij}\|_{C_X^{\alpha}(B(\bar{x},R_0))}$$

(by Proposition 3.15), and hence by $\|a_{ij}\|_{C_X^{\alpha}(\Omega)}$ (or more precisely, by $\|a_{ij}\|_{C_X^{\alpha}(\Omega'')}$ for some Ω'' such that $\Omega' \Subset \Omega'' \Subset \Omega$). □

7. L^p estimates

The logical structure of this section, as well as the general setting, is very similar to that of the previous one, following as closely as possible the strategy of [Bramanti and Brandolini 2000a]. The basic difference with the setting of Schauder estimates is the fact that here we start with representation formulas where the "frozen" point has finally been unfrozen; therefore, singular integrals with *variable* kernels are now involved, together with their commutators with VMO functions. This makes the singular integral part of the theory more involved.

The first step is contained in the following.

Theorem 7.1. *Let $\widetilde{B}(\bar{\xi}, R)$ be as in the previous section, and $p \in (1, \infty)$. There exists $R_0 < R$ such that, for every $u \in C_0^\infty(\widetilde{B}(\bar{\xi}, R_0))$,*

$$\|u\|_{S^{2,p}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},R_0))} \le c\{\|\widetilde{\mathcal{L}}u\|_{L^p(\widetilde{B}(\bar{\xi},R_0))} + \|u\|_{L^p(\widetilde{B}(\bar{\xi},R_0))}\} \tag{7-1}$$

for some constant c depending on $\{\widetilde{X}_i\}_{i=0}^q$, p, μ, and R; the number R_0 also depends on the local VMO moduli $\eta^*_{a_{ij},\Omega',\Omega}$.

The proof can be carried out exactly like in [Bramanti and Brandolini 2000a, Theorem 3.2], exploiting the representation formula proved in Theorem 4.21 and the results about singular integrals and commutators contained in Theorem 5.4.

Next, we have to remove the restriction to compactly supported functions.

Theorem 7.2. *Let $\widetilde{B}(\bar{\xi}, R)$ be as before. There exists $r_0 < R$ and, for any $r \le r_0$, there exists $c > 0$ such that, for any $u \in S^{2,p}_{\widetilde{X}}(\widetilde{B}(\bar{\xi}, r))$, we have*

$$\|u\|_{S^{2,p}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},r/2))} \le c\{\|\widetilde{\mathcal{L}}u\|_{L^p(\widetilde{B}(\bar{\xi},r))} + \|u\|_{L^p(\widetilde{B}(\bar{\xi},r))}\}.$$

The constants c, r_0 depend on $\{\widetilde{X}_i\}_{i=0}^q$, p, μ, R, and $\eta^*_{a_{ij},\Omega',\Omega}$; c also depends on r.

Analogously to what we have seen in Theorem 6.2, the proof of the above theorem relies on interpolation inequalities for Sobolev norms and the use of cutoff functions. Regarding cutoff functions, we need the following statement.

Lemma 7.3 (radial cutoff functions). *For any $\sigma \in (\frac{1}{2}, 1)$, $r > 0$ and $\xi \in \widetilde{B}(\bar{\xi}, r)$, there exists $\varphi \in C_0^\infty(\mathbb{R}^N)$ with the following properties.*

(i) *$\widetilde{B}_{\sigma r}(\xi) \prec \varphi \prec \widetilde{B}_{\sigma' r}(\xi)$ with $\sigma' = (1+\sigma)/2$ (this means that $\varphi = 1$ in $\widetilde{B}_{\sigma r}(\xi)$ and it is supported in $\widetilde{B}_{\sigma' r}(\xi)$).*

(ii) *For $i, j = 1, \dots, q$, we have*

$$|\widetilde{X}_i\varphi| \le \frac{c}{(1-\sigma)r}, \quad |\widetilde{X}_0\varphi|, |\widetilde{X}_i\widetilde{X}_j\varphi| \le \frac{c}{(1-\sigma)^2r^2}. \tag{7-2}$$

The above lemma, very similar to [Bramanti and Brandolini 2000a, Lemma 3.3], is actually contained in Lemma 6.3, but we prefer to state it explicitly because it is formulated in a slightly different notation, suitable to our application to L^p estimates.

Theorem 7.4 (interpolation inequality for Sobolev norms). *Let $\widetilde{B}(\bar{\xi}, R)$ be as before. For every $p \in (1, \infty)$, there exists $c > 0$ and $r_1 < R$ such that, for every $0 < \varepsilon \le 4r_1$, $u \in C_0^\infty(\widetilde{B}(\bar{\xi}, r_1))$,*

$$\|\widetilde{X}_i u\|_{L^p(\widetilde{B}(\bar{\xi},r_1))} \le \varepsilon\|Hu\|_{L^p(\widetilde{B}(\bar{\xi},r_1))} + \frac{c}{\varepsilon}\|u\|_{L^p(\widetilde{B}(\bar{\xi},r_1))} \tag{7-3}$$

for every $i = 1, \dots, q$, where $H := \sum_{i=1}^q \widetilde{X}_i^2 + \widetilde{X}_0$.

Proof. The proof of this proposition is adapted from [Bramanti and Brandolini 2000a, Theorem 3.6], but also improves that result, which is stated with a generic constant $c(\varepsilon)$ instead of c/ε.

Let r_1 be a small number to be fixed later. Like in the proof of Proposition 6.4, we can write, for any $u \in C_0^\infty(\widetilde{B}(\bar{\xi}, r_1))$ and $\xi \in \widetilde{B}(\bar{\xi}, r_1)$,

$$\widetilde{X}_i u(\xi) = SHu(\xi) + Tu(\xi),$$

where S and T are constant operators of type 1 and 0, respectively, over $\widetilde{B}(\bar{\xi}, 2r_1)$, provided $2r_1 < R$. (See the proof of Proposition 6.4 for the explanation of the term "constant operators of type λ".) Since

$$\|Tu\|_{L^p(\widetilde{B}(\bar{\xi},r_1))} \le c\|u\|_{L^p(\widetilde{B}(\bar{\xi},r_1))},$$

the result will follow if we prove that

$$\|SHu\|_{L^p(\widetilde{B}(\bar{\xi},r_1))} \le \varepsilon\|Hu\|_{L^p(\widetilde{B}(\bar{\xi},r_1))} + \frac{c}{\varepsilon}\|u\|_{L^p(\widetilde{B}(\bar{\xi},r_1))}. \tag{7-4}$$

Let $k(\xi, \eta)$ be the kernel of S, and, for any fixed $\xi \in \widetilde{B}(\bar{\xi}, r_1)$, let φ_ε be a cutoff function (as in Lemma 7.3) with

$$\widetilde{B}_{\varepsilon/2}(\xi) \prec \varphi_\varepsilon \prec \widetilde{B}_\varepsilon(\xi).$$

Let us split $SHu(\xi)$ as

$$\begin{aligned} SHu(\xi) &= \int_{\widetilde{B}(\bar{\xi},r_1),\rho(\xi,\eta)>\varepsilon/2} k(\xi,\eta)[1-\varphi_\varepsilon(\eta)]Hu(\eta)\,d\eta + \int_{\widetilde{B}(\bar{\xi},r_1),\rho(\xi,\eta)\le\varepsilon} k(\xi,\eta)Hu(\eta)\varphi_\varepsilon(\eta)\,d\eta \\ &=: I(\xi) + II(\xi). \end{aligned}$$

Then

$$\begin{aligned} |I(\xi)| &= \left| \int_{\widetilde{B}(\bar{\xi},r_1),\rho(\xi,\eta)>\varepsilon/2} H^T(k(\xi,\cdot)[1-\varphi_\varepsilon(\cdot)])(\eta)u(\eta)\,d\eta \right| \\ &\le \int_{\widetilde{B}(\bar{\xi},r_1),\rho(\xi,\eta)>\varepsilon/2} \Big\{ \big|[1-\varphi_\varepsilon]H^T k(\xi,\cdot)\big| + c\sum \big|\widetilde{X}_i[1-\varphi_\varepsilon]\cdot\widetilde{X}_j k(\xi,\cdot)\big| \\ &\qquad\qquad + \big|k(\xi,\cdot)H^T[1-\varphi_\varepsilon]|(\eta)|u(\eta)\big| \Big\}\,d\eta \\ &=: A(\xi) + B(\xi) + C(\xi). \end{aligned}$$

Recall that, for $i, j = 1, 2, \ldots, q$,

$$\begin{aligned} |k(\xi,\eta)| &\le \frac{c}{d(\xi,\eta)^{Q-1}}, \\ |\widetilde{X}_i k(\xi,\eta)| &\le \frac{c}{d(\xi,\eta)^{Q}}, \\ |H^T k(\xi,\cdot)(\eta)| &\le \frac{c}{d(\xi,\eta)^{Q+1}}, \\ |\widetilde{X}_i(1-\varphi_\varepsilon)(\eta)| &\le \frac{c}{\varepsilon}, \\ |H^T(1-\varphi_\varepsilon)(\eta)| &\le \frac{c}{\varepsilon^2}, \end{aligned}$$

and the derivatives of $(1-\varphi_\varepsilon)$ are supported in the annulus $\varepsilon/2 \le d(\xi,\eta) \le \varepsilon$. Since $\xi,\eta \in \widetilde{B}(\bar{\xi}, r_1)$, we have $d(\xi,\eta) < 2r_1$. Hence, letting k_0 be the integer such that $2^{k_0-1}\varepsilon < 2r_1 \le 2^{k_0}\varepsilon$, we have

$$\begin{aligned}
|A(\xi)| &\le c \sum_{k=0}^{k_0} \int_{2^{k-1}\varepsilon < \rho(\xi,\eta) \le 2^k\varepsilon} \frac{c}{d(\xi,\eta)^{Q+1}} |u(\eta)|\, d\eta \\
&\le c \sum_{k=0}^{k_0} \frac{1}{2^{k-1}\varepsilon} \frac{1}{(\varepsilon 2^{k-1})^Q} \int_{\rho(\xi,\eta)\le 2^k\varepsilon} |u(\eta)|\, d\eta \\
&\le \frac{c}{\varepsilon} \cdot \sup_{r \le 4r_1} \frac{1}{|\widetilde{B}(\xi,r)|} \int_{\widetilde{B}(\xi,r)} |u(\eta)|\, d\eta. \qquad (7\text{-}5)
\end{aligned}$$

We now have to recall the definition of a *local maximal function* $\mathcal{M}$ in a (metric) locally homogeneous space $(\Omega, \{\Omega_n\}, d, d\mu)$, given in [Bramanti and Zhu 2012]. Fix Ω_n, Ω_{n+1} (see Section 3C for the notation) and, for any $f \in L^1(\Omega_{n+1})$, define

$$\mathcal{M}_{\Omega_n,\Omega_{n+1}} f(x) = \sup_{r \le r_n} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f(y)|\, d\mu(y) \quad \text{for } x \in \Omega_n,$$

where $r_n = \frac{2}{5}\varepsilon_n$. Applying this definition to our situation where $4r_1 = r_n = \frac{2}{5}\varepsilon_n$, we get $\varepsilon_n = 10r_1$ and, for $\xi \in \widetilde{B}(\bar{\xi}, r_1)$, we have $\widetilde{B}(\xi, \varepsilon_n) \subset \widetilde{B}(\bar{\xi}, 11r_1)$. Therefore, by (7-5), we can write

$$|A(\xi)| \le \frac{c}{\varepsilon} \cdot \mathcal{M}_{\widetilde{B}(\bar{\xi},r_1),\widetilde{B}(\bar{\xi},11r_1)} u(\xi),$$

and, by [Bramanti and Zhu 2012, Theorem 8.3], we have

$$\|A\|_{L^p(\widetilde{B}(\bar{\xi},r_1))} \le \frac{c}{\varepsilon} \|u\|_{L^p(\widetilde{B}(\bar{\xi},11r_1))} = \frac{c}{\varepsilon} \|u\|_{L^p(\widetilde{B}(\bar{\xi},r_1))},$$

since $u \in C_0^\infty(\widetilde{B}(\bar{\xi}, r_1))$, provided $11r_1 < R$. Also,

$$\begin{aligned}
|B(\xi)| &\le c \int_{\frac{\varepsilon}{2} < \rho(\xi,\eta) \le \varepsilon,} \frac{1}{\varepsilon} \cdot \frac{1}{d(\xi,\eta)^Q} |u(\eta)|\, d\eta \le \frac{c}{\varepsilon^{Q+1}} \int_{\rho(\xi,\eta)\le\varepsilon} |u(\eta)|\, d\eta \\
&\le \frac{c}{\varepsilon} \cdot \sup_{r\le\varepsilon} \frac{1}{|\widetilde{B}(\xi,r)|} \int_{\widetilde{B}(\xi,r)} |u(\eta)|\, d\eta \le \frac{c}{\varepsilon} \cdot \mathcal{M}_{\widetilde{B}(\bar{\xi},r_1),\widetilde{B}(\bar{\xi},11r_1)} u(\xi)
\end{aligned}$$

provided $\varepsilon < 4r_1$. As before, we have

$$\|B\|_{L^p(\widetilde{B}(\bar{\xi},r_1))} \le \frac{c}{\varepsilon} \|u\|_{L^p(\widetilde{B}(\bar{\xi},r_1))}.$$

Finally,

$$|C(\xi)| \le c \int_{\varepsilon/2 < \rho(\xi,\eta) \le \varepsilon} \frac{1}{\varepsilon^2} \cdot \frac{1}{d(\xi,\eta)^{Q-1}} |u(\eta)| \eta\, dy \le \frac{c}{\varepsilon^{Q+1}} \int_{\rho(\xi,\eta)\le\varepsilon} |u(\eta)|\, d\eta.$$

Therefore, as for the term $B(\xi)$,

$$\|I\|_{L^p(\widetilde{B}(\bar{\xi},r_1))} \le \frac{c}{\varepsilon} \|u\|_{L^p(\widetilde{B}(\bar{\xi},r_1))}.$$

Let us bound II:

$$|II(\xi)| \le c \int_{\rho(\xi,\eta)\le\varepsilon} \frac{|Hu(\eta)|}{\rho(\xi,\eta)^{Q-1}}\,d\eta.$$

Then a computation similar to that of $C(\xi)$ gives

$$|II(\xi)| \le c\varepsilon \mathcal{M}_{\widetilde{B}(\bar{\xi},r_1),\widetilde{B}(\bar{\xi},11r_1)}u(\xi) \quad\text{and}\quad \|II\|_{L^p(\widetilde{B}(\bar{\xi},r_1))} \le c\varepsilon\|u\|_{L^p(\widetilde{B}(\bar{\xi},r_1))},$$

provided $\varepsilon < 4r_1$. □

Theorem 7.5. *For any* $u \in S^{2,p}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},r))$, $p\in[1,\infty)$, $0<r<r_1$ (*where* r_1 *is the number in Theorem 7.4*), *define the following quantities*:

$$\Phi_k = \sup_{1/2<\sigma<1}((1-\sigma)^k r^k \|\widetilde{D}^k u\|_{L^p(\widetilde{B}_{r\sigma})}) \quad \textit{for } k=0,1,2.$$

Then, *for any* $\delta>0$ (*small enough*),

$$\Phi_1 \le \delta\,\Phi_2 + \frac{c}{\delta}\Phi_0.$$

Proof. This result follows exactly as in [Bramanti and Brandolini 2000b, Theorem 21], exploiting the interpolation result for compactly supported functions (Theorem 7.4), cutoff functions (Lemma 7.3), and Proposition 3.19. □

Proof of Theorem 7.2. This proof is similar to that of theorem [Bramanti and Brandolini 2000b, Theorem 3]. Due to the different context, we include a complete proof for the convenience of the reader.

Pick $r_0=\min(R_0,r_1)$ where R_0 and r_1 are the numbers appearing in Theorems 7.1 and 7.4, respectively. For $r\le r_0$, let $u\in S^{2,p}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},r))$. Let φ be a cutoff function as in Lemma 7.3:

$$\widetilde{B}(\bar{\xi},\sigma r) \prec \varphi \prec \widetilde{B}(\bar{\xi},\sigma' r).$$

By Theorem 7.1, $\varphi u\in S^{2,p}_{\widetilde{X},0}(\widetilde{B}(\bar{\xi},r))$; then, by density, we can apply Theorem 7.1 to φu:

$$\|\varphi u\|_{S^{2,p}(\widetilde{B}(\bar{\xi},r))} \le c\{\|\widetilde{\mathcal{L}}(\varphi u)\|_{L^p(\widetilde{B}(\bar{\xi},r))} + \|\varphi u\|_{L^p(\widetilde{B}(\bar{\xi},r))}\}.$$

For $1\le i,\ j\le q$, from the above inequality we get

$$\|\widetilde{X}_i\widetilde{X}_j u\|_{L^p(\widetilde{B}_{\sigma r})} \le c\{\|\widetilde{\mathcal{L}}u\|_{L^p(\widetilde{B}_{\sigma' r})} + \frac{1}{(1-\sigma)r}\|\widetilde{D}u\|_{L^p(\widetilde{B}_{\sigma' r})} + \frac{1}{(1-\sigma)^2r^2}\|u\|_{L^p(\widetilde{B}_{\sigma' r})}\}.$$

Multiplying both sides by $(1-\sigma)^2r^2$, we get

$$(1-\sigma)^2r^2\,\|\,\widetilde{X}_i\widetilde{X}_j u\,\|_{L^p(\widetilde{B}_{\sigma r})} \le c\{(1-\sigma)^2r^2\|\widetilde{\mathcal{L}}u\|_{L^p(\widetilde{B}_{\sigma' r})} + (1-\sigma)r(\|\widetilde{D}u\|_{L^p(\widetilde{B}_{\sigma' r})}) + \|u\|_{L^p(\widetilde{B}_{\sigma' r})}\}. \tag{7-6}$$

Next, we compute $(1-\sigma)^2r^2\|\widetilde{X}_0u\|_{L^p(\widetilde{B}_{\sigma r})}$:

$$\begin{aligned}(1-\sigma)^2r^2\|\widetilde{X}_0u\|_{L^p(\widetilde{B}_{\sigma r})} &= (1-\sigma)^2r^2\|\widetilde{\mathcal{L}}u - \sum_{i,j=1}^{q}\tilde{a}_{ij}\widetilde{X}_i\widetilde{X}_j u\|_{L^p(\widetilde{B}_{\sigma r})}\\ &\le c(1-\sigma)^2r^2(\|\widetilde{\mathcal{L}}u\|_{L^p(\widetilde{B}_{\sigma r})} + \|\widetilde{X}_i\widetilde{X}_j u\|_{L^p(\widetilde{B}_{\sigma r})}).\end{aligned} \tag{7-7}$$

Combining (7-6) and (7-7), we have

$$(1-\sigma)^2r^2\|\widetilde{D}^2u\|_{L^p(\widetilde{B}_{\sigma r})} \le c\{(1-\sigma)^2r^2\|\widetilde{\mathscr{L}}u\|_{L^p(\widetilde{B}_{\sigma' r})} + (1-\sigma)r\|\widetilde{D}u\|_{L^p(\widetilde{B}_{\sigma' r})} + \|u\|_{L^p(\widetilde{B}_{\sigma' r})}\}. \tag{7-8}$$

Adding $(1-\sigma)r\|\widetilde{D}u\|_{L^p(\widetilde{B}_{\sigma r})}$ to both sides of (7-8),

$$\begin{aligned}(1-\sigma)r\|\widetilde{D}u\|_{L^p(\widetilde{B}_{\sigma r})} &+ (1-\sigma)^2r^2\|\widetilde{D}^2u\|_{L^p(\widetilde{B}_{\sigma r})}\\ &\le c\{(1-\sigma)^2r^2\|\widetilde{\mathscr{L}}u\|_{L^p(\widetilde{B}_{\sigma' r})} + (1-\sigma)r\|\widetilde{D}u\|_{L^p(\widetilde{B}_{\sigma' r})} + \|u\|_{L^p(\widetilde{B}_{\sigma' r})}\},\end{aligned} \tag{7-9}$$

which, by Theorem 7.5, is

$$\le c\{(1-\sigma)^2r^2\|\widetilde{\mathscr{L}}u\|_{L^p(\widetilde{B}_{\sigma' r})} + (\delta\Phi_2 + \frac{c}{\delta}\Phi_0) + \|u\|_{L^p(\widetilde{B}_{\sigma' r})}\}.$$

Choosing δ small enough, we have

$$\Phi_2 + \Phi_1 \le c\{r^2\|\widetilde{\mathscr{L}}u\|_{L^p(\widetilde{B}_r)} + \|u\|_{L^p(\widetilde{B}_r)}\}.$$

Then

$$r^2\|\widetilde{D}^2u\|_{L^p(\widetilde{B}(\bar{\xi},r/2))} + r\|\widetilde{D}u\|_{L^p(\widetilde{B}(\bar{\xi},r/2))} \le c\{r^2\|\widetilde{\mathscr{L}}u\|_{L^p(\widetilde{B}(\bar{\xi},r))} + \|u\|_{L^p(\widetilde{B}(\bar{\xi},r))}\},$$

hence

$$\|u\|_{S^{2,p}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},r/2))} \le c\{\|\widetilde{\mathscr{L}}u\|_{L^p(\widetilde{B}(\bar{\xi},r))} + \|u\|_{L^p(\widetilde{B}(\bar{\xi},r))}\},$$

which is the desired result. □

Proof of Theorem 2.2. This follows from Theorem 7.2 in a way which is analogous to that followed in Section 6 to prove Schauder estimates. Namely, fix $\bar{x} \in \Omega' \Subset \Omega$ and R such that in $B(\bar{x}, R) \subset \Omega$ all the construction of the previous two subsections (lifting to $\widetilde{B}(\bar{\xi}, R)$ and so on) can be performed. Let $r_0 < R$ as in Theorem 7.2, and let $u \in S^{2,p}_X(B(\bar{x}, r_0))$. By Theorem 3.20 we know that the function $\widetilde{u}(x,h) = u(x)$ belongs to $S^{2,p}_{\widetilde{X}}(B(\bar{\xi}, r_0))$, so we can apply to $\widetilde{u}$ the L^p estimates contained in Theorem 7.2. Combining this fact with the two estimates in Theorem 3.20, we get

$$\begin{aligned}\|u\|_{S^{2,\alpha}_X(B(\bar{x},\delta_0 r_0/2))} &\le c\|\widetilde{u}\|_{S^{2,\alpha}_{\widetilde{X}}(\widetilde{B}(\bar{\xi},r_0/2))}\\ &\le c(\|\widetilde{\mathscr{L}}\widetilde{u}\|_{L^p(\widetilde{B}(\bar{\xi},r_0))} + \|\widetilde{u}\|_{L^p(\widetilde{B}(\bar{\xi},r_0))})\\ &\le c(\|\mathscr{L}u\|_{L^p(B(\bar{x},r_0))} + \|u\|_{L^p(B(\bar{x},r_0))}),\end{aligned}$$

since $\widetilde{\mathscr{L}}\widetilde{u} = \widetilde{(\mathscr{L}u)}$. Next, let us choose a family of balls $B(x_i, r_0)$ in Ω such that

$$\Omega' \subset \bigcup_{i=1}^k B(x_i, \delta_0 r_0/2) \subset \bigcup_{i=1}^k B(x_i, r_0) \subset \Omega.$$

Therefore,

$$\begin{aligned}\|u\|_{S_X^{2,p}(\Omega')} \le \|u\|_{S_X^{2,p}(\cup B(x_i,\delta_0 r_0/2))} &\le \sum_{i=1}^{k} \|u\|_{S_X^{2,p}(B(x_i,\delta_0 r_0/2))}\\ &\le c\sum_{i=1}^{k}\{\|\mathscr{L}u\|_{L^p(B(x_i,r_0))}+\|u\|_{L^p(B(x_i,r_0))}\}\\ &\le c\{\|\mathscr{L}u\|_{L^p(\Omega)}+\|u\|_{L^p(\Omega)}\}\end{aligned}$$

with c also depending on r_0. □

Acknowledgements

This research was mainly carried out while Maochun Zhu was visiting the Department of Mathematics of Politecnico di Milano, which we wish to thank for the hospitality.

References

[Bramanti 2010] M. Bramanti, "Singular integrals in nonhomogeneous spaces: L^2 and L^p continuity from Hölder estimates", *Rev. Mat. Iberoam.* **26**:1 (2010), 347–366. MR 2011e:42021 Zbl 1205.42011

[Bramanti and Brandolini 2000a] M. Bramanti and L. Brandolini, "L^p estimates for nonvariational hypoelliptic operators with VMO coefficients", *Trans. Amer. Math. Soc.* **352**:2 (2000), 781–822. MR 2000c:35026 Zbl 0935.35037

[Bramanti and Brandolini 2000b] M. Bramanti and L. Brandolini, "L^p estimates for uniformly hypoelliptic operators with discontinuous coefficients on homogeneous groups", *Rend. Sem. Mat. Univ. Politec. Torino* **58**:4 (2000), 389–433. MR 2004c:35058

[Bramanti and Brandolini 2005] M. Bramanti and L. Brandolini, "Estimates of BMO type for singular integrals on spaces of homogeneous type and applications to hypoelliptic PDEs", *Rev. Mat. Iberoam.* **21**:2 (2005), 511–556. MR 2006f:35037 Zbl 1082.35060

[Bramanti and Brandolini 2007] M. Bramanti and L. Brandolini, "Schauder estimates for parabolic nondivergence operators of Hörmander type", *J. Differential Equations* **234**:1 (2007), 177–245. MR 2007k:35052 Zbl 1113.35033

[Bramanti and Zhu 2012] M. Bramanti and M. Zhu, "Local real analysis in locally homogeneous spaces", *Manuscripta Math.* **138**:3-4 (2012), 477–528. MR 2916323 Zbl 1246.42019

[Bramanti et al. 1996] M. Bramanti, M. C. Cerutti, and M. Manfredini, "L^p estimates for some ultraparabolic operators with discontinuous coefficients", *J. Math. Anal. Appl.* **200**:2 (1996), 332–354. MR 97a:35132 Zbl 0922.47039

[Bramanti et al. 2010] M. Bramanti, L. Brandolini, and M. Pedroni, "On the lifting and approximation theorem for nonsmooth vector fields", *Indiana Univ. Math. J.* **59**:6 (2010), 2093–2138. MR 2919750 Zbl 1258.53028

[Bramanti et al. 2013] M. Bramanti, L. Brandolini, and M. Pedroni, "Basic properties of nonsmooth Hörmander's vector fields and Poincaré's inequality", *Forum Math.* **25**:4 (2013), 703–769. MR 3089748 Zbl 06202429 arXiv 0809.2872

[Calderón and Zygmund 1957] A.-P. Calderón and A. Zygmund, "Singular integral operators and differential equations", *Amer. J. Math.* **79** (1957), 901–921. MR 20 #7196 Zbl 0081.33502

[Campanato 1963] S. Campanato, "Proprietà di Hölderianità di alcune classi di funzioni", *Ann. Scuola Norm. Sup. Pisa* (3) **17**:1–2 (1963), 175–188. MR 27 #6119 Zbl 0121.29201

[Chiarenza et al. 1991] F. Chiarenza, M. Frasca, and P. Longo, "Interior $W^{2,p}$ estimates for nondivergence elliptic equations with discontinuous coefficients", *Ricerche Mat.* **40**:1 (1991), 149–168. MR 93k:35051 Zbl 0772.35017

[Chiarenza et al. 1993] F. Chiarenza, M. Frasca, and P. Longo, "$W^{2,p}$-solvability of the Dirichlet problem for nondivergence elliptic equations with VMO coefficients", *Trans. Amer. Math. Soc.* **336**:2 (1993), 841–853. MR 93f:35232 Zbl 0818.35023

[Coifman and Weiss 1971] R. R. Coifman and G. Weiss, *Analyse harmonique non-commutative sur certains espaces homogènes: étude de certaines intégrales singulières*, Lecture Notes in Mathematics **242**, Springer, Berlin, 1971. MR 58 #17690 Zbl 0224.43006

[Coifman et al. 1976] R. R. Coifman, R. Rochberg, and G. Weiss, "Factorization theorems for Hardy spaces in several variables", *Ann. of Math.* (2) **103**:3 (1976), 611–635. MR 54 #843 Zbl 0326.32011

[Di Francesco and Polidoro 2006] M. Di Francesco and S. Polidoro, "Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form", *Adv. Differential Equations* **11**:11 (2006), 1261–1320. MR 2007h:35204 Zbl 1153.35312

[Folland 1975] G. B. Folland, "Subelliptic estimates and function spaces on nilpotent Lie groups", *Ark. Mat.* **13**:2 (1975), 161–207. MR 58 #13215 Zbl 0312.35026

[Folland and Stein 1974] G. B. Folland and E. M. Stein, "Estimates for the $\bar\partial_b$ complex and analysis on the Heisenberg group", *Comm. Pure Appl. Math.* **27** (1974), 429–522. MR 51 #3719 Zbl 0293.35012

[Franchi and Lanconelli 1983] B. Franchi and E. Lanconelli, "Hölder regularity theorem for a class of linear nonuniformly elliptic operators with measurable coefficients", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **10**:4 (1983), 523–541. MR 85k:35094 Zbl 0552.35032

[Gutiérrez and Lanconelli 2009] C. E. Gutiérrez and E. Lanconelli, "Schauder estimates for sub-elliptic equations", *J. Evol. Equ.* **9**:4 (2009), 707–726. MR 2011b:35076 Zbl 1239.35039

[Hörmander 1967] L. Hörmander, "Hypoelliptic second order differential equations", *Acta Math.* **119** (1967), 147–171. MR 36 #5526 Zbl 0156.10701

[Lanconelli and Polidoro 1994] E. Lanconelli and S. Polidoro, "On a class of hypoelliptic evolution operators", *Rend. Sem. Mat. Univ. Politec. Torino* **52**:1 (1994), 29–63. MR 95h:35044 Zbl 0811.35018

[Lanconelli et al. 2002] E. Lanconelli, A. Pascucci, and S. Polidoro, "Linear and nonlinear ultraparabolic equations of Kolmogorov type arising in diffusion theory and in finance", pp. 243–265 in *Nonlinear problems in mathematical physics and related topics, II*, edited by M. S. Birman et al., Int. Math. Ser. (N. Y.) **2**, Kluwer/Plenum, New York, 2002. MR 2004c:35238 Zbl 1032.35114

[Nagel et al. 1985] A. Nagel, E. M. Stein, and S. Wainger, "Balls and metrics defined by vector fields, I: Basic properties", *Acta Math.* **155**:1 (1985), 103–147. MR 86k:46049 Zbl 0578.32044

[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

[Rothschild and Stein 1976] L. P. Rothschild and E. M. Stein, "Hypoelliptic differential operators and nilpotent groups", *Acta Math.* **137**:3-4 (1976), 247–320. MR 55 #9171 Zbl 0346.35030

[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

[Tolsa 2001] X. Tolsa, "A proof of the weak (1, 1) inequality for singular integrals with non doubling measures based on a Calderón–Zygmund decomposition", *Publ. Mat.* **45**:1 (2001), 163–174. MR 2002d:42019 Zbl 0980.42012

Received 24 Dec 2011. Revised 5 Apr 2013. Accepted 13 Jun 2013.

MARCO BRAMANTI: marco.bramanti@polimi.it
Dipartimento di Matematica, Politecnico di Milano, Via Bonardi 9, I-20133 Milano, Italy

MAOCHUN ZHU: zhumaochun2006@126.com
Department of Applied Mathematics, Northwestern Polytechnical University, 127 West Youyi Road, Xi'an, 710072, China

ANALYSIS AND PDE
Vol. 6, No. 8, 2013
dx.doi.org/10.2140/apde.2013.6.1857

STRICHARTZ ESTIMATES FOR SCHRÖDINGER EQUATIONS WITH VARIABLE COEFFICIENTS AND UNBOUNDED POTENTIALS

HARUYA MIZUTANI

This paper is concerned with Schrödinger equations with variable coefficients and unbounded electromagnetic potentials, where the kinetic energy part is a long-range perturbation of the flat Laplacian and the electric (respectively magnetic) potential can grow subquadratically (respectively sublinearly) at spatial infinity. We prove sharp (local-in-time) Strichartz estimates, outside a large compact ball centered at the origin, for any admissible pair including the endpoint. Under the nontrapping condition on the Hamilton flow generated by the kinetic energy, global-in-space estimates are also studied. Finally, under the nontrapping condition, we prove Strichartz estimates with an arbitrarily small derivative loss without asymptotic flatness on the coefficients.

1. Introduction

We study sharp (local-in-time) Strichartz estimates for Schrödinger equations with variable coefficients and unbounded electromagnetic potentials. More precisely, we consider the Schrödinger operator

$$H = \frac{1}{2}\sum_{j,k=1}^{d}(-i\partial_j - A_j(x))g^{jk}(x)(-i\partial_k - A_k(x)) + V(x), \quad x \in \mathbb{R}^d,$$

where $d \geq 1$ is the spatial dimension. Throughout the paper we assume that g^{jk}, V, and A_j are smooth real-valued functions on $\mathbb{R}^d$ and that $(g^{jk}(x))_{j,k}$ is symmetric and positive definite:

$$\sum_{j,k=1}^{d} g^{jk}(x)\xi_j\xi_k \geq c|\xi|^2, \quad x, \xi \in \mathbb{R}^d,$$

with some $c > 0$. Moreover, we suppose the following condition holds.

Assumption 1.1. There exists $\mu \geq 0$ such that for any $\alpha \in \mathbb{Z}_+^d$,

$$\begin{aligned}
|\partial_x^\alpha(g^{jk}(x) - \delta_{jk})| &\leq C_\alpha \langle x\rangle^{-\mu-|\alpha|},\\
|\partial_x^\alpha A_j(x)| &\leq C_\alpha \langle x\rangle^{1-\mu-|\alpha|},\\
|\partial_x^\alpha V(x)| &\leq C_\alpha \langle x\rangle^{2-\mu-|\alpha|}, \quad x \in \mathbb{R}^d.
\end{aligned}$$

Then it is well known that H admits a unique self-adjoint realization on $L^2(\mathbb{R}^d)$, which we denote by the

The author was partially supported by GCOE "Fostering top leaders in mathematics", Kyoto University.
MSC2010: primary 35B45, 35Q41; secondary 35S30, 81Q20.
Keywords: Schrödinger equation, Strichartz estimates, asymptotically flat metric, unbounded potential, unbounded electromagnetic potentials.

same symbol H. By Stone's theorem, H generates a unique unitary propagator e^{-itH} on $L^2(\mathbb{R}^d)$ such that $u(t) = e^{-itH}\varphi$ is the solution to the Schrödinger equation

$$\begin{aligned} i\partial_t u(t) &= Hu(t), \quad t \in \mathbb{R}, \\ u|_{t=0} &= \varphi \in L^2(\mathbb{R}^d). \end{aligned}$$

In order to explain the purpose of the paper, we recall some known results. Let us first recall well-known properties of the free propagator e^{-itH_0}, where $H_0 = -\Delta/2$. The distribution kernel of e^{-itH_0} is given explicitly by $(2\pi it)^{-d/2}e^{i|x-y|^2/(2t)}$, and $e^{-itH_0}\varphi$ thus satisfies the dispersive estimate

$$\|e^{-itH_0}\varphi\|_{L^\infty(\mathbb{R}^d)} \le C|t|^{-d/2}\|\varphi\|_{L^1(\mathbb{R}^d)}, \quad t \neq 0.$$

Moreover, e^{-itH_0} enjoys the (global-in-time) Strichartz estimates

$$\|e^{-itH_0}\varphi\|_{L^p(\mathbb{R};L^q(\mathbb{R}^d))} \le C\|\varphi\|_{L^2(\mathbb{R}^d)},$$

where (p,q) satisfies the admissible condition

$$p \ge 2, \quad \frac{2}{p} = d\Big(\frac{1}{2} - \frac{1}{q}\Big), \quad (d,p,q) \neq (2,2,\infty). \tag{1-1}$$

Strichartz estimates imply that, for any $\varphi \in L^2$, $e^{-itH_0}\varphi \in \bigcap_{q\in Q_d} L^q$ for a.e. $t \in \mathbb{R}$, where $Q_1 = [2,\infty]$, $Q_2 = [2,\infty)$ and $Q_d = [2, 2d/(d-2)]$ for $d \ge 3$. These estimates can therefore be regarded as L^p-type smoothing properties of Schrödinger equations, and have been widely used in the study of nonlinear Schrödinger equations; see, for example, [Cazenave 2003]. Strichartz estimates for e^{-itH_0} were first proved in [Strichartz 1977] for a restricted pair of (p,q) with $p = q = 2(d+2)/d$, and have been generalized for (p,q) satisfying (1-1) and $p \neq 2$ in [Ginibre and Velo 1985]. The endpoint estimate $(p,q) = (2, 2d/(d-2))$ for $d \ge 3$ was obtained in [Keel and Tao 1998].

For Schrödinger operators with electromagnetic potentials, that is, $H = (1/2)(-i\partial_x - A)^2 + V$, (short-time) dispersive and (local-in-time) Strichartz estimates have been extended with potentials decaying at infinity [Yajima 1987] or growing at infinity [Fujiwara 1980; Yajima 1991]. In particular, it was shown in the last two references that if $g^{jk} = \delta_{jk}$, V and A satisfy Assumption 1.1 with $\mu \ge 0$, and all derivatives of the magnetic field $B = dA$ of short-range type, then $e^{-itH}\varphi$ satisfies (short-time) dispersive estimate

$$\|e^{-itH}\varphi\|_{L^\infty(\mathbb{R}^d)} \le C|t|^{-d/2}\|\varphi\|_{L^1(\mathbb{R}^d)},$$

for sufficiently small $t \neq 0$. Local-in-time Strichartz estimates, which have the form

$$\|e^{-itH}\varphi\|_{L^p([-T,T];L^q(\mathbb{R}^d))} \le C_T\|\varphi\|_{L^2(\mathbb{R}^d)}, \quad T > 0,$$

are immediate consequences of this estimate and the TT^*-argument in [Ginibre and Velo 1985] (see [Keel and Tao 1998] for the endpoint estimate). For the case with singular electric potentials or with supercritical electromagnetic potentials, we refer to [Yajima 1987; 1998; Yajima and Zhang 2004; D'Ancona and Fanelli 2009]. We mention that global-in-time dispersive and Strichartz estimates for scattering states have also been studied under suitable decaying conditions on potentials and assumptions for zero energy; see [Journé et al. 1991; Yajima 2005; Schlag 2007; Erdoğan et al. 2009; D'Ancona et al. 2010]. We also mention that there is no result on sharp global-in-time dispersive estimates for magnetic Schrödinger equations.

On the other hand, the influence of the geometry on the behavior of solutions to linear and nonlinear partial differential equations has been extensively studied. From this geometric viewpoint, sharp local-in-time Strichartz estimates for Schrödinger equations with variable coefficients (or, more generally, on manifolds) have recently been investigated by many authors under several conditions on the geometry; see, for example, [Staffilani and Tataru 2002; Burq et al. 2004; Robbiano and Zuily 2005; Hassell et al. 2006; Bouclet and Tzvetkov 2007; Bouclet 2011b; Burq et al. 2010; Mizutani 2012]. In [Staffilani and Tataru 2002; Robbiano and Zuily 2005; Bouclet and Tzvetkov 2007], the authors studied the case on the Euclidean space with nontrapping asymptotically flat metrics. The case on the nontrapping asymptotically conic manifold was studied in [Hassell et al. 2006; Mizutani 2012]. Bouclet [2011b] considered the case of a nontrapping asymptotically hyperbolic manifold. For the trapping case, it was shown in [Burq et al. 2004] that Strichartz estimates with a loss of derivative $1/p$ hold on any compact manifold without boundaries. They also proved that the loss $1/p$ is optimal in the case of $M = \mathbb{S}^d$. In [Bouclet and Tzvetkov 2007; Bouclet 2011b; Mizutani 2012], the authors proved sharp Strichartz estimates, outside a large compact set, without the nontrapping condition. It was shown in [Burq et al. 2010] that sharp Strichartz estimates still hold for the case with hyperbolic trapped trajectories of sufficiently small fractal dimension. We mention that there are also several works on global-in-time Strichartz estimates in the case of long-range perturbations of the flat Laplacian on $\mathbb{R}^d$ [Bouclet and Tzvetkov 2008; Tataru 2008; Marzuola et al. 2008].

While (local-in-time) Strichartz estimates are well studied for these two cases (at least under the nontrapping condition), the literature is sparser for the mixed case. In this paper we give a unified approach to a combination of these two kinds of results. More precisely, under Assumption 1.1 with $\mu > 0$, we prove

(1) sharp local-in-time Strichartz estimates, outside a large compact set centered at the origin, without the nontrapping condition, and

(2) global-in-space estimates with the nontrapping condition.

Under the nontrapping condition and Assumption 1.1 with $\mu \geq 0$, we also show local-in-time Strichartz estimates with an arbitrarily small derivative loss. We mention that all results include the endpoint estimates $(p, q) = (2, 2d/(d-2))$ for $d \geq 3$. This is a natural continuation of the author's previous work [Mizutani 2013], which was concerned with the nonendpoint estimates for the case with at most linearly growing potentials.

In the sequel, $F(*)$ denotes the characteristic function designated by $(*)$. We now state the main result.

Theorem 1.2 (Strichartz estimates near infinity). *Suppose that H satisfies Assumption 1.1 with $\mu > 0$. Then there exists $R_0 > 0$ such that for any $T > 0$, $p \geq 2$, $q < \infty$, $2/p = d(1/2 - 1/q)$, and $R \geq R_0$, we have*

$$\|F(|x| > R)e^{-itH}\varphi\|_{L^p([-T,T];L^q(\mathbb{R}^d))} \leq C_T\|\varphi\|_{L^2(\mathbb{R}^d)}, \tag{1-2}$$

where $C_T > 0$ may be taken uniformly with respect to R.

To state the result on global-in-space estimates, we recall the *nontrapping condition*. We denote by

$$k(x,\xi)=\frac{1}{2}\sum_{j,k=1}^{d} g^{jk}(x)\xi_j\xi_k,$$

the classical kinetic energy, and by $(y_0(t,x,\xi),\eta_0(t,x,\xi))$ the Hamilton flow generated by $k(x,\xi)$:

$$\dot{y}_0(t)=\partial_\xi k(y_0(t),\eta_0(t)),\quad \dot{\eta}_0(t)=-\partial_x k(y_0(t),\eta_0(t)),\quad (y_0(0),\eta_0(0))=(x,\xi).$$

The Hamiltonian vector field $H_k=\partial_\xi k\cdot\partial_x-\partial_x k\cdot\partial_\xi$ generated by k is complete on $\mathbb{R}^{2d}$ since (g^{jk}) satisfies the uniform elliptic condition. Hence $(y_0(t,x,\xi),\eta_0(t,x,\xi))$ exists for all $t\in\mathbb{R}$.

Definition 1.3. We say that $k(x,\xi)$ satisfies the nontrapping condition if, for any $(x,\xi)\in\mathbb{R}^{2d}$ with $\xi\neq 0$,

$$|y_0(t,x,\xi)|\to+\infty\quad\text{as } t\to\pm\infty. \tag{1-3}$$

To control the asymptotic behavior of the flow, we also impose the following condition, which is the classical analogue of Mourre's inequality.

Assumption 1.4 (convexity near infinity). There exists $f\in C^\infty(\mathbb{R}^d)$ satisfying $f\geq 1$ and $f\to+\infty$ as $|x|\to+\infty$ such that $\partial^\alpha f\in L^\infty(\mathbb{R}^d)$ for any $|\alpha|\geq 2$ and

$$H_k(H_k f)(x,\xi)\geq ck(x,\xi)$$

on $\{(x,\xi)\in\mathbb{R}^{2d}: f(x)\geq R\}$ for some positive constants $c, R>0$.

Note that if $|\partial_x g^{jk}(x)|=o(|x|^{-1})$ as $|x|\to+\infty$, Assumption 1.4 holds with $f(x)=1+|x|^2$. In particular, Assumption 1.1 with $\mu>0$ implies Assumption 1.4. Moreover, if $g^{jk}(x)=(1+a_1\sin(a_2\log r))\delta_{jk}$ for $a_1\in\mathbb{R}, a_2>0$ with $a_1^2(1+a_2^2)<1$ and for $r=|x|\gg 1$, then Assumption 1.4 holds with $f(r)=(\int_0^r(1+a_1\sin(a_2\log t))^{-1}\,dt)^2$. For more examples, we refer to [Doi 2005, Section 2].

Theorem 1.5 (global-in-space Strichartz estimates). *Suppose that H satisfies Assumption 1.1 with $\mu\geq 0$. Let $T>0$, $p\geq 2$, $q<\infty$, and $2/p=d(1/2-1/q)$. Then, for any $r>0$, there exists $C_{T,r}>0$ such that*

$$\|F(|x|<r)e^{-itH}\varphi\|_{L^p([-T,T];L^q(\mathbb{R}^d))}\leq C_{T,r}\|\langle H\rangle^{1/(2p)}\varphi\|_{L^2(\mathbb{R}^d)}. \tag{1-4}$$

If we assume in addition that $k(x,\xi)$ satisfies the nontrapping condition (1-3) *and Assumption 1.4*,

$$\|F(|x|<r)e^{-itH}\varphi\|_{L^p([-T,T];L^q(\mathbb{R}^d))}\leq C_{T,r}\|\varphi\|_{L^2(\mathbb{R}^d)}. \tag{1-5}$$

In particular, combining with Theorem 1.2, we have the (*global-in-space*) *Strichartz estimates*

$$\|e^{-itH}\varphi\|_{L^p([-T,T];L^q(\mathbb{R}^d))}\leq C_T\|\varphi\|_{L^2(\mathbb{R}^d)},$$

under the nontrapping condition (1-3), *provided that $\mu>0$.*

For the general case we have the following partial result.

Theorem 1.6 (near sharp estimates without asymptotic flatness). *Suppose H satisfies Assumption 1.1 with $\mu \geq 0$ and $k(x, \xi)$ satisfies the nontrapping condition* (1-3). *Assume also Assumption 1.4. Let $T > 0$, $p \geq 2$, $q < \infty$, and $2/p = d(1/2 - 1/q)$. Then, for any $\varepsilon > 0$, there exists $C_{T,\varepsilon} > 0$ such that*

$$\|e^{-itH}\varphi\|_{L^p([-T,T];L^q(\mathbb{R}^d))} \leq C_{T,\varepsilon}\|\langle H\rangle^{\varepsilon}\varphi\|_{L^2(\mathbb{R}^d)}.$$

Remark 1.7. (1) The estimates of forms (1-2), (1-4), and (1-5) have been proved [Staffilani and Tataru 2002; Bouclet and Tzvetkov 2007] when $A \equiv 0$ and V is of long-range type. Theorems 1.2 and 1.5 are therefore regarded as generalizations of their results for the case with growing electromagnetic potential perturbations.

(2) The only restriction for admissible pairs, in comparison to the flat case, is to exclude $(p, q) = (4, \infty)$ for $d = 1$, which is due to the use of the Littlewood–Paley decomposition.

(3) The missing derivative loss $\langle H\rangle^{\varepsilon}$ in Theorem 1.6 is due to the use of the following local smoothing effect, due to [Doi 2005]:

$$\|\langle x\rangle^{-1/2-\varepsilon}\langle D\rangle^{1/2}e^{-itH}\varphi\|_{L^2([-T,T];L^2(\mathbb{R}^d))} \leq C_{T,\varepsilon}\|\varphi\|_{L^2(\mathbb{R}^d)}.$$

It is well known that this estimate does not hold when $\varepsilon = 0$ even for $H = H_0$. We would expect that Theorem 1.2 still holds true for the case with critical electromagnetic potentials in the following sense:

$$\langle x\rangle^{-1}|\partial_x^{\alpha} A_j(x)| + \langle x\rangle^{-2}|\partial_x^{\alpha} V(x)| \leq C_{\alpha\beta}\langle x\rangle^{-|\alpha|},$$

(at least if g^{jk} satisfies the bounds in Assumption 1.1 with $\mu > 0$). However, this is beyond our techniques (see also Remark 4.2).

The rest of the paper is devoted to the proofs of Theorems 1.2, 1.5, and 1.6. Throughout the paper we use the following notations. $\langle x\rangle$ stands for $\sqrt{1+|x|^2}$. We write $L^q = L^q(\mathbb{R}^d)$ if there is no confusion. For Banach spaces X and Y, we denote by $\|\cdot\|_{X\to Y}$ the operator norm from X to Y. We write $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$ and denote the set of multi-indices by $\mathbb{Z}_+^d$. We denote by K the kinetic energy part of H and by H_0 the free Schrödinger operator:

$$K = -\frac{1}{2}\sum_{j,k=1}^{d} \partial_j g^{jk}(x)\partial_k, \quad H_0 = -\tfrac{1}{2}\Delta = -\frac{1}{2}\sum_{j=1}^{d}\partial_j^2.$$

We define the symbols $p(x, \xi)$ and $p_1(x, \xi)$ by

$$\begin{aligned} p(x,\xi) &= \frac{1}{2}\sum_{j,k=1}^{d} g^{jk}(x)(\xi_j - A_j(x))(\xi_k - A_k(x)) + V(x), \\ p_1(x,\xi) &= -\frac{i}{2}\sum_{j,k=1}^{d}\biggl(\frac{\partial g^{jk}}{\partial x_j}(x)(\xi_k - A_k(x)) - g^{jk}(x)\frac{\partial A_k}{\partial x_j}(x)\biggr). \end{aligned} \tag{1-6}$$

Assumption 1.1 implies

$$
\begin{aligned}
|\partial_x^\alpha \partial_\xi^\beta p(x,\xi)| &\le C_{\alpha\beta}\langle x\rangle^{-|\alpha|}\langle \xi\rangle^{-|\beta|}(|\xi|^2+\langle x\rangle^{2-\mu}),\\
|\partial_x^\alpha \partial_\xi^\beta p_1(x,\xi)| &\le C_{\alpha\beta}\langle x\rangle^{-|\alpha|}\langle \xi\rangle^{-|\beta|}(\langle x\rangle^{-1-\mu}|\xi|+\langle x\rangle^{-\mu}).
\end{aligned}
\tag{1-7}
$$

For $h \in (0,1]$ we consider $H^h := h^2 H$ as a semiclassical Schrödinger operator with h-dependent electromagnetic potentials $h^2 V$ and hA_j. The corresponding symbols p_h and $p_{1,h}$ are also defined by

$$
\begin{aligned}
p_h(x,\xi) &= \frac{1}{2}\sum_{j,k=1}^{d} g^{jk}(x)(\xi_j - hA_j(x))(\xi_k - hA_k(x)) + h^2 V(x),\\
p_{1,h}(x,\xi) &= -\frac{i}{2}\sum_{j,k=1}^{d}\biggl(\frac{\partial g^{jk}}{\partial x_j}(x)(\xi_k - hA_k(x)) - hg^{jk}(x)\frac{\partial A_k}{\partial x_j}(x)\biggr).
\end{aligned}
\tag{1-8}
$$

It is easy to see that $H = \mathrm{Op}(p) + \mathrm{Op}(p_1)$ and $H^h = \mathrm{Op}_h(p_h) + h\,\mathrm{Op}_h(p_{1,h})$.

Before starting the details of the proofs, we describe the main ideas. First we note that, since our Hamiltonian H is not bounded below, the Littlewood–Paley decomposition associated with H seems to be false for $p \neq 2$ in general. To overcome this difficulty, we consider the following partition of unity on the phase space $\mathbb{R}^{2d}$:

$$
\psi_\varepsilon(x,\xi) + \chi_\varepsilon(x,\xi) = 1,
$$

where ψ_ε is supported in $\{(x,\xi) : \langle x\rangle < \varepsilon|\xi|\}$ for some sufficiently small constant $\varepsilon > 0$. It is easy to see that the symbol $p(x,\xi)$ is elliptic on $\operatorname{supp}\psi_\varepsilon$:

$$
C^{-1}|\xi|^2 \le p(x,\xi) \le C|\xi|^2, \quad (x,\xi) \in \operatorname{supp}\psi_\varepsilon,
$$

and we can therefore prove a Littlewood–Paley type decomposition of the form

$$
\|\mathrm{Op}(\psi_\varepsilon)u\|_{L^q} \le C_q\|u\|_{L^2} + C_q\biggl(\sum_{\substack{h=2^{-j}\\ j\ge 0}} \|\mathrm{Op}_h(a_h)f(h^2H)u\|_{L^q}^2\biggr)^{1/2},
$$

where $2 \le q < \infty$, the sequence $\{f(h^2\,\cdot) : h = 2^{-j},\ j \ge 0\}$ is a 4-adic partition of unity on $[1,\infty)$, a_h is an appropriate h-dependent symbol supported in $\{|x| < 1/h,\ |\xi| \in I\}$ for some open interval $I \Subset (0,\infty)$, and $\mathrm{Op}(\psi_\varepsilon)$, $\mathrm{Op}_h(a_h)$ denote the corresponding pseudodifferential and semiclassical pseudodifferential operators, respectively.

Then the idea of the proof of Theorem 1.2 is as follows. In view of the above Littlewood–Paley estimate, the proof is reduced to proving Strichartz estimates for $F(|x| > R)\,\mathrm{Op}_h(a_h)e^{-itH}$ and $\mathrm{Op}(\chi_\varepsilon)e^{-itH}$. In order to prove Strichartz estimates for $F(|x| > R)\,\mathrm{Op}_h(a_h)e^{-itH}$, we use semiclassical approximations of Isozaki–Kitada type. However, we note that, because of the unboundedness of potentials with respect to x, it is difficult to directly construct such approximations. To overcome this difficulty, we introduce a modified Hamiltonian $\widetilde{H}$ [Yajima and Zhang 2004] so that $\widetilde{H} = H$ for $|x| \le L/h$ and $\widetilde{H} = K$ for $|x| \ge 2L/h$ for some constant $L \ge 1$. Then $\widetilde{H}^h = h^2\widetilde{H}$ can be regarded as a "long-range perturbation" of the semiclassical free Schrödinger operator $H_0^h = h^2 H_0$. We also introduce the corresponding modified symbol $\tilde{p}_h(x,\xi)$ so that $\tilde{p}_h(x,\xi)=p_h(x,\xi)$ for $|x|\le L/h$ and $\tilde{p}_h(x,\xi)=k(x,\xi)$ for $|x|\ge 2L/h$. Let $a_h^\pm$ be supported in outgoing and incoming regions $\{R<|x|<1/h,\ |\xi|\in I,\ \pm\hat{x}\cdot\hat{\xi}>1/2\}$, respectively, so that $F(|x|>R)a_h=a_h^+ + a_h^-$, where $\hat{x}=x/|x|$. Rescaling $t\mapsto th$, we first construct the semiclassical approximations for $e^{-it\widetilde{H}^h/h}\operatorname{Op}_h(a_h^\pm)^*$ of the forms

$$e^{-it\widetilde{H}^h/h}\operatorname{Op}_h(a_h^\pm)^*=J_h(S_h^\pm,b_h^\pm)e^{-itH_0^h/h}J_h(S_h^\pm,c_h^\pm)^*+O(h^N),\quad 0\le\pm t\le 1/h,$$

respectively, where $S_h^\pm$ solves the eikonal equation associated to $\tilde{p}_h$ and $J_h(S_h^\pm,b_h^\pm)$ and $J_h(S_h^\pm,c_h^\pm)$ are the associated semiclassical Fourier integral operators (FIOs). The method of the construction is similar to that of [Robert 1994]. On the other hand, we will see that if $L\ge 1$ is large enough, the Hamilton flow generated by $\tilde{p}_h$ with initial conditions in $\operatorname{supp} a_h^\pm$ cannot escape from $\{|x|\le L/h\}$ for $0<\pm t\le 1/h$, respectively, that is,

$$\pi_x(\exp tH_{\tilde{p}_h}(\operatorname{supp} a_h^\pm))\subset\{|x|\le L/h\},\quad 0<\pm t\le 1/h.$$

Since $\tilde{p}_h=p_h$ for $|x|\le L/h$, we have

$$\exp tH_{\tilde{p}_h}(\operatorname{supp} a_h^\pm)=\exp tH_{p_h}(\operatorname{supp} a_h^\pm),\quad 0<\pm t\le 1/h.$$

We can thus expect (at least formally) that the corresponding two quantum evolutions are approximately equivalent modulo some smoothing operator. We will prove the following rigorous justification of this formal consideration:

$$\|(e^{-itH^h/h}-e^{-it\widetilde{H}^h/h})\operatorname{Op}_h(a_h^\pm)^*\|_{L^2\to L^2}\le C_M h^M,\quad 0\le\pm t\le 1/h,\ M\ge 0,$$

where $H^h=h^2H$. By using such approximations for $e^{-itH^h/h}\operatorname{Op}_h(a_h^\pm)^*$, we prove local-in-time dispersive estimates for $\operatorname{Op}_h(a_h^\pm)e^{-itH}\operatorname{Op}_h(a_h^\pm)^*$:

$$\|\operatorname{Op}_h(a_h^\pm)e^{-itH}\operatorname{Op}_h(a_h^\pm)^*\|_{L^1\to L^\infty}\le C|t|^{-d/2},\quad 0<h\ll 1,\ 0<|t|<1.$$

Strichartz estimates follow from these estimates and the abstract theorem due to Keel and Tao [1998].

Strichartz estimates for $\operatorname{Op}(\chi_\varepsilon)e^{-itH}$ follow from the short-time dispersive estimate

$$\|\operatorname{Op}(\chi_\varepsilon)e^{-itH}\operatorname{Op}(\chi_\varepsilon)^*\|_{L^1\to L^\infty}\le C_\varepsilon|t|^{-d/2},\quad 0<|t|<t_\varepsilon\ll 1.$$

To prove this, we first construct an approximation for $e^{-itH}\operatorname{Op}(\chi_\varepsilon)^*$ of the form

$$e^{-itH}\operatorname{Op}(\chi_\varepsilon)^*=J(\Psi,a)+O_{H^{-\gamma}\to H^\gamma}(1),\quad |t|<t_\varepsilon,\ \gamma>d/2,$$

where the phase function $\Psi=\Psi(t,x,\xi)$ is a solution to the time-dependent Hamilton–Jacobi equation associated to $p(x,\xi)$ and $J(\Psi,a)$ is the corresponding Fourier integral operator. In the construction, the fact that

$$|\partial_x^\alpha\partial_\xi^\beta p(x,\xi)|\le C_{\alpha\beta},\quad (x,\xi)\in\operatorname{supp}\chi_\varepsilon,\ |\alpha+\beta|\ge 2,$$

plays an important role. We note that if $(g^{jk})_{jk}-\mathrm{Id}_d\ne 0$ depends on x, these bounds do not hold without such a restriction of the support. Using these bounds, we construct the phase function $\Psi(t, x, \xi)$ such that

$$|\partial_x^\alpha \partial_\xi^\beta (\Psi(t, x, \xi) - x \cdot \xi + p(x, \xi))| \le C_{\alpha\beta} |t|^2 \langle x \rangle^{2-|\alpha+\beta|}.$$

Then we can follow a classical argument [Kitada and Kumano-go 1981] and construct the FIO $J(\Psi, a)$. By the composition formula, $\mathrm{Op}(\chi_\varepsilon) J(\Psi, a)$ is also an FIO and dispersive estimates for this operator follow from the standard stationary phase method. Finally, using an Egorov-type lemma, we prove that the remainder, $\mathrm{Op}(\chi_\varepsilon)(e^{-itH} \mathrm{Op}(\chi_\varepsilon)^* - J(\Psi, a))$, has a smooth kernel for sufficiently small t.

The proof of Theorem 1.5 is based on a standard idea [Staffilani and Tataru 2002]; see also [Burq et al. 2004; Bouclet and Tzvetkov 2007]. Strichartz estimates with loss of derivatives $\langle H \rangle^{1/(2p)}$ follow from semiclassical Strichartz estimates up to time scales of order h, which can be verified by the standard argument. Moreover, under the nontrapping condition, we will prove that the missing $1/p$ derivative loss can be recovered by using local smoothing effects [Doi 2005].

The proof of Theorem 1.6 is based on a slight modification of that of Theorem 1.5. By virtue of the Strichartz estimates for $\mathrm{Op}(\chi_\varepsilon) e^{-itH}$ and the Littlewood–Paley decomposition, it suffices to show

$$\|\mathrm{Op}_h(a_h) e^{-itH} \varphi\|_{L^p([-T,T];L^q)} \le C_T h^{-\varepsilon} \|\varphi\|_{L^2}, \quad 0 < h \ll 1.$$

To prove this estimate, we first prove semiclassical Strichartz estimates for $e^{-itH} \mathrm{Op}_h(a_h)^*$ up to time scales of order hR, where $R = \inf |\pi_x(\operatorname{supp} a_h)|$. The proof is based on a refinement of the standard WKB approximation for the semiclassical propagator $e^{-itH^h/h} \mathrm{Op}_h(a_h)^*$. Combining semiclassical Strichartz estimates with a partition of unity argument with respect to x, we will obtain the following Strichartz estimate with an inhomogeneous error term:

$$\|\mathrm{Op}_h(a_h) e^{-itH} \varphi\|_{L^p([-T,T];L^q)} \le C_T \|\varphi\|_{L^2} + C \|\langle x \rangle^{-1/2-\varepsilon} h^{-1/2-\varepsilon} \mathrm{Op}_h(a_h) e^{-itH} \varphi\|_{L^2([-T,T];L^2)},$$

for any $\varepsilon > 0$, which, combined with local smoothing effects, implies Theorem 1.6.

The paper is organized as follows. In Section 2 We record some known results on the semiclassical pseudodifferential calculus and prove the above Littlewood–Paley decomposition. Using dispersive estimates, which will be studied in Sections 4 and 5, we prove Theorem 1.2 in Section 3. We construct approximations of Isozaki–Kitada type and prove dispersive estimates for $\mathrm{Op}_h(a_h^\pm) e^{-itH} \mathrm{Op}_h(a_h^\pm)^*$ in Section 4. In Section 5 we discuss the dispersive estimates for $\mathrm{Op}(\chi_\varepsilon) e^{-itH} \mathrm{Op}(\chi_\varepsilon)^*$. The proofs of Theorems 1.5 and 1.6 are given in Sections 6 and 7, respectively.

2. Semiclassical functional calculus

Throughout this section we assume Assumption 1.1 with $\mu \ge 0$, that is,

$$|\partial_x^\alpha g^{jk}(x)| + \langle x \rangle^{-1} |\partial_x^\alpha A_j(x)| + \langle x \rangle^{-2} |\partial_x^\alpha V(x)| \le C_{\alpha\beta} \langle x \rangle^{-|\alpha|}. \tag{2-1}$$

The goal of this section is to prove a Littlewood–Paley type decomposition under a suitable restriction on the initial data. First we record (without proof) some known results on the pseudodifferential calculus which will be used throughout the paper. We refer to [Robert 1987; Martinez 2002] for the details of the proof.

Pseudodifferential calculus. For the metric $g = dx^2/\langle x\rangle^2 + d\xi^2/\langle\xi\rangle^2$ and a weight function $m(x,\xi)$ on the phase space $\mathbb{R}^{2d}$, we use Hörmander's symbol class notation $S(m, g)$, that is, $a \in S(m, g)$ if and only if $a \in C^\infty(\mathbb{R}^{2d})$ and

$$|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \le C_{\alpha\beta} m(x,\xi)\langle x\rangle^{-|\alpha|}\langle\xi\rangle^{-|\beta|}, \quad \alpha,\beta \in \mathbb{Z}_+^d.$$

To a symbol $a \in C^\infty(\mathbb{R}^{2d})$ and $h \in (0,1]$, we associate the semiclassical pseudodifferential operator (h-PDO for short) $\mathrm{Op}_h(a)$ defined by

$$\mathrm{Op}_h(a) f(x) = \frac{1}{(2\pi h)^d}\int e^{i(x-y)\cdot\xi/h} a(x,\xi) f(y)\, dy\, d\xi, \quad f \in \mathscr{S}(\mathbb{R}^d).$$

When $h = 1$ we write $\mathrm{Op}(a) = \mathrm{Op}_h(a)$ for simplicity. The Calderón–Vaillancourt theorem shows that for any symbol $a \in C^\infty(\mathbb{R}^{2d})$ satisfying $|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \le C_{\alpha\beta}$, $\mathrm{Op}_h(a)$ is extended to a bounded operator on $L^2(\mathbb{R}^d)$ uniformly with respect to $h \in (0,1]$. Moreover, for any symbol a satisfying

$$|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \le C_{\alpha\beta}\langle\xi\rangle^{-\gamma}, \quad \gamma > d,$$

$\mathrm{Op}_h(a)$ is extended to a bounded operator from $L^q(\mathbb{R}^d)$ to $L^r(\mathbb{R}^d)$ with the bounds

$$\|\mathrm{Op}_h(a)\|_{L^q\to L^r} \le C_{qr} h^{-d(1/q-1/r)}, \quad 1 \le q \le r \le \infty, \tag{2-2}$$

where $C_{qr} > 0$ is independent of $h \in (0,1]$. These bounds follow from the Schur lemma and an interpolation; see, for example, [Bouclet and Tzvetkov 2007, Proposition 2.4].

For two symbols $a \in S(m_1, g)$ and $b \in S(m_2, g)$, the composition $\mathrm{Op}_h(a)\,\mathrm{Op}_h(b)$ is also an h-PDO and is written in the form $\mathrm{Op}_h(c) = \mathrm{Op}_h(a)\,\mathrm{Op}_h(b)$ with a symbol $c \in S(m_1m_2, g)$ given by $c(x,\xi) = e^{ihD_\eta D_z} a(x,\eta) b(z,\xi)|_{z=x,\eta=\xi}$. Moreover, $c(x,\xi)$ has the expansion

$$c = \sum_{|\alpha|=0}^{N-1} \frac{h^{|\alpha|}}{i^{|\alpha|}\alpha!}\partial_\xi^\alpha a \cdot \partial_x^\alpha b + h^N r_N \quad \text{with } r_N \in S(\langle x\rangle^{-N}\langle\xi\rangle^{-N} m_1 m_2, g). \tag{2-3}$$

The symbol of the adjoint $\mathrm{Op}_h(a)^*$ is given by $a^*(x,\xi) = e^{ihD_\eta D_z} a(z,\eta)|_{z=x,\eta=\xi} \in S(m_1, g)$ which has the expansion

$$a^* = \sum_{|\alpha|=0}^{N-1} \frac{h^{|\alpha|}}{i^{|\alpha|}\alpha!}\partial_\xi^\alpha \partial_x^\alpha a + h^N r_N^* \quad \text{with } r_N^* \in S(\langle x\rangle^{-N}\langle\xi\rangle^{-N} m_1, g). \tag{2-4}$$

Littlewood–Paley decomposition. As we mentioned in the outline of the paper, H is not bounded below in general and hence we cannot expect that the Littlewood–Paley decomposition associated with H, which is of the form

$$\|u\|_{L^q} \le C_q\|u\|_{L^2} + C_q\biggl(\sum_{j=0}^\infty \|f(2^{-2j}H)u\|_{L^q}^2\biggr)^{1/2},$$

to hold if $q \neq 2$. The standard Littlewood–Paley decomposition associated with H_0 also does not work well in our case, since the commutator of H with the Littlewood–Paley projection $f(2^{-2j}H_0)$ can grow at spatial infinity. To overcome this difficulty, let us introduce an additional localization as follows. Given a parameter $\varepsilon > 0$ and a cut-off function $\varphi \in C_0^\infty(\mathbb{R}_+)$ such that $\varphi \equiv 1$ on $[0, 1/2]$ and $\operatorname{supp}\varphi \subset [0, 1]$, we define $\psi_\varepsilon(x, \xi)$ by

$$\psi_\varepsilon(x, \xi) = \varphi\Big(\frac{\langle x\rangle}{\varepsilon|\xi|}\Big).$$

It is easy to see that, for each $\varepsilon > 0$, $\psi_\varepsilon \in S(1, g)$ and is supported in $\{(x, \xi) \in \mathbb{R}^{2d} : \langle x\rangle < \varepsilon|\xi|\}$. Moreover, for sufficiently small $\varepsilon > 0$, $p(x, \xi)$ is uniformly elliptic on the support of ψ_ε and thus $\mathrm{Op}(\psi_\varepsilon)H$ is essentially bounded below.

In this subsection we prove a Littlewood–Paley type decomposition on the range of $\mathrm{Op}(\psi_\varepsilon)$. We begin with the following proposition which tells us that, for any $f \in C_0^\infty(\mathbb{R})$ and $h \in (0, 1]$, $\mathrm{Op}(\psi_\varepsilon)f(h^2H)$ is well approximated in terms of the h-PDO.

Proposition 2.1. *There exists $\varepsilon > 0$ such that, for any $f \in C_0^\infty(\mathbb{R})$ with $\operatorname{supp} f \Subset (0, \infty)$, we can construct bounded families $\{a_{h,j}\}_{h\in(0,1]} \subset \bigcap_{M\geq 0} S(\langle x\rangle^{-j}\langle\xi\rangle^{-M}, g)$, $j \geq 0$, such that*:

(1) *$a_{h,0}$ is given explicitly by $a_{h,0}(x, \xi) = \psi_\varepsilon(x, \xi/h)f(p_h(x, \xi))$. Moreover,*

$$\operatorname{supp} a_{h,j} \subset \operatorname{supp}\psi_\varepsilon(\cdot\,, \cdot/h) \cap \operatorname{supp} f(p_h) \subset \{(x, \xi) \in \mathbb{R}^{2d} : \langle x\rangle < 1/h,\ |\xi| \in I\},$$

for some relatively compact open interval $I \Subset (0, \infty)$. In particular, we have

$$\|\mathrm{Op}_h(a_{h,j})\|_{L^{q'}\to L^q} \leq C_{jqq'}h^{-d(1/q'-1/q)}, \quad 1 \leq q' \leq q \leq \infty,$$

uniformly in $h \in (0, 1]$.

(2) *For any integer $N > d+2$, we set $a_h = \sum_{j=0}^{N-1} h^j a_{h,j}$. Then*

$$\|\mathrm{Op}(\psi_\varepsilon)f(h^2H) - \mathrm{Op}_h(a_h)\|_{L^2\to L^q} \leq C_{qN}h^2, \quad 2 \leq q \leq \infty,$$

uniformly in $h \in (0, 1]$.

The following is an immediate consequence of this proposition.

Corollary 2.2. *For any $2 \leq q \leq \infty$ and $h \in (0, 1]$, $\mathrm{Op}(\psi_\varepsilon)f(h^2H)$ is bounded from $L^2(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$ and satisfies*

$$\|\mathrm{Op}(\psi_\varepsilon)f(h^2H)\|_{L^2\to L^q} \leq C_q h^{-d(1/2-1/q)},$$

where $C_q > 0$ is independent of $h \in (0, 1]$.

For the low energy part we have the following.

Lemma 2.3. *For any $f_0 \in C_0^\infty(\mathbb{R})$ and $2 \leq q \leq \infty$, we have*

$$\|\mathrm{Op}(\psi_\varepsilon)f_0(H)\|_{L^2\to L^q} \leq C_q.$$

Remark 2.4. If $V, A \equiv 0$, then Proposition 2.1, Corollary 2.2, and Lemma 2.3 hold without the additional term $\mathrm{Op}(\psi_\varepsilon)$. Moreover, in this case we see that the remainder satisfies

$$\|f(h^2H) - \mathrm{Op}_h(a_h)\|_{L^2\to L^q} \leq C_{qN}h^{N-d(1/2-1/q)}.$$

We refer to [Burq et al. 2004] (for the case on compact manifolds without boundary) and to [Bouclet and Tzvetkov 2007] (for the case with metric perturbations on $\mathbb{R}^d$). For more general cases with Laplace–Beltrami operators on noncompact manifolds with ends, we refer to [Bouclet 2010; 2011a]. Because of this result, we believe Proposition 2.1 is far from sharp. However, the bounds

$$\|\mathrm{Op}(\psi_\varepsilon)f(h^2H)-\mathrm{Op}_h(a_h)\|_{L^2\to L^q}\le C_{qN}h,\quad 2\le q\le\infty,$$

are sufficient to obtain our Littlewood–Paley type decomposition (Proposition 2.5). For more details, we refer to Burq, Gérard, and Tvzetkov [2004, Corollary 2.3].

Proof of Proposition 2.1. We write

$$\mathrm{Op}(\psi_\varepsilon)=\mathrm{Op}_h(\psi_{\varepsilon/h}),\quad h\in(0,1],$$

where $\psi_{\varepsilon/h}(x,\xi)=\psi_\varepsilon(x,\xi/h)$ satisfies $\operatorname{supp}\psi_{\varepsilon/h}\subset\{h\langle x\rangle<\varepsilon|\xi|\}$ and

$$|\partial_x^\alpha\partial_\xi^\beta\psi_{\varepsilon/h}(x,\xi)|\le C_{\alpha\beta\varepsilon}h^{-|\beta|}\langle x\rangle^{-|\alpha|}\langle\xi/h\rangle^{-|\beta|}\le C_{\alpha\beta\varepsilon}\langle x\rangle^{-|\alpha|}(h+|\xi|)^{-|\beta|}. \tag{2-5}$$

By using the Helffer–Sjöstrand formula [1989], we get

$$\mathrm{Op}_h(\psi_{\varepsilon/h})f(h^2H)=-\frac{1}{2\pi i}\int_{\mathbb{C}}\frac{\partial\tilde f}{\partial\bar z}(z)\,\mathrm{Op}_h(\psi_{\varepsilon/h})(h^2H-z)^{-1}\,dz\wedge d\bar z,$$

where $\tilde f(z)$ is an almost analytic extension of $f(\lambda)$. Since $f\in C_0^\infty(\mathbb{R})$, $\tilde f(z)$ is also compactly supported and satisfies

$$\partial_{\bar z}\tilde f(z)=O(|\operatorname{Im}z|^M)$$

for any $M>0$. We may assume $|z|\le C$ on $\operatorname{supp}\tilde f$ with some $C>0$. In order to use this formula, we shall construct a semiclassical approximation of $\mathrm{Op}_h(\psi_{\varepsilon/h})(h^2H-z)^{-1}$, in terms of the h-PDO, for $z\in\mathbb{C}\setminus[0,\infty)$ with $|z|\le C$. Although the method is based on the standard semiclassical parametrix construction (see, for example, [Robert 1987; Burq et al. 2004]), we give the details of the proof, since $\psi_{\varepsilon/h}$ is not uniformly bounded in $S(1,g)$ with respect to $h\in(0,1]$.

We first study the symbol of the resolvent $(h^2H-z)^{-1}$. Let p_h and $p_{1,h}$ be as in (1-8) so that $h^2H=\mathrm{Op}_h(p_h)+h\,\mathrm{Op}_h(p_{1,h})$. Since

$$h|A(x)|\lesssim|\xi|,\quad h^2|V(x)|\lesssim|\xi|^2,$$

on $\operatorname{supp}\psi_{\varepsilon/h}$, we obtain by (1-7) that

$$|\partial_x^\alpha\partial_\xi^\beta p_h(x,\xi)|\le C_{\alpha\beta}\langle x\rangle^{-|\alpha|}|\xi|^{2-|\beta|}\qquad\text{if }|\beta|\le2, \tag{2-6}$$

$$|\partial_x^\alpha\partial_\xi^\beta p_{1,h}(x,\xi)|\le C_{\alpha\beta}\langle x\rangle^{-1-|\alpha|}|\xi|^{1-|\beta|}\quad\text{if }|\beta|\le1, \tag{2-7}$$

uniformly in $(x,\xi)\in\operatorname{supp}\psi_{\varepsilon/h}$ and $h\in(0,1]$. Moreover, if $\varepsilon>0$ is sufficiently small, the uniform ellipticity of k implies that p_h is also uniformly elliptic on $\operatorname{supp}\psi_{\varepsilon/h}$:

$$C_1^{-2}|\xi|^2\le p_h(x,\xi)\le C_1^2|\xi|^2\quad\text{if }h\langle x\rangle<\varepsilon|\xi|,$$

with some $C_1 > 0$, which particularly implies

$$\frac{1}{|p_h(x,\xi)-z|} \lesssim \begin{cases} |\operatorname{Im} z|^{-1} & \text{if } |\xi| \le 2C_2, \\ \langle \xi \rangle^{-2} & \text{if } |\xi| \ge 2C_2 \end{cases} \tag{2-8}$$

for $(x,\xi) \in \operatorname{supp} \psi_{\varepsilon/h}$, $z \notin \mathbb{R}$, and $|z| \le C$, with some $C_2 > 0$.

Let us now consider a sequence of symbols $q_j^h = q_j^h(z,x,\xi)$ (depending holomorphically on $z \notin \mathbb{R}$) defined inductively by

$$\begin{aligned} q_0^h &= \frac{\psi_{\varepsilon/h}}{p_h - z}, \\ q_1^h &= -\frac{1}{p_h - z}\bigg(\sum_{|\alpha|=1} i^{-1}\partial_\xi^\alpha q_0^h \cdot \partial_x^\alpha p_h + q_0^h \cdot p_{1,h}\bigg), \\ q_j^h &= -\frac{1}{p_h - z}\bigg(\sum_{\substack{|\alpha|+k=j \\ |\alpha|\ge 1}} \frac{i^{-|\alpha|}}{\alpha!}\partial_\xi^\alpha q_k^h \cdot \partial_x^\alpha p_h + \sum_{|\alpha|+k=j-1} \frac{i^{-|\alpha|}}{\alpha!}\partial_\xi^\alpha q_k^h \cdot \partial_x^\alpha p_{1,h}\bigg), \quad j \ge 2. \end{aligned}$$

We then learn by (2-5), (2-6), and (2-8) that

$$\begin{aligned} |\partial_x^\alpha \partial_\xi^\beta q_0^h(z,x,\xi)| &\le C_{\alpha\beta\varepsilon} \begin{cases} \langle x \rangle^{-|\alpha|}(h+|\xi|)^{-|\beta|}|\operatorname{Im} z|^{-1-|\alpha+\beta|} & \text{if } |\xi| \le 2C_2, \\ \langle x \rangle^{-|\alpha|}\langle \xi \rangle^{-|\beta|-2} & \text{if } |\xi| \ge 2C_2, \end{cases} \\ &\le C_{\alpha\beta\varepsilon}\langle x \rangle^{-|\alpha|}(h+|\xi|)^{-|\beta|}|\operatorname{Im} z|^{-1-|\alpha+\beta|} \end{aligned} \tag{2-9}$$

for $z \notin \mathbb{R}$ with $|z| \le C$ and $h \in (0,1]$. Similarly, by using (2-6), (2-7), and (2-9), we obtain that if $h|\xi| \le 2C_2$,

$$\begin{aligned} &|\partial_x^\alpha \partial_\xi^\beta q_1^h(z,x,\xi)| \\ &\quad \le C_{\alpha\beta\varepsilon}\big(\langle x \rangle^{-1-|\alpha|}(h+|\xi|)^{-1-|\beta|}|\xi|^2|\operatorname{Im} z|^{-3-|\alpha+\beta|} + \langle x \rangle^{-1-|\alpha|}(h+|\xi|)^{-|\beta|}(h+|\xi|)|\operatorname{Im} z|^{-2-|\alpha+\beta|}\big) \\ &\quad \le C_{\alpha\beta\varepsilon}(h+|\xi|)^2\langle x \rangle^{-1-|\alpha|}(h+|\xi|)^{-1-|\beta|}|\operatorname{Im} z|^{-3-|\alpha+\beta|}, \end{aligned}$$

for $z \notin \mathbb{R}$ with $|z| \le C$ and $h \in (0,1]$. Here note that, in this case, $(h+|\xi|)^{-1}$ may have a singularity at $\xi = 0$ as $h \to +0$. In order to prove the remainder estimate, we will remove this singularity by using a rescaling $\xi \mapsto h\xi$ (see the estimates (2-12)). For $h|\xi| \ge 2C_2$, q_1^h does not have such a singularity and satisfies

$$|\partial_x^\alpha \partial_\xi^\beta q_1^h(z,x,\xi)| \le C_{\alpha\beta\varepsilon}\langle x \rangle^{-1-|\alpha|}\langle \xi \rangle^{-|\beta|-4}|\xi| \le C_{\alpha\beta\varepsilon}\langle x \rangle^{-1-|\alpha|}\langle \xi \rangle^{-|\beta|-3}$$

uniformly in $z \notin \mathbb{R}$ with $|z| \le C$ and $h \in (0,1]$. Since $1 \lesssim h+|\xi|$ if $h|\xi| \gtrsim 1$, summarizing these, we get

$$|\partial_x^\alpha \partial_\xi^\beta q_1^h(z,x,\xi)| \le C_{\alpha\beta\varepsilon}\langle x \rangle^{-1-|\alpha|}(h+|\xi|)^{1-|\beta|}|\operatorname{Im} z|^{-3-|\alpha+\beta|}, \quad z \notin \mathbb{R},\ |z| \le C,\ h \in (0,1].$$

The estimates (2-9) and a direct computation also show that q_1^h is of the form

$$q_1^h = q_{11}^h(p_h - z)^{-3} + q_{10}^h(p_h - z)^{-2},$$

where q_{1k}^h are supported in $\operatorname{supp} \psi_{\varepsilon/h}$, are independent of z, and satisfy

$$|\partial_x^\alpha \partial_\xi^\beta q_{1k}^h(x,\xi)| \le C_{\alpha\beta\varepsilon}\langle x \rangle^{-1-|\alpha|}(h+|\xi|)^{-|\beta|}\langle \xi \rangle^{N_1(k)}, \quad h \in (0,1],$$

with some positive integer $N_1(k) > 0$. For $j \ge 2$, an induction argument yields that

$$|\partial_x^\alpha \partial_\xi^\beta q_j^h(z,x,\xi)| \le C_{\alpha\beta\varepsilon} \langle x\rangle^{-j-|\alpha|}(h+|\xi|)^{2-j-|\beta|} |\operatorname{Im} z|^{-2j-1-|\alpha+\beta|}, \quad j \ge 2, \tag{2-10}$$

for $z \notin \mathbb{R}$ with $|z| \le C$ and $h \in (0,1]$. It also follows from an induction on j that there exists a sequence of z-independent symbols $(q_{jk}^h)_{k=0}^j$ supported in $\operatorname{supp} \psi_{\varepsilon/h}$ and satisfying

$$|\partial_x^\alpha \partial_\xi^\beta q_{jk}^h(x,\xi)| \le C_{\alpha\beta\varepsilon} \langle x\rangle^{-j-|\alpha|}(h+|\xi|)^{-|\beta|} \langle\xi\rangle^{N_j(k)} \tag{2-11}$$

with some $N_j(k) > 0$, such that q_j^h is of the form

$$q_j^h = \sum_{k=0}^j q_{jk}^h (p_h - z)^{-j-k-1}.$$

Rescaling $\xi \mapsto h\xi$, we learn by (2-9) and (2-10) that

$$q_0^h(z,x,h\xi) \in S(1,g), \quad h^j q_j^h(z,x,h\xi) \in S(h^2\langle x\rangle^{-j}\langle\xi\rangle^{2-j}, g),$$

with uniform bounds in h and polynomially bounds in $|\operatorname{Im} z|^{-1}$. Then, by the construction of q_j^h, the standard symbolic calculus (not in the semiclassical regime), and the fact that

$$\operatorname{Op}(h^j q_j^h(z,x,h\xi)) = h^j \operatorname{Op}_h(q_j^h),$$

we obtain

$$\operatorname{Op}(\psi_\varepsilon) = \sum_{j=0}^{N-1} h^j \operatorname{Op}_h(q_j^h)(h^2 H - z) + h^2 \operatorname{Op}(r_{h,N,z}), \quad N \ge 1,$$

with some $r_{h,N,z} \in S(\langle x\rangle^{-N}\langle\xi\rangle^{2-N}, g)$ satisfying

$$|\partial_x^\alpha \partial_\xi^\beta r_{h,N,z}(x,\xi)| \le C_{\alpha\beta\varepsilon N} \langle x\rangle^{-N-|\alpha|}\langle\xi\rangle^{2-N-|\beta|} |\operatorname{Im} z|^{-2N-1-|\alpha+\beta|}, \tag{2-12}$$

where $C_{\alpha\beta\varepsilon N} > 0$ may be taken uniformly in $h \in (0,1]$, $z \in \mathbb{C} \setminus \mathbb{R}$ with $|z| \le C$ and $x, \xi \in \mathbb{R}^d$.

We now use the Helffer–Sjöstrand formula to obtain

$$\operatorname{Op}(\psi_\varepsilon) f(h^2 H) = \sum_{j=0}^{N-1} h^j \operatorname{Op}_h(a_{h,j}) + h^2 R(h,N),$$

where

$$a_{h,0}(x,\xi) = \psi_{\varepsilon/h}(x,\xi)(f \circ p_h)(x,\xi),$$

$$a_{h,j}(x,\xi) = \sum_{k=0}^j \frac{(-1)^{k+j}}{(k+j)!} q_{jk}^h(x,\xi)(f^{(j+k)} \circ p_h)(x,\xi), \ 1 \le j \le N-1,$$

$$R(h,N) = -\frac{1}{2\pi i} \int_{\mathbb{C}} \frac{\partial \tilde f}{\partial \bar z}(z) \operatorname{Op}_h(r_{h,N,z})(h^2 H - z)^{-1} \, dz \wedge d\bar z.$$

Since $\operatorname{supp} q_{jk} \subset \operatorname{supp} \psi_{\varepsilon/h} \subset \{h\langle x\rangle < \varepsilon|\xi|\}$ and p_h is uniformly elliptic (that is, $p_h \approx |\xi|^2$) on the latter region, taking $\varepsilon > 0$ smaller if necessary, we have

$$a_{h,j} \subset \operatorname{supp} \psi_{\varepsilon/h} \cap \operatorname{supp} f(p_h) \subset \{(x,\xi) : |x| < 1/h, \ C_0^{-1} \le |\xi| \le C_0\}$$

with some positive constant $C_0 > 0$, which, combined with (2-11), implies that $\{a_{h,j}\}_{h\in(0,1]}$ is bounded in $\bigcap_{M\geq 0} S(\langle x\rangle^{-j}\langle\xi\rangle^{-M}, g)$, since $h+|\xi| \gtrsim \langle\xi\rangle$ on $\operatorname{supp}\psi_{\varepsilon/h}\cap\operatorname{supp} f(p_h)$. By virtue of (2-2), we also obtain

$$\|\mathrm{Op}_h(a_{h,j})\|_{L^{q'}\to L^q} \leq C_{jqq'}h^{-d(1/q'-1/q)}, \quad h\in(0,1],\ 1\leq q'\leq q\leq\infty.$$

Finally, we prove the estimate on the remainder $R(h,N)$. If we choose $N > d+2$, then (2-12) and (2-2) (with $h=1$) imply

$$\|\mathrm{Op}(r_{h,N,z})\|_{L^2\to L^q} \leq C_{qN}\,|\mathrm{Im}\,z|^{-n(N,q)}, \quad 2\leq q\leq\infty,$$

with some positive integer $n(N,q)\geq 2N+1$, where $C_{qN}>0$ is independent of h. Using the bounds $\|(h^2H-z)^{-1}\|_{L^2\to L^2}\leq|\mathrm{Im}\,z|^{-1}$, $|\partial_{\bar z}\tilde f(z)|\leq C_M|\mathrm{Im}\,z|^M$ for any $M\geq 0$ and the fact that $\tilde f$ is compactly supported, we conclude that

$$\begin{aligned}\|R(h,N)\|_{L^2\to L^q} &\leq C_M\int_{\operatorname{supp}\tilde f}|\mathrm{Im}\,z|^M\|\mathrm{Op}(r_{h,N,z})\|_{L^2\to L^q}\|(h^2H-z)^{-1}\|_{L^2\to L^2}\,dz\wedge d\bar z\\ &\leq C_{MNq}\int_{\operatorname{supp}\tilde f}|\mathrm{Im}\,z|^{M-n(N,q)-1}\,dz\wedge d\bar z\\ &\leq C_{MNq},\end{aligned}$$

provided that M is large enough. This completes the proof. □

Proof of Lemma 2.3. By the same argument as above with $h=1$, we can see that

$$\mathrm{Op}(\psi_\varepsilon)f_0(H)=\sum_{j=0}^{N-1}\mathrm{Op}(a_j)+R(N)$$

where $a_j\in\bigcap_{M\geq 0}S(\langle x\rangle^{-j}\langle\xi\rangle^{-M}, g)$ are supported in

$$\operatorname{supp}\psi_\varepsilon\cap\operatorname{supp} f_0(p)\subset\{(x,\xi)\in\mathbb{R}^{2d}:\langle x\rangle<\varepsilon|\xi|,\ |\xi|\lesssim 1\}$$

and $R(N)$ satisfies

$$\|R(h,N)\|_{L^2\to L^q}\leq C_{Nq}, \quad 2\leq q\leq\infty,$$

if $N>d+2$. The assertion then follows from (2-2). □

Consider a 4-adic partition of unity

$$f_0(\lambda)+\sum_h f(h^2\lambda)=1, \quad \lambda\in\mathbb{R},$$

where $f_0, f\in C_0^\infty(\mathbb{R})$ with $\operatorname{supp} f_0\subset[-1,1]$, $\operatorname{supp} f\subset[1/4,4]$ and $\sum_h$ means that, in the sum, h takes all negative powers of 2 as values, that is, $\sum_h=\sum_{h=2^{-j},j\geq 0}$. Let $F\in C_0^\infty(\mathbb{R})$ be such that $\operatorname{supp} F\subset[1/8,8]$ and $F\equiv 1$ on $\operatorname{supp} f$. The spectral decomposition theorem implies

$$1=f_0(H)+\sum_h f(h^2H)=f_0(H)+\sum_h F(h^2H)f(h^2H).$$

Let $a_h \in S(1, g)$ be as in Proposition 2.1 with $f = F$. Using Proposition 2.1, we obtain a Littlewood–Paley type estimates on a range of $\mathrm{Op}(\psi_\varepsilon)$.

Proposition 2.5. *For any $2 \le q < \infty$,*

$$\|\mathrm{Op}(\psi_\varepsilon)u\|_{L^q(\mathbb{R}^d)} \le C_q\|u\|_{L^2(\mathbb{R}^d)} + C_q\biggl(\sum_h \|\mathrm{Op}_h(a_h) f(h^2H)u\|_{L^q(\mathbb{R}^d)}\biggr)^{1/2}.$$

Proof. The proof is the same as that of [Burq et al. 2004, Corollary 2.3] and we omit the details. □

Corollary 2.6. *Let $\varepsilon > 0$ and ψ_ε be as above and $\chi_\varepsilon = 1 - \psi_\varepsilon$. Let $\rho \in C^\infty(\mathbb{R}^d)$ be such that*

$$|\partial_x^\alpha \rho(x)| \le C_\alpha \langle x\rangle^{-|\alpha|}, \quad \alpha \in \mathbb{Z}_+^d.$$

Then, for any $T > 0$ and any (p, q) satisfying $p \ge 2$, $q < \infty$ and $2/p = d(1/2-1/q)$, there exists $C_T > 0$ such that

$$\begin{aligned}\|\rho e^{-itH}\varphi\|_{L^p([-T,T];L^q(\mathbb{R}^d))} &\le C_T\|\varphi\|_{L^2(\mathbb{R}^d)} + C\|\mathrm{Op}(\chi_\varepsilon)e^{-itH}\varphi\|_{L^p([-T,T];L^q(\mathbb{R}^d))} \\ &\quad + C\biggl(\sum_h \|\mathrm{Op}_h(a_h)e^{-itH} f(h^2H)\varphi\|^2_{L^p([-T,T];L^q(\mathbb{R}^d))}\biggr)^{1/2},\end{aligned}$$

where a_h is given by Proposition 2.1 with ψ_ε replaced by $\rho\psi_\varepsilon$. In particular, $a_h(x, \xi)$ is supported in $\operatorname{supp} \rho(x)\psi(x, \xi/h)F(p_h(x, \xi))$.

Proof. This proposition follows from the L^2-boundedness of e^{-itH}, Propositions 2.1 and 2.5 (with ψ_ε replaced by $\rho\psi_\varepsilon$), and the Minkowski inequality. □

3. Proof of Theorem 1.2

In this section we prove Theorem 1.2 under Assumption 1.1 with $\mu > 0$. We first state two key estimates which we will prove in later sections. For $R > 0$, an open interval $I \Subset (0, \infty)$ and $\sigma \in (-1, 1)$, we define the outgoing and incoming regions $\Gamma^\pm(R, I, \sigma)$ by

$$\Gamma^\pm(R, I, \sigma) := \biggl\{(x, \xi) \in \mathbb{R}^{2d} : |x| > R,\ |\xi| \in I,\ \pm\frac{x\cdot\xi}{|x||\xi|} > -\sigma\biggr\},$$

respectively. We then have the following (local-in-time) dispersive estimates.

