

# OPTIMAL CONSTANTS IN NONTRAPPING RESOLVENT ESTIMATES AND APPLICATIONS IN NUMERICAL ANALYSIS

JEFFREY GALKOWSKI, EUAN A. SPENCE, AND JARED WUNSCH

ABSTRACT. We study the resolvent for nontrapping obstacles on manifolds with Euclidean ends. It is well known that for such manifolds, the outgoing resolvent satisfies  $\|\chi R(k)\chi\|_{L^2 \rightarrow L^2} \leq Ck^{-1}$  for  $k > 1$ , but the constant  $C$  has been little studied. We show that, for high frequencies, the constant is bounded above by  $2/\pi$  times the length of the longest generalized bicharacteristic of  $|\xi|_g^2 - 1$  remaining in the support of  $\chi$ . We show that this estimate is optimal in the case of manifolds without boundary. We then explore the implications of this result for the numerical analysis of the Helmholtz equation.

## 1. INTRODUCTION

Let  $(M, g)$  be a manifold with Euclidean ends and  $\Omega \Subset M$  an obstacle with smooth boundary. Assume that all Melrose-Sjöstrand generalized bicharacteristics (i.e., geodesics) escape to infinity. Let  $\Delta_{\Omega, g}$  be the Dirichlet realization of the Laplacian on  $M \setminus \Omega$ . It is well known that for any  $\chi \in C_c^\infty(M)$ , there exists a constant  $C > 0$  and  $k_0 > 0$  so that

$$\|\chi(-\Delta_{\Omega, g} - k^2 - i0)^{-1}\chi\|_{L^2(M \setminus \Omega) \rightarrow L^2(M \setminus \Omega)} \leq C|k|^{-1}, \quad k > k_0.$$

In this paper we study how the constant  $C > 0$  depends on the classical dynamics on  $((M, g), \Omega)$ .

Suppose that there are no geodesics tangent to  $\partial\Omega$  to infinite order and let  $\varphi_t : S^*M \rightarrow S^*M$  denote the Melrose-Sjöstrand generalized bicharacteristic flow [29, Section 24.3]. Next, fix  $R_1 > 0$  so that  $\Omega \subset B(0, R_1)$  and  $(M, g)$  is Euclidean outside  $B(0, R_1)$ . Then define for any  $R \geq R_1$ ,

$$(1.1) \quad L(g, \Omega, R) := \inf \{t > 0 \mid \varphi_t(S_{B(0, R)}^*M) \cap S_{B(0, R)}^*M = \emptyset\}.$$

(We will omit the  $\Omega$  from the notation when  $\Omega = \emptyset$ .)

In the statement of the following estimates, we will use a family of Sobolev spaces with appropriate semiclassical scaling, using the global definition for  $s \in \mathbb{R}$ ,

$$(1.2) \quad \|u\|_{H^s(M \setminus \Omega)}^2 = \langle (-\Delta_{\Omega, g} + k^2)^s u, u \rangle.$$

**Theorem 1.** *Let  $(M, g)$  be a manifold with Euclidean ends with  $g \in C^{1,1}$ . Suppose that  $\Omega \Subset M$  has smooth boundary,  $g$  is  $C^\infty$  near  $\Omega$ , and  $\Omega$  is nowhere tangent to the geodesic flow to infinite order. Assume the generalized geodesic flow on  $M \setminus \Omega$  is nontrapping. Then for every  $R > R_1$ ,  $\chi \in C_c^\infty(B(0, R); [0, 1])$ , there exists  $k_0 > 0$  so that for  $k > k_0$ ,*

$$(1.3) \quad \|\chi(-\Delta_{\Omega, g} - k^2 - i0)^{-1}\chi\|_{L^2(M \setminus \Omega) \rightarrow L^2(M \setminus \Omega)} \leq \frac{2L(g, \Omega, R)}{\pi k}.$$

More generally, for  $0 \leq s \leq 2$ ,

$$(1.4) \quad \|\chi(-\Delta_{\Omega, g} - k^2 - i0)^{-1}\chi\|_{L^2(M \setminus \Omega) \rightarrow H^s(M \setminus \Omega)} \leq \frac{2^{\frac{s}{2}+1}L(g, \Omega, R)}{\pi} k^{s-1}.$$

---

Date: October 30, 2019.

The constant  $k_0$  may be chosen uniformly as  $g$  varies within a sufficiently small open neighborhood in  $\mathcal{C}^{2,\alpha}$  ( $\alpha > 0$ ) of a given nontrapping metric in  $\mathcal{C}^{2,\alpha}$  (where all metrics are taken smooth near  $\partial\Omega$ ).

Conversely, when  $\Omega = \emptyset$  and  $g \in \mathcal{C}^\infty$ , for every  $R' > R_1$  with  $B(0, R') \subset \{\chi \equiv 1\}$ , there is  $k_0 > 0$  so that for  $k > k_0$ ,

$$(1.5) \quad \|\chi(-\Delta_g - k^2 - i0)^{-1}\chi\|_{L^2(M \setminus \Omega) \rightarrow L^2(M \setminus \Omega)} \geq \frac{2L(g, R')}{\pi k}.$$

Notice that in the interior of  $M \setminus \Omega$ , the Melrose-Sjöstrand flow is equal to the geodesic flow except that, on the unit cotangent bundle, the speed of the flow is 2 rather than 1. Thus, in the case where  $\Omega = \emptyset$ , the theorem states that the growth of the resolvent as a map on  $L^2$  as  $k \rightarrow \infty$  is controlled above and below by  $1/\pi$  times the length of the longest geodesic contained entirely in  $B(0, R_1)$ .

**1.1. More general operators.** In §6.2 below, we recall that several important physical applications of the Helmholtz equation involve an operator that differs slightly from  $-\Delta_{\Omega, g}$ , namely the divergence form operator  $-\sum \partial_i g^{ij} \partial_j$  which is self-adjoint with respect to the Euclidean volume form on  $\mathbb{R}^n$ . In order to deal with this and similar operators, we will prove a slightly more general result.

**Theorem 2.** *Let  $(M, g)$  be a manifold with Euclidean ends,  $g \in \mathcal{C}^{1,1}$ . Suppose that  $\Omega \subset M$  has smooth boundary,  $g$  is  $\mathcal{C}^\infty$  in a neighborhood of  $\Omega$ , and  $\Omega$  is nowhere tangent to the geodesic flow to infinite order. Assume the generalized geodesic flow on  $M \setminus \Omega$  is nontrapping. Suppose that  $P(g) \in \text{Diff}^2(M)$  so that*

$$(1.6) \quad P(g) + \Delta_g = \sum_j L_j \partial_{x_j} + L$$

where  $L_j, L \in \mathcal{C}_c^{0,\alpha}(B(0, R_1))$ , for some  $\alpha > 0$ . Suppose further that  $\nu$  is a density on  $M$  so that  $P_\Omega(g)$  is self adjoint with respect to  $L^2(M \setminus \Omega; \nu)$  where  $P_\Omega(g)$  is the Dirichlet realization of  $P(g)$ . Let

$$(1.7) \quad \|u\|_{H_\nu^s(M \setminus \Omega)}^2 := \langle (P_\Omega(g) + k^2)^s u, u \rangle_{L^2(M \setminus \Omega; \nu)}.$$

Then for every  $R > R_1$ ,  $\chi \in \mathcal{C}_c^\infty(B(0, R); [0, 1])$ , there exists  $k_0 > 0$  so that for  $k > k_0$  and  $0 \leq s \leq 2$ ,

$$(1.8) \quad \|\chi(P_\Omega(g) - k^2 - i0)^{-1}\chi\|_{L_\nu^2(M \setminus \Omega) \rightarrow H_\nu^s(M \setminus \Omega)} \leq \frac{2^{\frac{s}{2}+1} L(g, \Omega, R)}{\pi} k^{s-1}.$$

The constant  $k_0$  may be chosen uniformly as  $g$  varies within sufficiently small open neighborhoods of in  $\mathcal{C}^{2,\alpha}$  ( $\alpha > 0$ ) of a given nontrapping metric in  $\mathcal{C}^{2,\alpha}$  and  $L_j, L$  vary in small neighborhoods in  $\mathcal{C}^{0,\alpha}$  (where all  $g, L_j, L$  are taken smooth near  $\partial\Omega$  and the subset is assumed to be contained in a small open neighborhood in  $\mathcal{C}^N$  for some sufficiently large  $N$  near  $\partial\Omega$ ).

**1.2. Motivation from, and applications to, numerical analysis.** In the last few years, there has been growing interest from the numerical-analysis community in proving bounds on solutions of the Helmholtz equation where the constants are explicit in the metric (i.e, the coefficients); see [7, 17, 4, 40, 43, 37, 26, 27]. Almost all of these previously-obtained bounds used variants of the Morawetz commutator  $x \cdot \nabla$ , and thus are restricted to star-shaped domains and certain classes of coefficients (although this class includes discontinuous coefficients; see, e.g., [26]); the exception are the 1-d bounds in [43], which use the fact that the solution of the Helmholtz equation with piecewise-constant coefficients in 1-d can be expressed in terms of the solution of a linear system of algebraic equations.

This interest from the numerical-analysis community is because the analysis of *any* numerical method for solving the Helmholtz equation with variable coefficients requires a resolvent estimate, and if the constant in the resolvent estimate is not given explicitly in terms of the coefficients, then the numerical analysis will not be explicit in the coefficients. For example, having proved a bound explicit in the coefficients, [17, 27] were then concerned with analyzing standard finite-element methods applied to Helmholtz problems with variable coefficients, and [7, 17, 4, 40] were concerned with designing and analyzing methods tailored to the coefficients.

The resolvent estimate (1.8) will therefore be a fundamental ingredient in the numerical analysis of variable-coefficient Helmholtz problems in nontrapping scenarios. In §6 we illustrate this fact by proving an error estimate for the finite-element method applied to the variable-coefficient Helmholtz equation posed in the exterior of a nontrapping Dirichlet obstacle, with this estimate explicit in  $k$ , the coefficients, and the parameters of the discretization; see Theorem 3 below. The key point is that Theorem 3 shows how the condition on the discretization for the error estimate to hold depends on the length of the longest ray, showing that this condition becomes more restrictive as the length of the longest ray grows.

In §6 we also briefly outline the implications of the estimate (1.8) for (i) preconditioning finite-element discretizations of the Helmholtz equation (see Remark 6.8), and (ii) “uncertainty quantification” of the Helmholtz equation (see Remark 6.9).

**Acknowledgements.** The first author was supported by NSF Postdoctoral Research Fellowship DMS-1502661 and thanks Maciej Zworski for helpful conversations. The second author was supported by EPSRC grant EP/R005591/1. The third author was partially supported by NSF grant DMS-1600023. The authors are grateful to an anonymous referee for helpful comments on the manuscript.

## 2. MANIFOLDS WITH EUCLIDEAN ENDS

We now define the notion of a manifold with Euclidean ends. Note that the canonical example of a manifold with Euclidean ends is the space  $\mathbb{R}^n$  with a metric  $g$  so that  $I - g$  has compact support. In order to allow more general topologies, we define the general notion of a manifold with Euclidean ends.

*Definition 2.1.*  $(M, g)$  is an  $n$ -dimensional *manifold with Euclidean ends* if it is a non-compact, complete Riemannian manifold such that

- there exists a function  $r \in C^\infty(M; \mathbb{R})$  such that the sets  $\{r \leq c\}$  are compact for all  $c$ , and
- there exists  $R_1 > 0$  such that  $\{r \geq R_1\}$  is the disjoint union of finitely many components, each of which is isometric to  $\mathbb{R}^n \setminus B(0, R_1)$  with the Euclidean metric, and the pullback of  $r$  under the isometry is the Euclidean norm.

The connected components of  $\{r \geq R_1\}$  are called the *infinite ends* of  $M$ . The notation  $B(0, R)$  is defined for  $R \geq R_1$  and has the following meaning:

$$B(0, R) := \{r < R\}.$$

**2.1. The outgoing resolvent on manifolds with Euclidean ends.** We now review (following the treatment in [20, Section 4.2]) some properties of the outgoing resolvent on a manifold with Euclidean ends. Let  $E_i$ ,  $i = 1, \dots, m$  be the infinite ends of  $M$ . and let  $R_0(k)$  denote the free resolvent on  $\mathbb{R}^n$ . That is  $R_0(k) : L^2_{\text{comp}}(\mathbb{R}^n) \rightarrow L^2_{\text{loc}}(\mathbb{R}^n)$  is the meromorphic continuation of  $(-\Delta - k^2)^{-1}$  from the half plane  $\text{Im } k > 0$ . Define  $\tilde{R}_0(k) : L^2_{\text{comp}}(M) \rightarrow L^2_{\text{loc}}(M)$  by

$$(2.1) \quad \tilde{R}_0(k)f := \sum_{i=1}^m 1_{E_i} R_0(k) 1_{E_i} f.$$

Next, let  $\chi_i \in C_c^\infty(M; [0, 1])$ ,  $i = 0, \dots, 3$  so that

$$(2.2) \quad \chi_i \equiv 1 \text{ on } \text{supp } \chi_{i-1}, \quad \text{supp}(1 - \chi_i) \subset M \setminus \overline{B(0, R_1)}, \quad \text{supp } \chi_i \subset B(0, R)$$

for some  $R > R_1$ . Then for  $k_0 \in \mathbb{C}$  with  $\text{Im } k_0 \gg 1$ , let

$$Q(k, k_0) := Q_0(k) + Q_1(k_0) \\ Q_0(k) := (1 - \chi_0)\tilde{R}_0(k)(1 - \chi_1), \quad Q_1(k) := \chi_2(-\Delta_g - k_0^2)^{-1}\chi_1.$$

following the proof in [20, Section 4.2], we can write

$$(2.3) \quad R(k) := (P_\Omega(g) - k^2)^{-1} = Q(k, k_0)(I + K(k, k_0)\chi_3)^{-1}(I - K(k, k_0)(1 - \chi_3)).$$

where  $K(k, k_0) : L^2(M) \rightarrow L_{\text{comp}}^2(M)$  and in particular,  $(1 - \chi_3)K(k, k_0) = 0$ . This implies that  $(1 - \chi_3)R(k)$  lies in the image of  $\tilde{R}_0(k)$ .

### 3. THE CASE OF A MANIFOLD WITHOUT BOUNDARY

To illustrate the methods we begin by proving (1.3) in the case of a manifold without boundary. The idea of the proof is identical when the boundary is non-empty, but the proofs of propagation statements are more involved. Thus we take  $\partial M = \emptyset$  throughout this section.

**3.1. Defect measures.** We will argue by contradiction. Let  $h = k^{-1}$  and  $P(h) := h^2 P_\Omega(g) - 1$ . We also write  $L^2$  for  $L^2_\nu(M \setminus \Omega)$  and  $H^s$  for  $H^s_\nu(M \setminus \Omega)$ .

*Remark 3.1.* We will sometimes write  $\|\cdot\|_{H^s_h}$  for  $h^s\|\cdot\|_{H^s}$ ; since  $\|\cdot\|_{H^s}$  has a semiclassical scaling built into it, this means that

$$\|u\|_{H^s_h}^2 = \langle (h^2 \Delta_{\Omega, g} + 1)^s u, u \rangle.$$

Note that if  $\chi_0, \chi_1 \in C_c^\infty(M; [0, 1])$  and  $\chi_1 \equiv 1$  on  $\text{supp } \chi_0$ , then

$$\frac{\|\chi_0(P - i0)^{-1}\chi_0 f\|_{L^2}}{\|f\|_{L^2}} = \frac{\|\chi_0\chi_1(P - i0)^{-1}\chi_1\chi_0 f\|_{L^2}}{\|f\|_{L^2}} \leq \frac{\|\chi_0\chi_1(P - i0)^{-1}\chi_1\chi_0 f\|_{L^2}}{\|\chi_0 f\|_{L^2}} \\ \leq \frac{\|\chi_1(P - i0)^{-1}\chi_1\chi_0 f\|_{L^2}}{\|\chi_0 f\|_{L^2}} \leq \|\chi_1(P - i0)^{-1}\chi_1\|_{L^2 \rightarrow L^2}.$$

In particular,

$$\|\chi_0(P - i0)^{-1}\chi_0\|_{L^2 \rightarrow L^2} \leq \|\chi_1(P - i0)^{-1}\chi_1\|_{L^2 \rightarrow L^2}.$$

and, since  $R > R_1$ , we may assume without loss of generality that  $\chi \equiv 1$  on  $B(0, R_1)$ .

If (1.3) fails, there exists a sequence of discrete values of  $h = h_j \downarrow 0$  and a sequence  $0 \neq f(h) \in L^2$  such that

$$\|\chi(P - i0)^{-1}\chi h f(h)\|_{L^2} = 1$$

and

$$\lim_{h \rightarrow 0} \frac{\|\chi(P - i0)^{-1}\chi h f(h)\|_{L^2}}{\|f(h)\|_{L^2}} = M \geq 2L(g, R).$$

(note that we allow  $M = \infty$ .) There exists  $R' < R$  such that we still have  $\text{supp } \chi \subset B(0, R')$ , hence we have the strict inequality  $M > L(g, R')$ . Let

$$u(h) = (P - i0)^{-1}\chi h f \in L^2_{\text{loc}}.$$

so that  $\|\chi u\|_{L^2} = 1$ .

**Lemma 3.2.** *For all  $\tilde{\chi} \in C_c^\infty(M)$ , there exists  $C > 0$  so that  $\|\tilde{\chi} u\|_{L^2} \leq C$ .*

*Proof.* Let  $\chi_i \in \mathcal{C}_c^\infty(M; [0, 1])$ ,  $i = 0, 1$  so that  $\chi \equiv 1$  on  $\text{supp } \chi_1$  and  $\chi_1 \equiv 1$  on  $\text{supp } \chi_0$ , and  $\chi_0 \equiv 1$  on  $B(0, R_1)$ . Then

$$(1 - \chi_0)P(1 - \chi_1) = \sum_i (1 - \chi_0)(-h^2\Delta_{\mathbb{R}^n} - 1)1_{E_i}(1 - \chi_1)$$

where  $E_i$  denotes the  $i^{\text{th}}$  Euclidean end of  $M$ .

Then,

$$\begin{aligned} P(1 - \chi_1)u &= \sum_i (1 - \chi_0)1_{E_i}(-h^2\Delta_{\mathbb{R}^n} - 1)1_{E_i}(1 - \chi_1)u \\ &= (1 - \chi_1)\chi hf - \sum_i (1 - \chi_0)[-h^2 1_{E_i}\Delta_{\mathbb{R}^n} 1_{E_i}, \chi_1]u \\ &= (1 - \chi_1)\chi hf - \sum_i [-h^2 1_{E_i}\Delta_{\mathbb{R}^n} 1_{E_i}, \chi_1]u. \end{aligned}$$

Next extend  $u1_{E_i}$  by 0 to a function  $v_i$  on  $\mathbb{R}^n$  so that  $v_i \equiv u1_{E_i}$  on  $\mathbb{R}^n \setminus B(0, R_1)$ . Then

$$(-h^2\Delta_{\mathbb{R}^n} - 1)(1 - \chi_1)1_{E_i}v_i = 1_{E_i}(1 - \chi_1)\chi hf - [h^2\Delta_{\mathbb{R}^n}, 1_{E_i}\chi_1]v_i.$$

Since by (2.3)  $v_i$  is  $h^{-1}$  outgoing,

$$(1 - \chi_1)1_{E_i}v_i = h^{-2}R_0(h^{-1})(1_{E_i}(1 - \chi_1)\chi hf - [h^2\Delta_{\mathbb{R}^n}, 1_{E_i}\chi_1]v_i).$$

In particular, since  $\chi \equiv 1$  on  $\text{supp } \chi_1$ ,

$$\|(1 - \chi_1)1_{E_i}v_i\|_{L^2} \leq C(\|f\|_{L^2} + \|[h^2\Delta_{\mathbb{R}^n}, 1_{E_i}\chi_1]v_i\|_{H^{-1}}) \leq C(\|f\|_{L^2} + \|\chi u\|_{L^2}).$$

Now,

$$(1 - \chi_1)u = (1 - \chi_1)\sum_i 1_{E_i}v_i.$$

Therefore,

$$\|(1 - \chi_1)u\|_{L^2} \leq C(\|f\|_{L^2} + \|\chi u\|_{L^2}).$$

Let  $\tilde{\chi} \in \mathcal{C}_c^\infty(M; [0, 1])$ . Then using again that  $\chi \equiv 1$  on  $\text{supp } \chi_1$ ,

$$\begin{aligned} \limsup_{h \rightarrow 0} \|\tilde{\chi}u\|_{L^2} &\leq \limsup_{h \rightarrow 0} (\|\tilde{\chi}(1 - \chi)u\|_{L^2} + \|\tilde{\chi}\chi u\|_{L^2}) \\ &\leq \limsup_{h \rightarrow 0} (\|\tilde{\chi}(1 - \chi_1)u\|_{L^2} + \|\chi u\|_{L^2}) \\ &\leq C \limsup_{h \rightarrow 0} (\|f\|_{L^2} + \|\chi u\|_{L^2}) \leq C\left(\frac{1}{2L(g, R')} + 1\right) \end{aligned}$$

completing the proof of the lemma.  $\square$

By Lemma 3.2,  $u$  is uniformly bounded in  $L_{\text{loc}}^2$  and, taking subsequences, we may assume that  $u$  has defect measure  $\mu$ ,  $\chi f$  has defect measure  $\alpha$ , and  $u$  and  $\chi f$  have joint defect measure  $\mu^j$ ; in other words, for  $h = h_j$  in this chosen subsequence, and for every  $a \in \mathcal{C}_c^\infty(T^*M)$ ,

$$(3.1) \quad \begin{aligned} \lim_{h \downarrow 0} \langle a(x, hD)u, u \rangle &= \int a d\mu \\ \lim_{h \downarrow 0} \langle a(x, hD)\chi f, \chi f \rangle &= \int a d\alpha \\ \lim_{h \downarrow 0} \langle a(x, hD)\chi f, u \rangle &= \int a d\mu^j. \end{aligned}$$

Note that we use a quantization procedure that sends symbols with compact support in  $x$  to operators with compactly supported kernel.

The Cauchy-Schwarz inequality gives us a simple inequality satisfied by these three measures. We use the notation

$$\mu(a) \equiv \int a \, d\mu$$

for the pairing of a function and a measure.

**Lemma 3.3.** *For any  $a \in \mathcal{C}_c^0(T^*M; \mathbb{R})$ .*

$$|\mu^j(a)| \leq \sqrt{\mu(|a|)} \sqrt{\alpha(|a|)}.$$

*Proof.* If  $b = \sqrt{a}$  is in  $\mathcal{C}_c^\infty(T^*M)$ ,

$$\begin{aligned} \langle a(x, hD)\chi f, u \rangle &= \langle b^2(x, hD)\chi f, u \rangle \\ &= \langle b(x, hD)\chi f, b(x, hD)u \rangle + O(h) \\ &\leq \sqrt{\langle b(x, hD)\chi f, b(x, hD)\chi f \rangle} \sqrt{\langle b(x, hD)u, b(x, hD)u \rangle} + O(h) \\ &\leq \left( \int b^2 \, d\alpha \right)^{1/2} \left( \int b^2 \, d\mu \right)^{1/2} + o(1), \end{aligned}$$

which proves the desired inequality by letting  $h \rightarrow 0$ .

