Adv Comput Math (2023) 49:27
https://doi.org/10.1007/s10444-023-10025-3

A simple proof that the *hp*-FEM does not suffer from the pollution effect for the constant-coefficient full-space Helmholtz equation

E. A. Spence[1]

Received: 19 September 2022 / Accepted: 23 February 2023 / Published online: 10 April 2023

Abstract

In d dimensions, accurately approximating an arbitrary function oscillating with frequency $\lesssim k$ requires $\sim k^d$ degrees of freedom. A numerical method for solving the Helmholtz equation (with wavenumber k) suffers from the pollution effect if, as $k \to \infty$, the total number of degrees of freedom needed to maintain accuracy grows faster than this natural threshold. While the h-version of the finite element method (FEM) (where accuracy is increased by decreasing the meshwidth h and keeping the polynomial degree p fixed) suffers from the pollution effect, the hp-FEM (where accuracy is increased by decreasing the meshwidth h and increasing the polynomial degree p) does not suffer from the pollution effect. The heart of the proof of this result is a PDE result splitting the solution of the Helmholtz equation into "high" and "low" frequency components. This result for the constant-coefficient Helmholtz equation in full space (i.e. in $\mathbb{R}^d$) was originally proved in Melenk and Sauter (*Math. Comp* **79**(272), 1871–1914, 2010). In this paper, we prove this result using *only* integration by parts and elementary properties of the Fourier transform. The proof in this paper is motivated by the recent proof in Lafontaine et al. (*Comp. Math. Appl.* **113**, 59–69, 2022) of this splitting for the variable-coefficient Helmholtz equation in full space use the more-sophisticated tools of semiclassical pseudodifferential operators.

Keywords Helmholtz equation · High frequency · Finite element method · Pollution effect

Mathematics Subject Classification (2010) 35J05 · 65N12 · 65N15 · 65N30

Communicated by: Jon Wilkening

✉ E. A. Spence
e.a.spence@bath.ac.uk

[1] Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK

1 Introduction and motivation

When computing approximations with the finite element method to the solution of the Helmholtz equation

$$\Delta u + k^2 u = -f, \tag{1.1}$$

with wavenumber $k > 0$, a fundamental question is:

> How quickly must the meshwidth h decrease with k and/or the polynomial degree p increase with k to maintain accuracy as $k \to \infty$?

This question has been the subject of sustained interest since the late 1980s, with initially answers obtained for 1-d problems [1, 29, 30], and now answers obtained for 2- and 3-d problems in general geometries, both for "standard" FEMs [14, 16, 19, 24, 32–36, 52, 55] and for variations of these, e.g. discontinuous Galerkin methods [17, 18, 37, 54] and multiscale methods [3, 7, 9, 10, 20, 45, 46]. Moreover, there is large current interest in this question when the Helmholtz equation (1.1) is replaced by its variable-coefficient generalisation $\nabla \cdot (A\nabla u) + k^2 nu = 0$ [4, 6, 11, 21, 22, 25, 27, 32, 38] or even the time-harmonic Maxwell equations [39, 40, 44].

A highlight of this body of research is the result from [4, 16, 21, 22, 32, 34, 36, 37] that the hp-FEM does not suffer from the *pollution effect*; i.e. accuracy can be maintained with a choice of the number of degrees of freedom growing like k^d, where "accuracy" here means that the computed solution is *quasi-optimal* (see (3.2) below). This is contrast to the h-version of the FEM which, e.g. with $p = 1$, needs the total number of degrees of freedom to grow like k^{2d} to maintain accuracy in this sense. Having the number of degrees of freedom growing like k^d is the natural threshold for this problem since an oscillatory function in d dimensions with frequency $\lesssim k$ requires $\sim k^d$ degrees of freedom to be well-approximated by piecewise polynomials; this is expected from the Nyquist–Shannon–Whittaker sampling theorem in 1-d [50, 53], and from the recent results in general dimension in [23].

The proofs that the hp-FEM does not suffer from the pollution effect consist of the following three ingredients.

1. Sufficient conditions for FEM solutions to be quasi-optimal originating from the ideas of Schatz [49] (related to the classic "Aubin–Nitsche trick"), and then developed by Sauter [48].
2. Results from [34, Appendices B and C] about how well the hp-FEM spaces approximate analytic functions.
3. A PDE result splitting the solution of the Helmholtz equation into "high" and "low" frequency components.

The motivation for the present paper was the realisation that, for the constant-coefficient Helmholtz equation (1.1) posed in $\mathbb{R}^d$, given a bound on the solution in terms of the data (which can be proved using essentially only integration by parts), the splitting in point 3 above can be proved using *only* elementary properties of the Fourier transform, thus making the key ideas behind this body of work accessible to a wide audience.

The proof in this paper is motivated by the recent proof in [32] of this splitting for the variable-coefficient Helmholtz equation in $\mathbb{R}^d$; the proof in [32] uses the more-sophisticated tools of semiclassical pseudodifferential operators that reduce to the elementary ones above in the constant-coefficient case — we discuss this further in Section 7.6.

We highlight that the splitting in point 3 above, and also the results of points 1 and 2, are all concerned with the *analysis* of the *hp*-FEM; i.e. they influence the implementation of the *hp*-FEM only in that that together they give a prescription of how to tie h and p to k to ensure that the FEM solution is quasi-optimal, uniformly in k, as $k \to \infty$.

Plan of the paper Section 2 recalls basic facts about the Helmholtz equation. Sections 3, 4, and 5 concern points 1, 2, and 3 above, respectively. Section 5 states the splitting for the constant-coefficient full-space Helmholtz equation, and then uses the results of points 1–3 to prove that the *hp*-FEM does not suffer from the pollution effect when applied to this problem. Section 6 recaps the basic properties of the Fourier transform, and then Section 7 proves the splitting stated in Section 5 using the material in Section 6.

2 The Helmholtz equation

2.1 The model Helmholtz problem

As in [34], we consider the following model Helmholtz problem. Given $f \in L^2(\mathbb{R}^d)$ with compact support, let $u \in H^1_{\mathrm{loc}}(\mathbb{R}^d)$ be the solution of

$$k^{-2}\Delta u + u = -f \quad \text{in } \mathbb{R}^d, \;\; d = 2, 3, \tag{2.1}$$

that is *outgoing* in the sense that it satisfies the *Sommerfeld radiation condition*

$$k^{-1}\partial_r u(x) - \mathrm{i}u(x) = o\big(r^{-(d-1)/2}\big) \tag{2.2}$$

as $r := |x| \to \infty$, uniformly in $\widehat{x} := x/r$. Rellich's uniqueness theorem (see, e.g. [13, Theorem 3.13]) implies that the solution of (2.1)–(2.2) is unique. For this particular model problem, the solution can be written down explicitly as an integral of f against the fundamental solution of the Helmholtz equation (see (7.10) below), with then mapping properties of this integral operator (see, e.g. [41, Theorem 6.1]) showing existence of the solution to (2.1)–(2.2). Observe that we have multiplied the Helmholtz equation (1.1) by k^{-2} and rescaled the right-hand side f; we see below how this rescaling by k^{-2} allows us to keep better track of the k-dependence.

2.2 The variational formulation of the Helmholtz equation

Let $R > 0$ be large enough so that $\operatorname{supp} f \subset B_R := \{x \in \mathbb{R}^d : |x| < R\}$. The variational formulation of (2.1)–(2.2) is then

$$\text{find } \widetilde{u} \in H^1(B_R) \text{ such that } \; a(\widetilde{u}, v) = F(v) \;\; \text{for all } v \in H^1(B_R), \tag{2.3}$$

where

$$a(\widetilde{u}, v) := \int_{B_R} \Big(k^{-2} \nabla \widetilde{u} \cdot \overline{\nabla v} - \widetilde{u}\overline{v} \Big) - k^{-1} \big\langle \mathrm{DtN}_k \widetilde{u}, v \big\rangle_{\partial B_R} \tag{2.4}$$

and

$$F(v) := \int_{B_R} f\, \overline{v}.$$

The operator DtN_k in (2.4) is the Dirichlet-to-Neumann map for the outgoing solution of the Helmholtz equation in the exterior of the ball B_R; i.e. given $g \in H^{1/2}(\partial B_R)$, let v be the unique outgoing solution to

$$(-k^{-2}\Delta - 1)v = 0 \quad \text{in } \mathbb{R}^d \setminus \overline{B_R} \quad \text{and} \quad v = g \text{ on } \partial B_R;$$

then $\mathrm{DtN}_k g := k^{-1}\partial_r v$, i.e.

$$\mathrm{DtN}_k g(\theta) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \frac{H_n^{(1)'}(kR)}{H_n^{(1)}(kR)} \exp(\mathrm{i}n\theta) \int_0^{2\pi} \mathrm{e}^{-\mathrm{i}n\theta} g(R,\theta)\, \mathrm{d}\theta;$$

for the analogous expression when $d = 3$, see, e.g. [34, Equations 3.7 and 3.10].