Proposition 3.1. *Suppose that H satisfies Assumption 1.1 with $\mu > 0$. Let $I \Subset (0, \infty)$ and $\sigma \in (-1, 1)$. Then, for sufficiently large $R \ge 1$, small $h_0 > 0$, and any symbols $a_h^\pm \in S(1, g)$ supported in $\Gamma^\pm(R, I, \sigma) \cap \{x : |x| < 1/h\}$, we have*

$$\|\mathrm{Op}_h(a_h^\pm)e^{-itH}\,\mathrm{Op}_h(a_h^\pm)^*\|_{L^1\to L^\infty} \le C|t|^{-d/2}, \quad 0 < |t| \le 1,$$

uniformly with respect to $h \in (0, h_0]$.

We prove this proposition in Section 4. In the region $\{|x| \gtrsim |\xi|\}$, we have the following (short-time) dispersive estimates.

Proposition 3.2. *Suppose that H satisfies Assumption 1.1 with $\mu \ge 0$. Let us fix arbitrarily $\varepsilon > 0$. Then there exists $t_\varepsilon > 0$ such that, for any symbol $\chi_\varepsilon \in S(1, g)$ supported in $\{(x, \xi) : \langle x\rangle \ge \varepsilon|\xi|\}$, we have*

$$\|\mathrm{Op}(\chi_\varepsilon)e^{-itH}\,\mathrm{Op}(\chi_\varepsilon)^*\|_{L^1\to L^\infty} \le C_\varepsilon|t|^{-d/2}, \quad 0 < |t| \le t_\varepsilon.$$

We prove this proposition in Section 5.

Proof of Theorem 1.2. Taking $\rho \in C^\infty(\mathbb{R}^d)$ so that $0 \le \rho(x) \le 1$, $\rho(x) = 1$ for $|x| \ge 1$ and $\rho(x) = 0$ for $|x| \le 1/2$, we set $\rho_R(x) = \rho(x/R)$. In order to prove Theorem 1.2, it suffices to show

$$\|\rho_R e^{-itH}\varphi\|_{L^p([-T,T];L^q(\mathbb{R}^d))} \le C_T\|\varphi\|_{L^2(\mathbb{R}^d)},$$

for sufficiently large $R \ge 1$. We may also assume without loss of generality that $T > 0$ is sufficiently small. Indeed, if the above estimate holds on $[-T_0, T_0]$ with some $T_0 > 0$, we obtain by the unitarity of e^{-itH} on L^2 that, for any $T > T_0$,

$$\begin{aligned}\|\rho_R e^{-itH}\varphi\|^p_{L^p([-T,T];L^q(\mathbb{R}^d))} &\lesssim \sum_{k=-[T/T_0]}^{[T/T_0]+1}\|\rho_R e^{-itH}e^{-i(k+1)H}\varphi\|^p_{L^p([-T_0,T_0];L^q(\mathbb{R}^d))}\\ &\lesssim (T/T_0)C^p_{T_0}\|\varphi\|^p_{L^2(\mathbb{R}^d)}.\end{aligned}$$

Let a_h be as in Proposition 2.1. Replacing ψ_ε with $\rho_R\psi_\varepsilon$ and taking $\varepsilon > 0$ smaller if necessary, we may assume without loss of generality that $\operatorname{supp} a_h \subset \{(x, \xi) : R < |x| < 1/h, |\xi| \in I\}$ for some open interval $I \Subset (0, \infty)$. Choosing $\theta^\pm \in C^\infty([-1, 1])$ so that $\theta^+ + \theta^- = 1$, $\theta^+ = 1$ on $[1/2, 1]$ and $\theta^+ = 0$ on $[-1, -1/2]$, we set $a^\pm_h(x, \xi) = a_h(x, \xi)\theta^\pm(\hat{x}\cdot\hat{\xi})$, where $\hat{x} = x/|x|$. It is clear that $\{a^\pm_h\}_{h\in(0,1]}$ is bounded in $S(1, g)$ and $\operatorname{supp} a^\pm_h \subset \Gamma^\pm(R, I, 1/2) \cap \{x : |x| < 1/h\}$, and that $a_h = a^+_h + a^-_h$. We now apply Proposition 3.1 to $a^\pm_h$ and obtain the local-in-time dispersive estimate for $\mathrm{Op}_h(a^\pm_h)e^{-itH}\,\mathrm{Op}_h(a^\pm_h)^*$ (uniformly in $h \in (0, h_0]$), which, combined with the L^2-boundedness of $\mathrm{Op}_h(a^\pm_h)e^{-itH}$ and the abstract theorem [Keel and Tao 1998], implies the following Strichartz estimates for $\mathrm{Op}_h(a_h)e^{-itH}$:

$$\begin{aligned}\|\mathrm{Op}_h(a_h)e^{-itH}\varphi\|_{L^p([-1,1];L^q(\mathbb{R}^d))} &\le \sum_\pm\|\mathrm{Op}_h(a^\pm_h)e^{-itH}\varphi\|_{L^p([-1,1];L^q(\mathbb{R}^d))}\\ &\le C\|\varphi\|_{L^2(\mathbb{R}^d)},\end{aligned}$$

uniformly with respect to $h \in (0, h_0]$. Since $\mathrm{Op}_h(a_h)$ is bounded from $L^2(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$ with the bound of order $O(h^{-d(1/2-1/q)})$, for $h_0 < h \le 1$, we have

$$\sum_{h_0<h\le1}\|\mathrm{Op}_h(a_h)e^{-itH}f(h^2H)\varphi\|^2_{L^p([-1,1];L^q(\mathbb{R}^d))} \le C(h_0)\|\varphi\|^2_{L^2(\mathbb{R}^d)}$$

with some $C(h_0) > 0$. Using these two bounds, we obtain

$$\begin{aligned}\sum_h\|\mathrm{Op}_h(a_h)e^{-itH}f(h^2H)\varphi\|^2_{L^p([-1,1];L^q(\mathbb{R}^d))} &\le C\sum_{0<h<h_0}\|f(h^2H)\varphi\|^2_{L^2(\mathbb{R}^d))} + C(h_0)\|\varphi\|^2_{L^2(\mathbb{R}^d)}\\ &\le C\|\varphi\|^2_{L^2(\mathbb{R}^d)}.\end{aligned}$$

On the other hand, Strichartz estimates for $\mathrm{Op}(\chi_\varepsilon)e^{-itH}$ are an immediate consequence of Proposition 3.2. Together with Corollary 2.6, this completes the proof. □

4. Semiclassical approximations for outgoing propagators

Throughout this section we assume Assumption 1.1 with $\mu > 0$. Here we study the behavior of

$$e^{-itH}\,\mathrm{Op}_h(a_h^{\pm})^*,$$

where $a_h^{\pm} \in S(1, g)$ are supported in $\Gamma^{\pm}(R, I, \sigma) \cap \{|x| < 1/h\}$, respectively. The main goal of this section is to prove Proposition 3.1. For simplicity, we consider the outgoing propagator $e^{-itH}\,\mathrm{Op}_h(a_h^{+})^*$ for $0 \le t \le 1$ only, and the proof for the incoming case is analogous.

In order to prove dispersive estimates, we construct a semiclassical approximation for the outgoing propagator $e^{-itH}\,\mathrm{Op}_h(a_h^{+})^*$ by using the method of Isozaki–Kitada. Namely, rescaling $t \mapsto th$ and setting $H^h = h^2 H$, $H_0^h = -h^2\Delta/2$, we consider an approximation for the semiclassical propagator $e^{-itH^h/h}\,\mathrm{Op}_h(a_h^{+})^*$ of the form

$$e^{-itH^h/h}\,\mathrm{Op}_h(a_h^{+})^* = J_h(S_h^{+}, b_h^{+})e^{-itH_0^h/h}J_h(S_h^{+}, c_h^{+})^* + O(h^N), \quad 0 \le t \le h^{-1},$$

where S_h^{+} solves a suitable eikonal equation in the outgoing region and $J(S_h^{+}, w)$ is the corresponding semiclassical Fourier integral operator (h-FIO for short):

$$J_h(S_h^{+}, w)f(x) = (2\pi h)^{-d}\int e^{i(S_h^{+}(x,\xi)-y\cdot\xi)/h}w(x, \xi)f(y)\,dy\,d\xi.$$

Such approximations (uniformly in time) have been studied for Schrödinger operators with long-range potentials [Robert and Tamura 1987] and for the case of long-range metric perturbations [Robert 1987; 1994; Bouclet and Tzvetkov 2007]. We also refer to the original paper by Isozaki and Kitada [1985], in which the existence and asymptotic completeness of modified wave operators (with time-independent modifiers) were established for the case of Schrödinger operators with long-range potentials. We note that, in these cases, we do not need the additional restriction of the initial data in $\{|x| < 1/h\}$. On the other hand, in [Mizutani 2013], we constructed such approximations (locally in time) for the case with long-range metric perturbations, combined with potentials growing subquadratically at infinity, under the additional restriction on the initial data into $\{|x| < 1/h\}$.

As we mentioned in the outline of the paper, we first construct an approximation for the modified propagator $e^{-it\widetilde{H}^h/h}$, where $\widetilde{H}^h$ is defined as follows. Taking arbitrarily a cut-off function $\psi \in C_0^\infty(\mathbb{R}^d)$ such that $0 \le \psi \le 1$, $\psi \equiv 1$ for $|x| \le 1/2$ and $\psi \equiv 0$ for $|x| \ge 1$, we define truncated electric and magnetic potentials, V_h and $A_h = (A_{h,j})_j$ by $V_h(x) := \psi(hx/L)V(x)$ and $A_{h,j}(x) = \psi(hx/L)A_j(x)$, respectively. It is easy to see that

$$V_h \equiv V,\ \ A_{h,j} \equiv A_j \text{ on } \{|x| \le L/(2h)\}, \quad \operatorname{supp} A_{h,j},\ \operatorname{supp} V_h \subset \{|x| \le L/h\},$$

and that, for any $\alpha \in \mathbb{Z}_+^d$, there exists $C_{L,\alpha} > 0$, independent of x, h, such that

$$h^2|\partial_x^\alpha V_h(x)| + h|\partial_x^\alpha A_h(x)| \le C_{\alpha,L}\langle x\rangle^{-\mu-|\alpha|}. \tag{4-1}$$

Let us define $\widetilde{H}^h$ by

$$\widetilde{H}^h = \frac{1}{2}\sum_{j,k=1}^{d}(-ih\partial_j - hA_{h,j}(x))g^{jk}(x)(-ih\partial_k - hA_{h,k}(x)) + h^2V_h(x).$$

We consider $\widetilde{H}^h$ as a "semiclassical" Schrödinger operator with h-dependent electromagnetic potentials h^2V_h and hA_h. By virtue of the estimates on g^{jk}, A_h, and V_h, $\widetilde{H}^h$ can be regarded as a long-range perturbation of the semiclassical free Schrödinger operator $H_0^h = -h^2\Delta/2$. Such a type modification has been used to prove Strichartz estimates and local smoothing effects (with loss of derivatives) for Schrödinger equations with superquadratic potentials; see [Yajima and Zhang 2004, Section 4]. Let us denote by $\tilde{p}_h$ the corresponding modified symbol

$$\tilde{p}_h(x,\xi) = \frac{1}{2}\sum_{j,k=1}^{d} g^{jk}(x)(\xi_j - hA_{h,j}(x))(\xi_k - hA_{h,k}(x)) + h^2V_h(x). \tag{4-2}$$

The following proposition provides the existence of the phase function of h-FIOs.

Proposition 4.1 [Robert 1994]. *Fix an open interval $I \Subset (0,\infty)$, $-1 < \sigma < 1$ and $L > 0$. Then there exist $R_0, h_0 > 0$ and a family of smooth and real-valued functions*

$$\{S_h^+ : 0 < h \le h_0,\, R \ge R_0\} \subset C^\infty(\mathbb{R}^{2d} : \mathbb{R})$$

satisfying the eikonal equation associated to $\tilde{p}_h$:

$$\tilde{p}_h(x, \partial_x S_h^+(x,\xi)) = |\xi|^2/2, \quad (x,\xi) \in \Gamma^+(R, I, \sigma), \tag{4-3}$$

such that

$$|S_h^+(x,\xi) - x\cdot\xi| \le C\langle x\rangle^{1-\mu}, \quad x,\xi \in \mathbb{R}^d. \tag{4-4}$$

Moreover, for any $|\alpha+\beta| \ge 1$,

$$|\partial_x^\alpha \partial_\xi^\beta (S_h^+(x,\xi) - x\cdot\xi)| \le C_{\alpha\beta}\min\{R^{1-\mu-|\alpha|}, \langle x\rangle^{1-\mu-|\alpha|}\}, \quad x,\xi \in \mathbb{R}^d. \tag{4-5}$$

Here $C, C_{\alpha\beta} > 0$ are independent of x, ξ, R, and h.

Proof. Since h^2V_h and hA_h are of long-range type uniformly with respect to $h \in (0,1]$ (the constant $C_{L,\alpha}$ in (4-1) can be taken independently of h), the proof is the same as that of [Robert 1994, Proposition 4.1], and we omit it. For the R dependence, we refer to [Bouclet and Tzvetkov 2007, Proposition 3.1]. □

Remark 4.2. The crucial point to obtain the estimates (4-4) and (4-5) is the uniform bound (4-1), and we do not have to use the support properties of A_h and V_h. Suppose that A and V satisfy $\langle x\rangle^{-1}|\partial_x^\alpha A(x)| + \langle x\rangle^{-2}|\partial_x^\alpha V(x)| \le C_{\alpha\beta}\langle x\rangle^{-|\alpha|}$, and g^{jk} satisfies Assumption 1.1 with $\mu \ge 0$. Then there exists $L > 0$, independent of h, such that if $0 < L \le L_0$, we can still construct the solution S_h^+ to (4-3) by using the support properties of A_h and V_h. However, in this case, $S_h^+ - x\cdot\xi$ behaves like $\langle x\rangle^{1-\mu}h^{-1}$ as $h \to 0$, and we cannot obtain the uniform L^2-boundedness of the corresponding h-FIO. This is one of the reasons why we exclude the critical case $\mu = 0$.

To the phase S_h^+ and an amplitude $a \in S(1, g)$, we associate the h-FIO defined by

$$J_h(S_h^+, a) f(x) = (2\pi h)^{-d} \int e^{i(S_h^+(x,\xi) - y\cdot\xi)/h} a(x, \xi) f(y)\, dy\, d\xi.$$

Using (4-5), for sufficiently large $R > 0$, we have

$$|\partial_\xi \otimes \partial_x S_h^+(x, \xi) - \mathrm{Id}\,| \le C\langle R\rangle^{-\mu} < \tfrac{1}{2}, \quad |\partial_x^\alpha \partial_\xi^\beta S_h^+(x, \xi)| \le C_{\alpha\beta} \text{ for } |\alpha + \beta| \ge 2,$$

uniformly in $h \in (0, h_0]$. Therefore, the standard L^2-boundedness of FIOs implies that $J_h(S_h^+, a)$ is uniformly bounded on $L^2(\mathbb{R}^d)$ with respect to $h \in (0, h_0]$.

We now construct the outgoing approximation for $e^{-it\widetilde{H}^h/h}$.

Theorem 4.3. *Let us fix arbitrarily open intervals $I \Subset I_0 \Subset I_1 \Subset I_2 \Subset (0, \infty)$, $-1 < \sigma < \sigma_0 < \sigma_1 < \sigma_2 < 1$ and $L > 0$. Let R_0 and h_0 be as in Proposition 4.1 with I, σ replaced by I_2, σ_2, respectively. Then, for every integer $N \ge 0$, the following hold uniformly with respect to $R \ge R_0$ and $h \in (0, h_0]$.*

(1) *There exists a symbol*

$$b_h^+ = \sum_{j=0}^{N-1} h^j b_{h,j}^+ \quad \text{with } b_{h,j}^+ \in S(\langle x\rangle^{-j}\langle\xi\rangle^{-j}, g),\ \operatorname{supp} b_{h,j}^+ \subset \Gamma^+(R^{1/3}, I_1, \sigma_1),$$

such that, for any $a^+ \in S(1, g)$ with $\operatorname{supp} a^+ \subset \Gamma^+(R, I, \sigma)$, we can find

$$c_h^+ = \sum_{j=0}^{N-1} h^j c_{h,j}^+ \quad \text{with } c_{h,j}^+ \in S(\langle x\rangle^{-j}\langle\xi\rangle^{-j}, g),\ \operatorname{supp} c_{h,j}^+ \subset \Gamma^+(R^{1/2}, I_0, \sigma_0),$$

such that, for all $0 \le t \le h^{-1}$, $e^{-it\widetilde{H}^h/h} \operatorname{Op}_h(a^+)^*$ can be brought to the form

$$e^{-it\widetilde{H}^h/h} \operatorname{Op}_h(a^+)^* = J_h(S_h^+, b_h^+) e^{-itH_0^h/h} J_h(S_h^+, c_h^+)^* + Q_{\mathrm{IK}}^+(t, h, N),$$

where $J_h(S_h^+, w)$, $w = b_h^+, c_h^+$, are h-FIOs associated to the phase S_h^+ defined in Proposition 4.1 with R, I, and σ replaced by $R^{1/4}$, I_2, and σ_2, respectively. Moreover, for any integer $s \ge 0$ with $2s \le N - 1$, the remainder $Q_{\mathrm{IK}}^+(t, h, N)$ satisfies

$$\|\langle D\rangle^s Q_{\mathrm{IK}}^+(t, h, N)\langle D\rangle^s\|_{L^2\to L^2} \le C_{Ns} h^{N-2s-1}, \tag{4-6}$$

uniformly with respect to $h \in (0, h_0]$ and $0 \le t \le h^{-1}$.

(2) *Let $K_{S_h^+}(t, x, y)$ be the distribution kernel of $J_h(S_h^+, b_h^+) e^{-itH_0^h/h} J_h(S_h^+, c_h^+)^*$. Then $K_{S_h^+}$ satisfies the dispersive estimate*

$$|K_{S_h^+}(t, x, y)| \le C|th|^{-d/2}, \tag{4-7}$$

uniformly with respect to $h \in (0, h_0]$, $x, y \in \mathbb{R}^d$ and $0 \le t \le h^{-1}$.

Proof. This theorem is basically known; hence we omit the proof. For the construction of the amplitudes b_h^+ and c_h^+, we refer to [Robert 1994, Section 4]; see also [Bouclet and Tzvetkov 2007, Section 3]. The remainder estimate (4-6) can be proved by the same argument as that in [Bouclet and Tzvetkov 2007, Proposition 3.3, Lemma 3.4] combined with the simple estimate

$$\|\langle D\rangle^s(\widetilde{H}^h+C_1)^{-s/2}\|_{L^2\to L^2}\le C_s h^{-s},\quad s\ge 0.$$

where $C_1>0$ is a large constant. Note that this estimate follow from the obvious bounds

$$\|\langle D\rangle^s\langle hD\rangle^{-s}\|_{L^2\to L^2}\le C_s h^{-s},\quad s\ge 0,$$

and the fact that $(\tilde{p}_h+h\tilde{p}_{1,h}+C_1)^{-s/2}\in S(\langle\xi\rangle^{-s},g)$ since $\tilde{p}_h+h\tilde{p}_{1,h}+C_1$ is uniformly elliptic for sufficiently large $C_1>0$. The dispersive estimate (4-7) can be verified by the same argument as that in [Bouclet and Tzvetkov 2007, Lemma 4.4]. □

The following lemma, which has been essentially proved in [Mizutani 2013], tells us that we can still construct the semiclassical approximation for the original propagator $e^{-itH^h/h}$ if we restrict the support of initial data in the region $\Gamma^+(R,J,\sigma)\cap\{x:|x|<h^{-1}\}$.

Lemma 4.4. *Suppose that $\{a_h^+\}_{h\in(0,1]}$ is a bounded set in $S(1,g)$ with symbols supported in*

$$\Gamma^+(R,I,\sigma)\cap\{x:|x|<h^{-1}\}.$$

There exists $L>1$ such that, for any $M,s\ge 0$, $h\in(0,h_0]$ and $0\le t\le h^{-1}$, we have

$$\|(e^{-itH^h/h}-e^{-it\widetilde{H}^h/h})\operatorname{Op}_h(a_h^+)^*\langle D\rangle^s\|_{L^2\to L^2}\le C_{M,s}h^{M-s},$$

where $C_{M,s}>0$ is independent of h and t.

In order to prove this lemma, we need the following.

Lemma 4.5. *Let $f_h\in C^\infty(\mathbb{R}^d)$ be such that for any $\alpha\in\mathbb{Z}_+^d$,*

$$|\partial_x^\alpha f_h(x)|\le C_\alpha$$

uniformly with respect to $h\in(0,h_0]$ and such that $\operatorname{supp} f_h\subset\{|x|\ge L/(2h)\}$. *Let $L>1$ be large enough. Then, under the conditions in Lemma 4.4, we have*

$$\|f_h(x)\langle D\rangle^\gamma e^{-it\widetilde{H}^h/h}\operatorname{Op}_h(a_h^+)^*\langle D\rangle^s\|_{L^2\to L^2}\le C_{M,s,\gamma}h^{M-s-\gamma},$$

for any $s,\gamma\ge 0$ and $M\ge 0$, uniformly with respect $h\in(0,h_0]$ and $0\le t\le 1/h$.

Proof. We apply Theorem 4.3 to $e^{-it\widetilde{H}^h/h}\operatorname{Op}_h(a_h^+)^*$ and obtain

$$e^{-it\widetilde{H}^h/h}\operatorname{Op}_h(a_h^+)^*=J_h(S_h^+,b_h^+)e^{-itH_0^h/h}J_h(S_h^+,c_h^+)^*+Q_{\mathrm{IK}}^+(t,h,N).$$

By virtue of (4-6), the remainder $f_h(x)\langle D\rangle^\gamma Q_{\mathrm{IK}}^+(t,h,N)\langle D\rangle^s$ is bounded on $L^2(\mathbb{R}^d)$ with the norm dominated by $C_{Ns\gamma}h^{N-\gamma-s-1}$, uniformly with respect $h\in(0,h_0]$ and $t\in[0,1/h]$. On the other hand, by virtue of (4-5), the phase of $K_{S_h^+}(t,x,y)$, which is given by

$$\Phi_h^+(t,x,y,\xi)=S_h^+(x,\xi)-\tfrac{1}{2}t|\xi|^2-S_h^+(y,\xi),$$

satisfies $\partial_\xi\Phi_h^+(t,x,y,\xi)=(x-y)(\mathrm{Id}+O(R^{-\mu/4}))-t\xi$. Here we recall that

$$\operatorname{supp} c_h^+ \subset \{(y,\xi)\in\mathbb{R}^{2d} : a_h^+(y,\partial_\xi S_h^+(y,\xi))\neq 0\};$$

see [Mizutani 2013, Lemma 3.2] and its proof. In particular, $c_h^+(y,\xi)$ vanishes in the region $\{y : |y|\geq 1/h\}$. We now set $L=4\sqrt{\sup I_2}+2$, where I_2 is given in Theorem 4.3. Since $|x|\geq L/(2h)$, $|y|<1/h$, and $|\xi|^2\in I_2$ on the support of the amplitude $f_h(x)b_h^+(x,\xi)\overline{c_h^+(y,\xi)}$, we obtain

$$|\partial_\xi\Phi_h^+(t,x,y,\xi)|>c(1+|x|+|y|+|\xi|+t+h^{-1}),\quad 0\leq t\leq h^{-1},$$

for some universal constant $c>0$. The assertion now follows from an integration by parts and the L^2-boundedness of h-FIOs. □

Proof of Lemma 4.4. The Duhamel formula yields

$$\begin{aligned}(e^{-itH^h/h}&-e^{-it\widetilde{H}^h/h})\\ &=-\frac{i}{h}\int_0^t e^{-i(t-s)H^h/h}W_0^h e^{-is\widetilde{H}^h/h}\,ds\\ &=-\frac{i}{h}\int_0^t e^{-i(t-s)H^h/h}e^{-is\widetilde{H}^h/h}W_0^h\,ds+\frac{1}{h^2}\int_0^t e^{-i(t-s)H^h/h}\int_0^s e^{-i(s-\tau)\widetilde{H}^h/h}[\widetilde{H}^h,W_0^h]e^{-i\tau\widetilde{H}^h/h}\,d\tau\,ds,\end{aligned}$$

where $W_0^h:=H^h-\widetilde{H}^h$ consists of two parts,

$$\frac{ih^2}{2}\sum_{j,k}(\partial_j g^{jk}(1-\psi(hx/L))A_k+(1-\psi(hx/L))A_j g^{jk}\partial_k)$$

and

$$\frac{h^2}{2}\sum_{j,k}(1-\psi(hx/L))^2 g^{jk}A_jA_k+h^2(1-\psi(hx/L))V.$$

In particular, W_0^h is a first order differential operator of the form

$$h^2\sum_{|\alpha|=1}f_\alpha^h(x)\partial_x^\alpha+h^2 f_0^h(x),$$

where f_α^h, f_0^h are supported in $\{|x|\geq L/(2h)\}$ and satisfy

$$|\partial_x^\beta f_\alpha^h(x)|\leq C_{\alpha\beta}\langle x\rangle^{1-\mu-|\beta|},\quad |\partial_x^\beta f_0^h(x)|\leq C_{\alpha\beta}\langle x\rangle^{2-\mu-|\beta|}. \tag{4-8}$$

Since $\{|x|\geq L/(2h)\}\cap\pi_x(\operatorname{supp} a_h^+)=\varnothing$ if $L>1$, we have

$$\|W_0^h\operatorname{Op}_h(a_h^+)^*\langle D\rangle^s\|_{L^2\to L^2}\leq C_{M,s}h^{M-s},\quad M\geq 0,\ s\in\mathbb{R}.$$

Therefore, the first term of the right-hand side of the above Duhamel formula satisfies the desired estimates since $e^{-itH^h/h}$ and $e^{-it\widehat{H}^h/h}$ are unitary on L^2.

We next study the second term. Again by the Duhamel formula, we have

$$[\widetilde{H}^h,W_0^h]e^{-i\tau\widetilde{H}^h/h}=e^{-i\tau\widetilde{H}^h/h}[\widetilde{H}^h,W_0^h]+\frac{i}{h}\int_0^\tau e^{-i(\tau-u)\widetilde{H}^h/h}[\widetilde{H}^h,[\widetilde{H}^h,W_0^h]]e^{-iu\widetilde{H}^h/h}\,du.$$

Since the coefficients of the commutator $[\widetilde{H}^h, W_0^h]$ are supported in $\{|x| \geq L/(2h)\}$, the support property of a_h^+ again implies that $[\widetilde{H}^h, W_0^h] \operatorname{Op}_h(a_h^+)^* \langle D \rangle^s = O_{L^2 \to L^2}(h^{M-s})$ for any $M \geq 0$ and $s \in \mathbb{R}$. Furthermore, by virtue of (4-1), (4-8), and the symbolic calculus, the coefficients of $[\widetilde{H}^h, [\widetilde{H}^h, W_0^h]]$ are uniformly bounded in x and supported in $\{|x| \geq L/(2h)\}$. We now apply Lemma 4.5 to

$$[\widetilde{H}^h, [\widetilde{H}^h, W_0^h]] e^{-iu\widetilde{H}^h/h} \operatorname{Op}_h(a_h^+)^*$$

and obtain the assertion. □

Proof of Proposition 3.1. Rescaling $t \to th$, it suffices to show

$$\|\operatorname{Op}_h(a_h^+) e^{-itH^h/h} \operatorname{Op}_h(a_h^+)^*\|_{L^1 \to L^\infty} \leq C_\varepsilon |th|^{-d/2}, \quad 0 < |t| \leq h^{-1},$$

where $H^h = h^2 H$. Let $A_h(x, y)$ be the distribution kernel of $\operatorname{Op}_h(a_h^+)$:

$$A_h(x, y) = (2\pi h)^{-d} \int e^{i(x-y)\cdot\xi/h} a_h^+(x, \xi)\, d\xi.$$

Since $a_h^+ \in S(1, g)$ is compactly supported in I with respect to ξ, we easily see that

$$\sup_x \int |A_h(x, y)|\, dy + \sup_y \int |A_h(x, y)|\, dx \leq C, \quad h \in (0, 1].$$

Moreover, since $\langle \xi \rangle^s a_h^+ \langle \xi \rangle^\gamma \in S(1, g)$ for any s, γ, we have

$$\|\langle D \rangle^s \operatorname{Op}_h(a_h^+) \langle D \rangle^\gamma\|_{L^2 \to L^2} \leq C_s h^{-s-\gamma}. \tag{4-9}$$

Combining these two estimates with Theorem 4.3 and Lemma 4.4, we can write

$$\operatorname{Op}_h(a_h^+) e^{-itH^h/h} \operatorname{Op}_h(a_h^+)^* = K_1(t, h, N) + K_2(t, h, N),$$

where

$$K_1(t, h, N) = \operatorname{Op}_h(a_h^+) J_h(S_h^+, b_h^+) e^{-itH_0^h/h} J_h(S_h^+, c_h^+)^*,$$
$$K_2(t, h, N) = \operatorname{Op}_h(a_h^+) Q_{\mathrm{IK}}^+(t, h, N) + \operatorname{Op}_h(a_h^+)(e^{-itH^h/h} - e^{-it\widetilde{H}^h/h}) \operatorname{Op}_h(a_h^+)^*.$$

By (4-7), the distribution kernel of $K_1(t, h, N)$, which we denote by $K_1(t, x, y)$, satisfies

$$|K_1(t, x, y)| \leq \int |A_h(x, z)| |K_{S_h^+}(t, z, y)|\, dz \leq C_N |th|^{-d/2}, \quad 0 < t \leq h^{-1},$$

uniformly in $h \in (0, h_0]$. On the other hand, (4-6), Lemma 4.4, and (4-9) imply

$$\|\langle D \rangle^s K_2(t, h, N) \langle D \rangle^s\|_{L^2 \to L^2} \leq C_{N,s} h^{N-2s-1}.$$

If we choose $N \geq d + 2$ and $s > d/2$, it follows from the Sobolev embedding that the distribution kernel of $K_2(t, h, N)$ is uniformly bounded in $\mathbb{R}^{2d}$ with respect to $h \in (0, h_0]$ and $0 < t \leq h^{-1}$. Therefore, $\operatorname{Op}_h(a_h^+) e^{-itH^h/h} \operatorname{Op}_h(a_h^+)^*$ has the distribution kernel $K(t, x, y)$ satisfying dispersive estimates for $0 < t \leq h^{-1}$:

$$|K(t, x, y)| \leq C_N |th|^{-d/2}, \quad x, y \in \mathbb{R}^d. \tag{4-10}$$

Finally, using the relation

$$\mathrm{Op}_h(a_h^+)e^{-itH^h/h}\,\mathrm{Op}_h(a_h^+)^* = (\,\mathrm{Op}_h(a_h^+)e^{itH^h/h}\,\mathrm{Op}_h(a_h^+)^*)^*,$$

we learn that $K(t,x,y)=\overline{K(-t,y,x)}$ and (4-10) also holds for $0<-t\le h^{-1}$. For the incoming case, the proof is analogous and we omit it. $\square$

5. Fourier integral operators with the time dependent phase

Throughout this section we assume Assumption 1.1 with $\mu\ge 0$. Consider a symbol $\chi_\varepsilon\in S(1,g)$ supported in a region

$$\Omega(\varepsilon):=\{(x,\xi)\in\mathbb{R}^{2d}:\langle x\rangle>\varepsilon|\xi|/2\},$$

where $\varepsilon>0$ is an arbitrarily small fixed constant. In this section we prove the dispersive estimate

$$\|\mathrm{Op}(\chi_\varepsilon)e^{-itH}\,\mathrm{Op}(\chi_\varepsilon)^*\|_{L^1\to L^\infty}\le C_\varepsilon|t|^{-d/2},\quad 0<|t|\le t_\varepsilon,$$

where $t_\varepsilon>0$ is a small constant depending on ε. This estimate, combined with the L^2-boundedness of $\mathrm{Op}(\chi_\varepsilon)$ and e^{-itH}, implies the Strichartz estimates for $\mathrm{Op}(\chi_\varepsilon)e^{-itH}$.

Let us give a short summary of the steps of the proof. Choose $\chi_\varepsilon^*\in S(1,g)$ so that $\operatorname{supp}\chi_\varepsilon^*=\operatorname{supp}\chi_\varepsilon$ and $\mathrm{Op}(\chi_\varepsilon)^*=\mathrm{Op}(\chi_\varepsilon^*)+\mathrm{Op}(r_N)$ with some $r_N\in S(\langle x\rangle^{-N}\langle\xi\rangle^{-N},g)$ for sufficiently large $N>d/2$. We first construct an approximation for $e^{-itH}\,\mathrm{Op}(\chi_\varepsilon^*)$ in terms of the FIO with a time dependent phase

$$J(\Psi,b^N)f(x)=\frac{1}{(2\pi)^d}\int e^{i(\Psi(t,x,\xi)-y\cdot\xi)}b(t,x,\xi)f(y)\,dy\,d\xi,$$

where Ψ is a generating function of the Hamilton flow associated to $p(x,\xi)$ and $(\partial_\xi\Psi,\xi)\mapsto(x,\partial_x\Psi)$ is the corresponding canonical map, and the amplitude

$$b=b_0+b_2+\cdots+b_{N-1}$$

solves the corresponding transport equations. Although such parametrix constructions are well known as WKB approximations (at least if χ_ε^* is compactly supported in ξ and the time scale depends on the size of frequency), we give the details of the proof since, in the present case, $\operatorname{supp}\chi_\varepsilon^*$ is not compact with respect to ξ and t_ε is independent of the size of frequency. The crucial point is that $p(x,\xi)$ is of *quadratic type* on $\Omega(\varepsilon)$:

$$|\partial_x^\alpha\partial_\xi^\beta p(x,\xi)|\le C_{\alpha\beta},\quad (x,\xi)\in\Omega(\varepsilon),\ |\alpha+\beta|\ge 2,$$

which allows us to follow a classical argument (see, for example, [Kitada and Kumano-go 1981]) and construct the approximation for $|t|<t_\varepsilon$ if $t_\varepsilon>0$ is small enough. The composition $\mathrm{Op}(\chi_\varepsilon)J(\Psi,b)$ is also an FIO with the same phase, and a standard stationary phase method can be used to prove dispersive estimates for $0<|t|<t_\varepsilon$. It remains to obtain the $L^1\to L^\infty$ bounds of the remainders $\mathrm{Op}(\chi_\varepsilon)e^{-itH}\,\mathrm{Op}(r_N)$ and $\mathrm{Op}(\chi_\varepsilon)e^{-itH}(\,\mathrm{Op}(\chi_\varepsilon^*)-J(\Psi,b^N))$. If e^{-itH} maps from the Sobolev space $H^{d/2}(\mathbb{R}^d)$ to itself, then $L^1\to L^\infty$ bounds are direct consequences of the Sobolev embedding and L^2-boundedness of PDOs. However, our Hamiltonian H is not bounded below (on $\{|x|\gtrsim|\xi|\}$) and such a property does not hold in

general. To overcome this difficulty, we use an Egorov-type lemma as follows. By the Sobolev embedding and the Littlewood–Paley decomposition, the proof is reduced to that of the estimate

$$\sum_{j\geq 0} \|2^{j\gamma} S_j(D)\,\mathrm{Op}(\chi_\varepsilon)e^{-itH}\,\mathrm{Op}(r_N)\langle D\rangle^\gamma f\|_{L^2}^2 \leq C\|f\|_{L^2}^2, \tag{5-1}$$

where $\gamma > d/2$ and S_j is a dyadic partition of unity. Then we will prove that there exists $\eta_j(t,\cdot,\cdot)\in S(1,g)$ such that

$$2^j \leq C(1+|x|+|\xi|) \quad \text{on } \operatorname{supp}\eta_j(t),$$

and that

$$S_j(D)\,\mathrm{Op}(\chi_\varepsilon)e^{-itH} = e^{-itH}\,\mathrm{Op}(\eta_j(t)) + O_{L^2\to L^2}(2^{-jN}), \quad |t|<t_\varepsilon \ll 1.$$

Choosing $\delta>0$ with $\gamma+\delta\leq N/2$, we learn that $2^{j(\gamma+\delta)}\eta_j(t)r_N\langle\xi\rangle^\gamma\in S(1,g)$, and hence (5-1) holds. $\mathrm{Op}(\chi_\varepsilon)e^{-itH}(\mathrm{Op}(\chi_\varepsilon^*)-J(\Psi,b))$ can be controlled similarly.

Short-time behavior of the Hamilton flow. We now discuss the classical mechanics generated by $p(x,\xi)$. We denote by $(X(t),\Xi(t))=(X(t,x,\xi),\Xi(t,x,\xi))$ the solution to the Hamilton equations

$$\begin{cases} \dot X_j = \dfrac{\partial p}{\partial \xi_j}(X,\Xi) = \displaystyle\sum_k g^{jk}(X)(\Xi_k - A_k(X)),\\ \dot\Xi_j = -\dfrac{\partial p}{\partial x_j}(X,\Xi)\\ \quad = -\dfrac12\displaystyle\sum_{k,l}\dfrac{\partial g^{kl}}{\partial x_j}(X)(\Xi_k-A_k(X))(\Xi_l-A_l(X)) + \sum_{k,l} g^{kl}(X)\dfrac{\partial A_k}{\partial x_j}(X)(\Xi_l-A_l(X)) - \dfrac{\partial V}{\partial x_j}(X) \end{cases}$$

with the initial condition $(X(0),\Xi(0))=(x,\xi)$, where $\dot f=\partial_t f$. We first observe that the flow conserves the energy:

$$p(x,\xi)=p(X(t),\Xi(t)),$$

which, combined with the uniform ellipticity of g^{jk}, implies

$$\begin{aligned}|\Xi(t)-A(X(t))|^2 &\lesssim p(X(t),\Xi(t))-V(X(t))\\ &= p(x,\xi)-V(X(t))\\ &\lesssim |\xi-A(x)|^2+|V(x)|+|V(X(t))|,\end{aligned}$$

and hence $|\Xi(t)|\lesssim|\xi|+\langle x\rangle+\langle X(t)\rangle$. By the Hamilton equation, we then have

$$|\dot X(t)|+|\dot\Xi(t)|\leq C(1+|\xi|+|x|+|X(t)|+|\Xi(t)|).$$

Applying Gronwall's inequality to this estimate, we obtain an a priori bound:

$$|X(t)-x|+|\Xi(t)-\xi|\leq C_T|t|(1+|x|+|\xi|), \quad |t|\leq T,\ x,\xi\in\mathbb{R}^d.$$

Using this estimate, we obtain more precise behavior of the flow with initial conditions in $\Omega(\varepsilon)$.

Lemma 5.1. *Let $\varepsilon > 0$. Then, for sufficiently small $t_\varepsilon > 0$ and all $\alpha, \beta \in \mathbb{Z}_+^d$,*

$$|\partial_x^\alpha \partial_\xi^\beta (X(t,x,\xi) - x)| + |\partial_x^\alpha \partial_\xi^\beta (\Xi(t,x,\xi) - \xi| \le C_{\alpha\beta\varepsilon} |t| \langle x \rangle^{1-|\alpha+\beta|},$$

uniformly with respect to $(t,x,\xi) \in (-t_\varepsilon, t_\varepsilon) \times \Omega(\varepsilon)$.

Proof. We only consider the case with $t \ge 0$, the proof for the opposite case is similar. Let $(x,\xi) \in \Omega(\varepsilon)$. First we remark that, for sufficiently small $t_\varepsilon > 0$,

$$|x|/2 \le |X(t,x,\xi)| \le 2\langle x \rangle, \quad |t| \le t_\varepsilon. \tag{5-2}$$

For $|\alpha+\beta| = 0$, the assertion is obvious. We let $|\alpha+\beta| = 1$ and differentiate the Hamilton equations with respect to $\partial_x^\alpha \partial_\xi^\beta$:

$$\frac{d}{dt}\begin{pmatrix} \partial_x^\alpha \partial_\xi^\beta X \\ \partial_x^\alpha \partial_\xi^\beta \Xi \end{pmatrix} = \begin{pmatrix} \partial_x \partial_\xi p(X,\Xi) & \partial_\xi^2 p(X,\Xi) \\ -\partial_x^2 p(X,\Xi) & -\partial_\xi \partial_x p(X,\Xi) \end{pmatrix} \begin{pmatrix} \partial_x^\alpha \partial_\xi^\beta X \\ \partial_x^\alpha \partial_\xi^\beta \Xi \end{pmatrix}. \tag{5-3}$$

Using (5-2), we learn that $p(X(t), \Xi(t))$ is of quadratic type in $\Omega(\varepsilon)$:

$$|(\partial_x^\alpha \partial_\xi^\beta p)(X(t), \Xi(t))| \le C_{\alpha\beta\varepsilon} \langle x \rangle^{2-|\alpha+\beta|}, \quad (t,x,\xi) \in (-t_\varepsilon, t_\varepsilon) \times \Omega(\varepsilon).$$

Hence all entries of the above matrix are uniformly bounded in $(t,x,\xi) \in (-t_\varepsilon, t_\varepsilon) \times \Omega(\varepsilon)$. Taking $t_\varepsilon > 0$ smaller if necessary, integrating (5-3) with respect to t, and applying Gronwall's inequality, we have the assertion with $|\alpha+\beta| = 1$. For $|\alpha+\beta| \ge 2$, we prove the estimate for $\partial_{\xi_1}^2 X(t)$ and $\partial_{\xi_1}^2 \Xi(t)$ only, where $\xi = (\xi_1, \xi_2, \dots, \xi_d)$. Proofs for other cases are similar, and proofs for higher derivatives follow from an induction on $|\alpha+\beta|$. By the Hamilton equation, we learn

$$\frac{d}{dt}\partial_{\xi_1}^2 X(t) = \partial_x \partial_\xi p(X(t), \Xi(t)) \partial_{\xi_1}^2 X(t) + \partial_\xi^2 p(X(t), \Xi(t)) \partial_{\xi_1}^2 \Xi(t) + Q(X(t), \Xi(t)),$$

where $Q(X(t), \Xi(t))$ satisfies

$$\begin{aligned} |Q(X(t), \Xi(t))| &\le C_\varepsilon \sum_{|\alpha+\beta|=3, |\beta| \ge 1} |(\partial_x^\alpha \partial_\xi^\beta p)(X(t), \Xi(t))| \, |\partial_{\xi_1} X(t)|^{|\alpha|} |\partial_{\xi_1} \Xi(t)|^{|\beta|} \\ &\le C_\varepsilon \langle x \rangle^{-1}. \end{aligned}$$

We similarly obtain

$$\frac{d}{dt}\partial_{\xi_1}^2 \Xi(t) = -\partial_x^2 p(X(t), \Xi(t)) \partial_{\xi_1}^2 X(t) - \partial_\xi \partial_x p(X(t), \Xi(t)) \partial_{\xi_1}^2 \Xi(t) + O(\langle x \rangle^{-1}).$$

Applying Gronwall's inequality, we have the desired estimates. □

Lemma 5.2. (1) *Let $t_\varepsilon > 0$ be small enough. Then, for any $|t| < t_\varepsilon$, the map*

$$g(t) : (x,\xi) \mapsto (X(t,x,\xi), \xi)$$

is a diffeomorphism from $\Omega(\varepsilon/2)$ onto its range, and satisfies

$$\Omega(\varepsilon) \subset g(t, \Omega(\varepsilon/2)) \quad \textit{for all } |t| < t_\varepsilon.$$

(2) *Let* $\Omega(\varepsilon) \ni (x,\xi) \mapsto (Y(t,x,\xi),\xi) \in \Omega(\varepsilon/2)$ *be the inverse map of* $g(t)$. *Then* $Y(t,x,\xi)$ *and* $\Xi(t,Y(t,x,\xi),\xi)$ *satisfy the same estimates as those for* $X(t,x,\xi)$ *and* $\Xi(t,x,\xi)$ *of Lemma 5.1*, *respectively*:

$$|\partial_x^\alpha \partial_\xi^\beta (Y(t,x,\xi)-x)| + |\partial_x^\alpha \partial_\xi^\beta (\Xi(t,Y(t,x,\xi),\xi)-\xi| \le C_{\alpha\beta\varepsilon}|t|\langle x\rangle^{1-|\alpha+\beta|},$$

uniformly with respect to $(t,x,\xi) \in (-t_\varepsilon,t_\varepsilon)\times\Omega(\varepsilon)$.

Proof. Choosing a cutoff function $\rho \in S(1,g)$ such that $0 \le \rho \le 1$, $\operatorname{supp}\rho \subset \Omega(\varepsilon/3)$, and $\rho \equiv 1$ on $\Omega(\varepsilon/2)$, we modify $g(t)$ as follows:

$$g_\rho(t,x,\xi) = (X_\rho(t,x,\xi),\xi), \quad X_\rho(t,x,\xi) = (1-\rho(x,\xi))x + \rho(x,\xi)X(t,x,\xi).$$

It is easy to see that, for $(t,x,\xi) \in (-t_\varepsilon,t_\varepsilon)\times\Omega(\varepsilon/2)$, $g_\rho(t,x,\xi)$ is smooth and Lemma 5.1 implies

$$\begin{aligned} |\partial_x^\alpha \partial_\xi^\beta g_\rho(t,x,\xi)| &\le C_{\alpha\beta\varepsilon}, \quad |\alpha+\beta| \ge 1, \\ |J(g_\rho)(t,x,\xi) - \mathrm{Id}| &\le C_\varepsilon t_\varepsilon, \end{aligned}$$

where $J(g_\rho)$ is the Jacobi matrix with respect to (x,ξ) and the constant $C_\varepsilon > 0$ is independent of t, x, and ξ. Choosing $t_\varepsilon > 0$ so small that $C_\varepsilon t_\varepsilon < 1/2$, and applying the Hadamard global inverse mapping theorem, we see that, for any fixed $|t| < t_\varepsilon$, $g_\rho(t)$ is a diffeomorphism from $\mathbb{R}^{2d}$ onto itself. By definition, $g(t)$ is diffeomorphic from $\Omega(\varepsilon/2)$ onto its range. Since $g_\rho(t)$ is bijective, it remains to check that

$$\Omega(\varepsilon)^c \supset g_\rho(t,\Omega(\varepsilon/2)^c), \quad |t| < t_\varepsilon.$$

Suppose that $(x,\xi) \in \Omega(\varepsilon/2)^c$. If $(x,\xi) \in \Omega(\varepsilon/3)^c$, the assertion is obvious since $g_\rho(t) \equiv \mathrm{Id}$ outside $\Omega(\varepsilon/3)$. If $(x,\xi) \in \Omega(\varepsilon/3)\setminus\Omega(\varepsilon/2)$, then, by Lemma 5.1 and the support property of ρ, we have

$$|X_\rho(t,x,\xi)| \le |x| + \rho(x,\xi)|(X(t,x,\xi)-x)| \le (\varepsilon/2 + C_0 t_\varepsilon)\langle\xi\rangle$$

for some $C_0 > 0$ independent of x, ξ, and t_ε. Choosing $t_\varepsilon < \varepsilon/(2C_0)$, we obtain the assertion.

We next prove the estimates on $Y(t)$. Since $(Y(t,x,\xi),\xi) \in \Omega(\varepsilon/2)$, we learn

$$\begin{aligned} |Y(t,x,\xi)-x| &= |X(0,Y(t,x,\xi),\xi) - X(t,Y(t,x,\xi),\xi)| \\ &\le \sup_{(x,\xi)\in\Omega(\varepsilon/2)} |X(t,x,\xi)-x| \\ &\le C_\varepsilon|t|\langle x\rangle. \end{aligned}$$

For $\alpha,\beta \in \mathbb{Z}_+^d$ with $|\alpha+\beta| = 1$, apply $\partial_x^\alpha\partial_\xi^\beta$ to the equality $x = X(t,Y(t,x,\xi),\xi)$. We then have the equality

$$A(t,Z(t,x,\xi))\partial_x^\alpha\partial_\xi^\beta(Y(t,x,\xi)-x) = \partial_y^\alpha\partial_\eta^\beta(y - X(t,y,\eta))|_{(y,\eta)=Z(t,x,\xi)},$$

where $Z(t,x,\xi) = (Y(t,x,\xi),\xi)$ and $A(t,Z) = (\partial_x X)(t,Z)$ is a $d\times d$ matrix. By Lemma 5.1 and a similar argument to that in the proof of Lemma 5.2(1), we learn that $A(t,Z(t,x,\xi))$ is invertible if $t_\varepsilon > 0$ is small enough, and that $A(t,Z(t,x,\xi))$ and $A(t,Z(t,x,\xi))^{-1}$ are bounded uniformly in

$(t, x, \xi) \in (-t_\varepsilon, t_\varepsilon) \times \Omega(\varepsilon/2)$. Therefore,

$$\begin{aligned}|\partial_x^\alpha \partial_\xi^\beta (Y(t, x, \xi) - x)| &\le C_{\alpha\beta} \sup_{(x,\xi)\in\Omega(\varepsilon/2)} |\partial_x^\alpha \partial_\xi^\beta (x - X(t, x, \xi))| \\ &\le C_{\alpha\beta} |t| \langle x\rangle^{1-|\alpha+\beta|}.\end{aligned}$$

Proofs for higher derivatives are obtained by induction in $|\alpha+\beta|$ and proofs for $\Xi(t, Y(t, x, \xi), \xi)$ are similar. □

***The parametrix for* $\mathbf{Op}(\chi_\varepsilon)e^{-itH}\,\mathbf{Op}(\chi_\varepsilon)^*$.** Before starting the construction of parametrix, we prepare two lemmas. The following Egorov-type theorem will be used to control the remainder term. We write $\exp tH_p(x, \xi) = (X(t, x, \xi), \Xi(t, x, \xi))$.

Lemma 5.3. *For* $h \in (0, 1]$, *consider a* h*-dependent symbol* $\eta_h \in S(1, g)$ *such that* $\operatorname{supp}\eta_h \subset \Omega(\varepsilon) \cap \{1/(2h) < |\xi| < 2/h\}$. *Then, for sufficiently small* $t_\varepsilon > 0$, *independent of* h, *and any integer* $N \ge 0$, *there exists a bounded family of symbols*

$$\{\eta_h^N(t, \cdot, \cdot) : |t| < t_\varepsilon,\ 0 < h \le 1\} \subset S(1, g)$$

such that

$$\operatorname{supp}\eta_h^N(t, \cdot, \cdot) \subset \exp(-t)H_p(\operatorname{supp}\eta_h)$$

and

$$\|e^{itH}\operatorname{Op}(\eta_h)e^{-itH} - \operatorname{Op}(\eta_h^N(t))\|_{L^2\to L^2} \le C_{N\varepsilon}h^N,$$

uniformly with respect to $0 < h \le 1$ *and* $|t| < t_\varepsilon$.

Proof. Let $\eta_h^0(t, x, \xi) = \eta_h(\exp tH_p(x, \xi)) = \eta_h(X(t, x, \xi), \Xi(t, x, \xi))$. It is easy to see that

$$\operatorname{supp}\eta_h^0 \subset \exp(-t)H_p(\operatorname{supp}\eta_h).$$

Moreover, Lemma 5.1 implies that $\{\eta_h^0 : |t| < t_\varepsilon, 0 < h \le 1\}$ is a bounded subset of $S(1, g)$. By a direct computation, η_h^0 solves

$$\partial_t\eta_h^0 = \{p, \eta_h^0\}, \quad \eta_h^0|_{t=0} = \eta_h,$$

where $\{\cdot, \cdot\}$ is the Poisson bracket. Then, by standard pseudodifferential calculus, there exists a bounded set $\{r_h^0(t, \cdot, \cdot) : 0 \le t < t_\varepsilon, 0 < h \le 1\} \subset S(1, g)$ with $\operatorname{supp} r_h^0 \subset \exp(-t)H_p(\operatorname{supp}\eta_h)$ such that

$$\frac{d}{dt}\operatorname{Op}(\eta_h^0) = i[H, \operatorname{Op}(\eta_h^0)] + h\operatorname{Op}(r_h^0).$$

We next set

$$\eta_h^1(t, x, \xi) = \int_0^t r_h^0(s, X(t-s, x, \xi), \Xi(t-s, x, \xi))\,ds.$$

Again, we learn that $\{\eta_h^1(t, \cdot, \cdot) : |t| < t_\varepsilon, 0 < h \le 1\} \subset S(1, g)$ is also bounded and that

$$\operatorname{supp}\eta_h^1 \subset \exp(-t)H_p(\operatorname{supp}\eta_h)$$

for all $|t| < t_\varepsilon$ and $0 < h \le 1$. Moreover, η_h^1 solves

$$\partial_t \eta_h^1 = \{p, \eta_h^1\} + r_h^0, \quad \eta_h^1|_{t=0} = 0,$$

which implies

$$\frac{d}{dt}\operatorname{Op}(\eta_h^0 + h\eta_h^1) = i[H, \operatorname{Op}(\eta_h^0 + h\eta_h^1)] + h^2 \operatorname{Op}(r_h^1)$$

with some $\{r_h^1 : 0 \le t < t_\varepsilon, 0 < h \le 1\} \subset S(1, g)$ and $\operatorname{supp} r_h^1 \subset \exp(-t)H_p(\operatorname{supp}\eta_h)$. Iterating this procedure and putting $\eta_h^N = \sum_{j=0}^{N-1} h^j \eta_h^j$, we obtain the assertion. $\square$

Using this lemma, we have the following.

Lemma 5.4. *Let $\varepsilon > 0$. Then, for any symbol $\chi_\varepsilon \in S(1, g)$ with $\operatorname{supp}\chi_\varepsilon \subset \Omega(\varepsilon)$ and any integer $N \ge 1$, there exists $\chi_\varepsilon^* \in S(1, g)$ with $\operatorname{supp}\chi_\varepsilon^* \subset \Omega(\varepsilon)$ such that for any $\gamma < N/2$,*

$$\sup_{|t|<t_\varepsilon} \|\operatorname{Op}(\chi_\varepsilon)e^{-itH}\operatorname{Op}(\chi_\varepsilon)^* - \operatorname{Op}(\chi_\varepsilon)e^{-itH}\operatorname{Op}(\chi_\varepsilon^*)\|_{H^{-\gamma}(\mathbb{R}^d)\to H^\gamma(\mathbb{R}^d)} \le C_{N\gamma\varepsilon}.$$

Proof. By the expansion formula (2-4), there exists $\chi_\varepsilon^* \in S(1, g)$ with $\operatorname{supp}\chi_\varepsilon^* \subset \Omega(\varepsilon)$ such that

$$\operatorname{Op}(\chi_\varepsilon)^* = \operatorname{Op}(\chi_\varepsilon^*) + \operatorname{Op}(r_0(N))$$

with some $r_0(N) \in S(\langle x\rangle^{-N}\langle\xi\rangle^{-N}, g)$. For $\delta > 0$ with $2\gamma + \delta \le N$, we split

$$\langle D\rangle^\gamma \operatorname{Op}(\chi_\varepsilon)e^{-itH}\operatorname{Op}(r_0(N))\langle D\rangle^\gamma = \langle D\rangle^\gamma \operatorname{Op}(\chi_\varepsilon)e^{-itH}\langle D\rangle^{-\gamma-\delta}\langle x\rangle^{-\gamma-\delta}\cdot\langle x\rangle^{\gamma+\delta}\langle D\rangle^{\gamma+\delta}\operatorname{Op}(r_0(N))\langle D\rangle^\gamma.$$

Since $\langle x\rangle^{\gamma+\delta}\langle\xi\rangle^{\gamma+\delta}r_0(N)\langle\xi\rangle^\gamma \in S(1, g)$, $\langle x\rangle^{\gamma+\delta}\langle D\rangle^{\gamma+\delta}\operatorname{Op}(r_0(N))\langle D\rangle^\gamma$ is bounded on L^2. In order to prove the L^2-boundedness of the first term of the right hand side, we use the standard Littlewood–Paley decomposition and Lemma 5.3 as follows. Consider a dyadic partition of unity with respect to the frequency:

$$\sum_{j=0}^\infty S_j(D) = 1,$$

where $S_j(\xi) = S(2^{-j}\xi)$, $j \ge 1$, with some $S \in C_0^\infty(\mathbb{R}^d)$ supported in $\{1/2 < |\xi| < 2\}$ and $S_0 \in C_0^\infty(\mathbb{R}^d)$ supported in $\{|\xi| < 1\}$. Then

$$\|\langle D\rangle^\gamma \operatorname{Op}(\chi_\varepsilon)e^{-itH}\langle D\rangle^{-\gamma-\delta}\langle x\rangle^{-\gamma-\delta}f\|_{L^2} \le C\biggl(\sum_{j=0}^\infty \|2^{j\gamma}S_j(D)\operatorname{Op}(\chi_\varepsilon)e^{-itH}\langle D\rangle^{-\gamma-\delta}\langle x\rangle^{-\gamma-\delta}f\|_{L^2}^2\biggr)^{1/2}.$$

By the expansion formula (2-3), there exists a sequence of symbols $\eta_j \in S(1, g)$ supported in

$$\Omega(\varepsilon) \cap \{2^{j-1} < |\xi| < 2^{j+1}\}$$

such that

$$S_j(D)\operatorname{Op}(\chi_\varepsilon) = \operatorname{Op}(\eta_j) + Q_1(j, N), \quad \|Q_1(j, N)\|_{L^2\to L^2} = O(2^{-jN}).$$

We then learn from Lemma 5.3 with $h = 2^{-j}$ that there exists $\{\eta_j^N(t) : |t| < t_\varepsilon\} \subset S(1, g)$ such that

$$\operatorname{Op}(\eta_j)e^{-itH} = e^{-itH}\operatorname{Op}(\eta_j^N(t)) + Q_2(t, j, N), \quad \sup_{|t|<t_\varepsilon} \|Q_2(t, j, N)\|_{L^2\to L^2} = O(2^{-jN}).$$

Since $N \geq \gamma + \delta$, the remainder satisfies

$$\sup_{|t|<t_\varepsilon} \|2^{j\gamma}(Q_1(j, N)e^{-itH} + Q_2(t, j, N))\langle D\rangle^{-\gamma-\delta}\langle x\rangle^{-\gamma-\delta} f\|_{L^2}^2 \leq C2^{-2j\delta}\|f\|_{L^2}^2.$$

Suppose that $(x, \xi) \in \operatorname{supp} \eta_j^N(t)$. Since $\operatorname{supp} \eta_j^N(t) \subset \exp(-t)H_p(\operatorname{supp} \eta_j)$, we have

$$|X(t, x, \xi)| > \varepsilon\langle \Xi(t, x, \xi)\rangle, \quad 2^{j-1} < |\Xi(t, x, \xi)| < 2^{j+1}.$$

Using Lemma 5.1 with the initial data $(X(t, x, \xi), \Xi(t, x, \xi))$, we learn

$$|x - X(t, x, \xi)| + |\xi - \Xi(t, x, \xi)| \leq Ct_\varepsilon\langle X(t, x, \xi)\rangle, \quad |t| < t_\varepsilon.$$

Combining these two estimates, we see that

$$2^j \leq C(1 + |x| + |\xi|), \quad (x, \xi) \in \operatorname{supp} \eta_j^N(t), \ |t| < t_\varepsilon,$$

where the constant $C > 0$ is independent of x, ξ, and t, provided that $t_\varepsilon > 0$ is small enough. Therefore, $2^{j(\gamma+\delta)}\eta_j^N(t)\langle\xi\rangle^{-\gamma-\delta}\langle x\rangle^{-\gamma-\delta} \in S(1, g)$ and the corresponding PDO is bounded on L^2. Finally, we obtain

$$\begin{aligned}
\sum_{j=0}^{\infty} &\|2^{j\gamma}\operatorname{Op}(\eta_j)e^{-itH}\langle D\rangle^{-\gamma-\delta}\langle x\rangle^{-\gamma-\delta} f\|_{L^2}^2 \\
&\leq C\sum_{j=0}^{\infty}\bigl(\|2^{-j\delta}2^{j(\gamma+\delta)}\operatorname{Op}(\eta_j^N(t))\langle D\rangle^{-\gamma-\delta}\langle x\rangle^{-\gamma-\delta} f\|_{L^2}^2 + 2^{-2j\delta}\|f\|_{L^2}^2\bigr) \\
&\leq C\sum_{j=0}^{\infty} 2^{-2j\delta}\|f\|_{L^2}^2 \\
&\leq C\|f\|_{L^2}^2,
\end{aligned}$$

□

We now consider a parametrix construction of $\operatorname{Op}(\chi_\varepsilon)e^{-itH}\operatorname{Op}(\chi_\varepsilon^*)$. Let us first make the following ansatz:

$$v(t, x) = \frac{1}{(2\pi)^d}\int e^{i(\Psi(t,x,\xi)-y\cdot\xi)} b^N(t, x, \xi) f(y)\,dy\,d\xi,$$

where $b^N = \sum_{j=0}^{N-1} b_j$. In order to approximately solve the Schrödinger equation

$$i\partial_t v(t) = Hv(t), \quad v|_{t=0} = \operatorname{Op}(\chi_\varepsilon^*)\varphi,$$

the phase function Ψ and the amplitude b^N should satisfy respectively the Hamilton–Jacobi equation

$$\partial_t\Psi + p(x, \partial_x\Psi) = 0, \quad \Psi|_{t=0} = x\cdot\xi \tag{5-4}$$

and the transport equations

$$\begin{cases} \partial_t b_0 + \mathscr{X} \cdot \partial_x b_0 + \mathscr{Y} b_0 = 0, & b_0|_{t=0} = \chi_\varepsilon, \\ \partial_t b_j + \mathscr{X} \cdot \partial_x b_j + \mathscr{Y} b_j + iK b_{j-1} = 0, & b_j|_{t=0} = 0, \ 1 \le j \le N-1, \end{cases} \tag{5-5}$$

where K is the kinetic part of H, and the vector field $\mathscr{X}$ and function $\mathscr{Y}$ are defined by

$$\mathscr{X}_j(t, x, \xi) := (\partial_{\xi_j} p)(x, \partial_x \Psi(t, x, \xi)), \quad j = 1, \dots, d,$$
$$\mathscr{Y}(t, x, \xi) := [k(x, \partial_x)\Psi + p_1(x, \partial_x \Psi)](t, x, \xi).$$

Here p, p_1 are given by (1-6). We first construct the phase function Ψ.

Proposition 5.5. *Let us fix $\varepsilon > 0$ arbitrarily. Then, for sufficiently small $t_\varepsilon > 0$, we can construct a smooth and real-valued function $\Psi \in C^\infty((-t_\varepsilon, t_\varepsilon) \times \mathbb{R}^{2d}; \mathbb{R})$ which solves the Hamilton–Jacobi equation* (5-4) *for $(x, \xi) \in \Omega(\varepsilon)$ and $|t| \le t_\varepsilon$. Moreover, for all $\alpha, \beta \in \mathbb{Z}_+^d$, $x, \xi \in \mathbb{R}^d$ and $|t| \le t_\varepsilon$,*

$$|\partial_x^\alpha \partial_\xi^\beta (\Psi(t, x, \xi) - x \cdot \xi + tp(x, \xi)| \le C_{\alpha\beta\varepsilon} |t|^2 \langle x \rangle^{2-|\alpha+\beta|}, \tag{5-6}$$

where $C_{\alpha\beta\varepsilon} > 0$ is independent of x, ξ and t.

Proof. We consider the case when $t \ge 0$, and the proof for $t \le 0$ is similar. We first define the action integral $\widetilde{\Psi}(t, x, \xi)$ on $[0, t_\varepsilon) \times \Omega(\varepsilon/2)$ by

$$\widetilde{\Psi}(t, x, \xi) := x \cdot \xi + \int_0^t L(X(s, Y(t, x, \xi), \xi), \Xi(s, Y(t, x, \xi), \xi))\, ds,$$

where $L(x, \xi) = \xi \cdot \partial_\xi p(x, \xi) - p(x, \xi)$ is the Lagrangian associated to $p(x, \xi)$, and X, Ξ, and Y are given by Lemma 5.2(2) with ε replaced by $\varepsilon/2$. The smoothness of $\widetilde{\Psi}(t, x, \xi)$ follows from corresponding properties of $X(t)$, $\Xi(t)$, and $Y(t)$. It is well known that $\widetilde{\Psi}(t, x, \xi)$ solves the Hamilton–Jacobi equation

$$\partial_t \widetilde{\Psi}(t, x, \xi) + p(x, \partial_x \widetilde{\Psi}(t, x, \xi)) = 0, \quad \Psi|_{t=0} = x \cdot \xi,$$

for $(x, \xi) \in \Omega(\varepsilon/2)$, and satisfies

$$\partial_x \widetilde{\Psi}(t, x, \xi) = \Xi(t, Y(t, x, \xi), \xi), \quad \partial_\xi \widetilde{\Psi}(t, x, \xi) = Y(t, x, \xi).$$

Lemma 5.2(2) shows that $p(Y(t, x, \xi), \xi)$ is of quadratic type:

$$|\partial_x^\alpha \partial_\xi^\beta p(Y(t, x, \xi), \xi)| \le C_{\alpha\beta\varepsilon} \langle x \rangle^{2-|\alpha+\beta|}, \quad (t, x, \xi) \in [0, t_\varepsilon) \times \Omega(\varepsilon/2),$$

which, combined with the energy conservation

$$p(x, \partial_x \widetilde{\Psi}(t, x, \xi)) = p(Y(t, x, \xi), \xi),$$

implies

$$|\partial_x^\alpha \partial_\xi^\beta (\widetilde{\Psi}(t, x, \xi) - x \cdot \xi)| \le C_{\alpha\beta\varepsilon} |t| \langle x \rangle^{2-|\alpha+\beta|}, \quad (t, x, \xi) \in [0, t_\varepsilon) \times \Omega(\varepsilon/2).$$

We similarly obtain, for $(t, x, \xi) \in [0, t_\varepsilon) \times \Omega(\varepsilon/2)$,

$$|p(x, \partial_x \widetilde{\Psi}(t, x, \xi)) - p(x, \xi)| = \left| (\partial_x \widetilde{\Psi}(t, x, \xi) - \xi) \cdot \int_0^1 (\partial_\xi p)(x, \theta \partial_x \widetilde{\Psi}(t, x, \xi) + (1-\theta)\xi)\, d\theta \right|$$
$$\le C_\varepsilon |t| \langle x \rangle^2,$$

and, more generally,

$$|\partial_x^\alpha \partial_\xi^\beta \big(p(x, \partial_x \widetilde{\Psi}(t, x, \xi)) - p(x, \xi) \big)| \le C_{\alpha\beta\varepsilon} |t| \langle x \rangle^{2-|\alpha+\beta|}.$$

Therefore, integrating the Hamilton–Jacobi equation with respect to t, we have

$$|\partial_x^\alpha \partial_\xi^\beta (\widetilde{\Psi}(t, x, \xi) - x \cdot \xi + t p(x, \xi))| \le C_{\alpha\beta\varepsilon} |t|^2 \langle x \rangle^{2-|\alpha+\beta|}.$$

Finally, choosing a cutoff function $\rho \in S(1, g)$ so that $0 \le \rho \le 1$, $\rho \equiv 1$ on $\Omega(\varepsilon)$, and $\operatorname{supp} \rho \subset \Omega(\varepsilon/2)$, we define

$$\Psi(t, x, \xi) := x \cdot \xi - t p(x, \xi) + \rho(x, \xi)(\widetilde{\Psi}(t, x, \xi) - x \cdot \xi + t p(x, \xi)).$$

$\Psi(t, x, \xi)$ clearly satisfies the statement of Proposition 5.5. □

Using the phase function constructed in Proposition 5.5, we can define the FIO $J(\Psi, a) : \mathscr{S} \to \mathscr{S}'$ by

$$J(\Psi, a) f(x) = \frac{1}{(2\pi)^d} \int e^{i(\Psi(t,x,\xi) - y\cdot\xi)} a(x, \xi) f(y)\, dy\, d\xi, \qquad f \in \mathscr{S}(\mathbb{R}^d),$$

where $a \in S(1, g)$. Moreover, we have the following.

Lemma 5.6. *Let $t_\varepsilon > 0$ be small enough. Then, for any bounded family of symbols*

$$\{a(t) : |t| < t_\varepsilon\} \subset S(1, g),$$

$J(\Psi, a)$ is bounded on $L^2(\mathbb{R}^d)$ uniformly with respect to $|t| < t_\varepsilon$:

$$\sup_{|t| \le t_\varepsilon} \|J(\Psi, a)\|_{L^2 \to L^2} \le C_\varepsilon.$$

Proof. For sufficiently small $t_\varepsilon > 0$, the estimates (5-6) imply

$$|(\partial_\xi \otimes \partial_x \Psi)(t, x, \xi) - \mathrm{Id}\,| \le C_\varepsilon t_\varepsilon < \tfrac{1}{2}, \quad |\partial_x^\alpha \partial_\xi^\beta \Psi(t, x, \xi)| \le C_{\alpha\beta\varepsilon} \quad \text{for } |\alpha + \beta| \ge 2,$$

uniformly with respect to $(t, x, \xi) \in (-t_\varepsilon, t_\varepsilon) \times \mathbb{R}^{2d}$. Therefore, the assertion is a consequence of the standard L^2-boundedness of FIOs, or, equivalently, Kuranishi's trick and the L^2-boundedness of PDOs; see, for example, [Robert 1987; Mizutani 2013, Lemma 4.2]. □

We next construct the amplitude.

Proposition 5.7. *Let $\Psi(t, x, \xi)$ be as in Proposition 5.5 with ε replaced by $\varepsilon/3$. Then, for any integer $N \ge 0$, there exist families of symbols $\{b_j(t, \cdot, \cdot) : |t| < t_\varepsilon\} \subset S(\langle x \rangle^{-j} \langle \xi \rangle^{-j}, g)$, $j = 0, 1, 2, \ldots, N-1$, such that* $\operatorname{supp} b_j(t, \cdot, \cdot) \subset \Omega(\varepsilon/2)$ *and b_j solve the transport equations* (5-5).

Proof. We consider the case $t \ge 0$ only. Symbols b_j can be constructed by a standard method of characteristics along the flow generated by $\mathscr{X}(t,x,\xi)$ as follows. First note that Assumption 1.1 and (5-6) imply that

$$|\partial_x^\alpha \partial_\xi^\beta \mathscr{X}(t,x,\xi)| \le C_{\alpha\beta\varepsilon}\langle x\rangle^{1-|\alpha+\beta|}, \tag{5-7}$$

$$|\partial_x^\alpha \partial_\xi^\beta \mathscr{Y}(t,x,\xi)| \le C_{\alpha\beta\varepsilon}\langle x\rangle^{-|\alpha+\beta|}, \tag{5-8}$$

uniformly with respect to $0 \le t \le t_\varepsilon$ and $(x,\xi) \in \Omega(\varepsilon/3)$. For all $0 \le s, t \le t_\varepsilon$, we consider the solution to the ODE

$$\partial_t z(t,s,x,\xi) = \mathscr{X}(t, z(t,s,x,\xi), \xi), \quad z(s,s) = x.$$

We learn from (5-7) and an argument as in the proof of Lemma 5.1 that $z(t,s)$ is well defined for $0 \le s,t \le t_\varepsilon$ and $(x,\xi) \in \Omega(\varepsilon/3)$, and that

$$|\partial_x^\alpha \partial_\xi^\beta (z(t,s,x,\xi) - x)| \le C_{\alpha\beta\varepsilon} t_\varepsilon \langle x\rangle^{1-|\alpha+\beta|}, \quad (x,\xi) \in \Omega(\varepsilon/3). \tag{5-9}$$

Then $b_j(t)$ are defined inductively by

$$b_0(t,x,\xi) = \chi_\varepsilon^*(z(0,t,x,\xi),\xi) \exp\left(\int_0^t \mathscr{Y}(s, z(s,t,x,\xi),\xi)\,ds\right),$$

$$b_j(t,x,\xi) = -\int_0^t (iKb_{j-1})(s, z(s,t,x,\xi),\xi) \exp\left(\int_u^t \mathscr{Y}(u, z(u,t,x,\xi),\xi)\,du\right) ds.$$

Since $\operatorname{supp} \chi_\varepsilon^* \subset \Omega(\varepsilon)$, by (5-9) and an argument as in the proof of Lemma 5.2(1), we see that $b_j(t,x,\xi)$ is smooth with respect to (x,ξ) and that $\partial_x^\alpha \partial_\xi^\beta b_j(t,x,\xi)$ are supported in $\Omega(\varepsilon/2)$ for all $0 \le t \le t_\varepsilon$. Thus, if we extend b_j on $\mathbb{R}^{2d}$ so that $b_j(t,x,\xi) = 0$ outside $\Omega(\varepsilon/2)$, then b_j is still smooth in (x,ξ). Furthermore, we learn by (5-8) and (5-9) that $\{b_j(t,\cdot,\cdot) : t \in [0,t_\varepsilon],\ 0 \le j \le N-1\}$ is a bounded set in $S(\langle x\rangle^{-j}\langle\xi\rangle^{-j}, g)$. Finally, a standard Hamilton–Jacobi theory shows that $b_j(t)$ solve the transport equations (5-5). $\square$

We now state the main result in this section.

Theorem 5.8. *Fix $\varepsilon > 0$ arbitrarily. Then, for any sufficiently small $t_\varepsilon > 0$, any nonnegative integer $N \ge 0$ and any symbol $\chi_\varepsilon \in S(1,g)$ supported in $\Omega(\varepsilon)$, we can find a bounded family of symbols $\{a^N(t,\cdot,\cdot) : |t| < t_\varepsilon\} \subset S(1,g)$ such that $\operatorname{Op}(\chi_\varepsilon)e^{-itH}\operatorname{Op}(\chi_\varepsilon)^*$ can be brought to the form*

$$\operatorname{Op}(\chi_\varepsilon)e^{-itH}\operatorname{Op}(\chi_\varepsilon)^* = J(\Psi, a^N) + Q(t,N),$$

where $J(\Psi,a^N)$ is the FIO with the phase $\Psi(t,x,\xi)$ constructed in Proposition 5.5 with ε replaced by $\varepsilon/3$. The distribution kernel of $J(\Psi,a^N)$, which we denote by $K_{\Psi,a^N}(t,x,y)$, satisfies the dispersive estimate

$$|K_{\Psi,a^N}(t,x,y)| \le C_{N,\varepsilon}|t|^{-d/2}, \quad 0 < |t| < t_\varepsilon,\ x,\xi \in \mathbb{R}^d.$$

Moreover, for any $\gamma \ge 0$ with $N > 2\gamma$, the remainder $Q(t,N)$ satisfies

$$\|\langle D\rangle^\gamma Q(t,N)\langle D\rangle^\gamma\|_{L^2\to L^2} \le C_{N\gamma\varepsilon}|t|, \quad |t| < t_\varepsilon. \tag{5-10}$$

In particular, if we choose $N \geq d+1$, the distribution kernel of $Q(t,N)$ is uniformly bounded in $\mathbb{R}^{2d}$ with respect to $|t| < t_\varepsilon$. Hence

$$\|\mathrm{Op}(\chi_\varepsilon)e^{-itH}\,\mathrm{Op}(\chi_\varepsilon)^*\|_{L^1\to L^\infty} \leq C_\varepsilon |t|^{-d/2}, \quad 0<|t|<t_\varepsilon.$$

Proof. We consider the case when $t \geq 0$ and the proof for the opposite case is similar. By virtue of Lemma 5.4, we may replace $\mathrm{Op}(\chi_\varepsilon)^*$ by $\mathrm{Op}(\chi_\varepsilon^*)$ for some $\chi_\varepsilon^* \in S(1,g)$ supported in $\Omega(\varepsilon)$, without loss of generality. Let $b^N = \sum_{j=0}^{N-1} b_j$ with b_j constructed in Proposition 5.7. Since $J(\Psi, b^N)|_{t=0} = \mathrm{Op}(\chi_\varepsilon^*)$, we have the Duhamel formula

$$\mathrm{Op}(\chi_\varepsilon)e^{-itH}\,\mathrm{Op}(\chi_\varepsilon^*) = \mathrm{Op}(\chi_\varepsilon)J(\Psi,b^N) - i\int_0^t \mathrm{Op}(\chi_\varepsilon)e^{-i(t-s)H}(D_t+H)J(\Psi,b^N)|_{t=s}\,ds.$$

Estimates on the remainder. It suffices to show that

$$\sup_{|t|<t_\varepsilon} \|\langle D\rangle^\gamma\,\mathrm{Op}(\chi_\varepsilon)e^{-itH}(D_t+H)J(\Psi,b^N)\langle D\rangle^\gamma\|_{L^2\to L^2} \leq C_{N\gamma\varepsilon}.$$

Since Ψ, b_j solve the Hamilton–Jacobi equation (5-4) and transport equations (5-5), respectively, a direct computation yields

$$e^{-i\Psi(t,x,\xi)}(D_t+H)\left(e^{i\Psi(t,x,\xi)}\sum_{j=0}^{N-1} b_j(t,x,\xi)\right) = r_N(t,x,\xi),$$

with some $\{r_N(t,\cdot,\cdot): 0\leq t\leq t_\varepsilon\} \subset S(\langle x\rangle^{-N}\langle\xi\rangle^{-N}, g)$. In particular,

$$(D_t+H)J(\Psi,b^N) = J(\Psi,r_N).$$

A standard L^2-boundedness of FIOs then implies

$$\sup_{|t|<t_\varepsilon} \|\langle x\rangle^{\gamma+\delta}\langle D\rangle^{\gamma+\delta}J(\Psi,r_N)\langle D\rangle^\gamma\|_{L^2\to L^2} \leq C_{N\gamma\delta},$$

for any $\gamma,\delta \geq 0$ with $2\gamma+\delta \leq N$. Since, in the proof of Lemma 5.4, we already proved that

$$\sup_{|t|\leq t_\varepsilon} \|\langle D\rangle^\gamma\,\mathrm{Op}(\chi_\varepsilon)e^{-itH}\langle D\rangle^{-\gamma-\delta}\langle x\rangle^{-\gamma-\delta}\|_{L^2\to L^2} \leq C_{\gamma\delta},$$

we obtain the desired estimate.

Dispersive estimates. By the composition formula of PDOs and FIOs (cf. [Robert 1987]),

$$\mathrm{Op}(\chi_\varepsilon)J(\Psi,b^N)$$

is also an FIO with the same phase Ψ and the amplitude

$$a^N(t,x,\xi) = \frac{1}{(2\pi)^d}\int e^{iy\cdot\eta}\chi_\varepsilon(x,\eta+\widetilde{\Xi}(t,x,y,\xi))b^N(t,x+y,\xi)\,dy\,d\eta,$$

where $\widetilde{\Xi}(t,x,y,\xi) = \int_0^1 (\partial_x\Psi)(t,y+\lambda(x-y),\xi)\,d\lambda$. By virtue of (5-6), $\widetilde{\Xi}$ satisfies

$$|\partial_x^\alpha\partial_y^{\alpha'}\partial_\xi^\beta(\widetilde{\Xi}(t,x,y,\xi)-\xi)| \leq C_{\alpha\alpha'\beta}|t|, \quad |\alpha+\alpha'+\beta|\geq 1.$$

Combining this with the relations $\chi_\varepsilon, b^N \in S(1,g)$, $\operatorname{supp}\chi_\varepsilon \subset \Omega(\varepsilon)$, and $\operatorname{supp} b^N(t,\cdot,\cdot) \subset \Omega(\varepsilon/2)$, we see that $\{a^N : 0 \le t < t_\varepsilon\}$ is bounded in $S(1,g)$. The distribution kernel of $J(\Psi, a^N)$ is given by

$$K_{\Psi,a^N}(t,x,y) = \frac{1}{(2\pi)^d}\int e^{i(\Psi(t,x,\xi)-y\cdot\xi)} a^N(t,x,\xi)\,d\xi.$$

By virtue of Proposition 5.5, we have

$$\sup_{|t|\le t_\varepsilon} |\partial_x^\alpha \partial_y^\beta \partial_\xi^\gamma(\Psi(t,x,\xi) - y\cdot\xi)| \le C_{\alpha\beta\gamma}, \quad |\alpha+\beta+\gamma| \ge 2,$$
$$\partial_\xi^2 \Psi(t,x,\xi) = -t(g^{jk}(x))_{j,k} + O(t^2), \qquad |t| \to 0.$$

As a consequence, since $g^{jk}(x)$ is uniformly elliptic, the phase function $\Psi(t,x,\xi) - y\cdot\xi$ has a unique nondegenerate critical point for all $|t| < t_\varepsilon$ and we can apply the stationary phase method to $K_{\Psi,a^N}(t,x,y)$, provided that $t_\varepsilon > 0$ is small enough. Therefore,

$$|K_{\Psi,a^N}(t,x,y)| \le C|t|^{-d/2}, \quad 0 < |t| \le t_\varepsilon,\ x,\xi \in \mathbb{R}^d. \qquad \square$$

6. Proof of Theorem 1.5

We now give the proof of Theorem 1.5. Suppose that H satisfies Assumption 1.1 with $\mu \ge 0$. In view of Corollary 2.6, (1-4) is a consequence of the following proposition.

Proposition 6.1. *For any symbol $a \in C_0^\infty(\mathbb{R}^{2d})$ and $T > 0$,*

$$\|\mathrm{Op}_h(a)e^{-itH}\varphi\|_{L^p([-T,T];L^q(\mathbb{R}^d))} \le C_T h^{-1/p}\|\varphi\|_{L^2(\mathbb{R}^d)},$$

uniformly with respect to $h \in (0,1]$, provided that (p,q) satisfies (1-1).

Proof. This proposition follows from the standard WKB approximation for $e^{-itH}\,\mathrm{Op}_h(a)$ up to time scales of order $1/h$. The proof is essentially the same as that in the case for the Laplace–Beltrami operator on compact manifolds without boundaries [Burq et al. 2004, Section 2]. We omit the details. $\square$

Using this proposition, we have the semiclassical Strichartz estimates with inhomogeneous error terms.

Proposition 6.2. *Let $a \in C_0^\infty(\mathbb{R}^{2d})$. Then, for any $T > 0$ and any (p,q) satisfying the admissible condition* (1-1),

$$\begin{aligned}\|\mathrm{Op}_h(a)e^{-itH}\varphi\|_{L^p([-T,T];L^q(\mathbb{R}^d))} &\\ \le C_T\|\mathrm{Op}_h(a)\varphi\|_{L^2(\mathbb{R}^d)} &+ C_T h\|\varphi\|_{L^2(\mathbb{R}^d)} + Ch^{-1/2}\|\mathrm{Op}_h(a)e^{-itH}\varphi\|_{L^2([-T,T];L^2(\mathbb{R}^d))}\\ &+ Ch^{1/2}\|[\mathrm{Op}_h(a),H]e^{-itH}\varphi\|_{L^2([-T,T];L^2(\mathbb{R}^d))},\end{aligned}$$

uniformly with respect to $h \in (0,1]$.

This proposition has been proved by [Bouclet and Tzvetkov 2007] for the case with $V, A \equiv 0$. We give a refinement of this proposition with its proof in Section 7.

Next, we shall prove that if $k(x,\xi)$ satisfies the nontrapping condition (1-3), the missing $1/p$ derivative can be recovered. We first recall the local smoothing effects for Schrödinger operators proved by Doi [2005]. For any $s\in\mathbb{R}$, we set $\mathcal{B}^s := \{f \in L^2(\mathbb{R}^d) : \langle x\rangle^s f,\, \langle D\rangle^s f \in L^2(\mathbb{R}^d)\}$. Define a symbol $e_s(x,\xi)$ by

$$e_s(x,\xi) := (k(x,\xi)+|x|^2+L(s))^{s/2} \in S((1+|x|+|\xi|)^s, g),$$

where $L(s)>1$ is a large constant depending on s. We denote by E_s its Weyl quantization,

$$E_s f(x) = \mathrm{Op}^w(e_s) f(x) = \frac{1}{(2\pi)^d}\int e^{i(x-y)\cdot\xi} e_s\Big(\frac{x+y}{2},\xi\Big) f(y)\,dy\,d\xi.$$

Then, for any $s\in\mathbb{R}$, there exists $L(s)>0$ such that E_s is a homeomorphism from $\mathcal{B}^{r+s}$ to $\mathcal{B}^r$ for all $r\in\mathbb{R}$, and $(E_s)^{-1}$ is still a Weyl quantization of a symbol in $S((1+|x|+|\xi|)^{-s}, g)$; see, [Doi 2005, Lemma 4.1].

Proposition 6.3 (the local smoothing effects [Doi 2005]). *Suppose that $k(x,\xi)$ satisfies the nontrapping condition* (1-3) *and Assumption 1.4. Then, for any $T>0$ and $\sigma>0$, there exists $C_{T,\sigma}>0$ such that*

$$\|\langle x\rangle^{-1/2-\sigma} E_{1/2} e^{-itH}\varphi\|_{L^2([-T,T];L^2(\mathbb{R}^d))} \le C_{T,\sigma}\|\varphi\|_{L^2(\mathbb{R}^d)}. \tag{6-1}$$

Remark 6.4. (6-1) implies a standard local smoothing effect,

$$\|\langle x\rangle^{-1/2-\sigma}\langle D\rangle^{1/2} e^{-itH}\varphi\|_{L^2([-T,T];L^2(\mathbb{R}^d))} \le C_{T,\sigma}\|\varphi\|_{L^2(\mathbb{R}^d)}. \tag{6-2}$$

Indeed, we compute

$$\begin{aligned}
\langle x\rangle^{-1/2-\sigma}\langle D\rangle^{1/2} &= \langle D\rangle^{1/2}\langle x\rangle^{-1/2-\sigma} + [\langle D\rangle^{1/2}, \langle x\rangle^{-1/2-\sigma}]\\
&= \langle D\rangle^{1/2}(E_{1/2})^{-1}E_{1/2}\langle x\rangle^{-1/2-\sigma} + [\langle D\rangle^{1/2}, \langle x\rangle^{-1/2-\sigma}]\\
&= \langle D\rangle^{1/2}(E_{1/2})^{-1}(\langle x\rangle^{-1/2-\sigma}E_{1/2} + [E_{1/2}, \langle x\rangle^{-1/2-\sigma}]) + [\langle D\rangle^{1/2}, \langle x\rangle^{-1/2-\sigma}].
\end{aligned}$$

It is easy to see that $\langle D\rangle^{1/2}(E_{1/2})^{-1}$, $[E_{1/2}, \langle x\rangle^{-1/2-\sigma}]$, and $[\langle D\rangle^{1/2}, \langle x\rangle^{-1/2-\sigma}]$ are bounded on $L^2(\mathbb{R}^d)$ since their symbols belong to $S(1,g)$. Therefore, (6-1) implies (6-2).

Proof of (1-5) *of Theorem 1.5.* It is clear that (1-5) follows from Proposition 6.2, (6-2), and Corollary 2.6, since a is compactly supported with respect to x and $\{a,p\}\in S(\langle\xi\rangle, g)$, where $p=p(x,\xi)$. □

7. Near sharp Strichartz estimates without asymptotic flatness

This section is devoted to proving Theorem 1.6. We may assume $\mu=0$ without loss of generality.

Proposition 7.1. *Let $I\Subset(0,\infty)$ be a relatively compact open interval and $C_0>1$. Then there exist $\delta_0, h_0>0$ such that for any $0<\delta\le\delta_0$, $0<h\le h_0$, $1\le R\le 1/h$, and any symbol $a_h\in S(1,g)$ supported in $\{(x,\xi): R<|x|<C_0/h,\ |\xi|\in I\}$, we have*

$$\|\mathrm{Op}_h(a_h)e^{-itH}\,\mathrm{Op}_h(a_h)^*\|_{L^1\to L^\infty} \le C_\delta|t|^{-d/2}, \quad 0<|t|<\delta hR, \tag{7-1}$$

where $C_\delta>0$ may be taken uniformly with respect to h and R.

Remark 7.2. When $|t| > 0$ in (7-1) is small and independent of R, (7-1) is well known and the proof is given by the standard method of the short-time WKB approximation for $e^{-itH^h/h}\operatorname{Op}_h(a_h)^*$; see, for example, [Burq et al. 2004].

For $h \in (0, 1]$, $R \geq 1$, an open interval $I \Subset (0, \infty)$, and $C_0 > 1$, we set

$$\Gamma(R, h, I) := \{(x, \xi) \in \mathbb{R}^{2d} : R < |x| < C_0/h,\ |\xi| \in I\}.$$

Equation (7-1) is a consequence of the same argument as in the proof of Proposition 3.1 and the following proposition.

Proposition 7.3. *Let $I \Subset I_1 \Subset (0, \infty)$ and $C_0 > 1$. Then there exist $\delta_0, h_0 > 0$ such that the following hold for any $0 < \delta \leq \delta_0$, $0 < h \leq h_0$, and $1 \leq R \leq C_0/h$.*

(1) *There exists $\Phi_h(t, x, \xi) \in C^\infty((-\delta R, \delta R) \times \mathbb{R}^{2d})$ such that Φ_h solves the Hamilton–Jacobi equation*

$$\begin{cases} \partial_t \Phi_h(t, x, \xi) = -p_h(x, \partial_x \Phi_h(t, x, \xi)), & |t| < \delta R,\ (x, \xi) \in \Gamma(R/2, h/2, I_1), \\ \Phi_h(0, x, \xi) = x \cdot \xi, & (x, \xi) \in \Gamma(R/2, h/2, I_1). \end{cases} \tag{7-2}$$

Furthermore, we have

$$|\partial_x^\alpha \partial_\xi^\beta (\Phi_h(t, x, \xi) - x \cdot \xi + t p_h(x, \xi))| \leq C_{\alpha\beta} R^{-|\alpha|} h |t|^2, \quad \alpha, \beta \in \mathbb{Z}_+^d, \tag{7-3}$$

uniformly with respect to $x, \xi \in \mathbb{R}^d$, $h \in (0, h_0]$, $0 \leq R \leq C_0/h$, and $|t| < \delta R$.

(2) *For any $a_h \in S(1, g)$ with $\operatorname{supp} a_h \subset \Gamma(R, h, I)$ and any integer $N \geq 0$, we can find $b_h^N(t, \cdot, \cdot) \in S(1, g)$ such that*

$$e^{-it\widetilde{H}^h/h}\operatorname{Op}_h(a_h)^* = J_h(\Phi_h, b_h^N) + Q_{\mathrm{WKB}}(t, h, N),$$

where $J_h(\Phi_h, b_h^N)$ is the h-FIO with phase function Φ_h and amplitude b_h^N, and its distribution kernel satisfies

$$|K_{\mathrm{WKB}}(t, h, x, y)| \leq C|th|^{-d/2}, \quad h \in (0, h_0],\ 0 < |t| \leq \delta R,\ x, \xi \in \mathbb{R}^d. \tag{7-4}$$

Moreover the remainder $Q_{\mathrm{WKB}}(t, h, N)$ satisfies

$$\|\langle D\rangle^s Q_{\mathrm{WKB}}(t, h, N)\langle D\rangle^s\|_{L^2 \to L^2} \leq C_{N,s} h^{N-2s}|t|, \quad h \in (0, h_0],\ |t| \leq \delta R.$$

Sketch of proof. The proof is similar to that of Theorem 5.8; in particular, the proof of the second claim is completely the same. Thus, we just outline the construction of Φ_h. We may assume $C_0 = 1$ without loss of generality. Denote by (X_h, Ξ_h) the Hamilton flow generated by p_h. To construct the phase function, the most important step is to study the inverse map of $(x, \xi) \mapsto (X_h(t, x, \xi), \xi)$. Choose an open interval $\tilde{I}_1$ so that $I_1 \Subset \tilde{I}_1 \Subset (0, \infty)$. The following bound was proved in [Mizutani 2013]:

$$|\partial_x^\alpha \partial_\xi^\beta (X_h(t, x, \xi) - x)| + \langle x\rangle |\partial_x^\alpha \partial_\xi^\beta (\Xi_h(t, x, \xi) - \xi)| \leq C_{\alpha\beta} \langle x\rangle^{-|\alpha|}|t|$$

for $(x, \xi) \in \Gamma(R/3, h/3, \tilde{I}_1)$ and $|t| \leq \delta R$. For sufficiently small $\delta > 0$ and for any fixed $|t| \leq \delta R$, this implies

$$|\partial_x X_h(t) - \mathrm{Id}\,| \leq C R^{-1}|t| \leq C\delta < \tfrac{1}{2}.$$

By the same argument as that in the proof of Lemma 5.2, the map $(x, \xi) \mapsto (X_h(t, x, \xi), \xi)$ is a diffeomorphism from $\Gamma(R/3, h/3, \tilde{I}_1)$ onto its range and the corresponding inverse $(x, \xi) \mapsto (Y_h(t, x, \xi), \xi)$ is well-defined for $|t| < \delta R$ and $(x, \xi) \in \Gamma(R/2, h/2, I_1)$. Moreover, $Y_h(t)$ satisfies an estimate like the one for $X_h(t)$:

$$|\partial_x^\alpha \partial_\xi^\beta (Y_h(t, x, \xi) - x)| \le C_{\alpha\beta} \langle x \rangle^{-|\alpha|} |t|, \quad |t| < \delta R,\ (x, \xi) \in \Gamma(R/2, h/2, I_1).$$

We now define Φ_h by

$$\Phi_h(t, x, \xi) := x \cdot \xi + \int_0^t L_h\big(X_h(s, Y(t, x, \xi), \xi),\, \Xi(s, y(t, x, \xi), \xi)\big)\, ds,$$

where $L_h = \xi \cdot \partial_\xi p_h - p_h$. By the standard Hamilton–Jacobi theory, Φ_h solves (7-2). Moreover, using the energy conservation $p_h(x, \partial_x \Phi_h(t)) = p_h(Y_h(t), \xi)$ and the above estimates on X_h, Ξ_h, and Y_h, we see that

$$\begin{aligned}
|p_h(x, \partial_x \Phi_h(t)) - p_h(x, \xi)| &= |p_h(Y_h(t), \xi) - p_h(x, \xi)| \\
&\le |Y_h(t) - x| \left| \int_0^\lambda (\partial_x p_h)(\lambda Y_h(t) - (1-\lambda)x, \xi)\, d\lambda \right| \\
&\le C|y(t) - x|(h + h^2\langle x \rangle^2) \\
&\le Ch|t|
\end{aligned}$$

and that

$$|\partial_x^\alpha \partial_\xi^\beta (p_h(x, \partial_x \Phi_h) - p_h(x, \xi))| \le C_{\alpha\beta} \langle x \rangle^{-|\alpha|} h|t|.$$

Using these estimates, we can check that Φ_h satisfies (7-3). Finally, we extend Φ_h to the whole space so that $\Phi_h(t, x, \xi) = x \cdot \xi - t p_h(x, \xi)$ outside $\Gamma(R/3, h/3, \tilde{I}_1)$. □

Using Proposition 7.1, we obtain a refinement of Proposition 6.2.