Next, let  $a \in \mathcal{C}_c^\infty(T^*M)$  with  $a \geq 0$ ,  $\psi_i \in \mathcal{C}_c^\infty(T^*M; [0, 1])$ ,  $i = 0, 1$  have  $\psi_i \equiv 1$  in a neighborhood of  $\text{supp } a$  and  $\psi_1 \equiv 1$  in a neighborhood of  $\text{supp } \psi_0$ . Then we apply the previous result to  $a_\epsilon = \psi_0^2(a + \epsilon\psi_1)$  and let  $\epsilon \downarrow 0$  to see that  $|\mu^j(a)| \leq \sqrt{\mu(a)} \sqrt{\alpha(a)}$ .

Now, letting  $0 \leq a \in \mathcal{C}_c^0(T^*M)$ ,  $0 \leq a_n \in \mathcal{C}_c^\infty(T^*M)$  with  $a_n \rightarrow a$  uniformly. Then we have

$$\mu^j(a_n) \rightarrow \mu^j(a), \quad \mu(a_n) \rightarrow \mu(a), \quad \alpha(a_n) \rightarrow \alpha(a)$$

In particular,

$$|\mu^j(a)| \leq \sqrt{\mu(a)} \sqrt{\alpha(a)}, \quad 0 \leq a \in \mathcal{C}_c^0(T^*M).$$

Next, for  $a \in \mathcal{C}_c^0(T^*M; \mathbb{R})$ , write  $a = a_+ - a_-$  with  $0 \leq a_\pm \in \mathcal{C}_c^0(T^*M)$ . Then,

$$\begin{aligned} |\mu^j(a)| &\leq |\mu^j(a_+)| + |\mu^j(a_-)| \leq \sqrt{\mu(a_+)} \sqrt{\alpha(a_+)} + \sqrt{\mu(a_-)} \sqrt{\alpha(a_-)} \\ &\leq \sqrt{\mu(|a|)} \sqrt{\alpha(|a|)}. \end{aligned}$$

□

Since  $Pu = h\chi f$ , for  $a \in \mathcal{C}_c^\infty(T^*M; \mathbb{R})$ ,

$$(3.2) \quad \begin{aligned} ih^{-1} \langle [P, a(x, hD)]u, u \rangle &= ih^{-1} (\langle a(x, hD)u, Pu \rangle - \langle a(x, hD)Pu, u \rangle) \\ &= 2 \text{Im} \langle a(x, hD)\chi f, u \rangle. \end{aligned}$$

Sending  $h \rightarrow 0$  yields

$$(3.3) \quad \mu(H_p a) = 2 \text{Im} \mu^j(a).$$

For the proof of the following standard result, we use the formula (2.3) relating  $R(k) := (P_\Omega(g) - k^2 - i0)^{-1}$  to the free outgoing resolvent on  $\mathbb{R}^n$  and refer the reader to, e.g., [9, Proposition 3.5].

**Lemma 3.4.** *Let*

$$\mathcal{I} := \left\{ \rho \in S^*M \mid \bigcup_{t \geq 0} \varphi_{-t}(\rho) \cap \text{supp } \chi = \emptyset \right\}$$

*be the directly incoming set. Then  $\mu(\mathcal{I}) = 0$ .*

Next, we show that  $\mu$  is supported on the characteristic variety and that  $u$  is oscillating at frequency roughly  $h^{-1}$ .

**Lemma 3.5.** *The measure  $\mu$  is supported on  $S^*M$ . In addition, for  $b \in S^1$ ,*

$$\lim_{h \rightarrow 0} \|b(x, hD)\chi u\|_{L^2}^2 = \mu(|b|^2\chi^2).$$

*Proof.* Suppose that  $a \in \mathcal{C}_c^\infty(T^*M)$  with  $a \equiv 0$  in a neighborhood of  $|\xi|_g = 1$ . Then there exists  $E \in \Psi^{-2}$  compactly supported so that

$$a(x, hD) = EP + O_{L_{\text{loc}}^2 \rightarrow L_{\text{comp}}^2}(h^\infty).$$

Therefore,

$$\langle a(x, hD)u, u \rangle = \langle Eh\chi f, u \rangle + O(h^\infty) \rightarrow 0.$$

Hence,  $\mu(a) = 0$ . In particular,  $\text{supp } \mu \subset S^*M$ .

Fix  $\psi \in \mathcal{C}_c^\infty(T^*M)$  with  $\psi \equiv 1$  on  $\text{supp } \chi \cap \{|\xi|_g^2 \leq 2\}$ . Then, there exists  $E \in \Psi^0$  compactly supported so that

$$\chi(x)b(x, hD)^*b(x, hD)\chi(x)(1 - \psi(x, hD)) = EP + O_{L_{\text{loc}}^2 \rightarrow L_{\text{comp}}^2}(h^\infty).$$

In particular,

$$\begin{aligned} \|b(x, hD)\chi u\|_{L^2}^2 &= \langle \chi(x)b(x, hD)^*b(x, hD)\chi(x)u, u \rangle \\ &= \langle \chi(x)b(x, hD)^*b(x, hD)\chi(x)\psi(x, hD)u, u \rangle \\ &\quad + \langle \chi(x)b(x, hD)^*b(x, hD)\chi(x)(1 - \psi(x, hD))u, u \rangle \\ &= \langle \chi(x)b(x, hD)^*b(x, hD)\chi(x)\psi(x, hD)u, u \rangle + \langle hE\chi f, u \rangle + O(h^\infty) \\ &\rightarrow \mu(|b|^2\chi^2\psi) = \mu(|b|^2\chi^2). \end{aligned}$$

□

**3.2. Hölder continuous metrics.** We now make the necessary adjustments to allow the metric  $g$  to be Hölder continuous. We refer the reader to [47, Chapter 3, Section 11] for an analogous account of propagation of singularities for the wave equation as well as a review of the history of propagation of singularities theorems for operators with rough coefficients. Stronger results (i.e., with weaker regularity hypotheses) are probably possible in line with the work of Burq–Zuily [11, Remark 3.3], but low regularity is not our main focus here.

**Lemma 3.6.** *Let  $g_0, L_0, L_{j,0}$  satisfy the hypotheses of Theorem 2. Suppose that  $g(h) \in \mathcal{C}^{1,\alpha}$  and  $L(h), L_j(h) \in \mathcal{C}^{0,\alpha}$  for some  $\alpha > 0$  satisfy*

$$\begin{aligned} \lim_{h \rightarrow 0} \|g(h) - g_0\|_{\mathcal{C}^{1,\alpha}} &= 0, \\ \lim_{h \rightarrow 0} \|L(h) - L_0\|_{\mathcal{C}^{0,\alpha}} &= 0, \\ \lim_{h \rightarrow 0} \|L_j(h) - L_{j,0}\|_{\mathcal{C}^{0,\alpha}} &= 0, \quad j = 1, \dots, n. \end{aligned}$$

*Then the measure  $\mu$  is supported in  $S^*M$ , for  $b \in S^1$ ,*

$$(3.4) \quad \mu(|b|^2\chi^2) = \lim_{h \rightarrow 0} \|Op_h(b)\chi u\|_{L^2}^2$$

*and for  $a \in \mathcal{C}_c^\infty(T^*M)$ ,*

$$(3.5) \quad \mu(H_p a) = 2 \text{Im } \mu^j(a).$$

*where  $p = |\xi|_{g_0}^2 - 1$ .*

This lemma (together with the results of Section 3.3 below) suffices to prove our estimate provided the bicharacteristic flow is unique. Note that the lemma only requires the hypothesis that  $g \in \mathcal{C}^{1,\alpha}$  for  $\alpha > 0$  and that this regularity suffices to prove existence of solutions to Hamilton's equations. However, it is not adequate for proving uniqueness. Hence Theorem 1

is only stated for  $\alpha = 1$ , which is sufficient to guarantee the existence of a well-defined single-valued bicharacteristic flow. (A quantitative result using the dynamics of the multi-valued flow for  $\alpha < 1$  would be an interesting direction for further research.)

To prove Lemma 3.6, we will need a more general set of symbol classes. Let  $r \in \mathbb{N}$ ,  $0 < \alpha < 1$ ,  $0 \leq \rho < 1$ . We say that  $p(x, \xi) \in \mathcal{C}^{r, \alpha} S_\rho^m$  if

$$\|D_\xi^\beta p(\cdot, \xi)\|_{\mathcal{C}^{r, \alpha}} \leq C_\beta h^{-r\rho} \langle \xi \rangle^{m-|\beta|}.$$

We say  $p \in S_\rho^m$  if  $p \in \bigcap_r \mathcal{C}^{r, \alpha} S_\rho^m$ . We need the following boundedness property for operators with symbol in  $\mathcal{C}^{r, \alpha} S^m$ .

**Lemma 3.7.** *For  $r \geq 0$  and  $\alpha > 0$ , the map  $Op_h : \mathcal{C}^{r, \alpha} S^m \rightarrow \mathcal{L}(H_h^m, L^2)$  with  $p \mapsto p(x, hD)$  is continuous with norm bounded independently of  $h$ .*

*Proof.* It is enough to show that  $Op_h : \mathcal{C}^{r, \alpha} S^0 \rightarrow \mathcal{L}(L^2, L^2)$  is continuous with norm bounded independently of  $h$ . For this, we rescale to  $h = 1$ . That is, let  $T_h : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$  be given by  $(T_h u)(x) = h^{\frac{n}{4}} u(h^{\frac{1}{2}} x)$ . Then  $T_h$  is unitary and

$$Op_h(a)u = T_h^* Op_1(a_h) T_h u$$

where

$$a_h(x, \xi) = a(h^{\frac{1}{2}} x, h^{\frac{1}{2}} \xi).$$

Now, for  $a \in \mathcal{C}^{r, \alpha} S^m$ ,

$$\|D_\xi^\beta a_h(\cdot, \xi)\|_{\mathcal{C}^{r, \alpha}} \leq C_\beta h^{\frac{|\beta|}{2}} \langle h^{1/2} \xi \rangle^{-|\beta|}.$$

In particular,  $a_h$  is uniformly bounded in  $\mathcal{C}^{r, \alpha} S^m$ . Therefore, by [48, Chapter 13, Theorem 9.1]  $Op_1(a_h) : L^2 \rightarrow L^2$  uniformly in  $h$  with norm bounded by a finite sum of  $\mathcal{C}^{r, \alpha} S^0$  seminorms. Since  $T_h$  is unitary, this completes the proof.  $\square$

**Lemma 3.8.** *Let  $a \in \mathcal{C}_c^\infty(T^*M)$ ,  $m \in \mathbb{R}$ , and  $0 < \alpha < 1$*

$$\mathcal{C}^{1, \alpha} S^m \ni p \mapsto h^{-1} [p(x, hD), a(x, hD)] \in \mathcal{L}(L^2, L^2)$$

*is continuous with norm bounded independently of  $h$ .*

*Proof.* Let  $p \in \mathcal{C}^{1, \alpha} S^m$  and choose  $\rho \in (2/(2 + \alpha), 1)$  so that  $(1 + \alpha/2)\rho > 1$ . Let  $\psi \in \mathcal{C}_c^\infty(\mathbb{R})$  with  $\psi \equiv 1$  on  $[-2, 2]$ . Define  $p_h(x, \xi) = (\psi(h^\rho |D_x|)p)(x, \xi)$ . Then,  $p_h \in S_\rho^m$ . Moreover,

$$\begin{aligned} |D_x D_\xi^\beta p_h(x, \xi)| &\leq C_\beta \langle \xi \rangle^{m-|\beta|}, \\ |D_x^\gamma D_\xi^\beta p_h(x, \xi)| &\leq C_\beta h^{-\rho(|\gamma|-1)} \langle \xi \rangle^{m-|\beta|}, \quad |\gamma| > 1. \end{aligned}$$

Note that the symbol classes  $S_\rho^m$  have a symbol calculus and, in particular,

$$(3.6) \quad h^{-1} [Op_h(p_h), a(x, hD)] = \frac{1}{i} Op_h(\{p_h, a\}) + O(h^{1-\rho})_{L^2 \rightarrow L^2}.$$

Next, note that by the characterization of Hölder spaces by Littlewood–Paley decomposition [48, Chapter 13, Theorems 8.1, 8.2],

$$\|(p - p_h)(\cdot, \xi)\|_{\mathcal{C}^{0, \alpha/2}} \leq h^{\rho(1 + \frac{\alpha}{2})} \|p(\cdot, \xi)\|_{\mathcal{C}^{1, \alpha}} = o(h) \|p(\cdot, \xi)\|_{\mathcal{C}^{1, \alpha}}$$

In particular,  $p - p_h \in o(h) \mathcal{C}^{0, \alpha/2} S^2$  and hence

$$h^{-1} [Op_h(p - p_h), a(x, hD)] = o(1)_{L^2 \rightarrow L^2}.$$

Combining this with (3.6) completes the proof.  $\square$



*Proof of Lemma 3.6.* We start with the proof of (3.4). Note that  $P = Op_h(p_0) + hOp_h(p_1)$  where  $p_0(h) = |\xi|_{g(h)}^2 - 1 \in \mathcal{C}^{1,\alpha}S^2$  and  $p_1(h) \in \mathcal{C}^{0,\alpha}S^1$  uniformly as  $h \rightarrow 0$ .

Fix  $\psi \in \mathcal{C}_c^\infty(\mathbb{R})$  with  $\psi \equiv 1$  on  $[-2, 2]$  and let  $\epsilon > 0$ . Let  $p_{i,\epsilon}(x, \xi) = (\psi(\epsilon|D_x|)p_i)(x, \xi)$ . Then,  $p_{i,\epsilon} \in S^{2-i}$  and again by Littlewood–Paley,

$$\|D_\xi^\beta(p_0 - p_{0,\epsilon})(\cdot, \xi)\|_{\mathcal{C}^{0,\alpha}} \leq C_\beta \epsilon \langle \xi \rangle^{2-|\beta|}.$$

In particular, for  $\epsilon > 0$  small and  $h < h_0$  small,  $p_\epsilon$  is elliptic on  $\{|\xi|_g - 1\} \geq C\epsilon$ . Therefore, for  $a \in S^0(T^*M)$  supported away from  $p = 0$ , there is  $\epsilon > 0$ ,  $h_0$  small enough so that  $p_{0,\epsilon}$  is elliptic on  $\text{supp } a$ . Hence there is  $E_\epsilon \in \Psi^{-\infty}$  (uniformly for  $\epsilon > 0$  small) so that

$$a(x, hD) = E_\epsilon Op_h(p_{0,\epsilon}) + O(h^\infty)_{\Psi^{-\infty}}.$$

In particular,

$$a(x, hD) = E_\epsilon(P - Op_h(p_0 - p_{0,\epsilon}) - hOp_h(p_1)) + O(h^\infty)_{\Psi^{-\infty}}$$

Therefore,

$$\begin{aligned} \mu(a) &= \lim_{h \rightarrow 0} \langle E_\epsilon P u, u \rangle - \langle E_\epsilon [Op_h(p_0 - p_{0,\epsilon}) + hOp_h(p_1)] u, u \rangle \\ &= \lim_{h \rightarrow 0} \langle E_\epsilon h \chi f, u \rangle + O(\epsilon) = O(\epsilon) \end{aligned}$$

Since the left-hand side does not depend on  $\epsilon$ , sending  $\epsilon \rightarrow 0$ ,  $\mu(a) = 0$  and hence  $\text{supp } \mu \subset S^*M$ .

Next, fix  $K > 0$  large enough so that for  $h < h_0$ ,  $\{|\xi|_{g(h)} \leq 2\} \subset \{|\xi| \leq K\}$ . We apply the previous argument with  $a(x, hD) = \chi(x)b(x, hD)^*b(x, hD)\chi(x)(1 - \psi(K^{-1}|hD|))$ . In particular, there exists  $E_\epsilon \in \Psi^0$  compactly supported so that

$$\chi(x)b(x, hD)^*b(x, hD)\chi(x)(1 - \psi(|hD|_g)) = E_\epsilon P + O_{L_{\text{loc}}^2 \rightarrow L_{\text{comp}}^2}(\epsilon).$$

Therefore,

$$\begin{aligned} \lim_{h \rightarrow 0} \|b(x, hD)\chi u\|_{L^2}^2 &= \lim_{h \rightarrow 0} \langle \chi(x)b(x, hD)^*b(x, hD)\chi(x)u, u \rangle \\ &= \lim_{h \rightarrow 0} \langle \chi(x)b(x, hD)^*b(x, hD)\chi(x)\psi(K^{-1}|hD|)u, u \rangle \\ &\quad + \lim_{h \rightarrow 0} \langle \chi(x)b(x, hD)^*b(x, hD)\chi(x)(1 - \psi(K^{-1}|hD|))u, u \rangle \\ &= \lim_{h \rightarrow 0} \langle \chi(x)b(x, hD)^*b(x, hD)\chi(x)\psi(K^{-1}|hD|)u, u \rangle \\ &\quad + \lim_{h \rightarrow 0} \langle hE_\epsilon f, u \rangle + O(\epsilon) \\ &= \mu(|b|^2 \chi^2 \psi(|\xi|_g)) + O(\epsilon) = \mu(|b|^2 \chi^2) + O(\epsilon). \end{aligned}$$

Again, since the left hand side is independent of  $\epsilon$ , this proves (3.4).

We now prove (3.5). Let  $a \in \mathcal{C}_c^\infty(T^*M; \mathbb{R})$ . Then,

$$\frac{i}{h} \langle [P, a(x, hD)]u, u \rangle = i[(a(x, hD)u, h\chi f) - (ha(x, hD)\chi f, u)] \rightarrow 2 \text{Im } \mu^j(a).$$

Therefore, it remains to show that

$$(3.7) \quad \frac{i}{h} \langle [P, A]u, u \rangle \rightarrow \mu(H_p a).$$

First, observe that by Lemma 3.7 and the fact that

$$\|D_\xi^\beta(p_1 - p_{1,\epsilon})(\cdot, \xi)\|_{\mathcal{C}^{0, \frac{\alpha}{2}}} \leq C_\beta \epsilon^{\frac{\alpha}{2}} \langle \xi \rangle^{1-|\beta|},$$

we obtain

$$\|ha(x, hD)Op_h(p_1 - p_{1,\epsilon})\| + \|hOp_h(p_1 - p_{1,\epsilon})a(x, hD)u\|_{L^2} \leq Ch\epsilon^{\frac{\alpha}{2}}.$$

Therefore,

$$(3.8) \quad \|[hOp_h(p_1), a(x, hD)u]\|_{L^2} \leq Ch\epsilon^{\frac{\alpha}{2}} \|u\|_{L^2} + O_\epsilon(h^2).$$

Next, observe that by Lemma 3.8,

$$(3.9) \quad \|[Op_h(p_0 - p_{0,\epsilon}), a(x, hD)u]\|_{L^2} \leq C\epsilon h \|u\|_{L^2}.$$

Finally, with  $q_\epsilon := \lim_{h \rightarrow 0} p_{0,\epsilon}$ ,

$$(3.10) \quad \lim_{h \rightarrow 0} \langle \frac{i}{h} [Op_h(p_{0,\epsilon}), a(x, hD)u], u \rangle = \mu(H_{q_\epsilon} a).$$

Combining (3.8), (3.9), and (3.10) gives

$$\left| \lim_{h \rightarrow 0} \frac{i}{h} \langle [P, A]u, u \rangle - \mu(H_{q_\epsilon} a) \right| \leq C\epsilon^{\frac{\alpha}{2}}$$

Since  $p_0 \in \mathcal{C}^{1,\alpha} S^2$  uniformly in  $h$ ,  $H_{q_\epsilon} \rightarrow H_p$ . In particular, sending  $\epsilon \rightarrow 0$  and applying the dominated convergence theorem gives (3.7).  $\square$

**3.3. Appearance of the Volterra Operator.** To complete the proof of (1.3) in the boundaryless case, we prove the following measure-theoretic proposition. This lemma will be modified slightly in the case of a manifold with boundary to account for the fact that the generalized bicharacteristic flow is not generated by a vector field.

**Proposition 3.9.** *Suppose that  $\mu$ ,  $\mu^j$ , and  $\alpha$  are Radon measures on  $T^*M$  with  $\alpha$  finite and  $\mu$  finite on compact subsets of  $T^*M$ . Let  $X$  be a continuous vector field on  $T^*M$  and  $\varphi_t : T^*M \rightarrow T^*M$  be given by  $\varphi_t(q) = \exp(tX)(q)$ . Let  $\Sigma \subset T^*M$  be a hypersurface transverse to  $X$  so that the map  $F : \mathbb{R} \times \Sigma \ni (t, q) \mapsto \varphi_t(q) \in T^*M$  is homeomorphism onto its image. Suppose that  $\mu(F((-\infty, 0) \times \Sigma)) = 0$  and for  $a \in \mathcal{C}_c^0(T^*M; \mathbb{R})$ ,*

$$(3.11) \quad |\mu^j(a)| \leq \sqrt{\mu(|a|)\alpha(|a|)}.$$

Furthermore, for  $\psi \in \mathcal{C}_c^1(\mathbb{R})$ ,  $a \in \mathcal{C}_c^0(\Sigma)$ , let  $f_{\psi,a}(F(t, q)) = \psi(t)a(q)$  and assume that

$$(3.12) \quad \mu(f_{\partial_t \psi, a}) = 2 \operatorname{Im} \mu^j(f_{\psi, a}).$$

Then,

$$\mu(F([0, L] \times \Sigma)) \leq \frac{4L^2}{\pi^2} \alpha(F([0, L] \times \Sigma)).$$

*Proof.* For  $0 \leq a \in \mathcal{C}_c^0(T^*M)$ , define the Radon measures  $\mu_a^\sharp$ ,  $\operatorname{Im} \mu^{j,\sharp}$ , and  $\alpha_a^\sharp$  on  $\mathbb{R}$  by

$$\mu_a^\sharp(\psi) = \mu(f_{\psi, a}), \quad \mu_a^{j,\sharp}(\psi) = \mu^j(f_{\psi, a}), \quad \alpha_a^\sharp(\psi) = \mu(f_{\psi, a}).$$

**Lemma 3.10.** *The measure  $\mu^\sharp$  is absolutely continuous with respect to Lebesgue measure and*

$$\mu_a^\sharp = \dot{\mu}_a^\sharp dt$$

with  $|\dot{\mu}_a^\sharp| \leq C \sup |a|$ .