Green's identity and the definitions of DtN_k and $a(\cdot,\cdot)$ imply that if u is a solution of (2.1)–(2.2), then $u|_{B_R}$ is a solution of the variational problem (2.3). Conversely, if $\widetilde{u}$ is a solution of this variational problem, then there exists a solution u of (2.1)–(2.2) such that $u|_{B_R} = \widetilde{u}$; thus, the solution of the variational problem (2.3) exists and is unique. Because of this equivalence result, and for simplicity, from now on, we denote by u *both* the solution of (2.1)–(2.2) *and* the solution of the variational problem (2.3).

Remark 2.1 (Approximating DtN_k) Implementing the operator DtN_k appearing in $a(\cdot,\cdot)$ (2.4) is computationally expensive, and so in practice one seeks to approximate this operator by, e.g. imposing an absorbing boundary condition on ∂B_R, or using a perfectly-matched layer (PML). For simplicity, in this paper (as in [34]), we analyse the FEM assuming that DtN_k is realised exactly.

2.3 The k-dependence of the Helmholtz solution operator

We work with the weighted norm

$$\|v\|^2_{H^m_k(B_R)} := \sum_{0\le|\alpha|\le m} \Big\| (k^{-1}\partial)^\alpha v \Big\|^2_{L^2(B_R)}. \tag{2.5}$$

with special case $\|v\|^2_{H^1_k(B_R)} := \|k^{-1}\nabla v\|^2_{L^2(B_R)} + \|v\|^2_{L^2(B_R)}$. The rationale for using these norms is that if a function v oscillates with frequency k, then we expect $|(k^{-1}\partial)^\alpha v| \sim |v|$ for all α; this is true, e.g. if $v(x) = \exp(\mathrm{i}kx \cdot a)$.

Let $C_{\mathrm{sol}}(k, R)$ be the operator norm of the map $L^2(B_R) \ni f \mapsto u \in H^1(B_R)$, where u is the outgoing solution of the Helmholtz equation (2.1); i.e. given k and R, $C_{\mathrm{sol}}(k, R)$ is such that given $f \in L^2(B_R)$ if u is the solution of (2.1)–(2.2), then

$$\|u\|_{H^1_k(B_R)} \le C_{\mathrm{sol}} \, \|f\|_{L^2(B_R)}. \tag{2.6}$$

For the simple model problem (2.1)–(2.2), the following bound on C_{sol} can be obtained by multiplying the Helmholtz equation (2.1) by a judiciously chosen test function and integrating by parts.

Theorem 2.2 (Morawetz bound on C_{sol}) *For all $k > 0$ and $R > 0$,*

$$C_{\mathrm{sol}} \leq 2kR\sqrt{1 + \left(\frac{d-1}{2kR}\right)^2}. \tag{2.7}$$

This bound was essentially proved in [42, 43] (although the bound does not quite appear in this form in those papers) (see also [8, Lemma 3.5] and [28, Equations 1.9 and 1.10]). The details of the proof of this bound are not needed in the rest of the paper, but for completeness (and since they involve essentially only integrating by parts), we include them in Appendix A.

The bound (2.7) is sharp in its kR dependence for kR large; indeed, by considering $u(x) = \mathrm{e}^{\mathrm{i}kx_1}\chi(|x|/R)$ for $\chi \in C^\infty$ supported in $[0, 1)$, one can show that given $k_0, R_0 > 0$, there exists $C > 0$ such that $C_{\mathrm{sol}} \geq CkR$ for all $k \geq k_0$ and $R \geq R_0$.

Corollary 2.3 (Bound on H^2_k norm of the Helmholtz solution) *Given $k_0, R_0 > 0$, there exists $C > 0$ such that the solution u of* (2.1)–(2.2) *with* $\mathrm{supp} f \subset B_R$ *with $R \geq R_0$ satisfies*

$$\|u\|_{H^2_k(B_R)} \leq CkR\,\|f\|_{L^2(B_R)} \quad \text{for all } k \geq k_0. \tag{2.8}$$

Comparing the bounds (2.7) and (2.8) shows the advantage of working in the particular weighted norms (2.5) — the k-dependence of the solution operator is the same regardless of the spaces it maps between.

Sketch proof of Corollary 2.3 Corollary 6.4 below uses the Fourier transform to prove H^2 regularity of the solution of $(-k^{-2}\Delta + 1)v = g \in L^2(\mathbb{R}^d)$. To apply this bound to the solution of (2.1)–(2.2), let $\varphi \in C^\infty_{\mathrm{comp}}(\mathbb{R}^d, [0, 1])$ be equal to one on B_R and vanish outside B_{2R}, and then let $v := \varphi u$ so that $g = (-k^{-2}\Delta + 1)(\varphi u) = (-k^{-2}\Delta - 1)(\varphi u) + 2\varphi u$. The bound (2.8) then follows by bounding $\|g\|_{L^2(B_{2R})}$ in terms of $\|f\|_{L^2(B_R)}$ using the PDE (2.1) and the bounds (2.6) and (2.7). □

Remark 2.4 Many papers on the numerical analysis of the Helmholtz equation use the weighted H^1 norm $\|u\|^2_{H^1_k(B_R)} := \|\nabla u\|^2_{L^2(B_R)} + k^2\|u\|^2_{L^2(B_R)}$; we use (2.5) instead since weighting the jth derivative by k^{-j} is easier to keep track of than weighting it by k^{-j+1} (especially for high derivatives).

3 The Galerkin method and sufficient conditions for quasioptimality

Let $\mathcal{H}_N$ be a finite-dimensional subspace of $H^1(B_R)$. The Galerkin method applied to the variational problem (2.3) is

$$\text{find } u_N \in \mathcal{H}_N \text{ such that } \; a(u_N, v_N) = F(v_N) \; \text{ for all } v_N \in \mathcal{H}_N. \tag{3.1}$$

The FEM is the Galerkin method (3.1) with $\mathcal{H}_N$ consisting of piecewise polynomials (we describe in Section 4 the specific assumptions we make on $\mathcal{H}_N$).

Given a sequence of finite-dimensional spaces $\{\mathcal{H}_N\}_{N=1}^{\infty}$, a standard error bound one seeks to prove on the sequence of Galerkin solutions $\{u_N\}_{N=1}^{\infty}$ is the following *quasioptimal* error bound: there exists a $C_{\rm qo} > 0$ and $N_0 \in \mathbb{N}$ such that, for $N \geq N_0$,

$$\|u - u_N\|_{H^1_k(B_R)} \leq C_{\rm qo} \min_{v_N \in \mathcal{H}_N} \|u - v_N\|_{H^1_k(B_R)} . \tag{3.2}$$

The main result of this section is Lemma 3.4 below, giving sufficient conditions for quasioptimality. This result crucially relies on the following properties of the sesquilinear form $a(\cdot, \cdot)$ and the following properties of the adjoint solution operator.

Lemma 3.1 (Properties of $a(\cdot, \cdot)$)

(i) *(Continuity) Given $k_0, R_0 > 0$ there exists $C_{\rm cont} > 0$ such that for all $k \geq k_0$ and $R \geq R_0$,*

$$|a(u, v)| \leq C_{\rm cont} \|u\|_{H^1_k(B_R)} \|v\|_{H^1_k(B_R)} \quad \textit{for all } u, v \in H^1(B_R).$$

(ii) *(Gårding inequality)*

$$\Re a(v, v) \geq \|v\|^2_{H^1_k(B_R)} - 2\|v\|^2_{L^2(B_R)} \quad \textit{for all } v \in H^1(B_R).$$

References for the proof Part (i) follows from the Cauchy–Schwarz inequality and boundedness of DtN_k; see [34, Lemma 3.3, Part 1]. Part (ii) follows from the fact that $-\Re\langle \mathrm{DtN}_k \phi, \phi\rangle_{\partial B_R} \geq 0$ for all $\phi \in H^{1/2}(\partial B_R)$; see [34, Lemma 3.3, Part 2]. □

Definition 3.2 (Adjoint solution operator $\mathcal{S}^*$) Given $f \in L^2(B_R)$, let $\mathcal{S}^* f \in H^1(B_R)$ be defined by

$$a(v, \mathcal{S}^* f) = (v, f)_{L^2(B_R)} \text{ for all } v \in H^1(B_R). \tag{3.3}$$

The following lemma shows that our knowledge about outgoing Helmholtz solutions immediately gives us knowledge about $\mathcal{S}^*$.

Lemma 3.3 *If $\mathcal{S}^*$ is defined as in (3.3) then*

$$a(\overline{\mathcal{S}^* f}, v) = (\overline{f}, v)_{L^2(B_R)} \quad \textit{for all } v \in H^1(B_R).$$

i.e. $\mathcal{S}^* f$ is the complex-conjugate of the outgoing Helmholtz solution with data $\overline{f}$.