Proposition 7.4. *Let* $0 < R \le 1/h$ *and let* $a_h \in S(1, g)$ *be supported in* $\{(x, \xi) : R < |x| < 1/h,\ |\xi| \in I\}$. *Then, for any* $T > 0$ *and* (p, q) *satisfying the admissible condition* (1-1),

$$\begin{aligned}
&\|\mathrm{Op}_h(a_h) e^{-itH} \varphi\|_{L^p([-T,T]; L^q(\mathbb{R}^d))} \\
&\quad \le C_T \|\mathrm{Op}_h(a_h)\varphi\|_{L^2(\mathbb{R}^d)} + C_T h \|\varphi\|_{L^2(\mathbb{R}^d)} + C_T (hR)^{-1/2} \|\mathrm{Op}_h(a_h) e^{-itH}\varphi\|_{L^2([-T,T]; L^2(\mathbb{R}^d))} \\
&\qquad + C_T (hR)^{1/2} \|[H, \mathrm{Op}_h(a_h)] e^{-itH} \varphi\|_{L^2([-T,T]; L^2(\mathbb{R}^d))},
\end{aligned}$$

uniformly with respect to $h \in (0, h_0]$.

Proof. The proof is similar to that of [Bouclet and Tzvetkov 2007, Proposition 5.4]. By time reversal invariance we can restrict our considerations to the interval $[0, T]$. We may assume $T \ge hR$ without loss of generality and split $[0, T]$ as follows: $[0, T] = J_0 \cup J_1 \cup \cdots \cup J_N$, where $J_j = [jhR, (j+1)hR]$, $0 \le j \le N-1$, and $J_N = [T - \delta hR, T]$. For $j = 0$, we have the Duhamel formula

$$\mathrm{Op}_h(a_h) e^{-itH} = e^{-itH}\, \mathrm{Op}_h(a_h) - i \int_0^t e^{-i(t-s)H} [\mathrm{Op}_h(a_h), H] e^{-isH}\, ds, \quad t \in J_0.$$

Here we choose $b_h \in S(1, g)$ so that $b_h \equiv 1$ on $\operatorname{supp} a$ and b_h is supported in a sufficiently small neighborhood of $\operatorname{supp} a_h$. By Proposition 7.1, $\mathrm{Op}_h(b_h)e^{-i(t-s)H}\mathrm{Op}_h(b_h)^*$ satisfies dispersive estimates (7-1) for $0<|t-s|<\delta hR$ with some $\delta>0$ small enough. Using the Keel–Tao theorem [1998] and the unitarity of e^{-itH}, we then learn that for any interval J_R of size $|J_R|\le 2hR$, the following homogeneous and inhomogeneous Strichartz estimates hold uniformly with respect to $h\in(0,h_0]$:

$$\|\mathrm{Op}_h(b_h)e^{-itH}\varphi\|_{L^p(J_R;L^q(\mathbb{R}^d))}\le C\|\varphi\|_{L^2(\mathbb{R}^d)}, \tag{7-5}$$

$$\left\|\int_0^t F(s\in J_R)\,\mathrm{Op}_h(b_h)e^{-i(t-s)H}\mathrm{Op}_h(b_h)^*g(s)\,ds\right\|_{L^p(J_R;L^q(\mathbb{R}^d))}\le C\|g\|_{L^1(J_R;L^2(\mathbb{R}^d))}, \tag{7-6}$$

where $F(s\in J_R)$ is the characteristic function of J_R and (p,q) satisfies the admissible condition (1-1). On the other hand, using the expansions (2-3) and (2-4), we see that for any $M\ge 0$,

$$\begin{aligned}\mathrm{Op}_h(a_h)&=\mathrm{Op}_h(b_h)\,\mathrm{Op}_h(a_h)+h^M\,\mathrm{Op}_h(r_{1,h})\\&=\mathrm{Op}_h(b_h)^*\,\mathrm{Op}_h(a_h)+h^M\,\mathrm{Op}_h(r_{2,h}),\\ [\mathrm{Op}_h(a_h),H]&=\mathrm{Op}_h(b_h)^*[\mathrm{Op}_h(a_h),H]+h^M\,\mathrm{Op}_h(r_{3,h}),\end{aligned}$$

with some $\{r_{l,h}\}_{h\in(0,1]}$, $l=1,2,3$, which are bounded in $S(\langle x\rangle^{-M}\langle\xi\rangle^{-M},g)$. Therefore, we can write

$$\begin{aligned}&\mathrm{Op}_h(a_h)e^{-itH}\\&\quad=\mathrm{Op}_h(b_h)e^{-itH}\,\mathrm{Op}_h(a_h)-i\int_0^t\mathrm{Op}_h(b_h)e^{-i(t-s)H}\,\mathrm{Op}_h(b_h)^*[\mathrm{Op}_h(a_h),H]e^{-isH}\,ds+Q(t,h,M),\end{aligned}$$

where the remainder $Q(t,h,M)$ satisfies

$$\|Q(t,h,M)\|_{L^2\to L^q}\le C_Mh^{M-1-d(1/2-1/q)},\quad 2\le q\le\infty,$$

uniformly in $h\in(0,1]$. Combining this estimate with (7-5) and (7-6), we obtain

$$\begin{aligned}\|\mathrm{Op}_h(a_h)e^{-itH}\varphi\|_{L^p(J_0;L^q)}&\le C\|\mathrm{Op}_h(a_h)\varphi\|_{L^2}+Ch\|\varphi\|_{L^2}+C\|[\mathrm{Op}_h(a_h),H]e^{-itH}\varphi\|_{L^1(J_0;L^2)}\\&\le C\|\mathrm{Op}_h(a_h)\varphi\|_{L^2}+Ch\|\varphi\|_{L^2}+C(hR)^{1/2}\|[\mathrm{Op}_h(a_h),H]e^{-itH}\varphi\|_{L^2(J_0;L^2)}.\end{aligned}$$

We similarly obtain the same bound for $j=N$:

$$\|\mathrm{Op}_h(a_h)e^{-itH}\varphi\|_{L^p(J_N;L^q)}\le C\|\mathrm{Op}_h(a_h)\varphi\|_{L^2}+Ch\|\varphi\|_{L^2}+C(hR)^{1/2}\|[\mathrm{Op}_h(a_h),H]e^{-itH}\varphi\|_{L^2(J_N;L^2)}.$$

For $j=1,2,\dots,N-1$, taking $\theta\in C_0^\infty(\mathbb{R})$ so that $\theta\equiv 1$ on $[-1/2,1/2]$ and $\operatorname{supp}\theta\subset[-1,1]$, we set $\theta_j(t)=\theta(t/(hR)-j-1/2))$. It is easy to see that $\theta_j\equiv 1$ on J_j and $\operatorname{supp}\theta_j\subset\widetilde{J}_j=J_j+[-hR/2,hR/2]$. We consider $v_j=\theta_j(t)\,\mathrm{Op}_h(a_h)e^{-itH}\varphi$, which solves

$$i\partial_t v_j=Hv_j+\theta_j'\,\mathrm{Op}_h(a_h)e^{-itH}\varphi+\theta_j[\mathrm{Op}_h(a_h),H]e^{-itH}\varphi,\quad v_j|_{t=0}=0.$$

An argument as above and the Duhamel formula then imply that, for any $t\in\widetilde{J}_j$ and $M\ge 0$, v_j satisfies

$$v_j=-i\int_0^t\mathrm{Op}_h(b_h)e^{-i(t-s)H}\,\mathrm{Op}_h(b_h)^*\big(\theta_j'(s)\,\mathrm{Op}_h(a_h)+\theta_j(s)[\mathrm{Op}_h(a_h),H]\big)e^{-isH}\varphi\,ds+\widetilde{Q}(t,h,M),$$

where the remainder $\widetilde{Q}(t,h,M)$ satisfies

$$\|\widetilde{Q}(t,h,M)\|_{L^2\to L^q}\le C_Mh^{M-1-d(1/2-1/q)},\quad 2\le q\le\infty,$$

uniformly in $h \in (0,1]$ and $t \in \widetilde{J}_j$. Taking $M \geq 0$ large enough, we learn

$$\begin{aligned}
&\|v_j\|_{L^p(J_j;L^q)} \\
&\quad\leq Ch^2\|\varphi\|_{L^2} + C(hR)^{-1}\|\mathrm{Op}_h(a_h)e^{-itH}\varphi\|_{L^1(\widetilde{J}_j;L^2)} + C\|[\mathrm{Op}_h(a_h), H]e^{-itH}\varphi\|_{L^1(\widetilde{J}_j;L^2)} \\
&\quad\leq Ch^2\|\varphi\|_{L^2} + C(hR)^{-1/2}\|\mathrm{Op}_h(a_h)e^{-itH}\varphi\|_{L^2(\widetilde{J}_j;L^2)} + C(hR)^{1/2}\|[\mathrm{Op}_h(a_h), H]e^{-itH}\varphi\|_{L^2(\widetilde{J}_j;L^2)}.
\end{aligned}$$

Since $N \leq T/h$ and $p \geq 2$, summing over $j = 0, 1, \dots, N$, we have the assertion by Minkowski's inequality. □

Proof of Theorem 1.6. In view of Corollary 2.6, Theorem 1.5, and Proposition 3.2, it suffices to show that, for any $a_h \in S(1, g)$ with

$$\operatorname{supp} a_h \in \{(x,\xi) : 2 \leq |x| \leq 1/h,\ |\xi| \in I\}$$

and any $\varepsilon > 0$,

$$\sum_h \|\mathrm{Op}_h(a_h)e^{-itH} f(h^2H)\varphi\|^2_{L^p([-T,T];L^q)} \leq C_{T,\varepsilon}\|\langle H\rangle^\varepsilon\varphi\|^2_{L^2}.$$

Let us consider a dyadic partition of unity:

$$\sum_{1\leq j\leq j_h} \chi(2^{-j}x) = 1, \quad 2 \leq |x| \leq 1/h,$$

where $\chi \in C_0^\infty(\mathbb{R}^d)$ with

$$\operatorname{supp}\chi \subset \{1/2 < |x| < 2\}$$

and $j_h \leq [\log(1/h)] + 1$. We set

$$\chi_j(x) = \chi(2^{-j}x).$$

Proposition 7.4 then implies

$$\begin{aligned}
&\|\chi_j \mathrm{Op}_h(a_h)e^{-itH}\varphi\|_{L^p([-T,T];L^q)} \\
&\qquad\leq C_T\|\chi_j \mathrm{Op}_h(a_h)\varphi\|_{L^2} + C_T h\|\varphi\|_{L^2} + C_T(h2^j)^{-1/2}\|\chi_j \mathrm{Op}_h(a_h)e^{-itH}\varphi\|_{L^2([-T,T];L^2)} \\
&\qquad\qquad + C_T(h2^j)^{1/2}\|[\chi_j \mathrm{Op}_h(a_h), H]e^{-itH}\varphi\|_{L^2([-T,T];L^2)}.
\end{aligned}$$

Since $2^{j-1} \leq |x| \leq 2^{j+1}$ and $|x| \leq 1/h$ on $\operatorname{supp}\chi_j a_h$, we have, for any $\varepsilon \geq 0$,

$$(h2^j)^{-1/2}\|\chi_j \mathrm{Op}_h(a_h)e^{-itH}\varphi\|_{L^2([-T,T];L^2)} \leq C\|\chi_j\langle x\rangle^{-1/2-\varepsilon}h^{-1/2-\varepsilon}\mathrm{Op}_h(a_h)e^{-itH}\varphi\|_{L^2([-T,T];L^2)}.$$

Since $\{\chi_j a_h, p\} \in S(\langle x\rangle^{-1}\langle\xi\rangle, g)$, we similarly obtain

$$\begin{aligned}
&(h2^j)^{1/2}\|\chi_j[\mathrm{Op}_h(a_h), H]e^{-itH}\varphi\|_{L^2([-T,T];L^2)} \\
&\qquad\qquad \leq \|\widetilde{\chi}_j\langle x\rangle^{-1/2-\varepsilon}h^{-1/2-\varepsilon}\mathrm{Op}_h(b_h)e^{-itH}\varphi\|_{L^2([-T,T];L^2)} + C_T h\|\varphi\|_{L^2},
\end{aligned}$$

where $\widetilde{\chi}_j(x) = \widetilde{\chi}(2^{-j}x)$ for some $\widetilde{\chi} \in C_0^\infty(\mathbb{R}^d)$ satisfying $\widetilde{\chi} \equiv 1$ on $[1/2, 2]$ and $\operatorname{supp}\widetilde{\chi} \subset [1/4, 4]$, and $b_h \in S(1, g)$ is supported in a neighborhood of $\operatorname{supp} a_h$ so that $b_h \equiv 1$ on $\operatorname{supp} a_h$. Summing over

$1 \le j \le j_h$ and using the local smoothing effect (6-2), since $p, q \ge 2$, we obtain

$$
\begin{aligned}
\|\mathrm{Op}_h(a_h)e^{-itH}\varphi\|^2_{L^p([-T,T];L^q)}
&\le \sum_{1\le j\le j_h} \|\chi_j \mathrm{Op}_h(a_h)e^{-itH}\varphi\|^2_{L^p([-T,T];L^q)}\\
&\le C_T \sum_{1\le j\le j_h} (\|\chi_j \mathrm{Op}_h(a_h)\varphi\|^2_{L^2} + h\|\varphi\|^2_{L^2})\\
&\qquad + C \sum_{1\le j\le j_h} \|\widetilde{\chi}_j \langle x\rangle^{-1/2-\varepsilon} h^{-1/2-\varepsilon} \mathrm{Op}_h(a_h+b_h)e^{-itH}\varphi\|^2_{L^2([-T,T];L^2)}\\
&\le C_T\|\varphi\|^2_{L^2} + C\|\langle x\rangle^{-1/2-\varepsilon} h^{-1/2-\varepsilon} \mathrm{Op}_h(a_h+b_h)e^{-itH}\varphi\|^2_{L^2([-T,T];L^2)}\\
&\le C_{T,\varepsilon} h^{-2\varepsilon}\|\varphi\|^2_{L^2},
\end{aligned}
$$

which implies

$$
\sum_h \|\mathrm{Op}_h(a_h)e^{-itH} f(h^2H)\varphi\|^2_{L^p([-T,T];L^q)} \le C_{T,\varepsilon} \sum_h h^{-2\varepsilon}\|f(h^2H)\varphi\|^2_{L^2} \le C_{T,\varepsilon}\|\langle H\rangle^{\varepsilon/2}\varphi\|^2_{L^2}.
$$

This completes the proof. □

Acknowledgements

The author thanks Professor Erik Skibsted for valuable discussions and for hospitality at Institut for Matematiske Fag, Aarhus Universitet, where a part of this work was carried out. He also thanks the referee for very carefully reading the manuscript and for providing valuable suggestions, which substantially helped to improve the quality of the paper.

References

[Bouclet 2010] J.-M. Bouclet, "Littlewood–Paley decompositions on manifolds with ends", *Bull. Soc. Math. France* **138**:1 (2010), 1–37. MR 2011e:42020 Zbl 1198.42013

[Bouclet 2011a] J.-M. Bouclet, "Semi-classical functional calculus on manifolds with ends and weighted L^p estimates", *Ann. Inst. Fourier* (*Grenoble*) **61**:3 (2011), 1181–1223. MR 2918727 Zbl 1236.58033

[Bouclet 2011b] J.-M. Bouclet, "Strichartz estimates on asymptotically hyperbolic manifolds", *Anal. PDE* **4**:1 (2011), 1–84. MR 2012i:58022 Zbl 1230.35027

[Bouclet and Tzvetkov 2007] J.-M. Bouclet and N. Tzvetkov, "Strichartz estimates for long range perturbations", *Amer. J. Math.* **129**:6 (2007), 1565–1609. MR 2009d:58035 Zbl 1154.35077

[Bouclet and Tzvetkov 2008] J.-M. Bouclet and N. Tzvetkov, "On global Strichartz estimates for non-trapping metrics", *J. Funct. Anal.* **254**:6 (2008), 1661–1682. MR 2009d:35039 Zbl 1168.35005

[Burq et al. 2004] N. Burq, P. Gérard, and N. Tzvetkov, "Strichartz inequalities and the nonlinear Schrödinger equation on compact manifolds", *Amer. J. Math.* **126**:3 (2004), 569–605. MR 2005h:58036 Zbl 1067.58027

[Burq et al. 2010] N. Burq, C. Guillarmou, and A. Hassell, "Strichartz estimates without loss on manifolds with hyperbolic trapped geodesics", *Geom. Funct. Anal.* **20**:3 (2010), 627–656. MR 2012f:58068 Zbl 1206.58009

[Cazenave 2003] T. Cazenave, *Semilinear Schrödinger equations*, Courant Lecture Notes in Mathematics **10**, University Courant Institute of Mathematical Sciences, New York, 2003. MR 2004j:35266 Zbl 1055.35003

[D'Ancona and Fanelli 2009] P. D'Ancona and L. Fanelli, "Smoothing estimates for the Schrödinger equation with unbounded potentials", *J. Differential Equations* **246**:12 (2009), 4552–4567. MR 2011a:35084 Zbl 1173.35031

[D'Ancona et al. 2010] P. D'Ancona, L. Fanelli, L. Vega, and N. Visciglia, "Endpoint Strichartz estimates for the magnetic Schrödinger equation", *J. Funct. Anal.* **258**:10 (2010), 3227–3240. MR 2011e:35325 Zbl 1188.81061

[Doi 2005] S.-i. Doi, "Smoothness of solutions for Schrödinger equations with unbounded potentials", *Publ. Res. Inst. Math. Sci.* **41**:1 (2005), 175–221. MR 2006d:35043 Zbl 1082.35054

[Erdoğan et al. 2009] M. B. Erdoğan, M. Goldberg, and W. Schlag, "Strichartz and smoothing estimates for Schrödinger operators with almost critical magnetic potentials in three and higher dimensions", *Forum Math.* **21**:4 (2009), 687–722. MR 2010j:35091 Zbl 1181.35208

[Fujiwara 1980] D. Fujiwara, "Remarks on convergence of the Feynman path integrals", *Duke Math. J.* **47**:3 (1980), 559–600. MR 83c:81030 Zbl 0457.35026

[Ginibre and Velo 1985] J. Ginibre and G. Velo, "The global Cauchy problem for the nonlinear Schrödinger equation revisited", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **2**:4 (1985), 309–327. MR 87b:35150 Zbl 0586.35042

[Hassell et al. 2006] A. Hassell, T. Tao, and J. Wunsch, "Sharp Strichartz estimates on nontrapping asymptotically conic manifolds", *Amer. J. Math.* **128**:4 (2006), 963–1024. MR 2007d:58053 Zbl 1177.58019

[Helffer and Sjöstrand 1989] B. Helffer and J. Sjöstrand, "Équation de Schrödinger avec champ magnétique et équation de Harper", pp. 118–197 in *Schrödinger operators* (Sønderborg, 1988), edited by H. Holden and A. Jensen, Lecture Notes in Phys. **345**, Springer, Berlin, 1989. MR 91g:35078 Zbl 0699.35189

[Isozaki and Kitada 1985] H. Isozaki and H. Kitada, "Modified wave operators with time-independent modifiers", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **32**:1 (1985), 77–104. MR 86j:35125 Zbl 0582.35036

[Journé et al. 1991] J.-L. Journé, A. Soffer, and C. D. Sogge, "Decay estimates for Schrödinger operators", *Comm. Pure Appl. Math.* **44**:5 (1991), 573–604. MR 93d:35034 Zbl 0743.35008

[Keel and Tao 1998] M. Keel and T. Tao, "Endpoint Strichartz estimates", *Amer. J. Math.* **120**:5 (1998), 955–980. MR 2000d:35018 Zbl 0922.35028

[Kitada and Kumano-go 1981] H. Kitada and H. Kumano-go, "A family of Fourier integral operators and the fundamental solution for a Schrödinger equation", *Osaka J. Math.* **18**:2 (1981), 291–360. MR 83d:35171 Zbl 0472.35034

[Martinez 2002] A. Martinez, *An introduction to semiclassical and microlocal analysis*, Springer, New York, 2002. MR 2003b:35010 Zbl 0994.35003

[Marzuola et al. 2008] J. Marzuola, J. Metcalfe, and D. Tataru, "Strichartz estimates and local smoothing estimates for asymptotically flat Schrödinger equations", *J. Funct. Anal.* **255**:6 (2008), 1497–1553. MR 2011c:35485 Zbl 1180.35187

[Mizutani 2012] H. Mizutani, "Strichartz estimates for Schrödinger equations on scattering manifolds", *Comm. Partial Differential Equations* **37**:2 (2012), 169–224. MR 2876829 Zbl 1244.35020

[Mizutani 2013] H. Mizutani, "Strichartz estimates for Schrödinger equations with variable coefficients and potentials at most linear at spatial infinity", *J. Math. Soc. Japan* **65**:3 (2013), 687–721. MR 3084976 Zbl 1273.35232

[Robbiano and Zuily 2005] L. Robbiano and C. Zuily, *Strichartz estimates for Schrödinger equations with variable coefficients*, Mém. Soc. Math. Fr. (N.S.), 2005.

[Robert 1987] D. Robert, *Autour de l'approximation semi-classique*, Progress in Mathematics **68**, Birkhäuser, 1987. MR 89g:81016 Zbl 0621.35001

[Robert 1994] D. Robert, "Relative time-delay for perturbations of elliptic operators and semiclassical asymptotics", *J. Funct. Anal.* **126**:1 (1994), 36–82. MR 95j:35162 Zbl 0813.35073

[Robert and Tamura 1987] D. Robert and H. Tamura, "Semiclassical estimates for resolvents and asymptotics for total scattering cross-sections", *Ann. Inst. H. Poincaré Phys. Théor.* **46**:4 (1987), 415–442. MR 89b:81041 Zbl 0648.35066

[Schlag 2007] W. Schlag, "Dispersive estimates for Schrödinger operators: a survey", pp. 255–285 in *Mathematical aspects of nonlinear dispersive equations*, edited by B. J. et al., Ann. of Math. Stud. **163**, Princeton Univ. Press, 2007. MR 2009k:35043 Zbl 1143.35001

[Staffilani and Tataru 2002] G. Staffilani and D. Tataru, "Strichartz estimates for a Schrödinger operator with nonsmooth coefficients", *Comm. Partial Differential Equations* **27**:7-8 (2002), 1337–1372. MR 2003f:35248 Zbl 1010.35015

[Strichartz 1977] R. S. Strichartz, "Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations", *Duke Math. J.* **44**:3 (1977), 705–714. MR 58 #23577 Zbl 0372.35001

[Tataru 2008] D. Tataru, "Parametrices and dispersive estimates for Schrödinger operators with variable coefficients", *Amer. J. Math.* **130**:3 (2008), 571–634. MR 2009m:35077 Zbl 1159.35315

[Yajima 1987] K. Yajima, "Existence of solutions for Schrödinger evolution equations", *Comm. Math. Phys.* **110**:3 (1987), 415–426. MR 88e:35048 Zbl 0638.35036

[Yajima 1991] K. Yajima, "Schrödinger evolution equations with magnetic fields", *J. Analyse Math.* **56** (1991), 29–76. MR 94k:35270 Zbl 0739.35083

[Yajima 1998] K. Yajima, "Boundedness and continuity of the fundamental solution of the time dependent Schrödinger equation with singular potentials", *Tohoku Math. J.* (2) **50**:4 (1998), 577–595. MR 99k:35004 Zbl 0920.35007

[Yajima 2005] K. Yajima, "Dispersive estimates for Schrödinger equations with threshold resonance and eigenvalue", *Comm. Math. Phys.* **259**:2 (2005), 475–509. MR 2006h:35042 Zbl 1079.81021

[Yajima and Zhang 2004] K. Yajima and G. Zhang, "Local smoothing property and Strichartz inequality for Schrödinger equations with potentials superquadratic at infinity", *J. Differential Equations* **202**:1 (2004), 81–110. MR 2005f:35266 Zbl 1060.35121

Received 25 Feb 2012. Revised 24 Sep 2012. Accepted 19 Jan 2013.

HARUYA MIZUTANI: haruya@math.sci.osaka-u.ac.jp
Department of Mathematics, Graduate School of Science, Osaka University, Toyonaka 560-0043, Japan

and

Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan

ANALYSIS AND PDE
Vol. 6, No. 8, 2013
dx.doi.org/10.2140/apde.2013.6.1899

UNIFORMITY OF HARMONIC MAP HEAT FLOW AT INFINITE TIME

LONGZHI LIN

We show an energy convexity along any harmonic map heat flow with small initial energy and fixed boundary data on the unit 2-disk. In particular, this gives an affirmative answer to a question raised by W. Minicozzi asking whether such harmonic map heat flow converges uniformly in time strongly in the $W^{1,2}$-topology, as time goes to infinity, to the unique limiting harmonic map.

1. Introduction

Given a compact Riemannian manifold $\mathcal{M}$ and a closed (that is, compact and without boundary) Riemannian manifold $\mathcal{N}$ which is an isometrically embedded submanifold of $\mathbb{R}^n$, we can define the *Dirichlet* energy of a map $u \in W^{1,2}(\mathcal{M}, \mathcal{N})$:

$$\text{Energy}(u) = E(u) = \frac{1}{2}\int_{\mathcal{M}} |\nabla u|^2 \, dv_{\mathcal{M}}, \tag{1-1}$$

where $W^{1,2}(\mathcal{M}, \mathcal{N})$ is the class of maps

$$\left\{ u \in L^1_{\text{loc}}(\mathcal{M}, \mathbb{R}^n) : \int_{\mathcal{M}} |\nabla u|^2 \, dv_{\mathcal{M}} < +\infty, \, u(x) \in \mathcal{N} \text{ for a.e. } x \in \mathcal{M} \right\}.$$

The tension field $\tau(u) \in \Gamma(u^*(T\mathcal{N}))$ is the vector field along u representing the negative L^2-gradient of $E(u)$. A weakly harmonic map u from $\mathcal{M}$ to $\mathcal{N}$ is a critical point of the energy functional $E(u)$ in the distribution sense, that is, the tension field $\tau(u)$ vanishes, and it solves the Euler–Lagrange equation

$$-\Delta_{\mathcal{M}} u = \Pi(u)(\nabla u, \nabla u), \tag{1-2}$$

where $u = (u^1, \dots, u^n)$ and $\Pi(u)$ denotes the second fundamental form of $\mathcal{N} \hookrightarrow \mathbb{R}^n$ at the point u. We refer to this system of elliptic equations as the harmonic map equation.

A natural way to control the tension field for an energy minimizing sequence of maps and to get the existence of harmonic maps from $\mathcal{M}$ to $\mathcal{N}$ is to consider the initial (-boundary) value problem:

$$\begin{cases} u_t - \Delta_{\mathcal{M}} u = \Pi(u)(\nabla u, \nabla u) & \text{on } \mathcal{M} \times (0, T), \\ u(x, 0) = u_0(x) & \text{for } x \in \mathcal{M}, \\ u(x, t) = \chi(x) = u_0|_{\partial\mathcal{M}} & \text{for all } t \geq 0, \, x \in \partial\mathcal{M} \text{ if } \partial\mathcal{M} \neq \varnothing, \end{cases} \tag{1-3}$$

MSC2010: 53C44, 58E20.
Keywords: harmonic map heat flow, energy convexity, uniform convergence.

where $u = (u^1, \dots, u^n)$ and $T > 0$. We refer to this system of parabolic equations as the harmonic map heat flow, to the map u_0 as the initial data, and to the map χ as the boundary data. Given $u_0 \in W^{1,2}(\mathcal{M}, \mathcal{N})$ and $\chi = u_0|_{\partial\mathcal{M}} \in W^{1/2,2}(\partial\mathcal{M}, \mathcal{N})$, we define $u \in W^{1,2}(\mathcal{M} \times [0, T], \mathcal{N})$ to be the weak solution of (1-3) if

$$\int_0^T \int_{\mathcal{M}} \langle u_t, \xi\rangle + \langle \nabla u, \nabla \xi\rangle - \langle \Pi(u)(\nabla u, \nabla u), \xi\rangle \, dx\, dt = 0 \tag{1-4}$$

for any $\xi \in C_c^\infty(\mathcal{M} \times (0, T), \mathbb{R}^n)$.

In the fundamental paper where the harmonic map heat flow was first introduced, Eells and Sampson [1964] proved that the harmonic map heat flow exists for all time in the case where the source domain $\mathcal{M}$ (of arbitrary dimensions) is without boundary and the target manifold $\mathcal{N}$ has nonpositive sectional curvature. They also proved that there exists some sequence of times $t_i \nearrow +\infty$ such that

$$u_\infty = \lim_{i\to\infty} u(\,\cdot\,, t_i)$$

is a harmonic map from $\mathcal{M}$ to $\mathcal{N}$. The case in which the source domain $\mathcal{M}$ has boundary was dealt with in [Hamilton 1975] under the same curvature assumption on $\mathcal{N}$. The question of uniformity of the convergence in time of the flow considered by Eells and Sampson was left open at that stage, but it was settled later by Hartman. We shall state their results in the following theorem.

Theorem 1.1 [Eells and Sampson 1964; Hartman 1967]. *Suppose that $\mathcal{M}$ and $\mathcal{N}$ are two closed Riemannian manifolds and that $\mathcal{N}$ has nonpositive sectional curvature. Then, given any $u_0 \in C^1(\mathcal{M}, \mathcal{N})$, the harmonic map heat flow has a unique solution $u \in C^1(\mathcal{M} \times [0, \infty), \mathcal{N}) \cap C^\infty(\mathcal{M} \times (0, \infty), \mathcal{N})$. Moreover,*

$$u_\infty = \lim_{t\to\infty} u(\,\cdot\,, t) \tag{1-5}$$

exists uniformly in C^k-topology for all $k \geq 0$ and u_∞ is a harmonic map homotopic to u_0.

Other similar uniformity results were obtainable under various assumptions on the target manifold $\mathcal{N}$, such as being real analytic [Simon 1983] or admitting a strictly convex function; see also the interesting paper [Topping 1997] for harmonic map heat flow in a special case in which both the source and target manifolds are 2-spheres $\mathbb{S}^2$.

When the dimension of the source domain $\mathcal{M}$ is two, things are particularly interesting because the energy functional $E(u)$ and the harmonic map equation (1-2) are conformally invariant in this critical dimension. Regarding the harmonic map heat flow (1-3) from surfaces to a general closed target manifold $\mathcal{N}$, the first fundamental work was [Struwe 1985], dealing with the case $\partial\mathcal{M} = \varnothing$, where "bubbles" may occur and have been analyzed in detail. This result was then extended to the case $\partial\mathcal{M} \neq \varnothing$ with Dirichlet boundary condition in [Chang 1989]. If the initial energy $E(u_0)$ is sufficiently small, it is well known by now that the weak solution of (1-3) is smooth (in the interior) by the results of [Freire 1995; 1996] using the so-called moving frame technique introduced by Hélein (see, for example, [Hélein 2002]). We will state their ε-regularity theorem required in this paper, and include an alternative proof of it for self-containedness; it uses the main tool of our current work, which we call Rivière's gauge decomposition (Theorem 3.7).

Theorem 1.2 ([Freire 1995; 1996]; cf. [Struwe 1985; Chang 1989; Wang 2012]). *Let $\mathcal{M}$ be a simply connected compact Riemannian surface and $\mathcal{N}$ a closed Riemannian manifold. There exists $\varepsilon_0 > 0$ depending only on $\mathcal{M}$ and $\mathcal{N}$ such that the following is true. For each initial data $u_0 \in W^{1,2}(\mathcal{M}, \mathcal{N})$ with $E(u_0) < \varepsilon_0$ and the boundary data $\chi = u_0|_{\partial\mathcal{M}}$ in the case where $\partial\mathcal{M} \neq \varnothing$, there exists a unique global weak solution $u \in W^{1,2}(\mathcal{M} \times [0, \infty), \mathcal{N})$ for which $E(u(\cdot, t))$ is nonincreasing in t. Also, u is smooth in $\mathcal{M} \times [1, \infty)$ and, for any $t_2 > t_1 \geq 1$, we have*

$$2 \int_{t_1}^{t_2} \int_{\mathcal{M}} |u_t|^2 = \int_{B_1} |\nabla u(\cdot, t_1)|^2 - \int_{B_1} |\nabla u(\cdot, t_2)|^2. \tag{1-6}$$

Moreover, there exists some sequence of times $t_i \nearrow +\infty$ such that

$$u_\infty = \lim_{i\to\infty} u(\cdot, t_i) \tag{1-7}$$

exists in the C^k-topology for any $k \geq 0$ and u_∞ is a harmonic map from $\mathcal{M}$ to $\mathcal{N}$.

Remark 1.3. In particular, in order to avoid the "bubble" (singularity) along the harmonic map heat flow, a priori we may choose $\varepsilon_0 < K_1 + K_2$ where

$$K_1 = \inf\{E(v) \mid v \in W^{1,2}(\mathcal{M}, \mathcal{N}) \text{ and } v|_{\partial\mathcal{M}} = \chi\}$$

and

$$K_2 = \inf\{E(v) \mid v : \mathbb{S}^2 \to \mathcal{N} \text{ is nonconstant and harmonic}\} > 0.$$

Remark 1.4. Freire's regularity results for harmonic map heat flow represent a parabolic version of the regularity theorem of Hélein stating that weakly harmonic maps from surfaces are regular; see, for example, [Hélein 2002].

A tempting question to ask is whether, for a general closed target manifold $\mathcal{N}$ (without additional geometric assumptions), one could establish uniformity results for the harmonic map heat flow similar to Theorem 1.1. In particular, is the convergence (1-7) in Theorem 1.2 uniform for all time in the natural $W^{1,2}$-topology, say? In view of the conformal invariance of the energy functional $E(u)$ in dimension two, the condition of small energy seems to be a natural candidate to work with in order to get such uniformity of the convergence in time for the flow. We will show in the following that this is indeed the case. In what follows we will concentrate on the case where the source domain $\mathcal{M}$ is a simply connected compact Riemannian surface with boundary. More precisely, we focus on domains which are conformally equivalent to the unit 2-disk $B_1 \subset \mathbb{R}^2$. From now on we will only work on B_1:

$$\begin{cases} u_t - \Delta u = \Pi(u)(\nabla u, \nabla u) & \text{on } B_1 \times (0, T), \\ u(x, 0) = u_0(x) & \text{for } x \in B_1, \\ u(x, t) = \chi(x) = u_0|_{\partial B_1} & \text{for all } t \geq 0 \text{ and } x \in \partial B_1, \end{cases} \tag{1-8}$$

where Δ is the usual Laplacian $\Delta = \sum_{i=1}^2 \partial^2/\partial x_i^2$ in $\mathbb{R}^2$. All the arguments could be easily modified to apply to the general case.

Notation 1.5. In what follows, $\nabla = (\partial_x, \partial_y)$ is the gradient operator in $\mathbb{R}^2$ and $\nabla^\perp = (-\partial_y, \partial_x)$ denotes the orthogonal gradient (that is, $\nabla^\perp$ is the ∇-operator rotated by $\pi/2$).

Now we state the main theorem of this paper.

Theorem 1.6. *Let $\mathcal{N}$ be a closed Riemannian manifold. There exist $\varepsilon_0, T_0 > 0$ depending only on $\mathcal{N}$ such that if $u \in W^{1,2}(B_1 \times [0,\infty), \mathcal{N})$ is a global weak solution to the harmonic map heat flow* (1-8) *with $E(u_0) < \varepsilon_0$, $E(u(\cdot,t))$ is nonincreasing in t, and $u(\cdot,t)|_{\partial B_1} = \chi$ for all $t \geq 0$, then, for all $t_2 > t_1 \geq T_0$, we have the energy convexity*

$$\frac{1}{4}\int_{B_1} |\nabla u(\cdot,t_1) - \nabla u(\cdot,t_2)|^2 \leq \int_{B_1} |\nabla u(\cdot,t_1)|^2 - \int_{B_1} |\nabla u(\cdot,t_2)|^2. \tag{1-9}$$

Remark 1.7. We do not know if the energy convexity (1-9) holds for all $t_2 > t_1 \geq 0$. In the following arguments we agree to let ε_0 be sufficiently small and T_0 be sufficiently large, as needed.

Our approach to the proof of Theorem 1.6 is based on the technique we call Rivière's gauge decomposition, introduced in [Rivière 2007]; see Section 3. Immediate applications of Theorem 1.6 are:

Corollary 1.8. *Let $\mathcal{N}$ be a closed Riemannian manifold. There exists $\varepsilon_0 > 0$ depending only on $\mathcal{N}$ such that if $u \in W^{1,2}(B_1 \times [0,\infty), \mathcal{N})$ is a global weak solution to the harmonic map heat flow* (1-8) *with $E(u_0) < \varepsilon_0$, $E(u(\cdot,t))$ is nonincreasing in t, and $u(\cdot,t)|_{\partial B_1} = \chi$ for all $t \geq 0$, then*

$$u(\cdot,t) \to u_\infty \quad \text{uniformly as } t \to +\infty \quad \text{strongly in } W^{1,2}(B_1, \mathbb{R}^n), \tag{1-10}$$

where u_∞ is the unique harmonic map with $E(u_\infty) < \varepsilon_0$ and boundary data χ.

Corollary 1.9. *Let $\mathcal{M}$ be a two dimensional domain that is conformally equivalent to B_1 and has smooth boundary, and let $\mathcal{N}$ be a closed Riemannian manifold. Suppose the initial energy $E(u_0) < \varepsilon_0$. Then the harmonic map heat flow* (1-3) *with initial data $u_0 \in C^{2,\alpha}(\overline{\mathcal{M}}, \mathcal{N})$ and boundary data $\chi \in C^{2,\alpha}(\partial\mathcal{M}, \mathcal{N})$, considered by Chang [1989], converges uniformly in time strongly in $W^{1,2}(\mathcal{M}, \mathcal{N})$ to the unique harmonic map $u_\infty \in C^{2,\alpha}_\chi(\overline{\mathcal{M}}, \mathcal{N})$.*

Remark 1.10. We do not know if a harmonic map heat flow can be nonuniform without the small energy assumption. In view of the nonuniqueness results of Brezis and Coron [1983] and Jost [1984] for harmonic maps (with large energy) sharing the same boundary data on ∂B_1, it is quite possible that the small energy assumption is necessary for the energy convexity and uniform convergence of the flow in Theorem 1.6 and Corollary 1.8 to hold.

Remark 1.11. Colding and Minicozzi [2008a] showed an energy convexity for weakly harmonic maps with small energy on B_1: there exists $\varepsilon_0 > 0$ such that if $u, v \in W^{1,2}(B_1, \mathcal{N})$ with $u|_{\partial B_1} = v|_{\partial B_1}$, $E(u) < \varepsilon_0$, and u is weakly harmonic, then we have the energy convexity

$$\frac{1}{2}\int_{B_1} |\nabla v - \nabla u|^2 \leq \int_{B_1} |\nabla v|^2 - \int_{B_1} |\nabla u|^2. \tag{1-11}$$

See [Lamm and Lin 2013] for an alternative proof of this energy convexity using the same techniques used in the present paper. A direct consequence of (1-11) is that u_∞ in Corollary 1.8 is unique in the class

$$\{v \in W^{1,2}(B_1, \mathbb{R}^n) : E(v) < \varepsilon_0 \text{ and } v|_{\partial B_1} = \chi\};$$

see [Colding and Minicozzi 2008a, Corollary 3.3].

The paper is organized as follows. In Section 2 we present some heuristic arguments and elaborate on the idea of the proof of the main theorem, Theorem 1.6. In Section 3 we review the main tool of our proof, namely, Rivière's gauge decomposition technique adapted to the case of harmonic map heat flow. In Section 4 we show improved estimates for Rivière's matrices B and P, which are the two key ingredients of our proof. We finish the proof of our main theorem in Section 5.

2. Heuristic arguments and the idea of the proof

In this section we will present some heuristic arguments and sketch the basic idea of the proof of Theorem 1.6. We will abbreviate $u(\,\cdot\,,t)$ to $u(t)$. In order to prove the energy convexity (1-9) along the harmonic map heat flow, that is, there exists some $T_0>0$ such that, for all $t_2>t_1\geq T_0$, we have

$$\frac{1}{4}\int_{B_1}|\nabla u(\,\cdot\,,t_1)-\nabla u(\,\cdot\,,t_2)|^2\leq\int_{B_1}|\nabla u(\,\cdot\,,t_1)|^2-\int_{B_1}|\nabla u(\,\cdot\,,t_2)|^2, \tag{2-1}$$

it suffices to show

$$\Psi\geq-\biggl(\int_{B_1}|\nabla u(t_1)|^2-\int_{B_1}|\nabla u(t_2)|^2\biggr)-\frac{1}{2}\int_{B_1}|\nabla u(t_1)-\nabla u(t_2)|^2, \tag{2-2}$$

where (using that $u(\,\cdot\,,t)|_{\partial B_1}=\chi$ for all $t\geq 0$ and the flow equation (1-8))

$$\begin{aligned}\Psi&:=\int_{B_1}|\nabla u(t_1)|^2-\int_{B_1}|\nabla u(t_2)|^2-\int_{B_1}|\nabla u(t_1)-\nabla u(t_2)|^2\\&=2\int_{B_1}\langle\nabla u(t_1)-\nabla u(t_2),\nabla u(t_2)\rangle\\&=-2\int_{B_1}\langle u(t_1)-u(t_2),u_t(t_2)-\Pi(u)(\nabla u,\nabla u)(t_2)\rangle.\end{aligned} \tag{2-3}$$

Now note that for any $p,q\in\mathcal{N}$, there exists some constant $C>0$ depending only on $\mathcal{N}$ such that $|(p-q)^{\perp}|\leq C|p-q|^2$, where the superscript $\perp$ denotes the normal component of a vector; see, for example, [Colding and Minicozzi 2008b, Lemma A.1]. Therefore, using the fact that $\Pi(u)(\nabla u,\nabla u)\perp T_u\mathcal{N}$ and the Cauchy–Schwarz inequality, (2-3) yields

$$\begin{aligned}\Psi&\geq-2\biggl(\int_{B_1}|u(t_1)-u(t_2)|^2\biggr)^{1/2}\biggl(\int_{B_1}|u_t(t_2)|^2\biggr)^{1/2}-C\int_{B_1}|(u(t_1)-u(t_2))^{\perp}|\,|\nabla u(t_2)|^2\\&\geq-2\sqrt{t_2-t_1}\biggl(\int_{t_1}^{t_2}\int_{B_1}|u_t|^2\biggr)^{1/2}\biggl(\int_{B_1}|u_t(t_2)|^2\biggr)^{1/2}-C\int_{B_1}|u(t_1)-u(t_2)|^2\,|\nabla u(t_2)|^2,\end{aligned}$$

where we also used the smoothness and compactness of the target manifold $\mathcal{N}$. Here and throughout the rest of the paper, $C>0$ will denote a universal constant depending only on $\mathcal{N}$ unless otherwise stated.

Since we have (1-6) and ε_0 can always be chosen sufficiently small, we know that (2-2) will be achieved if we can show the following two key propositions.

Proposition 2.1. *Let $u(x,t)$ be as in Theorem 1.6. Then there exists $T_0>0$ such that, for all $t_2>t_1\geq T_0$, we have*

$$\int_{B_1}|u_t(t_2)|^2\leq\frac{1}{t_2-t_1}\int_{t_1}^{t_2}\int_{B_1}|u_t|^2. \tag{2-4}$$

Remark 2.2. The key point of Proposition 2.1 is that (2-4) is valid for all $t_2>t_1\geq T_0$. We will see that, in fact, $\int_{B_1}|u_t(t)|^2$ is nonincreasing along the flow after T_0, which yields (2-4); cf. Lemma 2.5 and (5-16) below. A similar but weaker estimate was shown when the source domain of the heat flow is boundaryless [Struwe 1985, Equation (3.5)], which turned out to be the key estimate needed in Struwe's proof.

Proposition 2.3. *Let $u(x,t)$ be as in Theorem 1.6. Then there exists $T_0>0$ such that, for all $t_2>t_1\geq T_0$, we have*

$$\int_{B_1}|u(t_1)-u(t_2)|^2\,|\nabla u(t_2)|^2\leq C\varepsilon_0\int_{B_1}|\nabla u(t_1)-\nabla u(t_2)|^2. \tag{2-5}$$

If one were able to get

$$\|\nabla u(t_2)\|_{L^\infty(B_1)}\leq C\sqrt{\varepsilon_0}, \tag{2-6}$$

(2-5) would have been automatically true by Poincaré's inequality. However, without imposing any regularity information on the boundary data χ, it will be hopeless to get such a strong global pointwise gradient estimate. In fact, even if we look at the stationary case, that is, $W^{1,2}$ -weakly harmonic maps on B_1, it is easy to convince oneself that it is unreasonable to expect regularity with global estimates on the whole B_1 better than $W^{2,2}$ in general.

Nevertheless, not all hope is lost to show estimates (2-4) and (2-5). Indeed, the following lemma is true, which validates Proposition 2.3 under some extra assumptions.

Lemma 2.4. *Let $u(x,t)$ be as in Theorem 1.6 and suppose that, for all $t_2>t_1\geq T_0\geq 1$, we can solve the following Dirichlet problem for $\psi\in W_0^{1,2}\cap L^\infty(B_1)$:*

$$\begin{cases}\Delta\psi=|\nabla u(t_2)|^2 & \text{in } B_1,\\ \psi=0 & \text{on } \partial B_1\end{cases} \tag{2-7}$$

with the estimate

$$\|\psi\|_{L^\infty(B_1)}+\|\nabla\psi\|_{L^2(B_1)}\leq C\varepsilon_0. \tag{2-8}$$

Then Proposition 2.3 holds.

Proof. The proof is essentially taken from [Colding and Minicozzi 2008a]. Substituting (2-7) into the left-hand side of (2-5) yields (using also that $u(t_1)=u(t_2)=\chi$ on ∂B_1)

$$\begin{aligned}\int_{B_1}|u(t_1)-u(t_2)|^2\,|\nabla u(t_2)|^2&=\int_{B_1}|u(t_1)-u(t_2)|^2\Delta\psi\leq\int_{B_1}\big|\nabla|u(t_1)-u(t_2)|^2\big|\,|\nabla\psi|\\ &\leq 2\Big(\int_{B_1}|\nabla u(t_1)-\nabla u(t_2)|^2\Big)^{1/2}\Big(\int_{B_1}|u(t_1)-u(t_2)|^2\,|\nabla\psi|^2\Big)^{1/2},\end{aligned} \tag{2-9}$$

where we have applied Stokes' theorem to $\operatorname{div}(|u(t_1)-u(t_2)|^2\nabla\psi)$ and used the Cauchy–Schwarz inequality. Now, applying Stokes' theorem to $\operatorname{div}(|u(t_1)-u(t_2)|^2\psi\nabla\psi)$ and using that $\Delta\psi\geq 0$ and (2-9),

we have

$$\int_{B_1} |u(t_1)-u(t_2)|^2 |\nabla\psi|^2 \le \int_{B_1} |\psi|(|u(t_1)-u(t_2)|^2 \Delta\psi + \big|\nabla|u(t_1)-u(t_2)|^2\big| |\nabla\psi|)$$

$$\le 4\|\psi\|_{L^\infty}\left(\int_{B_1} |\nabla u(t_1)-\nabla u(t_2)|^2\right)^{1/2}\left(\int_{B_1} |u(t_1)-u(t_2)|^2|\nabla\psi|^2\right)^{1/2}, \quad (2\text{-}10)$$

so that

$$\left(\int_{B_1} |u(t_1)-u(t_2)|^2\,|\nabla\psi|^2\right)^{1/2} \le 4\|\psi\|_{L^\infty}\left(\int_{B_1} |\nabla u(t_1)-\nabla u(t_2)|^2\right)^{1/2}. \quad (2\text{-}11)$$

Finally, substituting (2-11) back into (2-9) and combining with (2-8) (and choosing ε_0 sufficiently small), yields

$$\int_{B_1} |u(t_1)-u(t_2)|^2\,|\nabla u(t_2)|^2 \le C\|\psi\|_{L^\infty}\int_{B_1} |\nabla u(t_1)-\nabla u(t_2)|^2 \le C\varepsilon_0 \int_{B_1} |\nabla u(t_1)-\nabla u(t_2)|^2,$$

which is just (2-5). □

Similarly, we can show the following lemma, which states, under some extra conditions, that $\int_{B_1} |u_t(t)|^2$ is nonincreasing along the harmonic map heat flow after some $T_0 > 0$ and Proposition 2.1 can be validated in this case.

Lemma 2.5. *Let $u(x,t)$ be as in Theorem 1.6. For any $t_2 > t_1 \ge T_0 \ge 1$, suppose that for any $t_0 \in [t_1, t_2]$ we can solve the following Dirichlet problem for $\psi \in W_0^{1,2} \cap L^\infty(B_1)$:*

$$\begin{cases} \Delta\psi = |\nabla u(t_0)|^2 & \text{in } B_1, \\ \psi = 0 & \text{on } \partial B_1, \end{cases} \quad (2\text{-}12)$$

with the estimate

$$\|\psi\|_{L^\infty(B_1)} + \|\nabla\psi\|_{L^2(B_1)} \le C\varepsilon_0. \quad (2\text{-}13)$$

Then we have

$$\int_{B_1} |u_t(t_2)|^2 \le \int_{B_1} |u_t(t_1)|^2. \quad (2\text{-}14)$$

In particular, Proposition 2.1 holds if (2-12) and (2-13) are valid for any $t_0 \in [t_1, t_2]$ and any $t_2 > t_1 \ge T_0 \ge 1$.

Proof. Differentiating the flow equation (1-8) with respect to t, multiplying with u_t, and integrating over $B_1 \times [t_1, t_2]$, we have

$$\frac{1}{2}\int_{t_1}^{t_2}\int_{B_1} \partial_t |u_t|^2 + \int_{t_1}^{t_2}\int_{B_1} |\nabla u_t|^2 \le C\int_{t_1}^{t_2}\int_{B_1} |u_t|^2\,|\nabla u|^2 + |u_t|\,|\nabla u|\,|\nabla u_t|$$

$$\le \frac{1}{2}\int_{t_1}^{t_2}\int_{B_1} |\nabla u_t|^2 + C\int_{t_1}^{t_2}\int_{B_1} |u_t|^2\,|\nabla u|^2. \quad (2\text{-}15)$$

Since (2-12) and (2-13) are valid for any $t_0 \in [t_1, t_2]$, we can use the same arguments as in the proof of Lemma 2.4 to get an estimate for $\int_{B_1} |u_t|^2\,|\nabla u|^2$ at the time t_0 slice. Indeed, similarly to (2-5) (that is, replacing $u(t_1)-u(t_2)$ by $u_t(t_0)$), for any $t_0\in[t_1,t_2]$, we have

$$\int_{B_1}|u_t|^2\,|\nabla u|^2(t_0)\le C\|\psi\|_{L^\infty}\int_{B_1}|\nabla u_t(t_0)|^2\le C\varepsilon_0\int_{B_1}|\nabla u_t(t_0)|^2. \tag{2-16}$$

Inserting (2-16) back into (2-15) (for any $t_0\in[t_1,t_2]$), we see that the right-hand side of (2-15) can be absorbed into the left-hand side if we choose ε_0 sufficiently small. This implies that we have (2-14) for any such $t_2>t_1\ge T_0$. In the above calculations, we should treat u_t as a difference quotient: $u_t(\,\cdot\,,t)=\lim_{h\to 0^+}(u(\,\cdot\,,t+h)-u(\,\cdot\,,t))/h$, which is zero on ∂B_1 for all $t\ge 1$; moreover, we have denoted $\nabla u_t(\,\cdot\,,t)=\lim_{h\to 0^+}(\nabla(u(\,\cdot\,,t+h)-u(\,\cdot\,,t)))/h$ and all the calculations are valid for any fixed $h>0$. We then we take $h\to 0^+$ to conclude (2-14).

If (2-12) and (2-13) are valid for any $t_0\in[t_1,t_2]$ and any $t_2>t_1\ge T_0\ge 1$, then, in view of (2-14), estimating by the mean value of $|u_t|^2$ over $B_1\times[t_1,t_2]$ gives Proposition 2.1. □

Therefore, everything boils down to validating the assumptions in Lemmas 2.4 and 2.5, that is, the existence of such functions ψ satisfying (2-7), (2-8) and (2-12), (2-13), respectively, for any $t_0\ge T_0$ for some $T_0\ge 1$. We point out that, a priori we only know that the energy density $|\nabla u(t)|^2$ lies in $L^1(B_1)$ with global estimate $\big\||\nabla u(t)|^2\big\|_{L^1(B_1)}\le\varepsilon_0$ for any fixed t. But L^1 is the borderline case in which the standard L^p-theory for the Dirichlet problem (2-7) with estimate (2-8) fails!

However, the following regularity theorem for boundary value problems in the local Hardy space $h^1(B_1)$ sheds new light on the problem of validating the assumptions in Lemmas 2.4 and 2.5. Here the local Hardy space $h^1(B_1)$ is a strict subspace of $L^1(B_1)$ and we will recall its definition in Definition 2.8 below.

Theorem 2.6 (cf. [Semmes 1994, Theorem 1.100; Chang et al. 1993, Theorem 5.1]). *Let $f\in h^1(B_1)$ such that $f\ge 0$ a.e. in B_1. Then there exists a function $\psi\in L^\infty\cap W_0^{1,2}(B_1)$ solving the Dirichlet problem*

$$\begin{cases}\Delta\psi=f & \text{in } B_1,\\ \psi=0 & \text{on } \partial B_1.\end{cases} \tag{2-17}$$

Moreover, there exists a constant $C>0$ such that

$$\|\psi\|_{L^\infty(B_1)}+\|\nabla\psi\|_{L^2(B_1)}\le C\|f\|_{h^1(B_1)}. \tag{2-18}$$

Proof. For self-containedness, we include an elementary proof of this theorem in Appendix A. □

Remark 2.7. This theorem can be thought of as a generalization of a result from [Müller 1990]; cf. Wente's lemma (Lemma 3.6). For a more general version of this theorem, we refer to Chang, Krantz, and Stein's work [Chang et al. 1993].

Definition 2.8 [Miyachi 1990]. Choose a Schwartz function $\phi\in C_0^\infty(B_1)$ such that $\int_{B_1}\phi\,dx=1$ and let $\phi_t(x)=t^{-2}\phi(x/t)$. For a measurable function f defined in B_1, we say that f lies in the local Hardy space $h^1(B_1)$ if the radial maximal function of f

$$f^*(x)=\sup_{0<t<1-|x|}\left|\int_{B_t(x)}\frac{1}{t^2}\phi\left(\frac{x-y}{t}\right)f(y)\,dy\right|(x)=\sup_{0<t<1-|x|}|\phi_t*f|(x) \tag{2-19}$$

belongs to $L^1(B_1)$ and we define

$$\|f\|_{h^1(B_1)} = \|f^*(x)\|_{L^1(B_1)}. \tag{2-20}$$

It follows immediately that $h^1(B_1)$ is a strict subspace of $L^1(B_1)$ and

$$\|f\|_{L^1(B_1)} \le \|f\|_{h^1(B_1)}.$$

It is also clear that if $f \in L^p(B_1)$ for some $p > 1$, then $\|f\|_{h^1(B_1)} \le C\|f\|_{L^p(B_1)}$.

We remark that the local Hardy spaces h^1 (or the global version $\mathscr{H}^1$) act as replacements for L^1 in Calderon–Zygmund estimates. Therefore, by Theorem 2.6, if we can somehow manage to obtain a "slightly" improved global estimate for $|\nabla u|^2$ from $L^1(B_1)$ to $h^1(B_1)$ for all $t_0 \ge T_0$, it will be sufficient to validate the assumptions in Lemmas 2.4 and 2.5. As mentioned above, the subtlety is that, without imposing any regularity information on the boundary data χ, global estimates are very difficult to obtain.

The rest of the paper is devoted to validating the assumptions in Lemmas 2.4 and 2.5. Namely, in view of Theorem 2.6, it suffices to show there exists $T_0 > 0$ such that

$$\big\| |\nabla u(t_0)|^2 \big\|_{h^1(B_1)} \le C\varepsilon_0 \quad \text{for any } t_0 \ge T_0. \tag{2-21}$$

The point here is that no pointwise estimate on ∇u such as (2-6) is needed, and instead, a (weaker) improved global integral estimate (2-21) will be sufficient and turns out to be the key to the proof of Theorem 1.6.

3. Analysis of harmonic map heat flow using Rivière's gauge

Regarding the regularity of weakly harmonic maps from surfaces, Hélein (see, for example, [Hélein 2002]) proved the interior regularity with the help of the so-called Coulomb or moving frame, and Qing [1995] showed the continuity up to the boundary in the case of continuous boundary data based on Hélein's technique. Rivière [2007] succeeded in writing the 2-dimensional conformally invariant nonlinear system of elliptic PDE's (which includes the weakly harmonic map equation (1-2)) in the form

$$-\Delta u^i = \Omega^i_j \cdot \nabla u^j, \quad i = 1, 2, \ldots, n, \quad \text{or} \quad -\Delta u = \Omega \cdot \nabla u \tag{3-1}$$

with $\Omega = (\Omega^i_j)_{1\le i,j\le n} \in L^2(B_1, \mathfrak{so}(n) \otimes \bigwedge^1\mathbb{R}^2)$ and $\Omega^i_j = -\Omega^j_i$ (antisymmetry). Here and throughout the paper, the Einstein summation convention is used. We refer to the system of equations (3-1) as Rivière's equation. This special form of the nonlinearity enabled Rivière to obtain a conservation law for this system of PDE's (see (3-8) below), which is accomplished via a technique that we call Rivière's gauge decomposition. More precisely, following the strategy of [Uhlenbeck 1982], Rivière [2007] used an algebraic feature of Ω — its antisymmetry — to construct $\xi \in W^{1,2}_0(B_1, \mathfrak{so}(n))$ and a gauge transformation matrix $P \in W^{1,2} \cap L^\infty(B_1, \mathrm{SO}(n))$ (which pointwise almost everywhere is an orthogonal matrix in $\mathbb{R}^{n\times n}$) satisfying some good properties.

Theorem 3.1 [Rivière 2007, Lemma A.3]. *There exist $\varepsilon > 0$ and $C > 0$ such that, for every Ω in $L^2(B_1, \mathfrak{so}(n) \otimes \bigwedge^1 \mathbb{R}^2)$ satisfying*

$$\int_{B_1} |\Omega|^2 \leq \varepsilon,$$

there exist $\xi \in W_0^{1,2}(B_1, \mathfrak{so}(n))$ and $P \in W^{1,2}(B_1, \mathrm{SO}(n))$ such that

$$\nabla^{\perp}\xi = P^T \nabla P + P^T \Omega P \text{ in } B_1 \quad \text{with } \xi = 0 \text{ on } \partial B_1, \tag{3-2}$$

and

$$\|\nabla \xi\|_{L^2(B_1)} + \|\nabla P\|_{L^2(B_1)} \leq C \|\Omega\|_{L^2(B_1)}. \tag{3-3}$$

Here the superscript T denotes the transpose of a matrix.

Remark 3.2. Multiplying both sides of (3-2) by P from the left gives (with indices and $1 \leq m, z \leq n$)

$$\nabla P^i_j = P^i_m \nabla^{\perp} \xi^m_j - \Omega^i_z P^z_j, \quad 1 \leq i, j \leq n. \tag{3-4}$$

Remark 3.3. Besides Uhlenbeck's method there is another way to construct the gauge transformation matrix P, namely, one can minimize the energy functional

$$E(R) = \int_{B_1} |R^T \nabla R + R^T \Omega R|^2 \tag{3-5}$$

among all $R \in W^{1,2}(B_1, \mathrm{SO}(n))$; see, for example, [Choné 1995; Schikorra 2010].

Another key result from Rivière's work is the following theorem, which was proved based on Theorem 3.1.

Theorem 3.4 [Rivière 2007, Theorem I.4]. *There exist $\varepsilon > 0$ and $C > 0$ such that, for every Ω in $L^2(B_1, \mathfrak{so}(n) \otimes \bigwedge^1 \mathbb{R}^2)$ satisfying*

$$\int_{B_1} |\Omega|^2 \leq \varepsilon,$$

there exist

$$\widehat{A} \in W^{1,2} \cap C^0(B_1, Gl_n(\mathbb{R})), \quad A = (\widehat{A} + \mathrm{Id}) P^T \in L^{\infty} \cap W^{1,2}(B_1, Gl_n(\mathbb{R})), \quad B \in W_0^{1,2}(B_1, M_n(\mathbb{R}))$$

such that

$$\nabla A - A\Omega = \nabla^{\perp} B \tag{3-6}$$

and

$$\|\widehat{A}\|_{W^{1,2}(B_1)} + \|\widehat{A}\|_{L^{\infty}(B_1)} + \|B\|_{W^{1,2}(B_1)} \leq C \|\Omega\|_{L^2(B_1)}. \tag{3-7}$$

Remark 3.5. Combining (3-6) with (3-1), one obtains the conservation law (in the distribution sense) for Rivière's equation, (3-1):

$$\operatorname{div}(A \nabla u + B \nabla^{\perp} u) = 0. \tag{3-8}$$

Equation (3-1), first considered in such generality in [Rivière 2007], generalizes a number of interesting equations appearing naturally in geometry, including the harmonic map equation (1-2), the H-surface equation, and, more generally, the Euler–Lagrange equation of any conformally invariant elliptic Lagrangian which is quadratic in the gradient. We remark that the harmonic map equation (1-2) can be written in the form of (3-1) if we set

$$\Omega := (\Omega^i_j)_{1\le i,j\le n}, \quad \text{where } \Omega^i_j := [\Pi^i(u)_{j,l} - \Pi^j(u)_{i,l}]\nabla u^l. \tag{3-9}$$

A central issue is the regularity of the weak solution u to this system of equations (3-1). Based on the conservation law (3-8), Rivière proved the (interior) continuity of any $W^{1,2}$ weak solution u to (3-1). This also resolved two conjectures by Heinz and Hildebrandt, respectively; see [Rivière 2007]. We point out that the harmonic map heat flow (1-8) on B_1 can be written in the form

$$u_t - \Delta u = \Omega \cdot \nabla u \quad \text{on } B_1 \times (0, T), \tag{3-10}$$

where Ω is as in (3-9).

The deep reason for Rivière's argument to work is that once the conservation law (3-8) is established, (3-1) can be rewritten in the form

$$\operatorname{div}(A\nabla u) = \nabla^\perp B \cdot \nabla u.$$

The right-hand side of this new equation lies in the Hardy space $\mathcal{H}^1$ by a result of Coifman, Lions, Meyer, and Semmes [Coifman et al. 1993]. Moreover, using a Hodge decomposition argument, one can show that u lies locally in $W^{2,1}$, which embeds into C^0 in two dimensions; cf. the proof of Theorem 3.7 below. The key to this fact is a special "compensation phenomena" for Jacobian determinants, first observed in [Wente 1969]. We will refer to the following lemma of Wente, for which an elementary proof can be found in [Brezis and Coron 1983; Hélein 2002, Theorem 3.1.2], and which will be the key ingredient of our proof.

Lemma 3.6 [Wente 1969]. *If $a, b \in W^{1,2}(B_1, \mathbb{R})$ and w is the solution of*

$$\begin{cases} \Delta w = \dfrac{\partial a}{\partial y}\dfrac{\partial b}{\partial x} - \dfrac{\partial a}{\partial x}\dfrac{\partial b}{\partial y} = \nabla a \cdot \nabla^\perp b & \text{in } B_1, \\ \quad w = 0 \ \text{ or } \ \dfrac{\partial w}{\partial \nu} = 0 & \text{on } \partial B_1, \end{cases} \tag{3-11}$$

then $w \in C^0 \cap W^{1,2}(B_1, \mathbb{R})$ and the estimate

$$\|w\|_{L^\infty(B_1)} + \|\nabla w\|_{L^2(B_1)} \le C\|\nabla a\|_{L^2(B_1)}\|\nabla b\|_{L^2(B_1)} \tag{3-12}$$

holds, where we choose $\int_{B_1} w = 0$ for the Neumann boundary data.

Now let $u(x,t) \in W^{1,2}(B_1 \times [0,\infty), \mathcal{N})$ be a global weak solution to the harmonic map heat flow (1-8) with $E(u_0) < \varepsilon_0$, $E(u(\cdot\,,t))$ nonincreasing in t, and $u(\cdot\,,t)|_{\partial B_1} = \chi$ for all $t \ge 0$ as in Theorem 1.6. First note that, for a.e. $t_0 \in (0,\infty)$, we have $u_t(t_0) \in L^2(B_1)$. Then, for any fixed t_0 such that $u_t(t_0) \in L^2(B_1)$, as in (3-9), we have

$$\Omega(t_0) = (\Omega^i_j(t_0))_{1\le i,j\le n}, \quad \text{where } \Omega^i_j(t_0) = [\Pi^i(u(t_0))_{j,l} - \Pi^j(u(t_0))_{i,l}]\nabla u^l(t_0).$$

We will express this by writing $\Omega(t_0) = \Pi(u(t_0))\nabla u(t_0)$. Moreover,

$$\int_{B_1} |\Omega(t_0)|^2 \le C E(u(t_0)) \le C\varepsilon_0. \tag{3-13}$$

Therefore Rivière's theorems on the existence of gauge (Theorems 3.1 and 3.4) apply to this time t_0 slice, and we find the existence of matrices $P(t_0) \in W^{1,2}(B_1, \mathrm{SO}(n))$,

$$A(t_0) = (\widehat{A}(t_0) + \mathrm{Id})P^T(t_0) \in L^\infty \cap W^{1,2}(B_1, Gl_n(\mathbb{R})),$$

and

$$B(t_0) \in W_0^{1,2}(B_1, M_n(\mathbb{R}))$$

such that

$$\nabla A(t_0) - A(t_0)\Omega(t_0) = \nabla^\perp B(t_0) \tag{3-14}$$

with the corresponding estimates (3-3) and (3-7).

Combining (3-14) with the harmonic map heat flow equation (3-10) yields (omitting the index t_0)

$$\begin{aligned} \operatorname{div}(A\nabla u + B\nabla^\perp u) &= \nabla A \cdot \nabla u + A\Delta u + \nabla B \cdot \nabla^\perp u \\ &= \nabla A \cdot \nabla u + A(-\Omega \cdot \nabla u + u_t) + \nabla B \cdot \nabla^\perp u \\ &= \nabla A \cdot \nabla u + (\nabla^\perp B - \nabla A) \cdot \nabla u + A u_t + \nabla B \cdot \nabla^\perp u \\ &= A u_t. \end{aligned} \tag{3-15}$$

We refer to (3-15) as an almost conservation law. By the results of [Coifman et al. 1993] and the standard L^p theory, (3-15) readily implies that $u(t_0) \in C^0(B_1, \mathbb{R}^n)$. In fact, we have the following ε-regularity theorem.

Theorem 3.7. *There exist $\varepsilon_0 > 0$ depending only on $\mathcal{N}$ such that if $u \in W^{1,2}(B_1 \times [0,\infty), \mathcal{N})$ is a global weak solution to the harmonic map heat flow* (1-8) *with $E(u_0) < \varepsilon_0$, $E(u(\cdot\,, t))$ nonincreasing in t, and $u(\cdot\,, t)|_{\partial B_1} = \chi$ for all $t \ge 0$, then $u \in C^\infty(B_1 \times [1,\infty), \mathcal{N})$.*

Proof. For any fixed t_0 such that $u_t(t_0) \in L^2(B_1)$, by Hodge decomposition (see, for example, [Iwaniec and Martin 2001, Corollary 10.5.1]), there exist $D(t_0), E(t_0) \in W^{1,2}(B_1, \mathbb{R}^n)$ such that (omitting the index t_0)

$$A\nabla u = \nabla D + \nabla^\perp E. \tag{3-16}$$

Note that (3-15) implies

$$\begin{cases} \operatorname{div}(A\nabla u) = -\nabla B \cdot \nabla^\perp u + A u_t, \\ \operatorname{curl}(A\nabla u) = \nabla^\perp A \cdot \nabla u. \end{cases} \tag{3-17}$$

Combining (3-16) and (3-17), we have

$$\begin{cases} \Delta D = -\nabla B \cdot \nabla^\perp u + A u_t, \\ \Delta E = \nabla^\perp A \cdot \nabla u. \end{cases} \tag{3-18}$$

Then, by the results of [Coifman et al. 1993] and via an extension argument, using the fact that $Au_t(t_0) \in L^2(B_1)$, we get $A\nabla u(t_0) \in W^{1,1}_{\mathrm{loc}}(B_1)$. Therefore $u(t_0) \in W^{2,1}_{\mathrm{loc}}(B_1)$, which embeds into $C^0(B_1)$.

Indeed, $A\nabla u(t_0) \in W^{1,1}_{\mathrm{loc}}(B_1)$ implies immediately that

$$\Omega(t_0) = \Pi(u(t_0))\nabla u(t_0) \in W^{1,1}_{\mathrm{loc}}(B_1). \tag{3-19}$$

Then, by [Rivière 2012, Theorem IV.4], the flow equation (3-10), (3-19), and the fact that $u_t(t_0) \in L^2(B_1)$ yield that $\nabla u(t_0) \in L^p_{\mathrm{loc}}(B_1)$ for some $p > 2$. Note that this is valid for a.e. $t_0 \in (0, \infty)$. Then, via a standard bootstrapping argument, we have $\nabla u \in L^q_{\mathrm{loc}}(B_1 \times [1, T])$ for all $q > 1$ and any $T > 1$ (see, for example, [Lieberman 1996]) and all higher order interior regularity follows. □

Again, we see that the "compensation phenomenon" enjoyed by the special Jacobian structure (see Lemma 3.6) has played an important role here, and these Wente-type estimates have many interesting applications, as in [Wente 1969; Brezis and Coron 1983; 1984; Tartar 1985; Coifman et al. 1993; Hélein 2002; Rivière 2007; 2008; 2011; Lamm and Lin 2013].

4. Improved estimates on the matrices B and P

Our main observation in this section is the existence of hidden Jacobian structures for ΔB and ΔP, valid only when Ω is of some special form: in our case, $\Omega = \Pi(u)\nabla u$. This will allow us to gain an improved global estimate for the matrix B and an improved local estimate for P. We start with the improved estimate for B.

Proposition 4.1. *Let $u(x, t)$ be as in Theorem 1.6. For any $t_0 \in [1, \infty)$, we have*

$$\|B(t_0)\|_{L^\infty(B_1)} \le C \int_{B_1} |\nabla u(t_0)|^2 \le C\varepsilon_0. \tag{4-1}$$

Proof. We recall that Ω is given by $\Pi(u)\nabla u$ as in (3-9) and therefore $\|\Omega(t_0)\|^2_{L^2(B_1)} \le C\varepsilon_0$ for all $t_0 \ge 1$. Now let ε_0 be so small that Theorems 3.1 and 3.4 apply. Taking the curl on both sides of (3-14) yields

$$\Delta B(t_0) = -\operatorname{curl}(A(t_0)\Pi(u(t_0))\nabla u(t_0)) = -\nabla u(t_0) \cdot \nabla^\perp(A(t_0)\Pi(u(t_0))). \tag{4-2}$$

Combining the Jacobian structure of the right-hand side of (4-2) with the zero boundary condition of B and estimates (3-3) and (3-7), Lemma 3.6 gives (4-1). Here we have also used $E(u(t)) < \varepsilon_0$ for all $t \ge 0$ and the smoothness and compactness of the target manifold $\mathcal{N}$. □

Next, as a step toward the improved local estimate on the matrix P, we show that ΔP also has a special Jacobian structure.

Lemma 4.2. *Let $u(x, t)$ be as in Theorem 1.6. For any $t_0 \in [1, \infty)$ such that $u_t(\,\cdot\,, t_0) \in L^2(B_1)$, there exist*

$$\xi(t_0) \in W^{1,2}_0(B_1, \mathfrak{so}(n)), \quad \eta(t_0) \in W^{1,2}(B_1, \mathbb{R}^n), \quad \zeta(t_0) \in W^{2,2}_0(B_1, \mathbb{R}^n)$$

and

$$Q_k(t_0), R_k(t_0) \in W^{1,2} \cap L^\infty(B_1, Gl_n(\mathbb{R})), \quad k = 1, \dots, n$$

with

$$\|\nabla\xi(t_0)\|_{L^2(B_1)} + \|\nabla\eta(t_0)\|_{L^2(B_1)} + \|\nabla\zeta(t_0)\|_{L^2(B_1)} + \sum_k \big(\|\nabla Q_k(t_0)\|_{L^2(B_1)} + \|\nabla R_k(t_0)\|_{L^2(B_1)}\big) \le C\sqrt{\varepsilon_0}$$

and

$$\|\zeta(t_0)\|_{W^{2,2}(B_1)} \le C\|u_t(t_0)\|_{L^2(B_1)} \tag{4-3}$$

such that

$$\Delta P(t_0) = \nabla P(t_0)\cdot\nabla^\perp\xi(t_0) + \nabla Q_k(t_0)\cdot\nabla^\perp\eta^k(t_0) + \nabla R_k(t_0)\cdot\nabla^\perp u^k(t_0) + \operatorname{div}(Q_k(t_0)\nabla\zeta^k(t_0)). \tag{4-4}$$

Proof. We omit the index t_0 in the proof. Hodge decomposition and the estimates for the L^∞-norms of A and B imply the existence of $\eta \in W^{1,2}(B_1, \mathbb{R}^n)$ and $\zeta \in W_0^{1,2}(B_1, \mathbb{R}^n)$ such that

$$\nabla^\perp\eta + \nabla\zeta = A\nabla u + B\nabla^\perp u \tag{4-5}$$

with

$$\|\nabla\eta\|_{L^2(B_1)} + \|\nabla\zeta\|_{L^2(B_1)} \le C\|\nabla u\|_{L^2(B_1)} \le C\sqrt{\varepsilon_0}. \tag{4-6}$$

Moreover, by the almost conservation law (3-15), we have

$$\Delta\zeta = Au_t \in L^2(B_1) \quad \text{and} \quad \zeta|_{\partial B_1} = 0,$$

which gives (4-3) by the standard L^p-theory. Multiplying both sides of (4-5) by A^{-1} from the left gives (with indices)

$$\nabla u^l = (A^{-1})^l_k\nabla^\perp\eta^k - (A^{-1}B)^l_k\nabla^\perp u^k + (A^{-1})^l_k\nabla\zeta^k, \quad l = 1, 2, \dots, n. \tag{4-7}$$

Taking the divergence on both sides of (3-4) yields

$$\Delta P^i_j = \nabla P^i_m\cdot\nabla^\perp\xi^m_j - \operatorname{div}(\Omega^i_z P^z_j), \quad 1 \le i, j \le n. \tag{4-8}$$

Since $\Omega^i_z = [\Pi^i(u)_{z,l} - \Pi^z(u)_{i,l}]\nabla u^l$, combining (4-7) and (4-8) gives

$$\begin{aligned}
\Delta P^i_j
&= \nabla P^i_m\cdot\nabla^\perp\xi^m_j - \operatorname{div}[(\Pi^i(u)_{z,l} - \Pi^z(u)_{i,l})((A^{-1})^l_k\nabla^\perp\eta^k - (A^{-1}B)^l_k\nabla^\perp u^k + (A^{-1})^l_k\nabla\zeta^k)P^z_j] \\
&= \nabla P^i_m\cdot\nabla^\perp\xi^m_j - \nabla[(\Pi^i(u)_{z,l} - \Pi^z(u)_{i,l})P^z_j(A^{-1})^l_k]\cdot\nabla^\perp\eta^k \\
&\quad + \nabla[(\Pi^i(u)_{z,l} - \Pi^z(u)_{i,l})P^z_j(A^{-1}B)^l_k]\cdot\nabla^\perp u^k - \operatorname{div}[((\Pi^i(u)_{z,l} - \Pi^z(u)_{i,l})P^z_j(A^{-1})^l_k)\nabla\zeta^k].
\end{aligned} \tag{4-9}$$

Defining

$$(Q_k)^i_j = -(\Pi^i(u)_{z,l} - \Pi^z(u)_{i,l})P^z_j(A^{-1})^l_k$$

and

$$(R_k)^i_j = (\Pi^i(u)_{z,l} - \Pi^z(u)_{i,l})P^z_j(A^{-1}B)^l_k,$$

where $1 \le k, i, j \le n$, completes the proof. □

Next we prove a local estimate on the oscillation of the matrix P based on Lemma 4.2. A key observation here is that whether a function is in the local Hardy space $h^1(B_1)$ essentially depends on its local behavior (see Definition 2.8). This local oscillation estimate on P provides important information that we need to control the local behavior of $|\nabla u|^2$. This point will become apparent in Section 5. As we shall see, the Jacobian structure of ΔP enters in a crucial way.

Lemma 4.3. *Let $u(x,t)$ be as in Theorem 1.6. For any $t_0 \in [1,\infty)$ such that $u_t(\,\cdot\,,t_0) \in L^2(B_1)$, any $x \in B_1$, any $r > 0$ such that $B_{2r}(x) \subset B_1$, and any $y \in B_r(x)$, we have*

$$|P(y,t_0) - P(x,t_0)| \le C(\sqrt{\varepsilon_0} + \|u_t(t_0)\|_{L^2(B_1)}). \tag{4-10}$$

Proof. We will omit the index t_0 in the proof. Let $\widetilde{P} \in W^{1,2}(B_1, M_n(\mathbb{R}))$ be the weak solution of

$$\begin{cases} \Delta \widetilde{P} = \nabla P \cdot \nabla^{\perp}\xi + \nabla Q_k \cdot \nabla^{\perp}\eta^k + \nabla R_k \cdot \nabla^{\perp} u^k + \operatorname{div}(Q_k \nabla \zeta^k) & \text{in } B_1, \\ \widetilde{P} = 0 & \text{on } \partial B_1, \end{cases}$$

where Q_k, R_k, η^k, and ζ^k are from Lemma 4.2.

Then, by Wente's lemma (Lemma 3.6) and the standard L^p-theory (and $W^{2,2}(B_1) \hookrightarrow C^0(B_1)$), we have $\widetilde{P} \in C^0(B_1, M_n(\mathbb{R}))$ and

$$\|\widetilde{P}\|_{L^\infty(B_1)} + \|\nabla \widetilde{P}\|_{L^2(B_1)} \le C(\varepsilon_0 + \|u_t(t_0)\|_{L^2(B_1)}). \tag{4-11}$$

Since

$$\Delta(P - \widetilde{P}) = 0 \quad \text{in } B_1,$$

we know that $V = P - \widetilde{P} \in C^\infty(B_1, M_n(\mathbb{R}))$ is harmonic. Now, for any $x \in B_1$, any $r > 0$ such that $B_{2r}(x) \subset B_1$, and any $y \in B_r(x)$, we have

$$\begin{aligned} |V(y) - V(x)| &\le Cr\|\nabla V\|_{L^\infty(B_r(x))} \le Cr^{-1}\|\nabla V\|_{L^1(B_{2r}(x))} \\ &\le C\|\nabla V\|_{L^2(B_{2r}(x))} \le C(\|\nabla P\|_{L^2(B_{2r}(x))} + \|\nabla \widetilde{P}\|_{L^2(B_{2r}(x))}) \\ &\le C(\sqrt{\varepsilon_0} + \|u_t(t_0)\|_{L^2(B_1)}), \end{aligned} \tag{4-12}$$

where we have used the mean value property of V and (4-11), (3-3). Combining (4-11) and (4-12) yields that, for any $x \in B_1$, any $r > 0$ such that $B_{2r}(x) \subset B_1$, and any $y \in B_r(x)$, we have

$$|P(y,t_0) - P(x,t_0)| \le C(\sqrt{\varepsilon_0} + \|u_t(t_0)\|_{L^2(B_1)}), \tag{4-13}$$

which gives the desired estimate (4-10). ☐

5. Validation of (2-21) and completion of the proof of Theorem 1.6

With the results so far at our disposal, we are now in a position to validate (2-21). As mention above, the local estimate on the oscillation of the transformation matrix P in Lemma 4.3 will be the key ingredient.

Lemma 5.1. *Let $u(x,t)$ be as in Theorem 1.6. For any $t_0 \in [1,\infty)$ such that $\|u_t(t_0)\|_{L^2(B_1)} < \sqrt{\varepsilon_0}$, we have*

$$|\nabla u(t_0)|^2 \in h^1(B_1) \tag{5-1}$$

with the estimate

$$\big\| |\nabla u(t_0)|^2 \big\|_{h^1(B_1)} \le C\varepsilon_0. \tag{5-2}$$

Remark 5.2. Lemma 5.1 continues to hold for the flow (3-10) with a more general Ω in the form $\Omega^i_j = \sum_{l=1}^n f^i_{jl}\nabla u^l + g^i_{jl}\nabla^\perp u^l$ (which includes $\Omega = \Omega^i_j = [\Pi^i(u)_{j,l} - \Pi^j(u)_{i,l}]\nabla u^l$ for the harmonic map heat flow as a special case); see [Lamm and Lin 2013]. Moreover, the condition $\|u_t(t_0)\|_{L^2(B_1)} < \varepsilon_0$ can be replaced by the fact that $\|u_t(t_0)\|_{L^p(B_1)}$ is sufficiently small for some $p > 1$.

Proof of Lemma 5.1. By the assumption $\|u_t(t_0)\|_{L^2(B_1)} < \sqrt{\varepsilon_0}$ and Lemma 4.3, for any $x \in B_1$, any $r > 0$ such that $B_{2r}(x) \subset B_1$, and any $y \in B_r(x)$, we have

$$|P(y, t_0) - P(x, t_0)| \le C\sqrt{\varepsilon_0}. \tag{5-3}$$

We will omit the index t_0 from now on. By Proposition 4.1 and Theorems 3.1 and 3.4, for any $x \in B_1$, any $r > 0$ such that $B_{2r}(x) \subset B_1$, and any $y \in B_r(x)$, we have (choosing ε_0 sufficiently small)

$$\begin{aligned} 0 &\le \tfrac{1}{2}|\nabla u|^2(y) \le (A\nabla u + B\nabla^\perp u)\cdot(P^T\nabla u)(y) \\ &= (A\nabla u + B\nabla^\perp u)\cdot[(P^T(x) + (P^T - P^T(x)))\nabla u](y), \end{aligned} \tag{5-4}$$

and therefore, by (4-5) and (5-3),

$$\begin{aligned} (\nabla^\perp\eta + \nabla\zeta)\cdot(P^T(x)\nabla u)(y) & \\ &= (A\nabla u + B\nabla^\perp u)\cdot(P^T(x)\nabla u)(y) \\ &\ge \tfrac{1}{2}|\nabla u|^2(y) - (A\nabla u + B\nabla^\perp u)\cdot[(P^T - P^T(x))\nabla u](y) \ge \tfrac{1}{4}|\nabla u|^2(y). \end{aligned} \tag{5-5}$$

Now we choose a function

$$\phi \in C_0^\infty(B_1) \text{ with } \phi \ge 0, \quad \operatorname{spt}(\phi) \subseteq B_{1/2}, \quad \phi = 2 \text{ on } B_{3/8}, \quad \text{and } \int_{B_1} \phi\, dx = 1. \tag{5-6}$$

Moreover, we additionally assume that $\|\nabla\phi\|_{L^\infty(B_1)} \le 100$. Using (4-3) and (5-5), one verifies directly that (by Definition 2.8)

$$\begin{aligned} \big\||\nabla u|^2\big\|_{h^1(B_1)} &= \int_{B_1} \sup_{0<t<1-|x|} \phi_t * |\nabla u|^2\, dx \\ &\le 4\int_{B_1} \sup_{0<t<1-|x|} \phi_t * ((\nabla^\perp\eta + \nabla\zeta)\cdot(P^T(x)\nabla u))\, dx \\ &= 4\int_{B_1} \sup_{0<t<1-|x|} \phi_t * [(P^T(x))_{ij}(\nabla^\perp\eta^i\cdot\nabla u^j + \nabla\zeta^i\cdot\nabla u^j)]\, dx \\ &\le C\sum_{i,j=1}^n (\|\nabla^\perp\eta^i\cdot\nabla u^j\|_{h^1(B_1)} + \|\nabla\zeta^i\|_{W^{1,2}(B_1)}\|\nabla u^j\|_{L^2(B_1)}) \\ &\le C\|\nabla^\perp\eta\|_{L^2(B_1)}\|\nabla u\|_{L^2(B_1)} + C\sqrt{\varepsilon_0}\|u_t\|_{L^2(B_1)} \le C\varepsilon_0, \end{aligned}$$

where we have used the relations

(1) $\nabla^\perp\eta^i\cdot\nabla u^j \in h^1(B_1)$ and $\|\nabla^\perp\eta^i\cdot\nabla u^j\|_{h^1(B_1)} \le C\|\nabla\eta\|_{L^2(B_1)}\|\nabla u\|_{L^2(B_1)}$ for all $i, j = 1, 2, \ldots, n$;

(2) $\|\nabla\zeta^i\cdot\nabla u^j\|_{L^p(B_1)} \le C\|\nabla\zeta^i\|_{W^{1,2}(B_1)}\|\nabla u^j\|_{L^2(B_1)}$ for any $1 < p < 2$ and $\|f\|_{h^1(B_1)} \le C\|f\|_{L^p(B_1)}$ for any $p > 1$.

To see (1), we first extend
$$\eta^i - \frac{1}{|B_1|}\int_{B_1}\eta^i \quad \text{and} \quad u^j - \frac{1}{|B_1|}\int_{B_1} u^j$$
from B_1 to $\mathbb{R}^2$, which yields the existence of $\tilde{\eta}^i, \tilde{u}^j \in W_c^{1,2}(\mathbb{R}^2)$ such that
$$\int_{\mathbb{R}^2} |\nabla\tilde{\eta}^i|^2 \le C\int_{B_1}|\nabla\eta^i|^2 \quad \text{and} \quad \int_{\mathbb{R}^2}|\nabla\tilde{u}^j|^2 \le C\int_{B_1}|\nabla u^j|^2 \tag{5-7}$$
and
$$\nabla\tilde{\eta}^i = \nabla\eta^i \quad \text{and} \quad \nabla\tilde{u}^j = \nabla u^j \quad \text{a.e. in } B_1. \tag{5-8}$$

Then, by the results of [Coifman et al. 1993], we know that
$$\begin{aligned}\|\nabla^{\perp}\tilde{\eta}^i\cdot\nabla\tilde{u}^j\|_{\mathscr{H}^1(\mathbb{R}^2)} &:= \int_{\mathbb{R}^2}\sup_{\phi\in\mathscr{T}}\sup_{t>0}\left|\int_{B_t(x)}\frac{1}{t^2}\phi\Big(\frac{x-y}{t}\Big)\Big(\nabla^{\perp}\tilde{\eta}^i\cdot\nabla\tilde{u}^j\Big)(y)\,dy\right|dx \\ &\le C\|\nabla\tilde{\eta}^i\|_{L^2(\mathbb{R}^2)}\|\nabla\tilde{u}^j\|_{L^2(\mathbb{R}^2)} \le C\|\nabla\eta\|_{L^2(B_1)}\|\nabla u\|_{L^2(B_1)},\end{aligned} \tag{5-9}$$
where $\mathscr{T} = \{\phi\in C^\infty(\mathbb{R}^2) : \operatorname{spt}(\phi)\subset B_1 \text{ and } \|\nabla\phi\|_{L^\infty}\le 100\}$. By (5-8), (5-9), and Definition 2.8, it is clear that
$$\begin{aligned}\|\nabla^{\perp}\eta^i\cdot\nabla u^j\|_{h^1(B_1)} &= \|\nabla^{\perp}\tilde{\eta}^i\cdot\nabla\tilde{u}^j\|_{h^1(B_1)} \\ &\le \|\nabla^{\perp}\tilde{\eta}^i\cdot\nabla\tilde{u}^j\|_{\mathscr{H}^1(\mathbb{R}^2)} \le C\|\nabla\eta\|_{L^2(B_1)}\|\nabla u\|_{L^2(B_1)}.\end{aligned} \tag{5-10}$$

This completes the proof of the lemma. □

Now, since $u(x,t)\in W^{1,2}\cap C^\infty(B_1\times[1,\infty),\mathcal{N})$ and the energy $E(u(\cdot,t))$ is nonincreasing along the flow as shown in (1-6), there exists $T_0\ge 1$ such that
$$\|u_t(T_0)\|_{L^2(B_1)} < \sqrt{\varepsilon_0}. \tag{5-11}$$
Then by Lemma 5.1 we know that $|\nabla u(T_0)|^2\in h^1(B_1)$ with estimate
$$\big\||\nabla u(T_0)|^2\big\|_{h^1(B_1)} \le C\varepsilon_0. \tag{5-12}$$

Therefore, in view of Lemma 5.1, in order to validate the global estimate (2-21) we are left to show
$$\|u_t(t_0)\|_{L^2(B_1)} < \sqrt{\varepsilon_0} \quad \text{for all } t_0\ge T_0. \tag{5-13}$$

We will next show this is indeed the case.

Lemma 5.3. *Let $u(x,t)$ be as in Theorem 1.6. Then there exists $T_0>0$ such that*
$$\|u_t(t_0)\|_{L^2(B_1)} < \sqrt{\varepsilon_0} \quad \textit{for all } t_0\ge T_0. \tag{5-14}$$

Proof. Let $T_0\ge 1$ be as in (5-11), so $\|u_t(T_0)\|_{L^2(B_1)}<\sqrt{\varepsilon_0}$. Since $u(x,t)\in W^{1,2}\cap C^\infty(B_1\times[1,\infty),\mathcal{N})$, and by the continuity of $\int_{B_1}|u_t(t)|^2$ in t, there exists $\delta=\delta(T_0,\varepsilon_0)>0$ such that, for any $t_0\in[T_0,T_0+\delta]$, we have
$$\|u_t(t_0)\|_{L^2(B_1)} < 2\sqrt{\varepsilon_0}. \tag{5-15}$$

Therefore, by our previous arguments (especially Theorem 2.6, Lemma 5.1 and (5-12) with T_0 replaced by t_0), Lemma 2.5 applies to any subinterval of $[T_0, T_0+\delta]$ and yields

$$\int_{B_1} |u_t(t_2)|^2 \le \int_{B_1} |u_t(t_1)|^2 \quad \text{for any } t_1, t_2 \text{ such that } T_0 \le t_1 < t_2 \le T_0+\delta. \tag{5-16}$$

This shows, instead of (5-15), for any $t_0 \in [T_0, T_0+\delta]$, we have

$$\|u_t(t_0)\|_{L^2(B_1)} \le \|u_t(T_0)\|_{L^2(B_1)} < \sqrt{\varepsilon_0}. \tag{5-17}$$

We can then continue and iterate this process beyond $T_0+\delta$ and we see that $\int_{B_1} |u_t(t)|^2$ is indeed nonincreasing along the flow after T_0. □

This completes the validation of (2-21) and therefore the assumptions in Lemmas 2.4 and 2.5 in view of Theorem 2.6, finishing the proof of our main Theorem 1.6 as shown in Section 2.