*Proof.* First, observe that for  $\psi \in \mathcal{C}_c^1(\mathbb{R})$ , (3.12) implies

$$\mu_a^\sharp(\partial_t \psi) = 2 \operatorname{Im} \mu^{j,\sharp}(f_{\psi, a}).$$

Let  $[b, c] \subset \mathbb{R}$ ,  $0 \leq \chi_\epsilon \in \mathcal{C}_c^\infty(\mathbb{R}; [0, 1])$  with  $\chi \equiv 1$  on  $[b, c]$  and  $\operatorname{supp} \chi_\epsilon \subset (b - \epsilon, c + \epsilon)$ . Define

$$\psi_\epsilon(t) = - \int \chi_\epsilon(s) ds + \int_{-\infty}^t \chi_\epsilon(s) ds.$$

Then, since  $\mu((-\infty, 0) \times \Sigma) = 0$ ,  $\mu^j(F((-\infty, 0) \times \Sigma)) = 0$  and hence

$$\begin{aligned} |\mu_a^\sharp(\chi_\epsilon)| &= 2 |\operatorname{Im} \mu^{j,\sharp}(1_{F([0, \infty) \times \Sigma]} f_{\psi_\epsilon, a})| \\ &\leq C \sup |a| \sup |\psi_\epsilon| \end{aligned}$$

$$\leq C \sup |a| \left| \int \chi_\epsilon ds \right| \leq C \sup |a| (c - b + 2\epsilon).$$

Since  $\chi_\epsilon \rightarrow 1_{[b,c]}$ , we have by the dominated convergence theorem,

$$|\mu_a^\sharp([a, b])| \leq C \sup |a| (c - b).$$

In particular,  $\mu_a^\sharp$  is absolutely continuous with respect to Lebesgue measure since the intervals generate the Borel sets on  $\mathbb{R}$ . Moreover, its Radon–Nikodym derivative is linear in  $a$  and bounded by  $C \sup |a|$ .  $\square$

By (3.11), since  $\mu_a^\sharp$  is absolutely continuous with respect to Lebesgue measure,  $\mu_a^{j,\sharp}$  is also. Let  $\dot{\mu}_a^\sharp$ ,  $\dot{\mu}_a^{j,\sharp}$ , and  $\dot{\alpha}_a^\sharp$  be the Radon–Nikodym derivatives of  $\mu_a^\sharp$ ,  $\mu_a^{j,\sharp}$ , and  $\alpha_a^\sharp$  respectively with respect to Lebesgue measure. In particular,

$$\mu_a^\sharp = \dot{\mu}_a^\sharp dt, \quad \text{Im } \mu_a^{j,\sharp} = \text{Im } \dot{\mu}_a^{j,\sharp} dt, \quad \alpha_a^\sharp = \dot{\alpha}_a^\sharp dt + \lambda$$

where  $\lambda \perp dt$ . Now, by (3.11),

$$\frac{1}{2r} \left| \mu_a^{j,\sharp}([t-r, t+r]) \right| \leq \frac{1}{2r} \sqrt{\mu_a^\sharp([t-r, t+r])} \sqrt{\alpha_a^\sharp([t-r, t+r])}$$

So, by the Lebesgue differentiation theorem [23, Theorem 3.22], for Lebesgue a.e.  $t$ ,

$$(3.13) \quad |\dot{\mu}_a^{j,\sharp}(t)| \leq \sqrt{\dot{\mu}_a^\sharp(t)} \sqrt{\dot{\alpha}_a^\sharp(t)}.$$

**Lemma 3.11.** *With  $\dot{\mu}_a^\sharp(t)$  as above,  $\dot{\mu}_a^{j,\sharp}(t)$  is continuous and*

$$(3.14) \quad \dot{\mu}_a^{j,\sharp}(t) \leq 2 \int_0^{\max(t,0)} |\text{Im } \dot{\mu}_a^{j,\sharp}(s)| ds \leq 2 \int_0^t \sqrt{\dot{\mu}_a^\sharp(s)} \sqrt{\dot{\alpha}_a^\sharp(s)} ds.$$

*Proof.* By (3.12), for  $\psi \in C_c^1(\mathbb{R})$ ,

$$(3.15) \quad \int \psi'(t) \dot{\mu}_a^\sharp(t) dt = 2 \int \psi(t) \text{Im } \dot{\mu}_a^{j,\sharp}(t) dt.$$

Let  $\chi \in C_c^\infty(\mathbb{R}; [0, 1])$  with  $\int \chi(s) ds \equiv 1$  and define  $\chi_{\epsilon,t'}(s) = \epsilon^{-1} \chi(\epsilon^{-1}(s - t'))$  and

$$\psi_{\epsilon,t'}(s) = -1 + \int_{-\infty}^t \chi_{\epsilon,t'}(s) ds.$$

Then by (3.15), together with the fact that  $\dot{\mu}_a^{j,\sharp}(t) = 0$  for  $t < 0$ ,

$$\int \chi_{\epsilon,t'}(t) \dot{\mu}_a^\sharp(t) dt = 2 \int_0^\infty \psi_{\epsilon,t'}(t) \text{Im } \dot{\mu}_a^{j,\sharp}(t) dt.$$

Applying the Lebesgue differentiation theorem on the left and dominated convergence on the right, we find for Lebesgue a.e.  $t'$

$$\dot{\mu}_a^\sharp(t') = -2 \int_0^{\max(t',0)} \text{Im } \dot{\mu}_a^{j,\sharp}(t) dt.$$

Note that this implies  $\dot{\mu}_a^\sharp(t')$  is continuous. Applying (3.13) gives (3.14).  $\square$

**Lemma 3.12.** *Let  $0 \leq b(t), c(t) \in L^2$  with  $b(t)$  continuous and*

$$b^2(t) \leq 2 \int_0^t b(s) c(s) ds.$$

*Then*

$$(3.16) \quad b(t) \leq \int_0^t c(s) ds.$$

*Proof.* Supposing (3.16) to be false, there exists  $\epsilon > 0$  and  $t > 0$  such that

$$b(t) \geq \int_0^t c(s)ds + \epsilon;$$

hence we may define

$$t_0 := \inf \left\{ t > 0 \mid \int_0^t c(s)ds + \epsilon \leq b(t) \right\}.$$

Then,

$$\begin{aligned} \left[ \int_0^{t_0} c(s)ds + \epsilon \right]^2 &= b^2(t_0) \leq 2 \int_0^{t_0} b(s)c(s)ds \\ &\leq 2 \int_0^{t_0} \int_0^s c(r)c(s)drds + 2\epsilon \int_0^{t_0} c(s)ds \\ &= \left[ \int_0^{t_0} \int_0^s c(r)c(s)drds + \int_0^{t_0} \int_s^{t_0} c(r)c(s)drds \right] + 2\epsilon \int_0^{t_0} c(s)ds \\ &= \left[ \int_0^{t_0} c(s)ds \right]^2 + 2\epsilon \int_0^{t_0} c(s)ds \\ &= \left[ \int_0^{t_0} c(s)ds + \epsilon \right]^2 - \epsilon^2 \end{aligned}$$

which is a contradiction.  $\square$

Using the continuity of  $\mu_a^\sharp$  together with (3.14) to apply Lemma 3.12 with  $b(t) = \sqrt{\mu_a^\sharp(t)}$ ,  $c(t) = \sqrt{\dot{\alpha}_a^\sharp(s)}$ , we have

$$\sqrt{\mu_a^\sharp(t)} \leq \int_0^t \sqrt{\dot{\alpha}_a^\sharp(s)}ds.$$

Now,

$$\begin{aligned} \mu(F([0, L] \times \Sigma)) &= \int_0^L \mu_1^\sharp(t)dt \leq \int_0^L \left| \int_0^t \sqrt{\dot{\alpha}_1^\sharp(s)}ds \right|^2 dt \\ &= \left\| V \sqrt{\dot{\alpha}_1^\sharp} \right\|_{L^2([0, L])}^2 \leq \|V\|_{L^2([0, L]) \rightarrow L^2([0, L])}^2 \int_0^L \dot{\alpha}_1^\sharp(t)dt \end{aligned}$$

where  $V : L^2([0, L]) \rightarrow L^2([0, L])$  is the Volterra operator. Since the Volterra operator acting on  $L^2([0, L])$  has norm  $(2L)/\pi$  ([28], Problem 188) and

$$\alpha(F([0, L] \times \Sigma)) = \int_0^L \dot{\alpha}_1^\sharp(t)dt,$$

we have

$$\mu(F([0, L] \times \Sigma)) \leq \frac{4L^2}{\pi^2} \alpha(F([0, L] \times \Sigma)).$$

$\square$

### 3.4. Completion of the proof of (1.3) in the boundaryless case. Let

$$\Sigma = \{ \rho \in T_{\partial B(0, R)}^* M : (H_p r)(\rho) < 0 \},$$

where  $r$  is the radial variable defined in the Euclidean end(s) as in Definition 2.1. Thus, these are the inward-pointing covectors over the boundary of the ball of radius  $R$ . By convexity of Euclidean balls,

$$\bigcup_{t>0} \exp(tH_p)(\Sigma) \supset \mathcal{I}^c, \quad \bigcup_{t \leq 0} \exp(tH_p)(\Sigma) \subset \mathcal{I},$$

and  $F$  as in Proposition 3.9 is a diffeomorphism onto its image.

Then, by (3.3) and Lemmas 3.3, 3.4, and 3.6,  $(\mu, \mu^j, \alpha, \Sigma)$  satisfy the hypotheses of Proposition 3.9 with  $\varphi_t = \exp(tH_p)$ . Therefore,

$$\mu\left(\bigcup_{t=0}^L \varphi_t(\Sigma \cap S^*M)\right) \leq \frac{4L^2}{\pi^2} \alpha(T^*M),$$

where we may take  $L = L(g, R')$ . So, since  $0 \leq \chi^2 \leq 1$  and

$$\text{supp } \chi \cap S^*M \subset \bigcup_{t=0}^{L(g, R')} \varphi_t(\Sigma \cap S^*M),$$

we conclude

$$\limsup_{h \rightarrow 0} \|\chi u\|_{L^2} = \sqrt{\mu(\chi^2)} \leq \frac{2L(g, R')}{\pi} \sqrt{\alpha(T^*M)} \leq \liminf \frac{2L(g, R')}{\pi} \|f(h)\|_{L^2}.$$

This completes the proof of (1.3) in the case of a manifold without boundary.

**3.5. Uniformity of  $k_0$  and Sobolev estimates.** In this section we prove the uniformity of  $k_0$  statement from Theorem 1, as well as the more general Sobolev mapping property (1.4). We will omit the terms  $L, L_j$  from this discussion for brevity, but will on the other hand prove a slightly more general statement about metric perturbations as follows.

We begin with uniformity in the case  $s = 0$ , i.e., in the basic  $L^2$  estimate. Fix  $R_1 > 0$  and let  $\mathcal{C}$  be a subset of  $\mathcal{C}^{1,1}$  metrics so that for  $g \in \mathcal{C}$ ,  $g$  is nontrapping and  $\text{supp}(I - g) \subset B(0, R_1)$ . Furthermore, suppose that for any  $\{g_k\}_{k=1}^\infty \subset \mathcal{C}$ , there exists a subsequence  $\{g_{k_m}\}_{m=1}^\infty$ ,  $\gamma > 0$  and  $g \in \mathcal{C}$  so that  $\|g_{k_m} - g\|_{\mathcal{C}^{1,\gamma}} \rightarrow 0$  and  $\lim_{m \rightarrow \infty} L(g_{k_m}, R_1) \geq L(g, R_1)$ . Let  $\chi \in \mathcal{C}_c^\infty(B(0, R); [0, 1])$  and  $R_1 < R' < R'' < R$  so that  $\text{supp } \chi \subset B(0, R')$ .

We start by showing that there exists  $k_0 > 0$  so that for all  $g \in \mathcal{C}$ , and  $k > k_0$ ,

$$(3.17) \quad \|\chi(P(g) - k^2 - i0)^{-1}\chi\|_{L^2 \rightarrow L^2} \leq \frac{2L(g, R'')}{\pi k}$$

Suppose not. Then since  $L(g, R') < L(g, R'')$ , there exists  $\{g_k\}_{k=1}^\infty \subset \mathcal{C}$ ,  $h_k \rightarrow 0$ ,  $f_k \in L^2(M)$ ,  $u_k \in H_{\text{loc}}^s(M)$ , with  $\|\chi u_k\|_{L^2} = 1$ , and  $\delta > 0$  so that

$$(3.18) \quad \begin{aligned} & (h_k^2 P(g_k) - 1)u_k = h_k \chi f_k, \\ & 1 = \|\chi u_k\|_{L^2} \geq \frac{2L(g_k, R') \|f_k\|_{L^2(M)}}{\pi} + \delta. \end{aligned}$$

Extracting subsequences, we may assume that  $g_k \rightarrow g$  in  $\mathcal{C}^{1,\gamma}$  for some  $\gamma > 0$  and  $\lim L(g_k, R') \geq L(g, R')$ . Note also that by Lemma 3.2,  $u_k$  is uniformly bounded in  $L_{\text{loc}}^2$  and hence, by extracting further subsequences, we may assume that  $u_k$  has defect measure  $\mu$ ,  $\chi f_k$  has defect measure  $\alpha$ , and  $u_k, \chi f_k$  has joint defect measure  $\mu^j$ . Let  $\Sigma$  denote the set of directly incoming points on  $T_{\partial B(0, R')}^*M$ . By Lemmas 3.3, 3.4, and 3.6,  $(\mu, \mu^j, \alpha, \Sigma)$  satisfy the hypotheses of Proposition 3.9 with  $\varphi_t = \exp(tH_p)$  and  $p = |\xi|_g^2 - 1$ . Therefore,

$$\mu\left(\bigcup_{t=0}^L \varphi_t(\Sigma \cap S^*M)\right) \leq \frac{4L^2}{\pi^2} \alpha(T^*M),$$

and we arrive at a contradiction just as above.

Finally, to obtain the bound (1.4), we begin by considering the case  $s = 2$ . We compute

$$(P(g) + k^2)\chi(P(g) - k^2 - i0)^{-1}\chi f = \chi^2 f + ([-\Delta_{\mathbb{R}^n}, \chi] + 2k^2\chi)(P(g) - k^2 - i0)^{-1}\chi f.$$

We next consider  $[-\Delta_{\mathbb{R}^n}, \chi]$  and show that there exists  $k_1 > 0$  so that for  $k > k_1$  and  $g \in \mathcal{C}$ ,

$$(3.19) \quad \|[-\Delta_{\mathbb{R}^n}, \chi](P - k^2 - i0)^{-1}\chi\|_{L^2 \rightarrow L^2} \leq \sup_{S^*M} |H_{|\xi|^2} \chi| \frac{2L(g_k, R'')}{\pi}.$$

Suppose that (3.19) does not hold. Then there is  $\delta > 0$  and a sequence  $h_k \rightarrow 0$ ,  $f_k \in L^2(M)$ ,  $u_k \in H^2(M)$  with  $\|f_k\|_{L^2} = 1$ ,  $u_k$  so that

$$\begin{aligned} (h_k^2 P(g_k) - 1)u_k &= \chi h_k f_k, \quad u_k|_{\partial M} = 0, \\ \|h_k^{-1}[-h_k^2 \Delta_{\mathbb{R}^n}, \chi]u_k\|_{L^2} &\geq \sup_{S^*M} |H_{|\xi|^2} \chi| \frac{2(L(g_k, R') + \delta)}{\pi} \end{aligned}$$

Let  $\tilde{\chi} \in C_c^\infty(B(0, R'); [0, 1])$  with  $\tilde{\chi} \equiv 1$  on  $\text{supp } \chi$ . As before, we may assume that  $u_k$  has a defect measure and then, by Lemma 4.2 applied with  $\chi$  replaced by  $\tilde{\chi} \in C_c^\infty(B(0, R'))$ ,

$$\begin{aligned} \sup_{S^*M} |H_{|\xi|^2} \chi|^2 \frac{4(L(g_k, R') + \delta)^2}{\pi^2} &\leq \|h^{-1}[-h^2 \Delta_{\mathbb{R}^n}, \chi]\tilde{\chi}u_k\|_{L^2}^2 \rightarrow \mu(|H_{|\xi|^2} \chi|^2 \tilde{\chi}^2) \\ &\leq \sup_{S^*M} |H_{|\xi|^2} \chi|^2 \mu(\tilde{\chi}^2) \leq \sup_{S^*M} |H_{|\xi|^2} \chi|^2 \frac{4(L(g, R'))^2}{\pi^2}. \end{aligned}$$

a contradiction.

Thus, by (3.17), for  $k > k_0$ , and any  $g \in \mathcal{C}$ ,

$$\begin{aligned} \|(P(g) + k^2)\chi(P(g) - k^2 - i0)^{-1}\chi f\|_{L^2} &\leq 2k^2 \|\chi(P(g) - k^2 - i0)^{-1}\chi f\|_{L^2} + \|\chi^2 f\| + \|[-\Delta_{\mathbb{R}^n}, \chi](P(g) - k^2 - i0)^{-1}\chi f\| \\ &\leq \frac{4L(g, R'')k}{\pi} \|\chi f\|_{L^2} + \|[-\Delta_{\mathbb{R}^n}, \chi](P(g) - k^2 - i0)^{-1}\chi f\|_{L^2} + \|\chi^2 f\|_{L^2}. \end{aligned}$$

Using (3.19), we have for  $k > k_1$ , and any  $g \in \mathcal{C}$ ,

$$\|[-\Delta_{\mathbb{R}^n}, \chi](P(g) - k^2 - i0)^{-1}\chi f\|_{L^2} \leq \sup_{S^*M} |H_{|\xi|^2} \chi| \frac{2L(g_k, R'')}{\pi} \|f\|_{L^2}.$$

In particular, letting

$$k > \max\left(k_0, k_1, \frac{\sup_{S^*M} |H_{|\xi|^2} \chi| 2L(g, R'') + \pi}{4(L(g, R) - L(g, R''))}\right),$$

we have for all  $g \in \mathcal{C}$ ,

$$(3.20) \quad \begin{aligned} \|\chi(P(g) - k^2 - i0)^{-1}\chi f\|_{H^2} &= \|(P(g) + k^2)\chi(P(g) - k^2 - i0)^{-1}\chi f\|_{L^2} \\ &\leq \frac{4L(g, R)k}{\pi} \|f\|_{L^2}. \end{aligned}$$

Since  $L(g, R'') \leq L(g, R)$ , the general case of  $0 \leq s \leq 2$  then follows by interpolation between (3.17) and (3.20)

If we choose  $\mathcal{C}$  to be a subset of an open neighborhood in the  $\mathcal{C}^{2,\alpha}$  topology of a given  $\mathcal{C}^{2,\alpha}$  nontrapping metric, then provided this neighborhood is chosen sufficiently small, all the metrics in  $\mathcal{C}$  are nontrapping, and the subsequence condition is guaranteed by the compactness of the embedding  $\mathcal{C}^{2,\alpha} \hookrightarrow \mathcal{C}^{1,\gamma}$ ; indeed we have subsequential convergence in  $\mathcal{C}^{1,1}$ , which ensures that  $\lim L(g_{k_m}, R_1) = L(g, R_1)$ . This choice of  $\mathcal{C}$  then gives the uniformity assertion from Theorem 1. This completes the proof of (1.3) in the case of a manifold without boundary.

## 4. MANIFOLDS WITH BOUNDARY

We now turn to the case of manifolds with boundary. Our treatment of the propagation of defect measures in this setting is motivated by previous work of Miller [36] and Burq–Lebeau [10]; see also [8].

Let  $\widetilde{M}$  be an open manifold extending  $M$  and extend  $P$  to an operator on  $\widetilde{M}$ . We then let  $\underline{u}$  be the extension of  $u$  by 0. As in the boundaryless case, we will assume that we have a sequence  $h = h_j \downarrow 0$  with

$$Pu(h) = h\chi f(h) \text{ in } M, \quad u|_{x_1=0} = 0,$$

and we take

$$\begin{aligned} \|f(h)\| &= 1, \\ \lim_{h \rightarrow 0} \|\chi u\|_{L^2} &= M > L(\Omega, g, R). \end{aligned}$$

Since propagation of singularities is a local consideration, we may employ Riemannian normal coordinates  $(x_1, x')$  in which  $\partial M$  is given by  $x_1 = 0$ ,  $M = \{x_1 > 0\}$ , and the operator  $P$  is given by

$$(4.1) \quad P = (hD_{x_1})^2 + R(x_1, x', hD_{x'}) + hE$$

for some self-adjoint  $E \in \text{Diff}_h^1$  and for  $R \in \text{Diff}_h^2$  with symbol  $r(x_1, x', \xi')$ .

We proceed as before, letting  $\mu$  be a defect measure of  $\underline{u}$ ,  $\alpha$  be a defect measure of  $\underline{\chi f}$ ,  $\mu^j$  a joint defect measure of  $\underline{u}$  and  $\underline{\chi f}$ . Finally, let  $\nu$  be the defect measures of  $hD_{x_1}u|_{x_1=0}$  on  $\partial M$ . Thus,

$$(4.2) \quad \begin{aligned} \lim_{h \downarrow 0} \langle a(x, hD)\underline{u}, \underline{u} \rangle &= \int a \, d\mu \\ \lim_{h \downarrow 0} \langle a(x, hD)\underline{\chi f}, \underline{\chi f} \rangle &= \int a \, d\alpha \\ \lim_{h \downarrow 0} \langle a(x, hD)\underline{\chi f}, \underline{u} \rangle &= \int a \, d\mu^j, \\ \lim_{h \downarrow 0} \langle b(x', hD')hD_{x_1}u, hD_{x_1}u \rangle_{\partial M} &= \int_{\partial M} b \, d\nu. \end{aligned}$$

4.1. Local study of the measure  $\mu$ .

**Lemma 4.1.** *For  $a = a_0(x, \xi') + a_1(x, \xi')\xi_1$  with  $a_i \in \mathcal{C}_c^\infty([0, \epsilon) \times T^*\partial M)$ , we have*

$$\mu(H_p a) = 2 \text{Im} \mu^j(a) - \nu(a_1|_{x_1=0}).$$

*Proof.* Here we work in a small neighborhood of the boundary and write

$$A = a_0(x, hD') + a_1(x, hD')hD_{x_1}.$$

Equation (3.2) must now be modified owing to boundary terms arising in integration by parts. In particular,

$$\begin{aligned} \langle (hD_{x_1})^2 Au, u \rangle_M \\ = ih(\langle hD_{x_1} Au, u \rangle_{L^2(\partial M)} + \langle Au, hD_{x_1} u \rangle_{L^2(\partial M)}) + \langle Au, (hD_{x_1})^2 u \rangle_{L^2(M)}. \end{aligned}$$

There is also a term arising when we integrate by parts to change  $\langle f, Au \rangle$  to  $\langle Au, f \rangle$  but this term is  $O(h)$ . Now,

$$\begin{aligned} \langle hD_{x_1} Au, u \rangle_{L^2(\partial M)} &= \langle a_0 hD_{x_1} u + a_1 (hD_{x_1})^2 u, u \rangle_{L^2(\partial M)} + o(1) \\ &= \langle a_0 hD_{x_1} u + a_1 (1 + h^2 \Delta_{\partial M}) u, u \rangle_{L^2(\partial M)} + \langle a_1 h\chi f, u \rangle_{\partial M} + o(1) \end{aligned}$$

Now, with Dirichlet boundary conditions, this is actually just 0. Moreover,

$$\langle Au, hD_{x_1}u \rangle_{L^2(\partial M)} = \langle [a_0 + a_1 hD_{x_1}]u, hD_{x_1}u \rangle = \langle a_1 hD_{x_1}u, hD_{x_1}u \rangle.$$

So (3.2) now reads

$$ih^{-1}\langle [P, A]u, u \rangle_M = \text{Im}\langle A\chi f, u \rangle_M - \langle a_1 hD_{x_1}u, hD_{x_1}u \rangle_{\partial M} + o(1). \quad \square$$

Now, we want to upgrade this to a statement for all functions in  $\mathcal{C}_c^\infty(T^*M)$ . First, we need to show that  $\mu$  is supported on  $S^*M$ .