Sketch proof Green's identity and the radiation condition (2.2) show that $\langle \mathrm{DtN}_k \psi, \overline{\phi}\rangle_{\partial B_R} = \langle \mathrm{DtN}_k \phi, \overline{\psi}\rangle_{\partial B_R}$ for all $\phi, \psi \in H^{1/2}(\partial B_R)$. This implies that $a(\overline{v}, u) = a(\overline{u}, v)$ for all u, v, which implies the result. □

Lemma 3.4 (Sufficient conditions for quasi-optimality) *If*

$$\eta(\mathcal{H}_N) := \sup_{0\neq f\in L^2(B_R)} \min_{v_N\in\mathcal{H}_N} \frac{\|\mathcal{S}^* f - v_N\|_{H^1_k(B_R)}}{\|f\|_{L^2(B_R)}} \leq \frac{1}{2C_{\mathrm{cont}}}, \tag{3.4}$$

then the Galerkin solution u_N to the variational problem (3.1) exists, is unique, and satisfies the quasi-optimal error bound

$$\|u - u_N\|_{H^1_k(B_R)} \leq 2C_{\mathrm{cont}} \left(\min_{v_N\in V_N} \|u - v_N\|_{H^1_k(B_R)} \right). \tag{3.5}$$

References for the proof See [48, Theorem 2.5] or [34, Theorem 4.3] (note that these references consider the Helmholtz equation in the form $\Delta u + k^2 u = -f$ and use the weighted norm discussed in Remark 2.4, but it is straightforward to convert the results to our setting). □

4 Recap of approximation results in *hp*-FEM spaces

The result that the *hp*-FEM does not suffer from the pollution effect is proved under the following two assumptions on the finite-dimensional subspaces; these assumptions describe how well (as a function of h and p) the spaces approximate functions with a given regularity.

We highlight immediately that both these assumptions are satisfied by *hp*-finite-element spaces with curved elements that fit ∂B_R exactly, provided that the triangulations are quasi-uniform and are constructed by refining a fixed triangulation that has analytic element maps (see Theorem 4.4 below). Nevertheless, we formulate these properties as specific assumptions to make it clear the actual properties of the subspaces that are needed in the proof of the result that the *hp*-FEM does not suffer from the pollution effect. Since we are ultimately thinking of the subspaces in these assumptions as *hp*-finite-element spaces, we denote the sequence of these subspaces as $\{\mathcal{H}_{h,p}\}_{h>0,p\in\mathbb{Z}^+}$.

Assumption 4.1 (Approximation in the finite-dimensional subspace of functions with finite regularity) *Let $\{\mathcal{H}_{h,p}\}_{h>0,p\in\mathbb{Z}^+}$ be a sequence of finite-dimensional subspaces of $H^1(B_R)$. Given s, d with d the spatial dimension and $s > d/2$, there exists $C_{\mathrm{approx}\,1} > 0$ such that if $v \in H^s(B_R)$ and $p \geq s-1$, then*

$$\min_{w_{h,p}\in\mathcal{H}_{h,p}} \left\| v - w_{h,p} \right\|_{H^1_k(B_R)} \leq C_{\mathrm{approx}\,1} \left(\frac{hk}{p}\right)^{s-1} \left(1 + \frac{hk}{p}\right) \|v\|_{H^s_k(B_R)}. \tag{4.1}$$

Discussion of Assumption 4.1 A standard polynomial approximation result is the following. For $d = 2, 3$, given $s \geq 2$ and $p \geq s-1$, there exists $C > 0$ such that if $v \in H^s(D)$ and $m = 0$ or 1, then

$$\left\| v - \mathcal{I}^h v \right\|_{H^m(D)} \leq C h^{s-m} |v|_{H^s(D)}, \tag{4.2}$$

where $\mathcal{I}^h$ is a global interpolation operator (see, e.g. [5, Equation 4.4.28], [12, Theorem 17.1]). The approximation result (4.1) is a generalisation of (4.2) which (i) makes explicit the dependence on p of the constant C in (4.2), and (ii) works in norms weighted with k.

Motivation for Assumption 4.3. Assumption 4.1 is about approximating in k-weighted norms an arbitrary function in H^s for some $s > 0$; note that the constant $C_{\text{approx}\,1}$ depends on s in an unspecified way, and so we cannot use the bound (4.1) for arbitrarily large s, and hence arbitrarily large p, since p is tied to s via $p \geq s-1$. The next assumption, Assumption 4.3, allows us to take arbitrarily large p; the price one pays is that the function being approximated must be analytic.

Before stating Assumption 4.3, we recall the relationship between derivative bounds and analyticity for families of functions depending on k.

Lemma 4.2 (k-explicit analyticity) *Let D be a bounded open subset of $\mathbb{R}^d$ and let $u \in C^\infty(D)$ be a family of functions depending on k.*

(i) If there exist $C, C_u > 0$, independent of α, such that

$$\left\| \partial^\alpha u \right\|_{L^2(D)} \leq C_u (Ck)^{|\alpha|} \quad \text{for all multiindices } \alpha, \tag{4.3}$$

then u is real analytic in D and its power series has infinite radius of convergence, i.e. u can be extended to an entire function on $\mathbb{R}^d$.

(ii) If there exist $C, C_u > 0$, independent of α, such that

$$\left\| \partial^\alpha u \right\|_{L^2(D)} \leq C_u (Ck)^{|\alpha|} |\alpha|! \quad \text{for all multiindices } \alpha,$$

then u is real analytic in D with radius of convergence of its power series proportional to $(Ck)^{-1}$.

(iii) If there exist $C, C_u > 0$, independent of α, such that

$$\left\| \partial^\alpha u \right\|_{L^2(D)} \leq C_u C^{|\alpha|} \max\left\{ |\alpha|, k \right\}^{|\alpha|} \quad \text{for all multiindices } \alpha,$$

then u is real analytic in D with radius of convergence of its power series proportional to C^{-1} and independent of k.

Sketch proof In each case, use the Sobolev embedding theorem (see, e.g. [41, Theorem 3.26]) to obtain a bound on $\|\partial^\alpha u\|_{L^\infty(D)}$, and then use this to bound the Lagrange form of the remainder in the Taylor series (see, e.g. [34, Proof of Lemma C.2]). □

In the rest of the paper, we only use the class of functions in part (i) of Lemma 4.2, but the classes in parts (ii) and (iii) are included for context.

Assumption 4.3 (Approximation in the finite-dimensional subspace of the class of functions in part (i) of Lemma 4.2) *Let $\{\mathcal{H}_{h,p}\}_{h>0,\,p\in\mathbb{Z}^+}$ be a sequence of finite-dimensional subspaces of $H^1(B_R)$, and suppose $v \in C^\infty(B_R)$ is such that, given $k_0 > 0$ there exists $C_1, C_2 > 0$ such that*

$$\left\| (k^{-1}\partial)^\alpha v \right\|_{L^2(B_R)} \leq C_1 (C_2)^{|\alpha|} \quad \text{for all } \alpha \text{ and for all } k \geq k_0. \tag{4.4}$$

Given $\widetilde{C}$, there exist $\sigma, C_{\mathrm{approx}_2} > 0$, depending on C_2 and $\widetilde{C}$ (but not C_1), such that, if $k \geq k_0$ and k, h, and p satisfy

$$\frac{h}{R} + \frac{hk}{p} \leq \widetilde{C}, \tag{4.5}$$

then

$$\min_{w_{h,p} \in \mathcal{H}_{h,p}} \left\| v - w_{h,p} \right\|_{H_k^1(B_R)} \leq C_1 C_{\mathrm{approx}_2} \left[\frac{1}{kR} \left(\frac{h/R}{h/R + \sigma} \right)^p + \left(\frac{hk}{\sigma p} \right)^p \right]. \tag{4.6}$$

Discussion of Assumption 4.3 The key points about the bound (4.6) are the following.

(i) For fixed p and k, as $h \to 0$, the right hand side of (4.6) is $O(h^p)$ (just like the right-hand side of (4.1) with $s = p + 1$).
(ii) If $hk/(\sigma p) < 1$ then the right-hand side of (4.6) decreases exponentially as $p \to \infty$.
(iii) The quantities h/R, hk/p, and kR are dimensionless, and thus the right-hand sides of (4.5) and (4.6) involve only dimensionless quantities.

Approximation spaces satisfying Assumptions 4.1 and 4.3 Let $(\mathcal{T}_h)_{0<h\leq h_0}$ (with h the maximum element diameter) be a sequence of triangulations of B_R, with each element $K \in \mathcal{T}_h$ the image of a reference element $\widehat{K}$ (a reference triangle in 2-d and a reference tetrahedron in 3-d) under the map $F_K : \widehat{K} \to K$. As is standard, we assume there are no hanging nodes and that the element maps of elements sharing an edge or face induce the same parametrisation on that edge or face. We consider the hp-finite-element spaces

$$\mathcal{H}^p(\mathcal{T}_h) := \left\{ v \in H^1(B_R) : \text{for each } K \in \mathcal{T}_h,\ v|_K \circ F_K \text{ is a polynomial of degree} \leq p \right\}. \tag{4.7}$$

Since B_R is curved, we consider triangulations with curved elements that fit ∂B_R exactly (thus avoiding the issue of analysing the non-conforming error coming from using simplicial triangulations (see, e.g. [5, Chapter 10])). Recall that the family $(\mathcal{T}_h)_{0<h\leq h_0}$ is *quasi-uniform* if there exists $C > 0$ such that

$$h := \max_{K\in\mathcal{T}_h} \mathrm{diam}(K) \leq C \min_{K\in\mathcal{T}_h} \mathrm{diam}(K) \quad \text{for all } 0 < h \leq h_0;$$

for such triangulations, the dimension of $\mathcal{H}^p(\mathcal{T}_h)$ is proportional to $(p/(h/R))^d$.