Appendix: A proof of Theorem 2.6

Proof. The idea of the proof follows [Semmes 1994, Proposition 1.68]. Since the Green's function of Δ on B_1 is given by $(1/(2\pi))\ln|x|$ for $x \in B_1$, we can write

$$\psi(x) = \frac{1}{2\pi}\int_{B_1} f(y)\left(\ln|x-y| - \ln\left(\left|\frac{x}{|x|} - |x|y\right|\right)\right)dy. \tag{A-1}$$

Let $\theta \in C_0^\infty(B_1)$ be a smooth bump function such that $0 \le \theta \le 1$, $\theta = 1$ in $B_{1/16}$ and $\operatorname{spt}(\theta) \subset B_{1/8}$. For $x \in B_1$, we define

$$l_x(y) := \sum_{j=0}^{\infty} \theta(2^j(1-|x|)^{-1}(x-y)) \quad \text{for } y \in B_1. \tag{A-2}$$

We claim that, for any $x, y \in B_1$,

$$-20\ln 2 \le \ln|x-y| - \ln\left(\left|\frac{x}{|x|} - |x|y\right|\right) + l_x(y)\ln 2 \le 20\ln 2. \tag{A-3}$$

To see this, it is clear that, for $x, y \in B_1$ such that

$$2^{-k} \le |x-y| \le 2^{-k+1}, \quad k \in \mathbb{N}_0, \tag{A-4}$$

we have

$$-k\ln 2 \le \ln|x-y| \le (-k+1)\ln 2. \tag{A-5}$$

Now note that

$$1-|x|-|x-y| \le 1-|x|+|x|-|y| = 1-|y| \le 1-|x|+|x-y|,$$

and therefore, for $x \in B_{1-2^{-i-1}} \setminus B_{1-2^{-i}}$, that is, $1-|x| \in [2^{-i-1}, 2^{-i}]$, $i \in \mathbb{N}_0$ (with $\bar{B}_0 = \varnothing$), and any $y \in B_1$ satisfying (A-4), we have

$$1-|y| \in \begin{cases} [2^{-i-1}-2^{-k+1}, 2^{-i}+2^{-k+1}] & \text{if } k \ge i+4, \\ [0, 2^{-i}+2^{-k+1}] & \text{if } k \le i+3. \end{cases}$$

We also have

$$\begin{aligned} 0 \le (1-|x|)(1-|y|) &\le (1-|x|^2)(1-|y|^2) \\ &= \left|\frac{x}{|x|}-|x|y\right|^2 - |x-y|^2 \le 2^2(1-|x|)(1-|y|), \end{aligned} \tag{A-6}$$

and thus

$$\left|\frac{x}{|x|}-|x|y\right|^2 - |x-y|^2 \in \begin{cases} [2^{-2i-2}-2^{-i-k},\, 2^{-2i+2}+2^{-i-k+3}] & \text{if } k \ge i+4, \\ [0,\, 2^{-2i+2}+2^{-i-k+3}] & \text{if } k \le i+3. \end{cases}$$

Combining this with (A-4), we get

$$\left|\frac{x}{|x|}-|x|y\right|^2 \in \begin{cases} [2^{-2i-2}-2^{-i-k}+2^{-2k},\, 2^{-2i+2}+2^{-i-k+3}+2^{-2k+2}] & \text{if } k \ge i+4, \\ [2^{-2k},\, 2^{-2i+2}+2^{-i-k+3}+2^{-2k+2}] & \text{if } k \le i+3. \end{cases}$$

Now, using the facts that for $k \ge i+4$ we have

$$2^{-2i-2}-2^{-i-k}+2^{-2k} \ge 2^{-2i-4} \quad \text{and} \quad 2^{-2i+2}+2^{-i-k+3}+2^{-2k+2} \le 2^{-2i+4}$$

and for $k \le i+3$ we have

$$2^{-2i+2}+2^{-i-k+3}+2^{-2k+2} \le 2^{-2k+10},$$

we arrive at

$$\left|\frac{x}{|x|}-|x|y\right|^2 \in \begin{cases} [2^{-2i-4},\, 2^{-2i+4}] & \text{if } k \ge i+4, \\ [2^{-2k},\, 2^{-2k+10}] & \text{if } k \le i+3, \end{cases}$$

and hence

$$-\ln\left|\frac{x}{|x|}-|x|y\right| \in \begin{cases} [(i-2)\ln 2,\, (i+2)\ln 2] & \text{if } k \ge i+4, \\ [(k-5)\ln 2,\, k\ln 2] & \text{if } k \le i+3. \end{cases} \tag{A-7}$$

Combining (A-5) and (A-7), we get

$$\ln|x-y| - \ln\left|\frac{x}{|x|}-|x|y\right| \in \begin{cases} [(-k+i-2)\ln 2,\, (-k+i+3)\ln 2] & \text{if } k \ge i+4, \\ [-5\ln 2,\, \ln 2] \quad (\text{in fact, } [-5\ln 2, 0]) & \text{if } k \le i+3 \end{cases} \tag{A-8}$$

for any $x \in B_{1-2^{-i-1}} \setminus B_{1-2^{-i}}$, $i \ge 0$, and any $y \in B_1$ satisfying (A-4) for some $k \ge 0$.

Now, for any $x \in B_{1-2^{-i-1}} \setminus B_{1-2^{-i}}$, $i \ge 0$, and any $y \in B_1$ satisfying (A-4), since $0 \le \theta \le 1$, $\theta = 1$ in $B_{1/16}$, and $\operatorname{spt}(\theta) \subset B_{1/8}$, we get that, for any $j \ge 0$,

$$\theta(2^j(1-|x|)^{-1}(x-y)) = 0 \quad \text{for } |x-y| \ge 2^{-j-3}(1-|x|) \in [2^{-j-i-4}, 2^{-j-i-3}]$$

and

$$\theta(2^j(1-|x|)^{-1}(x-y)) = 1 \quad \text{for } |x-y| \le 2^{-j-4}(1-|x|) \in [2^{-j-i-5}, 2^{-j-i-4}].$$

Combining with (A-4), we obtain

$$\theta(2^j(1-|x|)^{-1}(x-y)) = 0 \quad \text{for } j \ge k-i-3 \tag{A-9}$$

and

$$\theta(2^j(1-|x|)^{-1}(x-y)) = 1 \quad \text{if } k-1 \ge j+i+5 \quad (\text{that is } j \le k-i-6). \tag{A-10}$$

Hence, for any $x \in B_{1-2^{-i-1}} \setminus B_{1-2^{-i}}$, $i \geq 0$ and any $y \in B_1$ such that $2^{-k} \leq |x-y| \leq 2^{-k+1}$ for some $k = 0, 1, 2, \ldots$, (A-2), (A-9), and (A-10) imply

$$\begin{cases} k-i-10 \leq l_x(y) \leq k-i+10 & \text{if } k \geq i+4, \\ l_x(y) = 0 & \text{if } k \leq i+3. \end{cases} \tag{A-11}$$

Combining (A-8) and (A-11) gives (A-3).

Therefore, in order to obtain the L^∞-bound of ψ on B_1 as in (2-18), it suffices to bound $\int_{B_1} f(y) l_x(y)\, dy$, since we have (A-1), (A-3), and $\|f\|_{L^1(B_1)} \leq \|f\|_{h^1(B_1)}$.

In order to bound $\int_{B_1} f(y) l_x(y)\, dy$, we next claim that, for any $x \in B_1$, $j \geq 0$, and $z \in B_{2^{-j-4}(1-|x|)}(x)$, we have

$$\int_{B_1} f(y) 2^{2j+2}(1-|x|)^{-2} \theta(2^j(1-|x|)^{-1}(x-y))\, dy \leq \int_{B_t(z)} \frac{1}{t^2} \phi\left(\frac{z-y}{t}\right) f(y)\, dy, \tag{A-12}$$

where

$$t = 2^{-j-1}(1-|x|)$$

and ϕ is a nonnegative Schwartz function as in (5-6). To see (A-12), we first note that since $\operatorname{spt}(\theta) \subset B_{1/8}$, we have, for any $x \in B_1$ and $j \geq 0$,

$$\begin{aligned} &\int_{B_1} f(y) 2^{2j+2}(1-|x|)^{-2} \theta(2^j(1-|x|)^{-1}(x-y))\, dy \\ &\qquad = \int_{B_{2^{-j-3}(1-|x|)}(x)} f(y) 2^{2j+2}(1-|x|)^{-2} \theta(2^j(1-|x|)^{-1}(x-y))\, dy. \end{aligned} \tag{A-13}$$

Now since

$$\tfrac{3}{8}(2^{-j-1}) = 2^{-j-4} + 2^{-j-3}$$

and

$$2^{-j-4} + 2^{-j-1} = \tfrac{9}{16} 2^{-j} < 1$$

for any $j \geq 0$, we see that, for any $z \in B_{2^{-j-4}(1-|x|)}(x)$,

$$B_{2^{-j-3}(1-|x|)}(x) \subseteq B_{3t/8}(z) \subset B_t(z) = B_{2^{-j-1}(1-|x|)}(z) \subseteq B_1. \tag{A-14}$$

Using the relations $f \geq 0$, $0 \leq \theta \leq 1$, $\phi \geq 0$, and the fact that $\phi = 2$ on $B_{3/8}$, we conclude

$$\begin{aligned} &\int_{B_{2^{-j-3}(1-|x|)}(x)} f(y) 2^{2j+2}(1-|x|)^{-2} \theta(2^j(1-|x|)^{-1}(x-y))\, dy \\ &\qquad \leq \int_{B_{2^{-j-3}(1-|x|)}(x)} f(y) 2^{2j+2}(1-|x|)^{-2}\, dy \leq \int_{B_{3t/8}(z)} f(y) 2^{2j+2}(1-|x|)^{-2}\, dy \\ &\qquad \leq \int_{B_t(z)} f(y) 2^{2j+2}(1-|x|)^{-2} \phi\left(\frac{z-y}{t}\right) dy = \int_{B_t(z)} \frac{1}{t^2} \phi\left(\frac{z-y}{t}\right) f(y)\, dy. \end{aligned}$$

Combining this with (A-13) gives (A-12). Therefore, by (A-12) and the definition (2-19) of the radial maximal function f^*, for any $x \in B_1$ and $j \geq 0$, we have

$$\left|\int_{B_1} f(y)\theta(2^j(1-|x|)^{-1}(x-y))\,dy\right| \leq 2^{-2j-2}(1-|x|)^2 \inf_{z\in B_{2^{-j-4}(1-|x|)}(x)} f^*(z).$$

Therefore, by (A-2), for any $x \in B_1$, we have

$$\begin{aligned}\left|\int_{B_1} f(y)l_x(y)\,dy\right| &\leq \sum_{j=0}^{\infty}\left|\int_{B_1} f(y)\theta(2^j(1-|x|)^{-1}(x-y))\,dy\right| \\ &\leq \sum_{j=0}^{\infty} 2^{-2j-2}(1-|x|)^2 \inf_{z\in B_{2^{-j-4}(1-|x|)}(x)} f^*(z) \\ &\leq \frac{2^8}{3\pi}\sum_{j=0}^{\infty}\int_{B_{2^{-j-4}(1-|x|)}(x)\setminus B_{2^{-j-5}(1-|x|)}(x)} f^*(z)\,dz \\ &\leq \frac{2^8}{3\pi}\int_{B_1} f^*(z)\,dz \leq \frac{2^8}{3\pi}\|f\|_{h^1(B_1)}. \end{aligned} \tag{A-15}$$

Combining (A-6), (A-3), and (A-15) yields (using $\|f\|_{L^1(B_1)} \leq \|f\|_{h^1(B_1)}$)

$$|\psi(x)| = -\frac{1}{2\pi}\int_{B_1} f(y)\left(\ln|x-y| - \ln\left|\frac{x}{|x|} - |x|y\right|\right)dy \leq C\|f\|_{h^1(B_1)}.$$

This gives the desired L^∞-bound of ψ on B_1. The L^2-estimate for $\nabla\psi$ simply follows from an integration by parts argument. $\square$

Acknowledgement

The author thanks Bill Minicozzi for his continued guidance and support. The author also thanks his collaborator Tobias Lamm for all the valuable discussions.

References

[Brezis and Coron 1983] H. Brezis and J.-M. Coron, "Large solutions for harmonic maps in two dimensions", *Comm. Math. Phys.* **92**:2 (1983), 203–215. MR 85a:58022 Zbl 0532.58006

[Brezis and Coron 1984] H. Brezis and J.-M. Coron, "Multiple solutions of H-systems and Rellich's conjecture", *Comm. Pure Appl. Math.* **37**:2 (1984), 149–187. MR 85i:53010 Zbl 0537.49022

[Chang 1989] K.-C. Chang, "Heat flow and boundary value problem for harmonic maps", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **6**:5 (1989), 363–395. MR 90i:58037 Zbl 0687.58004

[Chang et al. 1993] D.-C. Chang, S. G. Krantz, and E. M. Stein, "H^p theory on a smooth domain in $\mathbb{R}^N$ and elliptic boundary value problems", *J. Funct. Anal.* **114**:2 (1993), 286–347. MR 94j:46032 Zbl 0804.35027

[Choné 1995] P. Choné, "A regularity result for critical points of conformally invariant functionals", *Potential Anal.* **4**:3 (1995), 269–296. MR 96e:58041 Zbl 0833.53010

[Coifman et al. 1993] R. Coifman, P.-L. Lions, Y. Meyer, and S. Semmes, "Compensated compactness and Hardy spaces", *J. Math. Pures Appl.* (9) **72**:3 (1993), 247–286. MR 95d:46033 Zbl 0864.42009

[Colding and Minicozzi 2008a] T. H. Colding and W. P. Minicozzi, II, "Width and finite extinction time of Ricci flow", *Geom. Topol.* **12**:5 (2008), 2537–2586. MR 2009k:53166 Zbl 1161.53352

[Colding and Minicozzi 2008b] T. H. Colding and W. P. Minicozzi, II, "Width and mean curvature flow", *Geom. Topol.* **12**:5 (2008), 2517–2535. MR 2009k:53165 Zbl 1165.53363

[Eells and Sampson 1964] J. Eells, Jr. and J. H. Sampson, "Harmonic mappings of Riemannian manifolds", *Amer. J. Math.* **86** (1964), 109–160. MR 29 #1603 Zbl 0122.40102

[Freire 1995] A. Freire, "Uniqueness for the harmonic map flow from surfaces to general targets", *Comment. Math. Helv.* **70**:2 (1995), 310–338. MR 96f:58045 Zbl 0831.58018

[Freire 1996] A. Freire, "Correction to: "Uniqueness for the harmonic map flow from surfaces to general targets" [Comment. Math. Helv. **70** (1995), no. 2, 310–338; MR1324632 (96f:58045)]", *Comment. Math. Helv.* **71**:2 (1996), 330–337. MR 97c:58032 Zbl 0851.58011

[Hamilton 1975] R. S. Hamilton, *Harmonic maps of manifolds with boundary*, Lecture Notes in Mathematics **471**, Springer, Berlin, 1975. MR 58 #2872 Zbl 0308.35003

[Hartman 1967] P. Hartman, "On homotopic harmonic maps", *Canad. J. Math.* **19** (1967), 673–687. MR 35 #4856 Zbl 0155.49705

[Hélein 2002] F. Hélein, *Harmonic maps, conservation laws and moving frames*, 2nd ed., Cambridge Tracts in Mathematics **150**, Cambridge University Press, 2002. MR 2003g:58024 Zbl 1010.58010

[Iwaniec and Martin 2001] T. Iwaniec and G. Martin, *Geometric function theory and non-linear analysis*, The Clarendon Press Oxford University Press, New York, 2001. MR 2003c:30001 Zbl 1045.30011

[Jost 1984] J. Jost, "The Dirichlet problem for harmonic maps from a surface with boundary onto a 2-sphere with nonconstant boundary values", *J. Differential Geom.* **19**:2 (1984), 393–401. MR 86b:58031 Zbl 0551.58012

[Lamm and Lin 2013] T. Lamm and L. Lin, "Estimates for the energy density of critical points of a class of conformally invariant variational problems", *Adv. Calc. Var.* **6**:4 (2013), 391–413. Zbl 06224215 arXiv 1202.5758

[Lieberman 1996] G. M. Lieberman, *Second order parabolic differential equations*, World Scientific, River Edge, NJ, 1996. MR 98k:35003 Zbl 0884.35001

[Miyachi 1990] A. Miyachi, "H^p spaces over open subsets of $\mathbf{R}^n$", *Studia Math.* **95**:3 (1990), 205–228. MR 91m:42022 Zbl 0716.42017

[Müller 1990] S. Müller, "Higher integrability of determinants and weak convergence in L^1", *J. Reine Angew. Math.* **412** (1990), 20–34. MR 92b:49026 Zbl 0713.49004

[Qing 1995] J. Qing, "On singularities of the heat flow for harmonic maps from surfaces into spheres", *Comm. Anal. Geom.* **3**:1-2 (1995), 297–315. MR 97c:58154 Zbl 0868.58021

[Rivière 2007] T. Rivière, "Conservation laws for conformally invariant variational problems", *Invent. Math.* **168**:1 (2007), 1–22. MR 2008d:58010 Zbl 1428.58010

[Rivière 2008] T. Rivière, "Analysis aspects of Willmore surfaces", *Invent. Math.* **174**:1 (2008), 1–45. MR 2009k:53154 Zbl 1155.53031

[Rivière 2011] T. Rivière, "The role of integrability by compensation in conformal geometric analysis", pp. 93–127 in *Analytic aspects of problems in Riemannian geometry: elliptic PDEs, solitons and computer imaging*, edited by P. Baird et al., Sémin. Congr. **22**, Soc. Math. France, Paris, 2011. MR 3060451

[Rivière 2012] T. Rivière, "Conformally invariant variational problems", preprint, 2012. arXiv 1206.2116

[Schikorra 2010] A. Schikorra, "A remark on gauge transformations and the moving frame method.", *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **27**:2 (2010), 503–515. Zbl 1187.35054

[Semmes 1994] S. Semmes, "A primer on Hardy spaces, and some remarks on a theorem of Evans and Müller", *Comm. Partial Differential Equations* **19**:1-2 (1994), 277–319. MR 94j:46038 Zbl 0836.35030

[Simon 1983] L. Simon, "Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems", *Ann. of Math.* (2) **118**:3 (1983), 525–571. MR 85b:58121 Zbl 0549.35071

[Struwe 1985] M. Struwe, "On the evolution of harmonic mappings of Riemannian surfaces", *Comment. Math. Helv.* **60**:4 (1985), 558–581. MR 87e:58056 Zbl 0595.58013

[Tartar 1985] L. Tartar, "Remarks on oscillations and Stokes' equation", pp. 24–31 in *Macroscopic modelling of turbulent flows* (Nice, 1984), edited by U. Frisch et al., Lecture Notes in Phys. **230**, Springer, Berlin, 1985. MR 815930 Zbl 0611.76042

[Topping 1997] P. M. Topping, "Rigidity in the harmonic map heat flow", *J. Differential Geom.* **45**:3 (1997), 593–610. MR 99d:58050 Zbl 0955.58013

[Uhlenbeck 1982] K. K. Uhlenbeck, "Connections with L^p bounds on curvature", *Comm. Math. Phys.* **83**:1 (1982), 31–42. MR 83e:53035 Zbl 0499.58019

[Wang 2012] L. Wang, "Harmonic map heat flow with rough boundary data", *Trans. Amer. Math. Soc.* **364**:10 (2012), 5265–5283. MR 2931329

[Wente 1969] H. C. Wente, "An existence theorem for surfaces of constant mean curvature", *J. Math. Anal. Appl.* **26** (1969), 318–344. MR 39 #4788 Zbl 0181.11501

Received 23 Jul 2012. Accepted 22 Aug 2013.

LONGZHI LIN: lzlin@math.rutgers.edu
Department of Mathematics, Rutgers University, 110 Frelinghuysen Road, Piscataway, New Jersey 08854-8019, United States

ANALYSIS AND PDE
Vol. 6, No. 8, 2013
dx.doi.org/10.2140/apde.2013.6.1923

A ROTATIONAL APPROACH TO TRIPLE POINT OBSTRUCTIONS

NOAH SNYDER

Subfactors where the initial branching point of the principal graph is 3-valent are subject to strong constraints called triple point obstructions. Since more complicated initial branches increase the index of the subfactor, triple point obstructions play a key role in the classification of small index subfactors. There are two strong triple point obstructions, called the triple-single obstruction and the quadratic tangles obstruction. Although these obstructions are very closely related, neither is strictly stronger. In this paper we give a more general triple point obstruction which subsumes both. The techniques are a mix of planar algebraic and connection-theoretic techniques with the key role played by the rotation operator.

1. Introduction

The principal graph of a subfactor begins with a type A string and then hits an initial branch point (unless the graph is A_k or A_∞). It is natural to stratify subfactors based on how complex this initial branch point is. Furthermore, complex initial branches increase the norm of the graph and thus the index of the subfactor. This means that small index subfactors can only have simple initial branches. The simplest possibility is an initial triple point (in this case the dual graph also begins with a triple point). Subfactors beginning with an initial triple point are subject to strong constraints known as triple point obstructions. For example, a triple point obstruction due to Ocneanu shows that as long as the index is greater than 4 the initial triple point must be at odd depth. These triple point obstructions play a crucial role in the classification of small index subfactors [Haagerup 1994; Morrison and Snyder 2012; Morrison et al. 2012; Izumi et al. 2012; Penneys and Tener 2012].

The current state of the art of triple point obstructions is given in our joint paper with S. Morrison, D. Penneys, and E. Peters [Morrison et al. 2012], but the status is somewhat unsatisfactory as there are two main results, neither of which is strictly stronger than the other. One result applies more generally and proves a certain inequality, while the other (due to V. F. R. Jones [2012]) has stricter assumptions but replaces the inequality with a finite list of values. The former is proved using connections and the latter using planar algebras. The main result of this paper is a mutual generalization of these two triple point obstructions, which proves the stronger conclusion using only the weaker assumptions. As one might expect, this paper uses a mix of connections and planar algebras following our earlier paper with M. Izumi, Jones, and Morrison [Izumi et al. 2012]. Furthermore, one can think of this argument as giving an alternate proof of the triple point obstruction from [Jones 2012].

MSC2010: 46L37.
Keywords: subfactors, planar algebras, connections.

Before stating the three relevant results, we fix some notation which we will use throughout the paper. Suppose that $N \subset M$ is an $n-1$ supertransitive finite index subfactor of index greater than 4 whose principal graphs begin with triple points, and let Γ and Γ' denote the principal and dual principal graphs. Let $[k]$ denote the quantum number $(\nu^k - \nu^{-k})/(\nu - \nu^{-1})$, where ν is a number such that the index is $[2]^2$. Let β and β' denote the initial triple points at depth $n-1$ (which is necessarily odd by Ocneanu's obstruction), let α_1 and γ_1 be the vertices at depth $n-2$, let α_2 and α_3 be the two vertices at depth n on Γ, and let γ_2 and γ_3 be the two vertices at depth n on Γ'. We will conflate vertices with the corresponding simple bimodules and the corresponding simple projections in the planar algebra. Assume without loss of generality that $\dim \alpha_2 \geq \dim \alpha_3$ and $\dim \gamma_2 \geq \dim \gamma_3$.

Theorem 1 (triple-single obstruction [Morrison et al. 2012, Theorem 3.5]). *If γ_3 is 1-valent, then*

$$\dim(\alpha_2) - \dim(\alpha_3) \leq 1.$$

Theorem 2 (quadratic tangles obstruction [Jones 2012]). *Suppose that γ_3 is 1-valent and that γ_2 is 3-valent; then*

$$r + \frac{1}{r} = \frac{\lambda + \lambda^{-1} + 2}{[n][n+2]} + 2,$$

where λ is the scalar by which rotation acts on the 1-dimensional perpendicular complement of Temperley–Lieb at depth n and $r = \dim(\alpha_2)/\dim(\alpha_3)$.

Since λ is a root of unity, we know that $-2 \leq \lambda + \lambda^{-1} \leq 2$. Hence the quadratic tangles obstruction gives an inequality, and (as observed by Zhengwei Liu) this inequality turns out to be precisely the one in the triple-single obstruction [Morrison et al. 2012, Lemma 3.3]. Thus the quadratic tangles obstruction is stronger (replacing an interval of possibilities with a finite list) when both apply, but the triple-single obstruction has a weaker assumption. The main result of this paper is the following mutual generalization of Theorems 1 and 2.

Theorem 3. *Suppose that γ_3 is 1-valent; then*

$$r + \frac{1}{r} = \frac{\lambda + \lambda^{-1} + 2}{[n][n+2]} + 2.$$

2. Background

We quickly summarize the key idea of [Izumi et al. 2012, §5.2], which is that the action of rotation on the planar algebra can be read off from the connection. Since rotational eigenvalues must be roots of unity, this gives highly nontrivial constraints on candidate connections. We assume that the reader is familiar with both planar algebras and connections; see [Izumi et al. 2012] for more detail.

Given a subfactor $N \subset M$ we get a certain collection of matrices called a connection. This connection depends on a choice of certain intertwiners, and thus is only well-defined up to gauge automorphisms. Let the branch matrix U denote the 3-by-3 matrix coming from the connection at the initial branch vertex of Γ. The key idea of Izumi et al. is that there is a canonical gauge choice for U, called the diagrammatic branch matrix, coming from the planar algebra. This choice is both easy to recognize and has nice properties, as captured by the following two results.

Lemma 4 [Izumi et al. 2012, Lemma 5.6]. *When n is odd the diagrammatic branch matrix is characterized within its gauge class by the property that all the entries in the first row and column are positive real numbers.*

Proposition 5. *Let U be the diagrammatic branch matrix for a subfactor with an initial triple point. Let x be an n-box in the perpendicular complement of Temperley–Lieb, and write $x = a_2(\alpha_2/\sqrt{\dim \alpha_2}) + a_3(\alpha_3/\sqrt{\dim \alpha_3})$. Let $(c_1, c_2, c_3) = U(0, a_2, a_3)$. Then $c_1 = 0$ and $c_2(\gamma_2/\sqrt{\dim \gamma_2}) + c_3(\gamma_3/\sqrt{\dim \gamma_3})$ is $\rho^{1/2}(x)$.*

Proof. This is a restatement of [Izumi et al. 2012, Corollary 5.3] in our special case. See [Izumi et al. 2012, pp. 18–19] for a worked example. □

In order to apply the previous proposition, we will want an explicit formula for vectors in the perpendicular complement to Temperley–Lieb in the n-box space and the action of rotation there. Recall that the rotation ρ preserves shading and thus is an endomorphism of each box space, while $\rho^{1/2}$ changes rotation and thus is a map from one box space to a different box space. We will use λ to denote the scalar by which ρ acts on the 1-dimensional perpendicular complement to Temperley–Lieb in the n-box space. Note that this is an n-th root of unity.

Lemma 6. *Let $r = \dim \alpha_2/\dim \alpha_3$ and $\check{r} = \dim \gamma_2/\dim \gamma_3$. Then $T = (1/\sqrt{r})\alpha_2 - \sqrt{r}\alpha_3$ and $\check{T} = (1/\sqrt{\check{r}})\gamma_2 - \sqrt{\check{r}}\gamma_3$ are each in the perpendicular complement of Temperley–Lieb.*

Furthermore $\rho^{1/2}(T) = \sqrt{\lambda}\check{T}$, where $\sqrt{\lambda}$ is some square root of the rotational eigenvalue for the action of rotation on the perpendicular compliment of Temperley–Lieb.

Proof. These calculations (with slightly different conventions) were done in an early version of [Jones 2012]. Seeing that T and $\check{T}$ are perpendicular to Temperley–Lieb is straightforward (you only need to work out their inner product with two specific Jones–Wenzl projections). Since half-click rotation preserves Temperley–Lieb and is an isometry, it also preserves the perpendicular complement of Temperley–Lieb. Thus $\rho^{1/2}(T)$ is some scalar multiple of $\check{T}$. To work out which scalar multiple this is you compute their norms. This tells you that the square of this scalar is λ. □

Remark 7. There are many square roots in this paper. Other than $\sqrt{\lambda}$, all square roots are positive square roots of positive numbers. Moreover $\sqrt{\lambda}$ will always be chosen such that the previous lemma works. In the final statement of the main theorem no $\sqrt{\lambda}$ appears, so this subtlety is not very important.

Combining the previous two results we have the following concrete statement, which will supply the main ingredient of our proof of Theorem 3.

Corollary 8. *The diagrammatic branch matrix U sends*

$$\left(0, \sqrt{\dim(\alpha_3)}, -\sqrt{\dim(\alpha_2)}\right) \mapsto \sqrt{\lambda}\left(0, \sqrt{\dim(\gamma_3)}, -\sqrt{\dim(\gamma_2)}\right).$$

3. Proof of Theorem 3

The idea of this argument is that having a 1-valent vertex allows us to solve for the branch matrix, and thus we can read off the rotational eigenvalue (since the diagrammatic branch matrix acts on the appropriate vectors by rotation). This gives an identity between the dimensions of objects and the rotational eigenvalue.

We begin with a quick calculation of the branch matrix following the proof of the triple-single obstruction [Morrison et al. 2012, Theorem 3.1]. Since α_1, γ_1, β, and β' are in the initial string their dimensions are $[n-1]$, $[n-1]$, $[n]$, and $[n]$, respectively. Since γ_3 is 1-valent, we have $\dim\gamma_2 = [n+2]/[2]$ and $\dim\gamma_3 = [n]/[2]$. Using the 1-valence of γ_3 the normalization condition on connections determines the magnitude of several of the entries in the branch matrix. Furthermore, unitarity of U allows us to work out several more of the entries. In particular, the branch matrix is gauge equivalent to the matrix below, where $p = \dim(\alpha_2)$ and $q = \dim(\alpha_3)$, where σ and τ are unknown phases, and where ? denotes unknown entries which will play no role in the calculation.

$$U = \begin{pmatrix} \dfrac{1}{[n]} & \dfrac{\sqrt{[n-1]p}}{[n]} & \dfrac{\sqrt{[n-1]q}}{[n]} \\ \sqrt{\dfrac{[n-1]}{[2][n]}} & \sigma\sqrt{\dfrac{p}{[2][n]}} & \tau\sqrt{\dfrac{q}{[2][n]}} \\ \sqrt{\dfrac{[n-1][n+2]}{[2][n]^2}} & ? & ? \end{pmatrix}.$$

The first row and column of this matrix are clearly positive, so by Lemma 4 we see that U is the diagrammatic branch matrix.

Remark 9. This matrix is the transpose of the matrix found in [Morrison et al. 2012] because the calculation there is done for Γ' instead of Γ. As shown in [Izumi et al. 2012], the diagrammatic branch matrices of Γ and Γ' are always transposes.

We would like to solve for σ and τ. Orthogonality of the first two rows of U tells us that

$$1 + \sigma p + \tau q = 0.$$

Although $1 + \sigma p + \tau q = 0$ is one equation in two unknowns, it actually determines σ and τ since they are phases:

$$\sigma = -\frac{1+\tau q}{p},$$

$$1 = \sigma\bar{\sigma} = \frac{1+\tau q}{p}\frac{1+\bar{\tau} q}{p} = \frac{1+(\tau+\bar{\tau})q+q^2}{p^2},$$

$$\tau + \bar{\tau} = \frac{p^2-q^2-1}{q}.$$

This determines the real part of τ, and thus τ itself. Similarly, $\sigma + \bar{\sigma} = (q^2 - p^2 - 1)/p$.

Now that we have a very explicit understanding of U we apply it to a rotational eigenvector. Corollary 8 tells us that U sends

$$(0, \sqrt{q}, -\sqrt{p}) \mapsto \sqrt{\lambda}\left(0, \sqrt{\frac{[n+2]}{[2]}}, -\sqrt{\frac{[n]}{[2]}}\right).$$

Looking at the middle coordinate of that identity, we see that

$$\sigma - \tau = \sqrt{\lambda}\sqrt{\frac{[n+2][n]}{pq}}.$$

Comparing the real parts of both sides yields

$$\begin{aligned}\left(\sqrt{\lambda}+\frac{1}{\sqrt{\lambda}}\right)\sqrt{\frac{[n+2][n]}{pq}} &= (\sigma+\bar{\sigma})-(\tau+\bar{\tau})\\ &= \frac{q^2-p^2-1}{p}-\frac{p^2-q^2-1}{q} = \frac{(q-p)((p+q)^2-1)}{pq}\\ &= \frac{(q-p)([n+1]^2-1)}{pq} = \frac{(q-p)([n][n+2])}{pq}.\end{aligned}$$

Squaring both sides and rearranging proves the theorem.

Remark 10. You might guess that $\sigma - \tau = \sqrt{\lambda}\sqrt{[n+2][n]/(pq)}$ would give a second condition coming from the imaginary parts. In fact there's no new information there, because the two sides automatically have the same norm.

Acknowledgements

The main ideas in this paper came out of joint work with Scott Morrison, and I would like to thank him for many helpful conversations. I would also like to thank Vaughan Jones, Dave Penneys, and Emily Peters. This work was supported by an NSF Postdoctoral Fellowship at Columbia University and DARPA grants HR0011-11-1-0001 and HR0011-12-1-0009.

References

[Haagerup 1994] U. Haagerup, "Principal graphs of subfactors in the index range $4 < [M:N] < 3+\sqrt{2}$", pp. 1–38 in *Subfactors* (Kyuzeso, 1993), edited by H. Araki et al., World Scientific, River Edge, NJ, 1994. MR 96d:46081 Zbl 0933.46058

[Izumi et al. 2012] M. Izumi, V. F. R. Jones, S. Morrison, and N. Snyder, "Subfactors of index less than 5, Part 3: Quadruple points", *Comm. Math. Phys.* **316**:2 (2012), 531–554. MR 2993924 Zbl 1272.46051 arXiv 1109.3190

[Jones 2012] V. F. R. Jones, "Quadratic tangles in planar algebras", *Duke Math. J.* **161**:12 (2012), 2257–2295. MR 2972458 Zbl 1257.46033 arXiv 1007.1158

[Morrison and Snyder 2012] S. Morrison and N. Snyder, "Subfactors of index less than 5, part 1: The principal graph odometer", *Comm. Math. Phys.* **312**:1 (2012), 1–35. MR 2914056 Zbl 1246.46055 arXiv 1007.1730

[Morrison et al. 2012] S. Morrison, D. Penneys, E. Peters, and N. Snyder, "Subfactors of index less than 5, Part 2: Triple points", *Internat. J. Math.* **23**:3 (2012), 1250016, 33. MR 2902285 Zbl 1246.46054 arXiv 1007.2240

[Penneys and Tener 2012] D. Penneys and J. E. Tener, "Subfactors of index less than 5, part 4: Vines", *Internat. J. Math.* **23**:3 (2012), 1250017, 18. MR 2902286 Zbl 1246.46056

Received 3 Oct 2012. Revised 22 Jan 2013. Accepted 8 Mar 2013.

NOAH SNYDER: nsnyder@math.indiana.edu
Mathematics Department, Indiana University, 831 E. Third St., Bloomington, IN 47401, United States

ANALYSIS AND PDE
Vol. 6, No. 8, 2013
dx.doi.org/10.2140/apde.2013.6.1929

ON THE ENERGY SUBCRITICAL, NONLINEAR WAVE EQUATION IN $\mathbb{R}^3$ WITH RADIAL DATA

RUIPENG SHEN

In this paper, we consider the wave equation in 3-dimensional space with an energy-subcritical nonlinearity, either in the focusing or defocusing case. We show that any radial solution of the equation which is bounded in the critical Sobolev space is globally defined in time and scatters. The proof depends on the compactness/rigidity argument, decay estimates for radial, "compact" solutions, gain of regularity arguments and the "channel of energy" method.

1. Introduction

In this paper we will consider the energy subcritical, nonlinear wave equation in $\mathbb{R}^3$ with radial initial data

$$\begin{cases} \partial_t^2 u - \Delta u = \pm|u|^{p-1}u, & (x,t) \in \mathbb{R}^3 \times \mathbb{R}, \\ u|_{t=0} = u_0 \in \dot{H}^{s_p}(\mathbb{R}^3), \\ \partial_t u|_{t=0} = u_1 \in \dot{H}^{s_p-1}(\mathbb{R}^3). \end{cases} \tag{1}$$

Here $3 < p < 5$ and

$$s_p = \frac{3}{2} - \frac{2}{p-1}.$$

The positive sign in the nonlinear term gives us the focusing case, while the negative sign indicates the defocusing case. The quantity

$$E(t) = \frac{1}{2}\int_{\mathbb{R}^3} (|\partial_t u(x,t)|^2 + |\nabla u(x,t)|^2)\,dx \mp \frac{1}{p+1}\int_{\mathbb{R}^3} |u(x,t)|^{p+1}\,dx \tag{2}$$

is called the energy of the solution. The energy is a constant in the whole lifespan of the solution, as long as it is well-defined. Note that the energy can be a negative number in the focusing case.

Previous results in the energy-critical case. In the energy-critical case, namely $p = 5$, the initial data is in the energy space $\dot{H}^1 \times L^2$. This automatically guarantees the existence of the energy by the Sobolev embedding $\dot{H}^1 \hookrightarrow L^6$. This kind of wave equations has been extensively studied. In the defocusing case, M. Grillakis [1990; 1992] proved the global existence and scattering of the solution with any $\dot{H}^1 \times L^2$ initial data. In the focusing case, however, the behavior of solutions is much more complicated. The solutions may scatter, blow up in finite time or even be independent of time. (See [Duyckaerts et al. 2013; Kenig and Merle 2008] for more details.) In particular, a solution independent of time is usually

MSC2010: 35L15, 35L71.
Keywords: wave equation, scattering, nonlinear, energy subcritical.

called a ground state or a soliton. This kind of solutions is actually the solutions of the elliptic equation $-\Delta W(x) = |W(x)|^{p-1}W(x)$. We can write down all the nontrivial radial solitons explicitly as

$$W(x) = \pm\frac{1}{\lambda^{1/2}}\left(1+\frac{|x|^2}{3\lambda^2}\right)^{-\frac{1}{2}}. \tag{3}$$

Here λ is an arbitrary positive parameter.

Energy subcritical case. We will consider the case $3 < p < 5$ in this paper; thus $1/2 < s_p < 1$. In this case the problem is critical in the space $\dot{H}^{s_p}(\mathbb{R}^3)\times\dot{H}^{s_p-1}(\mathbb{R}^3)$, because if $u(x,t)$ is a solution of (1) with initial data (u_0,u_1), then for any $\lambda > 0$, the function

$$\frac{1}{\lambda^{3/2-s_p}}u\left(\frac{x}{\lambda},\frac{t}{\lambda}\right)$$

is another solution of the (1) with the initial data

$$\left(\frac{1}{\lambda^{3/2-s_p}}u_0\left(\frac{x}{\lambda}\right),\frac{1}{\lambda^{5/2-s_p}}u_1\left(\frac{x}{\lambda}\right)\right),$$

which shares the same $\dot{H}^{s_p}\times\dot{H}^{s_p-1}$ norm as the original initial data (u_0,u_1). These scalings play an important role in our discussion of this problem.

Theorem 1.1 (main theorem). *Let u be a solution of the nonlinear wave equation* (1) *with radial initial data* $(u_0,u_1)\in\dot{H}^{s_p}\times\dot{H}^{s_p-1}(\mathbb{R}^3)$ *and a maximal lifespan I so that*

$$\sup_{t\in I}\|(u(t),\partial_t u(t))\|_{\dot{H}^{s_p}\times\dot{H}^{s_p-1}} < \infty. \tag{4}$$

Then u is global in time ($I=\mathbb{R}$) *and scatters*; *that is*,

$$\|u(x,t)\|_{S(\mathbb{R})} < \infty, \quad \text{or equivalently} \quad \|u(x,t)\|_{Y_{s_p}(\mathbb{R})} < \infty.$$

This is actually equivalent to saying that there exist two pairs (u_0^+,u_1^+) *and* (u_0^-,u_1^-) *in the space* $\dot{H}^{s_p}\times\dot{H}^{s_p-1}$ *such that*

$$\lim_{t\to\pm\infty}\left\|\left(u(t)-S(t)(u_0^\pm,u_1^\pm),\partial_t u(t)-\partial_t S(t)(u_0^\pm,u_1^\pm)\right)\right\|_{\dot{H}^{s_p}\times\dot{H}^{s_p-1}} = 0.$$

Here $S(t)(u_0^\pm,u_1^\pm)$ *is the solution of the linear wave equation with the initial data* $(u_0^\pm,u_1^\pm)$.

Please refer to Definition 2.4 for the S and Y_s norms. There are a couple of remarks on the main theorem.

- *The defocusing case.* As in the energy-critical case, we expect that the solutions always scatter as long as the initial data are in the critical Sobolev space. Besides the radial condition, the main theorem depends on the assumption (4), which is expected to be true for all solutions. Unfortunately, as far as the author knows, no one actually knows how to prove it without additional assumptions.

- *The focusing case.* In the focusing case, the solutions may blow up in finite time. (See Theorem 6.3, for instance.) Thus the assumption (4) is a meaningful and essential condition rather than a technical one. The main theorem gives us the following rough classification of radial solutions.

Proposition 1.2. *Let $u(t)$ be a solution of* (1) *in the focusing case with a maximal lifespan I and radial initial data $(u_0, u_1) \in \dot{H}^{s_p} \times \dot{H}^{s_p-1}(\mathbb{R}^3)$. Then one of the following holds for $u(x,t)$.*

(I) (*blow-up*) *The $\dot{H}^{s_p} \times \dot{H}^{s_p-1}$ norm of $(u(t), \partial_t u(t))$ blows up, namely*

$$\sup_{t\in I} \|(u(t), \partial_t u(t))\|_{\dot{H}^{s_p}\times\dot{H}^{s_p-1}} = +\infty.$$

(II) (*scattering*) *If the upper bound of the $\dot{H}^{s_p} \times \dot{H}^{s_p-1}$ norm above is finite instead, namely, the assumption* (4) *holds, then $u(t)$ is a global solution* (*i.e*, $I = \mathbb{R}$) *and scatters.*

Main idea in this paper. The main idea to establish Theorem 1.1 is to use the compactness/rigidity argument, namely to show:

(I) If the main theorem failed, it would break down at a minimal blow-up solution, which is almost periodic modulo scalings.

(II) The minimal blow-up solution is in the energy space.

(III) The minimal blow-up solution described above does not exist.

Step (I). The method of profile decomposition used here has been a standard way to deal with both the wave equation and the Schrödinger equation. Thus we will only give important statements instead of showing all the details. The other steps, however, depend on the specific problems. One could refer to [Bahouri and Gérard 1999] in order to understand what the profile decomposition is, and to [Kenig and Merle 2008; 2010; Killip and Visan 2010] in order to see why the profile decomposition leads to the existence of a minimal blow-up solution.

Step (II). We will combine the method used in my old paper [Shen 2011] and a method used in [Kenig and Merle 2011] on the supercritical case of the nonlinear wave equation in $\mathbb{R}^3$. The idea is to use the following fact. Given a radial solution $u(x,t)$ of the equation

$$\partial_t^2 u(x,t) - \Delta u(x,t) = F(x,t)$$

defined in the time interval I, if we define two functions $w, h : \mathbb{R}^+ \times I \to \mathbb{R}$, such that $w(|x|, t) = |x| u(x,t)$ and $h(|x|, t) = |x| F(x,t)$, then $w(r,t)$ is a solution of the one-dimensional wave equation $\partial_t^2 w(r,t) - \partial_r^2 w(r,t) = h(r,t)$. This makes it convenient to consider the integral

$$\int_{r_0\pm t}^{4r_0\pm t} |\partial_t w(r, t_0+t) \mp \partial_r w(r, t_0+t)|^2 \, dr.$$

as the parameter t changes.

Step (III). Given an energy estimate, all minimal blow-up solutions are not difficult to kill except for the soliton-like solutions in the focusing case. As I mentioned earlier, this kind of solutions actually exists in the energy-critical case. The ground states given in (3) are perfect examples. In the energy-subcritical case, however, the soliton does not exist at all. More precisely, none of the solutions of the corresponding elliptic equation is in the right space $\dot{H}^{s_p}$. This fact enables us to gain a contradiction by showing a soliton-like minimal blow-up solution must be a real soliton, which does not exist, using a new method introduced by T. Duyckaerts, C. E. Kenig and F. Merle. They classified all radial solutions of the energy-critical, focusing wave equation in their recent paper [Duyckaerts et al. 2013] using this "channel of energy" method.

Remark on the supercritical case. Simultaneously to this work, T. Duyckaerts et al. [2012] proved that results similar to ours also hold in the supercritical case $p > 5$ of the focusing wave equation, using the compactness/rigidity argument, a point-wise estimate on "compact" solutions obtained in [Kenig and Merle 2011] and the channel of energy method mentioned above.

2. Preliminary results

Notation. The following notation will be used throughout this paper.

- ($\lesssim$) The inequality $A \lesssim B$ means that there exists a constant c such that $A \le cB$. A subscript on $\lesssim$ implies that the constant c depends on the parameter(s) indicated but nothing else.
- (the smooth frequency cutoff) We use $P_{<A}$ and $P_{>A}$ for the standard smooth frequency cutoff operators. In particular, we use the following notation on u for convenience:

$$u_{<A} \doteq P_{<A}u, \quad u_{>A} \doteq P_{>A}u.$$

- (notation for radial functions) If $u(x,t)$ is radial in the space, then $u(r,t)$ represents the value $u(x,t)$ when $|x| = r$.
- (linear wave evolution) Let $(u_0, u_1) \in \dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^3)$ be a pair of initial data. Suppose $u(x,t)$ is the solution of the linear wave equation

$$\begin{cases} \partial_t^2 u - \Delta u = 0, & (x,t) \in \mathbb{R}^3 \times \mathbb{R}, \\ u|_{t=0} = u_0, \\ \partial_t u|_{t=0} = u_1. \end{cases}$$

  We will use the following notation to represent this solution u:

$$S(t_0)(u_0,u_1) = u(t_0), \quad \mathbf{S}(t_0)\begin{pmatrix} u_0 \\ u_1 \end{pmatrix} = \begin{pmatrix} u(t_0) \\ \partial_t u(t_0) \end{pmatrix}.$$

- (method of center cutoff) Let $(v_0, v_1) \in \dot{H}^1 \times L^2(\mathbb{R}^3 \setminus B(0,r))$ be a pair of radial functions. We define (with $R > r$)

$$(\Psi_R v_0)(x) = \begin{cases} v_0(x) & \text{if } |x| > R, \\ v_0(R) & \text{if } |x| \le R, \end{cases}$$

$$(\Psi_R v_1)(x) = \begin{cases} v_1(x) & \text{if } |x| > R, \\ 0 & \text{if } |x| \le R. \end{cases}$$

Local theory with $\dot{H}^{s_p} \times \dot{H}^{s_p-1}(\mathbb{R}^3)$ initial data. In this section, we will review the theory for the Cauchy problem of the nonlinear wave equation (1) with initial data in the critical Sobolev space $\dot{H}^{s_p} \times \dot{H}^{s_p-1}(\mathbb{R}^3)$. The same local theory works in both the focusing and defocusing cases. It can be also applied to the nonradial case.

Definition 2.1 (space-time norm). Let I be an interval of time. If $1 \le q, r < \infty$, the space-time norm is defined by

$$\|v(x,t)\|_{L^q L^r(I\times\mathbb{R}^3)} = \left(\int_I \left(\int_{\mathbb{R}^3} |v(x,t)|^r \, dx\right)^{q/r} dt\right)^{1/q},$$

$$\|v(x,t)\|_{L^\infty L^r(I\times\mathbb{R}^3)} = \inf\left\{ M > 0 : \left(\int_{\mathbb{R}^3} |v(x,t)|^r \, dx\right)^{1/r} < M, \text{ a.e. } t \in I \right\}.$$

This is used in the following Strichartz estimates.

Proposition 2.2 (generalized Strichartz inequalities; see Proposition 3.1 of [Ginibre and Velo 1995] — here we use the Sobolev version in $\mathbb{R}^3$). *Let $2 \le q_1, q_2 \le \infty$, $2 \le r_1, r_2 < \infty$ and $\rho_1, \rho_2, s \in \mathbb{R}$ with*

$$\begin{aligned} &1/q_i + 1/r_i \le 1/2 \quad \textit{for } i = 1, 2,\\ &1/q_1 + 3/r_1 = 3/2 - s + \rho_1,\\ &1/q_2 + 3/r_2 = 1/2 + s + \rho_2. \end{aligned}$$

Let u be the solution of the linear wave equation

$$\begin{cases} \partial_t^2 u - \Delta u = F(x,t), & (x,t) \in \mathbb{R}^3 \times \mathbb{R},\\ u|_{t=0} = u_0 \in \dot{H}^s(\mathbb{R}^3), \\ \partial_t u|_{t=0} = u_1 \in \dot{H}^{s-1}(\mathbb{R}^3). \end{cases} \tag{5}$$

Then we have

$$\|(u(T), \partial_t u(T))\|_{\dot{H}^s \times \dot{H}^{s-1}} + \|D_x^{\rho_1} u\|_{L^{q_1} L^{r_1}([0,T]\times\mathbb{R}^3)} \le C\big(\|(u_0,u_1)\|_{\dot{H}^s\times\dot{H}^{s-1}} + \|D_x^{-\rho_2} F(x,t)\|_{L^{\bar{q}_2} L^{\bar{r}_2}([0,T]\times\mathbb{R}^3)}\big).$$

The constant C does not depend on T.

Definition 2.3 (admissible pair). If $(q_1, r_1, s, \rho_1) = (q, r, m, 0)$ satisfies the conditions in Proposition 2.2, we say (q, r) is an m-admissible pair.

Definition 2.4. Fix $3 < p < 5$. We define the following norms with $s_p \le s \le 1$:

$$\begin{aligned} \|v(x,t)\|_{S(I)} &= \|v(x,t)\|_{L^{2(p-1)} L^{2(p-1)}(I\times\mathbb{R}^3)},\\ \|v(x,t)\|_{W(I)} &= \|v(x,t)\|_{L^4 L^4(I\times\mathbb{R}^3)},\\ \|v(x,t)\|_{Z_s(I)} &= \|v(x,t)\|_{L^{\frac{2}{s+1}} L^{\frac{2}{2-s}}(I\times\mathbb{R}^3)},\\ \|v(x,t)\|_{Y_s(I)} &= \|v(x,t)\|_{L^{\frac{2p}{s+1-(2p-2)(s-s_p)}} L^{\frac{2p}{2-s}}(I\times\mathbb{R}^3)}. \end{aligned}$$

Remark 2.5. By the Strichartz estimates, we have if $u(x,t)$ is the solution of

$$\begin{cases} \partial_t^2 u - \Delta u = F(x,t), & (x,t) \in \mathbb{R}^3 \times \mathbb{R}, \\ u|_{t=0} = u_0 \in \dot{H}^s(\mathbb{R}^3), \\ \partial_t u|_{t=0} = u_1 \in \dot{H}^{s-1}(\mathbb{R}^3). \end{cases}$$

then

$$\|(u(T), \partial_t u(T))\|_{\dot{H}^s \times \dot{H}^{s-1}} + \|u\|_{Y_s([0,T])} \le C(\|(u_0, u_1)\|_{\dot{H}^s \times \dot{H}^{s-1}} + \|F(x,t)\|_{Z_s([0,T])}).$$

Definition 2.6 (solutions). We say $u(t)(t \in I)$ is a solution of (1), if $(u(t), \partial_t u(t)) \in C(I; \dot{H}^{s_p} \times \dot{H}^{s_p - 1})$, with finite norms $\|u\|_{S(J)}$ and $\|D_x^{s_p - 1/2} u\|_{W(J)}$ for any bounded closed interval $J \subseteq I$ so that the integral equation

$$u(t) = S(t)(u_0, u_1) + \int_0^t \frac{\sin((t-\tau)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\, d\tau$$

holds for all time $t \in I$. Here $S(t)(u_0, u_1)$ is the solution of the linear wave equation with initial data (u_0, u_1) and

$$F(u) = \pm |u|^{p-1} u.$$

Remark 2.7. We can take another way to define the solutions by substituting $S(I)$ and $W(I)$ norms by a single $Y_{s_p}(I)$ norm. Using the Strichartz estimates, these two definitions are equivalent to each other.

By the Strichartz estimate and a fixed-point argument, we have the following theorems. (Our argument is similar to those in a lot of earlier papers. See, for instance, [Lindblad and Sogge 1995; Kenig and Merle 2008] for more details.)

Theorem 2.8 (local solution). *For any initial data $(u_0, u_1) \in \dot{H}^{s_p} \times \dot{H}^{s_p - 1}$, there is a maximal interval $(-T_-(u_0, u_1), T_+(u_0, u_1))$ in which the equation has a solution.*

Theorem 2.9 (scattering with small data). *There exists $\delta = \delta(p) > 0$ such that if the norm of the initial data $\|(u_0, u_1)\|_{\dot{H}^{s_p} \times \dot{H}^{s_p - 1}} < \delta$, then the Cauchy problem* (1) *has a global-in-time solution u with $\|u\|_{S(-\infty, +\infty)} < \infty$.*

Lemma 2.10 (standard finite blow-up criterion). *If $T_+ < \infty$, then $\|u\|_{S([0,T_+))} = \infty$.*

Theorem 2.11 (long-time perturbation theory; see [Colliander et al. 2008; Kenig and Merle 2008; 2006; 2011]). *Fix $3 < p < 5$. Let M, A, A' be positive constants. There exists $\varepsilon_0 = \varepsilon_0(M, A, A') > 0$ and $\beta > 0$ such that if $\varepsilon < \varepsilon_0$, then for any approximation solution $\tilde{u}$ defined on $\mathbb{R}^3 \times I$ $(0 \in I)$ and any initial data $(u_0, u_1) \in \dot{H}^{s_p} \times \dot{H}^{s_p - 1}$ satisfying*

$$(\partial_t^2 - \Delta)(\tilde{u}) - F(\tilde{u}) = e(x,t), \quad (x,t) \in \mathbb{R}^3 \times I,$$

$$\begin{cases} \sup_{t \in I} \|(\tilde{u}(t), \partial_t \tilde{u}(t))\|_{\dot{H}^{s_p} \times \dot{H}^{s_p - 1}} \le A, \\ \|\tilde{u}\|_{S(I)} \le M, \\ \|D_x^{s_p - 1/2} \tilde{u}\|_{W(J)} < \infty \quad \text{for each } J \Subset I, \end{cases} \tag{6}$$

$$\|(u_0 - \tilde{u}(0), u_1 - \partial_t \tilde{u}(0))\|_{\dot{H}^{s_p} \times \dot{H}^{s_p - 1}} \le A',$$

$$\|D_x^{s_p - \frac{1}{2}} e\|_{L_I^{4/3} L_x^{4/3}} + \|S(t)(u_0 - \tilde{u}(0), u_1 - \partial_t \tilde{u}(0))\|_{S(I)} \le \varepsilon,$$

there exists a solution of (1) *defined in the interval I with the initial data (u_0, u_1) and satisfying*

$$\|u\|_{S(I)} \le C(M, A, A'),$$
$$\sup_{t\in I} \|(u(t), \partial_t u(t)) - (\tilde{u}(t), \partial_t \tilde{u}(t))\|_{\dot{H}^{s_p}\times\dot{H}^{s_p-1}} \le C(M, A, A')(A' + \varepsilon + \varepsilon^{\beta}).$$

Theorem 2.12 (perturbation theory with Y_{s_p} norm). *Fix $3 < p < 5$. Let M be a positive constant. There exists a constant $\varepsilon_0 = \varepsilon_0(M) > 0$ such that if $\varepsilon < \varepsilon_0$, then for any approximation solution $\tilde{u}$ defined on $\mathbb{R}^3 \times I$ $(0 \in I)$ and any initial data $(u_0, u_1) \in \dot{H}^{s_p} \times \dot{H}^{s_p-1}$ satisfying*

$$(\partial_t^2 - \Delta)(\tilde{u}) - F(\tilde{u}) = e(x,t), \quad (x,t) \in \mathbb{R}^3 \times I,$$
$$\|\tilde{u}\|_{Y_{s_p}(I)} < M, \quad \|(\tilde{u}(0), \partial_t \tilde{u}(0))\|_{\dot{H}^{s_p}\times\dot{H}^{s_p-1}} < \infty,$$
$$\|e(x,t)\|_{Z_{s_p}(I)} + \|S(t)(u_0 - \tilde{u}(0), u_1 - \partial_t \tilde{u}(0))\|_{Y_{s_p}(I)} \le \varepsilon,$$

there exists a solution $u(x,t)$ of (1) *defined in the interval I with the initial data (u_0, u_1) and satisfying*

$$\|u(x,t) - \tilde{u}(x,t)\|_{Y_{s_p}(I)} < C(M)\varepsilon.$$
$$\sup_{t\in I}\left\| \begin{pmatrix} u(t) \\ \partial_t u(t) \end{pmatrix} - \begin{pmatrix} \tilde{u}(t) \\ \partial_t \tilde{u}(t) \end{pmatrix} - S(t)\begin{pmatrix} u_0 - \tilde{u}(0) \\ u_1 - \partial_t \tilde{u}(0) \end{pmatrix} \right\|_{\dot{H}^{s_p}\times\dot{H}^{s_p-1}} < C(M)\varepsilon.$$

Remark 2.13. If K is a compact subset of the space $\dot{H}^{s_p} \times \dot{H}^{s_p-1}$, then there exists $T = T(K) > 0$ such that $T_+(u_0, u_1) > T(K)$ for any $(u_0, u_1) \in K$. This is a direct result from perturbation theory.

Local theory with more regular initial data. Let $s \in (s_p, 1]$. By a similar fixed-point argument we can obtain the following results.

Theorem 2.14 (local solution with $\dot{H}^s \times \dot{H}^{s-1}$ initial data). *If $(u_0, u_1) \in \dot{H}^s \times \dot{H}^{s-1}$, then there is a maximal interval $(-T_-(u_0, u_1), T_+(u_0, u_1))$ in which the equation has a solution $u(x,t)$. In addition, we have*

$$T_-(u_0, u_1), T_+(u_0, u_1) > T_1 \doteq C_{s,p}(\|(u_0, u_1)\|_{\dot{H}^s\times\dot{H}^{s-1}})^{-1/(s-s_p)},$$
$$\|u(x,t)\|_{Y_s([-T_1, T_1])} \le C_{s,p}\|(u_0, u_1)\|_{\dot{H}^s\times\dot{H}^{s-1}}.$$

Theorem 2.15 (weak long-time perturbation theory). *Let $\tilde{u}$ be a solution of the equation* (1) *in the time interval $[0, T]$ with initial data $(\tilde{u}_0, \tilde{u}_1)$, so that*

$$\|(\tilde{u}_0, \tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}} < \infty, \quad \|\tilde{u}\|_{Y_s([0,T])} < M.$$

There exist two constants $\varepsilon_0(T, M), C(T, M) > 0$ such that if (u_0, u_1) is another pair of initial data with

$$\|(u_0 - \tilde{u}_0, u_1 - \tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}} < \varepsilon_0(T, M),$$

then there exists a solution u of the equation (1) *in the time interval $[0, T]$ with initial data (u_0, u_1) so that*

$$\|u - \tilde{u}\|_{Y_s([0,T])} \le C(T, M)\|(u_0 - \tilde{u}_0, u_1 - \tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}},$$
$$\sup_{t\in[0,T]} \|(u(t) - \tilde{u}(t), \partial_t u(t) - \partial_t \tilde{u}(t))\|_{\dot{H}^s\times\dot{H}^{s-1}} \le C(T, M)\|(u_0 - \tilde{u}_0, u_1 - \tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}}.$$

Technical results.

Lemma 2.16 (gluing of $\dot H^s$ functions). *Let $-1 \le s \le 1$. Suppose $f(x)$ is a tempered distribution defined on $\mathbb{R}^3$ such that* ($R > 0$)

$$f(x) = \begin{cases} f_1(x) & \text{for } x \in B(0,2R), \\ f_2(x) & \text{for } x \in \mathbb{R}^3 \setminus B(0,R), \end{cases}$$

with $f_1, f_2 \in \dot H^s(\mathbb{R}^3)$. Then f is in the space $\dot H^s(\mathbb{R}^3)$ with

$$\|f\|_{\dot H^s(\mathbb{R}^3)} \le C(s)\big(\|f_1\|_{\dot H^s(\mathbb{R}^3)} + \|f_2\|_{\dot H^s(\mathbb{R}^3)}\big).$$

Proof. By a dilation we can always assume $R = 1$. Let $\phi(x)$ be a smooth, radial, nonnegative function such that

$$\phi(x) = \begin{cases} 1 & \text{for } x \in B(0,1), \\ 0 & \text{for } x \in \mathbb{R}^3 \setminus B(0,2). \end{cases}$$

Let us define a linear operator: $P(f) = \phi(x) f$. We know this operator is bounded from $\dot H^1(\mathbb{R}^3)$ to $\dot H^1(\mathbb{R}^3)$, and from $L^2(\mathbb{R}^3)$ to $L^2(\mathbb{R}^3)$. Thus by an interpolation, this is a bounded operator from $\dot H^s$ to itself if $0 < s < 1$. By duality P is also bounded from $\dot H^s$ to itself if $-1 \le s \le 0$. In summary, P is a bounded operator from $\dot H^s$ to itself for each $-1 \le s \le 1$. Now we have

$$f = P f_1 + f_2 - P f_2$$

as a tempered distribution. Thus

$$\|f\|_{\dot H^s} \le \|Pf_1\|_{\dot H^s} + \|f_2\|_{\dot H^s} + \|Pf_2\|_{\dot H^s} \le (\|P\|_s + 1)(\|f_1\|_{\dot H^s} + \|f_2\|_{\dot H^s}). \qquad \square$$

Lemma 2.17. *Let $u(x,t)$ be a solution of the nonlinear wave equation* (1) *with the condition* (4). *Then for any $t_1, t_2 \in I$ and $t \in \mathbb{R}$, we have*

$$\left\| \begin{pmatrix} \displaystyle\int_{t_1}^{t_2} \frac{\sin((\tau - t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\, d\tau \\ \displaystyle -\int_{t_1}^{t_2} \cos((\tau - t)\sqrt{-\Delta}) F(u(\tau))\, d\tau \end{pmatrix} \right\|_{\dot H^{s_p} \times \dot H^{s_p - 1}} \lesssim 1. \tag{7}$$

Proof. It follows directly from the identity

$$\begin{pmatrix} \displaystyle\int_{t_1}^{t_2} \frac{\sin((\tau - t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\, d\tau \\ \displaystyle -\int_{t_1}^{t_2} \cos((\tau - t)\sqrt{-\Delta}) F(u(\tau))\, d\tau \end{pmatrix} = S(t - t_1)\begin{pmatrix} u(t_1) \\ \partial_t u(t_1) \end{pmatrix} - S(t - t_2)\begin{pmatrix} u(t_2) \\ \partial_t u(t_2) \end{pmatrix}. \qquad \square$$

Lemma 2.18 (see Lemma 3.2 of [Kenig and Merle 2011]). *Let $1/2 < s < 3/2$. If $u(y)$ is a radial $\dot H^s(\mathbb{R}^3)$ function, then*

$$|u(y)| \lesssim_s \frac{1}{|y|^{\frac{3}{2} - s}} \|u\|_{\dot H^s}. \tag{8}$$

Remark 2.19. This actually means that a radial $\dot{H}^s$ function is uniformly continuous in $\mathbb{R}^3 \setminus B(0, R)$ if $R > 0$.

Lemma 2.20. *Let $r_1, r_2 > 0$ and $t_0, t_1 \in \mathbb{R}$ so that $r_1 + r_2 \le t_1 - t_0$. Suppose (u_0, u_1) is a weak limit in the space $\dot{H}^{s_p} \times \dot{H}^{s_p-1}$:*

$$\begin{aligned} u_0 &= \lim_{T\to+\infty} \int_{t_1}^{T} \frac{\sin((t-t_0)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(t)\, dt, \\ u_1 &= -\lim_{T\to+\infty} \int_{t_1}^{T} \cos((t-t_0)\sqrt{-\Delta}) F(t)\, dt. \end{aligned} \tag{9}$$

Here $F(x,t)$ is a function defined in $[t_1, \infty) \times \mathbb{R}^3$ with a finite $Z_{s_p}([t_1, T])$ norm for each $T > t_1$. In addition, we have ($1/2 < s_1 \le 1$, $\chi(x,t)$ is the characteristic function of the region indicated)

$$S = \|\chi_{|x|>r_2+|t-t_1|}(x,t) F(x,t)\|_{L^1 L^{\frac{6}{5-2s_1}}([t_1,\infty)\times\mathbb{R}^3)} < +\infty. \tag{10}$$

Then there exists a pair $(\tilde{u}_0, \tilde{u}_1)$ with $\|(\tilde{u}_0, \tilde{u}_1)\|_{\dot{H}^{s_1}\times\dot{H}^{s_1-1}(\mathbb{R}^3)} \le C_{s_1} S$ and

$$(u_0, u_1) = (\tilde{u}_0, \tilde{u}_1) \text{ in the ball } B(0, r_1).$$

Proof. Let us define

$$\begin{aligned} u_{0,T} &= \int_{t_1}^{T} \frac{\sin((t-t_0)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(t)\, dt, \qquad & u_{1,T} &= -\int_{t_1}^{T} \cos((t-t_0)\sqrt{-\Delta}) F(t)\, dt, \\ \tilde{u}_{0,T} &= \int_{t_1}^{T} \frac{\sin((t-t_0)\sqrt{-\Delta})}{\sqrt{-\Delta}} (\chi F(t))\, dt, & \tilde{u}_{1,T} &= -\int_{t_1}^{T} \cos((t-t_0)\sqrt{-\Delta}) (\chi F(t))\, dt. \end{aligned}$$

By the Strichartz estimates and the assumption (10), we know the pair $(\tilde{u}_{0,T}, \tilde{u}_{1,T})$ converges strongly in $\dot{H}^{s_1} \times \dot{H}^{s_1-1}$ to a pair $(\tilde{u}_0, \tilde{u}_1)$ as $T \to +\infty$ so that

$$\|(\tilde{u}_0, \tilde{u}_1)\|_{\dot{H}^{s_1}\times\dot{H}^{s_1-1}(\mathbb{R}^3)} \le C_{s_1} S.$$

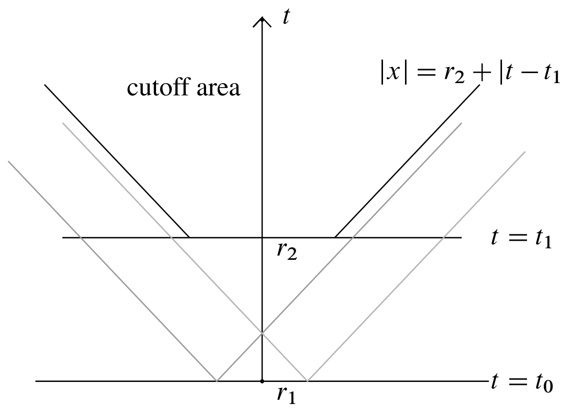

Figure 1. Illustration of proof.

In addition, we know the pair $(\tilde{u}_{0,T}, \tilde{u}_{1,T})$ is the same as $(u_{0,T}, u_{1,T})$ in the ball $B(0, r_1)$ by the strong Huygens principle. Figure 1 shows the region where the value of $F(x,t)$ may affect the value of the integrals in the ball $B(0,r_1)$. This region is disjoint with the cutoff area if $r_1 + r_2 \le t_1 - t_0$. As a result, the pair $(\tilde{u}_{0,T}, \tilde{u}_{1,T})$ converges to (u_0, u_1) weakly in the ball $B(0, r_1)$ as the pair $(u_{0,T}, u_{1,T})$ does. Considering both strong and weak convergence, we conclude that

$$(u_0, u_1) = (\tilde{u}_0, \tilde{u}_1) \quad \text{in the ball } B(0, r_1). \qquad \square$$

3. Compactness process

As we stated in the first section, the standard technique here is to show that if the main theorem failed, there would be a special minimal blow-up solution. In addition, this solution is almost periodic modulo symmetries.

Definition 3.1. A solution $u(x,t)$ of (1) is almost periodic modulo symmetries if there exists a positive function $\lambda(t)$ defined on its maximal lifespan I such that the set

$$\left\{ \left(\frac{1}{\lambda(t)^{3/2-s_p}} u\left(\frac{x}{\lambda(t)}, t\right), \frac{1}{\lambda(t)^{5/2-s_p}} \partial_t u\left(\frac{x}{\lambda(t)}, t\right) \right) : t \in I \right\}$$

is precompact in the space $\dot{H}^{s_p} \times \dot{H}^{s_p - 1}(\mathbb{R}^3)$. The function $\lambda(t)$ is called the frequency scale function, because the solution $u(t)$ at time t concentrates around the frequency $\lambda(t)$ by the compactness.

Remark 3.2. Here we use the radial condition, thus the only available symmetries are scalings. If we did not assume the radial condition, similar results would still hold but the symmetries would include translations besides scalings.

Existence of minimal blow-up solution.

Theorem 3.3 (minimal blow-up solution). *Assume that the main theorem failed. Then there would exist a solution $u(x,t)$ with a maximal lifespan I such that*

$$\sup_{t \in I} \|(u(t), \partial_t u(t))\|_{\dot{H}^{s_p} \times \dot{H}^{s_p - 1}} < \infty;$$

u blows up in the positive direction at time $T_+ \le +\infty$ with

$$\|u\|_{S([0, T_+))} = \infty.$$

In addition, u is almost periodic modulo scalings with a frequency scale function $\lambda(t)$. It is minimal in the following sense: if v is another solution with a maximal lifespan J and

$$\sup_{t \in J} \|(v(t), \partial_t v(t))\|_{\dot{H}^{s_p} \times \dot{H}^{s_p - 1}} < \sup_{t \in I} \|(u(t), \partial_t u(t))\|_{\dot{H}^{s_p} \times \dot{H}^{s_p - 1}},$$

then v is a global solution in time and scatters.

The main tool to obtain this result is the profile decomposition. One can follow the general argument in [Kenig and Merle 2010], which deals with the cubic defocusing NLS under similar assumptions.

Three enemies. Since the frequency scale function $\lambda(t)$ plays an important role in the further discussion, it is helpful if we could make additional assumptions on this function. It turns out that we can reduce the whole problem into the following three special cases. This method of three enemies was introduced in R. Killip, T. Tao and M. Visan's paper [Killip et al. 2009].

Theorem 3.4 (three enemies). *Suppose our main theorem failed. Then there would exist a minimal blow-up solution u satisfying all the conditions we mentioned in the previous theorem, so that one of the following three assumptions on its lifespan I and frequency scale function $\lambda(t)$ holds*:

(I) (soliton-like case) $I = \mathbb{R}$ *and* $\lambda(t) \equiv 1$.

(II) (high-to-low frequency cascade) $I = \mathbb{R}$, $\lambda(t) \leq 1$ *and*

$$\liminf_{t\to\pm\infty} \lambda(t) = 0.$$

(III) (self-similar case) $I = \mathbb{R}^+$ *and* $\lambda(t) = 1/t$.

The minimal blow-up solution u here could be different from the one we found in the previous theorem. But we can always manufacture a minimal blow-up solution in one of these three cases from the original one. One can follow the method used in [Killip et al. 2009] to verify this theorem.

Further compactness results. Fix a radial cutoff function $\varphi(x) \in C^\infty(\mathbb{R}^3)$ with the properties

$$\varphi(x) \begin{cases} = 0 & \text{for } |x| \leq 1/2, \\ \in [0,1] & \text{for } 1/2 \leq |x| \leq 1, \\ = 1 & \text{for } |x| \geq 1. \end{cases}$$

Given a minimal blow-up solution u mentioned above and its frequency scale function $\lambda(t)$, we have the following propositions by a compactness argument.

Proposition 3.5. *Let u be a minimal blow-up solution with a maximal lifespan I as above. There exist constants $d, C' > 0$ and $C_1 > 1$ independent of t such that*:

(I) *The interval $[t - d\lambda^{-1}(t), t + d\lambda^{-1}(t)] \subseteq I$ for all $t \in I$. In addition, we have*

$$\frac{1}{C_1}\lambda(t) \leq \lambda(t') \leq C_1\lambda(t) \tag{11}$$

for each $t' \in [t - d\lambda^{-1}(t), t + d\lambda^{-1}(t)]$.

(II) *The following estimate holds for each s_p-admissible pair (q, r) and each $t \in I$*:

$$\|u\|_{L^q L^r([t-d\lambda^{-1}(t), t+d\lambda^{-1}(t)]\times\mathbb{R}^3)} \leq C'.$$

Proposition 3.6. *Given $\varepsilon > 0$, there exists $R_1 = R_1(\varepsilon) > 0$ such that the inequality*

$$\left\| \left(\varphi\left(\frac{x}{R\lambda^{-1}(t)}\right) u(t), \varphi\left(\frac{x}{R\lambda^{-1}(t)}\right) \partial_t u(t) \right) \right\|_{\dot{H}^{s_p} \times \dot{H}^{s_p - 1}(\mathbb{R}^3)} \leq \varepsilon$$

holds for each $t \in I$ and $R > R_1(\varepsilon)$.

Proposition 3.7. *There exists two constants $R_0, \eta_0 > 0$, such that the inequality*

$$\int_t^{t+d\lambda^{-1}(t)} \int_{|x|<R_0\lambda^{-1}(t)} \frac{|u(x,\tau)|^{p+1}}{|x|}\,dx\,d\tau \geq \lambda(t)^{2-2s_p}\eta_0$$

holds for each $t \in I$. (The constant d is the same constant we used in Proposition 3.5.)

Proof. By a compactness argument we obtain that there exist $R_0, \eta_0 > 0$, so that for all $t \in I$,

$$\int_0^d \int_{|x|<R_0} \frac{\left(\frac{1}{\lambda(t)^{2/(p-1)}}|u(\lambda^{-1}(t)x, \lambda^{-1}(t)\tau + t)|\right)^{p+1}}{|x|}\,dx\,d\tau \geq \eta_0.$$

This implies

$$\int_0^d \int_{|x|<R_0} \frac{|u(\lambda^{-1}(t)x, \lambda^{-1}(t)\tau + t)|^{p+1}}{\lambda^{-1}(t)|x|} \frac{dx\,d\tau}{\lambda(t)^{\frac{2(p+1)}{p-1}+1}} \geq \eta_0.$$

$$\frac{1}{\lambda(t)^{4/(p-1)-1}} \int_0^d \int_{|x|<R_0} \frac{|u(\lambda^{-1}(t)x, \lambda^{-1}(t)\tau + t)|^{p+1}}{\lambda^{-1}(t)|x|} \frac{dx\,d\tau}{\lambda(t)^4} \geq \eta_0.$$

$$\begin{aligned}\int_t^{t+d\lambda^{-1}(t)} \int_{|x|<R_0\lambda^{-1}(t)} \frac{|u(x,\tau)|^{p+1}}{|x|}\,dx\,d\tau &\geq \lambda(t)^{4/(p-1)-1}\eta_0 \qquad (12)\\ &= \lambda(t)^{2-2s_p}\eta_0.\end{aligned}$$

This completes the proof. □

The Duhamel formula. The following formula will be frequently used in later sections.

Proposition 3.8 (Duhamel formula). *Let u be a minimal blow-up solution described above with a maximal lifespan $I = (T_-, \infty)$. Then we have*

$$\begin{aligned}
u(t) &= \lim_{T\to+\infty} \int_t^T \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\,d\tau,\\
\partial_t u(t) &= -\lim_{T\to+\infty} \int_t^T \cos((\tau-t)\sqrt{-\Delta}) F(u(\tau))\,d\tau;\\
u(t) &= \lim_{T\to T_-} \int_T^t \frac{\sin((t-\tau)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\,d\tau,\\
\partial_t u(t) &= \lim_{T\to T_-} \int_T^t \cos((t-\tau)\sqrt{-\Delta}) F(u(\tau))\,d\tau.
\end{aligned}$$

Given a time $t \in I$, these limits are weak limits in the space $\dot{H}^{s_p} \times \dot{H}^{s_p-1}$. If J is a closed interval compactly supported in I, then one can also understand the formula for $u(t)$ as a strong limit in the space $L^q L^r(J \times \mathbb{R}^3)$, as long as (q, r) is an s_p-admissible pair with $q \neq \infty$.

Remark 3.9. Actually we have

$$\begin{pmatrix} \displaystyle\int_t^T \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\,d\tau \\ \displaystyle -\int_t^T \cos((\tau-t)\sqrt{-\Delta}) F(u(\tau))\,d\tau \end{pmatrix} = \begin{pmatrix} u(t) \\ \partial_t u(t) \end{pmatrix} - S(t-T)\begin{pmatrix} u(T) \\ \partial_t u(T) \end{pmatrix}. \tag{13}$$

Thus we only need to show the corresponding limit of the last term is zero in order to verify this formula. See Lemma A.2 in the appendix for details.

4. Energy estimate near infinity

In this section, we will prove the following theorem for a minimal blow-up solution $u(x,t)$. The method was previously used in the supercritical case of the equation. (See [Kenig and Merle 2011] for more details.) In the supercritical case, by the Sobolev embedding, the energy automatically exists at least locally in the space, for any given time $t \in I$. In the subcritical case, however, we need to use the approximation techniques.

Theorem 4.1 (energy estimate near infinity). *Let $u(x,t)$ be a minimal blow-up solution as we found in the previous section. Then $(u(x,t), \partial_t u(x,t)) \in \dot{H}^1 \times L^2(\mathbb{R}^3 \setminus B(0,r))$ for each $r > 0$, $t \in I$. Actually we have*

$$\int_{r<|x|<4r} (|\nabla u(x,t)|^2 + |\partial_t u(x,t)|^2)\,dx \le C r^{-2(1-s_p)}. \tag{14}$$

The constant C depends on p and $\sup_{t\in I} \|(u(t), \partial_t u(t))\|_{\dot{H}^{s_p}\times \dot{H}^{s_p-1}}$ but nothing else.

Preliminary results.

Introduction to $w(r,t)$. Let $u(x,t)$ be a radial solution of the wave equation

$$\partial_t^2 u - \Delta u = F(x,t).$$

If we define $w(r,t), h(r,t) : \mathbb{R}^+ \times I \to \mathbb{R}$ so that

$$w(r,t) = ru(x,t), \quad h(r,t) = rF(x,t),$$

then we have $w(r,t)$ is the solution of the one-dimensional wave equation

$$\partial_t^2 w - \partial_r^2 w = h(r,t).$$

Lemma 4.2. *Let $(u(x,t_0), \partial_t u(x,t_0))$ be radial and in the energy space $\dot{H}^1 \times L^2$ locally. Then for any $0 < a < b < \infty$, we have that the identity*

$$\frac{1}{4\pi}\int_{a<|x|<b} (|\nabla u|^2 + |\partial_t u|^2)\,dx = \left(\int_a^b [(\partial_r w)^2 + (\partial_t w)^2]\,dr\right) + (au^2(a) - bu^2(b))$$

holds (if we take the value of the functions at time t_0).

Proof. By direct computation

$$\begin{aligned}\int_a^b [(\partial_r w)^2 + (\partial_t w)^2]\, dr &= \int_a^b [(r\partial_r u + u)^2 + (r\partial_t u)^2]\, dr\\ &= \int_a^b [r^2(\partial_r u)^2 + u^2 + r^2(\partial_t u)^2]\, dr + \int_a^b 2ru\, \partial_r u\, dr\\ &= \int_a^b [r^2(\partial_r u)^2 + r^2(\partial_t u)^2 + u^2]\, dr + \int_a^b r\, d(u^2)\\ &= \int_a^b r^2[(\partial_r u)^2 + (\partial_t u)^2]\, dr + [ru^2]_a^b\\ &= \frac{1}{4\pi}\int_{a<|x|<b} (|\nabla u|^2 + |\partial_t u|^2)\, dx + bu^2(b) - au^2(a). \qquad \square\end{aligned}$$

Lemma 4.3. *Let $w(r,t)$ be a solution to the equation*

$$\partial_t^2 w - \partial_r^2 w = h(r,t)$$

for $(r,t) \in \mathbb{R}^+ \times I$, so that $(w, \partial_t w) \in C(I; \dot{H}^1 \times L^2(R_1 < r < R_2))$ for any $0 < R_1 < R_2 < \infty$. Let us define

$$\begin{aligned} z_1(r,t) &= \partial_t w(r,t) - \partial_r w(r,t),\\ z_2(r,t) &= \partial_t w(r,t) + \partial_r w(r,t).\end{aligned}$$

Then we have (with $M > 0$)

$$\left|\left(\int_{r_0}^{4r_0} |z_1(r,t_0)|^2\, dr\right)^{\frac{1}{2}} - \left(\int_{r_0+M}^{4r_0+M} |z_1(r,t_0+M)|^2\, dr\right)^{\frac{1}{2}}\right| \leq \left(\int_{r_0}^{4r_0}\left(\int_0^M h(r+t, t_0+t)\, dt\right)^2 dr\right)^{\frac{1}{2}} \tag{15}$$

if $t_0, t_0 + M \in I$, and

$$\left|\left(\int_{r_0}^{4r_0} |z_2(r,t_0)|^2\, dr\right)^{\frac{1}{2}} - \left(\int_{r_0+M}^{4r_0+M} |z_2(r,t_0-M)|^2\, dr\right)^{\frac{1}{2}}\right| \leq \left(\int_{r_0}^{4r_0}\left(\int_0^M h(r+t, t_0-t)\, dt\right)^2 dr\right)^{\frac{1}{2}} \tag{16}$$

if $t_0, t_0 - M \in I$.

Proof. We will assume w has sufficient regularity, otherwise we only need to use the standard techniques of smooth approximation. Let us define

$$z(r,s) = (\partial_t - \partial_r) w(r+s, t_0+s).$$

We have

$$\partial_s z(r,s) = (\partial_t + \partial_r)(\partial_t - \partial_r) w(r+s, t_0+s) = h(r+s, t_0+s).$$

Thus

$$z(r,M) = z(r,0) + \int_0^M h(r+t, t_0+t)\,dt.$$

Applying the triangle inequality, we obtain the first inequality. The second inequality can be proved in a similar way. □

Smooth approximation.

Introduction. Let $u(x,t)$ be a minimal blow-up solution. Choose a smooth, nonnegative, radial function $\varphi(x,t)$ supported in the four-dimensional ball $B(0,1) \subset \mathbb{R}^4$ such that

$$\int_{\mathbb{R}^3\times\mathbb{R}} \varphi(x,t)\,dx\,dt = 1.$$

Let d be the number given in Proposition 3.5. If $\varepsilon < d$, we define (both the functions u and $F(u)$ are locally integrable)

$$\varphi_\varepsilon(x,t) = \frac{1}{\varepsilon^4}\varphi(x/\varepsilon, t/\varepsilon), \quad u_\varepsilon = u * \varphi_\varepsilon, \quad F_\varepsilon = F(u) * \varphi_\varepsilon.$$

This makes $u_\varepsilon(x,t)$ be a smooth solution of the linear wave equation

$$\partial_t^2 u_\varepsilon(x,t) - \Delta u_\varepsilon(x,t) = F_\varepsilon(x,t),$$

with the convergence (using the continuity of $(u(t), \partial_t u(t))$ in the space $\dot{H}^{s_p} \times \dot{H}^{s_p-1}$)

$$(u_\varepsilon(t_0), \partial_t u_\varepsilon(t_0)) \to (u(t_0), \partial_t u(t_0)) \quad \text{in the space } \dot{H}^{s_p} \times \dot{H}^{s_p-1} \text{ for each } t_0 \in I$$

and the estimate

$$\|(u_\varepsilon(t_0), \partial_t u_\varepsilon(t_0))\|_{\dot{H}^{s_p}\times\dot{H}^{s_p-1}} \le \sup_{t\in I} \|(u(t), \partial_t u(t))\|_{\dot{H}^{s_p}\times\dot{H}^{s_p-1}} \lesssim 1.$$

In addition, if $a - \varepsilon \in I$, we have

$$\|F_\varepsilon(x,t)\|_{Z_{s_p}([a,b])} < \infty.$$

Remark 4.4. We have to apply the smooth kernel on the whole nonlinear term, because if we just made the initial data smooth, we would not resume the compactness conditions of the minimal blow-up solution.

The Duhamel formula.

Lemma 4.5 (almost periodic property). *The set*

$$\left\{ \left(\frac{1}{\lambda(t)^{3/2-s_p}} u_\varepsilon\left(\frac{x}{\lambda(t)}, t\right), \frac{1}{\lambda(t)^{5/2-s_p}}\, \partial_t u_\varepsilon\left(\frac{x}{\lambda(t)}, t\right) \right) : t \in [d+1, \infty) \right\}$$

is precompact in the space $\dot{H}^{s_p} \times \dot{H}^{s_p-1}(\mathbb{R}^3)$ for each fixed $\varepsilon < d$. The number d here is the constant we obtained in Proposition 3.5.

Proof. Given a sequence $\{t_n\}$ we could assume without loss of generality that

$$\lambda(t_n) \to \lambda_0 \in [0,1],$$
$$\left(\frac{1}{\lambda(t_n)^{3/2-s_p}} u\left(\frac{x}{\lambda(t_n)}, t_n\right), \frac{1}{\lambda(t_n)^{5/2-s_p}}\, \partial_t u\left(\frac{x}{\lambda(t_n)}, t_n\right)\right) \to (u_0, u_1),$$

by extracting a subsequence if necessary. Let $\tilde{u}(x,t)$ be the solution of the equation (1) with initial data (u_0, u_1). By the long-time perturbation theory we know

$$\sup_{t\in[-d,d]} \left\| \begin{pmatrix} \frac{1}{\lambda(t_n)^{3/2-s_p}} u\left(\frac{x}{\lambda(t_n)}, t_n + \frac{t}{\lambda(t_n)}\right) \\ \frac{1}{\lambda(t_n)^{5/2-s_p}}\, \partial_t u\left(\frac{x}{\lambda(t_n)}, t_n + \frac{t}{\lambda(t_n)}\right) \end{pmatrix} - \begin{pmatrix} \tilde{u}(t) \\ \partial_t \tilde{u}(t) \end{pmatrix} \right\|_{\dot{H}^{s_p} \times \dot{H}^{s_p-1}} \to 0.$$

This implies

$$\begin{aligned}
\begin{pmatrix} \frac{1}{\lambda(t_n)^{3/2-s_p}} u_\varepsilon\left(\frac{x}{\lambda(t_n)}, t_n\right) \\ \frac{1}{\lambda(t_n)^{5/2-s_p}}\, \partial_t u_\varepsilon\left(\frac{x}{\lambda(t_n)}, t_n\right) \end{pmatrix} &= \left[\varphi_{\varepsilon\lambda(t_n)} * \begin{pmatrix} \frac{1}{\lambda(t_n)^{3/2-s_p}} u\left(\frac{\cdot}{\lambda(t_n)}, t_n + \frac{\cdot}{\lambda(t_n)}\right) \\ \frac{1}{\lambda(t_n)^{5/2-s_p}}\, \partial_t u\left(\frac{\cdot}{\lambda(t_n)}, t_n + \frac{\cdot}{\lambda(t_n)}\right) \end{pmatrix}\right]_{t=0} \\
&= \left[\varphi_{\varepsilon\lambda(t_n)} * \begin{pmatrix} \tilde{u} \\ \partial_t \tilde{u} \end{pmatrix}\right]_{t=0} + o(1) \\
&= \begin{cases} \left[\varphi_{\varepsilon\lambda_0} * \begin{pmatrix} \tilde{u} \\ \partial_t \tilde{u} \end{pmatrix}\right]_{t=0} + o(1) & \text{if } \lambda_0 > 0; \\ \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} + o(1) & \text{if } \lambda_0 = 0; \end{cases}
\end{aligned}$$

The error $o(1)$ tends to zero as $n \to \infty$ in the sense of the $\dot{H}^{s_p} \times \dot{H}^{s_p-1}$ norm. □

Lemma 4.6. *The Duhamel formula*

$$u_\varepsilon(t_0) = \int_{t_0}^{+\infty} \frac{\sin((\tau - t_0)\sqrt{-\Delta})}{\sqrt{-\Delta}} F_\varepsilon(x,\tau)\, d\tau,$$
$$\partial_t u_\varepsilon(t_0) = -\int_{t_0}^{+\infty} \cos((\tau - t_0)\sqrt{-\Delta}) F_\varepsilon(x,\tau)\, d\tau.$$

still holds for u_ε in the sense of weak limit if $\varepsilon < d$ and $t_0 - \varepsilon \in I$. In the soliton-like or high-to-low frequency cascade case, we can also establish the Duhamel formula in the negative time direction.

Proof. This lemma can be proved in exactly the same way as the original Duhamel formula (see Lemma A.2). The key ingredient is the almost periodic property we have just obtained above. □

Decay of u_ε and F_ε at infinity.

Lemma 4.7. *If $|x| > 10\varepsilon$, we have*

$$|u_\varepsilon(x,t)| \le \frac{C}{|x|^{2/(p-1)}}, \quad |F_\varepsilon(x,t)| \le \frac{C}{|x|^{2p/(p-1)}}.$$

The constant C depends only on p and the upper bound $\sup_{t\in I} \|(u(t), \partial_t u(t))\|_{\dot{H}^{s_p}\times\dot{H}^{s_p-1}}$.

Proof. This comes from the estimate (8) and an easy computation. □

Uniform estimate on u_ε. In this subsection, we will prove the following lemma. It implies Theorem 4.1 immediately by a limit process. The functions $w_\varepsilon(r,t)$ and $z_{i,\varepsilon}(r,t)$ below are defined as described earlier using $u_\varepsilon(x,t)$.