**Lemma 4.2.** *Let  $\mu$  be a defect measure associated to  $u$  as above. Then  $\text{supp } \mu \subset S^*M$  and  $\mu(\chi^2) = 1$ . Moreover, for  $b \in S^1(T^*M)$ , supported away from  $\partial M$ ,*

$$\lim_{h \rightarrow 0} \|Op_h(b)\chi u\|_{L^2}^2 = \mu(|b|^2\chi^2).$$

*Proof.* Note that

$$P\underline{u} = 1_{x_1 \geq 0}Pu + h(-\delta(x_1) \otimes h\partial_{x_1}u|_{x_1=0} + h\delta'(x_1) \otimes u|_{x_1=0})$$

Now, suppose  $a \in \mathcal{C}_c^\infty(T^*\widetilde{M})$  has  $\text{supp } a \cap \{p = 0\} = \emptyset$ . Then, there exist  $E \in \Psi^{-2}$  so that  $a(x, hD) = EP + O_{\mathcal{D}' \rightarrow \mathcal{C}^\infty}(h^\infty)$ . Therefore,

$$a(x, hD)\underline{u} = a(x, hD)1_{x_1 \geq 0}Pu + hE(-\delta(x_1) \otimes h\partial_{x_1}u|_{x_1=0})$$

Now, since  $E : H_h^s \rightarrow H_h^{s+2}$ , is bounded and  $\|h\partial_{x_1}u|_{x_1=0}\| \leq C$ , we have

$$a(x, hD)\underline{u} = a(x, hD)1_{x_1 \geq 0}Pu + O_{L^2}(h^{1/2}).$$

In particular, this implies that  $\mu(a) = 0$  completing the proof.

To see that  $\mu(\chi^2) = 1$ , let  $\psi \in \mathcal{C}_c^\infty(\mathbb{R})$  with  $\psi \equiv 1$  on  $|s| \leq 2$ . Then, as above there exists  $E \in S^{-2}$  so that

$$\chi(1 - \psi)(|hD|)\chi = EP + O_{\mathcal{D}' \rightarrow \mathcal{S}}(h^\infty)$$

and hence again

$$\chi(1 - \psi)(|hD|)\chi\underline{u} = O_{L^2}(h^{\frac{1}{2}}).$$

In particular,

$$\begin{aligned} 1 &= \langle \chi^2\underline{u}, \underline{u} \rangle = \langle \chi\psi(|hD|)\chi\underline{u}, \underline{u} \rangle + O(h^{\frac{1}{2}}) \\ &= \int \chi^2(x)\psi(|\xi|)d\mu = \int \chi^2 d\mu \end{aligned}$$

where in the last line we use the fact that  $\mu$  is supported on  $S^*M$ .

The last claim follows the same arguments as in the boundaryless case.  $\square$

We now introduce notation for the even and odd parts of smooth functions on  $T^*M$ .

*Definition 4.3.* For  $a \in \mathcal{C}^\infty(T^*M)$ , let

$$\begin{aligned} a_o(x, \xi_1, \xi') &= \frac{a(x, \xi_1, \xi') - a(x, -\xi_1, \xi')}{2\xi_1}, \\ a_e(x, \xi_1, \xi') &= \frac{a(x, \xi_1, \xi') + a(x, -\xi_1, \xi')}{2} \end{aligned}$$

Thus,  $a = a_e + \xi_1 a_o$  and  $a_e, a_o$  are both even functions of  $\xi_1$ .

Next we upgrade the test functions for which we can compute  $\mu(H_p a)$ .



**Lemma 4.4.** For  $a \in \mathcal{C}_c^\infty(T^*\widetilde{M})$ ,

$$\mu(H_p a) = 2 \operatorname{Im} \mu^j(a) - \dot{\nu}(a_o|_{x_1=0})$$

In particular, for  $a = a(x, x_1 \xi_1, \xi') \in \mathcal{C}_c^\infty({}^b T^*\overline{\mathbb{R}_+^n})$ ,

$$\mu(H_p a) = 2 \operatorname{Im} \mu^j(a).$$

*Remark 4.5.* Here,  ${}^b T^*\overline{\mathbb{R}_+^n}$  denotes the  $b$ -cotangent bundle of this local model for a manifold with boundary, constructed so that smooth functions on this space are simply functions of  $(x, x_1 \xi_1, \xi')$ . See Section 4.2 for a more precise description of this space and for additional references. We are regarding  $a$  here as extended to a smooth function on  $T^*\widetilde{M}$ ; the choice of extension is immaterial, as  $\mu$  is supported on  $M$ .

*Proof.* Since  $a_o$  and  $a_e$  are both even functions, we may write

$$a_{e/o}(x, \xi_1, \xi') = \tilde{a}_{e/o}(x, \xi_1^2, \xi')$$

for smooth functions  $\tilde{a}_{e/o}$ ; thus,

$$a = \tilde{a}_e(x, \xi_1^2, \xi') + \tilde{a}_o(x, \xi_1^2, \xi') \xi_1.$$

Now, since  $H_p$  is tangent to  $S^*\widetilde{M}$ ,  $H_p 1_{S^*\widetilde{M}} a = 1_{S^*\widetilde{M}} H_p a$ . In particular, since  $\operatorname{supp} \mu \subset S^*\widetilde{M}$ ,

$$\mu(H_p a) = \mu(H_p 1_{S^*\widetilde{M}} a).$$

Now,

$$S^*\widetilde{M} := \{\xi_1^2 - r(x, \xi') = 0\}.$$

Therefore,

$$1_{S^*\widetilde{M}} a = 1_{S^*\widetilde{M}} [\tilde{a}_e(x, r(x, \xi'), \xi') + \tilde{a}_o(x, r(x, \xi'), \xi') \xi_1].$$

Let  $b_{e/o} = a_{e/o}(x, r(x, \xi'), \xi')$ . Then we have

$$\begin{aligned} \mu(H_p a) &= \mu(H_p 1_{S^*\widetilde{M}} [b_e + b_o \xi_1]) \\ &= \mu(H_p (b_e + b_o \xi_1)) \\ &= 2 \operatorname{Im} \mu^j(b_e + b_o \xi_1) - \dot{\nu}(b_o|_{x_1=0}). \end{aligned}$$

Now, since  $\operatorname{supp} \mu \subset S^*M$ ,  $\operatorname{supp} \mu^j \subset S^*M$  (by Lemma 3.3) and hence

$$\mu^j(b_e + b_o \xi_1) = \mu^j(a).$$

Therefore, we have

$$\mu(H_p a) = 2 \operatorname{Im} \mu^j(a) - \dot{\nu}(b_o|_{x_1=0}).$$

□

Before proceeding further, we recall the decomposition of  $T_{\partial M}^*M$  into elliptic, hyperbolic, and glancing regions. In the notation of (4.1), with  $r$  the symbol of  $R(x_1, x', hD_{x'})$ , we write

$$\begin{aligned} (y, \xi_1, \xi') &\in \mathcal{E} \text{ if } r(0, x', \xi') > 0 \\ (y, \xi_1, \xi') &\in \mathcal{H} \text{ if } r(0, x', \xi') < 0 \\ (y, \xi_1, \xi') &\in \mathcal{G} \text{ if } r(0, x', \xi') = 0 \end{aligned}$$

to denote the elliptic resp. hyperbolic resp. glancing sets.

We further split the glancing set into subsets as follows: we let

$$\begin{aligned} \mathcal{G}_d &:= \mathcal{G} \cap \{H_p^2 x_1 > 0\}, \\ \mathcal{G}_g &:= \mathcal{G} \cap \{H_p^2 x_1 < 0\} \end{aligned}$$

these are the *diffractive* and *gliding* sets respectively. We will also employ the filtration of the remainder of the glancing set given by the order of contact with the boundary as follows:

$$\mathcal{G}^k := \{\rho \in \mathcal{G} \mid (H_p^j x_1)(\rho) = 0, \text{ for } 0 \leq j < k, \text{ and } (H_p^k x_1)(\rho) \neq 0\}.$$

(We remark that this definition differs from the  $G^k$  defined in [29, Section 24.3] in that our  $\mathcal{G}^k$  does not include higher-order glancing as well.)

With an open set  $U \subset {}^bT^*M$  fixed, we will also use the notation

$$\tilde{\mathcal{G}}^k := \{\rho \in U \mid \rho \text{ is connected to } \mathcal{G}^k \text{ by a generalized bicharacteristic}\}.$$

We further let

$$\tilde{\mathcal{G}}^{\geq k} := \bigcup_{k' \geq k} \tilde{\mathcal{G}}^{k'}.$$

### Near Glancing:

We now commence the study of the measure  $\mu$  near the glancing set  $\mathcal{G}$ . Here we follow an inductive strategy employed in [10].

**Lemma 4.6.** *There exists a positive measure  $\mu^\partial$  on the glancing set  $\mathcal{G}$  so that*

$$\mu = \mu 1_{x_1 > 0} + \delta(x_1) \otimes \delta(H_p x_1) \otimes \mu^\partial.$$

In particular,  $\mu(\mathcal{H}) = 0$ .

*Proof.* Let  $a_\epsilon = \varphi(x, \xi') \epsilon \chi(\epsilon^{-1} x_1)$  with  $\chi \in \mathcal{C}_c^\infty(\mathbb{R})$ ,  $\chi(0) = \chi'(0) = 1$ .

By Lemma 4.4,

$$\mu(H_p a_\epsilon) = 2 \operatorname{Im} \mu^j(a_\epsilon).$$

By the Dominated Convergence Theorem,

$$2 \operatorname{Im} \mu^j(a_\epsilon) \rightarrow 0$$

On the other hand,

$$H_p a_\epsilon = \chi'(\epsilon^{-1} x_1) \varphi(x, \xi') H_p x_1 + O(\epsilon).$$

So, by dominated convergence,

$$\mu(H_p a_\epsilon) \rightarrow \mu(1_{x_1=0} H_p x_1 \varphi).$$

In particular,

$$\mu 1_{x_1=0}(H_p x_1 \varphi) = 0$$

for any  $\varphi$  and hence  $\mu 1_{x_1=0}$  is supported on  $\mathcal{G}$ . Since it is a positive measure,  $\mu 1_{x_1=0}$  takes the form specified.  $\square$

**Lemma 4.7.** *We have*

$$-(H_p^2 x_1) \mu^\partial = 4 \dot{\nu} 1_{\mathcal{G}}.$$

In particular, since  $\mu^\partial$  and  $\dot{\nu}$  are positive measures,

$$\operatorname{supp} \mu^\partial \cup \operatorname{supp} \dot{\nu} \subset \{H_p^2 x_1 \leq 0\},$$

hence

$$\mu(\mathcal{G}_d) = 0,$$

and for  $E \subset \mathcal{G}$ ,

$$\dot{\nu}(E) = \dot{\nu}(E \cap \{H_p^2 x_1 < 0\}).$$

*Proof.* Since  $H_p x_1 = 2\xi_1$ ,

$$H_p(2a(x, \xi)\xi_1) = aH_p^2 x_1 + 2\xi_1 H_p a.$$

Let  $a_\epsilon = \chi(\epsilon^{-1}x_1)\chi(\epsilon^{-1}r(x, \xi'))2a(x, \xi')\xi_1$  where  $\chi \in \mathcal{C}_c^\infty(\mathbb{R})$  has  $\chi(0) = 1$ ,  $\chi'(0) = 0$ . Then

$$\begin{aligned} H_p a_\epsilon &= \epsilon^{-1}\chi'(\epsilon^{-1}x_1)\chi(\epsilon^{-1}r(x, \xi'))2a\xi_1 H_p x_1 \\ &\quad + \epsilon^{-1}\chi(\epsilon^{-1}x_1)\chi'(\epsilon^{-1}r(x, \xi'))2a\xi_1 H_p r \\ &\quad + 2\xi_1\chi(\epsilon^{-1}x_1)\chi(\epsilon^{-1}r(x, \xi'))H_p a + a\chi(\epsilon^{-1}x_1)\chi(\epsilon^{-1}r(x, \xi'))H_p^2 x_1 \\ &= a\chi(\epsilon^{-1}x_1)\chi(\epsilon^{-1}r(x, \xi'))H_p^2 x_1 + O(1)(|\chi'(\epsilon^{-1}x_1)| + |\chi'(\epsilon^{-1}r(x, \xi'))| + \epsilon^{1/2}) \end{aligned}$$

Here we have used that on  $S^*M$ ,  $H_p r = -H_p \xi_1^2 = O(\xi_1)$ . Then by dominated convergence,

$$\mu(H_p a_\epsilon) \rightarrow \frac{1}{2}\mu^\partial([H_p^2 x_1]a).$$

On the other hand, since  $a_\epsilon = O(\sqrt{\epsilon})$  on  $S^*M$ ,

$$\mu(H_p a_\epsilon) = 2\operatorname{Im} \mu^j(a_\epsilon) - \dot{\nu}(2a(0, x', \xi')\chi(\epsilon^{-1}r)) \rightarrow -\dot{\nu}(21_{\mathcal{G}}a)$$

So,  $-H_p^2 x_1 \mu^\partial = 4\dot{\nu}1_{\mathcal{G}}$  as claimed.  $\square$

#### 4.2. Invariance of the measure $\mu$ .

##### The $b$ -cotangent bundle, ${}^bT^*M$

In our discussion of the invariance of  $\mu$ , it will be convenient to have some facts about the  $b$ -cotangent bundle,  ${}^bT^*M$ . We refer the reader to [29, Section 18.3] (where the notation  $\tilde{T}^*M$  is used) and [49] for more details on the  ${}^bT^*M$ . Recall that  ${}^bT^*M$  is the dual to the vector bundle of vector fields tangent to  $\partial M$ . In local coordinates with  $\partial M = \{x_1 = 0\}$ , the bundle of vector fields tangent to  $M$  is generated over  $\mathcal{C}^\infty(M)$  by the vector fields

$$\{x_1\partial_{x_1}, \partial_{x_2}, \dots, \partial_{x_n}\}.$$

Over  $x_1 > 0$ ,  ${}^bT^*M$  is canonically diffeomorphic to  $T^*M$  with

$$(x, \zeta, \xi') \mapsto (x, \zeta/x_1, \xi').$$

In fact, we can identify  $T^*M$  as a subset of  ${}^bT^*M$  using the canonical projection map  $\pi : T^*M \rightarrow {}^bT^*M$  given in local coordinates by

$$\pi : (x, \xi_1, \xi') \mapsto (x, x_1\xi_1, \xi').$$

Using this map, we can push forward the measure  $\mu$  to a measure on  ${}^bT^*M$ . It is actually this pushforward  $\pi_*\mu$  for which we will show invariance under the generalized bicharacteristic flow.

Define

$$H_p^G = H_p + \frac{H_p^2 x_1}{H_{x_1}^2 p} H_{x_1}.$$

That is,  $H_p^G$  is the gliding vector field. Then  $H_p^G$  is tangent to  $\mathcal{G}$  (cf. [29, Section 24.3]). Let also  $\varphi_t : {}^bT^*M \rightarrow {}^bT^*M$  denote the generalized bicharacteristic flow on  ${}^bT^*M$ .

### Invariance

The aim of the rest of this section is to show that

$$(4.3) \quad \pi_*\mu(q \circ \varphi_t) - \pi_*\mu(q) = \int_0^t 2 \operatorname{Im} \pi_*\mu^j(q \circ \varphi_s) ds, \quad q \in \mathcal{C}_c^\infty({}^bT^*M).$$

The main lemma is as follows.

**Lemma 4.8.** *Let  $\mu$  and  $\mu_1$  be measures on  $T^*M$  supported by  $S^*M$  with  $\mu_1 \ll \mu$ . The following are equivalent:*

(1)

$$(4.4) \quad \pi_*\mu(q \circ \varphi_t) - \pi_*\mu(q) = \int_0^t \pi_*\mu_1(q \circ \varphi_s) ds, \quad q \in \mathcal{C}_c^\infty({}^bT^*M).$$

(2) For  $a = a(x_1, x', x_1\xi_1, \xi')$

$$(4.5) \quad \mu(H_p a) = \mu_1(a).$$

and  $\mu(\mathcal{G}_d \cup \mathcal{H}) = 0$ .

Once we prove this, Lemmas 4.4 and 4.7 imply that  $\mu$  satisfies condition (2) with  $\mu_1 = 2 \operatorname{Im} \mu^j$  and hence satisfies (4.3).

### Proof of Lemma 4.8

Here we follow [10, Section 3.3].

(1) IMPLIES (2):

Let  $a \in \mathcal{C}_c^\infty(T^*M)$  with  $a = a(x, x_1\xi_1, \xi')$ . Then there exists  $q \in \mathcal{C}_c^\infty({}^bT^*M)$  such that  $a = \pi^*q$ . Then,  $q \circ \varphi_s(\rho)$  is differentiable from the left and right for all  $s$  and

$$(4.6) \quad \partial_s^\mp(q \circ \varphi_s)(\rho) = H_p a(\pi_\pm^{-1}(\varphi_s(\rho))),$$

where  $\pi_\pm^{-1}$  denotes the outward and inward pointing inverses of  $\pi$ . In particular, for  $s$  such that  $\varphi_s(\rho) \notin \mathcal{H}$ ,

$$\partial_s(q \circ \varphi_s)(\rho) = H_p a(\varphi_s(\rho)).$$

Now,  $\pi_*\mu(\mathcal{H}) = 0$  since  $\pi_*\mu$  satisfies (4.4) and  $\mathcal{H}$  is transverse to the flow  $\varphi_s$ . Therefore, since  $\pi_*\mu$  satisfies (4.4), differentiating yields

$$\pi_*\mu(\partial_t q \circ \varphi_t|_{t=s}) = \pi_*\mu_1(q \circ \varphi_s)$$

and setting  $s = 0$ , gives

$$\mu(H_p a) = \mu_1(a).$$

Finally, since  $\mathcal{G}_d$  is transverse to the flow, and  $\pi_*\mu$  satisfies (1),  $\pi_*\mu(1_{\mathcal{G}_d}) = \mu(1_{\mathcal{G}_d}) = 0$ . This completes the proof of (2).  $\square$

(2) IMPLIES (1):

**Lemma 4.9.** *For  $b \in \mathcal{C}_c^\infty(T^*M)$ ,*

$$\mu(H_p b) = \mu_1(b) + \mu(1_{x_1=0} \partial_{x_1} r b_o|_{\xi_1^2=r(x, \xi')}) + u^0(b_o|_{x_1=0, \xi_1^2=r(0, x', \xi')})$$

where  $u^0$  is a measure supported  $\mathcal{H}$ ,  $1_{x_1=0}\mu$  is supported by  $\mathcal{G} \setminus \mathcal{G}_d$ , and

$$b_o = \frac{b(x, \xi_1, \xi') - b(x, -\xi_1, \xi')}{2\xi_1}$$

is the  $\xi_1$  odd part of  $b$ .

*Proof.* Write

$$b = b_e + \xi_1 b_o$$

where

$$b_e = \frac{b(x, \xi_1, \xi') + b(x, -\xi_1, \xi')}{2}.$$

Then,  $b_e = \tilde{b}_e(x, \xi_1^2, \xi')$ ,  $b_o = \tilde{b}_o(x, \xi_1^2, \xi')$  and so on  $S^*M$ ,

$$b_e = \tilde{b}_e(x, r(x, \xi'), \xi') =: a_e(x, \xi'), \quad b_o = \tilde{b}_o(x, r(x, \xi'), \xi') =: a_o(x, \xi').$$

So, since  $H_p$  is tangent to  $S^*M$ ,

$$\mu(H_p b) = \mu(H_p(a_e + \xi_1 a_o)) = \mu(H_p \xi_1 a_o) + \mu_1(a_e).$$

Now, let

$$(4.7) \quad \begin{aligned} \chi &\in \mathcal{C}_c^\infty(\mathbb{R}), \quad \chi \equiv 1 \text{ near } 0, \quad \chi' \leq 0 \text{ on } x_1 \geq 0, \\ \chi_\epsilon(x_1) &= \chi(\epsilon^{-1} x_1). \end{aligned}$$

Then,

$$\mu(H_p(1 - \chi_\epsilon)\xi_1 a_o) = \mu_1((1 - \chi_\epsilon)\xi_1 a_o) \rightarrow \mu_1(1_{x_1 > 0}\xi_1 a_o).$$

Now,

$$\mu(H_p(\chi_\epsilon \xi_1 a_o)) = \mu(\chi_\epsilon \xi_1 H_p a_o + \chi_\epsilon \partial_{x_1} r a_o + 2\epsilon^{-1} \chi'_\epsilon \xi_1^2 a_o) =: I_\epsilon + II_\epsilon + III_\epsilon.$$

By the dominated convergence theorem,

$$I_\epsilon \rightarrow \mu(1_{x_1=0}\xi_1 H_p a_o), \quad II_\epsilon \rightarrow \mu(1_{x_1=0}\partial_{x_1} r a_o).$$

Note that since  $\mu(\mathcal{H}) = 0$  and  $\xi_1 = 0$  on  $\{x_1 = 0\} \setminus \mathcal{H}$ , we have

$$(4.8) \quad I_\epsilon \rightarrow 0, \quad II_\epsilon \rightarrow \mu(1_{x_1=0}\partial_{x_1} r a_o).$$

Now, observe that

$$\mu_1((1 - \chi_\epsilon)\xi_1 a_o) = \mu(H_p(1 - \chi_\epsilon)\xi_1 a_o) = \mu((1 - \chi_\epsilon)H_p \xi_1 a_o) - III_\epsilon.$$

So,

$$III_\epsilon = \mu((1 - \chi_\epsilon)H_p \xi_1 a_o) - \mu_1((1 - \chi_\epsilon)\xi_1 a_o)$$

and in particular,

$$(4.9) \quad \lim_{\epsilon \rightarrow 0} III_\epsilon(a_o) = \mu(1_{x_1 > 0} H_p \xi_1 a_o) - \mu_1(1_{x_1 > 0} \xi_1 a_o).$$

On the other hand for  $a_o \geq 0$ ,

$$III_\epsilon(a_o) \leq 0.$$

Therefore,  $III_\epsilon$  is a family of measures with a weak limit. In particular,  $III_\epsilon \rightarrow u^0$  for some measure  $u_0$  supported in  $x_1 = 0$ . In fact, this also shows that

$$(4.10) \quad a_o \mapsto \mu(1_{x_1 > 0} H_p(\xi_1 a_o)) - \mu_1(\xi_1 a_o) =: u^0(a_o)$$

is a measure. We now check that  $u^0$  is supported in  $\mathcal{H}$ . Note that once we show this the proof will be complete, since by (4.8) and (4.9) together with  $\mu_1(\mathcal{H}) = 0$ , we then have

$$\begin{aligned} \mu(H_p b) &= \mu_1(a_e) + \mu_1(\xi_1 a_o) + \mu(1_{x_1=0}\partial_{x_1} r a_o) + u^0(a_o) \\ &= \mu_1(b) + \mu(1_{x_1=0}\partial_{x_1} r a_o) + u^0(a_o) \end{aligned}$$

as desired.