Theorem 4.4 (Conditions under which Assumptions 4.1 and 4.3 hold) *If $(\mathcal{T}_h)_{0<h\leq h_0}$ satisfies* [34, Assumption 5.2], *then $\mathcal{H}^p(\mathcal{T}_h)$ defined by* (4.7) *satisfies Assumptions 4.1 and 4.3.*

Informally, [34, Assumption 5.2] is that $(\mathcal{T}_h)_{0<h\leq h_0}$ is quasi-uniform with each element map F_K the composition of a affine map and an analytic map; [36, Remark 5.2] notes that $(\mathcal{T}_h)_{0<h\leq h_0}$ satisfying this assumption can be constructed by refining a fixed triangulation that has analytic element maps.

References for the proof of Theorem 4.4 That Assumption 4.1 holds follows from [34, Theorem B.4] (a result about approximation on the reference element) and a scaling argument (see [34, Bottom of Page 1895]). For Assumption 4.3, the bound (4.6) is proved in the course of [34, Proof of Theorem 5.5], see the last equation on [34, Page 1896]; note that (i) we have simplified this equation using the assumption (4.5), and (ii) the weighted H^1 norm in [34] is k times $\|\cdot\|_{H^1_k(B_R)}$ defined by (2.5). □

5 The splitting of the Helmholtz solution and the proof that the *hp*-FEM does not suffer from the pollution effect

5.1 Statement of the splitting

The crucial result used to prove that the hp-FEM applied to the problem (2.1)–(2.2) does not suffer from the pollution effect is the following (with the proof contained in Section 7).

Theorem 5.1 (Splitting of the Helmholtz solution) *Given $k_0, R_0 > 0$, there exists $C_{\mathrm{split},H^2}, C_{\mathrm{split},\mathcal{A}} > 0$ such that the following holds. Given $f \in L^2(B_R)$ with $R \geq R_0$, let u satisfy the Helmholtz equation (2.1) and the Sommerfeld radiation condition (2.2). Then*

$$u|_{B_R} = u_{H^2} + u_{\mathcal{A}} \tag{5.1}$$

where $u_{H^2} \in H^2(B_R)$ with

$$\left\|u_{H^2}\right\|_{H^2_k(B_R)} \leq C_{\mathrm{split},H^2} \left\|f\right\|_{L^2(B_R)} \qquad \text{for all } k \geq k_0 \tag{5.2}$$

and $u_{\mathcal{A}} \in C^\infty(B_R)$ with

$$\left\|(k^{-1}\partial)^\alpha u_{\mathcal{A}}\right\|_{L^2(B_R)} \leq C_{\mathrm{sol}}(k, 2R) \left(C_{\mathrm{split},\mathcal{A}}\right)^{|\alpha|} \left\|f\right\|_{L^2(B_R)} \tag{5.3}$$

for all α and for all $k \geq k_0$.

Discussion of the properties of u_{H^2} and $u_{\mathcal{A}}$ in Theorem 5.1 Recall that the solution u itself satisfies the bound (2.8); i.e. $\|u\|_{H^2_k(B_R)} \leq CkR\,\|f\|_{L^2(B_R)}$. Therefore,

(i) the bound (5.2) on u_{H^2} is one power of k better than the corresponding bound (2.8) on u, and
(ii) the bound (5.3) on $u_{\mathcal{A}}$ has the same k dependence as the bound (2.8) on u — both are governed by C_{sol} — although $u_{\mathcal{A}}$ is C^∞ (with each derivative incurring a power of k), and indeed analytic by Lemma 4.2.

We discuss in Section 5.3 why both of these points are crucial in proving that the hp-FEM does not suffer from the pollution effect.

We see in the proof of Theorem 5.1 in Section 7 that u_{H^2} corresponds to components of u with frequencies $\geq \lambda k$ and $u_{\mathcal{A}}$ corresponds to components of u with frequencies $\leq \lambda k$, where $\lambda > 1$, and the notion of "frequencies" is understood via the Fourier transform. We see in Section 7.4 below that u_{H^2} satisfies the property

(i) above because the Helmholtz operator is "well behaved" (in a sense made precise below) on frequencies $\geq \lambda k$ with $\lambda > 1$. We see in Section 7.3 that $u_{\mathcal{A}}$ satisfies the property (ii) above because a function with a compactly-supported Fourier transform is analytic.

5.2 The *hp*-FEM does not suffer from the pollution effect

Theorem 5.2 (Quasioptimality of the hp-FEM) *Suppose that $\{\mathcal{H}_{h,p}\}_{h>0, p\in\mathbb{Z}^+}$ satisfy Assumptions 4.1 and 4.3. Given $k_0, R_0 > 0$, there exist $C_1, C_2 > 0$ (independent of k, R, h, and p) such that the following holds. If u is the solution of the variational problem* (2.3)*, $k \geq k_0$,*

$$\frac{hk}{p} \leq C_1, \qquad \text{and} \qquad p \geq C_2 \log(kR), \tag{5.4}$$

then the Galerkin solution exists, is unique, and satisfies the quasi-optimal error bound (3.5)*.*

The pollution effect occurs when no choice of the number of degrees of freedom growing like $(kR)^d$ ensures that the quasi-optimal error bound (3.2) holds with $C_{\rm qo}$ independent of k (see [2, Definition 2.1] or [26, Equation 1.5] for more-precise statements of this). Since the number of degrees of freedom of $\mathcal{H}_{h,p}$ is proportional to $(p/(h/R))^d$, if h and p are chosen so that the inequalities (5.4) hold with equality, then the number of degrees of freedom of $\mathcal{H}_{h,p}$ is proportional to $(kR)^d$; i.e. Theorem 5.2 shows that *the hp-FEM does not suffer from the pollution effect.*

Proof of Theorem 5.2 The plan is to show that there exist $C_1, C_2 > 0$ such that if h, k, and p satisfy (5.4), then the inequality (3.4) holds; the result then follows from Lemma 3.4. By Lemma 3.3, we can consider $\mathcal{S}^* f$ to be the solution of (2.1)–(2.2); we then use Theorem 5.1 to split u into u_{H^2} and $u_{\mathcal{A}}$, approximate u_{H^2} using Assumption 4.1 (with $s = 2$), and approximate $u_{\mathcal{A}}$ by Assumption 4.3. By the bounds (4.1) and (5.2), there exists $v_{h,p}^{(1)} \in \mathcal{H}_{h,p}$ such that

$$\begin{aligned}\frac{\|u_{H^2} - v_{h,p}^{(1)}\|_{H_k^1(B_R)}}{\|f\|_{L^2(B_R)}} &\leq C_{\text{approx}_1}\left(1+\frac{hk}{p}\right)\left(\frac{hk}{p}\right) = \frac{\|u\|_{H_k^2(B_R)}}{\|f\|_{L^2(B_R)}} \\ &\leq C_{\text{approx}_1}\left(1+\frac{hk}{p}\right)\left(\frac{hk}{p}\right) C_{\text{split},H^2}.\end{aligned} \tag{5.5}$$

The bound (5.3) implies that $u_{\mathcal{A}}$ satisfies the conditions of Assumption 4.3 with $C_2 := C_{\text{split},\mathcal{A}}$ and $C_1 := C_{\text{sol}}(k, 2R)\,\|f\|_{L^2(B_R)}$. Therefore, by (4.6) and the bound (2.7) on C_{sol}, there exists $C > 0$ and $v_{h,p}^{(2)} \in \mathcal{H}_{h,p}$ such that

$$\frac{\|u_{\mathcal{A}} - v_{h,p}^{(2)}\|_{H_k^1(B_R)}}{\|f\|_{L^2(B_R)}} \leq C_{\text{approx}_2} C \left[\left(\frac{h}{h+\sigma}\right)^p + kR\left(\frac{hk}{\sigma p}\right)^p\right]. \tag{5.6}$$

Let $v_{h,p} := v_{h,p}^{(1)} + v_{h,p}^{(2)}$. Using the triangle inequality and the decomposition $u = u_{H^2} + u_{\mathcal{A}}$ on B_R, we obtain that

$$\eta(\mathcal{H}_{h,p}) \leq C_{\mathrm{approx}_1} \left(1 + \frac{hk}{p}\right) \left(\frac{hk}{p}\right) C_{\mathrm{split},H^2} + C_{\mathrm{approx}_2} C \left[\left(\frac{h}{h+\sigma}\right)^p + kR \left(\frac{hk}{\sigma p}\right)^p \right]. \tag{5.7}$$