Lemma 4.8. *Let $t_0 \in I$ and $r_0 > 0$. Then for sufficiently small ε, we have*

$$\int_{r_0<|x|<4r_0} (|\nabla u_\varepsilon(x,t_0)|^2 + |\partial_t u_\varepsilon(x,t_0)|^2)\, dx \le C r_0^{-2(1-s_p)}. \tag{17}$$

The constant C can be chosen in a way that it depends only on p and the upper bound

$$\sup_{t\in I} \|(u(t), \partial_t u(t))\|_{\dot H^{s_p}\times \dot H^{s_p-1}}.$$

Conversion to $w_\varepsilon(r,t)$. First choose $\varepsilon < \min\{r_0/10, d\}$. If the minimal blow-up solution is a self-similar one, we also require $\varepsilon < t_0/2$. Let us apply Lemmas 4.2 and 4.7. It is sufficient to show

$$\int_{r_0}^{4r_0} (|\partial_r w_\varepsilon(r,t_0)|^2 + |\partial_t w_\varepsilon(r,t_0)|^2)\, dr \le C r_0^{-2(1-s_p)}.$$

In other words,

$$\int_{r_0}^{4r_0} (|z_{1,\varepsilon}(r,t_0)|^2 + |z_{2,\varepsilon}(r,t_0)|^2)\, dr \le C r_0^{-2(1-s_p)}. \tag{18}$$

Expansion of $z_{1,\varepsilon}$. Let us break $(u_\varepsilon(t), \partial_t u_\varepsilon(t))$ into two pieces:

$$u_\varepsilon^{(1)}(t) = \int_t^{t_0+100r_0} \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F_\varepsilon(\tau)\, d\tau,$$
$$\partial_t u_\varepsilon^{(1)}(t) = -\int_t^{t_0+100r_0} \cos((\tau-t)\sqrt{-\Delta}) F_\varepsilon(\tau)\, d\tau,$$

and

$$u_\varepsilon^{(2)}(t) = \int_{t_0+100r_0}^{\infty} \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F_\varepsilon(\tau)\, d\tau,$$
$$\partial_t u_\varepsilon^{(2)}(t) = -\int_{t_0+100r_0}^{\infty} \cos((\tau-t)\sqrt{-\Delta}) F_\varepsilon(\tau)\, d\tau.$$

These are smooth functions, and we have

$$(u_\varepsilon(x,t_0), \partial_t u_\varepsilon(x,t_0)) = (u_\varepsilon^{(1)}(x,t_0), \partial_t u_\varepsilon^{(1)}(x,t_0)) + (u_\varepsilon^{(2)}(x,t_0), \partial_t u_\varepsilon^{(2)}(x,t_0)).$$

Defining $w_\varepsilon^{(j)}, z_{1,\varepsilon}^{(j)}$ accordingly for $j = 1,2$, we have

$$z_{1,\varepsilon}(x,t_0) = z_{1,\varepsilon}^{(1)}(x,t_0) + z_{1,\varepsilon}^{(2)}(x,t_0).$$

Short-time contribution. We have $u_\varepsilon^{(1)}$ satisfies the wave equation

$$\begin{cases}\partial_t^2 u_\varepsilon^{(1)} - \Delta u_\varepsilon^{(1)} = F_\varepsilon(x,t), & (x,t) \in \mathbb{R}^3 \times (t_0^-, +\infty),\\ u_\varepsilon^{(1)}|_{t=t_0+100r_0} = 0 \in \dot{H}^{s_p}(\mathbb{R}^3),\\ \partial_t u_\varepsilon^{(1)}|_{t=t_0+100r_0} = 0 \in \dot{H}^{s_p-1}(\mathbb{R}^3).\end{cases}$$

Thus $w_\varepsilon^{(1)}$ is a smooth solution of

$$\begin{cases}\partial_t^2 w_\varepsilon^{(1)} - \partial_r^2 w_\varepsilon^{(1)} = rF_\varepsilon(r,t), & (r,t) \in \mathbb{R}^+ \times (t_0^-, +\infty),\\ w_\varepsilon^{(1)}|_{t=t_0+100r_0} = 0,\\ \partial_t w_\varepsilon^{(1)}|_{t=t_0+100r_0} = 0.\end{cases}$$

Applying Lemmas 4.3 and 4.7, we obtain

$$\begin{aligned}\left(\int_{r_0}^{4r_0} |z_{1,\varepsilon}^{(1)}(r,t_0)|^2\,dr\right)^{\frac{1}{2}} &\le \left(\int_{r_0}^{4r_0}\left(\int_0^{100r_0}(t+r)F_\varepsilon(t+r,t+t_0)\,dt\right)^2 dr\right)^{\frac{1}{2}}\\ &\lesssim \left(\int_{r_0}^{4r_0}\left(\int_0^{100r_0}(t+r)\frac{1}{(t+r)^{2p/(p-1)}}\,dt\right)^2 dr\right)^{\frac{1}{2}}\\ &\lesssim \left(\int_{r_0}^{4r_0}\left(\int_0^{100r_0}\frac{1}{(t+r)^{1+2/(p-1)}}\,dt\right)^2 dr\right)^{\frac{1}{2}}\\ &\lesssim \left(\int_{r_0}^{4r_0}\frac{1}{r^{4/(p-1)}}\,dr\right)^{\frac{1}{2}}\\ &\lesssim \frac{1}{r_0^{1-s_p}}.\end{aligned}$$

Long-time contribution. Let us define a cutoff function $\chi(x,t)$ to be the characteristic function of the region $\{(x,t) : |x| > t - t_0 - 50r_0\}$. By Lemma 4.7, we know

$$\begin{aligned}\|\chi F_\varepsilon\|_{L^1L^2([t_0+100r_0,\infty)\times\mathbb{R}^3)} &= \int_{t_0+100r_0}^{\infty}\left(\int_{|x|>t-t_0-50r_0}|F_\varepsilon|^2\,dx\right)^{\frac{1}{2}} dt\\ &\lesssim \int_{t_0+100r_0}^{\infty}\left(\int_{|x|>t-t_0-50r_0}\frac{1}{|x|^{4p/(p-1)}}\,dx\right)^{\frac{1}{2}} dt\\ &\lesssim \int_{t_0+100r_0}^{\infty}\left(\frac{1}{|t-t_0-50r_0|^{1+4/(p-1)}}\right)^{\frac{1}{2}} dt\\ &\lesssim \int_{t_0+100r_0}^{\infty}\frac{1}{|t-t_0-50r_0|^{\frac{1}{2}+2/(p-1)}}\,dt\\ &\lesssim \frac{1}{r_0^{1-s_p}}.\end{aligned}$$

Applying Lemma 2.20, we obtain

$$\int_{r_0<|x|<4r_0} (|\nabla u_\varepsilon^{(2)}(x,t_0)|^2 + |\partial_t u_\varepsilon^{(2)}(x,t_0)|^2)\,dx \lesssim r_0^{2(s_p-1)}.$$

Applying Lemma 4.2 and using the fact (plus (8))

$$\begin{aligned}\|(u_\varepsilon^{(2)}(t_0),\partial_t u_\varepsilon^{(2)}(t_0))\|_{\dot H^{s_p}\times\dot H^{s_p-1}} &= \left\| S(-100r_0)\begin{pmatrix} u_\varepsilon(t_0+100r_0)\\ \partial_t u_\varepsilon(t_0+100r_0)\end{pmatrix}\right\|_{\dot H^{s_p}\times\dot H^{s_p-1}}\\ &= \|(u_\varepsilon(t_0+100r_0),\partial_t u_\varepsilon(t_0+100r_0))\|_{\dot H^{s_p}\times\dot H^{s_p-1}}\\ &\le \sup_I \|(u,\partial_t u)\|_{\dot H^{s_p}\times\dot H^{s_p-1}} \lesssim 1,\end{aligned}$$

we obtain

$$\int_{r_0}^{4r_0} (|\partial_r w_\varepsilon^{(2)}(r,t_0)|^2 + |\partial_t w_\varepsilon^{(2)}(r,t_0)|^2)\,dr \lesssim r_0^{2(s_p-1)},$$

$$\int_{r_0}^{4r_0} |z_{1,\varepsilon}^{(2)}(r,t_0)|^2\,dr \lesssim r_0^{2(s_p-1)}.$$

Combining with the estimate for $z_{1,\varepsilon}^{(1)}$, we have

$$\int_{r_0}^{4r_0} |z_{1,\varepsilon}(r,t_0)|^2\,dr \lesssim r_0^{2(s_p-1)}.$$

Estimate of $z_{2,\varepsilon}$. We also need to consider $z_{2,\varepsilon}$. In the soliton-like case or the high-to-low frequency cascade case, this can be done in exactly the same way as $z_{1,\varepsilon}$. Now let us consider the self-similar case.

Lemma 4.9. *Let u be a self-similar minimal blow-up solution. If $t_0 \le 0.3r_0$, then $(u(t_0),\partial_t u(t_0))$ is in $\dot H^1\times L^2(|x|>0.9r_0)$ with*

$$\int_{|x|>0.9r_0} (|\nabla u(x,t_0)|^2 + |\partial_t u(x,t_0)|^2)\,dx \lesssim r_0^{2(s_p-1)}.$$

Proof. We have (the Duhamel formula)

$$\begin{aligned} u(t_0) &= \int_{0^+}^{t_0} \frac{\sin((t_0-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(t)\,dt,\\ \partial_t u(t_0) &= \int_{0^+}^{t_0} \cos((t_0-t)\sqrt{-\Delta}) F(t)\,dt,\end{aligned}$$

and

$$\begin{aligned} \tilde u_0 &= \int_{0^+}^{t_0} \frac{\sin((t_0-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} \chi(|x|>0.5r_0)F(t)\,dt,\\ \tilde u_1 &= \int_{0^+}^{t_0} \cos((t_0-t)\sqrt{-\Delta})\chi(|x|>0.5r_0)F(t)\,dt.\end{aligned}$$

A straightforward computation shows $\|\chi F\|_{L^1L^2((0^+,t_0)\times\mathbb{R}^3)} \lesssim r_0^{s_p-1}$. This means $(\tilde{u}_0,\tilde{u}_1)$ is in the space $\dot{H}^1\times L^2(\mathbb{R}^3)$ with a norm $\lesssim r_0^{s_p-1}$. By the strong Huygens principle we can repeat the argument we used in Lemma 2.20 and obtain

$$(u(t_0),\partial_t u(t_0)) = (\tilde{u}_0,\tilde{u}_1) \text{ in the region } \mathbb{R}^3\setminus B(0,0.9r_0). \qquad \square$$

Lemma 4.10. *Let u be a self-similar solution. If $t_0 \le 0.2r_0$ and $\varepsilon < t_0/2$, then we have*

$$\int_{r_0<|x|<4r_0} (|\nabla u_\varepsilon(x,t_0)|^2 + |\partial_t u_\varepsilon(x,t_0)|^2)\,dx \lesssim r_0^{2(s_p-1)}.$$

Proof. We have $\nabla u_\varepsilon = \varphi_\varepsilon * \nabla u$, thus $|\nabla u_\varepsilon| \le \varphi_\varepsilon * |\nabla u|$. Thus (we have $\varepsilon < 0.1r_0$)

$$\int_{r_0<|x|<4r_0} |\nabla u_\varepsilon(x,t_0)|^2\,dx \le \sup_{t\in[t_0-\varepsilon,t_0+\varepsilon]} \int_{0.9r_0<|x|<4.1r_0} |\nabla u(x,t)|^2\,dx \lesssim r_0^{2(s_p-1)}$$

by our previous lemma. The other term can be estimated using the same method. $\square$

Remark 4.11. By Lemmas 4.2 and 4.7, this lemma implies (if $t_0 \le 0.2r_0$ and $\varepsilon < t_0/2$)

$$\int_{r_0}^{4r_0} (|\partial_r w_\varepsilon(r,t_0)|^2 + |\partial_t w_\varepsilon(r,t_0)|^2)\,dr \lesssim r_0^{2(s_p-1)}. \tag{19}$$

In the self-similar case, let us recall that we always choose $\varepsilon < \min\{r_0/10, t_0/2, d\}$. By Lemma 4.10 and Remark 4.11, we only need to consider the case $t_0 > 0.2r_0$ in order to estimate $z_{2,\varepsilon}$. Applying Lemma 4.3, we have

$$\begin{aligned}\left(\int_{r_0}^{4r_0} |z_{2,\varepsilon}(r,t_0)|^2\,dr\right)^{\frac12} \\ \le \left(\int_{t_0+0.8r_0}^{t_0+3.8r_0} |z_{2,\varepsilon}(r,0.2r_0)|^2\,dr\right)^{\frac12} + \left(\int_{r_0}^{4r_0}\left(\int_0^{t_0-0.2r_0} (t+r)F_\varepsilon(t+r,t_0-t)\,dt\right)^2 dr\right)^{\frac12}.\end{aligned}$$

The first term is dominated by $r_0^{s_p-1}$ because of (19). We can gain the same upper bound for the second term by a basic computation similar to the one we used for $z_{1,\varepsilon}$.

Conclusion. Now we combine the estimates for $z_{1,\varepsilon}$ and $z_{2,\varepsilon}$, thus concluding our Lemma 4.8.

Local energy estimate and its corollary. As mentioned earlier, we are able to establish Theorem 4.1 immediately by letting ε converge to zero. (See Lemma A.6 for details of this argument.) Furthermore, we can obtain the following proposition by applying Lemma 4.2 on u.

Proposition 4.12. *Let $u(x,t)$ be a minimal blow-up solution as above*; *we have*

$$\int_{r_0}^{4r_0} (|\partial_r w(r,t_0)|^2 + |\partial_t w(r,t_0)|^2)\,dr \lesssim r_0^{2(s_p-1)},$$
$$\int_{r_0}^{4r_0} (|z_1(r,t_0)|^2 + |z_2(r,t_0)|^2)\,dr \lesssim r_0^{2(s_p-1)}.$$

5. Recurrence process

In the previous section we found that the minimal blow-up solution is locally in the energy space. However, our goal is to gain a global energy estimate. This section features a recurrence process which helps us march toward higher regularity. We will prove the following lemma. Throughout the whole section we assume u satisfies all the conditions mentioned in the lemma.

Lemma 5.1. *Let $u(x,t)$ be a minimal blow-up solution of* (1) *as obtained in Section 3* (*compactness process*) *with a frequency scale function $\lambda(t)$. In addition, the set K is precompact in the space $\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^3)$ for some number $s \in [s_p, 1)$:*

$$K = \left\{ \left(\frac{1}{\lambda(t)^{3/2-s_p}} u\left(\frac{x}{\lambda(t)}, t\right), \frac{1}{\lambda(t)^{5/2-s_p}} \partial_t u\left(\frac{x}{\lambda(t)}, t\right) \right) : t \in I \right\}.$$

Then at least one of the following holds.

- *The solution u satisfies the energy estimate*

$$\|(u(t), \partial_t u(t))\|_{\dot{H}^1 \times L^2(\mathbb{R}^3)} \lesssim (\lambda(t))^{1-s_p}.$$

- *The set K is also precompact in the space $\dot{H}^{s+0.98\sigma_2(p)} \times \dot{H}^{s-1+0.98\sigma_2(p)}$. Here the number $\sigma_2(p) > 0$ depends on nothing but p.*

Remark 5.2. The compactness of K immediately gives the estimate

$$\|u(t), \partial_t u(t)\|_{\dot{H}^s \times \dot{H}^{s-1}} \lesssim (\lambda(t))^{s-s_p}, \quad t \in I.$$

Setup and technical lemmas.

Definition 5.3. Let us define

$$S(A) = \sup_{t \in I} (\lambda(t))^{s_p - s} \|u_{>\lambda(t)A}\|_{Y_s([t, t+d\lambda^{-1}(t)])},$$
$$N(A) = \sup_{t \in I} (\lambda(t))^{s_p - s} \|P_{>\lambda(t)A} F(u)\|_{Z_s([t, t+d\lambda^{-1}(t)])}.$$

Proposition 5.4. *The functions $S(A)$ and $N(A)$ are universally bounded for all $A > 0$ with the limit*

$$\lim_{A \to +\infty} S(A) = 0.$$

Proof. By our assumptions on compactness and Proposition 3.5 part (I), we obtain that the set

$$\left\{ \left(\frac{1}{\lambda(t)^{3/2-s_p}} u\left(\frac{x}{\lambda(t)}, t + \frac{\tau}{\lambda(t)}\right), \frac{1}{\lambda(t)^{5/2-s_p}} \partial_t u\left(\frac{x}{\lambda(t)}, t + \frac{\tau}{\lambda(t)}\right) \right) : \tau \in [0, d], t \in I \right\}$$

is precompact in the space $\dot{H}^s \times \dot{H}^{s-1}$. Applying either Proposition 3.5 part (II) (if $s = s_p$) or Theorem 2.14 (if $s > s_p$), we also have a bound independent of t:

$$\left\| \frac{1}{\lambda(t)^{3/2-s_p}} u\left(\frac{x}{\lambda(t)}, t + \frac{\tau}{\lambda(t)}\right) \right\|_{Y_s([0,d])} \lesssim 1. \tag{20}$$

Combining these facts with perturbation theory (Theorem 2.12 if $s = s_p$, or Theorem 2.15 if $s > s_p$), we have

$$\left\{ \frac{1}{\lambda(t)^{3/2-s_p}} u\left(\frac{x}{\lambda(t)}, t + \frac{\tau}{\lambda(t)}\right), \tau \in [0,d] : t \in I \right\}$$

is precompact in the space $Y_s([0,d])$. This immediately gives the uniform convergence for $t \in I$,

$$\left\| P_{>A} \frac{1}{\lambda(t)^{3/2-s_p}} u\left(\frac{x}{\lambda(t)}, t + \frac{\tau}{\lambda(t)}\right) \right\|_{Y_s([0,d])} \rightrightarrows 0, \quad \text{as } A \to \infty. \tag{21}$$

If we rescale the inequality (20) back, we obtain

$$(\lambda(t))^{s_p-s} \|u\|_{Y_s([t,t+d\lambda^{-1}(t)])} \lesssim 1 \Rightarrow (\lambda(t))^{s_p-s} \|F(u)\|_{Z_s([t,t+d\lambda^{-1}(t)])} \lesssim 1,$$

which implies that $S(A)$ and $N(A)$ are uniformly bounded. In a similar way we can show $S(A)$ converges to zero as $A \to \infty$, using the uniform convergence (21) above. □

Definition 5.5. Let us set

$$\Sigma(s,p) = s + 1 - (2p-2)(s - s_p)$$

for convenience. Thus the $Y_s(I)$ norm can also be written as $L^{2p/\Sigma(s,p)} L^{2p/(1-s)}(I \times \mathbb{R}^3)$ norm.

Lemma 5.6 (bilinear estimate). *Suppose u_i satisfies the linear wave equation on the time interval $I = [0,T]$, $i = 1,2$,*

$$\partial_t^2 u_i - \Delta u_i = F_i(x,t),$$

with the initial data $(u_i|_{t=0}, \partial_t u_i|_{t=0}) = (u_{0,i}, u_{1,i})$. Then

$$\begin{aligned} S &= \|(P_{>R} u_1)(P_{<r} u_2)\|_{L^{\frac{p}{\Sigma(s,p)}} L^{\frac{p}{2-s}}(I\times\mathbb{R}^3)} \\ &\lesssim \left(\frac{r}{R}\right)^{\sigma} \big(\|(u_{0,1}, u_{1,1})\|_{\dot H^s \times \dot H^{s-1}} + \|F_1\|_{Z_s(I)}\big) \times \big(\|(u_{0,2}, u_{1,2})\|_{\dot H^s \times \dot H^{s-1}} + \|F_2\|_{Z_s(I)}\big). \end{aligned}$$

Here the number σ is an arbitrary positive constant satisfying

$$\sigma \le 3\left(\frac{1}{2} - \frac{\Sigma(s,p)}{2p} - \frac{2-s}{2p}\right), \quad \sigma < 3 \times \frac{2-s}{2p}. \tag{22}$$

Remark 5.7. We can actually choose

$$\sigma = \sigma(p) = \frac{3\min\{p-3, 1\}}{2p} > 0.$$

This constant $\sigma(p)$ depends on nothing but p. This fact plays an important role in our discussion.

Proof. By the Strichartz estimate

$$\|(P_{>R}) u_1\|_{L^{\frac{2p}{\Sigma(s,p)}} L^{1/(\frac{2-s}{2p} + \frac{\sigma}{3})}} \lesssim \|(D_x^{-\sigma} P_{>R} u_{0,1}, D_x^{-\sigma} P_{>R} u_{1,1})\|_{\dot H^s \times \dot H^{s-1}} + \|D_x^{-\sigma} P_{>R} F_1\|_{Z_s(I)},$$

$$\|(P_{<r}) u_2\|_{L^{\frac{2p}{\Sigma(s,p)}} L^{1/(\frac{2-s}{2p} - \frac{\sigma}{3})}} \lesssim \|(D_x^{\sigma} P_{<r} u_{0,2}, D_x^{\sigma} P_{<r} u_{1,2})\|_{\dot H^s \times \dot H^{s-1}} + \|D_x^{\sigma} P_{<r} F_2\|_{Z_s(I)}.$$

Our choice of σ makes sure that the pairs above are admissible. Thus we have

$$
\begin{aligned}
&\|(P_{>R}u_1)(P_{<r}u_2)\|_{L^{\frac{p}{\Sigma(s,p)}}L^{\frac{p}{2-s}}} \\
&\quad\lesssim \|(P_{>R})u_1\|_{L^{\frac{2p}{\Sigma(s,p)}}L^{1/(\frac{2-s}{2p}+\frac{\sigma}{3})}}\|(P_{<r})u_2\|_{L^{\frac{2p}{\Sigma(s,p)}}L^{1/(\frac{2-s}{2p}-\frac{\sigma}{3})}} \\
&\quad\lesssim \big(\|(D_x^{-\sigma}P_{>R}u_{0,1}, D_x^{-\sigma}P_{>R}u_{1,1})\|_{\dot H^s\times\dot H^{s-1}} + \|D_x^{-\sigma}P_{>R}F_1\|_{Z_s(I)}\big) \\
&\qquad\qquad\qquad\times\big(\|(D_x^{\sigma}P_{<r}u_{0,2}, D_x^{\sigma}P_{<r}u_{1,2})\|_{\dot H^s\times\dot H^{s-1}} + \|D_x^{\sigma}P_{<r}F_2\|_{Z_s(I)}\big) \\
&\quad\lesssim \Big(\frac{1}{R}\Big)^{\sigma}\big(\|(P_{>R}u_{0,1}, P_{>R}u_{1,1})\|_{\dot H^s\times\dot H^{s-1}} + \|P_{>R}F_1\|_{Z_s(I)}\big) \\
&\qquad\qquad\qquad\times r^{\sigma}\big(\|(P_{<r}u_{0,2}, P_{<r}u_{1,2})\|_{\dot H^s\times\dot H^{s-1}} + \|P_{<r}F_2\|_{Z_s(I)}\big) \\
&\quad\lesssim \text{the right-hand side.}
\end{aligned}
$$
□

Lemma 5.8. *Let $u(x,t)$ be a function defined on $I\times\mathbb{R}^3$, such that $\hat u$ is supported in the ball $B(0,r)$ for each $t\in I$. Then*

$$
\|P_{>R}F(u(x,t))\|_{L^{\frac{2}{\Sigma(s,p)}}L^{\frac{2}{2-s}}(I\times\mathbb{R}^3)} \lesssim \Big(\frac{r}{R}\Big)^2\|u\|^p_{Y_s(I)}.
$$

Proof. We have

$$
\begin{aligned}
&\|P_{>R}F(u(x,t))\|_{L^{\frac{2}{\Sigma(s,p)}}L^{\frac{2}{2-s}}(I\times\mathbb{R}^3)} \\
&\qquad\lesssim \frac{1}{R^2}\|P_{>R}\Delta_x F(u(x,t))\|_{L^{\frac{2}{\Sigma(s,p)}}L^{\frac{2}{2-s}}(I\times\mathbb{R}^3)} \\
&\qquad\lesssim \frac{1}{R^2}\|\Delta_x F(u(x,t))\|_{L^{\frac{2}{\Sigma(s,p)}}L^{\frac{2}{2-s}}(I\times\mathbb{R}^3)} \\
&\qquad\lesssim \frac{1}{R^2}\|p(\Delta_x u)|u|^{p-1} + p(p-1)|\nabla_x u|^2|u|^{p-3}u\|_{L^{\frac{2}{\Sigma(s,p)}}L^{\frac{2}{2-s}}} \\
&\qquad\lesssim \frac{1}{R^2}\Big(\|\Delta_x u\|_{Y_s(I)}\|u\|^{p-1}_{Y_s(I)} + \|\nabla_x u\|^2_{Y_s(I)}\|u\|^{p-2}_{Y_s(I)}\Big) \\
&\qquad\lesssim \frac{r^2}{R^2}\|u\|^p_{Y_s(I)}.
\end{aligned}
$$
□

Lemma 5.9. *Let $v(t)$ be a long-time contribution in the Duhamel formula*

$$
v(t_0) = \int_{T_1}^{T_2}\frac{\sin((t-t_0)\sqrt{-\Delta})}{\sqrt{-\Delta}}F(u(t))\,dt.
$$

Then for any $t_0 < T_1 < T_2$, we have

$$
\|v(t_0)\|_{L^\infty(\mathbb{R}^3)} \lesssim (T_1-t_0)^{-2/(p-1)}.
$$

Proof. Using the explicit expression of the wave kernel in dimension 3, we obtain

$$\begin{aligned}\left|\left(\int_{T_1}^{T_2} \frac{\sin((t-t_0)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(t))\,dt\right)(x)\right| &= \left|\int_{T_1}^{T_2}\int_{|y-x|=t-t_0} \frac{1}{4\pi(t-t_0)} F(u(y,t))\,dS(y)\,dt\right| \\ &\lesssim \int_{T_1}^{T_2}\int_{|y-x|=t-t_0} \frac{1}{4\pi(t-t_0)} |u(y,t)|^p\,dS(y)\,dt \\ &\lesssim \int_{T_1}^{T_2}\int_{|y-x|=t-t_0} \frac{1}{(t-t_0)} \frac{1}{|y|^{\frac{2p}{p-1}}}\,dS(y)\,dt.\end{aligned}$$

In the last step, we use the estimate (8) for radial $\dot{H}^{s_p}$ functions. If $|x| \le \frac{1}{2}(T_1 - t_0)$, then on the sphere for the integral we have

$$|y| \ge |t-t_0| - |x| \ge \tfrac{1}{2}(t-t_0).$$

Thus for these small x, we obtain

$$\begin{aligned}\left|\left(\int_{T_1}^{T_2} \frac{\sin((t-t_0)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(t))\,dt\right)(x)\right| &\lesssim \int_{T_1}^{T_2}\int_{|y-x|=t-t_0} \frac{1}{(t-t_0)} \frac{1}{(t-t_0)^{2p/(p-1)}}\,dS(y)\,dt \\ &\lesssim \int_{T_1}^{T_2}\int_{|y-x|=t-t_0} \frac{1}{(t-t_0)^{3+2/(p-1)}}\,dS(y)\,dt \\ &\lesssim \int_{T_1}^{T_2} \frac{(t-t_0)^2}{(t-t_0)^{3+2/(p-1)}}\,dt \\ &\lesssim \int_{T_1}^{T_2} \frac{1}{(t-t_0)^{1+2/(p-1)}}\,dt \\ &\lesssim (T_1-t_0)^{-2/(p-1)}.\end{aligned}$$

On the other hand, if $x \ge \frac{1}{2}(T_1 - t_0)$, by the estimate on radial $\dot{H}^{s_p}$ functions (8) and Lemma 2.17, we have

$$\begin{aligned}\left|\left(\int_{T_1}^{T_2} \frac{\sin((t-t_0)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(t))\,dt\right)(x)\right| &\lesssim \frac{1}{|x|^{2/(p-1)}} \left\|\int_{T_1}^{T_2} \frac{\sin((t-t_0)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(t))\,dt\right\|_{\dot{H}^{s_p}} \\ &\lesssim \frac{1}{(T_1-t_0)^{2/(p-1)}}.\end{aligned}$$

Combining these two cases, we finish our proof. □

Lemma 5.10. *There exists a constant $\kappa = \kappa(p) \in (0,1)$ that depends only on p, so that for each $s \in [s_p, 1)$, there exists an s-admissible pair (q, r), with $q \neq \infty$ and*

$$\frac{\Sigma(s,p)}{2p} = \kappa \cdot 0 + (1-\kappa)\frac{1}{q}, \quad \frac{2-s}{2p} = \kappa\frac{3-2s}{6} + (1-\kappa)\frac{1}{r}.$$

Proof. We will choose $\kappa = 1-3/p \in (0, 0.4)$. Basic computation shows

$$\begin{aligned}
\frac{1}{q} &= \frac{\Sigma(s,p)}{2p(1-\kappa)} = \frac{s+1-(2p-2)(s-s_p)}{6} \in (0,1/3);\\
\frac{1}{r} &= \frac{2-s}{2p(1-\kappa)} - \frac{\kappa}{1-\kappa}\times\frac{3-2s}{6} = \frac{2-s}{6} - \frac{\kappa}{1-\kappa}\times\frac{3-2s}{6}\\
&\in \left(\frac{2-s}{6} - \frac{2}{3}\times\frac{3-2s}{6}, \frac{2-s}{6}\right)\\
&\subseteq \left(\frac{s}{18}, \frac{2-s}{6}\right)\\
&\subseteq (1/36, 1/4).
\end{aligned}$$

Thus we can solve two positive real numbers q, r so that the two identities hold. In addition, we have $q \in (3,\infty)$ and $r \in (4, 36)$. Furthermore, by adding the identities together, we obtain

$$\frac{3-(2p-2)(s-s_p)}{2p} = \kappa\frac{3-2s}{6} + (1-\kappa)\left(\frac{1}{q}+\frac{1}{r}\right).$$

This implies

$$\frac{1}{q}+\frac{1}{r} < \frac{3-(2p-2)(s-s_p)}{2p(1-\kappa)} = \frac{3-(2p-2)(s-s_p)}{6} \le 1/2.$$

Using the same method, one can show $1/q + 3/r = 3/2 - s$. In summary, (q,r) is an s-admissible pair. □

Lemma 5.11. *Given any s-admissible pair (q,r) with $q<\infty$ and three times $t_0 < t_1 < t_2$ in the maximal lifespan I of u, we have*

$$\lim_{T\to\infty}\left\|\int_{t_2}^{T}\frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}}F(u(\tau))\,d\tau\right\|_{L^qL^r([t_0,t_1]\times\mathbb{R}^3)} \le C(\lambda(t_2))^{s-s_p}.$$

The constant C does not depend on t_0, t_1 or t_2.

Proof. By Lemma A.5 and the identity

$$\int_{t_2}^{T}\frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}}F(u(\tau))\,d\tau = S(t-t_2)(u(t_2),\partial_t u(t_2)) - S(t-T)(u(T),\partial_t u(T)),$$

we have

$$\lim_{T\to\infty}\int_{t_2}^{T}\frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}}F(u(\tau))\,d\tau = S(t-t_2)(u(t_2),\partial_t u(t_2))$$

in the space $L^qL^r([t_0,t_1]\times\mathbb{R}^3)$. Thus

$$\begin{aligned}
\lim_{T\to\infty}\left\|\int_{t_2}^{T}\frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}}(u(\tau))\,d\tau\right\|_{L^qL^r([t_0,t_1])} &= \|S(t-t_2)(u(t_2),\partial_t u(t_2))\|_{L^qL^r([t_0,t_1]\times\mathbb{R}^3)}\\
&\lesssim \|(u(t_2),\partial_t u(t_2))\|_{\dot H^s\times\dot H^{s-1}}\\
&\lesssim (\lambda(t_2))^{s-s_p}.
\end{aligned}$$

□

Lemma 5.12. *Suppose $S(A)$ is a nonnegative function defined in $\mathbb{R}^+$ satisfying $S(A) \to 0$ as $A \to \infty$. In addition, there exist $0 < \alpha < \beta < 1$ and $l, \omega > 0$ with*

$$l\alpha + \beta > 1,$$

such that the inequality

$$S(A) \lesssim S(A^\beta) S^l(A^\alpha) + A^{-\omega} \tag{23}$$

is true for each sufficiently large A. Then

$$S(A) \lesssim A^{-\omega}$$

for each sufficiently large A.

Proof. Let us first choose two constants l^- and ω^-, which are slightly smaller than l and ω respectively, such that the inequality $l^-\alpha + \beta > 1$ still holds. By the conditions given, we can find a constant $A_0 \gg 1$, such that the following inequalities hold:

$$\begin{aligned} S(A) &\le \tfrac{1}{2} S(A^\beta) S^{l^-}(A^\alpha) + \tfrac{1}{2} A^{-\omega^-} && \text{if } A \ge A_0, \\ S(A) &< \tfrac{1}{2} && \text{if } A \ge A_0^\alpha. \end{aligned} \tag{24}$$

Using the second inequality above, we know the inequality

$$S(A) \le A^{-\omega_1} \tag{25}$$

holds for all $A \in [A_0^\alpha, A_0]$ if ω_1 is sufficiently small. Fix such a small constant $\omega_1 \le \omega^-$. We will show that the inequality (25) above holds for each $A \ge A_0^\alpha$ by an induction. We already know this is true for $A \in [A_0^\alpha, A_0]$. If $A \in [A_0, A_0^{1/\beta}]$, the inequality (24) implies

$$\begin{aligned} S(A) &\le \tfrac{1}{2} S(A^\beta) S^{l^-}(A^\alpha) + \tfrac{1}{2} A^{-\omega^-} \\ &\le \tfrac{1}{2} (A^\beta)^{-\omega_1} ((A^\alpha)^{-\omega_1})^{l^-} + \tfrac{1}{2} A^{-\omega^-} \\ &\le \tfrac{1}{2} (A^{-\omega_1})^{\beta + l^-\alpha} + \tfrac{1}{2} A^{-\omega_1} \\ &\le A^{-\omega_1}. \end{aligned}$$

Here we use the fact that $A^\alpha, A^\beta \in [A_0^\alpha, A_0]$ if A satisfies our assumption. Conducting an induction, we can show the inequality holds for each $A \in [A_0^{(1/\beta)^n}, A_0^{(1/\beta)^{n+1}}]$ if n is a nonnegative integer. In summary, the inequality (25) is true for each $A \ge A_0^\alpha$. Plugging this back in the original recurrence formula (23), we obtain for sufficiently large A,

$$S(A) \lesssim A^{-\omega_1(\beta + l\alpha)} + A^{-\omega} \lesssim A^{-\min\{\omega_1(\beta + l\alpha), \omega\}},$$

which indicates faster decay than $A^{-\omega_1}$. Iterating the argument if necessary, we gain the decay $S(A) \lesssim A^{-\omega}$ and finish the proof. $\square$

Recurrence formula. Under our setting in this section, given $0<\alpha<\beta<1$ and a small positive constant ε_1, we have the recurrence formula

$$N(A) \lesssim S(A^\beta)S^{p-1}(A^\alpha) + A^{-(\beta-\alpha)\sigma(p)} + A^{-2(1-\beta)}, \tag{26}$$

$$S(A) \lesssim N(A^{1-\varepsilon_1}) + A^{-\sigma_1(p)} \tag{27}$$

for sufficiently large A. The constants $\sigma(p), \sigma_1(p)$ depend on p but nothing else.

Proof of (26). In the following argument, all the space-time norms are taken in $[t, t+d\lambda^{-1}(t)] \times \mathbb{R}^3$:

$$\begin{aligned}
\|P_{>\lambda(t)A}F(u)\|_{Z_s} &\lesssim \lambda(t)^{-(p-1)(s-s_p)} \|P_{>\lambda(t)A}F(u)\|_{L^{\frac{2}{\Sigma(s,p)}}L^{\frac{2}{2-s}}} \\
&\le \lambda(t)^{-(p-1)(s-s_p)} \|P_{>\lambda(t)A}F(u_{\le A^\beta\lambda(t)})\|_{L^{\frac{2}{\Sigma(s,p)}}L^{\frac{2}{2-s}}} \\
&\qquad + \lambda(t)^{-(p-1)(s-s_p)} \|P_{>\lambda(t)A}(F(u)-F(u_{\le A^\beta\lambda(t)}))\|_{L^{\frac{2}{\Sigma(s,p)}}L^{\frac{2}{2-s}}} \\
&= \lambda(t)^{-(p-1)(s-s_p)}(I_1+I_2).
\end{aligned}$$

By Lemma 5.8, we have

$$I_1 \lesssim \left(\frac{A^\beta}{A}\right)^2 \|u\|_{Y_s}^p \lesssim (\lambda(t))^{p(s-s_p)} A^{-2(1-\beta)}.$$

In order to estimate I_2, we have (all unmarked norms are $L^{\frac{2}{\Sigma(s,p)}}L^{\frac{2}{2-s}}([t, t+d\lambda^{-1}(t)] \times \mathbb{R}^3)$ norms)

$$\begin{aligned}
I_2 &\le \left\| P_{>\lambda(t)A}\left[u_{>A^\beta\lambda(t)} \int_0^1 F'(u_{\le A^\beta\lambda(t)} + \tau u_{>A^\beta\lambda(t)})\,d\tau\right]\right\| \\
&\lesssim \left\| u_{>A^\beta\lambda(t)} \int_0^1 F'(u_{\le A^\beta\lambda(t)} + \tau u_{>A^\beta\lambda(t)})\,d\tau\right\| \\
&\lesssim \|I_{2,1}\| + \|I_{2,2}\|.
\end{aligned}$$

Here

$$\begin{aligned}
I_{2,1} &= u_{>A^\beta\lambda(t)} \int_0^1 F'(u_{\le A^\beta\lambda(t)} + \tau u_{>A^\beta\lambda(t)})\,d\tau \\
&\qquad\qquad - u_{>A^\beta\lambda(t)} \int_0^1 F'(u_{A^\alpha\lambda(t)<\cdot\le A^\beta\lambda(t)} + \tau u_{>A^\beta\lambda(t)})\,d\tau, \\
I_{2,2} &= u_{>A^\beta\lambda(t)} \int_0^1 F'(u_{A^\alpha\lambda(t)<\cdot\le A^\beta\lambda(t)} + \tau u_{>A^\beta\lambda(t)})\,d\tau.
\end{aligned}$$

We have

$$I_{2,1} = u_{>A^\beta\lambda(t)} u_{\le A^\alpha\lambda(t)} \times \int_0^1\int_0^1 F''(\tilde{\tau} u_{\le A^\alpha\lambda(t)} + u_{A^\alpha\lambda(t)<\cdot\le A^\beta\lambda(t)} + \tau u_{>A^\beta\lambda(t)})\,d\tau\,d\tilde{\tau}.$$

Applying the bilinear estimate (Lemma 5.6) on the term $u_{>A^\beta\lambda(t)}\, u_{\leq A^\alpha\lambda(t)}$, we obtain

$$\begin{aligned}
\|I_{2,1}\| &\lesssim \|u_{>A^\beta\lambda(t)} u_{\leq A^\alpha\lambda(t)}\|_{L^{\frac{p}{\Sigma(s,p)}} L^{\frac{p}{2-s}}} \\
&\quad \times \left\| \int_0^1 \int_0^1 F''(\tilde{\tau} u_{\leq A^\alpha\lambda(t)} + u_{A^\alpha\lambda(t)<\cdot\leq A^\beta\lambda(t)} + \tau u_{>A^\beta\lambda(t)})\, d\tau\, d\tilde{\tau} \right\|_{L^{\frac{2p}{(p-2)\Sigma(s,p)}} L^{\frac{2p}{(p-2)(2-s)}}} \\
&\lesssim \left[\left(\frac{A^\alpha\lambda(t)}{A^\beta\lambda(t)} \right)^{\sigma(p)} (\lambda(t))^{2(s-s_p)} \right] (\lambda(t))^{(p-2)(s-s_p)} \\
&\lesssim (\lambda(t))^{p(s-s_p)} A^{-(\beta-\alpha)\sigma(p)}.
\end{aligned}$$

On the other hand, we know that, for sufficiently large A,

$$\begin{aligned}
\|I_{2,2}\| &\lesssim \|u_{>A^\beta\lambda(t)}\|_{L^{\frac{2p}{\Sigma(s,p)}} L^{\frac{2p}{2-s}}} \left\| \int_0^1 F'(u_{A^\alpha\lambda(t)<\cdot\leq A^\beta\lambda(t)} + \tau u_{>A^\beta\lambda(t)})\, d\tau \right\|_{L^{\frac{2p}{(p-1)\Sigma(s,p)}} L^{\frac{2p}{(p-1)(2-s)}}} \\
&\lesssim (\lambda(t))^{s-s_p} S(A^\beta) [(\lambda(t))^{(p-1)(s-s_p)} S^{p-1}(A^\alpha)] \\
&\lesssim (\lambda(t))^{p(s-s_p)} S(A^\beta) S^{p-1}(A^\alpha).
\end{aligned}$$

Collecting all terms above, we have

$$\|P_{>\lambda(t)A}(F(u))\|_{Z_s} \lesssim (\lambda(t))^{s-s_p} [S(A^\beta)S^{p-1}(A^\alpha) + A^{-(\beta-\alpha)\sigma(p)} + A^{-2(1-\beta)}].$$

Multiplying both sides by $(\lambda(t))^{s_p-s}$ and taking sup for all $t \in I$, we obtain the first inequality.

Definition 5.13. Given $t_0 \in I$, define t_i recursively for $i \geq 1$ by

$$t_i = t_{i-1} + d\lambda^{-1}(t_{i-1}). \tag{28}$$

By the choice of d, all the t_i are in the maximal lifespan I. (See Proposition 3.5.)

Proof of (27). By the Strichartz estimate and the Duhamel formula (see Lemma A.5), we have

$$\begin{aligned}
\|u_{>\lambda(t_0)A}\|_{Y_s([t_0,t_1])} &= \left\| \int_t^\infty \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} P_{>\lambda(t_0)A} F(u(\tau))\, d\tau \right\|_{Y_s([t_0,t_1])} \\
&\leq \left\| \int_t^{t_2} \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} P_{>\lambda(t_0)A} F(u(\tau))\, d\tau \right\|_{Y_s([t_0,t_1])} \\
&\qquad + \liminf_{T\to\infty} \left\| \int_{t_2}^T \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} P_{>\lambda(t_0)A} F(u(\tau))\, d\tau \right\|_{Y_s([t_0,t_1])} \\
&\lesssim \|P_{>\lambda(t_0)A} F(u)\|_{Z_s([t_0,t_2]\times\mathbb{R}^3)} \\
&\qquad + \liminf_{T\to\infty} \left\| \int_{t_2}^T \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} P_{>\lambda(t_0)A} F(u(\tau))\, d\tau \right\|_{Y_s([t_0,t_1])} \\
&= I_1 + I_2.
\end{aligned}$$

The first term can be dominated by

$$\begin{aligned} I_1 &\lesssim \|P_{>\lambda(t_0)A}F(u)\|_{Z_s([t_0,t_1]\times\mathbb{R}^3)} + \|P_{>\lambda(t_0)A}F(u)\|_{Z_s([t_1,t_2]\times\mathbb{R}^3)} \\ &\lesssim (\lambda(t_0))^{s-s_p}N(A) + (\lambda(t_1))^{s-s_p}N\left(\frac{\lambda(t_0)}{\lambda(t_1)}A\right) \\ &\lesssim (\lambda(t_0))^{s-s_p}N(A^{1-\varepsilon_1}) \end{aligned}$$

for any small positive number ε_1 and sufficiently large $A > A_0(u,\varepsilon_1)$, because $\lambda(t_0)$ and $\lambda(t_1)$ are comparable to each other by the local compactness result (11). Now let us consider the term I_2. First of all, by Lemma 2.17, we have

$$\begin{aligned} &\left\|\int_{t_2}^T \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} P_{>\lambda(t_0)A}F(u(\tau))\,d\tau\right\|_{L^\infty L^2([t_0,t_1]\times\mathbb{R}^3)} \\ &\qquad \lesssim \frac{1}{(\lambda(t_0)A)^{s_p}}\left\|\int_{t_2}^T \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\,d\tau\right\|_{L^\infty_{[t_0,t_1]}\dot H^{s_p}(\mathbb{R}^3)} \\ &\qquad \lesssim \frac{1}{(\lambda(t_0)A)^{s_p}}. \end{aligned}$$

Using Lemma 5.9, we are also able to obtain

$$\begin{aligned} &\left\|\int_{t_2}^T \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} P_{>\lambda(t_0)A}F(u(\tau))\,d\tau\right\|_{L^\infty L^\infty([t_0,t_1]\times\mathbb{R}^3)} \\ &\qquad \lesssim \left\|\int_{t_2}^T \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\,d\tau\right\|_{L^\infty L^\infty([t_0,t_1]\times\mathbb{R}^3)} \\ &\qquad \lesssim (t_2-t_1)^{-2/(p-1)} \\ &\qquad \lesssim (\lambda(t_0))^{2/(p-1)}. \end{aligned}$$

By an interpolation between L^2 and L^∞, we have

$$\begin{aligned} &\left\|P_{>\lambda(t_0)A}\int_{t_2}^T \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\,d\tau\right\|_{L^\infty L^{\frac{6}{3-2s}}([t_0,t_1]\times\mathbb{R}^3)} \\ &\qquad \le \|\cdot\|^{2s/3}_{L^\infty L^\infty([t_0,t_1]\times\mathbb{R}^3)}\|\cdot\|^{(3-2s)/3}_{L^\infty L^2([t_0,t_1]\times\mathbb{R}^3)} \\ &\qquad \lesssim [\lambda(t_0)^{2/(p-1)}]^{2s/3}[(\lambda(t_0)A)^{-s_p}]^{(3-2s)/3} \\ &\qquad = (\lambda(t_0))^{s-s_p}A^{\frac{-s_p(3-2s)}{3}}. \end{aligned}$$

Next, we will use the interpolation again to gain an estimate of the Y_s norm. Let (q,r) be the admissible pair given by Lemma 5.10. Applying Lemma 5.11, we have

$$\lim_{T\to\infty}\left\|\int_{t_2}^T \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\,d\tau\right\|_{L^qL^r([t_0,t_1]\times\mathbb{R}^3)} \lesssim (\lambda(t_2))^{s-s_p} \lesssim (\lambda(t_0))^{s-s_p}.$$

Using this fact and the construction of (q, r), we obtain

$$
\begin{aligned}
I_2 &= \liminf_{T\to\infty} \left\| \int_{t_2}^{T} \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} P_{>\lambda(t_0)A} F(u(\tau))\, d\tau \right\|_{Y_s([t_0,t_1])} \\
&\le \liminf_{T\to\infty} \Bigg(\left\| \int_{t_2}^{T} \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} P_{>\lambda(t_0)A} F(u(\tau))\, d\tau \right\|_{L^\infty L^{\frac{6}{3-2s}}([t_0,t_1]\times\mathbb{R}^3)}^{\kappa(p)} \\
&\qquad\qquad \times \left\| \int_{t_2}^{T} \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} P_{>\lambda(t_0)A} F(u(\tau))\, d\tau \right\|_{L^q L^r([t_0,t_1]\times\mathbb{R}^3)}^{1-\kappa(p)} \Bigg) \\
&\lesssim \left[(\lambda(t_0))^{s-s_p} A^{\frac{-s_p(3-2s)}{3}} \right]^{\kappa(p)} \times \lim_{T\to\infty} \left\| \int_{t_2}^{T} \frac{\sin((\tau-t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\, d\tau \right\|_{L^q L^r}^{1-\kappa(p)} \\
&\lesssim \left[(\lambda(t_0))^{s-s_p} A^{\frac{-s_p(3-2s)}{3}} \right]^{\kappa(p)} (\lambda(t_0))^{(s-s_p)(1-\kappa(p))} \\
&\lesssim (\lambda(t_0))^{s-s_p} A^{\frac{-s_p\kappa(p)(3-2s)}{3}} \\
&\lesssim (\lambda(t_0))^{s-s_p} A^{-\sigma_1(p)}.
\end{aligned}
$$

Here $\sigma_1(p) = \kappa(p)/6$. It depends only on p. Combining our estimates on I_1 and I_2, we finish the proof of the second inequality.

Decay of $S(A)$ and $N(A)$ with applications.

Decay of $S(A)$ and $N(A)$. Plugging the first recurrence formula into the second one, we obtain

$$S(A) \lesssim S(A^{(1-\varepsilon_1)\beta}) S^{p-1}(A^{(1-\varepsilon_1)\alpha}) + A^{-\sigma(p)(1-\varepsilon_1)(\beta-\alpha)} + A^{-2(1-\varepsilon_1)(1-\beta)} + A^{-\sigma_1(p)}.$$

Choose α, β and ε_1 so that

$$(1-\varepsilon_1)\beta = 2/3, \quad (1-\varepsilon_1)\alpha = 1/3, \quad \varepsilon_1 = 1/10000. \tag{29}$$

Then we have

$$S(A) \lesssim S(A^{2/3}) S^{p-1}(A^{1/3}) + A^{-\sigma_2(p)}$$

for sufficiently large A. Here the positive number $\sigma_2(p)$, defined as

$$\sigma_2 = \min\{\sigma(p)/3, \sigma_1(p), 0.6\},$$

depends on p only. Applying Lemma 5.12, we have $S(A) \lesssim A^{-\sigma_2(p)}$ for sufficiently large A. Plugging this in the first recurrence formula, we have $N(A) \lesssim A^{-\sigma_2(p)}$ for large A. Observing that both $S(A)$ and $N(A)$ is uniformly bounded, we know these two decay estimates are actually valid for each $A > 0$. Now let us choose

$$s_1 = \min\Big\{1, s + \frac{99}{100}\sigma_2(p)\Big\},$$

and make the following definition.

Definition 5.14 (local contribution of the Duhamel formula). Assume $t' \in I$. Let us introduce the notation

$$v_{t'}(t) = \int_{t'}^{t'+d\lambda(t')^{-1}} \frac{\sin((\tau - t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\, d\tau;$$
$$\partial_t v_{t'}(t) = -\int_{t'}^{t'+d\lambda(t')^{-1}} \cos((\tau - t)\sqrt{-\Delta}) F(u(\tau))\, d\tau.$$

Estimate on local contribution. Given any $t \le t'$ and integer $k \ge 0$, we know

$$\begin{aligned}
&\|P_{\lambda(t')2^k<\cdot<\lambda(t')2^{k+1}}(v_{t'}(t), \partial_t v_{t'}(t))\|_{\dot H^{s_1}\times\dot H^{s_1-1}} \\
&\qquad\lesssim (\lambda(t')2^k)^{s_1-s}\|P_{\lambda(t')2^k<\cdot<\lambda(t')2^{k+1}}(v_{t'}(t), \partial_t v_{t'}(t))\|_{\dot H^{s}\times\dot H^{s-1}} \\
&\qquad\lesssim (\lambda(t')2^k)^{s_1-s}\|P_{>\lambda(t')2^k}(v_{t'}(t), \partial_t v_{t'}(t))\|_{\dot H^{s}\times\dot H^{s-1}} \\
&\qquad\lesssim (\lambda(t')2^k)^{s_1-s}\|P_{>\lambda(t')2^k}F(u)\|_{Z_s([t',t'+d\lambda(t')^{-1}])} \\
&\qquad\lesssim (\lambda(t')2^k)^{s_1-s}(\lambda(t'))^{s-s_p}N(2^k) \\
&\qquad\lesssim (\lambda(t'))^{s_1-s_p}(2^k)^{s_1-s-\sigma_2(p)}.
\end{aligned}$$

Summing for all $k \ge 0$, we have

$$\|P_{>\lambda(t')}(v_{t'}(t), \partial_t v_{t'}(t))\|_{\dot H^{s_1}\times\dot H^{s_1-1}} \lesssim (\lambda(t'))^{s_1-s_p}.$$

Combining this with the estimate

$$\begin{aligned}
\|P_{\le\lambda(t')}(v_{t'}(t), \partial_t v_{t'}(t))\|_{\dot H^{s_1}\times\dot H^{s_1-1}} &\lesssim (\lambda(t'))^{s_1-s_p}\|P_{\le\lambda(t')}(v_{t'}(t), \partial_t v_{t'}(t))\|_{\dot H^{s_p}\times\dot H^{s_p-1}} \\
&\lesssim (\lambda(t'))^{s_1-s_p},
\end{aligned}$$

we obtain

$$\|(v_{t'}(t), \partial_t v_{t'}(t))\|_{\dot H^{s_1}\times\dot H^{s_1-1}} \lesssim (\lambda(t'))^{s_1-s_p}. \tag{30}$$

Higher regularity. In this subsection we will show that $(u(x,t), \partial_t u(x,t)) \in \dot H^{s_1}\times\dot H^{s_1-1}(\mathbb{R}^3)$ for each $t \in I$. The idea is to deal with the "center" part and the "tail" part individually and then glue them together using Lemma 2.16.

Center estimate. Let us break the Duhamel formula into two pieces:

$$u^{(1)}(t) = \int_t^{t_1} \frac{\sin((\tau - t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\, d\tau,$$
$$u^{(2)}(t) = \int_{t_1}^{\infty} \frac{\sin((\tau - t)\sqrt{-\Delta})}{\sqrt{-\Delta}} F(u(\tau))\, d\tau.$$

Let χ be the characteristic function of the region $\{(x,t) : |x| > d\lambda^{-1}(t_0)/2 + |t-t_1|\}$. We have

$$\begin{aligned}
\|\chi F(u(t))\|_{L^1 L^{\frac{6}{5-2s_1}}([t_1,\infty)\times\mathbb{R}^3)} &= \int_{t_1}^{\infty} \left(\int_{|x|>\frac{d\lambda^{-1}(t_0)}{2}+|t-t_1|} (F(u))^{\frac{6}{5-2s_1}}\, dx \right)^{\frac{5-2s_1}{6}} dt \\
&\lesssim \int_{t_1}^{\infty} \left(\int_{|x|>\frac{d\lambda^{-1}(t_0)}{2}+|t-t_1|} \left(\frac{1}{|x|^{\frac{2p}{p-1}}} \right)^{\frac{6}{5-2s_1}} dx \right)^{\frac{5-2s_1}{6}} dt \\
&\lesssim \int_{t_1}^{\infty} \left(\left| \frac{d\lambda^{-1}(t_0)}{2} + t - t_1 \right|^{-\frac{2p}{p-1}\frac{6}{5-2s_1}+3} \right)^{\frac{5-2s_1}{6}} dt \\
&\lesssim \int_{t_1}^{\infty} \left(\frac{d\lambda^{-1}(t_0)}{2} + t - t_1 \right)^{s_p - s_1 - 1} dt \\
&\lesssim \lambda(t_0)^{s_1 - s_p}.
\end{aligned}$$

By Lemma 2.20, there exists a pair $(\tilde{u}_0, \tilde{u}_1)$ such that

$$\|(\tilde{u}_0, \tilde{u}_1)\|_{\dot{H}^{s_1}\times\dot{H}^{s_1-1}(\mathbb{R}^3)} \lesssim \lambda(t_0)^{s_1-s_p},$$

$$(u^{(2)}(t_0), \partial_t u^{(2)}(t_0)) = (\tilde{u}_0, \tilde{u}_1) \quad \text{in } B\left(0, \frac{d\lambda^{-1}(t_0)}{2}\right).$$

This implies

$$(u(t_0), \partial_t u(t_0)) = (\tilde{u}_0 + u^{(1)}(t_0), \tilde{u}_1 + \partial_t u^{(1)}(t_0)) \quad \text{in } B\left(0, \frac{d\lambda^{-1}(t_0)}{2}\right). \tag{31}$$

By (30), we have

$$\|(u^{(1)}(t_0), \partial_t u^{(1)}(t_0))\|_{\dot{H}^{s_1}\times\dot{H}^{s_1-1}} \lesssim \lambda(t_0)^{s_1-s_p}.$$

Combining this with the $\dot{H}^{s_1}\times\dot{H}^{s_1-1}$ bound of $(\tilde{u}_0, \tilde{u}_1)$, we have

$$\|(\tilde{u}_0 + u^{(1)}(t_0), \tilde{u}_1 + \partial_t u^{(1)}(t_0))\|_{\dot{H}^{s_1}\times\dot{H}^{s_1-1}} \lesssim \lambda(t_0)^{s_1-s_p}. \tag{32}$$

Tail estimate. Let $(u_0', u_1') = \Psi_{d\lambda^{-1}(t_0)/4}(u(t_0), \partial_t u(t_0))$, and

$$\frac{1}{q} = \frac{1}{2} + \frac{1-s_1}{3}.$$

By Theorem 4.1, if $r \geq d\lambda^{-1}(t_0)/4$, we have

$$\begin{aligned}
\left(\int_{r<|x|<4r} (|\nabla u_0'|^q + |u_1'|^q)\, dx \right)^{1/q} &\lesssim \left(\int_{r<|x|<4r} (|\nabla u_0'|^2 + |u_1'|^2)\, dx \right)^{\frac{1}{2}} (r^3)^{\frac{1}{q}-\frac{1}{2}} \\
&\lesssim r^{-(1-s_p)} (r^3)^{(1-s_1)/3} \\
&\lesssim r^{-(s_1-s_p)}.
\end{aligned}$$

Letting $r = 4^k d\lambda^{-1}(t_0)/4$ and summing for all $k \geq 0$, we obtain that the pair (u_0', u_1') is in the space $\dot{W}^{1,q} \times L^q(\mathbb{R}^3)$ with

$$\|(u_0', u_1')\|_{\dot{W}^{1,q} \times L^q(\mathbb{R}^3)} \lesssim (d\lambda(t_0)^{-1}/4)^{-(s_1 - s_p)} \lesssim (\lambda(t_0))^{s_1 - s_p}.$$

By the Sobolev embedding, we have

$$\|(u_0', u_1')\|_{\dot{H}^{s_1} \times \dot{H}^{s_1 - 1}(\mathbb{R}^3)} \lesssim (\lambda(t_0))^{s_1 - s_p}. \tag{33}$$

Combining the center estimate (32) and tail estimate (33) by Lemma 2.16, we have

$$\|(u(t_0), \partial_t u(t_0))\|_{\dot{H}^{s_1} \times \dot{H}^{s_1 - 1}(\mathbb{R}^3)} \lesssim (\lambda(t_0))^{s_1 - s_p}. \tag{34}$$

Conclusion. Now we can finish our proof of Lemma 5.1.

- Case 1 ($s_1 = 1$) The inequality (34) is exactly the energy estimate we are looking for.
- Case 2 ($s_1 < 1$) This means $s_1 = s + 0.99\sigma_2(p)$. As a result, the set

$$K = \left\{ \left(\frac{1}{\lambda(t)^{3/2 - s_p}} u\left(\frac{x}{\lambda(t)}, t\right), \frac{1}{\lambda(t)^{5/2 - s_p}} \partial_t u\left(\frac{x}{\lambda(t)}, t\right) \right) : t \in I \right\}$$

  is precompact in the space $\dot{H}^{s_p} \times \dot{H}^{s_p - 1}$, and bounded in the space $\dot{H}^{s + 0.99\sigma_2(p)} \times \dot{H}^{s - 1 + 0.99\sigma_2(p)}$, thus it is also precompact in the space $\dot{H}^{s + 0.98\sigma_2(p)} \times \dot{H}^{s - 1 + 0.98\sigma_2(p)}$ by an interpolation.

6. Global energy estimate and its corollary

Repeat the recurrence process we described in the previous section starting from the space $\dot{H}^{s_p} \times \dot{H}^{s_p - 1}$. Each time we either obtain the global energy estimate below or gain additional regularity by $0.98\sigma_2(p)$. However, this number depends on p only. As a result, the process has to stop at $\dot{H}^1 \times L^2$ after finite steps.

Proposition 6.1 (global energy estimate). *Let $u(x, t)$ be a minimal blow-up solution. Then $(u(t), \partial_t u(t))$ is in the energy space for each $t \in I$ with*

$$\|(u(t), \partial_t u(t))\|_{\dot{H}^1 \times L^2(\mathbb{R}^3)} \lesssim \lambda(t)^{1 - s_p}. \tag{35}$$

By the local theory, we actually obtain

$$(u(t), \partial_t u(t)) \in C(I; \dot{H}^1(\mathbb{R}^3) \times L^2(\mathbb{R}^3)).$$

Remark 6.2. By Lemma 4.2, we have, for any $0 < a < b < \infty$,

$$(\partial_r w(t), \partial_t w(t)) \in C(I; L^2 \times L^2([a, b])).$$

Self-similar and high-to-low frequency cascade cases. In both two cases, we can choose $t_i \to \infty$ such that $\lambda(t_i) \to 0$. This implies

$$\int_{\mathbb{R}^3} (|\nabla u(x,t_i)|^2 + |\partial_t u(x,t_i)|^2)\,dx \to 0.$$

By the Sobolev embedding, we have

$$\|u\|^{p+1}_{L^{p+1}(\mathbb{R}^3)} \le \|u\|^{p-1}_{L^{\frac{3}{2}(p-1)}(\mathbb{R}^3)} \|u\|^2_{L^6(\mathbb{R}^3)} \lesssim \|u\|^{p-1}_{\dot{H}^{s_p}(\mathbb{R}^3)} \|u\|^2_{\dot{H}^1(\mathbb{R}^3)}. \tag{36}$$

This implies $\|u(t_i)\|^{p+1}_{L^{p+1}(\mathbb{R}^3)} \to 0$. Using the definition of energy we have $E(t_i) \to 0$. On the other hand, we know the energy is a constant. Therefore the energy must be zero.

- *Defocusing case.* It is nothing to say, because in this case an energy zero means that the solution is identically zero.
- *Focusing case.* We can still solve the problem using the following theorem. By the fact that the energy is zero, the theorem claims that u blows up in finite time in both time directions. But this is a contradiction with our assumption $T_+ = \infty$.

Theorem 6.3 (see Theorem 3.1 in [Killip et al. 2014], nonpositive energy implies blowup). *Let $(u_0, u_1) \in (\dot{H}^1 \times L^2) \cap (\dot{H}^{s_p} \times \dot{H}^{s_p - 1})$ be initial data. Assume that (u_0, u_1) is not identically zero and satisfies $E(u_0, u_1) \le 0$. Then the maximal life-span solution to the nonlinear wave equation blows up both forward and backward in finite time.*

Soliton-like solutions in the defocusing case. Now let us consider the soliton-like solutions in the defocusing case. First we have a useful global integral estimate in the defocusing case.

Lemma 6.4 (see [Perthame and Vega 1999]; we use the 3-dimensional case). *Let u be a solution of* (1) *defined in a time interval* $[0, T]$ *with a finite energy*

$$E = \int_{\mathbb{R}^3} \left(\frac{1}{2}|\nabla_x u|^2 + \frac{1}{2}|\partial_t u|^2 + \frac{1}{p+1}|u(x)|^{p+1} \right) dx.$$

For any $R > 0$, we have

$$\begin{aligned}\frac{1}{2R}\int_0^T \int_{|x|<R} (|\nabla u|^2 + |\partial_t u|^2)\,dx\,dt + \frac{1}{2R^2}\int_0^T \int_{|x|=R} |u|^2\,d\sigma_R\,dt \\ + \frac{1}{2R}\frac{2p-4}{p+1}\int_0^T \int_{|x|<R} |u|^{p+1}\,dx\,dt + \frac{p-1}{p+1}\int_0^T \int_{|x|>R} \frac{|u|^{p+1}}{|x|}\,dx\,dt + \frac{2}{R^2}\int_{|x|<R} |u(T)|^2\,dx \\ \le 2E.\end{aligned}$$

Observing that each term on the left-hand side is nonnegative, we can obtain a uniform upper bound for the middle term in the second line above:

$$\int_0^T \int_{|x|>R} \frac{|u|^{p+1}}{|x|}\,dx\,dt \le \frac{2(p+1)}{p-1}E.$$

Letting R approach zero and T approach T_+, we have

$$\int_0^{T_+}\int_{\mathbb{R}^3} \frac{|u|^{p+1}}{|x|}\,dx\,dt \le \frac{2(p+1)}{p-1}E. \tag{37}$$

The energy E here is finite by our estimate (36). On the other hand, recalling our local compactness result Proposition 3.7, we obtain ($T_+ = \infty$)

$$\int_0^{\infty}\int_{\mathbb{R}^3} \frac{|u|^{p+1}}{|x|}\,dx\,dt = \infty.$$

This finishes our discussion in this case.

7. Further estimates in the soliton-like case

Let u be a soliton-like minimal blow-up solution. We will find additional decay of $u(x,t)$ as x tends to infinity. The method used here is similar to the one used in [Kenig and Merle 2011] for the supercritical case. Throughout this section $w(r,t)$, $h(r,t)$, $z_1(r,t)$ and $z_2(r,t)$ are defined as usual using $u(x,t)$. The argument in this section works in both the defocusing and focusing cases. But we are particularly interested in the focusing case, because the soliton-like solutions in the focusing case are the only solutions that still survive at this time.

Setup. Let $\varphi(x)$ be a smooth cutoff function in $\mathbb{R}^3$:

$$\varphi(x)\begin{cases} = 0 & \text{if } |x| \le \frac{1}{2},\\ \in [0,1] & \text{if } \frac{1}{2} \le |x| \le 1,\\ = 1 & \text{if } |x| \ge 1.\end{cases}$$

Then by Proposition 3.6 (compactness of u), $\|\varphi(x/R)u(x,t)\|_{\dot{H}^{s_p}}$ converges to zero uniformly in t as $R \to \infty$. Thus we have a positive function $g(r)$ so that $g(r)$ decreases to zero as r increases to infinity with

$$\|\varphi(x/R)u(x,t)\|_{\dot{H}^{s_p}} \le g(R).$$

This means for each $|x| \ge R$, we have

$$|u(x,t)| = |\varphi(x/R)u(x,t)| \le C\frac{\|\varphi(\cdot/R)u(\cdot,t)\|_{\dot{H}^{s_p}}}{|x|^{2/(p-1)}} \le \frac{Cg(R)}{|x|^{2/(p-1)}}.$$

Definition 7.1.
$$f_\beta(r) = \sup_{t\in\mathbb{R},|x|\ge r} |x|^\beta |u(x,t)|$$

for $\beta \in [2/(p-1), 1)$ and $r > 0$.

This is a nonincreasing function of r defined from $\mathbb{R}^+$ to $[0,\infty)\cup\{\infty\}$. Consider the set

$$U = \{\beta \in [2/(p-1), 1) : f_\beta(r) \to 0 \text{ as } r \to \infty\}.$$

This is not empty, since $2/(p-1)$ is in U. Due to the estimate

$$|x|^\beta |u(x,t)| \le C_p |x|^{\beta-\frac{2}{p-1}} \|u(\cdot,t)\|_{\dot{H}^{s_p}},$$

we know if $\beta \in U$, then $f_\beta(r)$ is a bounded function. By the definition of f_β, we have

$$|u(x,t)| \le \frac{f_\beta(r)}{|x|^\beta} \tag{38}$$

for any time $t \in \mathbb{R}$ and $|x| \ge r$. This is a meaningful inequality as long as $\beta \in U$.

Lemma 7.2. *Suppose u is a soliton-like minimal blow-up solution and $\beta \in U$. Then we have the local energy estimate on $w = ru$*

$$\left(\int_{r_0}^{4r_0} |\partial_t w(r,t_0)|^2 + |\partial_r w(r,t_0)|^2\, dr\right)^{\frac{1}{2}} \le C_p \frac{f_\beta^p(r_0)}{r_0^{p\beta-5/2}} \tag{39}$$

for any $r_0 > 0$ and $t_0 \in \mathbb{R}$. The constant C_p depends on p only.

Proof. Applying Lemma 4.3 to w, we have

$$\left(\int_{r_0}^{4r_0} |z_1(r,t_0)|^2\, dr\right)^{\frac{1}{2}} \le \left(\int_{r_0+M}^{4r_0+M} |z_1(r,t_0+M)|^2\, dr\right)^{\frac{1}{2}} + \left(\int_{r_0}^{4r_0}\left(\int_0^M h(r+t,t_0+t)\, dt\right)^2 dr\right)^{\frac{1}{2}}.$$

Let $M \to \infty$. Using Proposition 4.12 we have

$$\begin{aligned}
\left(\int_{r_0}^{4r_0} |z_1(r,t_0)|^2\, dr\right)^{\frac{1}{2}} &\le \limsup_{M\to\infty}\left(\int_{r_0}^{4r_0}\left(\int_0^M (r+t)F(u(r+t,t_0+t))\, dt\right)^2 dr\right)^{\frac{1}{2}}\\
&\le \limsup_{M\to\infty}\left(\int_{r_0}^{4r_0}\left(\int_0^M (r+t)\left(\frac{f_\beta(r_0)}{(r+t)^\beta}\right)^p dt\right)^2 dr\right)^{\frac{1}{2}}\\
&\lesssim_p \limsup_{M\to\infty}\left(\int_{r_0}^{4r_0}\left(\frac{f_\beta^p(r_0)}{r^{p\beta-2}}\right)^2 dr\right)^{\frac{1}{2}}\\
&\le f_\beta^p(r_0)\left(\int_{r_0}^{4r_0} \frac{1}{r^{2p\beta-4}}\, dr\right)^{\frac{1}{2}}\\
&\lesssim_p f_\beta^p(r_0)\left(\frac{1}{r_0^{2p\beta-5}}\right)^{\frac{1}{2}}\\
&\le f_\beta^p(r_0)\frac{1}{r_0^{p\beta-5/2}}.
\end{aligned}$$

Similarly we have

$$\left(\int_{r_0}^{4r_0} |z_2(r,t_0)|^2\, dr\right)^{\frac{1}{2}} \lesssim \frac{f_\beta^p(r_0)}{r_0^{p\beta-5/2}}.$$

Combining these two estimates we obtain the inequality (39). $\square$

Recurrence formula.

Lemma 7.3. *The function f_β defined above satisfies the recurrence formula*

$$f_\beta(r_0) \le \frac{1}{2}\Big[\Big(\frac{3}{2}\Big)^{1-\beta} + \Big(\frac{1}{2}\Big)^{1-\beta}\Big] f_\beta\Big(\frac{r_0}{2}\Big) + C_p f_\beta^p\Big(\frac{r_0}{2}\Big) r_0^{2-(p-1)\beta}. \tag{40}$$

Proof. We know $w = ru$ is a solution to the one-dimensional wave equation

$$\partial_t^2 w - \partial_r^2 w = r|u|^{p-1}u.$$

Using the explicit formula to solve this equation, we obtain

$$\begin{aligned} r_0 u(r_0, t_0) &= \frac{1}{2}\Big[\Big(r_0 + \frac{r_0}{2}\Big)u\Big(r_0 + \frac{r_0}{2}, t_0 - \frac{r_0}{2}\Big) + \Big(r_0 - \frac{r_0}{2}\Big)u\Big(r_0 - \frac{r_0}{2}, t_0 - \frac{r_0}{2}\Big)\Big] \\ &\qquad + \frac{1}{2}\int_{r_0 - \frac{r_0}{2}}^{r_0 + \frac{r_0}{2}} \partial_t w\Big(r, t_0 - \frac{r_0}{2}\Big)\, dr + \frac{1}{2}\int_0^{\frac{r_0}{2}} \int_{\frac{r_0}{2}+t}^{\frac{3r_0}{2}-t} r|u|^{p-1}u\Big(r, t_0 - \frac{r_0}{2} + t\Big)\, dr\, dt \\ &= I_1 + I_2 + I_3. \end{aligned}$$

By Cauchy–Schwartz and Lemma 7.2, we have

$$\begin{aligned} |I_2| &\le \frac{1}{2}\Big(\int_{\frac{r_0}{2}}^{\frac{3r_0}{2}} \Big|\partial_t w\Big(r, t_0 - \frac{r_0}{2}\Big)\Big|^2 dr\Big)^{\frac{1}{2}} \Big(\int_{\frac{r_0}{2}}^{\frac{3r_0}{2}} 1\, dr\Big)^{\frac{1}{2}} \\ &\le C_p \frac{f_\beta^p(r_0/2)}{r_0^{p\beta - 5/2}} r_0^{1/2} \\ &= C_p f_\beta^p\Big(\frac{r_0}{2}\Big) r_0^{3-p\beta}. \end{aligned}$$

Next we estimate I_3 using the estimate (38)

$$|I_3| \le \frac{1}{2}\int_0^{\frac{r_0}{2}} \int_{\frac{r_0}{2}+t}^{\frac{3r_0}{2}-t} r\Big(\frac{f_\beta(r_0/2)}{r^\beta}\Big)^p dr\, dt \le C_p \int_0^{\frac{r_0}{2}} r_0^2 \frac{f_\beta^p(r_0/2)}{r_0^{p\beta}}\, dt \le C_p f_\beta^p\Big(\frac{r_0}{2}\Big) r_0^{3-p\beta}.$$

At the same time, we know

$$|I_1| \le \frac{1}{2}\Big[\frac{3r_0}{2}\frac{f_\beta(3r_0/2)}{(3r_0/2)^\beta} + \frac{r_0}{2}\frac{f_\beta(r_0/2)}{(r_0/2)^\beta}\Big] = \frac{1}{2}\Big[\Big(\frac{3}{2}\Big)^{1-\beta} f_\beta\Big(\frac{3r_0}{2}\Big) + \Big(\frac{1}{2}\Big)^{1-\beta} f_\beta\Big(\frac{r_0}{2}\Big)\Big] r_0^{1-\beta}.$$

Combining these three terms and dividing both sides of the inequality by $r_0^{1-\beta}$, we obtain (replace r_0 by r)

$$r^\beta |u(r, t_0)| \le \frac{1}{2}\Big[\Big(\frac{3}{2}\Big)^{1-\beta} f_\beta\Big(\frac{3r}{2}\Big) + \Big(\frac{1}{2}\Big)^{1-\beta} f_\beta\Big(\frac{r}{2}\Big)\Big] + C_p f_\beta^p\Big(\frac{r}{2}\Big) r^{2-(p-1)\beta}.$$

Observing that the right-hand side is a nonincreasing function of r, we apply $\sup_{r \ge r_0}$ on both sides and obtain

$$f_\beta(r_0) \le \frac{1}{2}\Big[\Big(\frac{3}{2}\Big)^{1-\beta} f_\beta\Big(\frac{3r_0}{2}\Big) + \Big(\frac{1}{2}\Big)^{1-\beta} f_\beta\Big(\frac{r_0}{2}\Big)\Big] + C_p f_\beta^p\Big(\frac{r_0}{2}\Big) r_0^{2-(p-1)\beta}. \tag{41}$$

This completes the proof because we know $f_\beta(3r_0/2) \le f_\beta(r_0/2)$. □

Decay of $u(x,t)$.

Definition 7.4. Let us define $(2/(p-1) \le \beta < 1)$

$$g(\beta) = \tfrac{1}{2}\big[(\tfrac{3}{2})^{1-\beta} + (\tfrac{1}{2})^{1-\beta}\big] < 1.$$

Lemma 7.5. *If $\beta \in U$, then we have*

$$\left[\beta, \beta + \log_2 \frac{2}{1+g(\beta)}\right) \subseteq U.$$

Proof. Because $f_\beta(r) \to 0$ and $2-(p-1)\beta \le 0$, we know that there exists a large constant R, such that if $r_0 > R$, we have

$$C_p f_\beta^p\Big(\frac{r_0}{2}\Big) r_0^{2-(p-1)\beta} \le \frac{1-g(\beta)}{2} f_\beta\Big(\frac{r_0}{2}\Big).$$

Thus the inequality (40) gives, for $r_0 > R$,

$$f_\beta(r_0) \le \frac{g(\beta)+1}{2} f_\beta\Big(\frac{r_0}{2}\Big).$$

This implies

$$f_\beta(r) \le C r^{\log_2(\frac{g(\beta)+1}{2})}$$

for sufficiently large $r > R'$. As a result, for each $\beta_1 < \beta - \log_2(\frac{g(\beta)+1}{2}) \in (\beta, 1)$, we have

$$|x|^{\beta_1}|u(x,t)| \le f_\beta(|x|)|x|^{\beta_1-\beta} \le C|x|^{\beta_1-\beta+\log_2(\frac{g(\beta)+1}{2})} \to 0$$

as $|x| \to \infty$. This proves the lemma by our definition of f_{β_1} and U. □

Lemma 7.6. *Let U be defined as above. Then we have* $\sup U = 1$.

Proof. If this were false, we could assume $\sup U = \beta_0 < 1$. Then we have for each $\beta \in U$,

$$g(\beta) \le G_0 \doteq \max\Big\{g(\beta_0), g\Big(\frac{2}{p-1}\Big)\Big\} < 1$$

using the convexity of the function g. Thus $\log_2 \dfrac{2}{1+g(\beta)} \ge \log_2 \dfrac{2}{1+G_0} > 0$. By Lemma 7.5, we know

$$\left[\beta, \beta + \log_2 \frac{2}{1+G_0}\right) \subseteq U.$$

This gives us a contradiction as $\beta \to \sup U$. □

The following proposition is the main result of this section.

Proposition 7.7 (decay of u). *Let u be a soliton-like minimal blow-up solution. Then*

$$|u(x,t)| \le \frac{C_1}{|x|} \tag{42}$$

and

$$\int_{r<|x|<4r} (|\nabla u(x,t)|^2 + |\partial_t u(x,t)|^2)\,dx \le C_2 r^{-1}. \tag{43}$$

The constants C_1 and C_2 are independent of t, x or r.

Proof. Let β be a number slightly smaller than 1. Lemma 7.6 guarantees $\beta \in U$. As a result, we can obtain the following estimate using the conclusion of Lemma 7.2:

$$\int_{r_0}^{4r_0} |\partial_r w(r,t_0)|\,dr \le \left(\int_{r_0}^{4r_0} |\partial_r w(r,t_0)|^2\,dr\right)^{\frac{1}{2}} \left(\int_{r_0}^{4r_0} 1\,dr\right)^{\frac{1}{2}} \le \frac{C_p f_\beta^p(r_0)}{r_0^{p\beta-5/2}} r_0^{1/2} \le \frac{C_{p,\beta}}{r_0^{p\beta-3}}.$$

We can choose $\beta \in U$ so that $p\beta - 3 > 0$ by the fact $p > 3$. Thus we have

$$\int_1^\infty |\partial_r w(r,t_0)|\,dr \le C_{p,\beta}. \tag{44}$$

In addition, for $r \le 1$,

$$|w(r,t_0)| = r|u(r,t_0)| \le C_p \|u(t_0)\|_{\dot{H}^{s_p}} r^{1-\frac{2}{p-1}} \le C_p \|u(t_0)\|_{\dot{H}^{s_p}}.$$

Combining these two estimates above, we know that $|w(r,t)|$ is bounded by a universal constant C_1 for each pair (r,t). This gives us the first inequality in the conclusion by the definition $w = ru$. Plugging this in the definition of $f_\beta(r)$, we have

$$f_\beta(r_0) = \sup_{t\in\mathbb{R},|x|\ge r_0} |x|^\beta |u(x,t)| \le \sup_{t\in\mathbb{R},|x|\ge r_0} C_1 |x|^{\beta-1} = C_1 r_0^{\beta-1}.$$

Plugging this in (39), we obtain

$$\left(\int_{r_0}^{4r_0} |\partial_t w(r,t_0)|^2 + |\partial_r w(r,t_0)|^2\,dr\right)^{\frac{1}{2}} \lesssim \frac{1}{r_0^{p-5/2}}. \tag{45}$$

By Lemma 4.2, the combination of this estimate, Proposition 4.12 and the universal decay of u (42) indicates that the second inequality in the lemma is also true. $\square$

8. Death of soliton-like solution

Solitons in the focusing case. In order to kill the soliton-like minimal blow-up solutions, we need to consider the solitons of the wave equation. It turns out that there does not exist any soliton for our equation. The elliptic equation

$$-\Delta W(x) = |W(x)|^{p-1} W(x) \tag{46}$$

does admit a lot of nontrivial radial solutions. However, none of these solutions is in the space $\dot{H}^{s_p}$. Among these solutions we are particularly interested in the following solution W_0 which satisfies the condition $W_0(x) \sim 1/|x|$.

Proposition 8.1. *The elliptic equation* (46) *has a solution* $W_0(x)$ *such that*:

- $W_0(x)$ *is a radial and smooth solution in* $\mathbb{R}^3 \setminus \{0\}$.
- *The point* 0 *is a singularity of* $W_0(x)$.
- *The solution* $W_0(x)$ *is **not** in the space* $\dot{H}^{s_p}(\mathbb{R}^3)$.

- *Its behavior near infinity is given by* ($|x| > R_0$)

$$\left|W_0(x) - \frac{1}{|x|}\right| \le \frac{C}{|x|^{p-2}}, \quad |\nabla W_0(x)| \le \frac{C}{|x|^2}. \tag{47}$$

The next section has a complete discussion of this solution.

Idea to deal with the soliton-like solutions. We will show there does not exist a soliton-like minimal blow-up solution in the focusing case. This conclusion is natural because there is actually no soliton. However, to prove this result is not an easy task. We will use a method developed by T. Duyckaerts et al. as I mentioned at the beginning of this paper. In [Duyckaerts et al. 2013] they use this method to prove the soliton resolution conjecture for radial solutions of the focusing, energy-critical wave equation. The idea is to show that our soliton-like solution has to be so close to the solitons $\pm W_0(x)$ or their rescaled versions that they must be exactly the same. But the soliton we mentioned above is not in the right Sobolev space. This is a contradiction. In order to achieve this goal, we have to be able to understand the behavior of a minimal blow-up solution if it is close to our soliton $W_0(x)$.

Preliminary results. We first recall a lemma proved in [Duyckaerts et al. 2011].

Lemma 8.2 (energy channel). *Let $(v_0, v_1) \in \dot{H}^1 \times L^2$ be a pair of radial initial data. Suppose $v(x,t)$ is the solution of the linear wave equation with the given initial data (v_0, v_1). Let $w(r,t) = rv(r,t)$ as usual. Then for any $R > 0$, either the inequality*

$$\int_{|x|>R+t} (|\nabla v(x,t)|^2 + |\partial_t v(x,t)|^2)\, dx \ge 2\pi \int_R^\infty (|\partial_r w(r,0)|^2 + |\partial_t w(r,0)|^2)\, dr$$

holds for all $t > 0$, or the inequality

$$\int_{|x|>R-t} (|\nabla v(x,t)|^2 + |\partial_t v(x,t)|^2)\, dx \ge 2\pi \int_R^\infty (|\partial_r w(r,0)|^2 + |\partial_t w(r,0)|^2)\, dr$$

holds for all $t < 0$.

Definition 8.3. Let us define ($R > 0$)

$$V_R(x,t) = \begin{cases} W_0(R+|t|) & \text{if } |x| \le R+|t|, \\ W_0(|x|) & \text{if } |x| > R+|t|. \end{cases} \tag{48}$$

Lemma 8.4. *The following space-time norms of $V_R(x,t)$ are both finite for $R > 0$:*

$$\|V_R\|_{Y_{s_p}(\mathbb{R})} < \infty; \quad \|V_R\|_{L^{2p/(p-3)}L^{2p}(\mathbb{R}\times\mathbb{R}^3)} < \infty.$$

Furthermore, if R is sufficiently large, we have the estimate

$$\|V_R\|_{Y_{s_p}(\mathbb{R})} \lesssim R^{\frac{1}{2}-s_p}; \quad \|V_R\|_{L^{2p/(p-3)}L^{2p}(\mathbb{R}\times\mathbb{R}^3)} \lesssim R^{-\frac{1}{2}}. \tag{49}$$

Proof. By the estimate (47) in Proposition 8.1, we have

$$|W_0(x)| \le \frac{C_R}{|x|} \quad \text{if } |x| \ge R.$$

Thus, if $3/r + 1/q < 1$,

$$\begin{aligned}
\|V_R\|_{L^qL^r(\mathbb{R}\times\mathbb{R}^3)} &= \left(\int_{\mathbb{R}}\left(\int_{\mathbb{R}^3} |V_R(x,t)|^r\,dx\right)^{q/r} dt\right)^{1/q} \\
&\lesssim \left(\int_{\mathbb{R}}\left((R+|t|)^3|W_0(R+|t|)|^r + \int_{|x|>R+|t|} |W_0(x)|^r\,dx\right)^{q/r} dt\right)^{1/q} \\
&\lesssim C_R\left(\int_{\mathbb{R}}\left((R+|t|)^{3-r} + \int_{|x|>R+|t|} |x|^{-r}\,dx\right)^{q/r} dt\right)^{1/q} \\
&\lesssim_r C_R\left(\int_{\mathbb{R}}\left((R+|t|)^{3-r}\right)^{q/r} dt\right)^{1/q} \\
&\lesssim_{r,q} C_R(R^{(3-r)q/r+1})^{1/q} \\
&\lesssim_{r,q} C_R R^{\frac{3}{r}+\frac{1}{q}-1}.
\end{aligned}$$

This shows the norms in question are always finite. Furthermore, if R is sufficiently large, we can always choose $C_R = 2$. This finishes our proof by the computation above. ☐

Approximation theory.

Theorem 8.5. *Fix $3 < p < 5$. There exists a constant $\delta_0 > 0$, such that if $\delta < \delta_0$ and we have*

(i) *a function $V(x,t)$ with $\|V(x,t)\|_{Y_{s_p}(I)} < \delta$ (here I is a time interval containing 0), and*

(ii) *a pair of initial data (h_0, h_1) with*

$$\|(h_0,h_1)\|_{\dot H^1\times L^2(\mathbb{R}^3)} < \delta, \quad \|(h_0,h_1)\|_{\dot H^{s_p}\times \dot H^{s_p-1}(\mathbb{R}^3)} < \delta,$$

then the equation

$$\begin{cases} \partial_t^2 h - \Delta h = F(V+h) - F(V), & (x,t)\in\mathbb{R}^3\times I, \\ h|_{t=0} = h_0, \\ \partial_t h|_{t=0} = h_1 \end{cases}$$

has a unique solution $h(x,t)$ on $I\times\mathbb{R}^3$ so that

$$\|h\|_{Y_{s_p}(I)} \le C_p\delta,$$
$$\sup_{t\in I}\|(h,\partial_t h) - (h_L,\partial_t h_L)\|_{\dot H^1\times L^2} \le C_p\delta^{p-1}\|(h_0,h_1)\|_{\dot H^1\times L^2}.$$

Here $(h_L, \partial_t h_L)$ is the solution of the linear wave equation with initial data (h_0, h_1).

Proof. In this proof, C_p represents a constant that depends on p only. In different places C_p may represent different constants. We will also write Y instead of $Y_{s_p}(I)$ for convenience. By the Strichartz estimates, we have

$$\|F(V+h) - F(V)\|_{Z_{s_p}} \le C_p\|h\|_Y(\|h\|_Y^{p-1} + \|V\|_Y^{p-1}),$$
$$\|F(V+h^{(1)}) - F(V+h^{(2)})\|_{Z_{s_p}} \le C_p\|h^{(1)}-h^{(2)}\|_Y(\|h^{(1)}\|_Y^{p-1} + \|h^{(2)}\|_Y^{p-1} + \|V\|_Y^{p-1}).$$

In addition, if we choose a 1-admissible pair $\left(\frac{4p}{9-p}, \frac{4p}{p-3}\right)$, we also have

$$\|F(V+h)-F(V)\|_{L^1L^2} \le C_p \|h\|_{L^{\frac{4p}{9-p}}L^{\frac{4p}{p-3}}} (\|h\|_Y^{p-1} + \|V\|_Y^{p-1}),$$
$$\|F(V+h^{(1)})-F(V+h^{(2)})\|_{L^1L^2} \le C_p \|h^{(1)}-h^{(2)}\|_{L^{\frac{4p}{9-p}}L^{\frac{4p}{p-3}}} (\|h^{(1)}\|_Y^{p-1} + \|h^{(2)}\|_Y^{p-1} + \|V\|_Y^{p-1}).$$

By a fixed point argument, if δ is sufficiently small, we have a unique solution $h(x,t)$ defined on $I \times \mathbb{R}^3$, so that

$$\|h\|_Y \le C_p \delta, \quad \|h\|_{L^{\frac{4p}{9-p}}L^{\frac{4p}{p-3}}} \le C_p \|(h_0, h_1)\|_{\dot{H}^1 \times L^2}.$$

This immediately gives

$$\begin{aligned}\sup_{t\in I} \|(h, \partial_t h) - (h_L, \partial_t h_L)\|_{\dot{H}^1\times L^2} &\le C_p \|F(V+h)-F(V)\|_{L^1L^2} \\ &\le C_p \|h\|_{L^{\frac{4p}{9-p}}L^{\frac{4p}{p-3}}} (\|h\|_Y^{p-1} + \|V\|_Y^{p-1}) \\ &\le C_p \delta^{p-1} \|(h_0,h_1)\|_{\dot{H}^1\times L^2}.\end{aligned}$$

□

Match with $W_0(x)$. Using the estimate (45), we have

$$\int_{r_0}^{4r_0} |\partial_r w(r,t)|\, dr \lesssim \left(\int_{r_0}^{4r_0} |\partial_r w(r,t)|^2\, dr \right)^{\frac{1}{2}} r_0^{1/2} \lesssim \frac{1}{r_0^{p-3}}.$$

This means

$$\int_{r_0}^{\infty} |\partial_r w(r,t)|\, dr \lesssim \frac{1}{r_0^{p-3}}. \tag{50}$$

Thus we know the limit $\lim_{r\to\infty} w(r,t)$ exists for each t. In particular, the limit exists at $t=0$. There are two cases.

(I) If $\lim_{r\to\infty} w(r,0) = 0$. Then in the rest of this section, set $W(x) = 0$. By (50) we have

$$|w(r,0)| \lesssim \frac{1}{r^{p-3}}.$$

Thus

$$|u_0(x) - W(x)| = \frac{1}{|x|} |w(|x|,0)| \lesssim \frac{1}{|x|^{p-2}}.$$

(II) If $\lim_{r\to\infty} w(r,0) \neq 0$. Without loss, let us assume the limit is equal to 1. Otherwise we only need to apply some space-time dilation and/or multiplication by -1 on u. In the rest of this section, set $W(x) = W_0(x)$. By (50), we have

$$|w(r_0,0) - 1| \le \int_{r_0}^{\infty} |\partial_r w(r,0)|\, dr \lesssim \frac{1}{r_0^{p-3}}.$$

Dividing this inequality by r_0, we have

$$\left|u_0(x) - \frac{1}{|x|}\right| \lesssim \frac{1}{|x|^{p-2}}.$$

Combining this with our estimate for $W_0(x)$, we have for large x

$$|u_0(x) - W(x)| \lesssim \frac{1}{|x|^{p-2}}.$$

Identity near infinity.

Theorem 8.6. *Let $W(x) = W_0(x)$ or $W(x) = 0$. Suppose $u(x,t)$ is a global radial solution of the equation* (1) *with initial data $(u_0, u_1) \in \dot{H}^{s_p} \times \dot{H}^{s_p-1}(\mathbb{R}^3)$ satisfying the following conditions.*

(I) *The following inequality holds for each $t \in \mathbb{R}$ and $r > 0$:*

$$\int_{r<|x|<4r} (|\nabla u(x,t)|^2 + |\partial_t u(x,t)|^2)\, dx \le C_1 r^{-1}. \tag{51}$$

(II) *We have $u_0(x)$ and $W(x)$ are very close to each other as $|x|$ is large*:

$$|u_0(x) - W(x)| \lesssim \frac{1}{|x|^{p-2}}. \tag{52}$$

Then there exists $R_0 = R_0(C_1, p) \in (0, +\infty)$ such that the pair $(u_0(x) - W(x), u_1(x))$ is essentially supported in the ball $\bar{B}(0, R_0)$.

Remark 8.7. There are actually two separate theorems, and both can be proved in the same way. If $W(x) = W_0(x)$ (the primary case), then define V_{R_0} as usual in the proof below. Otherwise, if $W(x) = 0$, just make $V_{R_0} = 0$.