Let  $\chi, \chi_\epsilon$  be as in (4.7). Note that for each  $\epsilon > 0$ ,  $a_o \chi_\epsilon(x_1) \chi_{\sqrt{\epsilon}}(\xi_1)$  can be written as a smooth function of  $(x_1, x', x_1 \xi_1, \xi')$ . Using (4.10), we now decompose

$$u^0(a_o \chi_\epsilon(x_1) \chi_{\sqrt{\epsilon}}(\xi_1)) = u_1^0(\epsilon) + u_2^0(\epsilon) + u_3^0(\epsilon) + u_4^0(\epsilon)$$

where

$$\begin{aligned} u_1^0(\epsilon) &= \mu(1_{x_1>0}\chi_\epsilon(x_1)\chi_{\sqrt{\epsilon}}(\xi_1)\xi_1 H_p a_o), \\ u_2^0(\epsilon) &= \mu(1_{x_1>0}\epsilon^{-1}\chi'_\epsilon(x_1)\chi_{\sqrt{\epsilon}}(\xi_1)2\xi_1^2 a_o), \\ u_3^0(\epsilon) &= \mu(1_{x_1>0}\chi_\epsilon(x_1)(\xi_1\epsilon^{-\frac{1}{2}}\chi'_{\sqrt{\epsilon}}(\xi_1) + \chi_{\sqrt{\epsilon}}(\xi_1))\partial_{x_1} r a_o), \\ u_4^0(\epsilon) &= -\mu_1(\xi_1 a_o \chi_\epsilon(x_1)\chi_{\sqrt{\epsilon}}(\xi_1)). \end{aligned}$$

Clearly  $u_1^0(\epsilon) \rightarrow 0$  by dominated convergence. The same is true of  $u_2^0(\epsilon)$ : the function  $\chi'_\epsilon(x_1)\chi_{\sqrt{\epsilon}}(\xi_1)2\xi_1^2 a_o$  is uniformly bounded as  $\epsilon \downarrow 0$  since  $\xi_1^2 \leq C\epsilon$  on  $\text{supp } \chi_{\sqrt{\epsilon}}(\xi_1)$ . Next,

$$u_3^0(\epsilon) \rightarrow 0$$

as  $\xi_1\epsilon^{-\frac{1}{2}}\chi'_{\sqrt{\epsilon}}(\xi_1) + \chi_{\sqrt{\epsilon}}(\xi_1)$  likewise remains uniformly bounded. Finally,

$$u_4^0(\epsilon) \rightarrow -\mu_1(1_{x_1=0}1_{\xi_1=0}\xi_1 a_o) = 0.$$

In particular, this implies by the dominated convergence theorem that

$$u^0(a_o 1_{x_1=\xi_1=0}) = 0$$

and so  $u^0$  is supported by  $\mathcal{H}$ . □

### Invariance away from glancing

We now prove the integral invariance statement (1) away from glancing.

**Lemma 4.10.** *Suppose  $\tilde{a} = \tilde{a}(x, \sigma, \xi') \in \mathcal{C}_c^\infty$  and let  $a = \tilde{a}(x, x_1 \xi_1, \xi')$ . Furthermore, assume that for  $\rho \in \text{supp } a$ , on  $[-T, 0]$   $\varphi_{-t}(\rho)$  intersects  $x_1 = 0$  at most once and in a hyperbolic point. Then,*

$$\mu(a) - \mu(a \circ \varphi_{-T}) = \int_{-T}^0 \mu_1(a \circ \varphi_{-s}) ds.$$

*Proof.* Let  $-T(\rho)$  be the time that  $\varphi_{-T(\rho)}(\rho) \in \mathcal{H}$ . Then

$$\begin{aligned} \int_{-T}^0 H_p(a \circ \varphi_s)(\rho) ds &= \int_{-T}^{-T(\rho)-0} H_p(a \circ \varphi_s) ds + \int_{-T(\rho)+0}^0 H_p(a \circ \varphi_s) ds \\ &= a \circ \varphi_{-T(\rho)-0}(\rho) - a \circ \varphi_{-T}(\rho) + a - a \circ \varphi_{-T(\rho)+0}(\rho) \end{aligned}$$

Now, since at  $x_1 = 0$ ,  $a(0, x', \xi_1, \xi') = a(0, x', -\xi_1, \xi')$ ,  $a \circ \varphi_{-T(\rho)-0} = a \circ \varphi_{-T(\rho)+0}$  and we have

$$\int_{-T}^0 H_p(a \circ \varphi_s)(\rho) ds = a(\rho) - a \circ \varphi_{-T}(\rho).$$

Now, by Lemma 4.9

$$\begin{aligned} \int \int_{-T}^0 H_p(a \circ \varphi_s)(\rho) ds d\mu(\rho) &= (H_p^* \mu) \left( \int_{-T}^0 a \circ \varphi_s ds \right) \\ &= \mu_1 \left( \int_{-T}^0 a \circ \varphi_s ds \right) + u^0 \left[ \left( \int_{-T}^0 a \circ \varphi_s ds \Big|_{\mathcal{H}} \right)_o \right]. \end{aligned}$$

Note that

$$\rho \mapsto \int_{-T}^0 a \circ \varphi_s ds \Big|_{\mathcal{H}}(\rho)$$

is even under  $\xi_1 \mapsto -\xi_1$ , since points  $(x = 0, y, \pm\xi, \eta)$  over the hyperbolic set are both mapped to the same point in  $T^*M^\circ$  under the short-time flow; thus the  $u^0$  term vanishes. Therefore,

$$\int_{-T}^0 \int a \circ \varphi_{-s}(\rho) d\mu_1 ds = \mu_1 \left( \int_{-T}^0 a \circ \varphi_s ds \right) = \int a - a \circ \varphi_{-T} d\mu.$$

□

### Invariance on the glancing set

We now turn to analysis at the glancing set. Our strategy will be to prove (1) for all  $t < t_0$  fixed and small. To do this, we may break up the support of  $q$  by a partition of unity and work locally, assuming that  $q$  is replaced by a function  $a$  supported in a neighborhood  $U$  of a point  $\rho \in \mathcal{G}$ ; we will show that for each such  $U$ , if  $t_0$  is taken to be small, (1) holds with  $q$  replaced by  $a$ .

We will do this using a ‘layer stripping’ argument. First, we ‘strip away’ the hyperbolic layer using the invariance established in Lemma 4.10. In particular, recall that  $\tilde{\mathcal{G}}$  denotes the set of generalized bicharacteristics in  $U$  which encounter  $\mathcal{G}$ . Then we have already seen that (1) holds on  $\tilde{\mathcal{G}}^c$ . Since  $\tilde{\mathcal{G}}^c$  is open,  $\pi_* 1_{\tilde{\mathcal{G}}^c} \mu$  satisfies

$$(4.11) \quad \pi_* 1_{\tilde{\mathcal{G}}^c} \mu(q \circ \varphi_t) - \pi_* 1_{\tilde{\mathcal{G}}^c} \mu(q) = \int_0^t \pi_* 2 \operatorname{Im} 1_{\tilde{\mathcal{G}}^c} \mu_1(q \circ \varphi_s) ds, \quad q \in \mathcal{C}_c^\infty({}^b T^* M).$$

In particular, since (1) implies (2),  $1_{\tilde{\mathcal{G}}^c} \mu$  satisfies

$$(4.12) \quad 1_{\tilde{\mathcal{G}}^c} \mu(H_p a) = 1_{\tilde{\mathcal{G}}^c} \mu_1(a).$$

Subtracting this from (4.5) for  $\mu$ ,  $\mu 1_{\tilde{\mathcal{G}}}$  satisfies

$$(4.13) \quad \mu 1_{\tilde{\mathcal{G}}}(H_p a) = 1_{\tilde{\mathcal{G}}} \mu_1(a).$$

Thus we may turn our attention to studying the new measures  $\mu 1_{\tilde{\mathcal{G}}}$ ,  $1_{\tilde{\mathcal{G}}} \mu_1$ . By a slight abuse of notation, then *we assume henceforth that  $\mu$  and  $\mu_1$  are supported on  $\tilde{\mathcal{G}}$ .*

As we proceed through the proof, we will ‘strip away’ the higher order tangent layers. We start by showing that at points where the bicharacteristics of  $H_p$  are tangent to exactly order 2, the measure is invariant in the sense of (1). Once we have shown this, we can argue as we did to remove the hyperbolic set and replace the measure  $\mu$  by  $\mu 1_{\tilde{\mathcal{G}}^{\geq 3}}$ . We then prove invariance on  $\tilde{\mathcal{G}}^k$  by induction. As before, the fact that (1) implies (2) allows us to assume that  $\mu$  is supported on  $\tilde{\mathcal{G}}^{\geq k}$ . Then, since the flow is transverse to  $\mathcal{G}^k$  we may work with a measure supported on  $\tilde{\mathcal{G}}^k$  when showing invariance at order  $k$ . We will obtain this invariance by using the transversality of the flow to  $\mathcal{G}^k$  to show that there is no jump in the measure across these points. In particular, this argument will show that  $\mu$  accumulates no mass across  $\mathcal{G}^k$  and thus that the invariance continues to hold.

### The case of $\mathcal{G}_d$

We choose the neighborhood of  $U$  of  $\rho \in \tilde{\mathcal{G}}_d$  so small that  $U \cap (\mathcal{H} \cup \mathcal{G} \setminus \mathcal{G}_d) = \emptyset$ . Next, suppose that  $a \in \mathcal{C}_c^\infty(U)$ . Recall that

$$\mu(H_p a) = \mu_1(a) + \mu(1_{x_1=0} \partial_{x_1} r a_o|_{\xi_1^2=r(x,\xi')}) + u^0(a_o),$$

with  $u^0$  supported in  $\mathcal{H} \cap \operatorname{supp} \mu$  and  $\mu(\mathcal{G}_d) = 0$ . Since  $\operatorname{supp} \mu \subset \tilde{\mathcal{G}}$  and  $\operatorname{supp} a \subset U$ , we have

$$\mu(H_p a) = \mu_1(a).$$

and hence (4.4) holds since the generalized bicharacteristics through  $\mathcal{G}_d$  are bicharacteristics of  $H_p$ .

### The case of $\mathcal{G}_g$

Let  $\tilde{\mathcal{G}}_g$  be the set of generalized bicharacteristics encountering  $\mathcal{G}_g$  in  $U$ . Recall that we may assume  $\mu$  and  $\mu_1$  are supported on  $\tilde{\mathcal{G}}_g$ ; may further choose  $U$  small enough so that  $U \cap \tilde{\mathcal{G}}_g \subset \mathcal{G}_g$ , hence we may in fact assume at this stage that  $\mu$  and  $\mu_1$  are supported by  $\mathcal{G}_g$ .

Let  $a \in C_c^\infty(U)$ . For  $q = q(x_1, x', x_1 \xi_1, \xi')$ ,  $1_{\mathcal{G}} H_p^G q = 1_{\mathcal{G}} H_p q$ . Moreover, since  $H_p^G x_1 = H_p^G \xi_1 = 0$ , letting  $a_0 = a|_{\xi_1=0}$ ,

$$(4.14) \quad 1_{\mathcal{G}} H_p^G a = 1_{\mathcal{G}} H_p^G a_0 = 1_{\mathcal{G}} H_p a_0.$$

Now, since on  $U$ ,  $\text{supp } \mu \subset \mathcal{G}_g$

$$(4.15) \quad \mu(H_p^G a) = \mu(1_{\mathcal{G}_g} H_p^G a) = \mu(1_{\mathcal{G}_g} H_p^G a_0) = \mu(H_p a_0) = \mu_1(a_0) = \mu_1(1_{\mathcal{G}_g} a_0) = \mu_1(a)$$

Since  $H_p^G$  generates the bicharacteristic flow in  $\mathcal{G}_g$ , this implies (4.4).

*Remark 4.11.* While it may seem at first that the argument used for  $\mathcal{G}_g$  applies directly to higher order tangencies, we observe that it does not. In fact, the equality  $\mu(H_p q) = \mu_1(q)$  for  $q = q(x, x_1 \xi_1, \xi')$  a priori only implies that  $\mu$  mass can either ‘stick’ to the boundary as in  $\mathcal{G}_g$  or instantly detach from the boundary as in  $\mathcal{G}_d$ . It does not rule out either case. In  $\mathcal{G}_g$ , the fact that trajectories which detach from the boundary instantly leave  $\overline{M}$ , while in  $\mathcal{G}_d$ , the fact that  $\mu(\mathcal{G}_d) = 0$ , rule out detaching from and sticking to the boundary at  $\mathcal{G}_g$  and  $\mathcal{G}_d$  respectively. In order to determine whether mass at a point of higher order tangency sticks or detaches, we will use the transversality of the flow to such points to show that mass can neither accumulate nor dissipate.

### Higher order tangencies

Recall that

$$\mathcal{G}^k := \{\rho \in \mathcal{G} \mid (H_p^j x_1)(\rho) = 0, \text{ for } 0 \leq j < k \text{ and } (H_p^k x_1)(\rho) \neq 0\}.$$

We have shown that (4.3) holds for  $a \in C_c^\infty(T^*M)$  with  $\text{supp } a \cap \mathcal{G}^k = \emptyset$  for  $k \geq 3$ . We now proceed by induction on  $k$ .

Let  $r_0(x', \xi') = r(0, x', \xi')$  and  $r_1(x', \xi') = \partial_{x_1} r(0, x', \xi')$ . Then note that by [29, Lemma 24.3.1],

$$\mathcal{G}^k = \{x_1 = 0 = 1 - r(0, x', \xi'), H_{r_0}^j r_1 = 0, 0 \leq j < k - 2, H_{r_0}^{k-2} r_1 \neq 0\}.$$

For each  $k$ , let  $\tilde{\mathcal{G}}^k$  denote the set of generalized bicharacteristics encountering  $\mathcal{G}^k$  in a small fixed neighborhood of  $\mathcal{G}^k$ .

Suppose (4.4) holds for  $q \in C_c^\infty(T^*M)$  with  $\text{supp } q \cap \tilde{\mathcal{G}}^k = \emptyset$  for  $k \geq M - 1$ . We show that (4.4) holds for  $a \in C_c^\infty(T^*M)$  with  $a \cap \tilde{\mathcal{G}}^k = \emptyset$  for  $k \geq M$ .

Let  $\rho \in \mathcal{G}^M$  and  $U$  a neighborhood of  $\rho$  so that  $U \cap \mathcal{G}^k = \emptyset$  for  $k > M$ . As before, we may assume  $\mu$  and  $\mu_1$  are supported on  $\tilde{\mathcal{G}}^{\geq M}$ ; thus by our choice of  $U$ , we may in fact assume without loss of generality that  $\mu, \mu_1$  are supported on  $\tilde{\mathcal{G}}^M$ . Let  $(x_2, \rho)$  be coordinates on  $T^*(\partial M)$  so that  $\partial_{x_2} = H_{r_0}$ . Shrinking  $U$  if necessary, by the implicit function theorem there exists  $\Theta(\rho)$  such that on  $U$ ,

$$\tilde{\mathcal{G}}^M \cap \mathcal{G}^M = \{(x, \xi) \in \tilde{\mathcal{G}}^M \mid x_2 = \Theta(\rho)\}.$$

Let

$$\begin{aligned} \tilde{\mathcal{G}}_0 &:= \tilde{\mathcal{G}}^M \cap \mathcal{G}^M, & \tilde{\mathcal{G}}^+ &:= \{(x, \xi) \in \tilde{\mathcal{G}}^M \mid x_2 > \Theta(\rho)\}, \\ \tilde{\mathcal{G}}^- &:= \{(x, \xi) \in \tilde{\mathcal{G}}^M \mid x_2 < \Theta(\rho)\}. \end{aligned}$$

We write

$$\begin{aligned} \mu 1_{\tilde{\mathcal{G}}^M} &= \mu_- + \mu_+ + \Upsilon, \\ \mu_- &= 1_{\tilde{\mathcal{G}}^-} \mu, & \mu_+ &= 1_{\tilde{\mathcal{G}}^+} \mu, & \Upsilon &= 1_{\tilde{\mathcal{G}}_0} \mu. \end{aligned}$$

Then, for  $q = q(x_1, x', x_1 \xi_1, \xi')$  supported in  $U$ ,

$$(4.16) \quad \Upsilon(H_p q) = -\mu_-(H_p q) - \mu_+(H_p q) + \mu_1(q).$$



Let

$$\tilde{H}(\rho) = \begin{cases} H_p(\rho) & \rho \notin \mathcal{G}_g \\ H_p^G(\rho) & \rho \in \mathcal{G}_g; \end{cases}$$

this vector field generates the bicharacteristic flow on  $\tilde{\mathcal{G}}^M$ , so by our inductive hypothesis,

$$(4.17) \quad \tilde{H}^* \mu = \mu_1$$

on  $\tilde{\mathcal{G}}^M \setminus \mathcal{G}^M$ . Moreover, since on  $\tilde{\mathcal{G}}^0$  we have  $H_p^M x_1 \neq 0$ , the flow within  $U$  (perhaps after further shrinking that set) for forward and backward time is either in  $\{x_1 > 0\}$  or lies entirely in  $\mathcal{G}_g$  (with separate cases for forward/backward flow depending on the parity of  $M$  and the sign of  $H_p^M x_1$ ). In either case, the flow stays away from  $\mathcal{H}$ , hence is generated by  $\tilde{H}$  on  $\tilde{\mathcal{G}}^M \cap U$ .

Now as in Lemma 3.10 and the surrounding discussion, given any continuous  $a$  defined on  $\mathcal{G}_0$  and  $\psi \in \mathcal{C}_c^\infty(\mathbb{R})$ , we extend  $a$  to be constant on the flow, and set

$$\mu_a^\sharp(\psi) = \mu(\psi(s)a(\rho));$$

here we are using  $\rho \in \mathcal{G}_0$  together with  $s \in \mathbb{R}$  as local (non-smooth) coordinates on  $\tilde{\mathcal{G}}^M$  defined by the flow, mapping  $(-\delta, \delta) \times \tilde{\mathcal{G}}_0$  homeomorphically to  $\tilde{\mathcal{G}}^M$  via  $(s, \rho) \mapsto \exp(t\tilde{H})(\rho)$ . Just as in Lemma 3.10, the flow invariance established on  $\tilde{\mathcal{G}}^M \setminus \mathcal{G}_0$  implies that  $\mu_a^\sharp$  is AC with respect to Lebesgue measure on  $\mathbb{R} \setminus 0$ ; moreover the Radon–Nikodym derivative is itself bounded by a multiple of  $\sup|a|$ , hence is a function of  $s$  with values in measures on  $\mathcal{G}_0$ , denoted  $G(s)$ . Finally, let  $\text{tr } \dot{\mu}_\pm = G(0^\pm)$  denote the restrictions of these functions to  $\mathcal{G}_0$  from left and right; these are well defined since the relation (4.17) together with the arguments in Lemma 3.11 show that  $G$  is continuous on  $s \leq 0$  and on  $s \geq 0$  separately but that there may be a jump at  $s = 0$ .

Moreover, denoting the function  $\psi(s)a(\rho)$  by  $\psi \otimes a$ , we have for  $0 \notin \text{supp } \psi$ ,  $\mu(\psi' \otimes a) = \mu_1(\psi \otimes a)$ . So in particular,

$$\int [G(a)](s)\psi'(s)ds = \mu_1(\psi \otimes a).$$

Now, let  $\chi \in \mathcal{C}_c^\infty(\mathbb{R})$  with  $\chi \equiv 1$  on  $[-1, 1]$  and  $\text{supp } \chi \subset (-2, 2)$  and  $\chi_\epsilon(s) = \chi(\epsilon^{-1}s)$ . Then,

$$\begin{aligned} \mu((1 - \chi_\epsilon(s))\tilde{H}(\psi \otimes a)) &= \mu(\tilde{H}[(1 - \chi_\epsilon)\psi] \otimes a) + \mu([\psi \otimes a]\tilde{H}\chi_\epsilon) \\ &= \mu(\tilde{H}[(1 - \chi_\epsilon)\psi \otimes a]) + \mu(\epsilon^{-1}\psi\chi'(\epsilon^{-1}\cdot) \otimes a) \\ &= \mu_1((1 - \chi_\epsilon)\psi \otimes a) + \mu(\epsilon^{-1}\psi\chi'(\epsilon^{-1}\cdot) \otimes a) \end{aligned}$$

Then,

$$\begin{aligned} \mu(\epsilon^{-1}\psi\chi'(\epsilon^{-1}\cdot) \otimes a) &= \epsilon^{-1} \int_{\mathbb{R} \setminus 0} [G(a)](s)\chi'(\epsilon^{-1}s)\psi(s)ds \\ &= \psi(0)(G_a(0^-) - G_a(0^+)) - \int_{\mathbb{R} \setminus 0} G_a(s)\chi_\epsilon(s)\psi'(s)ds + \mu_1(\chi_\epsilon\psi \otimes a) \\ &\rightarrow \psi(0)(G_a(0^-) - G_a(0^+)) + \mu_1((\psi \otimes a)1_{\tilde{\mathcal{G}}_0}) \end{aligned}$$

In particular,

$$(4.18) \quad \begin{aligned} \mu_+(\tilde{H}\psi \otimes a) + \mu_-(\tilde{H}\psi \otimes a) &= \psi(0)([G(a)](0^-) - [G(a)](0^+)) + \mu_1(\psi \otimes a) \\ &= \text{tr } \dot{\mu}_+(\psi \otimes a) - \text{tr } \dot{\mu}_-(\psi \otimes a) + \mu_1(\psi \otimes a). \end{aligned}$$

Now we claim that we may also apply (4.16) to the function  $q = \psi \otimes a$ . We begin by approximating  $a$  with a smooth function on  $\mathcal{G}_0$ . Since  $\tilde{H}$  locally generates the flow and is at least Lipschitz across  $\mathcal{G}_0$ , we find that the flow on  $\tilde{\mathcal{G}}^M$  is at least  $\mathcal{C}^1$ . We note further that  $\tilde{\mathcal{G}}^M$  is without loss of generality disjoint from  $\mathcal{H}$  (after appropriately shrinking  $U$ ) since depending

on the parity and sign of  $H_p^M x_1$  along  $\mathcal{G}_0$  the flow is either gliding or enters  $x_1 > 0$  in each of  $\tilde{\mathcal{G}}^\pm$ . In particular,  $\pi(\tilde{\mathcal{G}}^M) \subset {}^bT^*M$  is  $\mathcal{C}^1$ . Moreover,

$$(-\delta, \delta) \times \mathcal{G}_0 \ni (s, q) \mapsto \varphi_s(q) \in \pi(\tilde{\mathcal{G}}^M)$$

are  $\mathcal{C}^1$  coordinates on  $\tilde{\mathcal{G}}^M$ . Hence  $b = \pi^* \psi \otimes a$  is  $\mathcal{C}^1$  on  $\pi(\tilde{\mathcal{G}}^M)$  and we may extend  $b$  to  $\tilde{b}$ , a  $\mathcal{C}^1$  function on  ${}^bT^*M$ . Equivalently, setting  $\tilde{q} = \pi^* \tilde{b}$ , we have  $\tilde{q} = \tilde{b}(x, x_1 \xi_1, \xi')$  with  $\tilde{b}$  a  $\mathcal{C}^1$  function of its arguments and  $\tilde{q}|_{\tilde{\mathcal{G}}^M} = \psi \otimes a$ .

Since  $\tilde{b}$  is  $\mathcal{C}^1$ , we may approximate it in  $\mathcal{C}^1$  by smooth functions  $\tilde{b}_\epsilon$  and hence we obtain using the dominated convergence theorem that (4.16) holds for  $\tilde{q}$ . Finally, since  $\tilde{q}|_{\tilde{\mathcal{G}}^M} = \psi \otimes a$  and  $\text{supp } \mu \subset \tilde{\mathcal{G}}^M$ , we obtain (4.16) for  $q = \psi \otimes a$ . Moreover, since  $\mu_\pm$  are each supported in the interior of  $\mathcal{G}_g$ , we can replace  $\tilde{H}$  by  $H_p$  in (4.18).