Therefore, to prove the bound (3.4) on $\eta(\mathcal{H}_{h,p})$, it is sufficient to prove that the right-hand sides of (5.5) and (5.6) are each $\leq C_{\mathrm{cont}}/4$. To do this, first recall from Lemma 3.1 that C_{cont} is independent of k. We then choose C_1 sufficiently small so that $C_1 \leq \min\{\widetilde{C}, \sigma\}$ (where $\widetilde{C}$ and σ are as in Assumption 4.3) and

$$C_{\mathrm{approx}_1} C_1 (1 + C_1) C_{\mathrm{split},H^2} \leq C_{\mathrm{cont}}/4;$$

observe that, since $hk/p \leq C_1$, this last inequality is the desired bound on the right-hand side of (5.5). Next let $\theta_1 := h/(\sigma + h)$ and $\theta_2 := C_1/\sigma$; observe that $\theta_1 < 1$ by definition, and $\theta_2 < 1$ by the definition of C_1. The right-hand side of (5.6) is then bounded by

$$C_{\mathrm{approx}_2} C \Big[(\theta_1)^p + kR(\theta_2)^p \Big].$$

Since $\theta_1, \theta_2 < 1$, if $p \geq C_2 \log(kR)$ for C_2 sufficiently large, then the decay of $(\theta_2)^p$ beats the growth of kR; thus, with C_2 sufficiently large, the right-hand side of (5.6) can be made $\leq C_{\mathrm{cont}}/4$ and the proof is complete. $\square$

5.3 Discussion of the insight the splitting gives into the pollution effect

This subsection discusses three natural questions:

1. Why is the splitting of Theorem 5.1 needed? That is, why does the no-pollution result not just follow from using the bound (2.8) on u itself to bound $\eta(\mathcal{H}_{h,p})$?
2. How are the properties of u_{H^2} and $u_{\mathcal{A}}$ used to prove Theorem 5.2 (the no-pollution result)?
3. Why does one need $p \to \infty$ to remove the pollution effect?

Regarding 1: inputting the bound (2.8) on u into the approximation result (4.1) in Assumption 4.1, we obtain that

$$\eta(\mathcal{H}_{h,p}) \leq C_{\mathrm{approx}_1} \left(1 + \frac{hk}{p}\right) \left(\frac{hk}{p}\right) CkR, \tag{5.8}$$

which leads to the condition "hk^2R/p sufficiently small" for quasioptimality. The condition "hk^2R sufficiently small" is indeed the observed sharp condition for quasioptimality of the h-FEM when $p = 1$ — see, e.g. [31, Fig. 8] — but in this case, the total number of degrees of freedom grows like $(kR)^{2d}$; i.e. the h-FEM with $p = 1$ suffers from the pollution effect.

Regarding 2 and 3: u_{H^2} satisfying the bound (5.2), which is kR better than the corresponding bound (2.8) on u, allows us to obtain the condition "hk/p sufficiently small" for making the first term on the right-hand side of (5.7) small, as opposed to the condition "hk^2R/p sufficiently small" for making the right-hand side of (5.8) small.

When p is fixed, to make the second term on the right-hand side of (5.7) small, we need "$kR(hk)^p$ sufficiently small", and the same condition is obtained if we use the approximation result of Assumption 4.1 to approximate $u_{\mathcal{A}}$ instead of that of Assumption 4.3 (i.e. if we ignore the fact that $u_{\mathcal{A}}$ is analytic, and just use that $u_{\mathcal{A}} \in H^s(B_R)$ for every $s > 0$). The recent polynomial-approximation results of [23] show that, for fixed p, $\eta(\mathcal{H}_{h,p}) \geq CkR(hk)^p$ (or, more generally, $\eta(\mathcal{H}_{h,p}) \geq CC_{\rm sol}(hk)^p$, where $C_{\rm sol}$ is defined by (2.6)). That is, for fixed p, the condition "$kR(hk)^p$ sufficiently small" is the sharp condition for ensuring that $\eta(\mathcal{H}_{h,p})$ is sufficiently small, and this condition is observed empirically to be the sharp condition required for the Galerkin method to be quasi-optimal with constant independent of k (see, e.g. [11, Figs. 3, 5, and 8] for $p = 1, 2, 3, 4$); i.e. the h-FEM suffers from the pollution effect.

Since $u_{\mathcal{A}}$ is analytic, we can take $p \to \infty$ in the second term on the right-hand side of (5.7), with, importantly, the approximation result of Assumption 4.3 controlling the dependence of the constant on p (as highlighted in the "Motivation for Assumption 4.3" paragraph in Section 4). The growth of kR is then removed by the exponential decrease of $(hk/\sigma p)^p$ when $hk/(\sigma p) < 1$. Note that if kR were replaced by $(kR)^M$ for any fixed $M > 0$, then this growth would also be removed by the exponential decrease of $(hk/\sigma p)^p$ when $hk/(\sigma p) < 1$. Since kR in (5.7) comes from $C_{\rm sol}$, the hp-FEM results in [4, 16, 21, 22, 32, 36, 37] all involve the assumption that $C_{\rm sol}$ is polynomially bounded in kR.

6 Recap of results about the Fourier transform and Fourier multipliers

6.1 The Fourier transform $\mathcal{F}_k$

Given $k > 0$, let

$$\mathcal{F}_k\phi(\xi) := \int_{\mathbb{R}^d} \exp\big(-{\rm i}kx\cdot\xi\big)\phi(x)\,{\rm d}x; \tag{6.1}$$

i.e. $\mathcal{F}_k$ is the standard Fourier transform with frequency variable scaled by k. The reason for including this scaling is that $\mathcal{F}_k$ is then tailor-made to work in the weighted Sobolev spaces H^s_k (see Section 6.2).

Let $\mathscr{S}(\mathbb{R}^d)$ be the Schwartz space of rapidly decreasing, C^∞ functions; i.e.

$$\mathscr{S}(\mathbb{R}^d) := \Big\{\phi \in C^\infty(\mathbb{R}^d) : \sup_{x\in\mathbb{R}^d}\big|x^\alpha\partial^\beta\phi(x)\big| < \infty \text{ for all multiindices } \alpha \text{ and } \beta\Big\}.$$

Let $\mathscr{S}^*(\mathbb{R}^d)$ be the space of continuous linear functionals on $\mathscr{S}(\mathbb{R}^d)$. Recall that $\mathcal{F}_k : \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ (see, e.g. [41, Page 72], [47, Proposition 13.15]); then $\mathcal{F}_k : \mathscr{S}^*(\mathbb{R}^d) \to \mathscr{S}^*(\mathbb{R}^d)$ via the definition that $\langle\mathcal{F}_k\phi, \psi\rangle := \langle\phi, \mathcal{F}_k\psi\rangle$ for $\phi \in \mathscr{S}^*(\mathbb{R}^d)$ and $\psi \in \mathscr{S}(\mathbb{R}^d)$, where $\langle\cdot,\cdot\rangle$ is the duality pairing between $\mathscr{S}^*(\mathbb{R}^d)$ and $\mathscr{S}(\mathbb{R}^d)$. In the next subsection, we consider $\langle\xi\rangle^s\mathcal{F}_k\phi$ for $\phi \in \mathscr{S}^*(\mathbb{R}^d)$ and $\langle\xi\rangle := (1+|\xi|^2)^{1/2}$; this is defined as an element of $\mathscr{S}^*(\mathbb{R}^d)$ by $\langle\langle\xi\rangle^s\mathcal{F}_k\phi, \psi\rangle := \langle\mathcal{F}_k\phi, \langle\xi\rangle^s\psi\rangle$ for $\psi \in \mathscr{S}(\mathbb{R}^d)$ (see, e.g. [47, Propositions 13.14 and 13.17]).

We recall the Fourier inversion theorem

$$\mathcal{F}_k^{-1}\psi(x) := \left(\frac{k}{2\pi}\right)^d \int_{\mathbb{R}^d} \exp\big(\mathrm{i}kx\cdot\xi\big)\psi(\xi)\,\mathrm{d}\xi,$$

the property

$$\mathcal{F}_k\Big(\big(-\mathrm{i}k^{-1}\partial\big)^\alpha\phi\Big)(\xi) = \xi^\alpha\,\mathcal{F}_k\phi(\xi), \tag{6.2}$$

and Plancherel's theorem

$$\|\phi\|_{L^2(\mathbb{R}^d)} = \left(\frac{k}{2\pi}\right)^{d/2}\|\mathcal{F}_k\phi\|_{L^2(\mathbb{R}^d)}. \tag{6.3}$$

6.2 Sobolev spaces weighted with *k*

The natural spaces in which to study solutions of the Helmholtz equation are Sobolev spaces with derivatives weighted with k, as in (2.5). On $\mathbb{R}^d$, these are naturally defined using $\mathcal{F}_k$. For $s\in\mathbb{R}$, let

$$H_k^s(\mathbb{R}^d) := \Big\{u\in\mathscr{S}^*(\mathbb{R}^d),\ \langle\xi\rangle^s\mathcal{F}_k u\in L^2(\mathbb{R}^d)\Big\},\qquad \text{where } \langle\xi\rangle := (1+|\xi|^2)^{1/2},$$

and let

$$|||u|||^2_{H_k^s(\mathbb{R}^d)} := \left(\frac{k}{2\pi}\right)^d\int_{\mathbb{R}^d}\langle\xi\rangle^{2s}|\mathcal{F}_k u(\xi)|^2\,\mathrm{d}\xi. \tag{6.4}$$

Because of (6.2), up to dimension-dependent constants, $|||u|||_{H_k^s(\mathbb{R}^d)}$ defined by (6.4) is equivalent to $\|u\|_{H_k^s(\mathbb{R}^d)}$ defined by (2.5) (with B_R replaced by $\mathbb{R}^d$).