Proof. Consider the functions

$$g_0 = \Psi_R(u_0 - W), \quad g_1 = \Psi_R u_1, \quad G(r) = u_0(r) - W(r),$$

for $R \ge R_0$, where the constant R_0 is to be determined later. Choose a small constant $\delta = \delta(p)$, so that it is smaller than the constant δ_0 in Theorem 8.5 and guarantees the number $C_p \delta^{p-1}$ in the conclusion of that theorem is smaller than $\varepsilon(p)$, which is a small number determined later in the argument below. By the condition (51) and the properties of $W(x)$, we know ($R > 1$)

$$\int_{\mathbb{R}^3} (|\nabla g_0|^2 + g_1^2)\, dx \lesssim_{C_1,p} R^{-1};$$

$$\int_{\mathbb{R}^3} \left(|\nabla g_0|^{3(p-1)/(p+1)} + g_1^{3(p-1)/(p+1)}\right) dx \lesssim_{C_1,p} R^{-3(p-3)/(p+1)}.$$

As a result, if $R_0 = R_0(C_1, p)$ is sufficiently large, the following inequalities hold as long as $R \ge R_0$ (we use the Sobolev embedding in order to obtain the second inequality):

$$\|(g_0, g_1)\|_{\dot{H}^1 \times L^2} \le \delta, \quad \|(g_0, g_1)\|_{\dot{H}^{s_p} \times \dot{H}^{s_p-1}} \le \delta, \quad \|V_{R_0}\|_{Y_{s_p}(\mathbb{R})} \le \delta.$$

Let g be the solution of

$$\partial_t^2 g - \Delta g = F(V_{R_0} + g) - F(V_{R_0})$$

with the initial data (g_0, g_1) and $\tilde{g}$ be the solution of the linear wave equation with the same initial data. On the other hand, we know $u(x,t) - W(x)$ is the solution of the equation

$$\partial_t^2 \tilde{u} - \Delta \tilde{u} = F(W + \tilde{u}) - F(W) \tag{53}$$

in the domain $\mathbb{R} \times (\mathbb{R}^3 \setminus \{0\})$ with the initial data $(u_0 - W, u_1)$. Let K be the domain

$$K = \{(x,t) : |x| > |t| + R\}.$$

Considering the fact $W(x) = V_{R_0}(x,t)$ in the region K and the construction of (g_0, g_1), we have

$$u(x,t) - W(x) = g(x,t), \quad \partial_t u(x,t) = \partial_t g(x,t)$$

in the domain K by the finite speed of propagation. Using our assumption (51) and the decay of $W(x)$ at infinity and considering the identity above, we have

$$\lim_{t \to \pm\infty} \int_{|x|>|t|+R} (|\nabla g(x,t)|^2 + |\partial_t g(x,t)|^2)\, dx \to 0. \tag{54}$$

Using Lemma 8.2, without loss of generality, let us assume for all $t > 0$

$$\int_{|x|>R+t} (|\nabla \tilde{g}(x,t)|^2 + |\partial_t \tilde{g}(x,t)|^2)\, dx \geq 2\pi \int_R^\infty (|\partial_r (r g_0(r,0))|^2 + r^2 |g_1(r,0)|^2)\, dr.$$

That is

$$\int_{|x|>R+t} (|\nabla \tilde{g}(x,t)|^2 + |\partial_t \tilde{g}(x,t)|^2)\, dx \geq \frac{1}{2}\left(\int_{|x|>R} (|\nabla g_0|^2 + g_1^2)\, dx\right) - 2\pi R g_0^2(R).$$

Combining this with (54), we have

$$\liminf_{t \to \infty} \left\| (g(x,t), \partial_t g(x,t)) - (\tilde{g}, \partial_t \tilde{g}) \right\|_{\dot{H}^1 \times L^2(|x|>R+t)} \geq \left(\frac{1}{2}\int_{|x|>R} (|\nabla g_0|^2 + g_1^2)\, dx - 2\pi R g_0^2(R)\right)^{\frac{1}{2}}.$$

On the other hand, we know that the inequality

$$\left\| (g(x,t), \partial_t g(x,t)) - (\tilde{g}, \partial_t \tilde{g}) \right\|_{\dot{H}^1 \times L^2} \leq C_p \delta^{p-1} \left\| (g_0, g_1) \right\|_{\dot{H}^1 \times L^2} \leq \varepsilon(p) \|(g_0, g_1)\|_{\dot{H}^1 \times L^2}$$

holds for each $t \in \mathbb{R}$, by Theorem 8.5. Considering both inequalities above, we have

$$\frac{1}{2}\int_{|x|>R} (|\nabla g_0|^2 + g_1^2)\, dx - 2\pi R g_0^2(R) \leq \varepsilon^2(p) \int_{|x|>R} (|\nabla g_0|^2 + g_1^2)\, dx.$$

Thus

$$\int_{|x|>R} (|\nabla g_0|^2 + g_1^2)\, dx \leq \frac{4\pi}{1 - 2\varepsilon^2(p)} R g_0^2(R). \tag{55}$$

We have

$$
\begin{aligned}
|g_0(mR) - g_0(R)| &\le \int_R^{mR} |\partial_r g_0|\,dr \\
&\le \left(\int_R^{mR} |r\partial_r g_0|^2\,dr\right)^{\frac{1}{2}} \left(\int_R^{mR} \frac{1}{r^2}\,dr\right)^{\frac{1}{2}} \\
&\le \left(\frac{1}{4\pi}\int_{|x|>R} (|\nabla g_0|^2 + g_1^2)\,dx\right)^{\frac{1}{2}} \left(\frac{1}{R} - \frac{1}{mR}\right)^{\frac{1}{2}} \\
&\le \left(\frac{Rg_0^2(R)}{1-2\varepsilon^2(p)}\right)^{\frac{1}{2}} \left(1 - \frac{1}{m}\right)^{\frac{1}{2}} R^{-\frac{1}{2}} \\
&\le \left(\frac{1-1/m}{1-2\varepsilon^2(p)}\right)^{\frac{1}{2}} |g_0(R)|.
\end{aligned}
$$

Since $p-2>1$, we can choose $k = k(p) \in \mathbb{Z}^+$ such that $(k+1)/k < p-2$. Let $m = 2^k$. Since

$$
(1-1/m)^{\frac{1}{2}} < 1 - \frac{1}{2m},
$$

we can choose $\varepsilon(p) > 0$ so small that

$$
\left(\frac{1-1/m}{1-2\varepsilon^2(p)}\right)^{\frac{1}{2}} \le 1 - \frac{1}{2m} = 1 - \frac{1}{2^{k+1}}.
$$

Plugging this into our estimate above, we obtain

$$
|g_0(2^k R) - g_0(R)| \le \left(1 - \frac{1}{2^{k+1}}\right)|g_0(R)|.
$$

Thus

$$
|g_0(2^k R)| \ge \frac{1}{2^{k+1}} |g_0(R)|.
$$

By the definition of g_0, this is the same as

$$
|G(2^k R)| \ge \frac{1}{2^{k+1}} |G(R)|.
$$

This inequality holds for all $R \ge R_0$. Now let us consider the value of $G(R_0)$. If $G(R_0) = 0$, let us choose $R = R_0$. Plugging $g_0(R)$ back in (55), we have $(g_0, g_1) = (0,0)$. This means that $(u_0 - W, u_1)$ is supported in $\bar{B}(0, R_0)$ and finishes the proof. If $|G(R_0)| > 0$, then we have

$$
|G(2^{kn} R_0)| \ge \frac{1}{(2^{kn})^{(k+1)/k}} |G(R_0)| > 0
$$

for each positive integer n. This contradicts the condition (52) because $(k+1)/k < p-2$ by our choice of k. □

Remark 8.8. If one feels uncomfortable about the singularity at zero in the equation (53), we could use the following center-cutoff version instead. Let φ be a smooth, radial, nonnegative function satisfying

$$\varphi(x) = \begin{cases} 1 & \text{if } |x| \geq 1, \\ \in [0,1] & \text{if } |x| \in (1/2, 1), \\ 0 & \text{if } |x| \leq 1/2. \end{cases}$$

Then $u(x,t) - \varphi(|x|/R_0)W_0(x)$ is a solution to the equation

$$\begin{cases} \partial_t^2 \tilde{u} - \Delta \tilde{u} = F(\varphi(|x|/R_0)W_0 + \tilde{u}) + \Delta(\varphi(|x|/R_0)W_0(x)), (x,t) \in \mathbb{R}^3 \times \mathbb{R}, \\ \tilde{u}|_{t=0} = u_0 - \varphi(|x|/R_0)W_0 \in \dot{H}^{s_p}(\mathbb{R}^3), \\ \partial_t \tilde{u}|_{t=0} = u_1 \in \dot{H}^{s_p - 1}(\mathbb{R}^3). \end{cases}$$

For any $T > 0$, we know

$$\|\varphi(|x|/R_0)W_0(x)\|_{Y_{s_p}([-T,T])} < \infty, \quad \|\Delta(\varphi(|x|/R_0)W_0(x))\|_{Z_{s_p}([-T,T])} < \infty.$$

In addition, the function $\Delta(\varphi(|x|/R_0)W_0(x)) = -F(W_0(x))$ in the region K. We can do the argument as usual in the proof above but avoid the singularity at zero with this new cutoff version of the equation (53). This method also works in the proof of Theorem 8.9, which will be introduced in the next subsection.

Application of the theorem. Now apply Theorem 8.6 to our soliton-like minimal blow-up solution. All the conditions are satisfied by our earlier argument. Thus $(u_0(x) - W(x), u_1(x))$ is supported in the closed ball of radius R_0 centered at the origin. In particular, because R_0 depends only on the constant C_1 and p, the same R_0 also works for other time t as long as the condition (52) is true at that time. But by the finite speed of propagation, we know $(u(x,t) - W(x), \partial_t u(x,t))$ is actually compactly supported in $\bar{B}(0, R_0 + |t|)$ at each time t. This means the condition (52) is always true at any given time. Thus the pair $(u(x,t) - W(x), \partial_t u(x,t))$ is essentially supported in the cylinder $\bar{B}(0, R_0) \times \mathbb{R}$.

Local radius analysis. Let us define the essential radius of the support of $(u(x,t) - W(x), \partial_t u(x,t))$ at time t as

$$R(t) = \min\{R \geq 0 : (u(x,t) - W(x), \partial_t u(x,t)) = (0,0) \text{ holds for } |x| > R\}.$$

This is well-defined for our minimal blow-up solution. Actually $R(t) \leq R_0$ holds for any $t \in \mathbb{R}$.

Theorem 8.9 (behavior of "compactly supported" solutions). *Let $W(x) = W_0(x)$ or $W(x) = 0$. Let $u(x,t)$ be a radial solution of the equation* (1) *in a time interval I containing* 0, *so that*

(I) $(u(x,t), \partial_t u(x,t)) \in C(I; \dot{H}^1(\mathbb{R}^3) \times L^2(\mathbb{R}^3))$.

(II) *The pair $(u(x,0) - W(x), \partial_t u(x,0))$ is compactly supported with an essential radius of support $R(0) > R_1 > 0$.*

Then there exists a constant $\tau = \tau(R_1, p)$, such that

$$R(t) = R(0) + |t|$$

holds either for each $t \in [0, \tau] \cap I$ or for each $t \in [-\tau, 0] \cap I$.

Remark 8.10. If $W(x) = W_0(x)$ (the primary case), then define V_{R_1} as usual in the proof. Otherwise if $W(x) = 0$, just make $V_{R_1} = 0$. In this case we can choose $\tau = \infty$. In the proof we use the notation (u_0, u_1) for the initial data $(u(x,0), \partial_t u(x,0))$.

Proof. By Lemma 8.4, we have $\|V_{R_1}\|_{Y_{s_p}(\mathbb{R})} < \infty$. Thus we can choose $\tau = \tau(R_1, p) > 0$ such that $\|V_{R_1}\|_{Y_{s_p}([-\tau,\tau])} < \delta$. Here δ is a small constant so that we can apply Theorem 8.5 and make the number $C_p \delta^{p-1}$ less than $1/100$ in that theorem. If $\varepsilon < R(0) - R_1$, let us consider a pair of initial data (g_0, g_1) for each $R \in (R(0) - \epsilon, R(0))$,

$$g_0 = \Psi_R(u_0 - W), \quad g_1 = \Psi_R u_1.$$

This pair $(g_0(x), g_1(x))$ is nonzero by the definition of $R(0)$. By our assumptions on (u_0, u_1), we know the inequalities

$$\|(g_0, g_1)\|_{\dot{H}^1 \times L^2} < \delta, \quad \|(g_0, g_1)\|_{\dot{H}^{s_p} \times \dot{H}^{s_p - 1}} < \delta$$

hold for each $R \in (R(0) - \epsilon, R(0))$ as long as ε is sufficiently small. (In order to obtain the second inequality we use the Sobolev embedding.) Furthermore, we have

$$\begin{aligned}
|g_0(R)| &= \left| g_0(R(0)) - \int_R^{R(0)} \partial_r g_0(r)\, dr \right| \le \int_R^{R(0)} |\partial_r g_0(r)|\, dr \\
&\le \left(\int_R^{R(0)} r^2 |\partial_r g_0(r)|^2\, dr \right)^{\frac{1}{2}} \left(\int_R^{R(0)} \frac{1}{r^2}\, dr \right)^{\frac{1}{2}} \\
&\le \left(\int_R^{R(0)} r^2 |\partial_r g_0(r)|^2\, dr \right)^{\frac{1}{2}} \left(\frac{R(0) - R}{R(0) R} \right)^{\frac{1}{2}} \\
&\le \left(\frac{\varepsilon}{R(0) R} \int_R^{R(0)} r^2 |\partial_r g_0(r)|^2\, dr \right)^{\frac{1}{2}}.
\end{aligned}$$

Thus

$$R g_0^2(R) \le \frac{\varepsilon}{R(0)} \int_R^{R(0)} r^2 |\partial_r g_0(r)|^2\, dr \le \frac{\varepsilon}{4\pi R(0)} \int_{R<|x|<R(0)} (|\nabla g_0(x)|^2 + |g_1(x)|^2)\, dx.$$

If ε is sufficiently small, we can apply Lemma 4.2 to obtain

$$\int_R^{R(0)} \left[|\partial_r (r g_0(r))|^2 + r^2 g_1(r)^2 \right] dr \ge \frac{0.99}{4\pi} \int_{R<|x|<R(0)} (|\nabla g_0(x)|^2 + |g_1(x)|^2)\, dx.$$

Let $\tilde{g}(x,t)$ be the solution to the linear wave equation with the initial data (g_0, g_1). By Lemma 8.2,

$$\begin{aligned}
\int_{|x|>R+|t|} \left(|\nabla \tilde{g}(x,t)|^2 + |\partial_t \tilde{g}(x,t)|^2 \right) dx &\ge 2\pi \int_R^{\infty} \left[|\partial_r (r g_0(r))|^2 + r^2 |g_1(r)|^2 \right] dr \\
&= 2\pi \int_R^{R(0)} \left[|\partial_r (r g_0(r))|^2 + r^2 |g_1(r)|^2 \right] dr \\
&\ge 0.49 \int_{R<|x|<R(0)} (|\nabla g_0(x)|^2 + |g_1(x)|^2)\, dx
\end{aligned}$$

holds either for each $t \geq 0$ or for each $t \leq 0$. Without loss of generality, let us choose $t \geq 0$; then we have

$$\|(\tilde{g}(x,t), \partial_t \tilde{g}(x,t))\|_{\dot{H}^1 \times L^2(|x|>R+t)} \geq 0.7 \|(g_0, g_1)\|_{\dot{H}^1 \times L^2(\mathbb{R}^3)}. \tag{56}$$

Let g be the solution of the equation

$$\begin{cases} \partial_t^2 g - \Delta g = F(V_{R_1} + g) - F(V_{R_1}), & (x,t) \in \mathbb{R}^3 \times [-\tau, \tau], \\ g|_{t=0} = g_0, \\ \partial_t g|_{t=0} = g_1. \end{cases}$$

By Theorem 8.5, we have

$$\|(g(x,t), \partial_t g(x,t)) - (\tilde{g}(x,t), \partial_t \tilde{g}(x,t))\|_{\dot{H}^1 \times L^2} \leq 0.01 \|(g_0, g_1)\|_{\dot{H}^1 \times L^2(\mathbb{R}^3)}$$

for each $t \in [-\tau, \tau]$. Combining this with (56), for $t \in [0, \tau]$ we obtain

$$\|(g(x,t), \partial_t g(x,t))\|_{\dot{H}^1 \times L^2(|x|>R+t)} \geq 0.69 \|(g_0, g_1)\|_{\dot{H}^1 \times L^2(\mathbb{R}^3)}. \tag{57}$$

In addition, we know $u(x,t) - W(x)$ is the solution of equation

$$\begin{cases} \partial_t^2 \tilde{u} - \Delta \tilde{u} = F(W(x) + \tilde{u}) - F(W(x)), \\ \tilde{u}|_{t=0} = u_0 - W, \\ \partial_t \tilde{u}|_{t=0} = u_1 \end{cases}$$

in $(\mathbb{R}^3 \setminus \{0\}) \times I$. The initial data of these two equations mentioned above is the same in the region $\{x : |x| \geq R\}$ and the nonlinear part is the same function in the region

$$K = \{(x,t) : |x| > R + t,\ t \in [0, \tau] \cap I\}.$$

Thus by the finite speed of propagation, we have $g(x,t) = u(x,t) - W(x)$ and $\partial_t g(x,t) = \partial_t u(x,t)$ in K. Plugging this in (57), we obtain

$$\|(u(x,t) - W(x), \partial_t u(x,t))\|_{\dot{H}^1 \times L^2(|x|>R+t)} \geq 0.69 \|(g_0, g_1)\|_{\dot{H}^1 \times L^2(\mathbb{R}^3)}$$

for each $t \in I \cap [0, \tau]$. Since $R < R(0)$, we know the right-hand side of the inequality above is positive by the definition of essential radius of support. Thus we have

$$R(t) \geq R + |t| \tag{58}$$

for all $t \in [0, \tau] \cap I$. Letting $R \to R(0)^-$, we obtain $R(t) \geq R(0) + |t|$. By the finite speed of propagation, we have $R(t) = R(0) + |t|$. □

Remark 8.11. For each $R \in (R(0) - \varepsilon, R(0))$, we know that the inequality (58) above holds either in the positive or negative time direction. It may work in different directions as we choose different values of R. However, we can always choose a sequence $R_i \to R(0)^-$ such that the inequality works in the same time direction for all the R_i. This is sufficient for us to conclude the theorem.

End of soliton-like solution. Now let us show $R(0) = 0$. If it were not zero, let $R_1 = R(0)/2$, and then apply Theorem 8.9. We have (without loss of generality) $R(t) = R(0) + t$ for each $t \in [0, \tau]$. Applying Theorem 8.9 again at $t = \tau$, we obtain

$$R(t) = R(0) + \tau + (t - \tau) = R(0) + t$$

for $t \in [\tau, 2\tau]$, because

(i) The same constant τ works by the inequality $R(\tau) > R(0) > R_1$.

(ii) The theorem may only work in the positive time direction, since we know the radius of support $R(t)$ decreases in the other direction.

Repeating this process, we have for each $t > 0$,

$$R(t) = R(0) + t.$$

But it is impossible since $R(t)$ is uniformly bounded by R_0. Therefore we must have $R(0) = 0$. But this means either $u_0 = W_0(x) \notin \dot{H}^{s_p}(\mathbb{R}^3)$ or $(u_0, u_1) = (0, 0)$. This is a contradiction.

9. The solution of the elliptic equation

In this section we will consider the elliptic equation

$$-\Delta W(x) = |W(x)|^{p-1} W(x), \tag{59}$$

and prove Proposition 8.1. It has infinitely many solutions. For example,

$$W_1(x) = C|x|^{-2/(p-1)}$$

is a solution if we choose an appropriate constant C. Since we are interested in radial solutions of this elliptic equation, we can assume $W(x) = y(|x|)$. Here the function $y(r)$ satisfies the following equation in $(0, \infty)$:

$$y''(r) + \frac{2}{r} y'(r) + |y|^{p-1} y(r) = 0. \tag{60}$$

Let us first show that the solution $W_0(x)$ we mentioned earlier in this paper exists.

Existence of $W_0(x)$.

The idea. We are seeking a solution with the property $W_0(x) \simeq 1/|x|$ as x is large. That is equivalent to $y(r) \simeq 1/r$. Let us define $\rho(r) = r y(r)$; then $\rho(r)$ satisfies

$$\rho''(r) = -\frac{F(\rho)}{r^{p-1}}, \quad F(\rho) = |\rho|^{p-1}\rho.$$

We expect $\rho(r) \simeq 1$ for large r, thus let us assume $\rho(r) = \phi(r) + 1$. The corresponding equation for $\phi(r)$ is given as

$$\phi''(r) = -\frac{F(\phi + 1)}{r^{p-1}}.$$

We will show the following facts:

(I) This equation has a solution in the interval $[R,\infty)$ with boundary conditions at infinity $\phi(+\infty) = \phi'(+\infty) = 0$, by a fixed-point argument.

(II) We can expand the domain of this solution to $\mathbb{R}^+$.

The fixed-point argument. Let us consider the metric space

$$K = \{\phi : \phi \in C([R,\infty);[-1,1]),\ \lim_{r\to+\infty} \phi(r) = 0\}$$

with the distance $d(\phi_1,\phi_2) = \sup_r |\phi_1(r) - \phi_2(r)|$. One can check K is complete. Let us define a map $L : K \to K$ by

$$L(\phi)(r) = \int_r^\infty \left(\int_s^\infty \left(-\frac{F(\phi(t)+1)}{t^{p-1}} \right) dt \right) ds.$$

We have

$$|L(\phi)(r)| \le \int_r^\infty \left(\int_s^\infty \frac{2^p}{t^{p-1}}\, dt \right) ds \le \frac{C_p}{r^{p-3}},$$

$$|L(\phi_1)(r) - L(\phi_2)(r)| \le C_p \int_r^\infty \left(\int_s^\infty \frac{d(\phi_1,\phi_2)}{t^{p-1}}\, dt \right) ds \le C_p \frac{d(\phi_1,\phi_2)}{r^{p-3}}.$$

Thus if $R > R(p)$ is a sufficiently large number, then L is a contraction map from K to itself. As a result, there exists a unique fixed point $\phi_0(r)$. This gives us a classic smooth solution of the ODE in $[R,\infty)$. We have $\phi_0(r) \lesssim r^{3-p}$ and its derivative $\phi_0'(r)$ satisfies

$$|\phi_0'(r)| = \left| \int_r^\infty \frac{F(\phi_0(t)+1)}{t^{p-1}}\, dt \right| \le \frac{C_p}{r^{p-2}}.$$

Expansion of the solution. Now let us solve the ODE backward from $r = R$. We need to show it will never break down before we approach $r = 0$. Actually we have

$$\frac{d}{dr}\left(\frac{|\phi_0+1|^{p+1}}{p+1} + \frac{r^{p-1}|\phi_0'|^2}{2} \right) = \frac{p-1}{2} r^{p-2} |\phi_0'|^2 \ge 0.$$

Thus we have that the inequality

$$\frac{|\phi_0(r)+1|^{p+1}}{p+1} + \frac{r^{p-1}|\phi_0'(r)|^2}{2} \le \frac{|\phi_0(R)+1|^{p+1}}{p+1} + \frac{R^{p-1}|\phi_0'(R)|^2}{2}$$

holds for all $0 < r \le R$ as long as the solution still exists at r. But this implies the solution will never break down at a positive r.

Properties of the solution. Now we can define

$$W_0(x) = \frac{\phi_0(|x|)+1}{|x|}.$$

This is a C^2, radial solution of our elliptic equation (59) for $|x| > 0$. Furthermore, we have for large x

$$\left|W_0(x) - \frac{1}{|x|}\right| = \frac{|\phi_0(|x|)|}{|x|} \le \frac{C_p}{|x|^{p-2}}, \quad |\nabla W_0(x)| = \left|\frac{r\phi_0'(r) - \phi_0(r) - 1}{r^2}\right|_{r=|x|} \le \frac{C_p}{|x|^2}.$$

Now the remaining task is to show $W_0(x)$ is not in the space $\dot{H}^{s_p}$. This implies $W_0(x)$ must have a singularity at 0. It turns out that it is not trivial. For instance, if we repeat the argument as above in the case $p = 5$, then the solution we obtain will be a smooth function in the whole space, as

$$W(x) = \frac{\sqrt{3}}{(1 + 3|x|^2)^{1/2}}.$$

Radial $\dot{H}^{s_p}$ solution does not exist. The following theorem shows that any nontrivial radial solution of our elliptic equation is not in the space $\dot{H}^{s_p}(\mathbb{R}^3)$. In particular, $W_0(x)$ is not in the space $\dot{H}^{s_p}(\mathbb{R}^3)$. Actually we have $\limsup_{x\to 0^+} |x|^\theta |W_0(x)| > 0$ by the argument below. This gives us a singularity at zero.

Theorem 9.1. *If* $3 < p < 5$, *then a radial* $\dot{H}^{s_p}(\mathbb{R}^3)$ *solution to the elliptic equation*

$$-\Delta W(x) = |W(x)|^{p-1} W(x)$$

must be the zero solution.

Remark 9.2. We always assume the function $y(r)$ has two continuous derivatives at any $r > 0$ in the proof below. Actually we can show any radial $\dot{H}^{s_p}$ solution of the elliptic equation must be in the space $C^2(\mathbb{R}^3 \setminus \{0\})$. First of all, a radial $\dot{H}^{s_p}$ function must be continuous except for $x = 0$. Using this fact and the regularity theory on the elliptic equation, we have the solution is C^2 except for $x = 0$.

Proof. The proof consists of three steps.

(I) (introduction to $r^\theta y(r)$) We assume $W(x) = y(|x|)$. The function $y(r)$ defined in $\mathbb{R}^+$ is a C^2 solution of

$$y''(r) + \frac{2}{r} y'(r) + |y|^{p-1} y(r) = 0.$$

Let us define another $C^2(\mathbb{R}^+)$ function

$$v(r) = r^\theta y(r), \quad \theta = \frac{2}{p-1}.$$

If $W(x) = y(|x|)$ is in the space $\dot{H}^{s_p}$, we then have $\lim_{r\to 0^+} v(r) = \lim_{r\to+\infty} v(r) = 0$ by Lemma A.7. Plugging $y(r) = r^{-\theta} v(r)$ in the equation for $y(r)$, we obtain an equation for $v(r)$,

$$r^2 v''(r) + \frac{2(p-3)}{p-1} r v'(r) - \frac{2(p-3)}{(p-1)^2} v(r) + |v|^{p-1} v(r) = 0.$$

Multiplying both sides by $v'(r)$, we obtain

$$\frac{d}{dr}\left(r^2 \frac{|v'(r)|^2}{2} - \frac{p-3}{(p-1)^2} v^2(r) + \frac{|v(r)|^{p+1}}{p+1}\right) = \frac{5-p}{p-1} r |v'(r)|^2 \ge 0. \tag{61}$$

(II) (the lower limit) If $v(r)$ is not the zero function, then the inequality

$$\liminf_{r\to+\infty} r^2|v'(r)|^2 > 0 \tag{62}$$

holds. If it failed, by considering the integral of (61) in the interval (ε, M) and letting $\varepsilon \to 0^+$ and $M \to +\infty$, we would have

$$\frac{5-p}{p-1}\int_0^\infty r|v'(r)|^2\,dr \le 0.$$

This means $v'(r) = 0$ everywhere, so $v(r) = 0$. But we assume it is not the zero function.

(III) (conclusion) If $W(x)$ were not identically zero, then $v(r)$ would be a nonzero function. By the limit (62), there exist $C > 0$ and $r_1 > 0$, such that if $r \in (r_1, \infty)$, the inequality $r^2|v'(r)|^2 > C$ holds. In other words, we have $|v'(r)| > \sqrt{C}r^{-1}$. This means $v'(r)$ does not change its sign in the interval (r_1, ∞) since it is a continuous function. Combining this fact with the lower bound of $|v'(r)|$, we know the limit of $v(r)$ does not exist at ∞. This gives us a contradiction. □

Further properties of the function $W_0(x)$. In this subsection, we will discover some additional properties of the soliton $W_0(x)$. Assume that $y(r)$ and $v(r)$ are defined in the same manner as the previous subsection.

- *$W_0(x)$ is a positive solution.* If this were not true, we could assume that $v(r_0) = 0$ for some $r_0 > 0$, because we know $v(r) > 0$ for sufficiently large r. Then by (61), we obtain

$$r^2\frac{|v'(r)|^2}{2} - \frac{p-3}{(p-1)^2}v^2(r) + \frac{|v(r)|^{p+1}}{p+1} \ge r_0^2\frac{|v'(r_0)|^2}{2} > 0 \tag{63}$$

for each $r > r_0$. However, the decay of $W_0(x)$ implies (if r is large) that

$$|v(r)| \lesssim r^{\theta-1}, \quad |v'(r)| = |\theta r^{\theta-1}y(r) + r^\theta y'(r)| \lesssim r^{\theta-2}.$$

This gives us a contradiction if we consider the limit of the left hand in the inequality (63) using these estimates.

- *$W_0(x)$ is smooth in $\mathbb{R}^3 \setminus \{0\}$.* Due to the fact that the function F is smooth in $\mathbb{R}^+$, a direct corollary follows that the function $W_0(x)$ is smooth everywhere except for $x = 0$.

Appendix

The Duhamel formula.

Lemma A.1. *Let $\frac{1}{2} < s \le 1$. If K is a compact subset of $\dot{H}^s \times \dot{H}^{s-1}$ with an s-admissible pair (q, r) so that $q \ne \infty$, then for each $\varepsilon > 0$, there exist two constants $M, \delta > 0$ such that*

$$\|S(t)(u_0,u_1)\|_{L^qL^r(J\times\mathbb{R}^3)} + \|S(t)(u_0,u_1)\|_{L^qL^r([M,\infty)\times\mathbb{R}^3)} + \|S(t)(u_0,u_1)\|_{L^qL^r((-\infty,M]\times\mathbb{R}^3)} < \varepsilon$$

holds for any $(u_0, u_1) \in K$ and any time interval J with a length $|J| \le \delta$.

Proof. Given $(u_0, u_1) \in \dot{H}^s \times \dot{H}^{s-1}$, it is clear that we are able to find $M, \delta > 0$ so that the inequality holds for this particular pair of initial data and any interval J with a length $|J| \le \delta$ by the fact $q < \infty$ and the Strichartz estimate

$$\|S(t)(u_0, u_1)\|_{L^q L^r(\mathbb{R}\times\mathbb{R}^3)} < \infty.$$

If K is a finite set, then we can find M and δ so that they work for each pair in K by taking a maximum over all M and a minimum over all δ. In the general case, we can just choose a finite subset $\{(u_{0,i}, u_{1,i})\}_{i=1,2,\ldots,n}$ of K such that for each $(u_0, u_1) \in K$, there exists a positive integer i with $1 \le i \le n$ and $\|S(t)(u_0 - u_{0,i}, u_1 - u_{1,i})\|_{L^q L^r(\mathbb{R}\times\mathbb{R}^3)} \le C\|(u_0 - u_{0,i}, u_1 - u_{1,i})\|_{\dot{H}^s\times\dot{H}^{s-1}} < 0.01\varepsilon$; and then use our result for a finite subset. □

Lemma A.2 (the Duhamel formula). *Let $u(x,t)$ be almost periodic modulo scaling in the interval $I = (T_-, \infty)$, namely the set*

$$K = \left\{ \left(\frac{1}{\lambda(t)^{3/2-s_p}} u\left(\frac{x}{\lambda(t)}, t\right), \frac{1}{\lambda(t)^{5/2-s_p}} \partial_t u\left(\frac{x}{\lambda(t)}, t\right) \right) : t \in I \right\}$$

is precompact in the space $\dot{H}^{s_p} \times \dot{H}^{s_p-1}(\mathbb{R}^3)$. Then for any time $t_0 \in \mathbb{R}$, any bounded closed interval $[a,b]$ and any s_p-admissible pair (q,r) with $q < \infty$, we have

$$\lim_{T\to+\infty} \|S(t-T)(u(T), \partial_t u(T))\|_{L^q L^r([a,b]\times\mathbb{R}^3)} = 0,$$

$$\textit{weak} \lim_{T\to+\infty} S(t_0 - T) \begin{pmatrix} u(T) \\ \partial_t u(T) \end{pmatrix} = 0.$$

Proof. We have

$$\begin{aligned} \|S(t-T)(u(T), \partial_t u(T))\|_{L^q L^r([a,b]\times\mathbb{R}^3)} &= \|S(t)(u(T), \partial_t u(T))\|_{L^q L^r([a-T,b-T]\times\mathbb{R}^3)} \\ &= \|S(t)(u_0^{(T)}, u_1^{(T)})\|_{L^q L^r([\lambda(T)(a-T),\lambda(T)(b-T)]\times\mathbb{R}^3)}; \end{aligned}$$

here

$$(u_0^{(T)}, u_1^{(T)}) = \left(\frac{1}{\lambda(T)^{3/2-s_p}} u\left(\frac{\cdot}{\lambda(T)}, T\right), \frac{1}{\lambda(T)^{5/2-s_p}} \partial_t u\left(\frac{\cdot}{\lambda(T)}, T\right) \right).$$

Given $\varepsilon > 0$, let M, δ be the constants as in Lemma A.1. It is clear that if T is sufficiently large, we have either ($\lambda(T)$ is small)

$$\lambda(T)(b-T) - \lambda(T)(a-T) = (b-a)\lambda(T) < \delta,$$

or ($\lambda(T)$ is large)

$$\lambda(T)(b-T) < -M.$$

In either case, by Lemma A.1 we have $\|S(t-T)(u(T), \partial_t u(T))\|_{L^q L^r([a,b]\times\mathbb{R}^3)} < \varepsilon$. This completes the proof of the first limit. In order to obtain the second limit, we only need to choose $t_1 \in (t_0, +\infty)$, set $[a,b] = [t_0, t_1]$ and apply Lemma A.4 below using the first limit and the identity

$$S(t-t_0)\left[S(t_0 - T) \begin{pmatrix} u(T) \\ \partial_t u(T) \end{pmatrix} \right] = S(t-T) \begin{pmatrix} u(T) \\ \partial_t u(T) \end{pmatrix}. \qquad □$$

Remark A.3. We can obtain the similar result in the negative time direction using exactly the same argument. This implies the corresponding Duhamel formula in the negative time direction.

- Soliton-like case or high-to-low frequency cascade case

$$\lim_{T\to-\infty} \|S(t-T)(u(T),\partial_t u(T))\|_{L^q L^r([a,b]\times\mathbb{R}^3)} = 0,$$

$$\text{weak} \lim_{T\to-\infty} S(t_0-T)\begin{pmatrix} u(T) \\ \partial_t u(T)\end{pmatrix} = 0.$$

- Self-similar case (let $a, t_0 > 0$)

$$\lim_{T\to 0^+} \|S(t-T)(u(T),\partial_t u(T))\|_{L^q L^r([a,b]\times\mathbb{R}^3)} = 0,$$

$$\text{weak} \lim_{T\to 0^+} S(t_0-T)\begin{pmatrix} u(T) \\ \partial_t u(T)\end{pmatrix} = 0.$$

Lemma A.4. *Suppose that $\{(u_{0,n},u_{1,n})\}_{n\in\mathbb{Z}}$ is a bounded sequence in $\dot{H}^s\times\dot{H}^{s-1}(\mathbb{R}^3)$ so that*

$$\lim_{n\to\infty} \|S(t)(u_{0,n},u_{1,n})\|_{L^q L^r([0,\mu]\times\mathbb{R}^3)} = 0.$$

Here (q,r) is an s-admissible pair and μ is a positive constant. Then we have the weak limit in $\dot{H}^s\times\dot{H}^{s-1}(\mathbb{R}^3)$

$$(u_{0,n},u_{1,n}) \rightharpoonup 0.$$

Proof. Let us suppose the conclusion were false. This means that there exists a subsequence (for which we use the same notation as the original sequence) that converges weakly to a nonzero limit $(\tilde{u}_0,\tilde{u}_1)$. We know the operator $P:\dot{H}^s\times\dot{H}^{s-1}\to L^qL^r([0,\mu]\times\mathbb{R}^3)$ defined by

$$P(u_0,u_1) = S(t)(u_0,u_1)$$

is bounded by the Strichartz estimate. This implies that we have the weak limit in $L^qL^r([0,\mu]\times\mathbb{R}^3)$

$$P(u_{0,n},u_{1,n}) \rightharpoonup P(\tilde{u}_0,\tilde{u}_1).$$

On the other hand, we know $P(u_{0,n},u_{1,n})$ converges to zero strongly. Thus $P(\tilde{u}_0,\tilde{u}_1)=0$. This means $(\tilde{u}_0,\tilde{u}_1)=0$, which is a contradiction. □

Lemma A.5. *Assume $s\in[s_p,1]$. Let $u(x,t)$ be defined on $I=(T_-,\infty)$ and almost periodic modulo scalings in $\dot{H}^s\times\dot{H}^{s-1}(\mathbb{R}^3)$, namely the set*

$$K = \left\{\left(\frac{1}{\lambda(t)^{3/2-s_p}}u\left(\frac{x}{\lambda(t)},t\right), \frac{1}{\lambda(t)^{5/2-s_p}}\partial_t u\left(\frac{x}{\lambda(t)},t\right)\right) : t\in I\right\}$$

is precompact in the space $\dot{H}^s\times\dot{H}^{s-1}(\mathbb{R}^3)$. In addition, $\lambda(t)\le 1$ when t is large. Then, for any closed interval $[a,b]$ and any s-admissible pair (q,r) with $q<\infty$, we have

$$\lim_{T\to+\infty} \|S(t-T)(u(T),\partial_t u(T))\|_{L^q L^r([a,b]\times\mathbb{R}^3)} = 0.$$

Proof. One could use the similar method as used in Lemma A.2 by observing

$$\begin{aligned}\|S(t-T)(u(T),\partial_t u(T))\|_{L^qL^r([a,b]\times\mathbb{R}^3)}&\\ &= \|S(t)(u(T),\partial_t u(T))\|_{L^qL^r([a-T,b-T]\times\mathbb{R}^3)}\\ &= (\lambda(T))^{s-s_p}\|S(t)(u_0^{(T)},u_1^{(T)})\|_{L^qL^r([\lambda(T)(a-T),\lambda(T)(b-T)]\times\mathbb{R}^3)}.\end{aligned}$$

Here

$$(u_0^{(T)},u_1^{(T)}) = \left(\frac{1}{\lambda(T)^{3/2-s_p}}u\left(\frac{\cdot}{\lambda(T)},T\right),\frac{1}{\lambda(T)^{5/2-s_p}}\partial_t u\left(\frac{\cdot}{\lambda(T)},T\right)\right).$$ □

Perturbation theory. In this subsection we will finish the proof of Theorem 2.12 and Theorem 2.15.

Proof of Theorem 2.12. Let us first prove the perturbation theory when M is sufficiently small. Let I_1 be the maximal lifespan of the solution $u(x,t)$ to the equation (1) with the given initial data (u_0,u_1) and assume $[0,T]\subseteq I\cap I_1$. By the Strichartz estimate, we have

$$\begin{aligned}\|\tilde u-u\|_{Y_{s_p}([0,T])} &\le \|S(t)(u_0-\tilde u(0),u_1-\tilde u(0))\|_{Y_{s_p}([0,T])}+C_p\|e+F(\tilde u)-F(u)\|_{Z_{s_p}([0,T])}\\ &\le \varepsilon + C_p\|e\|_{Z_{s_p}([0,T])}+C_p\|F(\tilde u)-F(u)\|_{Z_{s_p}([0,T])}\\ &\le \varepsilon + C_p\varepsilon + C_p\|\tilde u-u\|_{Y_{s_p}([0,T])}\big(\|\tilde u\|^{p-1}_{Y_{s_p}([0,T])}+\|\tilde u-u\|^{p-1}_{Y_{s_p}([0,T])}\big)\\ &\le C_p\varepsilon + C_p\|\tilde u-u\|_{Y_{s_p}([0,T])}\big(M^{p-1}+\|\tilde u-u\|^{p-1}_{Y_{s_p}([0,T])}\big).\end{aligned}$$

By a continuity argument in T, there exist $M_0=M_0(p)$ and $\varepsilon_0=\varepsilon_0(p)>0$ such that if $M\le M_0$ and $\varepsilon<\varepsilon_0$, we have

$$\|\tilde u-u\|_{Y_{s_p}([0,T])}\le C_p\varepsilon.$$

Observing that this estimate does not depend on the time T, we are actually able to conclude $I\subseteq I_1$ by the standard blow-up criterion and obtain

$$\|\tilde u-u\|_{Y_{s_p}(I)}\le C_p\varepsilon.$$

In addition, by the Strichartz estimate

$$\begin{aligned}\sup_{t\in I}\left\|\begin{pmatrix}u(t)\\ \partial_t u(t)\end{pmatrix}-\begin{pmatrix}\tilde u(t)\\ \partial_t\tilde u(t)\end{pmatrix}-S(t)\begin{pmatrix}u_0-\tilde u(0)\\ u_1-\partial_t\tilde u(0)\end{pmatrix}\right\|_{\dot H^{s_p}\times\dot H^{s_p-1}}&\\ &\le C_p\|F(u)-F(\tilde u)-e\|_{Z_{s_p}(I)}\\ &\le C_p\big(\|e\|_{Z_{s_p}(I)}+\|F(u)-F(\tilde u)\|_{Z_{s_p}(I)}\big)\\ &\le C_p\big[\varepsilon+\|u-\tilde u\|_{Y_{s_p}(I)}\big(\|\tilde u\|^{p-1}_{Y_{s_p}(I)}+\|u-\tilde u\|^{p-1}_{Y_{s_p}}\big)\big]\\ &\le C_p\varepsilon.\end{aligned}$$

This finishes the proof as M is sufficiently small. To deal with the general case, we can separate the time interval I into a finite number of subintervals $\{I_j\}$, so that $\|\tilde u\|_{Y_{s_p}(I_j)}<M_0$, and then iterate our argument above. □

Proof of Theorem 2.15. Let us first prove the perturbation theory when M and T are sufficiently small. Let I_1 be the maximal lifespan of the solution $u(x,t)$ to the equation (1) with the given initial data (u_0,u_1) and assume $[0,T_1]\subseteq[0,T]\cap I_1$. By the Strichartz estimate, we have

$$\begin{aligned}
\|\tilde{u}-u\|_{Y_s([0,T_1])} &\le \|S(t)(u_0-\tilde{u}_0,u_1-\tilde{u}_1)\|_{Y_s([0,T_1])}+C_{s,p}\|F(\tilde{u})-F(u)\|_{Z_s([0,T_1])}\\
&\le C_{s,p}\|(u_0-\tilde{u}_0,u_1-\tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}}+C_{s,p}\|F(\tilde{u})-F(u)\|_{Z_s([0,T_1])}\\
&\le C_{s,p}\|(u_0-\tilde{u}_0,u_1-\tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}}\\
&\qquad +C_{s,p}T_1^{(p-1)(s-s_p)}\|F(\tilde{u})-F(u)\|_{L^{\frac{2}{s+1-(2p-2)(s-s_p)}}L^{\frac{2}{2-s}}([0,T_1]\times\mathbb{R}^3)}\\
&\le C_{s,p}\|(u_0-\tilde{u}_0,u_1-\tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}}\\
&\qquad +C_{s,p}T_1^{(p-1)(s-s_p)}\|\tilde{u}-u\|_{Y_s([0,T_1])}\big(\|\tilde{u}-u\|^{p-1}_{Y_s([0,T_1])}+\|\tilde{u}\|^{p-1}_{Y_s([0,T_1])}\big)\\
&\le C_{s,p}\|(u_0-\tilde{u}_0,u_1-\tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}}\\
&\qquad +C_{s,p}T_1^{(p-1)(s-s_p)}\|\tilde{u}-u\|_{Y_s([0,T_1])}\big(\|\tilde{u}-u\|^{p-1}_{Y_s([0,T_1])}+M^{p-1}\big).
\end{aligned}$$

By a continuity argument in T_1, there exist $M_0=M_0(s,p)$ and $\varepsilon_0=\varepsilon_0(s,p)>0$ such that if $M\le M_0$, $T\le 1$ and

$$\|(u_0-\tilde{u}_0,u_1-\tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}}\le\varepsilon_0,$$

we have

$$\|\tilde{u}-u\|_{Y_s([0,T_1])}\le C_{s,p}\|(u_0-\tilde{u}_0,u_1-\tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}}.$$

Observing that this estimate does not depend on the time T_1 as long as $T_1\le T\le 1$, we are actually able to conclude $[0,T]\subseteq I_1$ by Theorem 2.14 and obtain

$$\|\tilde{u}-u\|_{Y_s([0,T])}\le C_{s,p}\|(u_0-\tilde{u}_0,u_1-\tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}}.$$

In addition, by the Strichartz estimate

$$\begin{aligned}
&\sup_{t\in[0,T]}\left\|\begin{pmatrix}u(t)\\ \partial_t u(t)\end{pmatrix}-\begin{pmatrix}\tilde{u}(t)\\ \partial_t\tilde{u}(t)\end{pmatrix}\right\|_{\dot{H}^s\times\dot{H}^{s-1}}\\
&\qquad\le\left\|S(t)\begin{pmatrix}u_0-\tilde{u}_0\\ u_1-\tilde{u}_1\end{pmatrix}\right\|_{\dot{H}^s\times\dot{H}^{s-1}}+C_{s,p}\|F(u)-F(\tilde{u})\|_{Z_s([0,T])}\\
&\qquad\le C_{s,p}\|(u_0-\tilde{u}_0,u_1-\tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}}\\
&\qquad\qquad +C_{s,p}T^{(p-1)(s-s_p)}\|\tilde{u}-u\|_{Y_s([0,T])}\big(\|\tilde{u}-u\|^{p-1}_{Y_s([0,T])}+\|\tilde{u}\|^{p-1}_{Y_s([0,T])}\big)\\
&\qquad\le C_{s,p}\|(u_0-\tilde{u}_0,u_1-\tilde{u}_1)\|_{\dot{H}^s\times\dot{H}^{s-1}}.
\end{aligned}$$

This finishes the proof as M and T are sufficiently small. To deal with the general case, we can separate the time interval $[0,T]$ into a finite number of subintervals $\{I_j\}$, so that $\|\tilde{u}\|_{Y_s(I_j)}\le M_0$ and $|I_j|\le 1$, then iterate our argument above. $\square$

Technical lemmas.

Lemma A.6. *Suppose that $(u_{0,\varepsilon}(x), u_{1,\varepsilon}(x))$ are radial, smooth pairs defined in $\mathbb{R}^3$ and converge to $(u_0(x), u_1(x))$ strongly in $\dot{H}^{s_p} \times \dot{H}^{s_p-1}(\mathbb{R}^3)$. In addition, we have*

$$\int_{r_0<|x|<4r_0} (|\nabla u_{0,\varepsilon}(x,t_0)|^2 + |u_{1,\varepsilon}(x,t_0)|^2)\, dx \le C$$

for each $\varepsilon < \varepsilon_0$. Then $(u_0(x), u_1(x))$ is in the space $\dot{H}^1 \times L^2(r < |x| < 4r)$ and satisfies

$$\int_{r_0<|x|<4r_0} (|\nabla u_0(x)|^2 + |u_1(x)|^2)\, dx \le C.$$

Proof. By the uniform bound of the integral, we can extract a sequence $\varepsilon_i \to 0$ so that $\partial_r u_{0,\varepsilon_i}(r)$ converges to $\tilde{u}_0'(r)$ weakly in $L^2(r_0, 4r_0)$, and u_{1,ε_i} converges to $\tilde{u}_1$ weakly in $L^2(r_0 < |x| < 4r_0)$. Define

$$\tilde{u}_0(r) = u_0(r_0) + \int_{r_0}^{r} \tilde{u}_0'(\tau)\, d\tau.$$

We have

$$\int_{r_0<|x|<4r_0} (|\nabla \tilde{u}_0(x)|^2 + |\tilde{u}_1(x)|^2)\, dx \le C.$$

By the strong and weak convergence, we have immediately $u_1 = \tilde{u}_1$ in the region $r_0 < |x| < 4r_0$. In order to conclude, we only need to show $u_0(r) = \tilde{u}_0(r)$. Observing $\int_{r_0}^{r_1} f(\tau)\, d\tau$ is a bounded linear functional in $L^2(r_0, 4r_0)$ for each $r_1 \in (r_0, 4r_0)$, we have

$$\begin{aligned}
\tilde{u}_0(r_1) &= u_0(r_0) + \int_{r_0}^{r_1} \tilde{u}_0'(\tau)\, d\tau \\
&= \lim_{i\to\infty} u_{0,\varepsilon_i}(r_0) + \lim_{i\to\infty} \int_{r_0}^{r_1} \partial_r u_{0,\varepsilon_i}(\tau)\, d\tau \\
&= \lim_{i\to\infty} \left(u_{0,\varepsilon_i}(r_0) + \int_{r_0}^{r_1} \partial_r u_{0,\varepsilon_i}(\tau)\, d\tau \right) \\
&= \lim_{i\to\infty} u_{0,\varepsilon_i}(r_1) \\
&= u_0(r_1).
\end{aligned}$$

This completes the proof. □

Lemma A.7. *Assume $\frac{1}{2} < s < \frac{3}{2}$. Given any radial $\dot{H}^s(\mathbb{R}^3)$ function f, we have*

$$\lim_{|x|\to 0^+} |x|^{\frac{3}{2}-s} f(x) = \lim_{|x|\to\infty} |x|^{\frac{3}{2}-s} f(x) = 0.$$

Proof. Let $s_1 \in (s, \frac{3}{2})$. Applying frequency cutoff techniques and using (8), we have

$$\begin{aligned}
|x|^{\frac{3}{2}-s} |(P_{>M} f)(x)| &\le C_s \|P_{>M} f\|_{\dot{H}^s}, \\
|x|^{\frac{3}{2}-s} |(P_{\le M} f)(x)| &\le C_{s_1} |x|^{s_1-s} \|P_{\le M} f\|_{\dot{H}^{s_1}},
\end{aligned}$$

for any fixed $M > 0$. Combining the higher and lower frequency parts, we obtain

$$\limsup_{|x|\to 0^+} |x|^{\frac{3}{2}-s} |f(x)| \le C_s \|P_{>M} f\|_{\dot{H}^s}.$$

This proves the first limit if we let $M \to +\infty$. We can prove the second limit in a similar way. $\square$

Acknowledgements

The majority of this work was completed when the author was a graduate student at the University of Chicago. I would like to express my sincere gratitude to my advisor Professor Carlos E. Kenig for the introduction of this topic and helpful discussions.

References

[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

[Colliander et al. 2008] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "Global well-posedness and scattering for the energy-critical nonlinear Schrödinger equation in $\mathbb{R}^3$", *Ann. of Math.* (2) **167**:3 (2008), 767–865. MR 2009f:35315

[Duyckaerts et al. 2011] T. Duyckaerts, C. Kenig, and F. Merle, "Universality of blow-up profile for small radial type II blow-up solutions of the energy-critical wave equation", *J. Eur. Math. Soc.* (*JEMS*) **13**:3 (2011), 533–599. MR 2012e:35160

[Duyckaerts et al. 2012] T. Duyckaerts, C. Kenig, and F. Merle, "Scattering for radial, bounded solutions of focusing supercritical wave equations", preprint, 2012. arXiv 1208.2158

[Duyckaerts et al. 2013] T. Duyckaerts, C. Kenig, and F. Merle, "Classification of radial solutions of the focusing, energy-critical wave equation", *Cambridge Journal of Mathematics* **1**:1 (2013), 75–144.

[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

[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

[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

[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 2008] C. E. Kenig and F. Merle, "Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation", *Acta Math.* **201**:2 (2008), 147–212. MR 2011a:35344

[Kenig and Merle 2010] C. E. Kenig and F. Merle, "Scattering for $\dot{H}^{1/2}$ bounded solutions to the cubic, defocusing NLS in 3 dimensions", *Trans. Amer. Math. Soc.* **362**:4 (2010), 1937–1962. MR 2011b:35486

[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 2012i:35244

[Killip and Visan 2010] R. Killip and M. Visan, "The focusing energy-critical nonlinear Schrödinger equation in dimensions five and higher", *Amer. J. Math.* **132**:2 (2010), 361–424. MR 2011e:35357

[Killip et al. 2009] R. Killip, T. Tao, and M. Visan, "The cubic nonlinear Schrödinger equation in two dimensions with radial data", *J. Eur. Math. Soc.* (*JEMS*) **11**:6 (2009), 1203–1258. MR 2010m:35487

[Killip et al. 2014] R. Killip, B. Stovall, and M. Visan, "Blowup behaviour for the nonlinear Klein–Gordon equation", *Math. Ann.* **358**:1-2 (2014), 289–350. MR 3157999

[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

[Perthame and Vega 1999] B. Perthame and L. Vega, "Morrey–Campanato estimates for Helmholtz equations", *J. Funct. Anal.* **164**:2 (1999), 340–355. MR 2000i:35023

[Shen 2011] R. Shen, "Global well-posedness and scattering of defocusing energy subcritical nonlinear wave equation in dimension 3 with radial data", preprint, 2011. arXiv 1111.2345

Received 22 Oct 2012. Accepted 5 Aug 2013.

RUIPENG SHEN: srpgow@gmail.com
Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S4K1, Canada

ANALYSIS AND PDE
Vol. 6, No. 8, 2013
dx.doi.org/10.2140/apde.2013.6.1989

GLOBAL WELL-POSEDNESS FOR THE NONLINEAR SCHRÖDINGER EQUATION WITH DERIVATIVE IN ENERGY SPACE

YIFEI WU

In this paper, we prove that there exists some small $\varepsilon_* > 0$ such that the derivative nonlinear Schrödinger equation (DNLS) is globally well-posed in the energy space, provided that the initial data $u_0 \in H^1(\mathbb{R})$ satisfies $\|u_0\|_{L^2} < \sqrt{2\pi} + \varepsilon_*$. This result shows us that there are no blow-up solutions whose masses slightly exceed 2π, even if their energies are negative. This phenomenon is much different from the behavior of the nonlinear Schrödinger equation with critical nonlinearity. The technique used is a variational argument together with the momentum conservation law. Further, for the DNLS on the half-line $\mathbb{R}^+$, we show the blow-up for the solution with negative energy.

1. Introduction

We study the following Cauchy problem of the nonlinear Schrödinger equation with derivative (DNLS):

$$\begin{cases} i\partial_t u + \partial_x^2 u = i\lambda\partial_x(|u|^2 u), & t \in \mathbb{R},\ x \in \mathbb{R}, \\ u(0, x) = u_0(x) \in H^1(\mathbb{R}), \end{cases} \tag{1-1}$$

where $\lambda \in \mathbb{R}$. It arises from studying the propagation of circularly polarized Alfvén waves in magnetized plasma with a constant magnetic field; see [Mio et al. 1976; Mjolhus 1976; Sulem and Sulem 1999] and the references therein.

This equation is L^2-critical in the sense that both the equation and the L^2-norm are invariant under the scaling transform

$$u_\alpha(t, x) = \alpha^{1/2} u(\alpha^2 t, \alpha x), \quad \alpha > 0.$$

It has the same scaling invariance as the quintic nonlinear Schrödinger equation,

$$i\partial_t u + \partial_x^2 u + \mu|u|^4 u = 0, \quad t \in \mathbb{R},\ x \in \mathbb{R},$$

and the quintic generalized Korteweg–de Vries equation,

$$\partial_t u + \partial_x^3 u + \mu\partial_x(u^5) = 0, \quad t \in \mathbb{R}, x \in \mathbb{R}.$$

One may always take $\lambda = 1$ in (1-1), since the general case can be reduced to this case by the following two transforms. First, we apply the transform

$$u(t, x) \mapsto \bar{u}(-t, x),$$

The author was partially supported by the NSF of China (number 11101042), the Chinese Postdoctoral Science Foundation (numbers 20110490018 and 2012T50068), and the Fundamental Research Funds for the Central Universities of China.
MSC2010: primary 35Q55; secondary 35A01, 35B44.
Keywords: nonlinear Schrödinger equation with derivative, global well-posedness, blow-up, half-line.

then reduce the equation to the case of $\lambda > 0$. Then we take the rescaling transform

$$u(t,x) \mapsto \frac{1}{\sqrt{\lambda}} u(t,x)$$

and reduce it to the case of $\lambda = 1$. So in this sense, (1-1) can always be regarded as the focusing equation. From now on, we always assume that $\lambda = 1$ in (1-1).

The H^1-solution of (1-1) obeys three conservation laws. The first is the conservation of the mass

$$M(u(t)) := \int_{\mathbb{R}} |u(t)|^2 \, dx = M(u_0); \tag{1-2}$$

the second is the conservation of energy

$$E_D(u(t)) := \int_{\mathbb{R}} \Big(|u_x(t)|^2 + \tfrac{3}{2} \operatorname{Im} |u(t)|^2 u(t) \overline{u_x(t)} + \tfrac{1}{2} |u(t)|^6 \Big) \, dx = E_D(u_0); \tag{1-3}$$

and the third is the conservation of momentum (see (3-4) below),

$$P_D(u(t)) := \operatorname{Im} \int_{\mathbb{R}} \bar{u}(t) u_x(t) \, dx - \frac{1}{2} \int_{\mathbb{R}} |u(t)|^4 \, dx = P_D(u_0). \tag{1-4}$$

Local well-posedness for the Cauchy problem (1-1) is well understood. It was proved for the energy space $H^1(\mathbb{R})$ in [Hayashi 1993; Hayashi and Ozawa 1992; 1994]; see also [Guo and Tan 1991] for an earlier result in smooth spaces. For rough data below the energy space, Takaoka [1999] proved local well-posedness in $H^s(\mathbb{R})$ for $s \geq \frac{1}{2}$. This result was shown to be sharp in the sense that the flow map fails to be uniformly C^0 for $s < \frac{1}{2}$; see [Biagioni and Linares 2001; Takaoka 2001].

The global well-posedness for (1-1) has also been widely studied. By using mass and energy conservation laws, and by developing the gauge transformations, Hayashi and Ozawa [Hayashi and Ozawa 1994; Ozawa 1996] proved that the problem (1-1) is globally well-posed in energy space $H^1(\mathbb{R})$ under the condition

$$\|u_0\|_{L^2} < \sqrt{2\pi}. \tag{1-5}$$

Further, for initial data of regularity below the energy space, Colliander et al. [2001; 2002] proved the global well-posedness for (1-1) in $H^s(\mathbb{R})$ for $s > \frac{1}{2}$, under the condition (1-5). Recently, Miao, Wu, and Xu [Miao et al. 2011] proved that (1-1) is globally well-posed in the critical space $H^{1/2}(\mathbb{R})$, also under the condition (1-5). For other work on the DNLS in the periodic case, see for example [Grünrock and Herr 2008; Herr 2006; Nahmod et al. 2012; Win 2010].

As mentioned above, all the results on global existence for initial data were obtained under the assumption (1-5). Since $\sqrt{2\pi}$ is just the mass of the ground state of the corresponding elliptic problem, the condition (1-5) was naturally used to keep the energy positive; see [Colliander et al. 2001; Miao et al. 2011] for examples. Now one may wonder what happens to the well-posedness for the solution when (1-5) is not fulfilled. Our first main result in this paper is to improve the assumption (1-5) and obtain the global well-posedness as follows.

Theorem 1.1. *There exists a small $\varepsilon_* > 0$ such that, for any $u_0 \in H^1(\mathbb{R})$ with*

$$\int_{\mathbb{R}} |u_0(x)|^2\,dx < 2\pi + \varepsilon_*, \tag{1-6}$$

the Cauchy problem (1-1) ($\lambda = 1$) *is globally well-posed in $H^1(\mathbb{R})$ and the solution u satisfies*

$$\|u\|_{L_t^\infty H_x^1} \le C(\varepsilon_*, \|u_0\|_{H^1}).$$

The technique used to prove Theorem 1.1 is a variational argument together with the momentum and energy conservation laws. The key ingredient is the momentum conservation law, rather than the energy conservation law, upon which many (subcritical) problems rely when studying the global existence. We argue by contradiction. Suppose that the solution of (1-1) blows up at finite/infinite time T and t_n is a time sequence tending to T such that $u(t_n)$ tends to infinity in $H^1(\mathbb{R})$ norm. Then, thanks to the energy conservation law and a variational lemma from Merle [2001], $u(t_n)$ is close to the ground state Q (see below for its definition) up to a spatial transformation, a phase rotation, and a scaling transformation. On the one hand, since $u(t_n)$ blows up at T, the scaling parameter λ_n decays to zero; on the other hand, the conservation of momentum prevents λ_n from tending to zero. This leads to a contradiction.

As mentioned above, Theorem 1.1 improves the smallness of the L^2-norm of the initial data of the previous works on global existence [Hayashi and Ozawa 1994; Ozawa 1996]. More importantly, it reveals some special features of the derivative nonlinear Schrödinger equation. As discussed before, the smallness condition (1-5) in the previous works is imposed to guarantee the positivity of the energy $E_D(u(t))$. Indeed, by using a variant gauge transformation

$$v(t,x) := e^{-(3/4)i\int_{-\infty}^{x}|u(t,y)|^2\,dy}u(t,x), \tag{1-7}$$

the energy is deduced to be

$$E_D(u(t)) = \|v_x(t)\|_{L_x^2}^2 - \tfrac{1}{16}\|v(t)\|_{L_x^6}^6 := E(v(t)), \tag{1-8}$$

and then the positivity of $E(v)$ is followed by the sharp Gagliardo–Nirenberg inequality (see [Weinstein 1982/83])

$$\|f\|_{L^6}^6 \le \frac{4}{\pi^2}\|f\|_{L^2}^4\|f_x\|_{L^2}^2. \tag{1-9}$$

Once the mass is greater than 2π, the positive energy can not be maintained. To see this, we first make use of the gauge transformation (1-7), and rewrite (1-1) as

$$i\partial_t v + \partial_x^2 v = \frac{i}{2}|v|^2 v_x - \frac{i}{2}v^2\bar{v}_x - \tfrac{3}{16}|v|^4 v. \tag{1-10}$$

Then there exists a standing wave $e^{it}Q$ of (1-10), where Q is the unique (up to some symmetries) positive solution of the elliptic equation

$$-Q_{xx} + Q - \tfrac{3}{16}Q^5 = 0.$$

This leads to the standing wave solution corresponding to (1-1),

$$R(t,x) := e^{it+(3/4)i\int_{-\infty}^{x}Q^2\,dy}Q(x).$$

So on the one hand, as a byproduct, our result implies the stability of the standing wave solution, which has been proved by Colin and Ohta [2006]. On the other hand,

$$\|Q\|_{L^2} = \sqrt{2\pi}, \quad E(Q) = 0,$$

and the Fréchet derivation of the functional $E(v)$ at Q satisfies $\delta E(Q) \cdot Q = -2\pi < 0$. These relations imply that there exists a u_0 such that u_0 obeys (1-6) and $E_D(u_0) < 0$. Therefore, there indeed exist global solutions with negative energy, as stated in Theorem 1.1. Obviously this is much different from the focusing, quintic nonlinear Schrödinger equation (3-1) and focusing, quintic generalized Korteweg–de Vries equation (3-2). For (3-1), Ogawa and Tsutsumi [1991] proved that the solutions with the initial data belonging to $H^1(\mathbb{R})$ and negative energy must blow up in finite time; for (3-2), Martel and Merle [Martel and Merle 2002; Merle 2001] proved that the solutions with the initial data belonging to $H^1(\mathbb{R})$, negative energy, and obeying some further decay conditions blow up in finite time. In Section 3 below we will discuss some differences among these three equations, in particular from the viewpoint of the virial arguments.

Moreover, the situation of the Cauchy problem and the initial boundary value problem of (1-1) are much different. We consider the following Cauchy–Dirichlet problem of the nonlinear Schrödinger equation with derivative on the half-line $\mathbb{R}^+$:

$$\begin{cases} i\partial_t u + \partial_x^2 u = i\partial_x(|u|^2 u), & t \in \mathbb{R},\ x \in (0, +\infty), \\ u(0, x) = u_0(x), \\ u(t, 0) = 0. \end{cases} \tag{1-11}$$

We show that under some assumptions, the solution must blow up in finite time if its energy is negative.

Theorem 1.2. *Let* $u_0 \in H^2(\mathbb{R}^+)$ *and* $xu_0 \in L^2(\mathbb{R}^+)$, *and let* u *be the corresponding solution of* (1-11) *which exists on the* (*right*) *maximal lifetime* $[0, T_*)$. *If* $E_D(u_0) < 0$, *then* $T_* < \infty$. *Moreover*, *there exists a constant* $C = C(u_0) > 0$ *such that*

$$\|u_x(t, x)\|_{L^2(\mathbb{R}^+)} \geq \frac{C}{\sqrt{T_* - t}} \to \infty \quad \text{as } t \nearrow T^*.$$

For related results on the blow-up solution to the DNLS equation on bounded domain with the Dirichlet condition, see [Tan 2004].

Lastly, we remark that it remains open for the DNLS equation (1-1) whether there exists an $H^1(\mathbb{R})$ initial data of much larger L^2-norm such that the corresponding solution blows up in finite time. Moreover, it may be interesting to study the existence of global rough solutions when the condition (1-5) on initial data is relaxed.

This paper is organized as follows. In Section 2, we present the gauge transformation and prove the virial identities of DNLS. In Section 3, we discuss the differences among the DNLS, the quintic NLS, and the quintic gKdV equations. In Section 4, we study the initial boundary value problem of the DNLS on the half-line and give the proof of Theorem 1.2. In Section 5, we prove Theorem 1.1.

2. Gauge transformations, virial identities

Gauge transformations. The gauge transformation is an important and very nice tool to study the nonlinear Schrödinger equation with derivative [Hayashi 1993; Hayashi and Ozawa 1992; 1994]. It gives some improvement of the nonlinearity. In this subsection, we present the various gauge transformations and their properties. See [Colliander et al. 2001; Ozawa 1996] for more details. We define

$$\mathscr{G}_a u(t,x) = e^{ia\int_{-\infty}^{x}|u(t,y)|^2\,dy}u(t,x).$$

Then $\mathscr{G}_a\mathscr{G}_{-a} = \mathrm{Id}$, the identity transform. For any function f,

$$\partial_x \mathscr{G}_a f = e^{ia\int_{-\infty}^{x}|f(t,y)|^2\,dy}(ia|f|^2 f + f_x). \tag{2-1}$$

Further, we have the following.

Lemma 2.1. *If u is the solution of* (1-1) (*where* $\lambda = 1$), $v = \mathscr{G}_a u$ *is the solution of the equation*

$$i\partial_t v + \partial_x^2 v - i2(a+1)|v|^2 v_x - i(2a+1)v^2\bar{v}_x + \tfrac{1}{2}a(2a+1)|v|^4 v = 0.$$

Moreover,

$$E_D(u) = \|\partial_x \mathscr{G}_a u\|_2^2 + (2a+\tfrac{3}{2})\,\mathrm{Im}\int_{\mathbb{R}} |\mathscr{G}_a u|^2 \mathscr{G}_a u \cdot \partial_x \overline{\mathscr{G}_a u}\,dx + (a^2+\tfrac{3}{2}a+\tfrac{1}{2})\int_{\mathbb{R}} |\mathscr{G}_a u|^6\,dx.$$

The proof of this lemma follows from a direct computation and is omitted.

To understand how the gauge transform improves the nonlinearity in the present form (1-1), we introduce the following two transforms used in [Hayashi and Ozawa 1994; Ozawa 1996]. Let

$$\phi = \mathscr{G}_{-1}u, \quad \psi = \mathscr{G}_{1/2}\partial_x \mathscr{G}_{-1/2}u.$$

Then (ϕ, ψ) solves the following system of nonlinear Schrödinger equations:

$$\begin{cases} i\partial_t \phi + \partial_x^2 \phi = -i\phi^2\bar{\psi}, \\ i\partial_t \psi + \partial_x^2 \psi = \psi^2\bar{\phi}. \end{cases} \tag{2-2}$$

Compared with the original equation (1-1), the system above has no loss of derivatives. Thus it is much more convenient to get the local solvability of (1-1) for suitable smooth data by considering the system (2-2) instead.

As mentioned above, it is convenient to consider $v = \mathscr{G}_{-3/4}u$. Then, by Lemma 2.1, the equation (1-1) of u reduces to (1-10), that is,

$$i\partial_t v + \partial_x^2 v = \tfrac{1}{2}i|v|^2 v_x - \tfrac{1}{2}2v^2\bar{v}_x - \tfrac{3}{16}|v|^4 v.$$

Moreover, the energy $E_D(u)$ in (1-3) is changed into $E(v)$ in (1-8). In the sequel we shall consider (1-10) and the energy (1-8) of v instead.

Virial identities. In this subsection, we discuss some virial identities for the nonlinear Schrödinger equation with derivative. Formally, one may find that the virial quantity of v is similar to that of the mass-critical nonlinear Schrödinger equation. However, it is in fact the difference that gives the different conclusions of these two equations. Let $\psi = \psi(x)$ be a smooth real function. Define

$$I(t) = \int_{\mathbb{R}} \psi |v(t)|^2 \, dx, \tag{2-3}$$

$$J(t) = 2 \operatorname{Im} \int_{\mathbb{R}} \psi \bar{v}(t) v_x(t) \, dx + \frac{1}{2} \int \psi |v(t)|^4 \, dx. \tag{2-4}$$

Lemma 2.2. *Let v be the solution of* (1-10) *with $v(0) = v_0 \in H^1(\mathbb{R})$, and let $\psi \in C^3$. Then*

$$I'(t) = 2 \operatorname{Im} \int_{\mathbb{R}} \psi' \bar{v}(t) v_x(t) \, dx, \tag{2-5}$$

$$J'(t) = 4 \int_{\mathbb{R}} \psi' \big(|v_x(t)|^2 - \tfrac{1}{16} |v(t)|^6\big) \, dx - \int_{\mathbb{R}} \psi''' |v(t)|^2 \, dx. \tag{2-6}$$

Proof. Employing the gauge transform

$$w(t, x) := \mathscr{G}_{-1/2} u(t, x) = \mathscr{G}_{1/4} v(t, x),$$

by Lemma 2.1, w obeys the equation

$$i w_t + w_{xx} = i |w|^2 w_x.$$

Moreover, since $v(t, x) = \mathscr{G}_{-1/4} w(t, x)$, by (2-1),

$$\partial_x v(t, x) = e^{-i(1/4) \int_{-\infty}^{x} |w(t,y)|^2 \, dy} \big(-\tfrac{1}{4} i |w|^2 w + w_x\big).$$

Thus we have

$$I(t) = \int_{\mathbb{R}} \psi |w(t)|^2 \, dx \quad \text{and} \quad J(t) = 2 \operatorname{Im} \int_{\mathbb{R}} \psi \bar{w}(t) w_x(t) \, dx.$$

Now, by a direct computation, we get

$$\begin{aligned} I'(t) &= 2 \operatorname{Re} \int_{\mathbb{R}} \psi \bar{w}(t, x) \partial_t w(t, x) \, dx = 2 \operatorname{Re} \int_{\mathbb{R}} \psi \bar{w} (i w_{xx} + |w|^2 w_x) \, dx \\ &= 2 \operatorname{Im} \int_{\mathbb{R}} \psi' \bar{w} w_x \, dx - \frac{1}{2} \int_{\mathbb{R}} \psi' |w|^4 \, dx. \end{aligned} \tag{2-7}$$

Applying (2-1) again,

$$\partial_x w(t, x) = e^{(1/4) i \int_{-\infty}^{x} |v(t,y)|^2 \, dy} \big(\tfrac{1}{4} i |v|^2 v + v_x\big). \tag{2-8}$$

This together with (2-7) gives (2-5). Now we turn to (2-6). For this, we get

$$\begin{aligned} J'(t) &= 2 \operatorname{Im} \int_{\mathbb{R}} \psi \bar{w}_t(t, x) w_x(t, x) \, dx + 2 \operatorname{Im} \int_{\mathbb{R}} \psi \bar{w}(t, x) w_{xt}(t, x) \, dx \\ &= -4 \operatorname{Im} \int_{\mathbb{R}} \psi w_t \bar{w}_x \, dx - 2 \operatorname{Im} \int_{\mathbb{R}} \psi' \bar{w} w_t \, dx \end{aligned}$$

$$
\begin{aligned}
&= -4\,\mathrm{Im}\int_{\mathbb{R}} \psi\,\bar{w}_x(iw_{xx}+|w|^2w_x)\,dx - 2\,\mathrm{Im}\int_{\mathbb{R}} \psi'\bar{w}(iw_{xx}+|w|^2w_x)\,dx \\
&= -4\,\mathrm{Re}\int_{\mathbb{R}} \psi\,\bar{w}_x w_{xx}\,dx - 2\,\mathrm{Re}\int_{\mathbb{R}} \psi'\bar{w}w_{xx}\,dx - 2\,\mathrm{Im}\int_{\mathbb{R}} \psi'|w|^2\bar{w}w_x\,dx \\
&= 4\int_{\mathbb{R}} \psi'|w_x|^2\,dx + 2\,\mathrm{Re}\int_{\mathbb{R}} \psi''\bar{w}w_x\,dx - 2\,\mathrm{Im}\int_{\mathbb{R}} \psi'|w|^2\bar{w}w_x\,dx \\
&= 4\int_{\mathbb{R}} \psi'|w_x|^2\,dx - \int_{\mathbb{R}} \psi'''|w|^2\,dx - 2\,\mathrm{Im}\int_{\mathbb{R}} \psi'|w|^2\bar{w}w_x\,dx. \qquad (2\text{-}9)
\end{aligned}
$$

Now, using (2-8), we have

$$|w_x|^2 = |v_x|^2 + \tfrac{1}{2}\,\mathrm{Im}(|v|^2\bar{v}v_x) + \tfrac{1}{16}|v|^6$$

and

$$|w|^2 = |v|^2, \quad \mathrm{Im}(|w|^2\bar{w}w_x) = \mathrm{Im}(|v|^2\bar{v}v_x) + \tfrac{1}{4}|v|^6.$$

These insert into (2-9) and we obtain (2-6). □

3. A comparison between DNLS, NLS-5, and gKdV-5

In this section, we discuss the nonlinear Schrödinger equation with derivative (1-10), the focusing, quintic nonlinear Schrödinger equation (NLS-5), which reads

$$i\partial_t u + \partial_x^2 u + \tfrac{3}{16}|u|^4 u = 0, \qquad (3\text{-}1)$$

and the focusing, quintic generalized Korteweg–de Vries equation (gKdV-5),

$$\partial_t u + \partial_x^3 u + \tfrac{3}{16}\partial_x(u^5) = 0. \qquad (3\text{-}2)$$

The first two equations have the same standing wave solutions as $e^{it}Q$, and the last one has a traveling wave solution $Q(x-t)$. These three equations have the same energies in the form of (1-8). So by the sharp Gagliardo–Nirenberg inequality, all of them are globally well-posed in $H^1(\mathbb{R})$ when the initial data $\|u_0\|_{L^2} < \|Q\|_{L^2} = \sqrt{2\pi}$.

Now we continue to discuss the difference between the first equation (DNLS) and the last two (NLS-5, gKdV-5).

First of all, we give some products from Lemma 2.2. We always assume that v is smooth enough. Taking $\psi = x$ and $\psi = x^2$, by (2-5), we have

$$\frac{d}{dt}\int_{\mathbb{R}} x|v(t)|^2\,dx = 2\,\mathrm{Im}\int_{\mathbb{R}} \bar{v}(t)v_x(t)\,dx$$

and

$$\frac{d}{dt}\int_{\mathbb{R}} x^2|v(t)|^2\,dx = 4\,\mathrm{Im}\int_{\mathbb{R}} x\bar{v}(t)v_x(t)\,dx, \qquad (3\text{-}3)$$

respectively. Note that these two identities resemble the corresponding identity of the mass-critical nonlinear Schrödinger equation (3-1).

Now we take $\psi = 1$ in (2-6), which gives the momentum conservation law,

$$P(v(t)) := \operatorname{Im}\int_{\mathbb{R}} \bar{v}(t)v_x(t)\,dx + \frac{1}{4}\int_{\mathbb{R}} |v(t)|^4\,dx = P(v_0). \tag{3-4}$$

Then, taking $\psi = x$, we have

$$\frac{d}{dt}\left(2\operatorname{Im}\int_{\mathbb{R}} x\bar{v}(t)v_x(t)\,dx + \frac{1}{2}\int_{\mathbb{R}} x|v(t)|^4\,dx\right) = 4E(v_0). \tag{3-5}$$

This equality is different from the situation of the mass-critical nonlinear Schrödinger equation (3-1). More precisely, for the solution u of (3-1) with the initial data u_0, we have

$$\frac{d}{dt}\left(2\operatorname{Im}\int_{\mathbb{R}} x\bar{u}(t)u_x(t)\,dx\right) = 4E(u_0). \tag{3-6}$$

Compared with the identity (3-6), there is an additional term $\frac{1}{2}\int x|v(t)|^4\,dx$ in (3-5). Indeed, for the solution of (3-1), combining with the same identity as in (3-3), one has

$$\frac{d^2}{dt^2}\int_{\mathbb{R}} x^2|u(t)|^2\,dx = 8E(u_0). \tag{3-7}$$

But this does not hold for the solution of (1-10). The "surplus" term $\frac{1}{2}\int x|v(t)|^4\,dx$ in (3-5) breaks the convexity of the variance. It is precisely this difference that leads to the distinct phenomena of the solutions of these two equations, at least at the technical level.

Using the virial identity (3-7), Glassey [1977] proved that the solution u of the mass-critical nonlinear Schrödinger equation

$$\partial_t u + \Delta u + |u|^{4/N}u = 0, \quad (t,x) \in \mathbb{R}\times\mathbb{R}^N,$$

blows up in finite time when $u_0 \in H^1(\mathbb{R}^N)$, $xu_0 \in L^2(\mathbb{R}^N)$, and $E(u_0) < 0$. Further, in the 1D case, Ogawa and Tsutsumi [1991] proved that the solutions of (3-1) blow up in finite time when $u_0 \in H^1(\mathbb{R})$ and $E(u_0) < 0$. See also [Du et al. 2013; Holmer and Roudenko 2010; Glangetas and Merle 1995; Nawa 1999], where all the solutions of the nonlinear Schrödinger equations with power nonlinearity blow up in finite time or infinite time if their energies are negative. However, Theorem 1.1 depicts a different scene, where there exist global and uniformly bounded solutions even if $E(v_0) < 0$.

The situation is also different from the mass-critical generalized KdV equation (3-2). The latter also has virial identity

$$\frac{d}{dt}\int_{\mathbb{R}} (x+t)|u(t)|^2\,dx = \int_{\mathbb{R}} u^2\,dx - 3\int_{\mathbb{R}} |u_x|^2\,dx - \frac{1}{3}\int_{\mathbb{R}} |u|^6\,dx.$$

The blow-up of the solutions to (3-2) also occurs when the initial data u_0 satisfies $E(u_0) < 0$, (1-6), and some decay conditions; see [Martel and Merle 2002; Merle 2001].

4. Blow-up for the DNLS on the half line

In this section, we use the virial identities obtained in Lemma 2.2 to study the blow-up solutions for the nonlinear Schrödinger equation with derivative on the half line. Consider the problem (1-11), and set

$$v(t,x)=\exp\biggl(-\tfrac{3}{4}i\int_0^x |u(t,y)|^2\,dy\biggr)u(t,x),$$

Using the gauge transformation, we see that v is the solution of

$$\begin{cases} i\partial_t v+\partial_x^2 v=\frac{1}{2}i|v|^2 v_x-\frac{1}{2}iv^2\bar{v}_x-\frac{3}{16}|v|^4 v, & t\in\mathbb{R},\ x\in(0,+\infty),\\ v(0,x)=v_0(x), \\ v(t,0)=0. \end{cases} \tag{4-1}$$

Note that after replacing the integral domain $\mathbb{R}$ by $\mathbb{R}^+$, the energy conservation law and all of the virial identities obtained in Section 2 also hold true for v.

Now using the virial identities and Glassey's argument [1977], we give the proof of Theorem 1.2.

Proof of Theorem 1.2. Let v be the solution to (4-1). Define

$$I(t)=\int_0^\infty x^2|v(t,x)|^2\,dx.$$

Then, by the identity analogous to (3-3), we have

$$I'(t)=4\operatorname{Im}\int_0^\infty x\bar{v}(t)v_x(t)\,dx=2\biggl(2\operatorname{Im}\int_0^\infty x\bar{v}(t)v_x(t)\,dx+\frac{1}{2}\int_0^\infty x|v(t)|^4\,dx\biggr)-\int_0^\infty x|v(t)|^4\,dx.$$

Now, by the identity analogous to (3-5), we get

$$\frac{d}{dt}\biggl(2\operatorname{Im}\int_0^\infty x\bar{v}(t)v_x(t)\,dx+\frac{1}{2}\int_0^\infty x|v(t)|^4\,dx\biggr)=4E(v_0).$$

Therefore, using these two identities, we obtain

$$I''(t)=8E(v_0)-\frac{d}{dt}\int_0^\infty x|v(t)|^4\,dx.$$

Integrating in time twice, we have

$$\begin{aligned} I(t)&=I(0)+I'(0)t+\int_0^t\int_0^s I''(\tau)\,d\tau\,ds\\ &=I(0)+I'(0)t+\int_0^t\int_0^s\biggl(8E(v_0)-\frac{d}{d\tau}\int_0^\infty x|v(\tau)|^4\,dx\biggr)\,d\tau\,ds\\ &=4E(v_0)t^2+\biggl(I'(0)+\int_0^\infty x|v_0|^4\,dx\biggr)t+I(0)-\int_0^t\int_0^\infty x|v(s)|^4\,dx\,ds\\ &\le 4E(v_0)t^2+\biggl(I'(0)+\int_0^\infty x|v_0|^4\,dx\biggr)t+I(0). \end{aligned} \tag{4-2}$$

Since $E(v_0)=E_D(u_0)<0$, there exists a finite time $T_*>0$ such that $I(T_*)=0$,

$$I(t)>0\quad\text{for } 0<t<T_*,$$

and

$$I(t)=O(T_*-t)\quad\text{as } t\nearrow T_*.$$

Note that

$$\int_0^\infty |v_0(x)|^2\,dx = \int_0^\infty |v(t,x)|^2\,dx = -2\,\mathrm{Re}\int_0^\infty x v(t,x)\overline{v_x}(t,x)\,dx$$
$$\le 2\|xv(t,x)\|_{L^2_x(\mathbb{R}^+)}\|v_x(t,x)\|_{L^2_x(\mathbb{R}^+)} = 2\sqrt{I(t)}\|v_x(t,\cdot)\|_{L^2(\mathbb{R}^+)}.$$

Then there is a constant $C = C(v_0) > 0$ such that

$$\|v_x(t,\cdot)\|_{L^2(\mathbb{R}^+)} \ge \frac{\int_0^\infty |v_0(x)|^2\,dx}{2\sqrt{I(t)}} \ge \frac{C}{\sqrt{T_* - t}}, \tag{4-3}$$

and the right-hand side goes to ∞ as $t \nearrow T^*$. Therefore, $v(t)$ blows up at time $T_* < +\infty$. Since

$$v_x = \exp\Bigl(-\tfrac{3}{4}i\int_0^x |u(t,y)|^2\,dy\Bigr)(-i\tfrac{3}{4}|u|^2u + u_x),$$

by the Gagliardo–Nirenberg inequality and the mass conservation law, there exists $C = C(u_0)$ such that

$$\|v_x(t,\cdot)\|_{L^2(\mathbb{R}^+)} \le \|u_x(t,\cdot)\|_{L^2(\mathbb{R}^+)} + \tfrac{3}{4}\|u(t,\cdot)\|^3_{L^6(\mathbb{R}^+)} \le C\|u_x(t,\cdot)\|_{L^2(\mathbb{R}^+)}.$$

Thus, by (4-3), this gives the analogous estimate on u. □

One may note from the proof that the key ingredient to obtain the blow-up result of the initial boundary value problem on the half-line case is the positivity of the "surplus" term $\int_0^\infty x|v(t)|^4\,dx$. This is not true for the Cauchy problem.

5. Proof of Theorem 1.1

Proof. Let $(-T_-(u_0), T_+(u_0))$ be the maximal lifespan of the solution u of (1-1). To prove Theorem 1.1, it is sufficient to obtain the (indeed uniformly) *a priori* estimate of the solutions on H^1-norm, that is,

$$\sup_{t\in(-T_-(u_0),T_+(u_0))} \|v_x(t)\|_{L^2} < +\infty.$$

Now we argue by contradiction and suppose that there exists a sequence $\{t_n\}$ with

$$t_n \to -T_-(u_0) \quad \text{or} \quad T_+(u_0)$$

such that

$$\|v_x(t_n)\|_{L^2} \to +\infty, \quad \text{as } n \to \infty. \tag{5-1}$$

Let

$$\lambda_n = \|Q_x\|_{L^2}/\|v_x(t_n)\|_{L^2} \tag{5-2}$$

and

$$w_n(x) = \lambda_n^{1/2} v(t_n, \lambda_n x). \tag{5-3}$$

Then, by (5-1),

$$\|\partial_x w_n\|_{L^2} = \|Q_x\|_{L^2} \quad \text{and} \quad \lambda_n \to 0, \ \text{as } n \to \infty.$$

First we have the following lemma.

Lemma 5.1. *For any $\varepsilon > 0$, there exists a small $\varepsilon_* = \varepsilon_*(\varepsilon) > 0$ such that if the function $f \in H^1(\mathbb{R})$ satisfies*

$$\int_{\mathbb{R}} |f(x)|^2\, dx < 2\pi + \varepsilon_*, \quad \|\partial_x f\|_{L^2} = \|\partial_x Q\|_{L^2}, \quad E(f) < \varepsilon_*,$$

then there exist $\gamma_0, x_0 \in \mathbb{R}$ such that

$$\|f - e^{-i\gamma_0} Q(\cdot - x_0)\|_{H^1} \le \varepsilon.$$

We put the proof of Lemma 5.1 at the end of this section and apply it to prove Theorem 1.1. Let $\varepsilon_0 > 0$ be a fixed small constant which will be chosen later, and let $\varepsilon_* = \varepsilon_*(\varepsilon_0) > 0$ be the number defined in Lemma 5.1. By (1-6), (5-3), and a simple computation,

$$\int_{\mathbb{R}} |w_n(x)|^2\, dx = \int_{\mathbb{R}} |v_0(x)|^2\, dx < 2\pi + \varepsilon_*,$$

and

$$\|\partial_x w_n\|_{L^2} = \|Q_x\|_{L^2}, \quad E(w_n) = \lambda_n^2 E(v_0) \to 0.$$

Then, by Lemma 5.1, we may inductively construct the sequences $\{\gamma_n\}$, $\{x_n\}$ which satisfy

$$\|w_n - e^{-i\gamma_n} Q(\cdot - x_n)\|_{H^1} \le \varepsilon_0 \quad \text{for any } n \ge n_0, \tag{5-4}$$

where $n_0 = n_0(\varepsilon_0)$ is a positive large number. Let

$$\varepsilon(t_n, x) = e^{i\gamma_n} w_n(x + x_n) - Q.$$

Then

$$w_n(x) = e^{-i\gamma_n} Q(x - x_n) + e^{-i\gamma_n} \varepsilon(t_n, x - x_n). \tag{5-5}$$

Therefore, by (5-3), (5-5), and (5-4), we have

$$v(t_n, x) = e^{-i\gamma_n} \lambda_n^{-1/2} (\varepsilon + Q)(t_n, \lambda_n^{-1} x - x_n), \quad \|\varepsilon(t_n)\|_{H^1} \le \varepsilon_0. \tag{5-6}$$

By the momentum and (5-6), one has

$$\begin{aligned}
P(v(t_n)) &= \operatorname{Im} \int_{\mathbb{R}} \bar{v}(t_n) v_x(t_n)\, dx + \frac{1}{4} \int_{\mathbb{R}} |v(t_n)|^4\, dx \\
&= \lambda_n^{-2} \operatorname{Im} \int_{\mathbb{R}} (\bar{\varepsilon} + Q)(t_n, \lambda_n^{-1} x - x_n) \cdot (\varepsilon_x + Q_x)(t_n, \lambda_n^{-1} x - x_n)\, dx \\
&\qquad\qquad + \tfrac{1}{4} \lambda_n^{-2} \int_{\mathbb{R}} |(\varepsilon + Q)(t_n, \lambda_n^{-1} x - x_n)|^4\, dx \\
&= \lambda_n^{-1} \operatorname{Im} \int_{\mathbb{R}} (\bar{\varepsilon}(t_n) + Q)(\varepsilon_x(t_n) + Q_x)\, dx + \tfrac{1}{4} \lambda_n^{-1} \int_{\mathbb{R}} |\varepsilon(t_n) + Q|^4\, dx \\
&= \lambda_n^{-1} \Bigl(\tfrac{1}{4} \|Q\|_{L^4}^4 + \operatorname{Im} \int_{\mathbb{R}} (Q_x \varepsilon(t_n) + Q \varepsilon_x(t_n) + \bar{\varepsilon} \varepsilon_x(t_n))\, dx + \frac{1}{4} \int_{\mathbb{R}} (|\varepsilon(t_n) + Q|^4 - Q^4)\, dx \Bigr) \\
&= \lambda_n^{-1} \bigl(\tfrac{1}{4} \|Q\|_{L^4}^4 + O(\|\varepsilon(t_n)\|_{H^1}) \bigr) \ge \lambda_n^{-1} \bigl(\tfrac{1}{4} \|Q\|_{L^4}^4 - C\varepsilon_0 \bigr).
\end{aligned} \tag{5-7}$$

Thus, by choosing ε_0 small enough such that $C\varepsilon_0 \le \frac{1}{8}\|Q\|_{L^4}^4$, one has $P(v(t_n)) \ge \lambda_n^{-1} \cdot \frac{1}{8}\|Q\|_{L^4}^4$. By the momentum conservation law, this proves that $P(v_0)\lambda_n \ge \frac{1}{8}\|Q\|_{L^4}^4$. That is, by (5-2),

$$\|v_x(t_n)\|_{L^2} \le 8P(v_0)\|Q_x\|_{L^2}/\|Q\|_{L^4}^4. \tag{5-8}$$

This violates (5-1). Therefore, we prove that there exists $C_0 = C_0(\varepsilon_*, \|v_0\|_{H^1})$, such that

$$\sup_{t\in\mathbb{R}} \|v_x(t)\|_{L^2} \le C_0.$$

Now, for the solution u of (1-1) (with $\lambda = 1$), we have $u = \mathcal{G}_{3/4}v$. Thus, by (2-1), we have

$$u_x = e^{i(3/4)\int_{-\infty}^x |v(t,y)|^2\,dy}\big(i\tfrac{3}{4}|v|^2 v + v_x\big).$$

Therefore, by (1-9) and the mass conservation law, for any $t \in \mathbb{R}$,

$$\|u_x(t)\|_{L^2} \le \|v_x(t)\|_{L^2} + \tfrac{3}{4}\|v(t)\|_{L^6}^3 \le \|v_x(t)\|_{L^2} + \frac{3}{2\pi}\|v(t)\|_{L^2}^2\|v_x(t)\|_{L^2} \le C_0\Big(1 + \frac{3}{2\pi}\|u_0\|_{L^2}^2\Big). \qquad \square$$

Proof of Lemma 5.1. The proof follows from the standard variational argument; see [Merle 2001; Weinstein 1986] for examples; see also [Banica 2004; Hmidi and Keraani 2005] for its applications. Here we prove it by using the profile decomposition (see [Gérard 1998] for example) for the sake of the completeness. Let $\{f_n\} \subset H^1(\mathbb{R})$ be any sequence satisfying

$$\|f_n\|_{L^2} \to \|Q\|_{L^2}, \quad \|\partial_x f_n\|_{L^2} = \|Q_x\|_{L^2}, \quad E(f_n) \to 0.$$

Then, by the profile decomposition, there exist $\{V^j\}$, $\{x_n^j\}$ such that, up to a subsequence,

$$f_n = \sum_{j=1}^{L} V^j(\cdot - x_n^j) + R_n^L,$$

where, for $j \ne k$, we have $|x_n^j - x_n^k| \to \infty$ as $n \to \infty$, and

$$\lim_{L\to\infty}\lim_{n\to\infty} \|R_n^L\|_{L^6} = 0. \tag{5-9}$$

Moreover,

$$\begin{aligned} \|f_n\|_{H^s}^2 &= \sum_{j=1}^{L} \|V^j\|_{H^s}^2 + \|R_n^L\|_{H^s}^2 + o_n(1) \quad \text{for } s = 0, 1,\\ E(f_n) &= \sum_{j=1}^{L} E(V^j) + E(R_n^L) + o_n(1). \end{aligned} \tag{5-10}$$

Since $\|f_n\|_{L^2} \to \|Q\|_{L^2}$, one has, by (5-10),

$$\|V^j\|_{L^2} \le \|Q\|_{L^2} \quad \text{for any } j \ge 1. \tag{5-11}$$

This implies, by the sharp Gagliardo–Nirenberg inequality (1-9), that $E(V^j) \ge 0$ for any $j \ge 1$. Further, by (5-9), one has

$$\lim_{L\to\infty}\lim_{n\to\infty} E(R_n^L) \ge 0.$$

Since $E(f_n) \to 0$, we have $E(V^j) = 0$ for any $j \geq 1$. Combining with (5-11) and (1-9), this again yields

$$\|V^j\|_{L^2} = \|Q\|_{L^2} \quad \text{or} \quad V^j = 0.$$

Since $\|f_n\|_{L^2} \to \|Q\|_{L^2}$, there exists exactly one j, say $j = 1$, such that

$$\|V^1\|_{L^2} = \|Q\|_{L^2}, \quad V^j = 0 \text{ for any } j \geq 2.$$

Moreover, by (5-10) and (1-9), when $n \to \infty$, we have $R_n^L \to 0$ in $L^2(\mathbb{R})$, and then further in $H^1(\mathbb{R})$. Therefore,

$$\|\partial_x V^1\|_{L^2} = \|Q_x\|_{L^2}, \quad E(V^1) = 0,$$

and $f_n \to V^1$ in $H^1(\mathbb{R})$ as $n \to \infty$. Now we note that V^1 attains the sharp Gagliardo–Nirenberg inequality (1-9). Thus, by the uniqueness of the minimizer of the Gagliardo–Nirenberg inequality [Weinstein 1982/83], we have $V^1 = e^{-i\gamma_0} Q(\cdot - x_0)$ for some $\gamma_0 \in \mathbb{R}$ and $x_0 \in \mathbb{R}$. This proves the lemma. □

Acknowledgements

The author thanks Professors Yongsheng Li and Changxing Miao for their help in writing this paper and for many valuable suggestions, and to Professor Tadahiro Oh for pointing out an important reference. The author is also grateful to the anonymous referees for helpful comments.

References

[Banica 2004] V. Banica, "Remarks on the blow-up for the Schrödinger equation with critical mass on a plane domain", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **3**:1 (2004), 139–170. MR 2005e:35209 Zbl 1170.35528

[Biagioni and Linares 2001] H. A. Biagioni and F. Linares, "Ill-posedness for the derivative Schrödinger and generalized Benjamin–Ono equations", *Trans. Amer. Math. Soc.* **353**:9 (2001), 3649–3659. MR 2002e:35215 Zbl 0970.35154

[Colin and Ohta 2006] M. Colin and M. Ohta, "Stability of solitary waves for derivative nonlinear Schrödinger equation", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **23**:5 (2006), 753–764. MR 2007e:35255 Zbl 1104.35050

[Colliander et al. 2001] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "Global well-posedness for Schrödinger equations with derivative", *SIAM J. Math. Anal.* **33**:3 (2001), 649–669. MR 2002j:35278 Zbl 1002.35113

[Colliander et al. 2002] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "A refined global well-posedness result for Schrödinger equations with derivative", *SIAM J. Math. Anal.* **34**:1 (2002), 64–86. MR 2004c:35381 Zbl 1034.35120

[Du et al. 2013] D. Du, Y. Wu, and K. Zhang, "On blow-up criterion for the nonlinear Schrödinger equation", preprint, 2013. arXiv 1309.6782

[Gérard 1998] P. Gérard, "Description du défaut de compacité de l'injection de Sobolev", *ESAIM Control Optim. Calc. Var.* **3** (1998), 213–233. MR 99h:46051 Zbl 0907.46027

[Glangetas and Merle 1995] L. Glangetas and F. Merle, "A geometrical approach of existence of blow up solutions in $H^1(\mathbb{R})$ for nonlinear Schrödinger equation", report R95031, Laboratoire d'Analyse Numérique, Univ. Pierre and Marie Curie, 1995.