Now (4.16) and (4.18) yield:

$$(4.19) \quad \Upsilon(H_p \psi \otimes a) = -\text{tr}(\dot{\mu}_+)(\psi \otimes a) + \text{tr}(\dot{\mu}_-)(\psi \otimes a)$$

and hence

$$|\Upsilon(H_p \psi \otimes a)| \leq C |\text{sup } \psi \otimes a|.$$

Since  $\Upsilon$  is supported at  $\tilde{\mathcal{G}}^0$ , and the collection of functions  $\psi \otimes a$  is dense, this implies that for all  $q$ ,

$$|\Upsilon(H_p q)| \leq C |\text{sup } q|.$$

Then, using that  $\Upsilon$  is a measure supported at  $\tilde{\mathcal{G}}_0$  and  $H_p$  is transverse to  $\tilde{\mathcal{G}}_0$ , this implies  $\Upsilon = 0$ . Finally, inserting  $\Upsilon = 0$  into (4.19) yields

$$\text{tr}(\dot{\mu}_-) = \text{tr}(\dot{\mu}_+).$$

Moreover, since  $\mu_1$  is absolutely continuous with respect to  $\mu$ ,  $\mu_1 1_{\tilde{\mathcal{G}}_0} = 0$ .

Let  $\psi_{t_0}(s) = \psi(s + t_0)$ . Then  $(\psi \otimes a) \circ \varphi_t = \psi_t \otimes a$  and

$$\begin{aligned} \mu(\psi_{t_0} \otimes a) - \mu(\psi \otimes a) &= \int_0^{t_0} \partial_t \int [G(a)](s) \psi(s+t) ds dt \\ &= \int_0^{t_0} \int_{-\infty}^0 [G(a)](s) \psi'(s+t) ds + \int_0^\infty [G(a)](s) \psi'(s+t) dt \\ &= \int_0^{t_0} \int_{-\infty}^\infty \mu_1(\psi_t \otimes a) ds + \text{tr } \dot{\mu}_-(a) \psi(t) - \text{tr } \dot{\mu}_+(a) \psi(t) dt \\ &= \int_0^{t_0} \mu_1(\psi_t \otimes a) dt \end{aligned}$$

In particular, we have

$$\mu((\psi \otimes a) \circ \varphi_{t_0}) - \mu(\psi \otimes a) = \int_0^{t_0} \mu_1([\psi \otimes a] \circ \varphi_t) dt.$$

Since the collection of functions of the form  $\psi \otimes a$  is dense in  $\mathcal{C}^0({}^bT^*M)$ , this completes the proof that  $\mu$  is invariant on  $\tilde{\mathcal{G}}^M$ .

As there are no infinite order tangencies,  $\mathcal{G} = \bigcup_M \mathcal{G}^M$  and this completes the proof of (4.4).  $\square$

Finally, to complete the proof of (1.3) we use Proposition 3.9 and the following lemma.

**Lemma 4.12.** *Suppose that  $\mu$  satisfies (4.3). Let  $\psi \in \mathcal{C}_c^1(\mathbb{R})$ ,  $\Sigma$  be a hypersurface transverse to the generalized bicharacteristic flow so that  $\varphi_t : \mathbb{R} \times \Sigma \rightarrow {}^bT^*M$  is a homeomorphism onto its image, and  $a \in \mathcal{C}_c^0(\Sigma)$ . Define  $f_{\psi,a}(\varphi_t(\rho)) = \psi(t)a(\rho)$ . Then*

$$\pi_* \mu(f_{\partial_t \psi, a}) = 2 \text{Im } \mu^j(f_{\psi, a}).$$

*Proof.* We have for  $q \in \mathcal{C}_c^\infty({}^bT^*M)$ ,

$$\partial_t \pi_* \mu(q \circ \varphi_t)|_{t=0} = 2 \operatorname{Im} \mu^j(q).$$

Therefore, it is enough to show that we can move the derivative inside  $\pi_* \mu$ . This follows from the dominated convergence theorem since  $\pi_* \mu(\mathcal{H}) = 0$ . Now approximation of  $f_{\psi,a}$  by smooth functions gives the result.  $\square$

In order to obtain the required uniformity, we adjust the set  $\mathcal{C}$  by fixing neighborhood  $U$  of  $\Omega$  in which we assume that for  $\{g_k\}_{k=1}^\infty \subset \mathcal{C}$ , we not only have  $\|g - g_{k_m}\|_{\mathcal{C}^{1,\gamma}} \rightarrow 0$  for some  $\gamma > 0$  but also,  $\|g - g_{k_m}\|_{\mathcal{C}^\infty(U)} \rightarrow 0$ . Arguing as in Section 3.5 then completes the proof of Theorem 2 when  $\Omega \neq \emptyset$ .

## 5. LOWER BOUNDS ON MANIFOLDS WITHOUT BOUNDARY

The idea behind our proof of the lower bounds in Theorem 1 is that near any segment of a geodesic  $\gamma$ , which is not trapped,  $P$  is *globally* microlocally equivalent to  $hD_{x_1}$  and hence it is enough to construct examples which saturate the  $L_{\text{comp}}^2 \rightarrow L_{\text{loc}}^2$  bounds for the solution operator for  $hD_{x_1}$  where we impose the condition that the solution be supported in  $x_1 \geq 0$ .

Let  $U$  be an open set. Assume there exists a null bicharacteristic curve  $\gamma$  such that

$$(5.1) \quad \{\pi(\gamma(s)), s \in (0, L)\} \subset U, \quad \pi(\gamma(s)) \notin U \text{ for } s \notin (0, L).$$

For any interval  $I$  (open or closed), let

$$\Gamma_I \equiv \{\gamma(s), s \in I\} \subset T^*M, \quad \Gamma_I^0 = \{x_1 \in I, \xi_1 = 0, x' = \xi' = 0\} \subset T^*\mathbb{R}^n$$

**Lemma 5.1.** *There exist neighborhoods  $V \subset T^*M$  of  $\Gamma_{[0,L]}$ ,  $U \subset T^*\mathbb{R}^n$  of  $\Gamma_{[0,L]}^0$  and a symplectomorphism  $\kappa : U \rightarrow V$  with*

$$\kappa : \Gamma_{[0,L]}^0 \rightarrow \Gamma_{[0,L]}, \quad \kappa^* p = \xi_1.$$

*Proof.* Let  $\rho_0 = \gamma(0)$ . Then by Darboux's theorem, there exist neighborhood  $V_1 \subset T^*M$  of  $\rho_0$ ,  $U_1 \subset T^*\mathbb{R}^n$  of 0 and a symplectomorphism  $\kappa_1 : U_1 \rightarrow V_1$  so that

$$\kappa_1^* p = \xi_1, \quad \kappa_1 : \Gamma_{[0,L]}^0 \cap U_1 \rightarrow \Gamma_{[0,L]} \cap V_1.$$

We will extend  $\kappa_1$  so that its image covers a neighborhood of  $\Gamma_{[0,L]}$ . For this, let  $\Sigma := \kappa_1(\{x_1 = 0\} \cap U_1)$ . Shrinking  $U_1$  if necessary, we assume that there is  $\epsilon > 0$  so that

$$(-\epsilon, L + \epsilon) \times \Sigma \ni (t, q) \mapsto \varphi_t(q) \in V \subset T^*M$$

is a diffeomorphism onto its image,  $V$ . Then, let

$$x_i(\varphi_t(q)) = x_i(\kappa_1(q)), \quad 1 < i \leq n, \quad \xi_i(\varphi_t(q)) = \xi_i(\kappa_1(q)), \quad 1 \leq i \leq n, \\ x_1(\varphi_t(q)) = t.$$

In particular, we have

$$H_p x_i = 0, \quad 1 < i \leq n, \quad H_p \xi_i = 0, \quad 1 \leq i \leq n, \quad H_p x_1 = 1.$$

By Jacobi's identity, for  $f, g \in \mathcal{C}^\infty(T^*M)$ ,

$$H_p \{g, f\} = \{g, H_p f\} - \{f, H_p g\}.$$

Therefore, since the corresponding identities hold at  $\Sigma$ ,

$$\{x_i, x_j\} = 0, \quad \{\xi_i, \xi_j\} = 0, \quad \{\xi_i, x_j\} = \delta_{ij}.$$

In particular, this implies that  $\kappa^{-1} : V \rightarrow T^*\mathbb{R}^n$

$$\kappa^{-1}(q) := (x(q), \xi(q))$$

is a symplectomorphism onto its image,  $U$ . Furthermore, since  $p$  is invariant under  $H_p$ ,  $\kappa^*p = \xi_1$  and hence also  $\kappa(\Gamma_{[0,L]}^0) = \Gamma_{[0,L]}$ .  $\square$

**Proposition 5.2.** *Let  $\gamma$ , a geodesic, and  $U \subset M$  satisfy (5.1). Then for any  $\epsilon > 0$  there exists  $\chi \in C_c^\infty(U)$  and  $f \in L^2$  with  $\text{supp } f \subset \{\chi = 1\}$  and  $u \in L_{\text{loc}}^2$  with*

$$Pu = f$$

and

$$\text{WF}_h u \subset \bigcup_{s \geq 0} [\exp(sH_p) \text{WF}_h f].$$

and where

$$\|\chi u\|_{L^2} \geq \left( \frac{2L - \epsilon}{\pi h} \right) \|f\|_{L^2}$$

Note that  $u$  is purely microlocally outgoing by the wavefront set statement, hence is given by the outgoing resolvent applied to  $f$  (modulo  $O(h^\infty)$ ).

*Proof.* We first let  $\kappa$  be a symplectomorphism from  $T^*\mathbb{R}^n$  into  $T^*M$  with

$$\kappa : \Gamma_{[0,L]}^0 \rightarrow \Gamma_{[0,L]}, \quad \kappa^*p = \xi_1,$$

where

$$\Gamma_I^0 = \{x_1 \in I, \xi_1 = 0, x' = \xi' = 0\};$$

this exists by Lemma 5.1.

Now quantize  $\kappa$  to a microlocally unitary FIO  $T : C_c^\infty(\mathbb{R}^n) \rightarrow C_c^\infty(M)$  with parametrix  $S$  such that  $ST - I, TS - I$  have no microsupport near  $\Gamma_{[0,L]}^0$  resp.  $\Gamma_{[0,L]}^0$  and such that

$$ThD_1 = PT + O(h^\infty) \text{ microlocally near } \Gamma_{[0,L]}^0.$$

(There are no obstructions to such a quantization as long as we work on a contractible neighborhood of  $\Gamma^0$ .) Assume without loss of generality that  $T$  is defined on a  $\delta_0$ -neighborhood of  $\Gamma_{[0,L]}^0$ .

Now for any  $\delta < \delta_0$  we define on  $\mathbb{R}^n$

$$f_0 = \begin{cases} \phi(x') \cos \pi(x_1 - \delta)/2(L - 2\delta), & x_1 \in [\delta, L - \delta], \\ 0, & \text{otherwise} \end{cases}$$

where  $\phi$  is smooth and compactly supported in  $B(0, \delta)$ , with

$$\int |\phi(x')|^2 dx' = 1.$$

Let

$$v_0 = \begin{cases} 0, & x_1 < \delta, \\ h^{-1} \frac{2(L-2\delta)}{\pi} \phi(x') \sin \pi(x_1 - \delta)/2(L - 2\delta), & x_1 \in [\delta, L - \delta], \\ h^{-1} \frac{2(L-2\delta)}{\pi} \phi(x'), & x_1 > L - \delta; \end{cases}$$

Thus of course  $hD_1 v_0 = f_0$  as distributions. Let  $\psi(x_1)$  be a smooth cutoff function supported in  $(0, L - \delta/4)$  and equal to 1 on  $[\delta, L - \delta/2]$ . Set  $u_0 = \psi(x_1)v_0$ ,  $w = Tu_0$ , and  $f = Tf_0$ . Then modulo  $O(h^\infty)$ ,

$$Pw \equiv PTu_0 \equiv ThD_{x_1} u_0 \equiv Tf_0 + Tr_0 \equiv f + r$$

where  $r_0$  is the error term coming from the  $hD_1(\psi)$  term. Note that

$$\text{WF } r_0 \subset \{x_1 \in [L - \delta/2, L - \delta/4], |x'| < \delta, \xi = 0\};$$

which is thus in an arbitrarily small neighborhood of the endpoint of  $\Gamma_{[0,L]}^0$ . Thus by the construction of  $T$ ,  $r \equiv Tr_0$  has wavefront set in an arbitrarily small neighborhood of  $\gamma(L)$  (as  $\delta \downarrow 0$ ).

Note also that we may construct a cutoff function  $\chi \in \mathcal{C}_c^\infty(U)$  such that

$$\chi = 1 \text{ on } \text{WF}_h f \text{ and } \chi = 0 \text{ on } \pi\{\exp(sH_p)(\text{WF}_h r), s \geq 0\} :$$

This follows from our hypotheses since on  $\mathbb{R}^n$   $f$  is supported on  $x_1 \in [\delta, L - \delta]$  in a small neighborhood of  $\Gamma^0$  while  $\text{WF}_h r_0 \subset x_1 \in [L - \delta/2, L - \delta/4]$ , and the forward flowout is thus contained (in these local coordinates) in  $\{x_1 \geq L - \delta/2\}$ . Since the Hamilton flow is non-radial and  $x_1$  maps to the flow parameter under  $\kappa$ , we know that  $\text{WF}_h r_0$  is separated from a neighborhood of  $\Gamma_{[0,L-\delta]}$  in *base* variables, not just in the cotangent bundle, and shrinking  $\delta$  if necessary we obtain the separation in the base variables. Note for use below that since  $\chi = 1$  on  $\text{WF}_h f$ , pulling back by  $T$  yields a fortiori  $\kappa^*(\chi) = 1$  on  $\text{WF}_h f_0$  and hence

$$(5.2) \quad \kappa^*(\chi) = 1 \text{ on } \{x_1 \in [\delta, L - \delta], |x'| < \delta, \xi = 0\}$$

Finally we set

$$u = w - (P - i0)^{-1}r$$

so that

$$Pu = f.$$

Owing to the propagation of singularities for the outgoing resolvent,  $\text{WF}_h(u - w)$  lies in the *forward flowout* of  $\text{WF}_h r$ ; by construction this  $\text{WF}_h r$  is disjoint from the support of  $\chi$  and by our geometric hypotheses, its forward flowout *remains* disjoint from  $U$  : here we use the fact that  $\pi\gamma(L + s) \notin U$  for  $s > 0$  and indeed this point escapes to infinity, hence  $\pi \exp(sH_p)(\gamma(L)) \notin \text{WF}_h f$  for all  $s > 0$ . By continuity, the same is true with  $\gamma(L)$  replaced by any point in  $\text{WF}_h r$  for  $\delta$  sufficiently small. Thus,

$$\chi u = \chi w + O(h^\infty).$$

By microlocal unitarity of  $T$  we compute by the Egorov Theorem

$$\begin{aligned} \|\chi u\|^2 &= \|\chi w\|^2 + O(h^\infty) \\ &= \|S\chi T S w\|^2 + O(h^\infty) \\ &= \|\text{Op}(\kappa^* \chi) u_0\|^2 + O(h) \\ &\geq \int_\delta^{1-\delta} \int_{B(0,\delta)} |v_0|^2 dx' dx_1 + O(h) \\ &\geq \int_\delta^{1-\delta} \int_{B(0,\delta)} h^{-2} \frac{4(L-2\delta)^2}{\pi^2} |\phi(x')|^2 \sin^2 \pi x_1 / 2(L-\delta) dx' dx_1 + O(h) \\ &= \frac{\pi}{4} \frac{1}{h^2 \mu^3} + O(h), \end{aligned}$$

where

$$\mu = \frac{\pi}{2(L-2\delta)};$$

here we have used the fact that  $\kappa^* \chi$  equals 1 on the zero section over  $[\delta, 1 - \delta] \times B(0, \delta)$ , while the semiclassical wavefront set of  $u_0$  lies within the zero section, hence the last inequality follows by existence of a microlocal parametrix for  $\text{Op}(\kappa^* \chi)$  on this set.

Meanwhile,

$$\|f\|^2 = \int_\delta^{1-\delta} \int_{B(0,\delta)} |\phi(x')|^2 \cos^2 \pi x_1 / 2(L-2\delta) dx' dx_1 = \frac{\pi}{4} \frac{1}{\mu}.$$

Thus

$$\frac{\|\chi u\|}{\|f\|} \geq \frac{1}{h\mu} = \frac{2(L-2\delta)}{\pi h}.$$

□

Finally, we apply Proposition 5.2 in the case of a nontrapping manifold with Euclidean ends and  $P = -h^2\Delta_g - 1$ . Let  $R_1 < R' < R''$  so that  $B(0, R'') \subset \{\chi \equiv 1\}$ . By the definition of  $L(g, R'')$ , there is a null-bicharacteristic (a speed-2 geodesic lifted to  $S^*M$ )  $\gamma$  with  $\pi\gamma(0) \in \partial B(0, R'')$ ,  $\pi\gamma(s) \in B(0, R'')$  for  $s \in (0, L(g, R''))$  and  $\gamma(s) \notin B(0, R'')$  for  $s > L(g, R'')$ . In particular, since  $L(g, R') < L(g, R'')$ , Proposition 5.2 applies to  $\gamma$  with  $U = B(0, R'')$  and  $\epsilon < 2L(g, R'') - 2L(g, R')$ . In particular, there exist  $f \in L^2(M)$  with  $\text{supp } f \subset B(0, R'')$  so that

$$(5.3) \quad Pu = hf, \quad \text{WF}_h u \subset \bigcup_{s \geq 0} [\exp(sH_p) \text{WF}_h f]$$

and

$$\|u\|_{L^2(B(0, R''))} \geq \frac{2L(g, R'') - 2L(g, R') + 2L(g, R')}{\pi} \|f\|_{L^2} = \frac{2L(g, R'')}{\pi} \|f\|_{L^2}.$$

Next, observe that (5.3) implies that

$$u = (-h^2\Delta_g - 1 - i0)^{-1}hf + O(h^\infty \|f\|_{L^2})_{L^2_{\text{loc}}}$$

In particular, letting  $k = h^{-1}$ ,

$$\|\chi(-\Delta_g - k^2 - i0)^{-1}f\|_{L^2} \geq \left( \frac{2L(g, R'')}{\pi k} - C_N k^{-N} \right) \|f\|_{L^2}$$

completing the proof of the lower bound in Theorem 1.

## 6. APPLICATION TO NUMERICAL ANALYSIS OF THE FINITE-ELEMENT METHOD

**6.1. Summary.** In this section, we focus on the implications the bound (1.8) has on the numerical analysis of solving the Helmholtz exterior Dirichlet problem by the finite-element method (FEM). We mention two other numerical-analysis applications of (1.8) in Remarks 6.8 and 6.9 at the end.

We consider the *h-version* of the FEM; i.e. the solution is approximated in spaces of piecewise polynomials of fixed degree on a mesh with meshwidth  $h_{\text{FEM}}$  (we use this notation to avoid a notational clash with the semiclassical parameter  $h := k^{-1}$  in the rest of the paper). The question of how fast  $h_{\text{FEM}}$  must decrease with  $k$  to maintain accuracy as  $k \rightarrow \infty$  was thoroughly investigated in the case of the *constant-coefficient* Helmholtz equation (i.e. (6.2) below with  $A = I$  and  $\nu = 1$ ) by [30, 31] when  $d = 1$  and by [33, 34, 35, 22] when  $d = 2, 3$ . For example, in the case of piecewise-linear polynomials, and when the solution of the boundary value problem is nontrapping, the FEM satisfies a *quasioptimal* error estimate (of the form (6.22) and (6.29) below) when

$$(6.1) \quad h_{\text{FEM}} k^2 \leq c$$

for a sufficiently small constant  $c$  (the case when the boundary value problem is trapping is more complicated; see [13, §1.4] for some initial results).

In this section, we use the bound (1.8) to prove the analogue of the result above in the case of the *variable-coefficient* Helmholtz equation ((6.2) below) posed in the exterior of a nontrapping Dirichlet obstacle (see Theorem 3 below). In particular, we show how the constant  $c$  in (6.1) depends on the constant in (1.8). The key point is that our bound on  $h_{\text{FEM}}$  shows explicitly how the constant  $c$  in (6.1) *decreases* (i.e. the requirement on  $h_{\text{FEM}}$  for quasioptimality becomes more stringent) as the length of the longest ray *increases*.

## 6.2. The Helmholtz exterior Dirichlet problem and FEM set-up.

**Motivation from applications.** Several important physical applications involve the PDE

$$(6.2) \quad \nabla \cdot (A \nabla u) + k^2 \nu u = -f$$

posed in  $\mathbb{R}^n \setminus \Omega$ , where  $A$  is a symmetric positive-definite matrix-valued function of position and  $\nu$  is a strictly-positive scalar-valued function of position. One example is via reductions of the time-harmonic Maxwell equations

$$(6.3) \quad \operatorname{curl} H + ik\varepsilon E = (ik)^{-1} J, \quad \operatorname{curl} E - ik\mu H = 0 \quad \text{in } \mathbb{R}^3,$$

when all the fields involved depend only on two Cartesian space variables, say  $x$  and  $y$ . In the transverse-magnetic (TM) mode,  $J$  and  $E$  are given by  $J = (0, 0, J_z(x, y))$  and  $E = (0, 0, E_z(x, y))$  so, when additionally the permittivity  $\varepsilon$  is a scalar and the permeability  $\mu$  satisfies

$$(6.4) \quad \mu = \begin{pmatrix} \tilde{\mu} & 0 \\ 0 & 1 \end{pmatrix}$$

for  $\tilde{\mu}$  a  $2 \times 2$  symmetric positive-definite matrix,  $E_z$  satisfies (6.2) with  $\nu = \varepsilon$ ,

$$(6.5) \quad A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^T (\tilde{\mu})^{-1} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

and  $f = J_z$ . Similarly, in the case of the transverse-electric (TE) mode,  $J$  and  $H$  are given by  $J = (J_x(x, y), J_y(x, y), 0)$  and  $H = (0, 0, H_z(x, y))$ , so that when  $\mu$  is a scalar and  $\varepsilon$  satisfies an equation analogous to (6.4),  $H_z$  satisfies (6.2) with  $A$  given by (6.5) with  $\tilde{\mu}$  replaced by  $\tilde{\varepsilon}$ ,  $\nu = \mu$ , and

$$f = -\frac{1}{ik} \nabla \cdot \left[ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (\tilde{\varepsilon})^{-1} \begin{pmatrix} J_x \\ J_y \end{pmatrix} \right].$$

Additionally, the so-called acoustic-approximation of the elastodynamic wave equation is (6.2) with  $A$  the inverse of the (scalar) density and  $\nu$  the inverse of the (scalar) bulk modulus; see, e.g., the derivation in [16, §1.2.6].