6.3 Fourier multipliers

The *Fourier multiplier* given by a function a is

$$\big(a(k^{-1}D)v\big)(x) := \mathcal{F}_k^{-1}\Big(a(\cdot)(\mathcal{F}_k v)(\cdot)\Big)(x); \tag{6.5}$$

i.e. we multiply the Fourier transform of v by a, and then apply the inverse Fourier transform. The rationale for the notation $a(k^{-1}D)$ is that, by (6.2), if $D := -\mathrm{i}\partial$ then $\mathcal{F}_k$ maps $k^{-1}D$ to ξ. Our motivation for studying Fourier multipliers is that (i) the operator $-k^{-2}\Delta - 1$ is one: by the derivative rule (6.2), $-k^{-2}\Delta-1 = p(k^{-1}D)$ where $p(\xi) := |\xi|^2-1$, and (ii) the functions u_{H^2} and $u_{\mathcal{A}}$ in our proof of Theorem 5.1 are defined by Fourier multipliers acting on u (see (7.3) below).

We say that a is a *Fourier symbol of order m* if there exists $C>0$ such that

$$|a(\xi)| \le C\langle\xi\rangle^m \qquad \text{for all } \xi\in\mathbb{R}^d, \tag{6.6}$$

where recall that $\langle\xi\rangle := (1+|\xi|^2)^{1/2}$. We use the (non-standard) notation that $a\in(FS)^m$.

Example 6.1 (Examples of Fourier symbols and multipliers)

(i) If $a(\xi)=1$, then $a\in(FS)^0$ with $a(k^{-1}D)v(x)=v(x)$ (since $\mathcal{F}_k^{-1}\mathcal{F}_k = I$).
(ii) If $p(\xi) := |\xi|^2-1$ then $p\in(FS)^2$ with $p(k^{-1}D)v = (-k^{-2}\Delta-1)v$.

(iii) If χ is bounded and has compact support, then $\chi \in (FS)^{-N}$ for all $N \geq 1$ and $1 - \chi \in (FS)^0$.

Theorem 6.2 (Composition and mapping properties of Fourier multipliers) *If $a \in (FS)^{m_a}$ and $b \in (FS)^{m_b}$ then the following properties hold.*

(i) $ab \in (FS)^{m_a+m_b}$.
(ii) $a(k^{-1}D)b(k^{-1}D) = (ab)(k^{-1}D) = (ba)(k^{-1}D) = b(k^{-1}D)a(k^{-1}D)$.
(iii) $a(k^{-1}D) : H_k^s(\mathbb{R}^d) \to H_k^{s-m_a}(\mathbb{R}^d)$ *and, with C the constant in (6.6), for all $s \in \mathbb{R}$ and $k > 0$,*

$$\left|\left|\left| a(k^{-1}D) \right|\right|\right|_{H_k^s(\mathbb{R}^d) \to H_k^{s-m_a}(\mathbb{R}^d)} \leq C; \tag{6.7}$$

i.e. $a(k^{-1}D)$ is bounded uniformly in both k and s as an operator from H_k^s to $H_k^{s-m_a}$.

Proof Part (i) follows directly from the definition (6.6), part (ii) follows directly from the definition (6.5) and the fact that multiplication in $\mathbb{C}$ is commutative, and part (iii) follows from the definitions of $\left|\left|\left| \cdot \right|\right|\right|_{H_k^s(\mathbb{R}^d)}$ (6.4) and $a(k^{-1}D)v$ (6.5). □

The key result from this section used in the proof of the bound Eq. 5.2 on u_{H^2} is the following.

Theorem 6.3 (Factoring an "elliptic" Fourier multiplier out of another) *Suppose that $a \in (FS)^{m_a}$, $b \in (FS)^{m_b}$, and there exists $c > 0$ such that*

$$|b(\xi)| \geq c\langle\xi\rangle^{m_b} \quad \textit{for } \xi \in \operatorname{supp} a. \tag{6.8}$$

Then

$$a(k^{-1}D) = q(k^{-1}D)b(k^{-1}D) \tag{6.9}$$

where $q \in (FS)^{m_a-m_b}$ is defined by $q(\xi) := a(\xi)/b(\xi)$.

Proof The fact that q is in $(FS)^{m_a-m_b}$ follows from the fact that $a \in (FS)^{m_a}$ and the bound (6.8). The result (6.9) then follows from part (ii) of Theorem 6.2. □

We now combine Theorem 6.3 and the mapping property (6.7) to obtain the following result (which we use in the proof of Corollary 2.3); we highlight that a similar combination of these results in used in Section 7 in the proof of the bound (5.2) on u_{H^2}.

Corollary 6.4 (Elliptic regularity) *If $(-k^{-2}\Delta + 1)v \in L^2(\mathbb{R}^d)$ then $v \in H^2(\mathbb{R}^d)$ with*

$$\left|\left|\left| v \right|\right|\right|_{H_k^2(\mathbb{R}^d)} \leq \left\| (-k^{-2}\Delta + 1)v \right\|_{L^2(\mathbb{R}^d)}.$$

Proof We apply Theorem 6.3 with $a(\xi) = 1$ and $b(\xi) = |\xi|^2 + 1$, so that $a(k^{-1}D) = I$ and $b(k^{-1}D) = -k^{-2}\Delta + 1$. The theorem implies that $q(\xi) := \langle\xi\rangle^{-2} \in (FS)^{-2}$ (this is also clear from the definition (6.6)). Then, by the mapping property (6.7),

$$\left\|\!\left|v\right|\!\right\|_{H_k^2(\mathbb{R}^d)} = \left\|\!\left|q(k^{-1}D)(-k^{-2}\Delta + 1)v\right|\!\right\|_{H_k^2(\mathbb{R}^d)} \leq \left\|(-k^{-2}\Delta + 1)v\right\|_{L^2(\mathbb{R}^d)}.$$

□

7 Proof of Theorem 5.1

7.1 Definition of high- and low-frequency cut-offs

Let

$$\chi_\lambda(\xi) = 1_{|\xi|\leq\lambda}(\xi) = \begin{cases} 1 & \text{for } |\xi| \leq \lambda, \\ 0 & \text{for } |\xi| > \lambda. \end{cases} \tag{7.1}$$

We define the low-frequency cut-off Π_L by

$$\Pi_L := \chi_\lambda(k^{-1}D); \qquad \text{i.e.} \qquad \Pi_L v = \mathcal{F}_k^{-1}\big(\chi_\lambda(\cdot)(\mathcal{F}_k v)(\cdot)\big), \tag{7.2}$$

by the definition (6.5) of a Fourier multiplier. We see that Π_L acting on a function v returns the frequencies of v that are $\leq \lambda$; hence, why we call it a low-frequency cut-off.[1] We define the high-frequency cut-off Π_H by

$$\Pi_H := I - \Pi_L = (1 - \chi_\lambda)(k^{-1}D);$$

i.e. Π_H acting on a function v returns the frequencies of v that are $\geq \lambda$.

7.2 Definition of u_{H^2} and $u_{\mathcal{A}}$ via the frequency cut-offs

Let u be as in Theorem 5.1; i.e. u is the outgoing solution of the Helmholtz equation (2.1). Let $\varphi \in C^\infty_{\text{comp}}(\mathbb{R}^d, [0, 1])$ be equal to one on B_R and vanish outside B_{2R}, and set

$$u_{\mathcal{A}} := \big(\Pi_L(\varphi u)\big)\big|_{B_R} \quad \text{and} \quad u_{H^2} := \big(\Pi_H(\varphi u)\big)\big|_{B_R}. \tag{7.3}$$

Since $\Pi_L + \Pi_H = I$, these definitions imply that, on B_R, $u_{\mathcal{A}} + u_{H^2} = \varphi u = u$; i.e. (5.1) holds. This splitting contains the arbitrary parameter λ; we fix this when proving the bound (5.2) on u_{H^2}.

7.3 Proof of the bound (5.3) on $u_{\mathcal{A}}$

The idea of the proof, in short, is that *a function with a compactly supported Fourier transform is analytic*. Plancherel's theorem (6.3), the derivative property (6.2), and the definition of Π_L (7.2) imply that

$$\left\|(k^{-1}\partial)^\alpha\big(\Pi_L\varphi u\big)\right\|_{L^2(\mathbb{R}^d)} = \left(\frac{k}{2\pi}\right)^{d/2} \|(\cdot)^\alpha \chi_\lambda(\cdot)\mathcal{F}_k(\varphi u)(\cdot)\|_{L^2(\mathbb{R}^d)}. \tag{7.4}$$

[1] The definition of $\mathcal{F}_k$ (6.1) implies that if $\mathcal{F}_k w$ is supported on $|\xi| \leq \lambda$, then the "standard" Fourier transform (i.e. with the transform variable not scaled by k) of w is supported for $|\zeta| \leq \lambda k$.