[Glassey 1977] R. T. Glassey, "On the blowing up of solutions to the Cauchy problem for nonlinear Schrödinger equations", *J. Math. Phys.* **18**:9 (1977), 1794–1797. MR 57 #842 Zbl 0372.35009

[Grünrock and Herr 2008] A. Grünrock and S. Herr, "Low regularity local well-posedness of the derivative nonlinear Schrödinger equation with periodic initial data", *SIAM J. Math. Anal.* **39**:6 (2008), 1890–1920. MR 2009a:35233 Zbl 1156.35471

[Guo and Tan 1991] B. L. Guo and S. B. Tan, "On smooth solutions to the initial value problem for the mixed nonlinear Schrödinger equations", *Proc. Roy. Soc. Edinburgh Sect. A* **119**:1-2 (1991), 31–45. MR 92i:35114 Zbl 0766.35051

[Hayashi 1993] N. Hayashi, "The initial value problem for the derivative nonlinear Schrödinger equation in the energy space", *Nonlinear Anal.* **20**:7 (1993), 823–833. MR 94c:35007 Zbl 0787.35099

[Hayashi and Ozawa 1992] N. Hayashi and T. Ozawa, "On the derivative nonlinear Schrödinger equation", *Phys. D* **55**:1-2 (1992), 14–36. MR 93h:35190 Zbl 0741.35081

[Hayashi and Ozawa 1994] N. Hayashi and T. Ozawa, "Finite energy solutions of nonlinear Schrödinger equations of derivative type", *SIAM J. Math. Anal.* **25**:6 (1994), 1488–1503. MR 95i:35272 Zbl 0809.35124

[Herr 2006] S. Herr, "On the Cauchy problem for the derivative nonlinear Schrödinger equation with periodic boundary condition", *Int. Math. Res. Not.* **2006** (2006), Art. ID 96763, 33. MR 2007e:35258 Zbl 1149.35074

[Hmidi and Keraani 2005] T. Hmidi and S. Keraani, "Blowup theory for the critical nonlinear Schrödinger equations revisited", *Int. Math. Res. Not.* **2005**:46 (2005), 2815–2828. MR 2007k:35464 Zbl 1126.35067

[Holmer and Roudenko 2010] J. Holmer and S. Roudenko, "Divergence of infinite-variance nonradial solutions to the 3D NLS equation", *Comm. Partial Differential Equations* **35**:5 (2010), 878–905. MR 2011m:35353 Zbl 1195.35277

[Martel and Merle 2002] Y. Martel and F. Merle, "Blow up in finite time and dynamics of blow up solutions for the L^2-critical generalized KdV equation", *J. Amer. Math. Soc.* **15**:3 (2002), 617–664. MR 2003c:35142 Zbl 0996.35064

[Merle 2001] F. Merle, "Existence of blow-up solutions in the energy space for the critical generalized KdV equation", *J. Amer. Math. Soc.* **14**:3 (2001), 555–578. MR 2002f:35193 Zbl 0970.35128

[Miao et al. 2011] C. Miao, Y. Wu, and G. Xu, "Global well-posedness for Schrödinger equation with derivative in $H^{\frac{1}{2}}(\mathbb{R})$", *J. Differential Equations* **251**:8 (2011), 2164–2195. MR 2012i:35372 Zbl 1227.35236

[Mio et al. 1976] K. Mio, T. Ogino, K. Minami, and S. Takeda, "Modified nonlinear Schrödinger equation for Alfvén waves propagating along the magnetic field in cold plasmas", *J. Phys. Soc. Japan* **41**:1 (1976), 265–271. MR 57 #2116

[Mjolhus 1976] E. Mjolhus, "On the modulational instability of hydromagnetic waves parallel to the magnetic field", *J. Plasma Phys.* **16** (1976), 321–334.

[Nahmod et al. 2012] A. R. Nahmod, T. Oh, L. Rey-Bellet, and G. Staffilani, "Invariant weighted Wiener measures and almost sure global well-posedness for the periodic derivative NLS", *J. Eur. Math. Soc.* (*JEMS*) **14**:4 (2012), 1275–1330. MR 2928851 Zbl 1251.35151

[Nawa 1999] H. Nawa, "Asymptotic and limiting profiles of blowup solutions of the nonlinear Schrödinger equation with critical power", *Comm. Pure Appl. Math.* **52**:2 (1999), 193–270. MR 99m:35235 Zbl 0964.37014

[Ogawa and Tsutsumi 1991] T. Ogawa and Y. Tsutsumi, "Blow-up of H^1 solutions for the one-dimensional nonlinear Schrödinger equation with critical power nonlinearity", *Proc. Amer. Math. Soc.* **111**:2 (1991), 487–496. MR 91f:35026 Zbl 0747.35004

[Ozawa 1996] T. Ozawa, "On the nonlinear Schrödinger equations of derivative type", *Indiana Univ. Math. J.* **45**:1 (1996), 137–163. MR 98b:35186 Zbl 0859.35117

[Sulem and Sulem 1999] C. Sulem and P.-L. Sulem, *The nonlinear Schrödinger equation: self-focusing and wave collapse*, Applied Mathematical Sciences **139**, Springer, New York, 1999. MR 2000f:35139 Zbl 0928.35157

[Takaoka 1999] H. Takaoka, "Well-posedness for the one-dimensional nonlinear Schrödinger equation with the derivative nonlinearity", *Adv. Differential Equations* **4**:4 (1999), 561–580. MR 2000e:35221 Zbl 0951.35125

[Takaoka 2001] H. Takaoka, "Global well-posedness for Schrödinger equations with derivative in a nonlinear term and data in low-order Sobolev spaces", *Electron. J. Differential Equations* **26** (2001), 23 pp. MR 2002f:35033 Zbl 0972.35140

[Tan 2004] S. B. Tan, "Blow-up solutions for mixed nonlinear Schrödinger equations", *Acta Math. Sin.* (*Engl. Ser.*) **20**:1 (2004), 115–124. MR 2005c:35268 Zbl 1061.35139

[Weinstein 1982/83] M. I. Weinstein, "Nonlinear Schrödinger equations and sharp interpolation estimates", *Comm. Math. Phys.* **87**:4 (1982/83), 567–576. MR 84d:35140 Zbl 0527.35023

[Weinstein 1986] M. I. Weinstein, "On the structure and formation of singularities in solutions to nonlinear dispersive evolution equations", *Comm. Partial Differential Equations* **11**:5 (1986), 545–565. MR 87i:35026 Zbl 0596.35022

[Win 2010] Y. Y. S. Win, "Global well-posedness of the derivative nonlinear Schrödinger equations on T", *Funkcial. Ekvac.* **53**:1 (2010), 51–88. MR 2011j:35229 Zbl 1194.35433

Received 9 Mar 2013. Revised 11 Sep 2013. Accepted 4 Oct 2013.

YIFEI WU: yifei@bnu.edu.cn
School of Mathematical Science, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing, 100875, China

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 6, No. 8, 2013
dx.doi.org/10.2140/apde.2013.6.2003

THE CALDERÓN PROBLEM WITH PARTIAL DATA ON MANIFOLDS AND APPLICATIONS

CARLOS KENIG AND MIKKO SALO

We consider Calderón's inverse problem with partial data in dimensions $n \geq 3$. If the inaccessible part of the boundary satisfies a (conformal) flatness condition in one direction, we show that this problem reduces to the invertibility of a broken geodesic ray transform. In Euclidean space, sets satisfying the flatness condition include parts of cylindrical sets, conical sets, and surfaces of revolution. We prove local uniqueness in the Calderón problem with partial data in admissible geometries, and global uniqueness under an additional concavity assumption. This work unifies two earlier approaches to this problem — one by Kenig, Sjöstrand, and Uhlmann, the other by Isakov — and extends both. The proofs are based on improved Carleman estimates with boundary terms, complex geometrical optics solutions involving reflected Gaussian beam quasimodes, and invertibility of (broken) geodesic ray transforms. This last topic raises questions of independent interest in integral geometry.

1. Introduction

This article is concerned with inverse problems where measurements are made only on part of the boundary. A typical example is the inverse problem of Calderón, where the objective is to determine the electrical conductivity of a medium from voltage and current measurements on its boundary. The mathematical formulation of this problem is as follows. Let $\Omega \subset \mathbb{R}^n$, $n \geq 2$, be a bounded domain with smooth boundary. Given a positive function $\gamma \in L^\infty(\Omega)$ (the electrical conductivity of the medium) and two open subsets Γ_D, Γ_N of $\partial\Omega$, consider the partial Cauchy data set

$$C_\gamma^{\Gamma_D,\Gamma_N} = \{(u|_{\Gamma_D}, \gamma\partial_\nu u|_{\Gamma_N}) : \operatorname{div}(\gamma\nabla u) = 0 \text{ in } \Omega,\ u \in H^1(\Omega),\ \operatorname{supp}(u|_{\partial\Omega}) \subset \Gamma_D\}.$$

MSC2010: primary 35R30; secondary 35J10, 58J32.
Keywords: Calderón problem, partial data, inverse problem.

The Calderón problem with partial data is to determine the conductivity γ from the knowledge of $C_\gamma^{\Gamma_D,\Gamma_N}$ for possibly very small sets Γ_D, Γ_N. Here ∂_ν is the normal derivative, and the conormal derivative $\gamma\partial_\nu u|_{\partial\Omega}$ is interpreted in the weak sense as an element of $H^{-1/2}(\partial\Omega)$.

A closely related problem is to determine a potential $q \in L^\infty(\Omega)$ from partial boundary measurements for the Schrödinger equation, given by the partial Cauchy data set

$$C_q^{\Gamma_D,\Gamma_N} = \{(u|_{\Gamma_D}, \partial_\nu u|_{\Gamma_N}) : (-\Delta+q) = 0 \text{ in } \Omega,\ u \in H_\Delta(\Omega),\ \operatorname{supp}(u|_{\partial\Omega}) \subset \Gamma_D\}.$$

Here we use the space

$$H_\Delta(\Omega) = \{u \in L^2(\Omega) : \Delta u \in L^2(\Omega)\},$$

and the trace $u|_{\partial\Omega}$ and normal derivative $\partial_\nu u|_{\partial\Omega}$ are in $H^{-1/2}(\partial\Omega)$ and $H^{-3/2}(\partial\Omega)$; see [Bukhgeim and Uhlmann 2002]. Above, one thinks of $u|_{\partial\Omega}$ as Dirichlet data prescribed only on Γ_D, and one measures the Neumann data of the corresponding solution on Γ_N. If $\Lambda_\gamma : H^{1/2}(\Omega) \to H^{-1/2}(\partial\Omega)$ is the Dirichlet-to-Neumann map (DN map) given by

$$\Lambda_\gamma : u|_{\partial\Omega} \mapsto \gamma\partial_\nu u|_{\partial\Omega}, \quad \text{where } u \in H^1(\Omega) \text{ solves } \operatorname{div}(\gamma\nabla u) = 0 \text{ in } \Omega,$$

then the partial Cauchy data set is a restriction of the graph of Λ_γ,

$$C_\gamma^{\Gamma_D,\Gamma_N} = \{(f|_{\Gamma_D}, \Lambda_\gamma f|_{\Gamma_N}) : f \in H^{1/2}(\partial\Omega),\ \operatorname{supp}(f) \subset \Gamma_D\}.$$

A similar interpretation is valid for $C_q^{\Gamma_D,\Gamma_N}$ provided that 0 is not a Dirichlet eigenvalue of $-\Delta+q$ in Ω.

The problems above are well studied questions in the theory of inverse problems. The case of full data ($\Gamma_D = \Gamma_N = \partial\Omega$) has received the most attention. Major results include [Sylvester and Uhlmann 1987; Haberman and Tataru 2013] in dimensions $n \geq 3$ and [Nachman 1996; Astala and Päivärinta 2006; Bukhgeim 2008] in the case $n = 2$. In particular, it is known that the set $C_\gamma^{\partial\Omega,\partial\Omega}$ determines uniquely a conductivity $\gamma \in C^1(\bar{\Omega})$ if $n \geq 3$ and a conductivity $\gamma \in L^\infty(\Omega)$ if $n = 2$. These results are based on the method of complex geometrical optics solutions developed in [Sylvester and Uhlmann 1987] for $n \geq 3$ and in [Nachman 1996; Bukhgeim 2008] in the case $n = 2$.

The partial data question where the sets Γ_D or Γ_N may not be the whole boundary has also attracted considerable attention. We mention here four approaches, each of which gives a slightly different partial data result. Formulated in terms of the Schrödinger problem, it is known that $C_q^{\Gamma_D,\Gamma_N}$ determines q in Ω in the following cases:

(1) $n \geq 3$, the set Γ_D is possibly very small, and Γ_N is slightly larger than $\partial\Omega \setminus \bar{\Gamma}_D$; proved by Kenig, Sjöstrand, and Uhlmann [Kenig et al. 2007].

(2) $n \geq 3$ and $\Gamma_D = \Gamma_N = \Gamma$, and $\partial\Omega \setminus \Gamma$ is either part of a hyperplane or part of a sphere; proved in [Isakov 2007].

(3) $n = 2$ and $\Gamma_D = \Gamma_N = \Gamma$, where Γ can be an arbitrary open subset of $\partial\Omega$; proved by Imanuvilov, Uhlmann, and Yamamoto [Imanuvilov et al. 2010].

(4) $n \geq 2$, linearized partial data problem, $\Gamma_D = \Gamma_N = \Gamma$, where Γ can be an arbitrary open subset of $\partial\Omega$; proved by dos Santos Ferreira, Kenig, Sjöstrand, and Uhlmann [Ferreira 2009b].

Approaches (1)–(3) also give a partial data result of determining γ from $C_\gamma^{\Gamma_D,\Gamma_N}$ with the same assumptions on the dimension and the sets Γ_D, Γ_N. In (4), the linearized partial data problem is to show injectivity of the Fréchet derivative of Λ_q at $q=0$ instead of injectivity of the full map $q \mapsto \Lambda_q$, when restricted to the sets Γ_D and Γ_N.

It is interesting that, although each of the four approaches is based on a version of complex geometrical optics solutions, the approaches are distinct in the sense that none of the above results is contained in any of the others. The result in [Kenig et al. 2007] uses Carleman estimates with boundary terms, given for special limiting weights, that allow one to control the solutions on parts of the boundary, whereas [Isakov 2007] is based on the full data arguments of [Sylvester and Uhlmann 1987] and a reflection argument. The result in [Imanuvilov et al. 2010] is a strong one that only requires Dirichlet and Neumann data on any small set, but the method involves complex analysis and Carleman weights with critical points and does not obviously extend to higher dimensions. Finally, [Ferreira 2009b] is based on analytic microlocal analysis but is so far restricted to the linearized problem.

Nevertheless, given that there exist several approaches to the same problem, one expects that a combination of ideas from different approaches might lead to improved partial data results. In this paper we unify the Carleman estimate approach of [Kenig et al. 2007] and the reflection approach of [Isakov 2007], and, in fact, we obtain the main results of both these papers as special cases.

The method also allows us to improve both approaches. Concerning [Isakov 2007], we are able to relax the hypothesis on the inaccessible part $\Gamma_i = \partial\Omega \setminus \Gamma$ of the boundary: instead of requiring Γ_i to be completely flat (or spherical), we can deal with Γ_i that satisfy a flatness condition only in one direction. Compared with [Kenig et al. 2007], we remove the need for measurements on certain parts of the boundary that are flat in one direction; and, in certain cases where $\partial\Omega$ may not have any symmetries, we eliminate the overlap of Γ_D and Γ_N needed in [Kenig et al. 2007]. The method eventually boils down to inverting geodesic ray transforms (possibly for broken geodesics). In some cases the invertibility of the ray transform is known, but in other cases it is not, and in these cases we obtain a reduction from the Calderón problem with partial data to integral geometry problems of independent interest.

The survey [Kenig and Salo 2013] describes earlier results on the Calderón problem with partial data and also the results in the present paper. However, we also list here some further references for partial data results, first for the case $n \geq 3$. The Carleman estimate approach was initiated in [Bukhgeim and Uhlmann 2002; Kenig et al. 2007]. Based on this approach, there are low regularity results [Knudsen 2006; Zhang 2012], results for other scalar equations [Ferreira 2007; Knudsen and Salo 2007; Chung 2012] and systems [Salo and Tzou 2010; Chung et al. 2013], stability results [Heck and Wang 2006], and reconstruction results [Nachman and Street 2010]. The reflection approach was introduced in [Isakov 2007], and has been employed for the Maxwell system [Caro et al. 2009]. Partial data results for slab geometries are given in [Li and Uhlmann 2010; Krupchyk et al. 2012]. Also, at the same time as this preprint was first submitted, a preprint of Imanuvilov and Yamamoto [2013a] appeared that independently proves a result similar to that in Section 3A in this paper.

In two dimensions, the main partial data result is that of [Imanuvilov et al. 2010], which has been extended in [Imanuvilov et al. 2011a; 2011b; Imanuvilov and Yamamoto 2012a; 2012b] to, respectively, more general equations, combinations of measurements on disjoint sets, less regular coefficients, and certain systems. An earlier result is [Astala et al. 2005]. In the case of Riemann surfaces with boundary, corresponding partial data results are given in [Guillarmou and Tzou 2011a; 2011b; Albin et al. 2013]. See also the surveys [Guillarmou and Tzou 2013; Imanuvilov and Yamamoto 2013b].

In the case when the conductivity is known near the boundary, the partial data problem can be reduced to the full data problem [Ammari and Uhlmann 2004; Alessandrini and Kim 2012; Hyvönen et al. 2012]. Also, we remark that in the corresponding problem for the wave equation, it has been known for a long time (see [Katchalov et al. 2001]) that measuring the Dirichlet and Neumann data of waves on an arbitrary open subset of the boundary is sufficient to determine the coefficients uniquely up to natural gauge transforms. Partial results for the case where Dirichlet and Neumann data are measured on disjoint sets are in [Lassas and Oksanen 2010; 2012].

The structure of this paper is as follows. Section 2 states our main partial data results in the setting of Riemannian manifolds, and Section 3 considers some consequences of the Calderón problem with partial data in Euclidean space. Section 4 gives a Carleman estimate that is used to control solutions on parts of the boundary, and Section 5 discusses a reflection approach that can be used as an alternative to Carleman estimates in some cases. In Section 6 we give the proofs of the local uniqueness results for simple transversal manifolds, based on complex geometrical optics solutions involving WKB type quasimodes. In Section 7 we discuss a more sophisticated quasimode construction based on reflected Gaussian beams, and in Section 8 we show how complex geometrical optics solutions involving reflected Gaussian beam quasimodes can be used to recover the broken ray transform of a potential from partial Cauchy data.

2. Statement of results

Our method is based on ideas developed for the anisotropic Calderón problem in [Ferreira 2009a], and even though much of the motivation comes from the Calderón problem with partial data in Euclidean domains, it is convenient to formulate our main results in the setting of manifolds. The Riemannian geometry notation we use is mostly that of [Ferreira 2009a].

Definition. Let (M, g) be a compact oriented Riemannian manifold with C^∞ boundary, and let $n = \dim(M) \geq 3$.

1. We say that (M, g) is *conformally transversally anisotropic* (or CTA) if
$$(M, g) \Subset (\mathbb{R} \times \widehat{M}_0, g), \quad g = c(e \oplus g_0),$$
where $(\widehat{M}_0, g_0)$ is some compact $(n-1)$-dimensional manifold with boundary, e is the Euclidean metric on the real line, and c is a smooth positive function in the cylinder $\mathbb{R} \times \widehat{M}_0$.
2. We say that (M, g) is *admissible* if it is CTA and additionally the transversal manifold $(\widehat{M}_0, g_0)$ is *simple*, meaning that the boundary $\partial \widehat{M}_0$ is strictly convex (the second fundamental form is positive definite) and for each $p \in \widehat{M}_0$, the exponential map $\exp_p$ is a diffeomorphism from its maximal domain of definition in $T_p \widehat{M}_0$ onto $\widehat{M}_0$.

The uniqueness results in [Ferreira 2009a] were given for admissible manifolds. In this paper we will give results both for admissible and CTA manifolds. In the main results, we will also assume that there is a compact $(n-1)$-dimensional manifold (M_0, g_0) with smooth boundary such that

$$(M, g) \subset (\mathbb{R} \times M_0, g) \Subset (\mathbb{R} \times \widehat{M}_0, g), \quad g = c(e \oplus g_0), \tag{2-1}$$

and the following intersection is nonempty:

$$\partial M \cap (\mathbb{R} \times \partial M_0) \neq \varnothing.$$

Under some conditions, it will be possible to ignore boundary measurements in the set $\partial M \cap (\mathbb{R} \times \partial M_0)$. In the results below, we will implicitly assume that the various manifolds satisfy (2-1), and if (M, g) is admissible, it is also assumed that $(\widehat{M}_0, g_0)$ is simple (but (M_0, g_0) need not be simple, since its boundary may not be strictly convex).

Write $x = (x_1, x')$ for points in $\mathbb{R} \times \widehat{M}_0$, where x_1 is the Euclidean coordinate. The approaches of [Kenig et al. 2007; Ferreira 2009a] are based on complex geometrical optics solutions of the form $u = e^{\tau\varphi}(m + r)$, where φ is a special *limiting Carleman weight*. We refer to the latter paper for the definition and properties of limiting Carleman weights on manifolds. For present purposes, we only mention that the functions $\varphi(x) = \pm x_1$ are natural limiting Carleman weights in the cylinder $(\mathbb{R} \times \widehat{M}_0, g)$.

The weight $\varphi(x) = x_1$ allows us to decompose the boundary ∂M as the disjoint union

$$\partial M = \partial M_+ \cup \partial M_- \cup \partial M_{\mathrm{tan}},$$

where

$$\partial M_\pm = \{x \in \partial M : \pm \partial_\nu \varphi(x) > 0\} \quad \text{and} \quad \partial M_{\mathrm{tan}} = \{x \in \partial M : \partial_\nu \varphi(x) = 0\}.$$

Here the normal derivative is understood with respect to the metric g. Note that $\partial_\nu \varphi = 0$ on $\mathbb{R} \times \partial M_0$ whenever $(M_0, g_0) \Subset (\widehat{M}_0, g_0)$. We think of $\partial M_{\mathrm{tan}}$ as being flat in one direction (the direction of the gradient of φ). For the sake of definiteness, the sets $\partial M_\pm = \partial M_\pm(\varphi)$ will refer to the weight $\varphi(x) = x_1$ in this section, but all results remain true when ∂M_+ and ∂M_- are interchanged (this amounts to replacing the weight x_1 by $-x_1$).

Next we give the local results for the Calderón problem with partial data on manifolds. In these results we say that a unit speed geodesic $\gamma : [0, L] \to M_0$ is *nontangential* if its endpoints are on ∂M_0, the vectors $\dot\gamma(0)$, $\dot\gamma(L)$ are nontangential, and $\gamma(t) \in M_0^{\mathrm{int}}$ for $0 < t < L$. We also define the partial Cauchy data set as

$$C_{g,q}^{\Gamma_D, \Gamma_N} = \{(u|_{\Gamma_D}, \partial_\nu u|_{\Gamma_N}) : (-\Delta_g + q) = 0 \text{ in } M,\ u \in H_{\Delta_g}(M), \operatorname{supp}(u|_{\partial M}) \subset \Gamma_D\},$$

where $H_{\Delta_g}(M) = \{u \in L^2(M) : \Delta_g u \in L^2(M)\}$ and $u|_{\partial M} \in H^{-1/2}(\partial M)$, $\partial_\nu u|_{\partial M} \in H^{-3/2}(\partial M)$ by the same arguments as in [Bukhgeim and Uhlmann 2002].

To explain the results, it is convenient to think in terms of the following special case.

Example. Let $M = M_{\mathrm{left}} \cup M_{\mathrm{mid}} \cup M_{\mathrm{right}}$ be a compact manifold with boundary consisting of three parts: $M_{\mathrm{mid}} = [a, b] \times M_0$ for some compact manifold (M_0, g_0) with boundary, $M_{\mathrm{left}} \subset \{x_1 < a\} \times M_0$, and

$M_{\text{right}} \subset \{x_1 > b\} \times M_0$. We also assume that $\partial M_- = M_{\text{left}} \cap \partial M$ and $\partial M_+ = M_{\text{right}} \cap \partial M$. In this case $\partial M_{\tan} = [a, b] \times \partial M_0$.

The methods developed in this paper suggest that it should suffice to measure Neumann data on ∂M_+ for Dirichlet data supported in ∂M_-, with no measurements required on $\partial M_{\tan}$. However, in the results below we need a part $\Gamma_a \subset \partial M_{\tan}$ that is accessible to measurements, and $\Gamma_i = \partial M_{\tan} \setminus \Gamma_a$ is the inaccessible part. Suppose for simplicity that

$$\Gamma_a = [a, b] \times E, \qquad \Gamma_i = [a, b] \times (\partial M_0 \setminus E)$$

for some nonempty open subset E of ∂M_0.

In this setting, Theorem 2.1 implies that from Neumann data measured near $\partial M_+ \cup \Gamma_a$ with Dirichlet data input near $\partial M_- \cup \Gamma_a$, one can determine certain integrals of the potential q in the set

$$\mathbb{R} \times \bigcup_{\gamma} \gamma([0, L]),$$

where the union is over all nontangential geodesics in M_0 with endpoints on E. Moreover, if the local ray transform is injective in this set in a suitable sense, one can determine the potential in this set by Theorem 2.2. Theorem 2.4 shows that one can go beyond this set and extract information about integrals of q over all nontangential broken rays with endpoints on E, and Theorem 2.3 gives a global uniqueness result in the case where $\partial M_{\tan}$ has zero measure.

Theorem 2.1. *Let (M, g) be an admissible manifold as in* (2-1), *and let $q_1, q_2 \in C(M)$. Let Γ_i be a closed subset of $\partial M_{\tan}$, and suppose that, for some nonempty open subset E of ∂M_0, one has*

$$\Gamma_i \subset \mathbb{R} \times (\partial M_0 \smallsetminus E).$$

Let $\Gamma_a = \partial M_{\tan} \smallsetminus \Gamma_i$, and assume that

$$C_{g,q_1}^{\Gamma_D, \Gamma_N} = C_{g,q_2}^{\Gamma_D, \Gamma_N},$$

where Γ_D and Γ_N are any open sets in ∂M such that $\Gamma_D \supset \partial M_- \cup \Gamma_a$ and $\Gamma_N \supset \partial M_+ \cup \Gamma_a$.

Given any nontangential geodesic $\gamma : [0, L] \to M_0$ with endpoints on E, and given any real number λ, one has

$$\int_0^L e^{-2\lambda t} (c(q_1 - q_2))\hat{\ }(2\lambda, \gamma(t))\, dt = 0.$$

Here $q_1 - q_2$ is extended by zero outside M, and $(\,\cdot\,)\hat{\ }$ denotes the Fourier transform in the x_1 variable.

The previous theorem allows us to conclude uniqueness of potentials in sets where the local ray transform is injective in the following sense.

Definition. Let (M_0, g_0) be a compact oriented manifold with smooth boundary, and let O be an open subset of M_0. We say that the local ray transform is injective on O if any function $f \in C(M_0)$ with

$$\int_\gamma f\, dt = 0 \quad \text{for all nontangential geodesics } \gamma \text{ contained in } O$$

must satisfy $f|_O = 0$.

Theorem 2.2. *Assume the conditions in Theorem 2.1. Then $q_1 = q_2$ in $M \cap (\mathbb{R} \times O)$ for any open subset O of M_0 such that the local ray transform is injective on O and $O \cap \partial M_0 \subset E$.*

The local ray transform is known to be injective in the next three cases (the second case will be used in Section 3):

1. $(M_0, g_0) = (\overline{\Omega}_0, e)$, where $\Omega_0 \subset \mathbb{R}^{n-1}$ is a bounded domain with C^∞ boundary, e is the Euclidean metric, E is an open subset of $\partial\Omega_0$, and O is the intersection of $\overline{\Omega}_0$ with the union of all hyperplanes in $\mathbb{R}^{n-1}$ that have $\partial\Omega_0 \setminus E$ on one side. The complement of this union is the intersection of half-spaces and is thus convex. If the integral of $f \in C(\overline{\Omega}_0)$, extended by zero to $\mathbb{R}^{n-1}$, vanishes over all line segments in O, the integral over all hyperplanes that do not meet $\partial\Omega_0 \setminus E$ also vanishes, and it follows from the Helgason support theorem [1999] that the local ray transform is injective on O.
2. $(M_0, g_0) \Subset (\widetilde{M}_0, g_0)$ are simple manifolds with real-analytic metric, and $\widetilde{\mathscr{F}}$ is an open set of nontangential geodesics in $(\widetilde{M}_0, g_0)$ such that any curve in $\widetilde{\mathscr{F}}$ can be deformed to a point on $\partial\widetilde{M}_0$ through curves in $\widetilde{\mathscr{F}}$. In such a case, by a result of Krishnan [2009] the local ray transform is injective on the set O of all points in M_0 that lie on some geodesic in $\widetilde{\mathscr{F}}$.
3. If $\dim(M_0) \geq 3$ and if ∂M_0 is strictly convex at a point $p \in \partial M_0$, then p has a neighborhood O in M_0 on which the local ray transform is injective. This is a result from [Uhlmann and Vasy 2012].

In Theorem 2.2, if the nontangential geodesics with endpoints on E cover a dense subset O of M_0 and if the local ray transform is injective in O, we obtain a global uniqueness result stating that $q_1 = q_2$ in M. An example of such a result under a concavity assumption is given in Section 3F.

The method for proving Theorems 2.1 and 2.2 also allows us to reduce the overlap for Γ_D and Γ_N needed in [Kenig et al. 2007]. An example of such a result is the following (a similar result was also proved in [Imanuvilov and Yamamoto 2013b]).

Theorem 2.3. *Let (M, g) be an admissible manifold and assume that $q_1, q_2 \in C(M)$. If $\partial M_{\mathrm{tan}}$ has zero measure in ∂M, then*

$$C_{g,q_1}^{\partial M_-, \partial M_+} = C_{g,q_2}^{\partial M_-, \partial M_+} \implies q_1 = q_2.$$

Next we wish to gather information on the potentials beyond the set that can be reached by transversal geodesics with endpoints on E. To do this, we will use broken geodesics in the transversal manifold that go inside M_0, reflect finitely many times, and eventually return to E.

Definition. Let (M_0, g_0) be a compact manifold with boundary.

(a) We call a continuous curve $\gamma : [a, b] \to M_0$ a *broken ray* if γ is obtained by following unit speed geodesics that are reflected according to geometrical optics (angle of incidence equals angle of reflection) whenever they hit a point of ∂M_0.

(b) A broken ray $\gamma : [0, L] \to M_0$ is called *nontangential* if $\dot{\gamma}(t)$ is nontangential whenever $\gamma(t) \in \partial M_0$, and additionally all points of reflection are distinct.

The next theorem is a generalization of Theorem 2.1 in the sense that it allows arbitrary transversal manifolds and recovers integrals over all nontangential broken rays (instead of just nontangential geodesics) with endpoints on E. However, it is stated with a weaker partial data condition.

Theorem 2.4. *Let (M, g) be a CTA manifold as in (2-1), and let $q_1, q_2 \in C(M)$. Let Γ_i be a closed subset of $\partial M_{\mathrm{tan}}$, and suppose that, for some nonempty open subset E of ∂M_0, one has*

$$\Gamma_i \subset \mathbb{R} \times (\partial M_0 \setminus E).$$

Let $\Gamma_a = \partial M_{\mathrm{tan}} \setminus \Gamma_i$, and assume that

$$C_{g,q_1}^{\Gamma_D,\Gamma_N} = C_{g,q_2}^{\Gamma_D,\Gamma_N},$$

where $\Gamma_D = \Gamma_N = \Gamma$ for some neighborhood Γ of the set $\overline{\partial M_+ \cup \partial M_- \cup \Gamma_a}$ in ∂M.

Given any nontangential broken ray $\gamma : [0, L] \to M_0$ with endpoints on E, and given any real number λ, one has

$$\int_0^L e^{-2\lambda t} (c(q_1 - q_2))\hat{\ }(2\lambda, \gamma(t))\, dt = 0.$$

Here $q_1 - q_2$ is extended by zero outside M, and $(\,\cdot\,)\hat{\ }$ denotes the Fourier transform in the x_1 variable.

It is natural to ask whether a function in M_0 is determined by its integrals over broken rays with endpoints in some subset E of ∂M_0 (that is, whether the broken ray transform is injective). Combined with Theorem 2.4 and with the proof of Theorem 2.2, such a result would imply unique recovery of the potential in the whole manifold M. However, it seems that there are very few results in this direction, except for the case where E is the whole boundary and the question reduces to the injectivity of the usual ray transform; see [Sharafutdinov 1994].

Eskin [2004] has proved injectivity in the case of Euclidean broken rays reflecting off several convex obstacles, with E being the boundary of a smooth domain enclosing all the obstacles, if the obstacles satisfy additional restrictions (in particular, the obstacles must have corner points and they cannot be smooth). Hubenthal [2013a; 2013b] and Ilmavirta [2013a; 2013b; 2013c] have given partial results for the broken ray transform in special geometries. See also [Florescu et al. 2011; Lozev 2013] for related results. However, the following question seems to be open even in convex Euclidean domains except when $E = \partial M_0$.

Question. *Let (M_0, g_0) be a simple manifold, let E be a nonempty open subset of ∂M_0, and assume that $f \in C(M_0)$ satisfies*

$$\int_0^L f(\gamma(t))\, dt = 0$$

for all nontangential broken rays $\gamma : [0, L] \to M_0$ with endpoints on E. Does this imply that $f = 0$?

3. The Euclidean case

In this section, we indicate some consequences of the previous results to the Calderón problem with partial data in Euclidean space. We assume that $\Omega \subset \mathbb{R}^3$ is a bounded domain with smooth boundary equipped with the Euclidean metric $g = e$, and $q_1, q_2 \in C(\bar{\Omega})$. We also assume that

$$C_{q_1}^{\Gamma,\Gamma} = C_{q_2}^{\Gamma,\Gamma},$$

where Γ is some strict open subset of $\partial\Omega$. Write

$$\Gamma_i = \partial\Omega \setminus \Gamma$$

for the inaccessible part of the boundary. The results in this section show that in cases where Γ_i satisfies certain geometric restrictions, it is possible to conclude that

$$q_1 = q_2 \quad \text{in } \bar{\Omega} \cap (\mathbb{R} \times O),$$

where the sets $O \subset \mathbb{R}^2$ will be described below.

Remark. We also obtain results for the conductivity equation by making a standard reduction to the Schrödinger equation. More precisely, if $\gamma_1, \gamma_2 \in C^2(\bar{\Omega})$ are positive functions such that $C_{\gamma_1}^{\Gamma,\Gamma} = C_{\gamma_2}^{\Gamma,\Gamma}$, the corresponding DN maps satisfy

$$\Lambda_{\gamma_1} f|_\Gamma = \Lambda_{\gamma_2} f|_\Gamma \quad \text{for } f \in H^{1/2}(\partial\Omega) \text{ with } \operatorname{supp}(f) \subset \Gamma.$$

Boundary determination [Kohn and Vogelius 1984; Sylvester and Uhlmann 1988] implies that

$$\gamma_1|_\Gamma = \gamma_2|_\Gamma, \quad \partial_\nu \gamma_1|_\Gamma = \partial_\nu \gamma_2|_\Gamma.$$

Writing $q_j = \Delta\gamma_j^{1/2}/\gamma_j^{1/2}$, the relation

$$\Lambda_{q_j} f = \gamma_j^{-1/2} \Lambda_{\gamma_j}(\gamma_j^{-1/2} f) + \tfrac{1}{2}\gamma_j^{-1}(\partial_\nu \gamma_j) f|_{\partial\Omega}$$

and the above conditions imply that the DN maps Λ_{q_j} for the Schrödinger equations satisfy

$$\Lambda_{q_1} f|_\Gamma = \Lambda_{q_2} f|_\Gamma \quad \text{for } f \in H^{1/2}(\partial\Omega) \text{ with } \operatorname{supp}(f) \subset \Gamma.$$

Thus $C_{q_1}^{\Gamma,\Gamma} = C_{q_2}^{\Gamma,\Gamma}$, and we obtain that

$$q_1 = q_2 \quad \text{in } \bar{\Omega} \cap (\mathbb{R} \times O).$$

Write $q = q_1 = q_2$ in $\bar{\Omega} \cap (\mathbb{R} \times O)$. Then $\gamma_1^{1/2}$ and $\gamma_2^{1/2}$ are both solutions of $(-\Delta + q)u = 0$ in $\bar{\Omega} \cap (\mathbb{R} \times O)$ having identical Cauchy data on Γ. It follows that $\gamma_1 = \gamma_2$ in any connected component of $\bar{\Omega} \cap (\mathbb{R} \times O)$ whose intersection with Γ contains a nonempty open subset of $\partial\Omega$.

In the following we will use some general facts on limiting Carleman weights from [Ferreira 2009a], where it was proved that any limiting Carleman weight in $\mathbb{R}^3$ has, up to translation, rotation and scaling, one of the following six forms:

$$x_1, \quad \log|x|, \quad \arg(x_1 + i x_2), \quad \frac{x_1}{|x|^2}, \quad \log\frac{|x+e_1|^2}{|x-e_1|^2}, \quad \arg(e^{i\theta}(x + i e_1)^2).$$

Here $\theta \in [0, 2\pi)$, and the argument function is defined by

$$\arg(z) = 2\arctan\frac{\operatorname{Im}(z)}{|z| + \operatorname{Re}(z)}, \quad z \in \mathbb{C} \setminus \{t \in \mathbb{R} : t \le 0\}.$$

It was also proved in Section 2 of [Ferreira 2009a] that if φ is a limiting Carleman weight near $(\overline{\Omega}, e)$, then $\nabla_{\tilde{g}}\varphi$ is a unit parallel vector field near $(\overline{\Omega}, \tilde{g})$ where

$$\tilde{g} = c^{-1}e, \quad c = |\nabla_e \varphi|_e^{-2}.$$

Furthermore, by the proof of Lemma A.5 of the same reference, if (y_1, y') are coordinates so that $\nabla_{\tilde{g}}\varphi = \partial_{y_1}$ and if the coordinates y' parametrize a 2-dimensional manifold S such that $\nabla_{\tilde{g}}\varphi$ is orthogonal to S with respect to the $\tilde{g}$ metric, then the metric has the form

$$\tilde{g}(y_1, y') = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{g}_0(y') \end{pmatrix},$$

where $\tilde{g}_0$ is the metric on S induced by $\tilde{g}$.

3A. ***Cylindrical sets.*** This case corresponds to the limiting Carleman weight $\varphi(x) = x_1$. Suppose that $\Omega \subset \mathbb{R} \times \Omega_0$, where Ω_0 is a bounded domain with smooth boundary in $\mathbb{R}^2$. Let E be an open subset of $\partial\Omega_0$, and assume that

$$\Gamma_i \subset \mathbb{R} \times (\partial\Omega_0 \setminus E).$$

If Ω_0 has strictly convex boundary, Theorem 2.2 and the result of [Krishnan 2009] imply that

$$q_1 = q_2 \quad \text{in } \overline{\Omega} \cap (\mathbb{R} \times O),$$

where O is the intersection of $\overline{\Omega}_0$ with the union of all lines in $\mathbb{R}^2$ that have $\partial\Omega_0 \setminus E$ on one side.

The above conclusion holds true also when Ω_0 does not have strictly convex boundary. To see this, let $\Omega_0 \Subset B \Subset \widetilde{B}$, where B and $\widetilde{B}$ are balls. The extensions of the line segments in O to $\widetilde{B}$ form a class $\widetilde{\mathscr{F}}$ such that any curve in $\widetilde{\mathscr{F}}$ can be deformed to a point through curves in $\widetilde{\mathscr{F}}$. It is then enough to extend $q_1 - q_2$ by zero to $\mathbb{R} \times \overline{B}$, and to use the proof of Theorem 2.2 with M_0 replaced by $\overline{B}$, together with [Krishnan 2009].

3B. ***Conical sets.*** Consider the limiting Carleman weight $\varphi(x) = \log|x|$. Suppose that $\overline{\Omega} \subset \{x_3 > 0\}$, let (S^2, g_0) be the sphere with its standard metric, let $S^2_+ = \{\omega \in S^2 : \omega_3 > 0\}$, and let (M_0, g_0) be a compact submanifold of (S^2_+, g_0) with smooth boundary. Let E be an open subset of ∂M_0, and assume that

$$\Gamma_i \subset \{r\omega : r > 0, \omega \in \partial M_0 \setminus E\}.$$

We have $c = |\nabla\varphi|^{-2} = |x|^2$ and $\tilde{g} = |x|^{-2}e$, $\nabla_{\tilde{g}}\varphi = x$. Choose coordinates so that

$$y_1 = \log|x|, \quad y' = x/|x|.$$

The coordinates y' parametrize the manifold S^2 and the metric $\tilde{g}_0$ on S^2 induced by $\tilde{g}$ is just the standard metric g_0. The discussion in the beginning of this section shows that

$$\tilde{g}(y_1, y') = \begin{pmatrix} 1 & 0 \\ 0 & g_0(y') \end{pmatrix}.$$

Now (M_0, g_0) is contained in some simple submanifold $(\widehat{M}_0, g_0)$ of the hemisphere (S^2_+, g_0) (just remove a neighborhood of the equator). Since geodesics in S^2_+ are restrictions of great circles, Theorem 2.2 and the local injectivity result [Krishnan 2009] imply, as in Section 3A, that

$$q_1 = q_2 \quad \text{in } \bar{\Omega} \cap \{r\omega : r > 0, \omega \in O\},$$

where O is the union of all great circle segments in S^2_+ such that $\partial M_0 \setminus E$ is on one side of the hyperplane containing the great circle segment.

3C. ***Surfaces of revolution.*** Let $\Omega \subset \mathbb{R}^3 \setminus \{x : x_1 \leq 0\}$, and consider the limiting Carleman weight

$$\varphi(x) = \arg(x_1 + ix_2).$$

Then

$$\nabla\varphi = \left(\frac{-x_2}{x_1^2 + x_2^2}, \frac{x_1}{x_1^2 + x_2^2}, 0\right)$$

and

$$c = x_1^2 + x_2^2, \quad \tilde{g} = \frac{1}{x_1^2 + x_2^2} e, \quad \nabla_{\tilde{g}}\varphi = (-x_2, x_1, 0).$$

We make the change of coordinates valid near $\bar{\Omega}$,

$$y_1 = \arg(x_1 + ix_2), \quad y_2 = \sqrt{x_1^2 + x_2^2}, \quad y_3 = x_3.$$

The coordinates y' parametrize the manifold $S = \{(x_1, 0, x_3) : x_1 > 0\}$ and $\nabla_{\tilde{g}}\varphi$ is orthogonal to S. Furthermore, we may also think of S as the set $\{(0, y_2, y_3) : y_2 > 0\}$, and the metric on S induced by $\tilde{g}$ is the hyperbolic metric $\tilde{g}_0 = (1/y_2^2)e$. The discussion in the beginning of this section shows that

$$\tilde{g}(y_1, y') = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{g}_0(y') \end{pmatrix}.$$

Let (M_0, g_0) be a compact submanifold of S with smooth boundary, let E be an open subset of ∂M_0. We think of M_0 as lying in $\{(x_1, 0, x_3) : x_1 > 0\}$. Now, assume that

$$\Gamma_i \subset \{R_\theta(\partial M_0 \setminus E) : \theta \in (-\pi, \pi)\},$$

where $R_\theta x = (\widetilde{R}_\theta(x_1, x_2)^t, x_3)^t$ and $\widetilde{R}_\theta$ rotates vectors in $\mathbb{R}^2$ by angle θ counterclockwise. That is, we assume that the inaccessible part Γ_i is contained in a surface of revolution obtained by rotating the boundary curve $\partial M_0 \setminus E$.

Now, the geodesics in S (and, after restriction, also in M_0) have either the form

$$(y_2(t), y_3(t)) = (R \sin t, R \cos t + \alpha),$$

where $t \in (0, \pi)$, $R > 0$, and $\alpha \in \mathbb{R}$, or the form $(y_2(t), y_3(t)) = (t, \alpha)$, where $t > 0$ and $\alpha \in \mathbb{R}$ (these are not unit speed parametrizations). In the x coordinates, these are either the half circles in the $\{x_2 = 0\}$ plane given by

$$(x_1(t), x_2(t), x_3(t)) = (R \sin t, 0, R \cos t + \alpha), \quad t \in (0, \pi),$$

or the lines

$$(x_1(t), x_2(t), x_3(t)) = (t, 0, \alpha), \quad t > 0.$$

Enclosing M_0 in some ball B in S, the manifold $(\bar{B}, g_0)$ is simple and Theorem 2.2 and [Krishnan 2009] imply, as in Section 3A, that

$$q_1 = q_2 \quad \text{in } \bar{\Omega} \cap \{R_\theta(O) : \theta \in (-\pi, \pi)\},$$

where O is the union of all geodesics in S that have $\partial M_0 \setminus E$ on one side.

3D. ***Other limiting Carleman weights.*** So far we have considered three of the six possible forms of limiting Carleman weights in $\mathbb{R}^3$. The fourth one, $\varphi(x) = x_1/|x|^2$, is the Kelvin transform of the linear weight, and corresponds to inaccessible parts of the boundary that are Kelvin transforms of cylindrical domains. In particular, if part of the cylindrical domain is on the hyperplane $\{x_3 = 1\}$, its Kelvin transform lies on the sphere centered at $(0, 0, 1/2)$ with radius $1/2$, and we recover the result of Isakov [2007] for domains where the inaccessible part is part of a sphere. The corresponding results for the remaining two limiting Carleman weights do not seem so easy to state and we omit them.

3E. ***Extension of Kenig, Sjöstrand, and Uhlmann's result.*** Now let $\Omega \subset \mathbb{R}^3$ be a bounded domain with smooth boundary, assume that 0 is not in the convex hull of $\bar{\Omega}$, and let $\varphi(x) = \log|x|$. Define

$$\partial\Omega_\pm = \{x \in \partial\Omega : \pm\partial_\nu\varphi(x) > 0\}, \quad \partial\Omega_{\tan} = \{x \in \partial\Omega : \partial_\nu\varphi(x) = 0\}.$$

It was proved in [Kenig et al. 2007] that whenever Γ_D is a neighborhood of $\partial\Omega_- \cup \partial\Omega_{\tan}$ and Γ_N is a neighborhood of $\partial\Omega_+ \cup \partial\Omega_{\tan}$, we have

$$C_{q_1}^{\Gamma_D, \Gamma_N} = C_{q_2}^{\Gamma_D, \Gamma_N} \implies q_1 = q_2.$$

In particular, Γ_D and Γ_N always need to overlap. This result is a consequence of the reduction given above for the logarithmic weight, Theorem 2.1 (the special case where $E = \partial\Omega_0$, so that $\Gamma_i = \varnothing$), and injectivity of the ray transform. If $\partial\Omega_{\tan}$ has zero measure in $\partial\Omega$, then Theorem 2.3 allows us to improve this result: we have

$$C_{q_1}^{\partial\Omega_-, \partial\Omega_+} = C_{q_2}^{\partial\Omega_-, \partial\Omega_+} \implies q_1 = q_2.$$

In this case, the sets where Dirichlet and Neumann data are measured are disjoint, but their union covers all of $\partial\Omega$ except for a set of measure zero. The result remains true if the roles of $\partial\Omega_+$ and $\partial\Omega_-$ are changed.

3F. ***Extension of Isakov's result.*** According to [Isakov 2007], the condition $C_{q_1}^{\Gamma,\Gamma} = C_{q_2}^{\Gamma,\Gamma}$ implies $q_1 = q_2$ in Ω if $\Omega \subset \{x_3 > 0\}$ and $\Gamma_i \subset \{x_3 = 0\}$, or if $\Omega \subset B$ for some ball B and $\Gamma_i \subset \partial B$. We have already recovered these results in Sections 3A and 3D, since in these cases the local injectivity set O is so large that the result $q_1 = q_2$ holds in all of Ω. Of course, the results above also extend [Isakov 2007], since we can conclude at least local uniqueness for potentials when the inaccessible part of the boundary satisfies a (conformal) flatness condition in only one direction, such as being part of a cylindrical set, a conical set, or a surface of revolution.

We also get global uniqueness if the local injectivity set O is sufficiently large. For instance, if

$$\Omega \subset \mathbb{R} \times \Omega_0, \quad \Gamma_i \subset \mathbb{R} \times (\partial\Omega_0 \setminus E),$$

where Ω_0 is a bounded domain with smooth boundary and E is a nonempty open subset of $\partial\Omega_0$, and if the lines in $\mathbb{R}^2$ that have $\partial\Omega_0 \setminus E$ on one side cover a dense subset of Ω_0, then $q_1 = q_2$ in Ω. One example of this situation is if

$$\Omega \subset \mathbb{R} \times \{(x_2, x_3) : x_3 > \eta(x_2)\}, \quad \Gamma_i \subset \mathbb{R} \times \{(x_2, x_3) : x_3 = \eta(x_2)\},$$

where $\eta : \mathbb{R} \to \mathbb{R}$ is a smooth concave function.

4. Carleman estimate

Let (M, g) be a CTA manifold, so (M, g) is compact with boundary and

$$(M, g) \Subset (\mathbb{R} \times M_0, g), \quad g = c(e \oplus g_0).$$

Here (M_0, g_0) is any compact $(n-1)$-dimensional manifold with boundary. We wish to prove a Carleman estimate with boundary terms for the conjugated operator $e^{\varphi/h}(-\Delta_g)e^{-\varphi/h}$ in M, where φ is the limiting Carleman weight $\varphi(x) = x_1$ or $\varphi(x) = -x_1$, and $h > 0$ is small. Following [Kenig et al. 2007], it is useful to consider a slightly modified weight

$$\varphi_\varepsilon = \varphi + hf_\varepsilon$$

where f_ε is a smooth real-valued function in M depending on a small parameter ε, with ε independent of h. The convexity of f_ε will lead to improved lower bounds in terms of the $L^2(M)$ norms of u and $h\nabla u$. On the other hand, the sign of $\partial_\nu \varphi_\varepsilon$ in the boundary term of the Carleman estimate will allow us to control functions on different parts of the boundary. Of special interest is the set $\partial M_{\tan}$, where $\partial_\nu \varphi = 0$, and in this set we have

$$\partial_\nu \varphi_\varepsilon|_{\partial M_{\tan}} = h \partial_\nu f_\varepsilon.$$

We would like to have $\partial_\nu f_\varepsilon < 0$ on $\partial M_{\tan}$. It is not easy to find a global convex function f_ε satisfying the last condition for a general set $\partial M_{\tan}$. However, splitting f_ε into a convex part whose normal derivative vanishes on $\partial M_{\tan}$ and another part which ensures the correct sign on $\partial M_{\tan}$ will give the required result. We will use semiclassical conventions in the next proof; see [Ferreira 2009a, Section 4; Zworski 2012] for more details. We also write $(v, w) = (v, w)_{L^2(M)}$, $\|v\| = \|v\|_{L^2(M)}$, and for $\Gamma \subset \partial M$ we write $(v, w)_\Gamma = (v, w)_{L^2(\Gamma)}$.

Proposition 4.1. *Let (M, g) be as above, let $\varphi(x) = \pm x_1$, and let κ be a smooth real-valued function in M so that $\partial_\nu \kappa = -1$ on ∂M. Also let $q \in L^\infty(M)$. There are constants $\varepsilon, C_0, h_0 > 0$ with $h_0 \leq \varepsilon/2 \leq 1$ such that, for the weight*

$$\varphi_\varepsilon = \varphi + \frac{h}{\varepsilon}\frac{\varphi^2}{2} + h\kappa,$$

where $0 < h \le h_0$, *one has*

$$\frac{h^3}{C_0}(|\partial_\nu\varphi_\varepsilon|\partial_\nu u, \partial_\nu u)_{\partial M_-(\varphi_\varepsilon)} + \frac{h^2}{C_0}(\|u\|^2 + \|h\nabla u\|^2) \\ \le \|e^{\varphi/h}(-h^2\Delta_g + h^2 q)(e^{-\varphi/h}u)\|^2 + h^3(|\partial_\nu\varphi_\varepsilon|\partial_\nu u, \partial_\nu u)_{\partial M_+(\varphi_\varepsilon)}$$

for any $u \in C^\infty(M)$ *with* $u|_{\partial M} = 0$.

Proof. Since $\varphi(x) = \pm x_1$ is a limiting Carleman weight in a manifold strictly containing M, the computations in the proof of [Ferreira 2009a, Theorem 4.1] apply and we can follow that proof. First of all, note that

$$c^{(n+2)/4}(-\Delta_g + q)u = (-\Delta_{c^{-1}g} + q_c)(c^{(n-2)/4}u),$$

where $q_c = cq + c^{(n+2)/4}\Delta_g(c^{-(n-2)/4})$. Thus, by replacing q with another potential, we may assume that $c = 1$ so that $g = e \oplus g_0$ and φ is a distance function in the g metric, that is, $|\nabla_g\varphi|_g = 1$.

Let $P_0 = -h^2\Delta_g$ and $P_{0,\varphi_\varepsilon} = e^{\varphi_\varepsilon/h}P_0 e^{-\varphi_\varepsilon/h}$. Then $P_{0,\varphi_\varepsilon} = A + iB$, where A and B are the formally self-adjoint operators

$$A = -h^2\Delta_g - |\nabla\varphi_\varepsilon|^2, \quad B = -2i\langle\nabla\varphi_\varepsilon, h\nabla\,\cdot\,\rangle - ih\Delta_g\varphi_\varepsilon.$$

Assume $u \in C^\infty(M)$ and $u|_{\partial M} = 0$. We have

$$\begin{aligned} \|P_{0,\varphi_\varepsilon}u\|^2 &= ((A+iB)u, (A+iB)u) \\ &= \|Au\|^2 + \|Bu\|^2 + i(Bu, Au) - i(Au, Bu) \\ &= \|Au\|^2 + \|Bu\|^2 + (i[A, B]u, u) - ih^2(Bu, \partial_\nu u)_{\partial M} \\ &= \|Au\|^2 + \|Bu\|^2 + (i[A, B]u, u) - 2h^3((\partial_\nu\varphi_\varepsilon)\partial_\nu u, \partial_\nu u)_{\partial M}. \end{aligned}$$

Define

$$\tilde\varphi_\varepsilon(x) = \varphi + \frac{h}{\varepsilon}\frac{\varphi^2}{2}.$$

Thus $\varphi_\varepsilon = \tilde\varphi_\varepsilon + h\kappa$. Let

$$\widetilde{A} = -h^2\Delta - |\nabla\tilde\varphi_\varepsilon|^2, \quad \widetilde{B} = -2i\langle\nabla\tilde\varphi_\varepsilon, h\nabla\,\cdot\,\rangle - ih\Delta\tilde\varphi_\varepsilon.$$

Since $\Delta\varphi_\varepsilon = \Delta\tilde\varphi_\varepsilon + h\Delta\kappa$ and $\nabla\varphi_\varepsilon = \nabla\tilde\varphi_\varepsilon + h\nabla\kappa$, we have

$$\begin{aligned} A &= \widetilde{A} + A_e, \quad A_e = -h^2|\nabla\kappa|^2 - 2h\langle\nabla\tilde\varphi_\varepsilon, \nabla\kappa\rangle, \\ B &= \widetilde{B} + B_e, \quad B_e = -2ih\langle\nabla\kappa, h\nabla\,\cdot\,\rangle - ih^2\Delta\kappa. \end{aligned}$$

Consequently,

$$i[A, B] = i[\widetilde{A}, \widetilde{B}] + i[\widetilde{A}, B_e] + i[A_e, \widetilde{B}] + i[A_e, B_e].$$

Recall from [Ferreira 2009a, p. 143] that

$$i[\widetilde{A}, \widetilde{B}] = \frac{4h^2}{\varepsilon}\left(1 + \frac{h}{\varepsilon}\varphi\right)^2 + h\widetilde{B}\beta\widetilde{B} + h^2 R,$$

where $\beta=(h/\varepsilon)(1+(h/\varepsilon)\varphi)^{-2}$ and R is a first order semiclassical differential operator whose coefficients are uniformly bounded with respect to h and ε if we assume that $h/\varepsilon\le 1/2$. Consider now

$$i[\widetilde{A},\, B_e]=i[-h^2\Delta-|\nabla\tilde{\varphi}_\varepsilon|^2,\, -2ih\langle\nabla\kappa,\, h\nabla\cdot\rangle-ih^2\Delta\kappa].$$

It is clear that this equals h^2Q, where Q is a second order semiclassical differential operator whose coefficients are uniformly bounded in h and ε. The terms $i[A_e,\,\widetilde{B}]$ and $i[A_e,\, B_e]$ are better. We thus have

$$i[A,\, B]=\frac{4h^2}{\varepsilon}\left(1+\frac{h}{\varepsilon}\varphi\right)^2+h\widetilde{B}\beta\widetilde{B}+h^2Q$$

for some Q as described above. It follows that

$$(i[A,\, B]u,\, u)=\frac{4h^2}{\varepsilon}\|(1+h\varphi/\varepsilon)u\|^2+h(\widetilde{B}\beta\widetilde{B}u,\, u)+h^2(Qu,\, u).$$

We will choose h_0 so small that $|h\varphi/\varepsilon|\le 1/2$ in M for $h\le h_0$. Since $u|_{\partial M}=0$, integration by parts gives

$$|h(\widetilde{B}\beta\widetilde{B}u,\, u)|\le C_1\frac{h^2}{\varepsilon}\|\widetilde{B}u\|^2,$$

Similarly,

$$|h^2(Qu,\, u)|\le C_2h^2(\|u\|^2+\|h\nabla u\|^2).$$

Putting this information together, we get

$$(i[A,\, B]u,\, u)\ge\frac{h^2}{\varepsilon}\|u\|^2-C_1\frac{h^2}{\varepsilon}\|\widetilde{B}u\|^2-C_2h^2(\|u\|^2+\|h\nabla u\|^2).$$

Next we revisit the term $\|Au\|^2$. Let K be a positive constant whose value will be specified later. Since $u|_{\partial M}=0$, integration by parts and Young's inequality give that

$$\begin{aligned}h^2\|h\nabla u\|^2&=h^2(-h^2\Delta u,\, u)=h^2(Au,\, u)+h^2(|\nabla\varphi_\varepsilon|^2u,\, u)\\&\le\frac{1}{2K}\|Au\|^2+\frac{Kh^4}{2}\|u\|^2+C_3h^2\|u\|^2,\end{aligned}$$

or

$$\|Au\|^2\ge 2Kh^2\|h\nabla u\|^2-K^2h^4\|u\|^2-2KC_3h^2\|u\|^2.$$

Also recall that $B-\widetilde{B}=B_e=-2ih\langle\nabla\kappa,\, h\nabla\cdot\rangle-ih^2\Delta\kappa$. Thus,

$$\|(B-\widetilde{B})u\|^2\le C_4h^2(\|u\|^2+\|h\nabla u\|^2).$$

Hence

$$\|\widetilde{B}u\|^2\le 2\|Bu\|^2+2C_4h^2(\|u\|^2+\|h\nabla u\|^2)$$

and

$$\|Bu\|^2\ge\tfrac{1}{2}\|\widetilde{B}u\|^2-C_4h^2(\|u\|^2+\|h\nabla u\|^2).$$

Putting our estimates together, we obtain

$$\begin{aligned}\|P_{0,\varphi_\varepsilon}u\|^2 \geq 2Kh^2\|h\nabla u\|^2 - K^2h^4\|u\|^2 - 2KC_3h^2\|u\|^2 + \tfrac{1}{2}\|\widetilde{B}u\|^2 - C_4h^2(\|u\|^2+\|h\nabla u\|^2) \\ + \frac{h^2}{\varepsilon}\|u\|^2 - C_1\frac{h^2}{\varepsilon}\|\widetilde{B}u\|^2 - C_2h^2(\|u\|^2+\|h\nabla u\|^2) - 2h^3((\partial_\nu\varphi_\varepsilon)\partial_\nu u, \partial_\nu u)_{\partial M}.\end{aligned}$$

At this point, we choose h_0 so small that

$$C_1h_0^2/\varepsilon \leq \tfrac{1}{4}.$$

We also make the choice

$$K = \frac{1}{\alpha\varepsilon},$$

where α is to be determined. Then, for $h \leq h_0$,

$$\begin{aligned}\|P_{0,\varphi_\varepsilon}u\|^2 \geq \frac{h^2}{\varepsilon}\left(\|u\|^2 + \frac{2}{\alpha}\|h\nabla u\|^2\right) - (C_2+C_4)h^2(\|u\|^2+\|h\nabla u\|^2) \\ - \frac{h^2}{\varepsilon}\frac{h^2}{\alpha^2\varepsilon}\|u\|^2 - \frac{h^2}{\varepsilon}\frac{2C_3}{\alpha}\|u\|^2 + \tfrac{1}{4}\|\widetilde{B}u\|^2 - 2h^3((\partial_\nu\varphi_\varepsilon)\partial_\nu u, \partial_\nu u)_{\partial M}.\end{aligned}$$

Choose first $\alpha = 4C_3$. It follows that

$$\|P_{0,\varphi_\varepsilon}u\|^2 \geq \frac{h^2}{2\varepsilon}\left(1-2\varepsilon(C_2+C_4)-\frac{2h^2}{\alpha^2\varepsilon}\right)\|u\|^2 + \frac{h^2}{\varepsilon}\left(\frac{2}{\alpha}-\varepsilon(C_2+C_4)\right)\|h\nabla u\|^2 - 2h^3((\partial_\nu\varphi_\varepsilon)\partial_\nu u, \partial_\nu u)_{\partial M}.$$

Next choose ε so that

$$\varepsilon = \min\left\{\frac{1}{4(C_2+C_4)}, \frac{1}{\alpha(C_2+C_4)}\right\}.$$

Finally, choose h_0 so it satisfies the restrictions made earlier, i.e., $h_0 \leq \dfrac{\varepsilon}{2}$, $h_0 \max\limits_{x\in M}|\varphi| \leq \dfrac{\varepsilon}{2}$, and $h_0^2 \leq \dfrac{\varepsilon}{4C_1}$, and additionally

$$\frac{2h_0^2}{\alpha^2\varepsilon} \leq \frac{1}{4}.$$

With these choices, we have

$$\|P_{0,\varphi_\varepsilon}u\|^2 \geq \frac{h^2}{8\varepsilon}\|u\|^2 + \frac{h^2}{\alpha\varepsilon}\|h\nabla u\|^2 - 2h^3((\partial_\nu\varphi_\varepsilon)\partial_\nu u, \partial_\nu u)_{\partial M}.$$

Adding a potential gives

$$\|P_{0,\varphi_\varepsilon}u\|^2 \leq 2\|(P_{0,\varphi_\varepsilon}+h^2q)u\|^2 + 2h^4\|q\|^2_{L^\infty(M)}\|u\|^2.$$

Choosing an even smaller value of h_0 depending on $\|q\|_{L^\infty(M)}$ if necessary, we obtain for $0 < h \leq h_0$ that

$$\|(P_{0,\varphi_\varepsilon}+h^2q)u\|^2 \geq \frac{h^2}{C_0}(\|u\|^2+\|h\nabla u\|^2) - 2h^3((\partial_\nu\varphi_\varepsilon)\partial_\nu u, \partial_\nu u)_{\partial M}.$$

Finally, we replace u by $e^{\varphi^2/2\varepsilon+\kappa}u$, where $u \in C^\infty(\overline{\Omega})$ and $u|_{\partial\Omega} = 0$, and use the fact that

$$1/C \leq e^{\varphi^2/2\varepsilon+\kappa} \leq C, \quad |\nabla(e^{\varphi^2/2\varepsilon+\kappa})| \leq C \quad \text{on } M.$$

The required estimate follows. $\square$

We now pass from φ_ε to φ in the boundary terms of the previous result, making use of the special properties of φ_ε on ∂M. Note that the factor h^4 in the boundary term on $\{x \in \partial M : -\delta < \partial_\nu\varphi(x) < h/3\}$ below is weaker than the factor h^3 in the other boundary terms. This follows from the fact that $\partial_\nu\varphi_\varepsilon = h\partial_\nu\kappa = -h$ in the set where $\partial_\nu\varphi$ vanishes, so one only has the weak lower bound.

Proposition 4.2. *Let (M, g) be as above, let $q \in L^\infty(M)$, and let $\varphi(x) = \pm x_1$. There exist constants $C_0, h_0 > 0$ such that, whenever $0 < h \le h_0$ and $\delta > 0$, one has*

$$\frac{\delta h^3}{C_0}\|\partial_\nu u\|^2_{L^2(\{\partial_\nu\varphi\le-\delta\})} + \frac{h^4}{C_0}\|\partial_\nu u\|^2_{L^2(\{-\delta<\partial_\nu\varphi<h/3\})} + \frac{h^2}{C_0}(\|u\|^2 + \|h\nabla u\|^2) \\ \le \|e^{\varphi/h}(-h^2\Delta_g + h^2 q)(e^{-\varphi/h}u)\|^2 + h^3\|\partial_\nu u\|^2_{L^2(\{\partial_\nu\varphi\ge h/3\})}$$

for any $u \in C^\infty(M)$ with $u|_{\partial M} = 0$.

Proof. Note that

$$\partial_\nu\varphi_\varepsilon = \Big(1 + \frac{h}{\varepsilon}\varphi\Big)\partial_\nu\varphi + h\partial_\nu\kappa = \Big(1 + \frac{h}{\varepsilon}\varphi\Big)\partial_\nu\varphi - h.$$

We choose h_0 so small that whenever $h \le h_0$, one has, for $x \in M$,

$$\tfrac{1}{2} \le 1 + \frac{h}{\varepsilon}\varphi(x) \le \tfrac{3}{2}.$$

On the set where $\partial_\nu\varphi(x) \le -\delta$, we have

$$|\partial_\nu\varphi_\varepsilon| \ge \delta/2.$$

If $-\delta < \partial_\nu\varphi < h/3$, we use the estimate

$$|\partial_\nu\varphi_\varepsilon| \ge h/2.$$

Moreover, $|\partial_\nu\varphi_\varepsilon| \le C_0$ on ∂M. Since $\{\partial_\nu\varphi < h/3\} \subset \{\partial_\nu\varphi_\varepsilon < 0\}$ and $\{\partial_\nu\varphi_\varepsilon \ge 0\} \subset \{\partial_\nu\varphi \ge h/3\}$, the result follows from Proposition 4.1 after replacing C_0 by some larger constant. □

We can now obtain a solvability result from the previous Carleman estimate in a standard way by duality; see [Bukhgeim and Uhlmann 2002; Kenig et al. 2007; Nachman and Street 2010]. There is a slight technical complication, since the solution will be in L^2 but not in H^1. To remedy this, we will work with the space

$$H_{\Delta_g}(M) = \{u \in L^2(M) : \Delta_g u \in L^2(M)\}$$

with norm $\|u\|_{H_\Delta} = \|u\|_{L^2} + \|\Delta u\|_{L^2}$. As in [Bukhgeim and Uhlmann 2002], we see that $H_\Delta(M)$ is a Hilbert space having $C^\infty(M)$ as a dense subset, and there is a well defined bounded trace operator from $H_\Delta(M)$ to $H^{-1/2}(\partial M)$ and a normal derivative operator from $H_\Delta(M)$ to $H^{-3/2}(\partial M)$. We also recall that if $u \in H_\Delta(M)$ and $u|_{\partial M} \in H^{3/2}(\partial M)$, then $u \in H^2(M)$.

Proposition 4.3. *Let (M, g) be as above, let $q \in L^\infty(M)$, and let $\varphi(x) = \pm x_1$. There exist constants $C_0, \tau_0 > 0$ such that when $\tau \ge \tau_0$ and $\delta > 0$, for any $f \in L^2(M)$ and $f_- \in L^2(S_- \cup S_0)$ there exists $u \in L^2(M)$ satisfying $e^{\tau\varphi}u \in H_{\Delta_g}(M)$ and $e^{\tau\varphi}u|_{\partial M} \in L^2(\partial M)$ such that*

$$e^{-\tau\varphi}(-\Delta_g + q)(e^{\tau\varphi}u) = f \text{ in } M, \quad e^{\tau\varphi}u|_{S_-\cup S_0} = e^{\tau\varphi}f_-,$$

and

$$\|u\|_{L^2(M)} \le C_0(\tau^{-1}\|f\|_{L^2(M)} + (\delta\tau)^{-1/2}\|f_-|_{S_-}\|_{L^2(S_-)} + \|f_-|_{S_0}\|_{L^2(S_0)}).$$

Here $S_\pm$ and S_0 are the subsets of ∂M defined by

$$S_- = \{\partial_\nu\varphi \le -\delta\}, \quad S_0 = \{-\delta < \partial_\nu\varphi < 1/(3\tau)\}, \quad S_+ = \{\partial_\nu\varphi \ge 1/(3\tau)\}.$$

Proof. Write $Lv = e^{\tau\varphi}(-\Delta_g + \bar{q})(e^{-\tau\varphi}v)$ and $\tau = 1/h$, $\tau_0 = 1/h_0$. We rewrite the Carleman estimate of Proposition 4.2 as

$$(\delta\tau)^{1/2}\|\partial_\nu v\|_{L^2(S_-)} + \|\partial_\nu v\|_{L^2(S_0)} + \tau\|v\| + \|\nabla v\| \le C_0\|Lv\| + C_0\tau^{1/2}\|\partial_\nu v\|_{L^2(S_+)}.$$

This is valid for any $\delta > 0$, provided that $\tau \ge \tau_0$ and $v \in C^\infty(M)$ with $v|_{\partial M} = 0$.

Consider the following subspace of $L^2(M) \times L^2(S_+)$:

$$X = \{(Lv, \partial_\nu v|_{S_+}) : v \in C^\infty(M),\ v|_{\partial M} = 0\}.$$

Any element of X is uniquely represented as $(Lv, \partial_\nu v|_{S_+})$, where $v|_{\partial M} = 0$ by the Carleman estimate. Define a linear functional $l : X \to \mathbb{C}$ by

$$l(Lv, \partial_\nu v|_{S_+}) = (v, f)_{L^2(M)} - (\partial_\nu v, f_-)_{L^2(S_-\cup S_0)}.$$

By the Carleman estimate, we have

$$\begin{aligned}|l(Lv, \partial_\nu v|_{S_+})| &\le \|v\|\|f\| + \|\partial_\nu v\|_{L^2(S_-)}\|f_-\|_{L^2(S_-)} + \|\partial_\nu v\|_{L^2(S_0)}\|f_-\|_{L^2(S_0)} \\ &\le C_0(\tau^{-1}\|f\| + (\delta\tau)^{-1/2}\|f_-\|_{L^2(S_-)} + \|f_-\|_{L^2(S_0)}) \times (\|Lv\| + \tau^{1/2}\|\partial_\nu v\|_{L^2(S_+)}).\end{aligned}$$

The Hahn–Banach theorem implies that l extends to a continuous linear functional

$$\bar{l} : L^2(M) \times \tau^{-1/2}L^2(S_+) \to \mathbb{C}$$

such that

$$\|\bar{l}\| \le C_0(\tau^{-1}\|f\| + (\delta\tau)^{-1/2}\|f_-\|_{L^2(S_-)} + \|f_-\|_{L^2(S_0)}).$$

By the Riesz representation theorem, there exist functions $u \in L^2(M)$ and $u_+ \in L^2(S_+)$ satisfying $\bar{l}(w, w_+) = (w, u)_{L^2(M)} + (w_+, u_+)_{L^2(S_+)}$. Moreover,

$$\|u\|_{L^2(M)} + \tau^{-1/2}\|u_+\|_{L^2(S_+)} \le C_0(\tau^{-1}\|f\| + (\delta\tau)^{-1/2}\|f_-\|_{L^2(S_-)} + \|f_-\|_{L^2(S_0)}).$$

If $v \in C^\infty(M)$ and $v|_{\partial M} = 0$, we have

$$(Lv, u)_{L^2(M)} + (\partial_\nu v, u_+)_{L^2(S_+)} = (v, f)_{L^2(M)} - (\partial_\nu v, f_-)_{L^2(S_-\cup S_0)}.$$

Choosing v compactly supported in M^{int}, it follows that $L^*u = f$, or

$$e^{-\tau\varphi}(-\Delta_g + q)(e^{\tau\varphi}u) = f \quad \text{in } M.$$

Furthermore, $e^{\tau\varphi}u \in H_\Delta(M)$.

If $w, v \in C^\infty(M)$ with $v|_{\partial M} = 0$, an integration by parts gives

$$(Lv, w) = -(e^{-\tau\varphi}\partial_\nu v, e^{\tau\varphi}w)_{L^2(\partial M)} + (v, L^* w).$$

Given our solution u, we choose $u_j \in C^\infty(M)$ so that $e^{\tau\varphi}u_j \to e^{\tau\varphi}u$ in $H_\Delta(M)$. Applying the above formula with $w = u_j$ and taking the limit, we see that

$$(Lv, u) = -(e^{-\tau\varphi}\partial_\nu v, e^{\tau\varphi}u)_{L^2(\partial M)} + (v, L^* u)$$

for $v \in C^\infty(M)$ with $v|_{\partial M} = 0$. Combining this with (4-1), using that $L^* u = f$ gives

$$(\partial_\nu v, f_-)_{L^2(S_- \cup S_0)} + (\partial_\nu v, u_+)_{L^2(S_+)} = (e^{-\tau\varphi}\partial_\nu v, e^{\tau\varphi}u)_{L^2(\partial M)}.$$

Since $\partial_\nu v$ can be chosen arbitrarily, it follows that $e^{\tau\varphi}u|_{S_-\cup S_0} = e^{\tau\varphi}f_-$ and $e^{\tau\varphi}u|_{S_+} = e^{\tau\varphi}u_+$. We also see that $e^{\tau\varphi}u|_{\partial M} \in L^2(\partial M)$. □

5. Reflection approach

In the previous section, we employed Carleman estimates and duality to obtain a solvability result (Proposition 4.3) that will be used to produce correction terms in complex geometrical optics solutions with prescribed behavior on parts of the boundary. In this section we give an alternative approach to the construction of correction terms vanishing on parts of the boundary. The method is based on a reflection argument. We extend the method of [Isakov 2007], which dealt with inaccessible parts that are part of a hyperplane, to the case of inaccessible parts that are part of the graph of a function independent of one of the variables. The results are less general than the ones in Section 4, and, for simplicity, will only be stated for domains in $\mathbb{R}^3$ with Euclidean metric, but on the other hand, the method is constructive and is based on direct Fourier arguments in the spirit of [Kenig et al. 2011a; Kenig et al. 2011b].

Let $\Omega \subset \mathbb{R}^3$ be a bounded domain with smooth boundary, and assume that

$$\Omega \subset \mathbb{R} \times \{(x_2, x_3) : x_3 > \eta(x_2)\},$$

where $\eta : \mathbb{R} \to \mathbb{R}$ is a smooth function. Also assume that Γ_0 is a closed subset of $\partial\Omega$ such that

$$\Gamma_0 \subset \mathbb{R} \times \{(x_2, x_3) : x_3 = \eta(x_2)\}.$$

We will show that if one has access to suitable amplitudes of complex geometrical optics solutions that vanish on Γ_0, it is possible to produce correction terms that also vanish on Γ_0.

Proposition 5.1. *Let Ω and Γ_0 be as above, and let $q \in L^\infty(\Omega)$. There are $C_0, \tau_0 > 0$ such that, for any τ with $|\tau| \geq \tau_0$ and for any $m \in H^2(\Omega)$ with $m|_{\Gamma_0} = 0$, the equation $(-\Delta + q)u = 0$ in Ω has a solution $u \in H^2(\Omega)$ of the form*

$$u = e^{-\tau x_1}(m + r)$$

such that $r|_{\Gamma_0} = 0$ and

$$\|r\|_{L^2(\Omega)} \leq \frac{C_0}{|\tau|}\|e^{\tau x_1}(-\Delta + q)(e^{-\tau x_1}m)\|_{L^2(\Omega)}.$$

The proof involves a reflection argument that reduces the construction of the correction term to the problem of solving a conjugated equation with anisotropic metric,

$$e^{\tau x_1}(-\Delta_{\hat{g}}+\hat{q})(e^{-\tau x_1}\hat{r})=\hat{f} \quad \text{in } \mathbb{R}\times\widehat{\Omega}_0,$$

where $\widehat{\Omega}_0\subset\mathbb{R}^2$ is a bounded open set and $\hat{g}$ is a metric of the form

$$\hat{g}(y_1,y')=\begin{pmatrix}1 & 0\\ 0 & \hat{g}_0(y')\end{pmatrix},$$

and where g_0 is smooth for $y_3\neq 0$ but only Lipschitz continuous across $\{y_3=0\}$. In three and higher dimensions, it is not known how to handle equations of this type with general Lipschitz coefficients in the second order part (the case of C^1 coefficients, and also Lipschitz coefficients with a smallness condition, is considered in [Haberman and Tataru 2013]). However, in our case, the singularity of $\hat{g}$ only appears in the lower right block $\hat{g}_0$, and this turns out not to be a problem.

The following is an analogue of [Kenig et al. 2011a, Proposition 4.1], the main difference being that the transversal metric is only Lipschitz. (With correct definitions, one could easily deal with L^∞ transversal metrics as well, but then the solution would only be in $H^1_{-\delta}(T)$.) Here we write (x_1,x') for coordinates in $T=\mathbb{R}\times M_0$, and for $\delta\in\mathbb{R}$ we consider the spaces

$$\|f\|_{L^2_\delta(T)}=\|\langle x_1\rangle^\delta f\|_{L^2(T)},\quad \|f\|_{H^1_\delta(T)}=\|f\|_{L^2_\delta(T)}+\|df\|_{L^2_\delta(T)}$$

with $\langle t\rangle=(1+t^2)^{1/2}$, and similarly for $H^2_\delta(T)$. We also write $\mathrm{Spec}(-\Delta_{g_0})$ for the set of Dirichlet eigenvalues of the Laplace–Beltrami operator $-\Delta_{g_0}$ in (M_0,g_0).

Proposition 5.2. *Let $T=\mathbb{R}\times M_0$ with metric $g=e\oplus g_0$, where (M_0,g_0) is a compact oriented manifold with smooth boundary and g_0 is a Lipschitz continuous Riemannian metric on M_0. Given any $q\in L^\infty_{comp}(T)$ and any $\delta>1/2$, there are constants $C_0,\tau_0>0$ such that whenever*

$$|\tau|\geq\tau_0 \quad \textit{and} \quad \tau^2\notin\mathrm{Spec}(-\Delta_{g_0}),$$

the equation

$$e^{\tau x_1}(-\Delta_g+q)(e^{-\tau x_1}r)=f \quad \textit{in } T$$

has a unique solution $r\in H^1_{-\delta}(T)$ with $r|_{\partial T}=0$ for any $f\in L^2_\delta(T)$. Moreover, $r\in H^2_{-\delta}(T)$, and one has the bounds

$$\|r\|_{L^2_{-\delta}(T)}\leq\frac{C_0}{|\tau|}\|f\|_{L^2_\delta(T)},\quad \|r\|_{H^1_{-\delta}(T)}\leq C_0\|f\|_{L^2_\delta(T)}.$$

Proof. The proof is almost exactly the same as the proof of [Kenig et al. 2011a, Proposition 4.1], and we only give the main idea. Since $\Delta_g=\partial_{x_1}^2+\Delta_{g_0}$, the equation that we need to solve is

$$(-\partial_{x_1}^2+2\tau\partial_{x_1}-\Delta_{g_0}-\tau^2+q)r=f \quad \text{in } T.$$

It is enough to consider $q=0$. The standard argument based on weak solutions shows that even when g_0 has very little regularity, there is an orthonormal basis of $L^2(M_0)$ consisting of Dirichlet eigenfunctions of $-\Delta_{g_0}$,

$$-\Delta_{g_0}\varphi_l = \lambda_l \varphi_l \text{ in } M_0, \quad \varphi_l \in H_0^1(M_0),$$

where $0 < \lambda_1 \le \lambda_2 \le \lambda_3 \le \cdots \to \infty$ are the Dirichlet eigenvalues of $-\Delta_{g_0}$ in M_0.

Considering the partial Fourier expansions

$$r(x_1, x') = \sum_{l=1}^{\infty} \tilde{r}(x_1, l)\varphi_l(x'), \quad f(x_1, x') = \sum_{l=1}^{\infty} \tilde{f}(x_1, l)\varphi_l(x'),$$

it is enough to solve

$$(-\partial_{x_1}^2 + 2\tau\partial_{x_1} + \lambda_l - \tau^2)\tilde{r}(\,\cdot\,, l) = \tilde{f}(\,\cdot\,, l) \quad \text{in } \mathbb{R} \text{ for all } l.$$

The condition $\tau^2 \notin \mathrm{Spec}(-\Delta_{g_0})$ allows us to solve these ordinary differential equations by the Fourier transform as in [Kenig et al. 2011a, Section 4], and the estimates given there imply that one obtains a unique solution $r \in H^1_{-\delta}(T)$ with $r|_{\partial T} = 0$ satisfying the required bounds. Elliptic H^2 regularity also works with Lipschitz g_0, and the argument in [Kenig et al. 2011a, Section 4] gives that $r \in H^2_{-\delta}(T)$. $\square$

Proof of Proposition 5.1. We begin by flattening Γ_0 via the map

$$\Phi : \mathbb{R}^3 \to \mathbb{R}^3, \quad (x_1, x_2, x_3) \mapsto (x_1, x_2, x_3 - \eta(x_2)).$$

Let $\widetilde{\Omega} = \Phi(\Omega)$, write y for coordinates in $\widetilde{\Omega}$, and let R be the reflection

$$R(y_1, y_2, y_3) = (y_1, y_2, -y_3).$$

Note that $\widetilde{\Omega} \subset \{y_3 > 0\}$. Consider the reflected domain $\widetilde{\Omega}^* = R(\widetilde{\Omega})$, so $\widetilde{\Omega}^* \subset \{y_3 < 0\}$, and let U the double domain $\widetilde{\Omega} \cup \Phi(\Gamma_0)^{\mathrm{int}} \cup \widetilde{\Omega}^*$.

Let $\Psi = \Phi^{-1}$, let $\tilde{g} = \Psi^* e$ be the metric in $\widetilde{\Omega}$ that is the pullback of the Euclidean metric in Ω, let $\tilde{q} = \Psi^* q$, and let $\tilde{m} = \Psi^* m$. In the double domain U, we use even reflection to define the quantities

$$\hat{g} = \begin{cases} \tilde{g} & \text{if } y_3 > 0, \\ R^*\tilde{g} & \text{if } y_3 < 0, \end{cases} \qquad \hat{q} = \begin{cases} \tilde{q} & \text{if } y_3 > 0, \\ R^*\tilde{q} & \text{if } y_3 < 0, \end{cases}$$

and odd reflection to define the amplitude

$$\hat{m} = \begin{cases} \frac{1}{2}\tilde{m} & \text{if } y_3 > 0, \\ -\frac{1}{2}R^*\tilde{m} & \text{if } y_3 < 0. \end{cases}$$

Since the flattening map Φ leaves x_1 intact, we have

$$\hat{g}(y_1, y') = \begin{pmatrix} 1 & 0 \\ 0 & \hat{g}_0(y') \end{pmatrix},$$

where $\hat{g}_0$ is a Lipschitz continuous metric only depending on y_2 and y_3. (In fact, $\tilde{g}$ and $\tilde{g}_0$ are well defined in $\{y_3 > 0\}$ by the flattening map Φ and the Euclidean metric in $\{x_3 > \eta(x_2)\}$.) Also, $\hat{q} \in L^\infty(U)$, and $\hat{m} \in H^2(U)$ by the boundary condition $m|_{\Gamma_0} = 0$ and by the properties of odd reflection.

We wish to find $\hat{r} \in H^1(U)$ satisfying

$$e^{\tau x_1}(-\Delta_{\hat{g}} + \hat{q})(e^{-\tau x_1}\hat{r}) = \hat{f},$$

where $\hat{f} = -e^{\tau x_1}(-\Delta_{\hat{g}} + \hat{q})(e^{-\tau x_1}\hat{m})$. Now

$$\begin{aligned}\|\hat{f}\|_{L^2(U)} &= \|\hat{f}\|_{L^2(\widetilde{\Omega})} + \|\hat{f}\|_{L^2(\widetilde{\Omega}^*)}\\ &= \|\Psi^*(e^{\tau x_1}(-\Delta + q)(e^{-\tau x_1}m)\|_{L^2(\widetilde{\Omega})} + \|R^*\Psi^*(e^{\tau x_1}(-\Delta + q)(e^{-\tau x_1}m)\|_{L^2(\widetilde{\Omega}^*)}\\ &\le C\|e^{\tau x_1}(-\Delta + q)(e^{-\tau x_1}m)\|_{L^2(\Omega)}.\end{aligned}$$

Choose a bounded open set $\widehat{\Omega}_0 \subset \mathbb{R}^2$ such that

$$U \Subset \mathbb{R} \times \widehat{\Omega}_0,$$

and let $\hat{g}_0$ be the metric in $\widehat{\Omega}_0$ that is the even extension of $\tilde{g}_0$ from $\{y_3 > 0\}$ to $\widehat{\Omega}_0$. Then $\hat{g}_0$ is smooth for $y_3 \neq 0$ and Lipschitz continuous across $\{y_3 = 0\}$. Extending $\hat{g}$ to $\mathbb{R} \times \widehat{\Omega}_0$ using the block structure and extending $\hat{q}$ and $\hat{f}$ by zero to $\mathbb{R} \times \widehat{\Omega}_0$, it is enough to find a solution $\hat{r} \in H^2_{loc}(\mathbb{R} \times \widehat{\Omega}_0)$ of the equation

$$e^{\tau x_1}(-\Delta_{\hat{g}} + \hat{q})(e^{-\tau x_1}\hat{r}) = \hat{f} \quad \text{in } \mathbb{R} \times \widehat{\Omega}_0. \tag{5-1}$$

Such a solution may be found by Proposition 5.2, and denoting by $\hat{r}$ its restriction to U, we have

$$\|\hat{r}\|_{L^2(U)} \le \frac{C}{|\tau|}\|\hat{f}\|_{L^2(U)}.$$

Now define

$$\hat{u} = e^{-\tau x_1}(\hat{m} + \hat{r}) \quad \text{in } U$$

and

$$\tilde{u} = \hat{u} - R^*\hat{u} \quad \text{in } \widetilde{\Omega}.$$

Then $(-\Delta_{\hat{g}} + \hat{q})\hat{u} = 0$ in U, and $(-\Delta_{\tilde{g}} + \tilde{q})\tilde{u} = 0$ in $\widetilde{\Omega}$ by the definition of $\hat{g}$ and $\hat{q}$ and using that $\hat{u} \in H^2(U)$. We also have

$$\tilde{u} = e^{-\tau x_1}(\hat{m} - R^*\hat{m} + \hat{r} - R^*\hat{r}) \quad \text{in } \widetilde{\Omega}.$$

But here $\hat{m} - R^*\hat{m}|_{\widetilde{\Omega}} = \tilde{m}$ by the definition of $\hat{m}$. Consequently, if we define $u = \Phi^*\tilde{u}$, then $(-\Delta + q)u = 0$ in Ω and

$$u = e^{-\tau x_1}(m + r) \quad \text{in } \Omega,$$

where $r = \Phi^*(\hat{r} - R^*\hat{r})$ satisfies

$$\|r\|_{L^2(\Omega)} \le C\|\hat{r}\|_{L^2(U)} \le \frac{C}{|\tau|}\|\hat{f}\|_{L^2(U)} \le \frac{C}{|\tau|}\|e^{\tau x_1}(-\Delta + q)(e^{-\tau x_1}m)\|_{L^2(\Omega)}. \qquad \square$$

Note how the odd reflection of the amplitude m in the proof ensured that the solution obtained by reflection is not the zero solution. We also remark that under certain conditions, the arguments in Sections 6 and 7 allow to construct amplitudes m vanishing on a part Γ_0 as above.

6. Local uniqueness on simple manifolds

In this section we prove Theorems 2.1–2.3. In these results the transversal manifold is assumed to be simple and we only use nonreflected geodesics. This case already illustrates the main features of the approach, and we can use a quasimode construction that is much easier than the Gaussian beam one used for nonsimple transversal manifolds and reflected geodesics.

The first observation is the usual integral identity.

Proposition 6.1. *If $\Gamma_D, \Gamma_N \subset \partial M$ are open and if $C_{g,q_1}^{\Gamma_D,\Gamma_N} = C_{g,q_2}^{\Gamma_D,\Gamma_N}$, then*

$$\int_M (q_1 - q_2) u_1 u_2 \, dV_g = 0$$

for any $u_j \in H_{\Delta_g}(M)$ satisfying $(-\Delta_g + q_j) u_j = 0$ in M and

$$\operatorname{supp}(u_1|_{\partial M}) \subset \Gamma_D, \quad \operatorname{supp}(u_2|_{\partial M}) \subset \Gamma_N.$$

Proof. Let u_j be as stated. Since $C_{g,q_1}^{\Gamma_D,\Gamma_N} = C_{g,q_2}^{\Gamma_D,\Gamma_N}$, there is a function $\tilde{u}_2 \in H_\Delta(M)$ with $(-\Delta + q_2)\tilde{u}_2 = 0$ in M, $\operatorname{supp}(\tilde{u}_2|_{\partial M}) \subset \Gamma_D$, and

$$(u_1|_{\Gamma_D}, \partial_\nu u_1|_{\Gamma_N}) = (\tilde{u}_2|_{\Gamma_D}, \partial_\nu \tilde{u}_2|_{\Gamma_N}).$$

Using that u_1, u_2, and $\tilde{u}_2$ are solutions, we have

$$\begin{aligned} \int_M (q_1 - q_2) u_1 u_2 \, dV &= \int_M [(\Delta u_1) u_2 - u_1 (\Delta u_2)] \, dV \\ &= \int_M [(\Delta(u_1 - \tilde{u}_2)) u_2 - (u_1 - \tilde{u}_2)(\Delta u_2)] \, dV. \end{aligned}$$

Now $u_1 - \tilde{u}_2|_{\partial M} = 0$, so in fact $u_1 - \tilde{u}_2 \in H^2(M)$ by the properties of the space $H_\Delta(M)$. Recall also that $C^\infty(M)$ is dense in $H_\Delta(M)$ and that $u_2|_{\partial M} \in H^{-1/2}(\partial M)$ and $\partial_\nu u_2|_{\partial M} \in H^{-3/2}(\partial M)$. These facts make it possible to integrate by parts, and we obtain that

$$\int_M (q_1 - q_2) u_1 u_2 \, dV = \int_{\partial M} [(\partial_\nu (u_1 - \tilde{u}_2)) u_2 - (u_1 - \tilde{u}_2)(\partial_\nu u_2)] \, dS$$

in the weak sense. The last expression vanishes since $\partial_\nu (u_1 - \tilde{u}_2)|_{\Gamma_N} = 0$ and $\operatorname{supp}(u_2|_{\partial M}) \subset \Gamma_N$. □

The next result will be used to pass from the metric $g = c(e \oplus g_0)$ to the slightly simpler metric $\tilde{g} = e \oplus g_0$.

Lemma 6.2. *Let c be a smooth positive function in M. Then $u \in H_{\Delta_g}(M)$ satisfies $(-\Delta_g + q)u = 0$ in M if and only if $\tilde{u} \in H_{\Delta_{\tilde{g}}}(M)$ satisfies $(-\Delta_{\tilde{g}} + \tilde{q})\tilde{u} = 0$ in M, where*

$$\tilde{g} = c^{-1} g, \quad \tilde{u} = c^{(n-2)/4} u, \quad \tilde{q} = c(q - c^{(n-2)/4} \Delta_g (c^{-(n-2)/4})).$$

Proof. This follows from the identity for $v \in C^\infty(M)$,

$$c^{(n+2)/4}(-\Delta_g + q)(c^{-(n-2)/4}v) = (-\Delta_{c^{-1}g} + c(q - c^{(n-2)/4}\Delta_g(c^{-(n-2)/4})))v,$$

upon approximating u or $\tilde{u}$ by smooth functions. □

Proof of Theorem 2.1. Let $\tilde{g} = e \oplus g_0$ and $\tilde{q}_j = c(q_j - c^{(n-2)/4}\Delta_g(c^{-(n-2)/4}))$. Let λ be a fixed real number, and consider the complex frequency

$$s = \tau + i\lambda,$$

where $\tau > 0$ will be large. We look for solutions

$$\tilde{u}_1 = e^{-sx_1}(v_s(x') + r_1),$$
$$\tilde{u}_2 = e^{sx_1}(v_s(x') + r_2)$$

of the equations $(-\Delta_{\tilde{g}} + \tilde{q}_1)\tilde{u}_1 = 0$, $(-\Delta_{\tilde{g}} + \overline{\tilde{q}}_2)\tilde{u}_2 = 0$ in M. Here v_s will be a quasimode for the Laplacian in (M_0, g_0) that concentrates near the given geodesic γ. Next we will construct a suitable solution $\tilde{u}_1$, and the case of $\tilde{u}_2$ will be analogous.