**Placing the PDE (6.2) in the framework of §1.1.** We let  $M := \mathbb{R}^n \setminus \Omega$ , where  $\Omega$  is a compact set such that its complement is connected,

$$g^{ij} := \frac{A_{ij}}{\nu}, \quad L_j := \sum_{i=1}^n \frac{g^{ij}}{\sqrt{\det g}} \partial_i (\sqrt{\det g}) + \sum_{i=1}^n \partial_i (g^{ij}) - \sum_{i=1}^n \frac{1}{\nu} \partial_i (A_{ij}), \quad L := 0,$$

and

$$P(g)\phi := -\frac{1}{\nu} \sum_{i,j=1}^n \partial_i (A_{ij} \partial_j \phi),$$

and then observe that (1.6) is satisfied, since

$$\Delta_g \phi := \sum_{i,j=1}^n \frac{1}{\sqrt{\det g}} \partial_i (\sqrt{\det g} g^{ij} \partial_j \phi).$$

With these definitions, the PDE (6.2) is  $(P(g) - k^2)u = f/\nu$ . The operator  $P_\Omega$  is then self-adjoint with respect to  $L^2(\mathbb{R}^n \setminus \Omega; \nu)$  and the  $H^1$  norm defined by (1.7) becomes

$$(6.6) \quad \|\phi\|_{H_\nu^1(\mathbb{R}^n \setminus \Omega)}^2 := \int_{\mathbb{R}^n \setminus \Omega} ((A \nabla \phi) \cdot \overline{\nabla \phi} + k^2 \nu |\phi|^2) dx.$$

*In this section (§6) only, to make contact with the standard numerical-analysis literature, we use non-semiclassically scaled Sobolev spaces; i.e., the norms on the spaces  $H^s$  do not include*

powers of  $k$  unless otherwise indicated. Indeed, in this section we write (6.6) as the weighted  $H^1$  norm

$$(6.7) \quad \|\phi\|_{H_{k,A,\nu}^1(\mathbb{R}^n \setminus \Omega)}^2 := \|A^{1/2} \nabla \phi\|_{L^2(\mathbb{R}^n \setminus \Omega)}^2 + k^2 \|\nu^{1/2} \phi\|_{L^2(\mathbb{R}^n \setminus \Omega)}^2,$$

and we use analogous notation for norms over subsets of  $\mathbb{R}^n \setminus \Omega$ . We also define the norms

$$\|\phi\|_{H^m(\mathbb{R}^n \setminus \Omega)}^2 := \sum_{|\alpha| \leq m} \int_{\mathbb{R}^n \setminus \Omega} |\partial^\alpha \phi|^2 dx;$$

i.e., if there are no  $A, \nu$  in the subscript, the  $H^m$  norm is the ‘standard’  $H^m$  norm on  $\mathbb{R}^n$ .

By (1.1),  $L(\nu A^{-1}, \Omega, R)$  is the length of the longest generalized bicharacteristic (informally, ray) in  $B(0, R) \setminus \Omega$ . Recall that, in the case  $\Omega = \emptyset$ , the (generalized) bicharacteristics are the projections in  $x$  of the solutions  $(x(s), \xi(s)) \in \mathbb{R}^n \times \mathbb{R}^n$  of the Hamiltonian system

$$(6.8) \quad \frac{dx_i}{ds}(s) = \frac{\partial}{\partial \xi_i} H(x(s), \xi(s)), \quad \frac{d\xi_i}{ds}(s) = -\frac{\partial}{\partial x_i} H(x(s), \xi(s)),$$

where the Hamiltonian  $H(x, \xi)$  is given by

$$(6.9) \quad H(x, \xi) := \frac{1}{\nu(x)} \sum_{i=1}^n \sum_{j=1}^n A_{ij}(x) \xi_i \xi_j - 1.$$

**Exterior Dirichlet problem and its variational formulation.** With  $\Omega$  a compact set such that its complement is connected, we define  $\Omega_+ := \mathbb{R}^n \setminus \Omega$ . Since the vast majority of numerical-analysis applications of the Helmholtz equation are in two and three dimensions, we restrict attention to  $d = 2, 3$ . Let  $\gamma : H_{\text{loc}}^1(\Omega) \rightarrow H^{1/2}(\partial\Omega)$  be the trace operator. Let  $A$  be a symmetric positive-definite matrix-valued function of position such that  $\text{supp}(I - A)$  is a compact subset of  $\mathbb{R}^n$ . Let  $\nu$  be a strictly-positive scalar-valued function of position such that  $\text{supp}(1 - \nu)$  is a compact subset of  $\mathbb{R}^n$ . Let  $A_{\min}$  and  $A_{\max}$  be such that

$$(6.10) \quad A_{\min} \leq A(x) \leq A_{\max} \quad \text{for all } x \in \Omega_+, \text{ in the sense of quadratic forms.}$$

and  $\nu_{\min}$  and  $\nu_{\max}$  be such that

$$(6.11) \quad \nu_{\min} \leq \nu(x) \leq \nu_{\max} \quad \text{for all } x \in \Omega_+.$$

We consider solving *both* the exterior Dirichlet problem

$$(6.12a) \quad \nabla \cdot (A \nabla u) + k^2 \nu u = -f \quad \text{in } \Omega_+,$$

$$(6.12b) \quad \gamma u = 0 \quad \text{on } \partial\Omega, \quad \text{and}$$

$$(6.12c) \quad \frac{\partial u}{\partial r}(x) - iku(x) = o\left(\frac{1}{r^{(d-1)/2}}\right) \quad \text{as } r := |x| \rightarrow \infty, \text{ uniformly in } \hat{x} := x/r,$$

and the sound-soft scattering problem

$$(6.13a) \quad \nabla \cdot (A \nabla u) + k^2 \nu u = 0 \quad \text{in } \Omega_+,$$

$$(6.13b) \quad \gamma u = 0 \quad \text{on } \partial\Omega, \text{ and}$$

$$(6.13c) \quad (u - u^I) \text{ satisfies the radiation condition (6.12c),}$$

where  $u^I$  is solution of  $\Delta u^I + k^2 u^I = 0$  (such as a plane wave or point source) that is smooth in a neighborhood of  $\text{supp}(I - A) \cup \text{supp}(1 - \nu) \cup \Omega$ .

The standard variational formulations of these problems are posed on  $\Omega_R := \Omega_+ \cap B(0, R) = B(0, R) \setminus \Omega$ , where  $R$  is chosen large enough such that  $\text{supp}(I - A)$ ,  $\text{supp}(1 - \nu)$ , and  $\text{supp } f$  are all compactly contained in  $B(0, R)$ , and also such that  $u^I$  is smooth in  $B(0, R)$ . Let

$$(6.14) \quad H_{0,D}^1(\Omega_R) := \{v \in H^1(\Omega_R) : \gamma v = 0 \text{ on } \partial\Omega\},$$



let  $\Gamma_R := \partial B(0, R)$ , and let  $T_R : H^{1/2}(\Gamma_R) \rightarrow H^{-1/2}(\Gamma_R)$  be the Dirichlet-to-Neumann (DtN) map for the equation  $\Delta u + k^2 u = 0$  posed in the exterior of  $B(0, R)$  with the Sommerfeld radiation condition (6.12c); the definition of  $T_R$  in terms of Hankel functions and polar coordinates (when  $d = 2$ )/spherical polar coordinates (when  $d = 3$ ) is given in, e.g., [12, Equations 3.5 and 3.6] [39, §2.6.3], [34, Equations 3.7 and 3.10]. The variational formulation of (6.12) is:

$$(6.15) \quad \text{find } u \in H_{0,D}^1(\Omega_R) \text{ such that } a(u, v) = F(v) \text{ for all } v \in H_{0,D}^1(\Omega_R),$$

where

$$(6.16) \quad a(u, v) := \int_{\Omega_R} \left( (A \nabla u) \cdot \overline{\nabla v} - k^2 \nu u \bar{v} \right) dx - \langle T_R(\gamma u), \gamma v \rangle_{\Gamma_R} \quad \text{and} \quad F(v) := \int_{\Omega_R} f \bar{v} dx,$$

and where  $\langle \cdot, \cdot \rangle_{\Gamma_R}$  denotes the duality pairing on  $\Gamma_R$  that is linear in the first argument and antilinear in the second. The variational formulation of (6.13) is (6.15) but with  $F(v)$  instead given by

$$F(v) := \int_{\Gamma_R} \left( \frac{\partial u^I}{\partial r} - T_R u^I \right) \gamma \bar{v} ds.$$

**Finite-element method.** Let  $\mathcal{T}_{h_{\text{FEM}}}$  be a family of triangulations of  $\Omega_R$  (in the sense of, e.g., [19, Page 67]) that is shape regular (see, e.g., [6, Definition 4.4.13], [19, Page 128]). Let

$$\mathcal{H}_{h_{\text{FEM}}} := \{v \in \mathcal{C}(\overline{\Omega_R}) : v|_K \text{ is a polynomial of degree 1 for each } K \in \mathcal{T}_{h_{\text{FEM}}}\},$$

and observe that the dimension of  $\mathcal{H}_{h_{\text{FEM}}}$  is proportional to  $h_{\text{FEM}}^{-n}$ . Our main result, Theorem 3 below, is valid when  $\Omega$  is  $\mathcal{C}^\infty$  (or, more precisely,  $\mathcal{C}^m$  for some large  $m$  not given explicitly). For such  $\Omega$  it is not possible to fit  $\partial\Omega$  exactly with simplicial elements (i.e. when each element of  $\mathcal{T}_{h_{\text{FEM}}}$  is a simplex), and fitting  $\partial\Omega$  with isoparametric elements (see, e.g., [19, Chapter VI]) or curved elements (see, e.g., [5]) is impractical, and therefore some analysis of non-conforming error is necessary; since this is very standard (see, e.g., [6, Chapter 10]), we ignore this issue here.

The finite-element method for the variational problem (6.15) is the Galerkin method applied to the variational problem (6.15), i.e.

$$(6.17) \quad \text{find } u_{h_{\text{FEM}}} \in \mathcal{H}_{h_{\text{FEM}}} \text{ such that } a(u_{h_{\text{FEM}}}, v_{h_{\text{FEM}}}) = F(v_{h_{\text{FEM}}}) \text{ for all } v_{h_{\text{FEM}}} \in \mathcal{H}_{h_{\text{FEM}}}.$$

**6.3. Main result.** Before stating the main result (Theorem 3 below) we define the following constants – all independent of  $k$  but dependent on one or more of  $A$ ,  $\nu$ ,  $\Omega$ ,  $R$ , and  $k_0$  – upon which the main result depends.

$C_{\text{int}}$ : Recall that the *nodal interpolant*  $I_{h_{\text{FEM}}} : \mathcal{C}(\overline{\Omega_R}) \rightarrow \mathcal{H}_{h_{\text{FEM}}}$  is well-defined for functions in  $H^2(\Omega_R)$  (for  $d = 2, 3$ ) and satisfies

$$(6.18) \quad \left\| \nu^{1/2}(v - I_{h_{\text{FEM}}} v) \right\|_{L^2(\Omega_R)} + h_{\text{FEM}} \left\| A^{1/2} \nabla(v - I_{h_{\text{FEM}}} v) \right\|_{L^2(\Omega_R)} \leq C_{\text{int}} (h_{\text{FEM}})^2 \|v\|_{H^2(\Omega_R)},$$

for all  $v \in H^2(\Omega_R)$ , for some constant  $C_{\text{int}} = C_{\text{int}}(A, \nu)$ . The only reason  $C_{\text{int}}$  depends on  $A$  and  $\nu$  is because of the weights  $A$  and  $\nu$  in the norms on the left-hand side of (6.18). Indeed, by, e.g., [6, Equation 4.4.28],

$$\|v - I_{h_{\text{FEM}}} v\|_{L^2(\Omega_R)} + h_{\text{FEM}} \|\nabla(v - I_{h_{\text{FEM}}} v)\|_{L^2(\Omega_R)} \leq \tilde{C}_{\text{int}} (h_{\text{FEM}})^2 \|v\|_{H^2(\Omega_R)},$$

for all  $v \in H^2(\Omega_R)$ , for some  $\tilde{C}_{\text{int}}$  that depends only on the shape-regularity constant of the mesh, and thus (6.18) holds with

$$C_{\text{int}}(A, \nu) := \tilde{C}_{\text{int}} \max \left\{ (A_{\text{max}})^{1/2}, (\nu_{\text{max}})^{1/2} \right\}.$$

$C_{\text{DtN}}$ : There exists  $C_{\text{DtN}} = C_{\text{DtN}}(A, \nu, R, k_0)$  such that

$$(6.19) \quad \left| \langle T_R(\gamma u), \gamma v \rangle_{\Gamma_R} \right| \leq C_{\text{DtN}} \|u\|_{H_{k,A,\nu}^1(\Omega_R)} \|v\|_{H_{k,A,\nu}^1(\Omega_R)} \quad \text{for all } u, v \in H^1(\Omega_R) \text{ and for all } k \geq k_0;$$

see [34, Lemma 3.3]. As above, the only reason  $C_{\text{DtN}}$  depends on  $A$  and  $\nu$  is because of the weights  $A$  and  $\nu$  in (6.7). Indeed, [34, Lemma 3.3] bounds the left-hand side of (6.19) by  $H^{1/2}(\Gamma_R)$  and  $L^2(\Gamma_R)$  norms of  $\gamma u$  and  $\gamma v$ , and then uses trace theorems to prove that

$$\left| \langle T_R(\gamma u), \gamma v \rangle_{\Gamma_R} \right| \leq \tilde{C}_{\text{DtN}}(R, k_0) \|u\|_{H_{k,I,1}^1(\Omega_R)} \|v\|_{H_{k,I,1}^1(\Omega_R)},$$

for some  $\tilde{C}_{\text{DtN}}(R, k_0)$  given explicitly in the proof of [34, Lemma 3.3]. Therefore, (6.19) holds with

$$C_{\text{DtN}}(A, \nu, R, k_0) := \tilde{C}_{\text{DtN}}(R, k_0) \max \left\{ \frac{1}{(A_{\min})^{1/2}}, \frac{1}{(\nu_{\min})^{1/2}} \right\}.$$

$C_{H^2}$ : There exists  $C_{H^2} = C_{H^2}(A, \Omega, R)$  such that, if  $\tilde{f} \in L^2(\Omega_{R+1})$  and  $v \in H^1(\Omega_{R+1})$  satisfy

$$\nabla \cdot (A \nabla v) = \tilde{f} \quad \text{in } \Omega_{R+1} \text{ and } \gamma v = 0 \text{ on } \partial\Omega,$$

then

$$(6.20) \quad \|v\|_{H^2(\Omega_R)} \leq C_{H^2} \left( \|A^{1/2} \nabla v\|_{L^2(\Omega_{R+1})} + \|v\|_{L^2(\Omega_{R+1})} + \|\tilde{f}\|_{L^2(\Omega_{R+1})} \right);$$

since  $A \in \mathcal{C}^{0,1}$  and  $\Omega \in \mathcal{C}^{1,1}$ , such a  $C_{H^2}$  exists by, e.g., [32, Theorem 4.18] and  $C_{H^2}$  depends on  $A_{\min}$  in (6.10), the  $\mathcal{C}^{0,1}$  norm of  $A$ , and the  $\mathcal{C}^{1,1}$  norm of the parametrisation of  $\partial\Omega$ . Note that the  $R+1$  in the norms on the right-hand side of (6.20) can be replaced by  $R+b$  for any  $b > 0$ , but the constant  $C_{H^2}$  blows up as  $b \rightarrow 0$ .

**Theorem 3** (Quasioptimality of the Galerkin method). *Let  $\Omega$ ,  $A$ , and  $\nu$  be as in §6.2; i.e.  $\Omega$  a compact set such that its complement  $\Omega_+$  is connected,  $A$  a symmetric positive-definite matrix-valued function of position such that  $\text{supp}(I - A)$  is a compact subset of  $B(0, R)$  and  $\nu$  a strictly-positive scalar-valued function of position such that  $\text{supp}(1 - \nu)$  is a compact subset of  $B(0, R)$ . Furthermore, let  $\Omega$  be  $\mathcal{C}^\infty$ , and let  $A, \nu \in \mathcal{C}^{1,1}(\Omega_R)$  be such that  $A$  and  $\nu$  are  $\mathcal{C}^\infty$  in a neighborhood of  $\partial\Omega$ . Finally assume that  $\Omega$  is nowhere tangent to infinite order to the geodesic flow generated by  $A$  and  $\nu$  via the Hamiltonian (6.9), and that this flow on  $\Omega_+$  is nontrapping.*

*There exists a  $k_0 > 0$  such that if*

$$(6.21) \quad 1 \geq h_{\text{FEM}} k^2 \sqrt{1 + (h_{\text{FEM}} k)^2} L(\nu A^{-1}, \Omega, R+2) C_{\text{int}} C_{H^2} (1 + C_{\text{DtN}}) \left( \frac{\nu_{\max}}{\nu_{\min}} \right)^{1/2} \frac{4\sqrt{2}}{\pi} \\ \times \left( (\nu_{\max})^{1/2} + \frac{1 + (\nu_{\min})^{1/2}}{k_0} + \frac{1}{k_0^2 (\nu_{\min})^{1/2}} \right),$$

*then the Galerkin equations (6.17) have a unique solution which satisfies*

$$(6.22) \quad \|u - u_{h_{\text{FEM}}}\|_{H_{k,A,\nu}^1(\Omega_R)} \leq 2(1 + C_{\text{DtN}}) \left( \min_{v_{h_{\text{FEM}}} \in \mathcal{H}_{h_{\text{FEM}}}} \|u - v_{h_{\text{FEM}}}\|_{H_{k,A,\nu}^1(\Omega_R)} \right).$$

*(Note that  $k_0$  depends on  $A$ ,  $\nu$ , and  $\Omega$  but—as in Theorem 2—is uniform on small  $\mathcal{C}^{2,\alpha}$  neighborhoods of  $A$  and  $\nu$ ; see Section 3.5.)*

*Remark 6.1* (The mesh threshold (6.21)). If one assumes that  $h_{\text{FEM}} k \leq C$ , then the mesh threshold (6.21) can be written in the form (6.1), i.e.  $h_{\text{FEM}} k^2 \leq c$ . The key point is that if  $A$ ,  $\nu$ , and  $\Omega$  are as in the statement of the theorem with the  $\mathcal{C}^{0,1}$  norms of  $A$  and  $\nu$  and the  $\mathcal{C}^{1,1}$  norm of  $\Omega$  all bounded by  $\tilde{C}$ , say, then all the constants in  $c$  apart from  $L(\nu A^{-1}, \Omega, R+2)$  (namely,  $A_{\max}$ ,  $C_{\text{int}}$ ,  $C_{H^2}$ ,  $C_{\text{DtN}}$ ) are bounded in terms of  $\tilde{C}$ ,  $A_{\min}$  and  $\nu_{\min}$ , but  $L(\nu A^{-1}, \Omega, R+2)$  can be arbitrarily large. Furthermore,  $c$  decreases as  $L(\nu A^{-1}, \Omega, R+2)$  increases; i.e., the condition

on the mesh threshold for quasioptimality becomes more restrictive as the length of the longest ray grows.

*Remark 6.2* (Quasioptimality for higher-order finite-element spaces). We have only considered the lowest-order conforming finite-element space of  $H^1$ , namely continuous piecewise-linear polynomials. In the case of the constant-coefficient Helmholtz equation, the mesh threshold for quasioptimality (analogous to (6.1)) has been determined by [34, 35, 22] for spaces of arbitrary order  $p$  (possibly depending on  $k$ ), with the threshold then explicit in  $k$ ,  $h$ , and  $p$ . These results are obtained by a careful splitting of the solution that does not immediately generalise to the case  $A \neq I$ . However, in the case of  $p$  fixed, the mesh thresholds for quasioptimality from [34, 35, 22] have recently been obtained in [18] using a simpler method (relying on well-known elliptic regularity results such as (6.20)), although the constants are not given explicitly in  $p$ . In the case when the DtN map is approximated by an impedance boundary condition, the results of [18] immediately apply to the case  $A \neq I$ , and they can in principle be extended to the full scattering problem considered here.

*Remark 6.3* (Comparison to existing results in the literature). The only existing results in the literature on quasioptimality (explicit in all parameters and coefficients) for the Galerkin method applied to the variable-coefficient Helmholtz equation are in [17] and [27]. The main similarity between these two works and the present paper is that they all use the ‘‘Schatz argument’’ described in Lemma 6.6 below (and pioneered for Helmholtz problems in [44, 3]). The main differences between [17, 27] and the present paper are that (i) both [17] and [27] consider the Helmholtz equation in a bounded domain ([17] in 1-d, [27] in 1-, 2-, or 3-d) with an impedance boundary condition (recall that this is the simplest-possible approximation to the DtN map operator  $T_R$ ), and (ii) to get an a priori bound explicit in  $k$  and the coefficients, both [17] and [27] impose conditions on the coefficients and the domain that are *stronger* than the analogue of nontrapping for the interior impedance problem (i.e., the assumption that every ray reaches the boundary in a uniform time).

**6.4. Proof of Theorem 3.** The heart of the proof is Lemma 6.6 below. This gives a condition for quasi-optimality to hold in terms of how well the solution of the adjoint problem is approximated by the finite-element space, and relies on the fact that  $a(\cdot, \cdot)$  satisfies a Gårding inequality. This argument essentially goes back to Schatz [45] (using the Aubin-Nitsche technique; see, e.g., the references in [46, Remark 26]) and was extensively used in the analysis of the FEM for Helmholtz problems by [44, 34, 35, 22].