The definition of χ_λ (7.1) implies that $\left|\xi^\alpha \chi_\lambda(\xi)\right| \le \lambda^{|\alpha|}$ for all $\xi \in \mathbb{R}^d$. Using this fact in (7.4) and then (in this order) Plancherel's theorem (6.3), the fact that $\varphi = 0$ outside B_{2R} and $\varphi \le 1$ inside B_{2R}, and the definition of $C_{\rm sol}$ (2.6), we find that

$$\left\|(k^{-1}\partial)^\alpha\big(\Pi_L\varphi u\big)\right\|_{L^2(\mathbb{R}^d)} \le \left(\frac{k}{2\pi}\right)^{d/2} \lambda^{|\alpha|}\left\|\mathcal{F}_k(\varphi u)\right\|_{L^2(\mathbb{R}^d)} \le \lambda^{|\alpha|}\left\|\varphi u\right\|_{L^2(\mathbb{R}^d)}$$
$$\le \lambda^{|\alpha|}\left\|u\right\|_{L^2(B_{2R})} \le \lambda^{|\alpha|} C_{\rm sol}(k,2R)\left\|f\right\|_{L^2(B_R)}. \quad (7.5)$$

By the definition (7.3) of $u_{\mathcal{A}}$,

$$\left\|(k^{-1}\partial)^\alpha u_{\mathcal{A}}\right\|_{L^2(B_R)} = \left\|(k^{-1}\partial)^\alpha\big(\Pi_L\varphi u\big)\right\|_{L^2(B_R)} \le \left\|(k^{-1}\partial)^\alpha\big(\Pi_L\varphi u\big)\right\|_{L^2(\mathbb{R}^d)};$$

the bound (5.3) then follows from (7.5) with $C_{\mathrm{split},\mathcal{A}} := \lambda$. We make two remarks.

- The fact that $\Pi_L\varphi u \in H^s_k(\mathbb{R}^d)$ for all s follows from the fact that $\chi_\lambda \in (FS)^{-N}$ for every $N > 0$ (from Part (iii) of Example 6.1) and the mapping property in Part (iii) of Theorem 6.2; we give the direct proof above, however, to have explicit control on the constants in the bound (to show that $\Pi_L\varphi u$ is actually analytic).
- The fact that $u_{\mathcal{A}}$ comes from a function with a compactly supported Fourier transform, and hence automatically is analytic, is one of the advantages of the current splitting compared to the original splitting in [34] (see the discussion in Section 7.5).

7.4 Proof of the bound (5.2) on u_{H^2}

The idea of the proof is to use Theorem 6.3, using the fact that *the Helmholtz operator is an elliptic Fourier multiplier (in the sense of* (6.8)*) on high frequencies*, and thus, if λ is sufficiently large, on the support of the high-frequency cut-off Π_H. By the definition of u_{H^2} (7.3) and the equivalence of $\|\cdot\|_{H^2_k}$ and $|||\cdot|||_{H^2_k}$ described in Section 6.2,

$$\left\|u_{H^2}\right\|_{H^2_k(B_R)} = \left\|\Pi_H(\varphi u)\right\|_{H^2_k(B_R)} \le \left\|\Pi_H(\varphi u)\right\|_{H^2_k(\mathbb{R}^d)} \lesssim \left|\left|\left|\Pi_H(\varphi u)\right|\right|\right|_{H^2_k(\mathbb{R}^d)},$$

where we use the notation that $A \lesssim B$ if there exists $C > 0$, independent of k and R, such that $A \le CB$. It is therefore sufficient to prove that

$$\left|\left|\left|\Pi_H(\varphi u)\right|\right|\right|_{H^2_k(\mathbb{R}^d)} \lesssim \left\|f\right\|_{L^2(B_R)} \qquad \text{for all } k \ge k_0. \quad (7.6)$$

Lemma 7.1 ("Ellipticity" of $p(\xi) := |\xi|^2 - 1$ when $|\xi| \ge \lambda > 1$) *If* $\lambda \ge \lambda_0 > 1$ *then there exists* $C > 0$ *such that*

$$\left||\xi|^2 - 1\right| \ge C\langle\xi\rangle^2 \quad \textit{for } |\xi| \ge \lambda. \quad (7.7)$$

Proof It is straightforward to check that (7.7) holds with $C := (1 + 2(\lambda_0^2 - 1)^{-1})^{-1}$. □

Corollary 7.2 (Factoring out $p(k^{-1}D)$ from $(1-\chi_\lambda)(k^{-1}D)$) *Let* $p(\xi) := |\xi|^2 - 1$. *If* $\lambda \geq \lambda_0 > 1$, *then there exists* $q \in (FS)^{-2}$ *such that*

$$(1-\chi_\lambda)(k^{-1}D) = q(k^{-1}D)\, p(k^{-1}D).$$

Proof By the definition (7.1) of χ_λ, $\operatorname{supp}(1-\chi_\lambda) = \{\xi : |\xi| \geq \lambda\}$. The result follows by applying Theorem 6.3 with $a(\xi) := (1-\chi_\lambda)(\xi)$ and $b(\xi) := p(\xi) = |\xi|^2 - 1$, since (7.7) implies that the inequality (6.8) holds (i.e. b is elliptic on $\operatorname{supp} a$). $\square$

We now use Corollary 7.2 and the mapping property (6.7) to prove the bound (7.6). By the definition of u_{H^2} (7.3) and Corollary 7.2,

$$\left\|\left|\Pi_H(\varphi u)\right|\right\|_{H^2_k} = \left\|\left|\big(1-\chi_\mu\big)(k^{-1}D)(\varphi u)\right|\right\|_{H^2_k} = \left\|\left|q(k^{-1}D)\, p(k^{-1}D)(\varphi u)\right|\right\|_{H^2_k}$$

with $q \in (FS)^{-2}$, and, by the mapping property (6.7),

$$\left\|\left|q(k^{-1}D)\, p(k^{-1}D)(\varphi u)\right|\right\|_{H^2_k(\mathbb{R}^d)} \lesssim \left\|p(k^{-1}D)(\varphi u)\right\|_{L^2(\mathbb{R}^d)};$$

the key point is that we now have the Helmholtz operator $p(k^{-1}D)$ on the right-hand side, and we can start to relate this side to $f = p(k^{-1}D)u$. For brevity, let $P := -k^{-2}\Delta - 1 = p(k^{-1}D)$. Then

$$\begin{aligned} \left\|\left|\Pi_H(\varphi u)\right|\right\|_{H^2_k(\mathbb{R}^d)} \lesssim \|P(\varphi u)\|_{L^2(\mathbb{R}^d)} &= \|\varphi P u + [P,\varphi]u\|_{L^2(\mathbb{R}^d)} \\ &\leq \|f\|_{L^2(\mathbb{R}^d)} + \|[P,\varphi]u\|_{L^2(\mathbb{R}^d)}, \end{aligned} \tag{7.8}$$

where we have used the fact that $\varphi \equiv 1$ on $\operatorname{supp} f$, and where the commutator $[A, B]$ is defined as $AB - BA$. By direct calculation,

$$[P,\varphi]u = -k^{-2}\big(u\Delta\varphi + 2\nabla\varphi\cdot\nabla u\big),$$

so that

$$\left\|[P,\varphi]u\right\|_{L^2(\mathbb{R}^d)} \lesssim (kR)^{-1}\,\|u\|_{H^1_k(B_{2R})}, \tag{7.9}$$

where we have used that the definition of φ in Section 7.2 implies that $|\nabla\varphi| \sim R^{-1}$ and $|\Delta\varphi| \sim R^{-2}$. Therefore, by combining (7.8) and (7.9), and using (2.8) (with R replaced by $2R$), we have

$$\|\Pi_H(\varphi u)\|_{H^2_k(\mathbb{R}^d)} \lesssim \|f\|_{L^2(\mathbb{R}^d)} + (kR)^{-1}\,\|u\|_{H^1_k(B_{2R})} \lesssim \|f\|_{L^2(\mathbb{R}^d)},$$

which is the result (7.6).