Since $\Delta_{\tilde{g}} = \partial_1^2 + \Delta_{g_0}$, the function $\tilde{u}_1$ is a solution if and only if

$$e^{sx_1}(-\Delta_{\tilde{g}} + \tilde{q}_1)(e^{-sx_1}r_1) = -(-\Delta_{g_0} + \tilde{q}_1 - s^2)v_s(x') \quad \text{in } M. \tag{6-1}$$

We want to choose $v_s \in C^\infty(M_0)$ to satisfy

$$\|v_s\|_{L^2(M_0)} = O(1), \quad \|(-\Delta_{g_0} - s^2)v_s\|_{L^2(M_0)} = O(1) \tag{6-2}$$

as $\tau \to \infty$. Looking for v_s in the form

$$v_s = e^{is\psi}a,$$

where $\psi, a \in C^\infty(M_0)$, a direct computation shows that

$$(-\Delta_{g_0} - s^2)v_s = e^{is\psi}\big(s^2[|d\psi|_{g_0}^2 - 1]a - is[2\langle d\psi, d\,\cdot\rangle_{g_0} + \Delta_{g_0}\psi]a - \Delta_{g_0}a\big).$$

Since (M_0, g_0) is simple, it is easy to find ψ and a so that the expressions in brackets will vanish and that the resulting quasimode v_s will concentrate near the geodesic γ. To do this, let $(\widehat{M}_0, g_0)$ be a simple manifold that is slightly larger than (M_0, g_0), extend γ as a geodesic in $\widehat{M}_0$, and choose $\varepsilon > 0$ such that $\gamma|_{(-2\varepsilon,0)\cup(L,L+2\varepsilon)}$ stays in $\widehat{M}_0 \setminus M_0$ (this is possible since γ is nontangential). Let $\omega = \gamma(-\varepsilon) \in \widehat{M}_0 \setminus M_0$, and let (r, θ) be polar normal coordinates in (M_0, g_0) with center ω. Then γ corresponds to the curve $r \mapsto (r, \theta_0)$ for some fixed $\theta_0 \in S^{n-2}$. We will choose

$$\psi(r, \theta) = r,$$
$$a(r, \theta) = |g_0(r, \theta)|^{-1/4}b(\theta),$$

where $|g_0|$ is the determinant of g_0, and b is a fixed function in $C^\infty(S^{n-2})$ that is supported so close to θ_0 such that $v_s|_{\partial M_0 \setminus E} = 0$. With these choices, we have, as in [Ferreira 2009a],

$$(-\Delta_{g_0} - s^2)v_s = -e^{is\psi}\Delta_{g_0}a.$$

Thus v_s satisfies the estimates (6-2), and also the estimate

$$\|v_s\|_{L^\infty(M_0)} = O(1).$$

We now go back to (6-1) and look for a solution in the form $r_1 = e^{i\lambda x_1} r_1'$ where r_1' satisfies

$$e^{\tau x_1}(-\Delta_{\tilde g} + \tilde q_1)(e^{-\tau x_1} r_1') = f \quad \text{in } M \tag{6-3}$$

with

$$f = -e^{-i\lambda x_1}(-\Delta_{g_0} + \tilde q_1 - s^2) v_s(x').$$

We also want to arrange that $\operatorname{supp}(\tilde u_1|_{\partial M}) \subset \Gamma_D$, where $\Gamma_D \supset \partial M_- \cup \Gamma_a$. For this purpose, let $\delta > 0$ be a small number to be fixed later, let $S_\pm$ and S_0 be the sets in Proposition 4.3 with Carleman weight $\varphi(x) = -x_1$, define

$$V^\delta = \{x \in S_- \cup S_0 : \operatorname{dist}_{\partial M}(x, \Gamma_i) < \delta \text{ or } x \in \partial M_+\},$$
$$\Gamma_a^\delta = (S_- \cup S_0) \setminus V_\delta,$$

and impose the boundary condition

$$e^{\tau\varphi} r_1'|_{S_- \cup S_0} = e^{\tau\varphi} f_-, \tag{6-4}$$

where

$$f_- = \begin{cases} -e^{-i\lambda x_1} v_s(x') & \text{on } V^\delta, \\ 0 & \text{on } \Gamma_a^\delta. \end{cases}$$

Note that $\partial M_+ \cup \partial M_{\mathrm{tan}}$ (these sets refer to the weight x_1) is in the interior of $S_- \cup S_0$ in ∂M.

We have seen that $\|f\|_{L^2(M)} = O(1)$ as $\tau \to \infty$. We also have

$$f_-|_{\partial M_{\mathrm{tan}}} = 0,$$

since $f_-|_{\Gamma_a^\delta \cap \partial M_{\mathrm{tan}}} = 0$ by definition and $f_-|_{\partial M_{\mathrm{tan}} \cap V_\delta} = 0$ for sufficiently small $\delta > 0$ by the construction of v_s and using that $\Gamma_i \subset \mathbb{R} \times (\partial M_0 \setminus E)$. Since $\|f_-\|_{L^\infty(S_- \cup S_0)} \lesssim 1$, we have

$$\|f_-\|_{L^2(S_-)} \lesssim \sigma(\{\partial_\nu x_1 \geq \delta\})$$

and

$$\|f_-\|_{L^2(S_0)} \lesssim \sigma(\{-1/(3\tau) < \partial_\nu x_1 < 0\} \cup \{0 < \partial_\nu x_1 < \delta\}),$$

where σ is the surface measure on ∂M. It follows from Proposition 4.3 that (6-3) has a solution r_1' satisfying the boundary condition (6-4), and having the estimate

$$\|r_1'\|_{L^2(M)} \lesssim \tau^{-1} + (\delta\tau)^{-1/2}\sigma(\{\partial_\nu x_1 \geq \delta\}) + \sigma(\{-1/(3\tau) < \partial_\nu x_1 < 0\}) + \sigma(\{0 < \partial_\nu x_1 < \delta\}).$$

The implied constants in the previous inequality are independent of τ and δ. By the basic properties of measures, for some constant $C_0 > 0$, we have

$$\|r_1'\|_{L^2(M)} \leq C_0[\tau^{-1} + (\delta\tau)^{-1/2} + o_{\tau\to\infty}(1) + o_{\delta\to 0}(1)].$$

Given $\varepsilon > 0$, we first choose δ so that $C_0 o_{\delta\to 0}(1) \leq \varepsilon/2$. After this, we choose $\tau > 0$ so large that $C_0(\tau^{-1} + (\delta\tau)^{-1/2} + o_{\tau\to\infty}(1)) \leq \varepsilon/2$. This shows that

$$\lim_{\tau\to\infty} \|r_1'(\,\cdot\,;\tau)\|_{L^2(M)} = 0.$$

Choosing r_1' as described above and choosing $r_1 = e^{i\lambda x_1} r_1'$, we have produced a solution $\tilde{u}_1 \in H_{\Delta_{\tilde{g}}}(M)$ of the equation $(-\Delta_{\tilde{g}} + \tilde{q}_1)\tilde{u}_1 = 0$ in M, having the form

$$\tilde{u}_1 = e^{-sx_1}(v_s(x') + r_1)$$

and satisfying

$$\operatorname{supp}(\tilde{u}_1|_{\partial M}) \subset \Gamma_D$$

and $\|r_1\|_{L^2(M)} = o(1)$ as $\tau \to \infty$. Repeating this construction for the Carleman weight $\varphi(x) = x_1$, we obtain a solution $\tilde{u}_2 \in H_{\Delta_{\tilde{g}}}(M)$ of the equation $(-\Delta_{\tilde{g}} + \overline{\tilde{q}}_2)\tilde{u}_2 = 0$ in M, having the form

$$\tilde{u}_2 = e^{sx_1}(v_s(x') + r_2)$$

and satisfying

$$\operatorname{supp}(\tilde{u}_2|_{\partial M}) \subset \Gamma_N$$

and $\|r_2\|_{L^2(M)} = o(1)$ as $\tau \to \infty$.

Writing $u_j = c^{-(n-2)/4}\tilde{u}_j$, Lemma 6.2 shows that $u_j \in H_{\Delta_g}(M)$ are solutions of $(-\Delta_g + q_1)u_1 = 0$ and $(-\Delta_g + \bar{q}_2)u_2 = 0$ in M. Then Proposition 6.1 implies that

$$\int_M (q_1 - q_2) u_1 \bar{u}_2 \, dV_g = 0.$$

We extend $q_1 - q_2$ by zero to $\mathbb{R} \times M_0$. Inserting the expressions for u_j, and using that $dV_g = c^{n/2}\, dx_1\, dV_{g_0}(x')$, we obtain

$$\int_{M_0}\int_{-\infty}^{\infty} (q_1 - q_2) c e^{-2i\lambda x_1}(|v_s(x')|^2 + v_s\bar{r}_2 + \bar{v}_s r_1 + r_1\bar{r}_2)\, dx_1\, dV_{g_0}(x') = 0.$$

Since $\|r_j\|_{L^2(M)} = o(1)$ as $\tau \to \infty$ and since $dV_{g_0} = |g_0|^{1/2}\, dr\, d\theta$ in the (r, θ) coordinates, it follows that

$$\int_{S^{n-2}}\int_0^{\infty} e^{-2\lambda r}(c(q_1 - q_2))\hat{\ }(2\lambda, r, \theta)|b(\theta)|^2\, dr\, d\theta = 0.$$

Varying b in $C^\infty(S^{n-2})$ so that the support of b is very close to θ_0, this implies that

$$\int_0^{\infty} e^{-2\lambda r}(c(q_1 - q_2))\hat{\ }(2\lambda, r, \theta_0)\, dr = 0.$$

Since γ was the curve $r \mapsto (r, \theta)$, this shows the result. $\square$

Proof of Theorem 2.2. Suppose that the local ray transform is injective on O and $O \cap \partial M \subset E$. By Theorem 2.1, we know that

$$\int_0^L e^{-2\lambda t} (c(q_1 - q_2))\hat{\ }(2\lambda, \gamma(t))\, dt = 0 \tag{6-5}$$

for any nontangential geodesic γ in O. Setting $\lambda = 0$ and using local injectivity of the ray transform, we obtain that

$$(c(q_1 - q_2))\hat{\ }(0, \cdot\,) = 0 \quad \text{in } O.$$

Going back to (6-5) and differentiating this identity with respect to λ, and then setting $\lambda = 0$ and using the vanishing of $(c(q_1 - q_2))\hat{\ }(0, \cdot\,)$ on O, it follows that

$$\int_0^L \frac{\partial}{\partial\lambda}[(c(q_1 - q_2))\hat{\ }](0, \gamma(t))\, dt = 0 \text{ in } O$$

for any nontangential geodesic in O. Local uniqueness for the ray transform again implies that

$$\frac{\partial}{\partial\lambda}[(c(q_1 - q_2))\hat{\ }](0, \cdot\,) = 0 \quad \text{in } O.$$

Iterating this argument by taking higher order derivatives of (6-5) shows that

$$\left(\frac{\partial}{\partial\lambda}\right)^k [(c(q_1 - q_2))\hat{\ }](0, \cdot\,) = 0 \quad \text{in } O$$

for any k. Since $c(q_1 - q_2)$ is compactly supported in x_1, its Fourier transform is analytic and we have

$$(c(q_1 - q_2))\hat{\ }(\lambda, \cdot\,) = 0 \quad \text{in } O \text{ for all } \lambda \in \mathbb{R}.$$

Inverting the Fourier transform and using that c is positive, we obtain that $q_1 = q_2$ in $M \cap (\mathbb{R} \times O)$. □

Proof of Theorem 2.3. Since (M, g) is admissible, we may assume that

$$(M, g) \subset (\mathbb{R} \times M_0, g), \qquad g = c(e \oplus g_0),$$

where (M_0, g_0) is simple. The argument is very similar to the proof of Theorem 2.1, and we only indicate the required changes. Up to the formula (6-3), the only change is that there is no restriction on $b \in C^\infty(S^{n-2})$ (we do not require v_s to vanish on any part of the boundary). The function r_1' is obtained as a solution of (6-3), but this time we want $\operatorname{supp}(\tilde{u}_1|_{\partial M}) \subset \partial M_-$. Fix $\delta > 0$. The boundary condition for $\tilde{u}_1$ is (6-4), where f_- is chosen to be

$$f_- = -e^{-i\lambda x_1} v_s(x') \quad \text{on } S_- \cup S_0.$$

We use Proposition 4.3 to solve for r_1'. We have $\|f\|_{L^2(M)} = O(1)$, and the bound $\|f_-\|_{L^\infty} \lesssim 1$ implies

$$\|f_-\|_{L^2(S_-)} \lesssim \sigma(\{\partial_\nu x_1 \geq \delta\})$$

and

$$\|f_-\|_{L^2(S_0)} \lesssim \sigma(\{-1/(3\tau) < \partial_\nu x_1 < 0\}) + \sigma(\partial M_{\tan}) + \sigma(\{0 < \partial_\nu x_1 < \delta\}).$$

Now we use that

$$\sigma(\partial M_{\tan}) = 0.$$

This shows that we obtain the same estimate for r_1' as before:

$$\|r_1'\|_{L^2(M)} \leq C_0[\tau^{-1} + (\delta\tau)^{-1/2} + o_{\tau\to\infty}(1) + o_{\delta\to 0}(1)].$$

We can now continue as in the proof of Theorem 2.1 to conclude that

$$\int_0^L e^{-2\lambda t}(c(q_1 - q_2))\hat{}\,(2\lambda, \gamma(t))\,dt = 0$$

for any $\lambda \in \mathbb{R}$ and for any nontangential geodesic in (M_0, g_0). The geodesic ray transform (with zero attenuation) is injective in (M_0, g_0) [Sharafutdinov 1994]. Following the proof of Theorem 2.2, but now using all the nontangential geodesics in (M_0, g_0), shows that $q_1 = q_2$ in M. □

7. Quasimodes concentrating near broken rays

In this section, to simplify notation, we write (M, g) instead of (M_0, g_0) and we assume that (M, g) is a compact oriented Riemannian manifold having smooth boundary and $\dim(M) = m \geq 2$. Suppose that E is a nonempty open subset of ∂M, and let $R = \partial M \setminus E$. We think of E as the observation set where geodesics can enter and exit, and R is the reflecting set. In the Calderón problem with partial data, we are led to consider attenuated broken ray transforms, where one integrates a function on M over broken geodesic rays that enter M at some point of E, reflect nontangentially at points of R, and then exit M at some point of E. The reflections will obey the law of geometric optics, so that a geodesic hitting the boundary in direction v will be continued by the geodesic in the reflected direction $\hat{v} = v - 2\langle v, \nu\rangle\nu$.

Given a slightly complex frequency $s = \tau + i\lambda$, we will construct corresponding quasimodes, or approximate eigenfunctions, that concentrate near a fixed nontangential broken ray.

Proposition 7.1. *Let $\gamma : [0, L] \to M$ be a nontangential broken ray with endpoints on E, and let λ be a fixed real number. For any $K > 0$, there is a family $\{v_s : s = \tau + i\lambda, \tau \geq 1\}$ in $C^\infty(M)$ such that, as $\tau \to \infty$,*

$$\|(-\Delta_g - s^2)v_s\|_{L^2(M)} = O(\tau^{-K}), \quad \|v_s\|_{L^2(M)} = O(1),$$

the boundary values of v_s satisfy

$$\|v_s\|_{L^2(R)} = O(\tau^{-K}), \quad \|v_s\|_{L^2(\partial M)} = O(1),$$

and, for any $\psi \in C(M)$,

$$\int_M |v_{\tau+i\lambda}|^2 \psi\, dV_g \to \int_0^L e^{-2\lambda t}\psi(\gamma(t))\,dt \quad \text{as } \tau \to \infty.$$

Let us begin by proving this result in the special case $E = \partial M$, so that $R = \varnothing$ and one does not need to worry about reflected rays. The next three preparatory lemmas describe a modified Fermi coordinate system that is very useful in this construction.

Lemma 7.2. *Let $(\widehat{M}, g)$ be a compact manifold without boundary, and let $\gamma : (a, b) \to \widehat{M}$ be a unit speed geodesic segment that has no loops. There are only finitely many times $t \in (a, b)$ such that γ intersects itself at $\gamma(t)$.*

Proof. Since γ has no loops, $(\gamma(t), \dot\gamma(t)) = (\gamma(t'), \dot\gamma(t'))$ implies $t = t'$. The first observation is that γ can only self-intersect transversally, since $(\gamma(t), \dot\gamma(t)) = (\gamma(t'), -\dot\gamma(t'))$ also implies $t = t'$ (if this would happen for $t < t'$, then, by uniqueness of geodesics, $\dot\gamma((t + t')/2) = -\dot\gamma((t + t')/2)$, which is impossible). Next note that if r is smaller than the injectivity radius of $(\widehat{M}, g)$, any two geodesic segments of length $\leq r$ can intersect transversally in at most one point (locally geodesics are close to straight lines). Partitioning (a, b) in disjoint intervals $\{J_k\}_{k=1}^K$ of length $\leq r$, we have an injective map

$$\{(t, t') \in (a, b)^2 Lt < t' \text{ and } \gamma(t) = \gamma(t')\} \mapsto \{(k, l) \in \{1, \dots, K\}^2 : t \in J_k,\ t' \in J_l\}.$$

Consequently, γ can only self-intersect finitely many times. $\square$

Lemma 7.3. *Let F be a C^1 map from a neighborhood of $(a, b) \times \{0\}$ in $\mathbb{R}^n$ into a smooth manifold such that $F|_{(a,b)\times\{0\}}$ is injective and $DF(t, 0)$ is invertible for $t \in (a, b)$. If $[a_0, b_0]$ is a closed subinterval of (a, b), then F is a C^1 diffeomorphism in some neighborhood of $[a_0, b_0] \times \{0\}$ in $\mathbb{R}^n$.*

Proof. For any $t \in [a_0, b_0]$, the inverse function theorem implies that there is $\varepsilon_t > 0$ such that $F|_{(t-3\varepsilon_t, t+3\varepsilon_t)\times B_{3\varepsilon_t}(0)}$ is a C^1 diffeomorphism. Since $[a_0, b_0]$ is covered by the intervals $(t - \varepsilon_t, t + \varepsilon_t)$, by compactness we have $[a_0, b_0] \subset \bigcup_{j=1}^N (t_j - \varepsilon_j, t_j + \varepsilon_j)$, where $F|_{(t_j-3\varepsilon_j, t_j+3\varepsilon_j)\times B_{3\varepsilon_j}(0)}$ is bijective. We can further assume (upon throwing away or shrinking some intervals if necessary) that the intervals $I_j = (t_j - \varepsilon_j, t_j + \varepsilon_j)$ satisfy $\bar I_j \cap \bar I_k = \varnothing$ unless $|j - k| \leq 1$. Since $\gamma(t) = F(t, 0)$ is injective, we also have $\gamma(\bar I_j) \cap \gamma(\bar I_k) = \varnothing$ unless $|j - k| \leq 1$.

Fix a Riemannian metric in the target manifold, and define

$$\delta = \inf\{\operatorname{dist}(\gamma(\bar I_j), \gamma(\bar I_k)) : |j - k| \geq 2\} > 0.$$

Let $U_j = I_j \times B_\varepsilon(0)$, where $\varepsilon < \min\{\varepsilon_1, \dots, \varepsilon_N\}$ is chosen so small that $F(U_j) \subset \{q : \operatorname{dist}(q, \gamma(\bar I_j)) < \delta/2\}$. Then $F(U_j) \cap F(U_k) = \varnothing$ unless $|j - k| \leq 1$. Define

$$U = \bigcup_{j=1}^N U_j.$$

To show that $F|_U$ is a C^1 diffeomorphism, it is enough to check injectivity. If $F(t, y) = F(t', y')$ for $(t, y), (t', y') \in U$, then, necessarily, $(t, y) \in U_j$, $(t', y') \in U_k$, where $|j - k| \leq 1$. We may assume that $\varepsilon_j \geq \varepsilon_k$. Since $F|_{(t_j-3\varepsilon_j, t_j+3\varepsilon_j)\times B_{3\varepsilon_j}(0)}$ is bijective, we obtain $(t, y) = (t', y')$. $\square$

Lemma 7.4. *Let $(\widehat{M}, g)$ be a compact manifold without boundary, and assume that $\gamma : (a, b) \to \widehat{M}$ is a unit speed geodesic segment with no loops. Given a closed subinterval $[a_0, b_0]$ of (a, b) such that $\gamma|_{[a_0,b_0]}$ self-intersects only at times t_j with $a_0 < t_1 < \cdots < t_N < b_0$ (set $t_0 = a_0$ and $t_{N+1} = b_0$), there is an open cover $\{(U_j, \varphi_j)\}_{j=0}^{N+1}$ of $\gamma([a_0, b_0])$ consisting of coordinate neighborhoods having the following properties.*

(1) *$\varphi_j(U_j) = I_j \times B$, where I_j are open intervals and $B = B(0, \delta)$ is an open ball in $\mathbb{R}^{n-1}$ where δ can be taken arbitrarily small.*

(2) *$\varphi_j(\gamma(t)) = (t, 0)$ for $t \in I_j$.*

(3) *t_j only belongs to I_j and $\bar{I}_j \cap \bar{I}_k = \varnothing$ unless $|j - k| \leq 1$.*

(4) *$\varphi_j = \varphi_k$ on $\varphi_j^{-1}((I_j \cap I_k) \times B)$.*

Further, if S is a hypersurface through $\gamma(a_0)$ that is transversal to $\dot{\gamma}(a_0)$, one can arrange for the map $y \mapsto \varphi_0^{-1}(a_0, y)$ to parametrize S near $\gamma(a_0)$.

Proof. We will use modified Fermi coordinates, constructed as follows. Let $\{v_1, \ldots, v_{n-1}\}$ be an orthonormal set of vectors in $T_{\gamma(a_0)}\widehat{M}$ such that $\{\dot{\gamma}(a_0), v_1, \ldots, v_{n-1}\}$ is a basis. (The case where $\{\dot{\gamma}(a_0), v_1, \ldots, v_{n-1}\}$ is an orthonormal basis corresponds to the usual Fermi coordinates.) Let $E_\alpha(t)$ be the parallel transport of v_α along the geodesic γ. Since $\dot{\gamma}(t)$ is also parallel along γ, the set $\{\dot{\gamma}(t), E_1(t), \ldots, E_{n-1}(t)\}$ is a basis of $T_{\gamma(t)}\widehat{M}$ for $t \in (a, b)$.

Define the function

$$F : (a, b) \times \mathbb{R}^{n-1} \to \widehat{M}, \quad F(t, y) = \exp_{\gamma(t)}(y^\alpha E_\alpha(t)).$$

Here exp is the exponential map in $(\widehat{M}, g)$ and α, β run from 1 to $n - 1$. Then $F(t, 0) = \gamma(t)$ and (with e_α the α-th coordinate vector)

$$\frac{\partial}{\partial s} F(t, se_\alpha)\big|_{s=0} = E_\alpha(t), \quad \frac{\partial}{\partial t} F(t, 0) = \dot{\gamma}(t).$$

Thus F is a C^∞ map near $(a, b) \times \{0\}$ such that $DF(t, 0)$ is invertible for $t \in (a, b)$.

In the case where γ does not self-intersect, $F|_{(a,b)\times\{0\}}$ is injective and Lemma 7.3 implies the existence of a single coordinate neighborhood of $\gamma([a_0, b_0])$ so that (1) and (2) are satisfied (then (3) and (4) are void). In the general case, by Lemma 7.2 the geodesic segment $\gamma|_{[a_0,b_0]}$ only self-intersects at finitely many times t_j with $a_0 < t_1 < \cdots < t_N < b_0$. For some sufficiently small δ, γ is injective on the intervals $(a, t_1 - \delta)$, $(t_1 - 2\delta, t_2 - \delta)$, $\ldots$, $(t_N - 2\delta, b)$ and each interval intersects at most two of the others. Restricting the map F above to suitable neighborhoods corresponding to these intervals (or slightly smaller ones) and using Lemma 7.3, we obtain the required coordinate charts with $\varphi_j = F^{-1}|_{U_j}$.

Let S be a hypersurface transversal to $\dot{\gamma}(a_0)$, and choose some parametrization $y \mapsto q(y)$ of S near $\gamma(a_0)$ satisfying $(\partial/\partial s)q(se_\alpha) = v_\alpha$. We will form a new chart $(\tilde{U}_0, \tilde{\varphi}_0)$ by modifying (U_0, φ_0) so that $y \mapsto \tilde{\varphi}_0^{-1}(a_0, y)$ parametrizes S near $\gamma(a_0)$.

We may assume that $a_0 = 0$, and write $F_0 = \varphi_0^{-1}$, $\widetilde{F}_0 = \tilde{\varphi}_0^{-1}$. It is enough to choose $\widetilde{F}_0 = F_0 \circ \Phi$, where Φ is a diffeomorphism near $\bar{I}_0 \times B$ such that

$$\begin{aligned}
&\Phi(t, 0) = (t, 0), \\
&\Phi(0, y) = F_0^{-1}(q(y)), \\
&\Phi(t, y) = (t, y) \quad \text{for } t > c \text{ with suitable } c > 0.
\end{aligned}$$

Write the components of $\tilde{q} = F_0^{-1} \circ q$ as Taylor series

$$\tilde{q}^j(y) = \tilde{q}^j(0) + \nabla \tilde{q}^j(0) \cdot y + H^j(y) y \cdot y,$$

where H^j are smooth matrices, and $j = 0, \dots, n-1$ (t is the 0-th variable). The properties of q imply that

$$\tilde{q}^j(0) = 0, \quad \partial_\beta \tilde{q}^0(0) = 0, \quad \partial_\beta \tilde{q}^\alpha(0) = \delta_\beta^\alpha.$$

We look for Φ in the form

$$\Phi^j(t, y) = f^j(t) + a^j(t) \cdot y + R^j(t, y) y \cdot y$$

for some smooth functions f^j, vectors a^j, and matrices R^j. The conditions for Φ motivate the choices

$$f^0(t) = t, \quad f^\alpha(t) = 0, \quad a_\beta^0(t) = 0, \quad a_\beta^\alpha(t) = \delta_\beta^\alpha.$$

We choose $R^j(t, y)$ to be a smooth matrix with $R^j(0, y) = H^j(y)$ and $R^j(t, y) = 0$ for $t > c$. Then $D\Phi(t, 0) = \mathrm{Id}$, and Lemma 7.3 ensures that Φ is a diffeomorphism near $\bar{I}_0 \times B$, possibly after decreasing B. □

The next result gives the construction of (nonreflected) Gaussian beam quasimodes associated with a finite length geodesic segment that enters and exits the domain nontangentially. To prepare for the reflected case, we also consider the possibility of prescribing the boundary values of the quasimode at least up to high order at a point. Recall that if f is a smooth function having a critical point at p, the Hessian of f at p is the quadratic form

$$\mathrm{Hess}_p(f)(\dot{\eta}(0), \dot{\eta}(0)) = (f \circ \eta)''(0),$$

where η is any smooth curve with $\eta(0) = p$.

Proposition 7.5. *Let $\gamma : [0, L] \to M$ be any unit speed geodesic in (M, g) such that $\gamma(0), \gamma(L) \in \partial M$, $\dot{\gamma}(0)$ and $\dot{\gamma}(L)$ are nontangential, and $\gamma(t) \in M^{int}$ for $0 < t < L$. Also let λ be a fixed real number. For any $K > 0$ there is a family $\{v_s : s = \tau + i\lambda, \tau \geq 1\}$ in $C^\infty(M)$ such that, as $\tau \to \infty$,*

$$\|(-\Delta_g - s^2) v_s\|_{L^2(M)} = O(\tau^{-K}), \quad \|v_s\|_{L^2(M)} = O(1), \quad \|v_s\|_{L^2(\partial M)} = O(1),$$

and, for any $\psi \in C(M)$,

$$\int_M |v_{\tau+i\lambda}|^2 \psi \, dV_g \to \int_0^L e^{-2\lambda t} \psi(\gamma(t)) \, dt \quad \text{as } \tau \to \infty. \tag{7-1}$$

Given any neighborhood of $\gamma([0, L])$, one can arrange for each v_s to be supported in this neighborhood, and away from the points where γ self-intersects one has $v_s = e^{is\Theta} a$ where Θ and a are smooth complex functions with

$$d\Theta(\gamma(t)) = \dot{\gamma}(t)^\flat, \quad a(\gamma(t)) \neq 0 \text{ for } \tau \text{ large.}$$

If γ does not self-intersect at $\gamma(0)$, the K-th order jets of $\Theta|_{\partial M}$ and $a|_{\partial M}$ can be prescribed freely at $\gamma(0)$ except for the following restrictions: $d\Theta(\gamma(0)) = \dot{\gamma}(0)^\flat$, the Hessian of $\mathrm{Im}(\Theta|_{\partial M})$ at $\gamma(0)$ is positive definite, and $a(\gamma(0)) \neq 0$.

Proof. We embed (M, g) in a compact manifold $(\widehat{M}, g)$ without boundary and extend γ as a unit speed geodesic in $\widehat{M}$. Choose $\varepsilon > 0$ so that $\gamma(t)$ lies in $\widehat{M} \setminus M$ and has no self-intersections in the interval $t \in [-2\varepsilon, 0) \cup (L, L + 2\varepsilon]$. We will construct a Gaussian beam quasimode in a neighborhood of $\gamma([-\varepsilon, L + \varepsilon])$.

Fix a point $p_0 = \gamma(t_0)$ on $\gamma([-\varepsilon, L + \varepsilon])$ and let (t, y) be any local coordinates near p_0, defined in $U = \{(t, y) : t \in I, |y| < \delta\}$ for some open interval I containing t_0, such that p_0 corresponds to $(t_0, 0)$ and the geodesic near p_0 is given by $\Gamma = \{(t, 0) : t \in I\}$. Write $x = (t, y)$, where $x_1 = t$ and $(x_2, \dots, x_m) = y$. We seek to find a quasimode v_s concentrated near Γ, having the form

$$v_s = e^{is\Theta} a,$$

where $s = \tau + i\lambda$, and Θ and a are smooth complex functions near Γ with a supported in $\{|y| < \delta/2\}$.

We compute

$$(-\Delta - s^2) v_s = f,$$

where

$$f = e^{is\Theta}\big(s^2[(\langle d\Theta, d\Theta\rangle - 1)a] - is[2\langle d\Theta, da\rangle + (\Delta\Theta)a] - \Delta a\big).$$

We first choose Θ so that

$$\langle d\Theta, d\Theta\rangle = 1 \quad \text{to } N\text{-th order on } \Gamma. \tag{7-2}$$

In fact, we look for Θ of the form $\Theta = \sum_{j=0}^N \Theta_j$ where

$$\Theta_j(t, y) = \sum_{|\gamma| = j} \frac{\Theta_{j,\gamma}(t)}{\gamma!} y^\gamma.$$

We also write $g^{jk} = \sum_{l=0}^N g_l^{jk} + g_{N+1}^{jk}$, where

$$g_l^{jk}(t, y) = \sum_{|\beta| = l} \frac{g_{l,\beta}^{jk}(t)}{\beta!} y^\beta, \quad g_{N+1}^{jk} = O(|y|^{N+1}).$$

Set $g_l^{jk} = 0$ for $l \geq N + 2$.

With the understanding that j, k run from 1 to m and α, β run from 2 to m, the main part of the argument will consist of finding suitable Θ_0, Θ_1, and Θ_2 in the following form:

$$\Theta_0(t) \text{ real-valued}, \quad \Theta_1(t) = \xi_\alpha(t) y^\alpha \text{ with } \xi_\alpha(t) \text{ real-valued}, \quad \Theta_2(t) = \tfrac{1}{2} H_{\alpha\beta}(t) y^\alpha y^\beta,$$

where $H(t) = (H_{\alpha\beta}(t))$ is a complex symmetric matrix, $H_{\alpha\beta} = H_{\beta\alpha}$, such that $\mathrm{Im}(H(t))$ is positive definite for all t. We also write

$$\xi_1(t) = \partial_t \Theta_0(t).$$

Since $\partial_t \Theta_0 = \xi_1$ and $\partial_\alpha \Theta_1 = \xi_\alpha$, we compute

$$
\begin{aligned}
&g^{jk}\partial_j\Theta\partial_k\Theta - 1 \\
&\quad = g^{11}(\partial_t\Theta_0 + \partial_t\Theta_1 + \cdots)(\partial_t\Theta_0 + \partial_t\Theta_1 + \cdots) + 2g^{1\alpha}(\partial_t\Theta_0 + \partial_t\Theta_1 + \cdots)(\partial_\alpha\Theta_1 + \partial_\alpha\Theta_2 + \cdots) \\
&\qquad\qquad + g^{\alpha\beta}(\partial_\alpha\Theta_1 + \partial_\alpha\Theta_2 + \cdots)(\partial_\beta\Theta_1 + \partial_\beta\Theta_2 + \cdots) - 1 \\
&\quad = g^{jk}\xi_j\xi_k + 2g^{11}\xi_1(\partial_t\Theta_1 + \cdots) + 2g^{1\alpha}\xi_1(\partial_\alpha\Theta_2 + \cdots) \\
&\qquad\qquad + 2g^{1\alpha}\xi_\alpha(\partial_t\Theta_1 + \cdots) + 2g^{\alpha\beta}\xi_\alpha(\partial_\beta\Theta_2 + \cdots) + g^{11}(\partial_t\Theta_1 + \cdots)(\partial_t\Theta_1 + \cdots) \\
&\qquad\qquad + 2g^{1\alpha}(\partial_t\Theta_1 + \cdots)(\partial_\alpha\Theta_2 + \cdots) + g^{\alpha\beta}(\partial_\alpha\Theta_2 + \cdots)(\partial_\beta\Theta_2 + \cdots) - 1 \\
&\quad = g^{jk}\xi_j\xi_k + 2g^{1k}\xi_k(\partial_t\Theta_1 + \cdots) + 2g^{\alpha k}\xi_k(\partial_\alpha\Theta_2 + \cdots) + g^{11}(\partial_t\Theta_1 + \cdots)(\partial_t\Theta_1 + \cdots) \\
&\qquad\qquad + 2g^{1\alpha}(\partial_t\Theta_1 + \cdots)(\partial_\alpha\Theta_2 + \cdots) + g^{\alpha\beta}(\partial_\alpha\Theta_2 + \cdots)(\partial_\beta\Theta_2 + \cdots) - 1.
\end{aligned}
$$

Writing $g^{jk} = g_0^{jk} + g_1^{jk} + \cdots$ and grouping like powers of y, we obtain

$$
\begin{aligned}
&g^{jk}\partial_j\Theta\partial_k\Theta - 1 \\
&= [g_0^{jk}\xi_j\xi_k - 1] + [g_1^{jk}\xi_j\xi_k + 2g_0^{1k}\xi_k\dot{\xi}_\beta y^\beta + 2g_0^{\alpha k}\xi_k H_{\alpha\beta}y^\beta] + (g_2^{jk} + \cdots)\xi_j\xi_k + 2g_0^{1k}\xi_k(\partial_t\Theta_2 + \cdots) \\
&\qquad + 2(g_1^{1k} + \cdots)\xi_k(\partial_t\Theta_1 + \cdots) + 2g_0^{\alpha k}\xi_k(\partial_\alpha\Theta_3 + \cdots) + 2(g_1^{\alpha k} + \cdots)\xi_k(\partial_\alpha\Theta_2 + \cdots) \\
&\qquad + g^{11}(\partial_t\Theta_1 + \cdots)(\partial_t\Theta_1 + \cdots) + 2g^{1\alpha}(\partial_t\Theta_1 + \cdots)(\partial_\alpha\Theta_2 + \cdots) \\
&\qquad + g^{\alpha\beta}(\partial_\alpha\Theta_2 + \cdots)(\partial_\beta\Theta_2 + \cdots). \qquad (7\text{-}3)
\end{aligned}
$$

We can make the two expressions in brackets vanish by choosing $\xi(t)$ to be part of the solution $(x(t), \xi(t))$ of the cogeodesic flow with Hamiltonian $h(x, \xi) = \frac{1}{2}g^{jk}(x)\xi_j\xi_k$,

$$
\begin{aligned}
\dot{x}^j(t) &= \partial_{\xi_j}h(x(t), \xi(t)), \\
\dot{\xi}_j(t) &= -\partial_{x_j}h(x(t), \xi(t)).
\end{aligned}
$$

There is a unique solution with $x(t_0) = p_0$ and $\xi(t_0) = \dot{\gamma}(t_0)^\flat$ (here we raise and lower indices with respect to the metric g). It follows that $x(t)$ is the unit speed geodesic $t \mapsto (t, 0)$, and $\xi^j(t) = \dot{x}^j(t)$. Then $g_0^{jk}\xi_j\xi_k = 1$, and with our choice of coordinates $\xi^1 = 1$ and $\xi^\alpha = 0$ so that also

$$g_0^{1k}\xi_k = 1, \quad g_0^{\alpha k}\xi_k = 0.$$

We further have

$$\dot{\xi}_\beta y^\beta = -\tfrac{1}{2}\partial_{x_\beta}g^{jk}(t, 0)\xi_j\xi_k y^\beta = -\tfrac{1}{2}g_1^{jk}\xi_j\xi_k.$$

Noting that ∂_1 has unit length, we have

$$\xi_1 = g_{1k}(t, 0)\xi^k = 1.$$

Since $\xi_\alpha = g_{\alpha k}(t, 0)\xi^k = g_{\alpha 1}(t, 0)$, we can therefore choose

$$
\begin{aligned}
\Theta_0(t) &= t, \\
\Theta_1(t) &= g_{\alpha 1}(t, 0)y^\alpha.
\end{aligned}
$$

Using these choices and the facts above, in (7-3), the expressions in brackets will indeed vanish, and one obtains

$$
\begin{aligned}
&g^{jk}\partial_j\Theta\partial_k\Theta-1\\
&=(g_2^{jk}+\cdots)\xi_j\xi_k+2(\partial_t\Theta_2+\cdots)+2(g_1^{1k}+\cdots)\xi_k(\partial_t\Theta_1+\cdots)+2(g_1^{\alpha k}+\cdots)\xi_k(\partial_\alpha\Theta_2+\cdots)\\
&\qquad+g^{11}(\partial_t\Theta_1+\cdots)(\partial_t\Theta_1+\cdots)+2g^{1\alpha}(\partial_t\Theta_1+\cdots)(\partial_\alpha\Theta_2+\cdots)+g^{\alpha\beta}(\partial_\alpha\Theta_2+\cdots)(\partial_\beta\Theta_2+\cdots)\\
&=\big(g_2^{jk}\xi_j\xi_k+2\partial_t\Theta_2+2g_1^{\alpha k}\xi_k\partial_\alpha\Theta_2+2g_0^{1\alpha}\partial_t\Theta_1\partial_\alpha\Theta_2+g_0^{\alpha\beta}\partial_\alpha\Theta_2\partial_\beta\Theta_2+2g_1^{1k}\xi_k\partial_t\Theta_1+g_0^{11}(\partial_t\Theta_1)^2\big)\\
&\qquad+\sum_{p=3}^{N}\Bigg(g_p^{jk}\xi_j\xi_k+2\partial_t\Theta_p+2g_1^{\alpha k}\xi_k\partial_\alpha\Theta_p+2g_0^{1\alpha}\partial_t\Theta_1\partial_\alpha\Theta_p+2g_0^{\alpha\beta}\partial_\alpha\Theta_2\partial_\beta\Theta_p\\
&\qquad\qquad+2\sum_{l=1}^{p-1}g_{p-l}^{1k}\xi_k\partial_t\Theta_l+2\sum_{l=2}^{p-1}g_{p-l+1}^{\alpha k}\xi_k\partial_\alpha\Theta_l+\sum_{l=0}^{p-2}g_l^{11}\sum_{\substack{r+s=p-l\\1\le r,s<p}}\partial_t\Theta_r\partial_t\Theta_s\\
&\qquad\qquad+\sum_{l=0}^{p-2}g_l^{1\alpha}\sum_{\substack{r+s=p-l+1\\1\le r<p\\2\le s<p}}\partial_t\Theta_r\partial_\alpha\Theta_s+\sum_{l=0}^{p-2}g_l^{\alpha\beta}\sum_{\substack{r+s=p-l+2\\2\le r,s<p}}\partial_\alpha\Theta_r\partial_\beta\Theta_s\Bigg)+O(|y|^{N+1}).
\end{aligned}
$$

We want to choose Θ_2 so that the first term in brackets vanishes. Recalling that we are looking for Θ_2 in the form $\Theta_2(t,y)=\frac{1}{2}H_{\alpha\beta}(t)y^\alpha y^\beta$, where $H(t)$ is a smooth complex symmetric matrix; it follows that H should satisfy the matrix equation

$$\dot H_{\alpha\beta}y^\alpha y^\beta+2g_1^{\gamma k}\xi_k H_{\gamma\beta}y^\beta+2g_0^{1\gamma}\partial_t\Theta_1 H_{\gamma\beta}y^\beta+g_0^{\gamma\delta}H_{\gamma\alpha}H_{\delta\beta}y^\alpha y^\beta=F_{\alpha\beta}y^\alpha y^\beta,$$

where $F(t)$ is a real-valued smooth symmetric matrix. This can be further written as the matrix Riccati equation

$$\dot H+BH+HB^t+HCH=F,$$

where $B(t)$ and $C(t)$ are real smooth matrices and C is symmetric. More precisely, since $g_1^{jk}=\partial_\alpha g^{jk}(t,0)y^\alpha$, we have

$$B_\alpha^\gamma=\partial_\alpha g^{\gamma k}(t,0)\xi_k+g_0^{1\gamma}\dot\xi_\alpha,\qquad C^{\gamma\delta}=g_0^{\gamma\delta}. \tag{7-4}$$

Choosing $H(t_0)=H_0$, where H_0 is a complex symmetric matrix with $\mathrm{Im}(H_0)$ positive definite, it follows that the Riccati equation has a unique smooth complex symmetric solution $H(t)$ with $\mathrm{Im}(H(t))$ positive definite; see [Katchalov et al. 2001]. This completes the construction of Θ_2. From the lower order terms we can find $\Theta_3,\ldots,\Theta_N$ successively by solving linear first order ODEs on Γ with prescribed initial conditions at t_0. In this way, we obtain a smooth Θ satisfying (7-2).

The next step is to find a such that

$$s[2\langle d\Theta,da\rangle+(\Delta\Theta)a]-i\Delta a=0\quad\text{to } N\text{-th order on }\Gamma.$$

We look for a in the form

$$a=\tau^{(m-1)/4}(a_0+s^{-1}a_{-1}+\cdots+s^{-N}a_{-N})\chi(y/\delta'),$$

where χ is a smooth function with $\chi = 1$ for $|y| \le 1/4$ and $\chi = 0$ for $|y| \ge 1/2$. Writing $\eta = \Delta\Theta$, it is sufficient to determine a_j so that

$$\begin{aligned} 2\langle d\Theta, da_0\rangle + \eta a_0 &= 0 \quad \text{to } N\text{-th order on } \Gamma, \\ 2\langle d\Theta, da_{-1}\rangle + \eta a_{-1} - i\Delta a_0 &= 0 \quad \text{to } N\text{-th order on } \Gamma, \\ &\vdots \\ 2\langle d\Theta, da_{-N}\rangle + \eta a_{-N} - i\Delta a_{-(N-1)} &= 0 \quad \text{to } N\text{-th order on } \Gamma. \end{aligned}$$

Consider $a_0 = a_{00} + \cdots + a_{0N}$, where $a_{0j}(t, y)$ is a polynomial of order j in y, and similarly let $\eta = \eta_0 + \cdots + \eta_N$. We compute

$$\begin{aligned} &2\langle d\Theta, da_0\rangle + \eta a_0 \\ &\quad = 2(g_0^{11} + \cdots)(\partial_t\Theta_0 + \cdots)(\partial_t a_{00} + \cdots) + 2(g_0^{1\alpha} + \cdots)(\partial_t\Theta_0 + \cdots)(\partial_\alpha a_{01} + \cdots) \\ &\qquad + 2(g_0^{1\alpha} + \cdots)(\partial_\alpha\Theta_1 + \cdots)(\partial_t a_{00} + \cdots) + 2(g_0^{\alpha\beta} + \cdots)(\partial_\beta\Theta_1 + \cdots)(\partial_\alpha a_{01} + \cdots) \\ &\qquad\qquad + (\eta_0 + \eta_1 + \cdots)(a_{00} + a_{01} + \cdots). \end{aligned}$$

Recalling that $\partial_t\Theta_0 = \xi_1 = 1$ and $\partial_\alpha\Theta_1 = \xi_\alpha$, where $g_0^{1j}\xi_j = 1$ and $g_0^{\alpha j}\xi_j = 0$, we obtain

$$\begin{aligned} &2\langle d\Theta, da_0\rangle + \eta a_0 \\ &= 2[g_0^{11}\xi_1 + g_0^{11}(\partial_t\Theta_1 + \cdots) + (g_1^{11} + \cdots)(\partial_t\Theta_0 + \cdots) + g_0^{1\alpha}\xi_\alpha + g_0^{1\alpha}(\partial_\alpha\Theta_2 + \cdots) + (g_1^{1\alpha} + \cdots)(\partial_\alpha\Theta_1 + \cdots)] \\ &\qquad\qquad \times (\partial_t a_{00} + \cdots) \\ &\quad + 2\big[g_0^{1\alpha}(\partial_t\Theta_1 + \cdots) + (g_1^{1\alpha} + \cdots)(\partial_t\Theta_0 + \cdots) + g_0^{\alpha\beta}(\partial_\beta\Theta_2 + \cdots) + (g_1^{\alpha\beta} + \cdots)(\partial_\beta\Theta_1 + \cdots)\big] \\ &\qquad\qquad \times (\partial_\alpha a_{01} + \cdots) + (\eta_0 + \eta_1 + \cdots)(a_{00} + a_{01} + \cdots) \\ &= [2\partial_t a_{00} + \eta_0 a_{00}] + \sum_{p=1}^{N} [2\partial_t a_{0p} + q_p^{\alpha\beta} y^\beta \partial_\alpha a_{0p} + \eta_0 a_{0p} + F_p] + O(|y|^{N+1}), \end{aligned}$$

where $q_p^{\alpha\beta}(t)$ are smooth functions only depending on g and Θ, and $F_p(t, y)$ is a polynomial of degree p in y that only depends on g, Θ, η, and $a_{00}, \dots, a_{0,p-1}$.

We want to choose a_{00} so that the first term in brackets vanishes, that is,

$$\partial_t a_{00} + t\tfrac{1}{2}\eta_0 a_{00} = 0.$$

This has the solution

$$a_{00}(t) = c_0 e^{-(1/2)\int_{t_0}^{t} \eta_0(s)\, ds}, \quad a_{00}(t_0) = c_0.$$

We obtain $a_{01}, \dots, a_{0N}$ successively by solving linear first order ODEs with prescribed initial conditions at t_0. The functions $a_1, \dots, a_N$ may be determined in a similar way so that the required equations are satisfied to N-th order on Γ. This completes the construction of a.

To review what has been achieved so far, we have constructed a function $v_s = e^{is\Theta}a$ in U, where

$$\begin{aligned}\Theta(t, y) &= t + \xi_\alpha(t)y^\alpha + \tfrac{1}{2}H_{\alpha\beta}(t)y^\alpha y^\beta + \widetilde{\Theta},\\ a(t, y) &= \tau^{(m-1)/4}(a_0 + s^{-1}a_{-1} + \cdots + s^{-N}a_{-N})\chi(y/\delta'),\\ a_0(t, 0) &= c_0 e^{-(1/2)\int_{t_0}^{t}\eta_0(s)\,ds}.\end{aligned}$$

Here $\widetilde{\Theta} = O(|y|^3)$ and $\widetilde{\Theta}$ and each a_j are independent of τ. Also, $f = (-\Delta - s^2)v_s$ has the form

$$f = e^{is\Theta}\tau^{(m-1)/4}(s^2h_2a + sh_1 + \cdots + s^{-(N-1)}h_{-(N-1)} - s^{-N}\Delta a_{-N})\chi(y/\delta') + e^{is\Theta}\tau^{(m-1)/4}sb\widetilde{\chi}(y/\delta'),$$

where, for each j, one has $h_j = 0$ to N-th order on Γ, b vanishes near Γ, and $\widetilde{\chi}$ is a smooth function with $\widetilde{\chi} = 0$ for $|y| \geq 1/2$. We also note that $d\Theta(\gamma(t)) = \dot{\gamma}(t)^\flat$ and $\mathrm{Hess}_{\gamma(t_0)}(\mathrm{Im}(\Theta|_{\{t=t_0\}})) = \mathrm{Im}(H(t_0))$.

To prove the norm estimates for v_s in U, note that

$$|e^{is\Theta}| = e^{-\lambda\,\mathrm{Re}\,\Theta}e^{-\tau\,\mathrm{Im}\,\Theta} = e^{-\lambda t}e^{-(1/2)\tau\,\mathrm{Im}(H(t))y\cdot y}e^{-\lambda O(|y|)}e^{-\tau O(|y|^3)}.$$

Here $\mathrm{Im}(H(t))y \cdot y \geq c|y|^2$ for $(t, y) \in U$, where $c > 0$ depends on H_0, g, and I. By decreasing δ' if necessary, this shows the following bound when t in a fixed compact set:

$$|v_s(t, y)| \lesssim \tau^{(m-1)/4}e^{-(1/4)c\tau|y|^2}\chi(y/\delta').$$

Integrating the square of this over U, we get, as $\tau \to \infty$,

$$\|v_s\|_{L^2(U)} \lesssim \|\tau^{(m-1)/4}e^{-(1/4)c\tau|y|^2}\|_{L^2(U)} = O(1).$$

Similarly, we have

$$\begin{aligned}\|(-\Delta - s^2)v_s\|_{L^2(U)} &\lesssim \|\tau^{(m-1)/4}e^{-(1/4)c\tau|y|^2}(\tau^2|y|^{N+1} + \tau^{-N})\|_{L^2(U)}\\ &= O(\tau^{(3-N)/2}).\end{aligned}$$

The norm estimates for v_s in U follow upon replacing N by $2K+3$.

For the $L^2(\partial M)$ estimate, if U contains a boundary point $x_0 = (t_0, 0) \in \partial M$, by assumption $(\partial/\partial t)|_{x_0}$ is transversal to ∂M. If ρ is a boundary defining function for M, so ∂M is given as the zero set $\rho(t, y) = 0$ near x_0 and $\nabla\rho = -\nu$ on ∂M, then $(\partial\rho/\partial t)(x_0) \neq 0$ and, by the implicit function theorem, there is a smooth function $y \mapsto t(y)$ near 0 such that ∂M is given by $\{(t(y), y) : |y| < r_0\}$ near x_0. The bound for v_s given above implies that, for δ' small,

$$\|v_s\|^2_{L^2(\partial M\cap U)} = \int_{|y|<r_0}|v_s(t(y), y)|\,dS(y) \lesssim \int_{|y|<r_0}\tau^{(m-1)/2}e^{-(1/2)c\tau|y|^2}\,dy = O(1).$$

At this point we can construct the quasimode v_s in M from the corresponding quasimodes defined on small pieces. Let $\gamma([-\varepsilon, L+\varepsilon])$ be covered by open sets $U^{(0)}, \ldots, U^{(N+1)}$ as in Lemma 7.4, and note that each $U^{(j)}$ corresponds to $I_j \times B(0, \delta)$ in the (t, y) coordinates. Suppose first that γ does not self-intersect at time $t = 0$. We find a quasimode $v^{(0)} = e^{is\Theta^{(0)}}a^{(0)}$ in $U^{(0)}$ by the above procedure, with some fixed initial conditions at $t = 0$ for the ODEs determining $\Theta^{(0)}$ and $a^{(0)}$. Choose some t_0' with $\gamma(t_0') \in U^{(0)} \cap U^{(1)}$, and construct a quasimode $v^{(1)} = e^{is\Theta^{(1)}}a^{(1)}$ in $U^{(1)}$ by choosing the initial conditions for the ODEs for $\Theta^{(1)}$ and $a^{(1)}$ at t_0' to be the corresponding values of $\Theta^{(0)}$ and $a^{(0)}$ at t_0'. Continuing in this way, we obtain $v^{(2)}, \dots, v^{(N+1)}$. If γ self-intersects at $t=0$, we start the construction from $v^{(1)}$ fixing initial conditions for the ODEs at $t=0$, and find $v^{(0)}$ by going backward.

Let $\{\widetilde{\chi}_j(t)\}$ be a partition of unity near $[-\varepsilon, L+\varepsilon]$ corresponding to the cover $\{I_j\}$, let $\chi_j(t,y) = \widetilde{\chi}_j(t)$ on $U^{(j)}$, and define

$$v_s = \sum_{j=0}^{N+1} \chi_j v^{(j)}.$$

Then v_s is smooth in $\widehat{M}$ and it is supported in a small neighborhood of $\gamma([-\varepsilon, L+\varepsilon])$. The important point is that since the ODEs for the phase functions and amplitudes have the same initial data in $U^{(j)}$ and in $U^{(j+1)}$, and since the local coordinates φ_j and φ_{j+1} coincide on $\varphi_j^{-1}((I_j \cap I_{j+1}) \times B)$, one actually has $v^{(j)} = v^{(j+1)}$ in $\varphi_j^{-1}((I_j \cap I_{j+1}) \times B)$. Letting $p_1, \dots, p_R$ be the points where γ intersects itself, we choose an open cover of $\operatorname{supp}(v_s) \cap M$,

$$\operatorname{supp}(v_s) \cap M \subset \left(\bigcup_{r=1}^{R} V_r \right) \cup \left(\bigcup_{j=0}^{N+1} (W_{j,0} \cup W_{j,1}) \right),$$

where V_r are small neighborhoods of the points p_r and $W_{j,0}, W_{j,1} \subset U^{(j)}$, such that

$$v_s|_{V_r} = \sum_{\gamma(t_j)=p_r} v^{(j)} \quad \text{and} \quad v_s|_{W_{j,l}} = v^{(j+l)}.$$

Since v_s is a finite sum of the $v^{(j)}$ in each case, the $L^2(M)$ bounds for v_s and $(-\Delta - s^2)v_s$ and the $L^2(\partial M)$ bounds for v_s follow from corresponding bounds for the $v^{(j)}$. The form of v_s near points where γ does not self-intersect and the possibility to prescribe the K-th order jets of $\Theta|_{\partial M}$ and $a|_{\partial M}$ at $\gamma(0)$ follow from the construction and Lemma 7.4.

To conclude the proof, using a partition of unity, it is enough to verify the limit (7-1) for any ψ supported in one of the sets $V_r \cap M$ or $W_{j,l} \cap M$. Further, we can choose the sets V_r to be so small that the real part of $d\Theta^{(j)} - d\Theta^{(k)}$ is nonvanishing near V_r if $\gamma(t_j) = \gamma(t_k) = p_r$ but $j \neq k$. This follows since

$$\operatorname{Re}(d\Theta^{(j)} - d\Theta^{(k)})(p_r) = \dot{\gamma}(t_j)^\flat - \dot{\gamma}(t_k)^\flat \neq 0.$$

Here we may need to decrease δ so that we still have an open cover.

Consider first the case where $\psi \in C_c(W_{j,l} \cap M)$. Here the support of ψ may reach ∂M, and we extend ψ by zero outside of $W_{j,l} \cap M$. Suppose that $v_s = e^{is\Theta}a$, where $\Theta = t + \xi_\alpha y^\alpha + \frac{1}{2}H(t)y \cdot y + O(|y|^3)$ and $a = \tau^{(m-1)/4}(a_0 + O(\tau^{-1}))\chi(y/\delta')$, and let $\rho = |g|^{1/2}$. Then

$$\begin{aligned}
&\int_M |v_{\tau+i\lambda}|^2 \psi \, dV_g \\
&\quad = \int_{-\infty}^{\infty} \int_{\mathbb{R}^{m-1}} e^{-2\lambda t} e^{-\tau \operatorname{Im}(H(t)) y\cdot y} e^{\tau O(|y|^3)} e^{O(|y|)} \tau^{(m-1)/2} (|a_0|^2 + O(\tau^{-1})) \chi(y/\delta')^2 \psi \rho \, dt \, dy \\
&\quad = \int_{-\infty}^{\infty} e^{-2\lambda t} \int_{\mathbb{R}^{m-1}} e^{-\operatorname{Im}(H(t)) y\cdot y} e^{\tau^{-1/2} O(|y|^3)} e^{\tau^{-1/2} O(|y|)} (|a_0(t, \tau^{-1/2} y)|^2 \\
&\qquad\qquad + O(\tau^{-1})) \chi(y/\tau^{1/2}\delta')^2 \psi(t, \tau^{-1/2} y) \rho(t, \tau^{-1/2} y) \, dt \, dy.
\end{aligned}$$

Since $\mathrm{Im}(H(t))$ is positive definite and δ' is sufficiently small, the term $e^{-\mathrm{Im}(H(t))y\cdot y}$ dominates the other exponentials and one obtains

$$\lim_{\tau\to\infty}\int_M |v_{\tau+i\lambda}|^2\psi\,dV_g=\int_0^L e^{-2\lambda t}\biggl(\int_{\mathbb{R}^{m-1}} e^{-\mathrm{Im}(H(t))y\cdot y}\,dy\biggr)|a_0(t,0)|^2\psi(t,0)\rho(t,0)\,dt.$$

Evaluating the integral over y gives

$$\lim_{\tau\to\infty}\int_M |v_{\tau+i\lambda}|^2\psi\,dV_g=C_m\int_0^L e^{-2\lambda t}\frac{|a_0(t,0)|^2\rho(t,0)}{\sqrt{\det\mathrm{Im}(H(t))}}\psi(t,0)\,dt.$$

We will prove below that

$$\frac{|a_0(t,0)|^2\rho(t,0)}{\sqrt{\det\mathrm{Im}(H(t))}}=\text{const.} \tag{7-5}$$

The limit (7-1) will follow upon dividing the family $\{v_s\}$ by a suitable constant.

If $\psi\in C_c(V_r\cap M)$ (again $\mathrm{supp}(\psi)$ may extend up to ∂M), we have

$$v_s|_{V_r}=\sum_{\gamma(t_j)=p_r} v^{(j)},$$

so that on V_r

$$|v_s|^2=\sum_{\gamma(t_j)=p_r}|v^{(j)}|^2+\sum_{\substack{\gamma(t_j)=\gamma(t_k)=p_r\\ j\neq k}} v^{(j)}\overline{v^{(k)}}.$$

We arranged earlier for $\mathrm{Re}(d\Theta^{(j)}-d\Theta^{(k)})$ to be nonvanishing near V_r if $\gamma(t_j)=\gamma(t_k)=p_r$ but $j\neq k$. Thus the cross terms give rise to terms of the form

$$\int_{V_r\cap M} v^{(j)}\overline{v^{(k)}}\psi\,dV=\int_{V_r\cap M} e^{i\tau\varphi}w^{(j)}\overline{w^{(k)}}\psi\,dV,$$

where $\varphi=\mathrm{Re}(\Theta^{(j)}-\Theta^{(k)})$ has nonvanishing gradient in V_r, and $w^{(l)}=e^{is\,\mathrm{Im}(\Theta^{(l)})}e^{-\lambda\,\mathrm{Re}(\Phi^{(l)})}a^{(l)}$. We wish to prove that

$$\lim_{\tau\to\infty}\int_{V_r\cap M} e^{i\tau\varphi}w^{(j)}\overline{w^{(k)}}\psi\,dV=0,\quad j\neq k, \tag{7-6}$$

showing that the cross terms vanish in the limit and the previous computation for $|v^{(l)}|^2$ shows the limit (7-1) also when ψ is supported in some $V_r\cap M$. To show (7-6), let $\varepsilon>0$ and decompose $\psi=\psi_1+\psi_2$, where $\psi_1\in C_c^\infty(V_r\cap M)$ and $\|\psi_2\|_{L^\infty(V_r\cap M)}\leq\varepsilon$. Then

$$\biggl|\int_{V_r\cap M} e^{i\tau\varphi}w^{(j)}\overline{w^{(k)}}\psi_2\,dV\biggr|\lesssim\|w^{(j)}\|_{L^2}\|w^{(k)}\|_{L^2}\|\psi_2\|_{L^\infty}\lesssim\varepsilon,$$

since $\|w^{(l)}\|_{L^2}\lesssim\|v^{(l)}\|_{L^2}\lesssim 1$. For the smooth part ψ_1, we employ a nonstationary phase argument and integrate by parts using that

$$e^{i\tau\varphi}=\frac{1}{i\tau}L(e^{i\tau\varphi}),\quad Lw=\langle|d\varphi|^{-2}d\varphi,dw\rangle.$$

This gives

$$\int_{V_r\cap M} e^{i\tau\varphi} w^{(j)}\overline{w^{(k)}}\psi_1\,dV = \int_{\partial M} \frac{\partial_\nu\varphi}{i\tau|d\varphi|^2} v^{(j)}\overline{v^{(k)}}\psi_1\,dS + \frac{1}{i\tau}\int_{V_r\cap M} e^{i\tau\varphi} L^t(w^{(j)}\overline{w^{(k)}}\psi_1)\,dV.$$

Since $\|v^{(l)}\|_{L^2(\partial M)} = O(1)$, the boundary term can be made arbitrarily small as $\tau\to\infty$. As for the last term, the worst behavior is when the transpose L^t acts on $e^{is\operatorname{Im}(\Theta^{(l)})}$, and these terms have bounds of the form

$$\big\| |d(\operatorname{Im}(\Theta^{(j)}))| v^{(j)}\big\|_{L^2}\|v^{(k)}\|_{L^2}\|\psi_1\|_{L^\infty}.$$

Here $|d(\operatorname{Im}(\Theta^{(j)}))| \lesssim |y|$ if (t,y) are coordinates along the geodesic segment corresponding to $v^{(j)}$, and the computation above for $\|v^{(j)}\|_{L^2}$ shows that

$$\big\| |d(\operatorname{Im}(\Theta^{(j)}))| v^{(j)}\big\|_{L^2}\|v^{(k)}\|_{L^2}\|\psi_1\|_{L^\infty} \lesssim \tau^{-1/2}.$$

This finishes the proof of (7-6) and also of (7-1).

It remains to show (7-5). We have

$$|a_0(t,0)|^2\rho(t,0) = |c_0|^2 e^{-\int_{t_0}^t \operatorname{Re}(\eta_0)(s)\,ds}|g(t,0)|^{1/2}.$$

Note that $\eta_0(t)$ is given by

$$\begin{aligned}
&\eta_0(t)\\
&\quad= \Delta\Theta(t,0)\\
&\quad= (g^{jk}\partial_{jk}\Theta + \partial_j g^{jk}\partial_k\Theta + |g|^{-1/2}\partial_j(|g|^{1/2})g^{jk}\partial_k\Theta)(t,0)\\
&\quad= g^{11}\partial_t^2\Theta_0 + 2g^{1\alpha}\partial_{t\alpha}\Theta_1 + g^{\alpha\beta}\partial_{\alpha\beta}\Theta_2 + \partial_j g^{j1}\partial_t\Theta_0 + \partial_j g^{j\alpha}\partial_\alpha\Theta_1 + \tfrac{1}{2}\partial_j(\log|g|)(g^{j1}\partial_t\Theta_0 + g^{j\alpha}\partial_\alpha\Theta_1)\\
&\quad= 2g^{1\alpha}\dot{\xi}_\alpha + g^{\alpha\beta}H_{\alpha\beta} + (\partial_j g^{jk})\xi_k + \tfrac{1}{2}\partial_j(\log|g|)g^{jk}\xi_k.
\end{aligned}$$

The conditions $g^{jk}\xi_k = \delta_1^j$ and $g^{1\alpha}\dot{\xi}_\alpha = g^{1k}\dot{\xi}_k = -(\partial_t g^{1k})\xi_k$ at $(t,0)$, together with the general fact that $\partial_t(\log|g|) = -g_{jk}\partial_t g^{jk}$, imply that

$$\eta_0(t) = g^{1\alpha}\dot{\xi}_\alpha + g^{\alpha\beta}H_{\alpha\beta} + (\partial_\alpha g^{\alpha k})\xi_k + \tfrac{1}{2}\partial_t(\log|g|).$$

Recalling the definition of the B and C matrices in (7-4), this says precisely that

$$\begin{aligned}
\eta_0(t) &= B_\alpha^\alpha + C^{\alpha\gamma}H_{\gamma\alpha} + \tfrac{1}{2}\partial_t(\log|g|)\\
&= \operatorname{tr}(B(t) + C(t)H(t)) + \tfrac{1}{2}\partial_t(\log|g|).
\end{aligned}$$

Consequently, $|a_0(t,0)|^2\rho(t,0) = c_0' e^{-\int_{t_0}^t \operatorname{tr}(B(s)+C(s)\operatorname{Re}(H(s)))\,ds}$. On the other hand, by [Katchalov et al. 2001, Lemma 2.58], solutions of the matrix Riccati equation have the property that

$$\det\operatorname{Im}(H(t)) = \det\operatorname{Im}(H(t_0))e^{-2\int_{t_0}^t \operatorname{tr}(B(s)+C(s)\operatorname{Re}(H(s)))\,ds}.$$

This proves the result. □

The proof of Proposition 7.1 now follows quickly from the way we have set up the previous result.

Proof of Proposition 7.1. Let $\gamma : [0, L] \to M$ be a nontangential broken ray with endpoints on E, and let $0 < t_1 < \cdots < t_N < L$ be the times of reflection. Let $v_s^{(0)}$ be a Gaussian beam quasimode as in Proposition 7.5 associated with the geodesic $\gamma|_{[0,t_1]}$. We will construct another Gaussian beam quasimode $v_s^{(1)}$ associated with $\gamma|_{[t_1,t_2]}$ such that $v_s^{(0)} - v_s^{(1)}|_{\partial M}$ will be small near $\gamma(t_1)$.

In fact, by Proposition 7.5 we have $v_s^{(j)} = e^{is\Theta^{(j)}} a^{(j)}$ near $\gamma(t_1)$, and we can choose the K-th order jet of $\Theta^{(1)}|_{\partial M}$ at $\gamma(t_1)$ to be equal to that of $\Theta^{(0)}|_{\partial M}$ with the following exception: we always have

$$d(\Theta^{(0)})|_{\gamma(t_1)} = \dot{\gamma}(t_1-)^\flat,$$
$$d(\Theta^{(1)})|_{\gamma(t_1)} = \dot{\gamma}(t_1+)^\flat.$$

It follows that

$$d(\Theta^{(0)}|_{\partial M})|_{\gamma(t_1)} = \dot{\gamma}(t_1-)^\flat_{\tan},$$
$$d(\Theta^{(1)}|_{\partial M})|_{\gamma(t_1)} = \dot{\gamma}(t_1+)^\flat_{\tan},$$

where we have taken the projections to the cotangent space of ∂M at $\gamma(t_1)$. But by the rule that the angle of incidence equals angle of reflection, $\dot{\gamma}(t_1-)^\flat_{\tan}$ equals $\dot{\gamma}(t_1+)^\flat_{\tan}$. Thus the K-th order jets of $\Theta^{(0)}|_{\partial M}$ and $\Theta^{(1)}|_{\partial M}$ actually coincide at $\gamma(t_1)$, and by Proposition 7.5 we can also arrange for the K-th order jets of $a^{(0)}|_{\partial M}$ and $a^{(1)}|_{\partial M}$ to coincide at $\gamma(t_1)$.

Write $f_s = v_s^{(0)} - v_s^{(1)}|_{\partial M}$, and let (t, y) be coordinates near $\gamma(t_1)$ such that ∂M is parametrized by $y \mapsto (t_1, y)$ and $\gamma(t_1)$ corresponds to $(t_1, 0)$. Recall that $v_s^{(j)}$ are supported in small tubular neighborhoods of the corresponding geodesic segments. By the above considerations and the construction of $\Theta^{(j)}$ and $a^{(j)}$, and dropping the variable t_1 from the notations, the restrictions of $\Theta^{(j)}$ and $a^{(j)}$ to ∂M satisfy

$$\Theta^{(j)}(y) = \Theta(y) + \Xi^{(j)}(y), \quad a^{(j)}(y) = a(y) + b^{(j)}(y),$$

where Θ is a polynomial of order K, $a = \tau^{(m-1)/4}\tilde{a}\chi(y/\delta')$, where $\tilde{a}$ is a polynomial of order K, and $|\Xi^{(j)}(y)| \leq C|y|^{K+1}$ and $|b^{(j)}(y)| \leq C\tau^{(m-1)/4}|y|^{K+1}\chi(y/\delta')$ on $\operatorname{supp}(\chi(\,\cdot\,/\delta'))$, where χ is a cutoff function and δ' is a constant independent of τ that can be chosen as small as we want (these initially depend on j, but since there are finitely many reflections, we can choose them independently of j). Here Θ and $\Xi^{(j)}$ are independent of τ, and a and $b^{(j)}$ are mildly τ-dependent and satisfy uniform bounds with respect to τ. Then

$$f_s = e^{is\Theta}((e^{is\Xi^{(0)}} - e^{is\Xi^{(1)}})a + e^{is\Xi^{(0)}}b^{(0)} - e^{is\Xi^{(1)}}b^{(1)}).$$

We have

$$e^{is\Xi^{(0)}} - e^{is\Xi^{(1)}} = is(\Xi^{(0)} - \Xi^{(1)}) \int_0^1 e^{is(r\Xi^{(0)}+(1-r)\Xi^{(1)})}\,dr$$

and, consequently, near $y = 0$,

$$|e^{is\Xi^{(0)}} - e^{is\Xi^{(1)}}| \leq C\tau|y|^{K+1}e^{C\tau|y|^{K+1}}.$$

Thus, near $y = 0$,

$$|f_s(y)| \leq C\tau^{(m-1)/4}e^{-\tau\,\mathrm{Im}(\Theta)}\tau|y|^{K+1}e^{C\tau|y|^{K+1}}\chi(y/\delta').$$

Using that the Hessian of $\mathrm{Im}(\Theta)$ at 0 is positive definite and choosing δ' sufficiently small, we have

$$|f_s(y)| \le C\tau^{(m-1)/4} e^{-c\tau|y|^2} \tau|y|^{K+1}\chi(y/\delta').$$

Integrating the square of $|f_s|$ over $\mathbb{R}^{m-1}$ and changing y to $\tau^{-1/2}y$, we obtain

$$\|f_s\|_{L^2(R_1)} = O(\tau^{-(K-1)/2}),$$

where R_1 is a small neighborhood of $\gamma(t_1)$ on ∂M containing the set of interest.

Repeating this construction for the other points of reflection, we end up with a quasimode

$$v_s = \sum_{j=0}^{N} (-1)^j v_s^{(j)}$$

that is supported in a small neighborhood of the broken ray γ. Since all points of reflection are distinct, we can arrange that the quasimode satisfies

$$\|v_s|_R\|_{L^2(R)} = O(\tau^{-(K-1)/2}).$$

It also satisfies

$$\|(-\Delta - s^2)v_s\|_{L^2(M)} = O(\tau^{-K}), \quad \|v_s\|_{L^2(M)} = O(1).$$

Replacing K by $2K+1$, we have proved all the other statements in the proposition except for the expression of the limit measure. To do this, we consider the finitely many points where the full broken ray γ self-intersects or reflects, and decompose the terms $v_s^{(j)}$ as in the proof of Proposition 7.5 to parts living in small neighborhoods of the self-intersection and reflection points and parts away from these points. Now all self-intersection points are in the interior or on E and all self-intersections must be transversal, and also all reflections are transversal. Consequently, when forming $|v_s|^2$, the cross terms arising from different parts living near the same self-intersection or reflection point contribute an $o(1)$ term by nonstationary phase as in the proof of Proposition 7.5. Thus the limit measure of $|v_s|^2\, dV_g$ is indeed the measure $e^{-2\lambda t}\delta_\gamma$, where δ_γ is the delta function of the broken ray γ. □

8. Recovering the broken ray transform

In this section we give the proof of Theorem 2.4 concerning the recovery of integrals over broken rays.

Proof of Theorem 2.4. The proof is very similar to the proof of Theorem 2.1, except that we use reflected Gaussian beam quasimodes instead of WKB type quasimodes. Let $\gamma : [0, L] \to M_0$ be a nontangential broken ray with endpoints on E, and let $\lambda > 0$. Also let $\tilde{g} = e \oplus g_0$ and $\tilde{q}_j = c(q_j - c^{(n-2)/4}\Delta_g(c^{-(n-2)/4}))$. Consider the complex frequency

$$s = \tau + i\lambda,$$

where $\tau > 0$ will be large. We look for solutions

$$\tilde{u}_1 = e^{-sx_1}(v_s(x') + r_1),$$
$$\tilde{u}_2 = e^{sx_1}(v_s(x') + r_2)$$

of the equations $(-\Delta_{\tilde{g}} + \tilde{q}_1)\tilde{u}_1 = 0$, $(-\Delta_{\tilde{g}} + \overline{\tilde{q}}_2)\tilde{u}_2 = 0$ in M. Here $v_s \in C^\infty(M_0)$ is the quasimode constructed in Proposition 7.1 that concentrates near the given broken ray γ and is small on $\partial M_0 \setminus E$.

Since $\Delta_{\tilde{g}} = \partial_1^2 + \Delta_{g_0}$, the function $\tilde{u}_1$ is a solution if and only if

$$e^{sx_1}(-\Delta_{\tilde{g}} + \tilde{q}_1)(e^{-sx_1} r_1) = -(-\Delta_{g_0} + \tilde{q}_1 - s^2) v_s(x') \quad \text{in } M.$$

We look for a solution in the form $r_1 = e^{i\lambda x_1} r_1'$ where r_1' satisfies

$$e^{\tau x_1}(-\Delta_{\tilde{g}} + \tilde{q}_1)(e^{-\tau x_1} r_1') = f \quad \text{in } M$$

with

$$f = -e^{-i\lambda x_1}(-\Delta_{g_0} + \tilde{q}_1 - s^2) v_s(x').$$

To arrange that $\tilde{u}_1|_{\Gamma_i} = 0$, fix some small $\delta > 0$, let $S_\pm$ and S_0 be the sets in Proposition 4.3 with Carleman weight $\varphi(x) = -x_1$, and consider the boundary condition

$$e^{\tau\varphi} r_1'|_{S_- \cup S_0} = e^{\tau\varphi} f_-,$$

where

$$f_- = \begin{cases} -e^{-i\lambda x_1} v_s(x') & \text{on } \Gamma_i, \\ 0 & \text{on } (S_- \cup S_0) \setminus \Gamma_i. \end{cases}$$

For any fixed $K > 0$, by Proposition 7.1 and by the condition that $\Gamma_i \subset \mathbb{R} \times (\partial M_0 \setminus E)$, we may assume that the following bounds are valid:

$$\|f\|_{L^2(M)} = O(1), \quad \|f_-\|_{L^2(S_-)} = 0, \quad \|f_-\|_{L^2(S_0)} = O(\tau^{-K}).$$

It follows from Proposition 4.3 that there is a solution r_1' satisfying the above boundary condition and having the estimate

$$\|r_1'\|_{L^2(M)} = O(\tau^{-1}).$$

Choosing r_1' as described above and choosing $r_1 = e^{i\lambda x_1} r_1'$, we have produced a solution $\tilde{u}_1 \in H_{\Delta_{\tilde{g}}}(M)$ of the equation $(-\Delta_{\tilde{g}} + \tilde{q}_1)\tilde{u}_1 = 0$ in M, having the form

$$\tilde{u}_1 = e^{-sx_1}(v_s(x') + r_1)$$

and satisfying

$$\operatorname{supp}(\tilde{u}_1|_{\partial M}) \subset \overline{\partial M_+ \cup \partial M_- \cup \Gamma_a}$$

and $\|r_1\|_{L^2(M)} = O(\tau^{-1})$ as $\tau \to \infty$. Repeating this construction for the Carleman weight $\varphi(x) = x_1$, we obtain a solution $\tilde{u}_2 \in H_{\Delta_{\tilde{g}}}(M)$ of the equation $(-\Delta_{\tilde{g}} + \overline{\tilde{q}}_2)\tilde{u}_2 = 0$ in M, having the form

$$\tilde{u}_2 = e^{sx_1}(v_s(x') + r_2)$$

satisfying the same support condition and bound for $\|r_2\|_{L^2(M)}$.

Writing $u_j = c^{-(n-2)/4}\tilde{u}_j$, Lemma 6.2 shows that $u_j \in H_{\Delta_g}(M)$ are solutions of $(-\Delta_g + q_1)u_1 = 0$ and $(-\Delta_g + \overline{q}_2)u_2 = 0$ in M. Then Proposition 6.1 implies that

$$\int_M (q_1 - q_2) u_1 \bar{u}_2 \, dV_g = 0.$$

We extend $q_1 - q_2$ by zero to $\mathbb{R} \times M_0$. Inserting the expressions for u_j and using the equality $dV_g = c^{n/2}\, dx_1\, dV_{g_0}(x')$, we obtain

$$\int_{M_0} \int_{-\infty}^{\infty} (q_1 - q_2) c e^{-2i\lambda x_1} (|v_s(x')|^2 + v_s \bar{r}_2 + \bar{v}_s r_1 + r_1 \bar{r}_2)\, dx_1\, dV_{g_0}(x') = 0.$$

Since $\|r_j\|_{L^2(M)} = O(\tau^{-1})$ as $\tau \to \infty$, Proposition 7.1 implies that

$$\int_0^L e^{-2\lambda t} (c(q_1 - q_2))\hat{\ }(2\lambda, \gamma(t))\, dt = 0.$$

This concludes the proof. $\square$

Acknowledgements

Kenig is partly supported by the NSF. Salo is supported in part by the Academy of Finland and an ERC Starting Grant. Salo thanks David dos Santos Ferreira, Yaroslav Kurylev, and Matti Lassas for several helpful discussions and for allowing the authors to use arguments from [Ferreira 2013], which was in preparation simultaneously with this manuscript and involved Gaussian beam quasimodes and a reduction from the attenuated ray transform to the usual ray transform. Salo also thanks the Department of Mathematics of the University of Chicago, where part of this work was carried out.

References

[Albin et al. 2013] P. Albin, C. Guillarmou, L. Tzou, and G. Uhlmann, "Inverse boundary problems for systems in two dimensions", *Ann. Henri Poincaré* **14**:6 (2013), 1551–1571. MR 3085925 Zbl 06195029

[Alessandrini and Kim 2012] G. Alessandrini and K. Kim, "Single-logarithmic stability for the Calderón problem with local data", *J. Inverse Ill-Posed Probl.* **20**:4 (2012), 389–400. MR 2984474

[Ammari and Uhlmann 2004] H. Ammari and G. Uhlmann, "Reconstruction of the potential from partial Cauchy data for the Schrödinger equation", *Indiana Univ. Math. J.* **53**:1 (2004), 169–183. MR 2005h:35351 Zbl 1051.35103

[Astala and Päivärinta 2006] K. Astala and L. Päivärinta, "Calderón's inverse conductivity problem in the plane", *Ann. of Math.* (2) **163**:1 (2006), 265–299. MR 2007b:30019 Zbl 1111.35004

[Astala et al. 2005] K. Astala, L. Päivärinta, and M. Lassas, "Calderón's inverse problem for anisotropic conductivity in the plane", *Comm. Partial Differential Equations* **30**:1-3 (2005), 207–224. MR 2005k:35421 Zbl 1129.35483

[Bukhgeim 2008] A. L. Bukhgeim, "Recovering a potential from Cauchy data in the two-dimensional case", *J. Inverse Ill-Posed Probl.* **16**:1 (2008), 19–33. MR 2008m:30049 Zbl 1142.30018

[Bukhgeim and Uhlmann 2002] A. L. Bukhgeim and G. Uhlmann, "Recovering a potential from partial Cauchy data", *Comm. Partial Differential Equations* **27**:3-4 (2002), 653–668. MR 2003d:35262 Zbl 0998.35063

[Caro et al. 2009] P. Caro, P. Ola, and M. Salo, "Inverse boundary value problem for Maxwell equations with local data", *Comm. Partial Differential Equations* **34**:10-12 (2009), 1425–1464. MR 2010m:35558 Zbl 1185.35321

[Chung 2012] F. J.-H. Chung, *A partial data result for the magnetic Schrodinger inverse problem*, Ph.D. Thesis, The University of Chicago, 2012, Available at http://search.proquest.com/docview/1032548281. MR 3054984

[Chung et al. 2013] F. Chung, M. Salo, and L. Tzou, "Partial data inverse problems for the Hodge Laplacian", preprint, 2013. arXiv 1310.4616

[Eskin 2004] G. Eskin, "Inverse boundary value problems in domains with several obstacles", *Inverse Problems* **20**:5 (2004), 1497–1516. MR 2005h:35358 Zbl 1074.35086

[Ferreira 2007] D. dos Santos Ferreira, C. E. Kenig, J. Sjöstrand, and G. Uhlmann, "Determining a magnetic Schrödinger operator from partial Cauchy data", *Comm. Math. Phys.* **271**:2 (2007), 467–488. MR 2008a:35044 Zbl 1148.35096

[Ferreira 2009a] D. dos Santos Ferreira, C. E. Kenig, M. Salo, and G. Uhlmann, "Limiting Carleman weights and anisotropic inverse problems", *Invent. Math.* **178**:1 (2009), 119–171. MR 2010h:58033 Zbl 1181.35327

[Ferreira 2009b] D. dos Santos Ferreira, C. E. Kenig, J. Sjöstrand, and G. Uhlmann, "On the linearized local Calderón problem", *Math. Res. Lett.* **16**:6 (2009), 955–970. MR 2011d:35515 Zbl 1198.31003

[Ferreira 2013] D. dos Santos Ferreira, Y. Kurylev, M. Lassas, and M. Salo, "The Calderon problem in transversally anisotropic geometries", preprint, 2013. arXiv 1305.1273

[Florescu et al. 2011] L. Florescu, V. A. Markel, and J. C. Schotland, "Inversion formulas for the broken-ray Radon transform", *Inverse Problems* **27**:2 (2011), 025002, 13. MR 2011h:45019 Zbl 1211.65168

[Guillarmou and Tzou 2011a] C. Guillarmou and L. Tzou, "Calderón inverse problem with partial data on Riemann surfaces", *Duke Math. J.* **158**:1 (2011), 83–120. MR 2012f:35574 Zbl 1222.35212

[Guillarmou and Tzou 2011b] C. Guillarmou and L. Tzou, "Identification of a connection from Cauchy data on a Riemann surface with boundary", *Geom. Funct. Anal.* **21**:2 (2011), 393–418. MR 2012k:58027 Zbl 1260.58011

[Guillarmou and Tzou 2013] C. Guillarmou and L. Tzou, "The Calderón inverse problem in two dimensions", pp. 119–166 in *Inverse problems and applications: inside out II*, edited by G. Uhlmann, Cambridge University Press, 2013.

[Haberman and Tataru 2013] B. Haberman and D. Tataru, "Uniqueness in Calderón's problem with Lipschitz conductivities", *Duke Math. J.* **162**:3 (2013), 496–516. MR 3024091 Zbl 1260.35251

[Heck and Wang 2006] H. Heck and J.-N. Wang, "Stability estimates for the inverse boundary value problem by partial Cauchy data", *Inverse Problems* **22**:5 (2006), 1787–1796. MR 2007g:35270 Zbl 1106.35133

[Helgason 1999] S. Helgason, *The Radon transform*, 2nd ed., Progress in Mathematics **5**, Birkhäuser, Boston, MA, 1999. MR 2000m:44003 Zbl 0932.43011

[Hubenthal 2013a] M. Hubenthal, "The broken ray transform on the square", preprint, 2013. arXiv 1302.6193

[Hubenthal 2013b] M. Hubenthal, "The broken ray transform in n dimensions", preprint, 2013. arXiv 1310.7156

[Hyvönen et al. 2012] N. Hyvönen, P. Piiroinen, and O. Seiskari, "Point measurements for a Neumann-to-Dirichlet map and the Calderón problem in the plane", *SIAM J. Math. Anal.* **44**:5 (2012), 3526–3536. MR 3023421 Zbl 1257.35197

[Ilmavirta 2013a] J. Ilmavirta, "Broken ray tomography in the disc", *Inverse Problems* **29**:3 (2013), 035008, 17. MR 3040563 Zbl 06158616

[Ilmavirta 2013b] J. Ilmavirta, "A reflection approach to the broken ray transform", preprint, 2013. arXiv 1306.0341

[Ilmavirta 2013c] J. Ilmavirta, "Boundary reconstruction for the broken ray transform", preprint, 2013. arXiv 1310.2025

[Imanuvilov and Yamamoto 2012a] O. Y. Imanuvilov and M. Yamamoto, "Inverse boundary value problem for Schrödinger equation in two dimensions", *SIAM J. Math. Anal.* **44**:3 (2012), 1333–1339. MR 2982714 Zbl 1273.35316

[Imanuvilov and Yamamoto 2012b] O. Y. Imanuvilov and M. Yamamoto, "Inverse problem by Cauchy data on an arbitrary sub-boundary for systems of elliptic equations", *Inverse Problems* **28**:9 (2012), 095015, 30. MR 2972464 Zbl 1250.35184

[Imanuvilov and Yamamoto 2013a] O. Y. Imanuvilov and M. Yamamoto, "Inverse boundary value problem for the Schrödinger equation in a cylindrical domain by partial boundary data", *Inverse Problems* **29**:4 (2013), 045002, 8. MR 3041540 Zbl 1273.35317

[Imanuvilov and Yamamoto 2013b] O. Y. Imanuvilov and M. Yamamoto, "Uniqueness for Inverse Boundary Value Problems by Dirichlet-to-Neumann Map on Subboundaries", *Milan J. Math.* **81**:2 (2013), 187–258. MR 3129784 Zbl 06236315

[Imanuvilov et al. 2010] O. Y. Imanuvilov, G. Uhlmann, and M. Yamamoto, "The Calderón problem with partial data in two dimensions", *J. Amer. Math. Soc.* **23**:3 (2010), 655–691. MR 2012c:35472 Zbl 1201.35183

[Imanuvilov et al. 2011a] O. Y. Imanuvilov, G. Uhlmann, and M. Yamamoto, "Determination of second-order elliptic operators in two dimensions from partial Cauchy data", *Proc. Natl. Acad. Sci. USA* **108**:2 (2011), 467–472. MR 2012a:35364 Zbl 1256.35203

[Imanuvilov et al. 2011b] O. Y. Imanuvilov, G. Uhlmann, and M. Yamamoto, "Inverse boundary value problem by measuring Dirichlet data and Neumann data on disjoint sets", *Inverse Problems* **27**:8 (2011), 085007, 26. MR 2012c:78002 Zbl 1222.35213

[Isakov 2007] V. Isakov, "On uniqueness in the inverse conductivity problem with local data", *Inverse Probl. Imaging* **1**:1 (2007), 95–105. MR 2007m:35273 Zbl 1125.35113

[Katchalov et al. 2001] A. Katchalov, Y. Kurylev, and M. Lassas, *Inverse boundary spectral problems*, Monographs and Surveys in Pure and Applied Mathematics **123**, Chapman & Hall/CRC, Boca Raton, FL, 2001. MR 2003e:58045 Zbl 1037.35098

[Kenig and Salo 2013] C. Kenig and M. Salo, "Recent progress in the Calderón problem with partial data", preprint, 2013. arXiv 1302.4218

[Kenig et al. 2007] C. E. Kenig, J. Sjöstrand, and G. Uhlmann, "The Calderón problem with partial data", *Ann. of Math.* (2) **165**:2 (2007), 567–591. MR 2008k:35498 Zbl 1127.35079

[Kenig et al. 2011a] C. E. Kenig, M. Salo, and G. Uhlmann, "Inverse problems for the anisotropic Maxwell equations", *Duke Math. J.* **157**:2 (2011), 369–419. MR 2012d:35408 Zbl 1226.35086

[Kenig et al. 2011b] C. E. Kenig, M. Salo, and G. Uhlmann, "Reconstructions from boundary measurements on admissible manifolds", *Inverse Probl. Imaging* **5**:4 (2011), 859–877. MR 2012k:58038 Zbl 1237.35160

[Knudsen 2006] K. Knudsen, "The Calderón problem with partial data for less smooth conductivities", *Comm. Partial Differential Equations* **31**:1-3 (2006), 57–71. MR 2006k:35303 Zbl 1091.35116

[Knudsen and Salo 2007] K. Knudsen and M. Salo, "Determining nonsmooth first order terms from partial boundary measurements", *Inverse Probl. Imaging* **1**:2 (2007), 349–369. MR 2008k:35500 Zbl 1122.35152

[Kohn and Vogelius 1984] R. Kohn and M. Vogelius, "Determining conductivity by boundary measurements", *Comm. Pure Appl. Math.* **37**:3 (1984), 289–298. MR 85f:80008 Zbl 0595.35092

[Krishnan 2009] V. P. Krishnan, "A support theorem for the geodesic ray transform on functions", *J. Fourier Anal. Appl.* **15**:4 (2009), 515–520. MR 2010i:53146 Zbl 1186.53089

[Krupchyk et al. 2012] K. Krupchyk, M. Lassas, and G. Uhlmann, "Inverse problems with partial data for a magnetic Schrödinger operator in an infinite slab and on a bounded domain", *Comm. Math. Phys.* **312**:1 (2012), 87–126. MR 2914058 Zbl 1238.35188

[Lassas and Oksanen 2010] M. Lassas and L. Oksanen, "An inverse problem for a wave equation with sources and observations on disjoint sets", *Inverse Problems* **26**:8 (2010), 085012, 19. MR 2011m:35422 Zbl 1197.35323

[Lassas and Oksanen 2012] M. Lassas and L. Oksanen, "Inverse problem for the Riemannian wave equation with Dirichlet data and Neumann data on disjoint sets", preprint, 2012. arXiv 1208.2105

[Li and Uhlmann 2010] X. Li and G. Uhlmann, "Inverse problems with partial data in a slab", *Inverse Probl. Imaging* **4**:3 (2010), 449–462. MR 2011f:35361 Zbl 1200.35331

[Lozev 2013] K. Lozev, "Mathematical modeling and solution of the tomography problem in domains with reflecting obstacles", *J. Phys.: Conf. Ser.* **410** (2013).

[Nachman 1996] A. I. Nachman, "Global uniqueness for a two-dimensional inverse boundary value problem", *Ann. of Math.* (2) **143**:1 (1996), 71–96. MR 96k:35189 Zbl 0857.35135

[Nachman and Street 2010] A. Nachman and B. Street, "Reconstruction in the Calderón problem with partial data", *Comm. Partial Differential Equations* **35**:2 (2010), 375–390. MR 2012b:35368 Zbl 1186.35242

[Salo and Tzou 2010] M. Salo and L. Tzou, "Inverse problems with partial data for a Dirac system: a Carleman estimate approach", *Adv. Math.* **225**:1 (2010), 487–513. MR 2011g:35432 Zbl 1197.35329

[Sharafutdinov 1994] V. A. Sharafutdinov, *Integral geometry of tensor fields*, VSP, Utrecht, 1994. MR 97h:53077

[Sylvester and Uhlmann 1987] J. Sylvester and G. Uhlmann, "A global uniqueness theorem for an inverse boundary value problem", *Ann. of Math.* (2) **125**:1 (1987), 153–169. MR 88b:35205 Zbl 0625.35078

[Sylvester and Uhlmann 1988] J. Sylvester and G. Uhlmann, "Inverse boundary value problems at the boundary—continuous dependence", *Comm. Pure Appl. Math.* **41**:2 (1988), 197–219. MR 89f:35213 Zbl 0632.35074

[Uhlmann and Vasy 2012] G. Uhlmann and A. Vasy, "The inverse problem for the local geodesic ray transform", preprint, 2012. arXiv 1210.2084

[Zhang 2012] G. Zhang, "Uniqueness in the Calderón problem with partial data for less smooth conductivities", *Inverse Problems* **28**:10 (2012), 105008, 18. MR 2987911 Zbl 1256.35207

[Zworski 2012] M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics **138**, American Mathematical Society, Providence, RI, 2012. MR 2952218 Zbl 1252.58001

Received 6 May 2013. Accepted 13 Nov 2013.

CARLOS KENIG: cek@math.uchicago.edu
Department of Mathematics, University of Chicago, 5734 S. University Avenue, Chicago, Illinois 60637, United States

MIKKO SALO: mikko.j.salo@jyu.fi
Department of Mathematics and Statistics, University of Jyväskylä, FI-40014 Jyväskylä, Finland

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 6 No. 8 2013

2157-5045(2013)6:8;1-A