Before stating Lemma 6.6 we need to introduce some notation. Let  $C_{\text{cont}} = C_{\text{cont}}(A, \nu, R, k_0)$  be the *continuity constant* of the sesquilinear form  $a(\cdot, \cdot)$  (defined in (6.16)) in the norm  $\|\cdot\|_{H_{k,A,\nu}^1(\Omega_R)}$ ; i.e.

$$a(u, v) \leq C_{\text{cont}} \|u\|_{H_{k,A,\nu}^1(\Omega_R)} \|v\|_{H_{k,A,\nu}^1(\Omega_R)} \quad \text{for all } u, v \in H_{0,D}^1(\Omega_R).$$

By the Cauchy-Schwarz inequality and (6.19) we have

$$(6.23) \quad C_{\text{cont}} \leq 1 + C_{\text{DtN}}.$$

*Definition 6.4* (The adjoint sesquilinear form  $a^*(\cdot, \cdot)$ ). The adjoint sesquilinear form,  $a^*(u, v)$ , to the sesquilinear form  $a(\cdot, \cdot)$  defined in (6.16) is given by

$$(6.24) \quad a^*(u, v) := \overline{a(v, u)} = \int_{\Omega_R} \left( (A \nabla u) \cdot \overline{\nabla v} - k^2 \nu u \overline{v} \right) - \langle \gamma u, T_R(\gamma v) \rangle_{\Gamma_R}.$$

**Lemma 6.5.** *Given  $F \in (H_{0,D}^1(\Omega_R))'$ , if  $u$  is the solution to the variational problem*

$$(6.25) \quad a^*(u, v) = F(v) \quad \text{for all } v \in H_{0,D}^1(\Omega_R),$$

then  $\bar{u}$  satisfies

$$(6.26) \quad a(\bar{u}, v) = \overline{F(\bar{v})} \quad \text{for all } v \in H_{0,D}^1(\Omega_R).$$

*Proof of Lemma 6.5.* By (6.25),

$$\overline{a^*(u, \bar{v})} = \overline{F(\bar{v})} \quad \text{for all } v \in H_{0,D}^1(\Omega_R).$$

The result then follows from the definition of  $a^*(\cdot, \cdot)$  and the following property of the DtN map  $T_R$ :

$$\langle T_R \psi, \bar{\phi} \rangle_\Gamma = \langle T_R \phi, \bar{\psi} \rangle_\Gamma \quad \text{for all } \phi, \psi \in H^{1/2}(\Gamma_R).$$

This property follows from the fact that, if  $u$  and  $v$  are solutions of the homogeneous Helmholtz equation  $\Delta u + k^2 u = 0$  in  $\mathbb{R}^n \setminus \overline{B(0, R)}$ , both satisfying the Sommerfeld radiation condition (6.12c), then

$$\int_{\Gamma_R} (\gamma u) \frac{\partial v}{\partial n} = \int_{\Gamma_R} (\gamma v) \frac{\partial u}{\partial n}$$

by Green's second identity; see, e.g., [46, Lemma 6.13].  $\square$

**Lemma 6.6** (Conditions for quasi-optimality). *Given  $f \in L^2(\Omega_R)$ , let  $S^* f$  be the solution of the adjoint problem (6.25) with  $F(v)$  defined in (6.16). Let*

$$(6.27) \quad \eta(\mathcal{H}_{h_{\text{FEM}}}) := \sup_{0 \neq f \in L^2(\Omega_R)} \min_{v_{h_{\text{FEM}}} \in \mathcal{H}_{h_{\text{FEM}}}} \frac{\|S^* f - v_{h_{\text{FEM}}}\|_{H_{k,A,\nu}^1(\Omega_R)}}{\|f\|_{L^2(\Omega_R)}}.$$

If

$$(6.28) \quad \eta(\mathcal{H}_{h_{\text{FEM}}}) \leq \frac{1}{2C_{\text{cont}}(\nu_{\max})^{1/2} k}$$

then the Galerkin equations (6.17) have a unique solution which satisfies

$$(6.29) \quad \|u - u_{h_{\text{FEM}}}\|_{H_{k,A,\nu}^1(\Omega_R)} \leq 2C_{\text{cont}} \left( \min_{v_{h_{\text{FEM}}} \in \mathcal{H}_{h_{\text{FEM}}}} \|u - v_{h_{\text{FEM}}}\|_{H_{k,A,\nu}^1(\Omega_R)} \right).$$

*Proof.* Since

$$(6.30) \quad \text{Re}(-\langle T_R \phi, \phi \rangle_{\Gamma_R}) \geq 0 \quad \text{for all } \phi \in H^{1/2}(\Gamma_R) \text{ and for all } k \geq k_0$$

(see [12, Corollary 3.1] or [39, Theorem 2.6.4]),  $a(\cdot, \cdot)$  satisfies the Gårding inequality

$$\text{Re}(a(v, v)) \geq \|v\|_{H_{k,A,\nu}^1(\Omega_R)}^2 - 2k^2 \nu_{\max} \|v\|_{L^2(\Omega_R)}^2$$

and the result follows from the account of the Schatz argument in, e.g., [46, Theorem 6.32].  $\square$

The condition (6.28) on  $\eta(\mathcal{H}_{h_{\text{FEM}}})$  is implicitly a condition on the meshwidth  $h_{\text{FEM}}$ . To make this condition explicit, we observe that the polynomial-approximation bound (6.18) implies that a bound on  $\eta(\mathcal{H}_{h_{\text{FEM}}})$  (defined by (6.27)) can be obtained from a bound on the  $H^2$  norm of the (adjoint of the) exterior Dirichlet problem.

**Lemma 6.7** (Bound on  $H^2$  norm). *If  $\Omega$ ,  $A$ , and  $\nu$  are as in Theorem 3, then there exists a  $k_0 > 0$  such that the solution of the exterior Dirichlet problem with  $f$ ,  $I - A$ , and  $1 - \nu$  supported in  $\Omega_R$  satisfies*

$$(6.31) \quad \|u\|_{H^2(\Omega_R)} \leq k C_{H^2} \frac{2\sqrt{2}}{\pi(\nu_{\min})^{1/2}} L(\nu A^{-1}, \Omega, R+2) \left( (\nu_{\max})^{1/2} + \frac{1 + (\nu_{\min})^{1/2}}{k_0} + \frac{1}{k_0^2 (\nu_{\min})^{1/2}} \right) \|f\|_{L^2(\Omega_R)}.$$

*Proof.* Choosing  $\chi$  such that  $\text{supp } \chi \subset B_{R+2}$  but  $\chi = 1$  on  $B_{R+1}$ , Theorem 2 with  $s = 1$  implies that there exists  $k_0 > 0$  such that

$$(6.32) \quad \|A^{1/2} \nabla u\|_{L^2(\Omega_{R+1})}^2 + k^2 \|\nu^{1/2} v\|_{L^2(\Omega_{R+1})}^2 \leq \left( \frac{2\sqrt{2}}{\pi} L(\nu A^{-1}, \Omega, R+2) \right)^2 \frac{1}{\nu_{\min}} \|f\|_{L^2(\Omega_+)}^2,$$

for all  $k \geq k_0$ . Since the  $\chi$  in our argument can be chosen independent of  $A$  and  $\nu$ , the constant  $k_0$  can be chosen uniformly as  $A$  and  $\nu$  vary within a sufficiently small open neighborhood in  $C^{2,\alpha}$  ( $\alpha > 0$ ), just as in Theorem 2.

We then apply (6.20) with  $v = u$ ,  $\tilde{f} = -k^2 \nu u - f$ , and recall that  $\text{supp } f \subset \Omega_R$ , to obtain

$$\|u\|_{H^2(\Omega_R)} \leq C_{H^2} \left( \frac{2\sqrt{2}}{\pi(\nu_{\min})^{1/2}} L(\nu A^{-1}, \Omega, R+2) \left( k(\nu_{\max})^{1/2} + 1 + \frac{1}{k(\nu_{\min})^{1/2}} \right) + 1 \right) \|f\|_{L^2(\Omega_R)}.$$

The result (6.31) then follows by noting that  $L(\nu A^{-1}, \Omega, R+2) \geq 2$ ; this last inequality holds since the geodesic needs to at least travel from the boundary of the ball of radius  $R+2$  into the ball of radius  $R$  and out again (a distance of at least 4), and in this annular region the metric is Euclidean and the flow has speed 2.  $\square$

We can now prove Theorem 3.

*Proof of Theorem 3.* Using the polynomial approximation result (6.18) and the definitions of  $\eta(\mathcal{H}_{h_{\text{FEM}}})$  (6.27) and  $\|\cdot\|_{H_{k,A,\nu}^1(\Omega_R)}$  (6.7), we have

$$\eta(\mathcal{H}_{h_{\text{FEM}}}) \leq C_{\text{int}} h_{\text{FEM}} \sqrt{1 + (h_{\text{FEM}} k)^2} \sup_{0 \neq f \in L^2(\Omega_R)} \frac{\|S^* f\|_{H^2(\Omega_R)}}{\|f\|_{L^2(\Omega_R)}}.$$

By Lemma 6.5, the solution of the adjoint exterior Dirichlet problem with data  $f$  is the complex conjugate of the solution of the exterior Dirichlet problem with data  $\bar{f}$ . Therefore, the bound (6.31) for the solution of the exterior Dirichlet problem also holds for the solution of the adjoint problem, and implies that

$$\begin{aligned} \eta(\mathcal{H}_{h_{\text{FEM}}}) &\leq C_{\text{int}} h_{\text{FEM}} \sqrt{1 + (h_{\text{FEM}} k)^2} k C_{H^2} \frac{2\sqrt{2}}{\pi(\nu_{\min})^{1/2}} L(\nu A^{-1}, \Omega, R+2) \\ &\quad \times \left( (\nu_{\max})^{1/2} + \frac{1 + (\nu_{\min})^{1/2}}{k_0} + \frac{1}{k_0^2 (\nu_{\min})^{1/2}}; \right) \end{aligned}$$

the result then follows from using this bound on  $\eta(\mathcal{H}_{h_{\text{FEM}}})$  in Lemma 6.6.  $\square$

*Remark 6.8* (How the bound (1.8) can be used in the analysis of preconditioning strategies). The Galerkin method (6.17) is equivalent to a linear system of equations; denote the matrix of this linear system by  $\mathbf{A}$ . Linear systems involving  $\mathbf{A}$  are difficult to solve because (a) the dimension of  $\mathbf{A}$  is proportional to  $h_{\text{FEM}}^{-n}$  and (by Theorem 3)  $h_{\text{FEM}}$  must decrease like  $k^{-2}$  for the Galerkin solution to be quasioptimal, therefore  $\mathbf{A}$  is large, and (b) since  $\mathbf{A}$  is large and sparse (when using standard piecewise-polynomial bases of  $\mathcal{H}_{h_{\text{FEM}}}$ ), iterative methods are usually used to solve the linear system, but  $\mathbf{A}$  is both non-normal (in general) and sign-indefinite when  $k$  is sufficiently large, and the efficient iterative solution of linear systems involving such matrices is difficult. One therefore seeks to *precondition*  $\mathbf{A}$ ; i.e., to find a  $\mathbf{B}$  such that (i)  $\mathbf{B}^{-1}$  approximates  $\mathbf{A}^{-1}$  and (ii) the action of  $\mathbf{B}^{-1}$  is cheap to compute, and one then applies the iterative solver to  $\mathbf{B}^{-1} \mathbf{A}$ .

A very popular and successful preconditioner for  $\mathbf{A}$  is based on choosing  $\mathbf{B}^{-1}$  to be a cheap approximation of  $(\mathbf{A}_\alpha)^{-1}$ , where  $\mathbf{A}_\alpha$  is the Galerkin matrix arising from the exterior Dirichlet problem with  $k^2 \mapsto k^2 + i\alpha$ , i.e., with artificial absorption  $\alpha$  added. The rationale behind this method (introduced in [21]) is that the larger  $\alpha$  is, the less oscillatory the problem is, and hence

the easier it is to find a cheap approximation to  $(\mathbf{A}_\alpha)^{-1}$ . However, the larger  $\alpha$  is, the further  $(\mathbf{A}_\alpha)^{-1}$  is from  $\mathbf{A}^{-1}$ , hence the question of what is the optimal  $\alpha$  is nontrivial.

Although one of the advantages of preconditioning with absorption, compared to other Helmholtz-preconditioning strategies, is its easy applicability to variable-coefficient problems (see, e.g., [41]), the only existing analysis is for constant-coefficient Helmholtz problems (i.e. (6.2) with  $A = I$  and  $\nu = 1$ ). As a component of determining the optimal  $\alpha$ , [24, Theorem 1.4] proved that, when  $A = I$ ,  $\nu = 1$ ,  $\Omega$  is starshaped, and a quasi-optimal error estimate (similar to (6.22)) holds for Galerkin discretizations of the problem with absorption, there exists  $C > 0$  (independent of  $k$  and  $h_{\text{FEM}}$ , but dependent on  $\Omega$  and  $\mathcal{T}_{h_{\text{FEM}}}$ ) such that

$$(6.33) \quad \|\mathbf{I} - (\mathbf{A}_\alpha)^{-1}\mathbf{A}\| \leq C\alpha/k;$$

i.e. choosing  $\alpha$  to be a sufficiently small multiple of  $k$  guarantees that  $(\mathbf{A}_\alpha)^{-1}$  is a good approximation to  $\mathbf{A}^{-1}$ , uniformly as  $k \rightarrow \infty$ . Using the bound (1.8) in the arguments of [24] one can show that the bound (6.33) holds when  $A$ ,  $\nu$ , and  $\Omega$  satisfy the conditions of Theorem 3 with  $C$  a constant multiple of  $L(\nu A^{-1}, \Omega, R + 1)$ ; i.e., as the length of the longest ray increases, less absorption is allowed for  $(\mathbf{A}_\alpha)^{-1}$  to be a good approximation to  $\mathbf{A}^{-1}$ . This result is then consistent with the numerical experiments in [24, Tables 8 and 9]: the geometry corresponding to Table 9 supports longer rays than the geometry corresponding to Table 8, and the numerical results in the tables (the number of iterations required to solve  $(\mathbf{A}_\alpha)^{-1}\mathbf{A}$  with GMRES) show that a lower amount of absorption is allowed in the former case than in the latter for  $(\mathbf{A}_\alpha)^{-1}$  to be a good preconditioner for  $\mathbf{A}$ .

*Remark 6.9* (How the bound (1.8) can be used in ‘‘Uncertainty Quantification’’). In the last 10 years there has been a surge of interest in ‘‘Uncertainty Quantification (UQ)’’ of PDEs, understood as theory and algorithms for computing statistics of quantities of interest involving PDEs *either* posed on a random domain *or* having random coefficients.

There is a large literature on UQ for the Poisson equation

$$(6.34) \quad -\nabla \cdot (A(\omega)\nabla u(\omega)) = f(\omega),$$

(where  $\omega$  is an element of the underlying probability space) due, in part, to its large number of applications (e.g. in modelling groundwater flow). The fact that a priori bounds on the solution of (6.34) that are explicit in the coefficient  $A$  can easily be obtained is the starting point for the rigorous analysis of UQ algorithms; see e.g. [2, 1, 25, 38, 14, 15]. For example, for (6.34) posed in a bounded Lipschitz domain  $D$  with homogeneous Dirichlet boundary conditions and  $A \in L^\infty(D)$  with  $A_{\min} > 0$  (in the sense of quadratic forms as in (6.10)), the Lax-Milgram theorem implies that

$$(6.35) \quad \|u\|_{H^1(D)} \leq \frac{C_D}{A_{\min}} \|f\|_{L^2(D)},$$

where  $C_D$  is the constant appearing in the Poincaré inequality  $\|v\|_{H^1(D)} \leq C_D^{1/2} \|\nabla v\|_{L^2(D)}$  for all  $v \in H_0^1(D)$ . In contrast, there has been essentially no rigorous theory of UQ for the Helmholtz equation with large  $k$  because a priori bounds on the solution that are explicit in both  $k$  and the coefficients have not been available.

The recent paper [42] presented general measure-theory arguments that convert a bound on the (deterministic) Helmholtz equation that is explicit in both  $k$  and the coefficients into a bound, and associated well-posedness result, on the Helmholtz equation with random coefficients (and random data). The paper [42] used as input to these general arguments the deterministic bounds from [26] for star-shaped Lipschitz  $\Omega$  and coefficients satisfying radial monotonicity-like conditions which are stronger than nontrapping. We highlight that the bound (1.8) can be used with the arguments of [42] to obtain well-posedness results and a priori bounds on the

Helmholtz equation (6.2) where the coefficients and domain are such that the problem is almost surely nontrapping.

## REFERENCES

- [1] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 45(3):1005–1034, 2007.
- [2] I. Babuska, R. Tempone, and G. E. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM Journal on Numerical Analysis*, 42(2):800–825, 2004.
- [3] L. Banjai and S. Sauter. A refined Galerkin error and stability analysis for highly indefinite variational problems. *SIAM Journal on Numerical Analysis*, 45(1):37–53, 2007.
- [4] H. Barucq, T. Chaumont-Frelet, and C. Gout. Stability analysis of heterogeneous Helmholtz problems and finite element solution based on propagation media approximation. *Math. Comp.*, 86(307):2129–2157, 2017.
- [5] C. Bernardi. Optimal finite-element interpolation on curved domains. *SIAM Journal on Numerical Analysis*, 26(5):1212–1240, 1989.
- [6] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, 3rd edition, 2008.
- [7] D. L. Brown, D. Gallistl, and D. Peterseim. Multiscale Petrov-Galerkin method for high-frequency heterogeneous Helmholtz equations. In *Meshfree Methods for Partial Differential Equations VIII*, pages 85–115. Springer, 2017.
- [8] N. Burq. Mesures semi-classiques et mesures de défaut. *Séminaire Bourbaki*, (245):Exp. No. 826, 4, 167–195, 1997. Vol. 1996/97.
- [9] N. Burq. Semi-classical estimates for the resolvent in nontrapping geometries. *Int. Math. Res. Not.*, (5):221–241, 2002.
- [10] N. Burq and G. Lebeau. Mesures de défaut de compacité, application au système de Lamé. *Ann. Sci. École Norm. Sup. (4)*, 34(6):817–870, 2001.
- [11] Nicolas Burq and Claude Zuily. Laplace eigenfunctions and damped wave equation on product manifolds. *Appl. Math. Res. Express. AMRX*, (2):296–310, 2015.
- [12] S. N. Chandler-Wilde and P. Monk. Wave-number-explicit bounds in time-harmonic scattering. *SIAM Journal on Mathematical Analysis*, 39(5):1428–1455, 2008.
- [13] S. N. Chandler-Wilde, E. A. Spence, A. Gibbs, and V. P. Smyshlyaev. High-frequency bounds for the Helmholtz equation under parabolic trapping and applications in numerical analysis. *arXiv preprint arXiv:1708.08415*, 2017.
- [14] J. Charrier. Strong and weak error estimates for elliptic partial differential equations with random coefficients. *SIAM Journal on Numerical Analysis*, 50(1):216–246, 2012.
- [15] J. Charrier, R. Scheichl, and A. L. Teckentrup. Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods. *SIAM Journal on Numerical Analysis*, 51(1):322–352, 2013.
- [16] T. Chaumont Frelet. *Approximation par éléments finis de problèmes d’Helmholtz pour la propagation d’ondes sismiques*. PhD thesis, Rouen, INSA, 2015.
- [17] T. Chaumont-Frelet. On high order methods for the heterogeneous Helmholtz equation. *Computers & Mathematics with Applications*, 72(9):2203–2225, 2016.
- [18] T. Chaumont-Frelet and S. Nicaise. Wavenumber explicit convergence analysis for finite element discretizations of general wave propagation problem. *IMA J. Num. Anal.*, to appear, 2019.
- [19] P. G. Ciarlet. Basic error estimates for elliptic problems. In *Handbook of numerical analysis, Vol. II*, Handb. Numer. Anal., II, pages 17–351. North-Holland, Amsterdam, 1991.
- [20] Semyon Dyatlov and Maciej Zworski. *Mathematical theory of scattering resonances*. AMS, 2019. <http://math.mit.edu/~dyatlov/res/>.
- [21] Y. A. Erlangga, C. W. Oosterlee, and C. Vuik. A novel multigrid based preconditioner for heterogeneous Helmholtz problems. *SIAM J. Sci. Comp.*, 27:1471–1492, 2006.
- [22] S. Esterhazy and J. M. Melenk. On stability of discretizations of the Helmholtz equation. In I. G. Graham, Th. Y. Hou, O. Lakkis, and R. Scheichl, editors, *Numerical Analysis of Multiscale Problems*, volume 83 of *Lecture Notes in Computational Science and Engineering*, pages 285–324. Springer, 2012.
- [23] G. B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.
- [24] M. J. Gander, I. G. Graham, and E. A. Spence. Applying GMRES to the Helmholtz equation with shifted Laplacian preconditioning: what is the largest shift for which wavenumber-independent convergence is guaranteed? *Numerische Mathematik*, 131(3):567–614, 2015.



- [25] C. J. Gittelsohn. Stochastic galerkin discretization of the log-normal isotropic diffusion problem. *Mathematical Models and Methods in Applied Sciences*, 20(02):237–263, 2010.
- [26] I. G. Graham, O. R. Pembrey, and E. A. Spence. The Helmholtz equation in heterogeneous media: a priori bounds, well-posedness, and resonances. *Journal on Differential Equations*, to appear, 2018.
- [27] I. G. Graham and S. A. Sauter. Stability and finite element error analysis for the Helmholtz equation with variable coefficients. *Math. Comp.*, to appear, 2019.
- [28] P. R. Halmos. *A Hilbert space problem book*. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London, 1967.
- [29] L. Hörmander. *The analysis of linear partial differential operators. III*. Classics in Mathematics. Springer, Berlin, 2007. Pseudo-differential operators, Reprint of the 1994 edition.
- [30] F. Ihlenburg and I. Babuška. Finite element solution of the Helmholtz equation with high wave number Part I: The h-version of the FEM. *Computers & Mathematics with Applications*, 30(9):9–37, 1995.
- [31] F. Ihlenburg and I. Babuška. Finite element solution of the Helmholtz equation with high wave number part II: the hp version of the FEM. *SIAM Journal on Numerical Analysis*, 34(1):315–358, 1997.
- [32] W. C. H. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, 2000.
- [33] J. M. Melenk. *On generalized finite element methods*. PhD thesis, The University of Maryland, 1995.
- [34] J. M. Melenk and S. Sauter. Convergence analysis for finite element discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions. *Math. Comp.*, 79(272):1871–1914, 2010.
- [35] J. M. Melenk and S. Sauter. Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation. *SIAM J. Numer. Anal.*, 49:1210–1243, 2011.
- [36] L. Miller. Refraction of high-frequency waves density by sharp interfaces and semiclassical measures at the boundary. *J. Math. Pures Appl. (9)*, 79(3):227–269, 2000.
- [37] A. Moiola and E. A. Spence. Acoustic transmission problems: wavenumber-explicit bounds and resonance-free regions. *Mathematical Models & Methods in Applied Sciences*, 29(2):317–354, 2019.
- [38] A. Mugler and H.-J. Starkloff. On elliptic partial differential equations with random coefficients. *Studia Universitatis Babeş-Bolyai, Mathematica*, 56(2), 2011.
- [39] J. C. Nédélec. *Acoustic and electromagnetic equations: integral representations for harmonic problems*. Springer Verlag, 2001.
- [40] M. Ohlberger and B. Verfurth. A new heterogeneous multiscale method for the Helmholtz equation with high contrast. *Multiscale Modeling & Simulation*, 16(1):385–411, 2018.
- [41] C. W. Oosterlee, C. Vuik, W. A. Mulder, and R.-E. Plessix. Shifted-Laplacian preconditioners for heterogeneous Helmholtz problems. In B. Koren and C. Vuik, editors, *Advanced Computational Methods in Science and Engineering*, volume 71 of *Lecture Notes in Computational Science and Engineering*, pages 21–46. Springer, 2010.
- [42] O. R. Pembrey and E. A. Spence. The Helmholtz equation in random media: well-posedness and a priori bounds. *arXiv preprint 1805.00282*, 2018.
- [43] S. Sauter and C. Torres. Stability estimate for the Helmholtz equation with rapidly jumping coefficients. *Zeitschrift für angewandte Mathematik und Physik*, 69:139, 2018.
- [44] S. A. Sauter. A refined finite element convergence theory for highly indefinite Helmholtz problems. *Computing*, 78(2):101–115, 2006.
- [45] A. H. Schatz. An observation concerning Ritz-Galerkin methods with indefinite bilinear forms. *Mathematics of Computation*, 28(128):959–962, 1974.
- [46] E. A. Spence. Overview of Variational Formulations for Linear Elliptic PDEs. In A. S. Fokas and B. Pelloni, editors, *Unified transform method for boundary value problems: applications and advances*, pages 93–159. SIAM, 2015.
- [47] M. E. Taylor. *Tools for PDE*, volume 81 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000. Pseudodifferential operators, paradifferential operators, and layer potentials.
- [48] M. E. Taylor. *Partial differential equations III. Nonlinear equations*, volume 117 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2011.
- [49] A. Vasy. Propagation of singularities for the wave equation on manifolds with corners. *Ann. of Math. (2)*, 168(3):749–812, 2008.



DEPARTMENT OF MATHEMATICS, UNIVERSITY COLLEGE LONDON, LONDON, UK  
*E-mail address:* `j.galkowski@ucl.ac.uk`

DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF BATH, BATH, UK  
*E-mail address:* `E.A.Spence@bath.ac.uk`

DEPARTMENT OF MATHEMATICS, NORTHWESTERN UNIVERSITY, EVANSTON, IL, USA  
*E-mail address:* `jwunsch@math.northwestern.edu`