7.5 Discussion of the original proof of Theorem 5.1 in [34]

The original proof of Theorem 5.1 in [34, Section 3.2] is also based on the idea of frequency cut-offs using the indicator function (7.1). However, in [34, Section 3.2], the frequency cut-offs are applied to the data f (as opposed to φu in our case). Furthermore, the analysis in [34, Section 3.2] is based on writing the solution u to the Helmholtz eqaution (2.1) satisfying the Sommerfeld radiation condition (2.2) as

$$u(x) = k^2 \int_{B_R} \Phi_k(x,y)\, f(y)\,\mathrm{d}y, \tag{7.10}$$

where $\Phi_k(x,y)$ is the fundamental solution (satisfying $(\Delta_y+k^2)\Phi_k(x,y) = -\delta(x-y)$) defined by

$$\Phi_k(x,y) := \frac{\mathrm{i}}{4}H_0^{(1)}\big(k|x-y|\big), \quad d=2, \quad := \frac{\mathrm{e}^{\mathrm{i}k|x-y|}}{4\pi|x-y|}, \quad d=3.$$

The proof in [34, Section 3.2] then considers the function

$$v_\mu(x) := k^2\int_{B_R}\Phi_k(x,y)\,\mu(|x-y|)\,f(y)\,\mathrm{d}y,$$

where $\mu \in C^\infty_{\mathrm{comp}}(\mathbb{R})$ with $\mu|_{[0,2R]} = 1$, $\operatorname{supp}\mu \subset [0,4R]$, and additionally the first and second derivatives of μ satisfying certain bounds (see [34, Equation 3.27]). This definition of μ implies that $v_\mu|_{B_R} = u|_{B_R}$. The advantage of studying v_μ instead of u is that v is the convolution of a compactly-supported kernel with f; thus, v has compact support and its Fourier transform is well defined. (In contrast, in the present paper, we first multiply u by φ to ensure that $\mathcal{F}_k(\varphi u)$ makes sense.) The analysis in [34, Section 3.2 and Appendix A] then proceeds by studying the compactly supported integral kernel in v_μ and the Fourier transform of this kernel, with identities and bounds on Bessel and Hankel functions needed when both $d=2$ and $d=3$.

7.6 Generalising Theorem 5.1 to more-complicated Helmholtz problems

The proof of Theorem 5.1 above can be generalised to more-complicated Helmholtz problems, involving variable coefficients and obstacles, using pseudodifferential operators. Indeed, whereas the operator $-k^{-2}\Delta - 1$ is a Fourier multiplier (by Part (ii) of Example 6.1), its variable coefficient analogue $-k^{-2}\nabla\cdot(A\nabla) - n$ is not; i.e. one cannot write the Fourier transform of $-k^{-2}\nabla\cdot(A\nabla v) - nv$ in terms of the Fourier transform of v. Nevertheless, [32] proves Theorem 5.1, using the same ideas in the proof above, with the Helmholtz equation (2.1) replaced by

$$k^{-2}\nabla\cdot(A\nabla u) + nu = -f \quad \text{in } \mathbb{R}^d. \tag{7.11}$$

This is achieved by replacing the Fourier multipliers in Section 6 by *pseudodifferential operators* (indeed, recall that one of the motivations for the development of the theory of pseudodifferential operators in the 1960s was to study variable-coefficient PDEs such as (7.11) using Fourier analysis). In particular, the pseudodifferential generalisation of the key result of Theorem 6.3 is the so-called elliptic parametrix (see, e.g. [15, Proposition E.32]).

These ideas can also be used to prove an analogue of Theorem 5.1 when the Helmholtz equation (with or without variable coefficients) is posed not in $\mathbb{R}^d$ but outside an impenetrable obstacle. In this case, the Fourier transform (6.1) is no longer available; nevertheless, the *functional calculus* (essentially the idea of expanding in eigenfunctions of the differential operator) can be used to define Fourier-type transforms, tailored to the particular problem. The ideas of the proof of Theorem 5.1 can then be implemented in this situation (albeit with more technicalities, and the requirement that the obstacle is analytic) (see [21, 22]).

Remark 7.3 (The smoothness of the frequency cut-offs) The theory of pseudodifferential operators is easiest when the symbols (i.e. the generalisation of the Fourier symbols in (6.6)) are C^∞. To ensure this, [32] assumes that the coefficients A and n in (7.11) are C^∞, and uses a smooth frequency cut-off; i.e. the indicator function in (7.1) is replaced by $\chi_\lambda \in C^\infty_{\mathrm{comp}}(\mathbb{R}^d; [0,1])$ such that $\chi_\lambda(\xi) = 1$ for $|\xi| \leq \lambda$ and $\chi_\lambda(\xi) = 0$ for, say, $|\xi| \geq 2\lambda$. The reader can check that the proof in Section 7 goes through as before with this smooth cut-off (with $C_{\mathrm{split},\mathcal{A}}$ in Section 7.3 changed from λ to 2λ).

Appendix A. Proof of Theorem 2.2

Lemma A.1 *(*Morawetz identity for the Helmholtz operator [43, Section I.2]*) If*

$$\mathcal{L}v := k^{-2}\Delta v + v \quad \text{and} \quad \mathcal{M}_{\beta,\alpha} v := x\cdot\nabla v - \mathrm{i}k\beta v + \alpha v, \tag{A.1}$$

with β and α real-valued C^1 functions, then

$$\begin{aligned}2\Re\big(\overline{\mathcal{M}_{\beta,\alpha} v}\,\mathcal{L}v\big) &= \nabla\cdot\Big[2k^{-1}\Re\big(\overline{\mathcal{M}_{\beta,\alpha} v}\,k^{-1}\nabla v\big) + \big(|v|^2 - k^{-2}|\nabla v|^2\big)x\Big]\\ &\quad -2\Re\big(\overline{v}\,(\mathrm{i}\nabla\beta + k^{-1}\nabla\alpha)\cdot k^{-1}\nabla v\big) - \big(d-2\alpha\big)|v|^2 - \big(2\alpha - d + 2\big)k^{-2}|\nabla v|^2.\end{aligned} \tag{A.2}$$

Proof This follows in a straightforward (but slightly involved) way by expanding the divergence on the right-hand side; for this done step-by-step, see, e.g. [51, Proof of Lemma 2.1]. □

The idea of the proof of Theorem 2.2 is to integrate the identity (A.2) over $\mathbb{R}^d$ with $v = u$, $\alpha = (d-1)/2$, and β defined piecewise as $\beta = R$ for $r \leq R$ and $\beta = r$ for $r \geq R$. The choice $\beta =$ constant and $\alpha = (d-1)/2$ means that the non-divergence terms on the right-hand side of (A.2) become $-|u|^2 - k^{-2}|\nabla u|^2$; this is where we get $\|u\|^2_{H^1_k(B_R)}$ from. The choice $\beta = r$ deals with the contribution from infinity (although this is not immediately clear from (A.2)). We therefore first look at the special case of (A.2) with $\beta = r$.

Lemma A.2 *(*Special case of (A.2) [42, Equation 1.2]*) With $\mathcal{L}v$ and $\mathcal{M}_{\beta,\alpha}v$ as in Lemma A.1, show that if $\alpha \in \mathbb{R}$, then, with $v_r = x\cdot\nabla v/r$,*

$$\begin{aligned}2\Re\big(\overline{\mathcal{M}_{r,\alpha} v}\mathcal{L}v\big) &= \nabla\cdot\Big[2k^{-1}\Re\big(\overline{\mathcal{M}_{r,\alpha} v}\,k^{-1}\nabla v\big) + \Big(|v|^2 - k^{-2}|\nabla v|^2\Big)\,x\Big] - \big|k^{-1}v_r - \mathrm{i}v\big|^2\\ &\quad +\big(2\alpha - (d-1)\big)\big(|v|^2 - k^{-2}|\nabla v|^2\big) - k^{-2}\big(|\nabla v|^2 - |v_r|^2\big).\end{aligned} \tag{A.3}$$

Proof This follows from (A.2) by choosing $\beta = r$ and writing the term involving $\nabla\beta$ as

$$-2\Re\big(\mathrm{i}\overline{v}k^{-1}v_r\big) = -2\Re\big(\overline{(-\mathrm{i}v)}k^{-1}v_r\big) = |v|^2 + k^{-2}|v_r|^2 - \big|k^{-1}v_r - \mathrm{i}v\big|^2$$

using the identity $-2\Re(z_1\overline{z_2}) = |z_1|^2 + |z_2|^2 - |z_1 + z_2|^2$. □

To integrate (A.3) over $\mathbb{R}^d \setminus B_R$, we integrate it over $B_{R_1} \setminus B_R$ and then send $R_1 \to \infty$. In preparation for this, we look at the boundary term on ∂B_{R_1}.

Lemma A.3 *Let*

$$Q_{r,\alpha}(v) := 2k^{-1}\Re\big(\overline{\mathcal{M}_{r,\alpha} v}\, k^{-1}\nabla v\big) + \big(|v|^2 - k^{-2}|\nabla v|^2\big)x. \tag{A.4}$$

If u is an outgoing solution of $\mathcal{L}u = 0$ in $\mathbb{R}^d \setminus \overline{B_{R_0}}$, then, for all $\alpha \in \mathbb{R}$,

$$\int_{\Gamma_{\partial B_{R_1}}} Q_{R_1,\alpha}(u) \cdot \widehat{x} \to 0 \quad \text{as } R_1 \to \infty. \tag{A.5}$$

The proof of Lemma A.3 requires the following classic result.

Theorem A.4 (Atkinson–Wilcox expansion (see, e.g. [13, Theorem 3.7])) *If $u \in H^1_{loc}(\mathbb{R}^d \setminus \overline{B_{R_0}})$ is an outgoing solution of $k^{-2}\Delta u + u = 0$ in $\mathbb{R}^d \setminus \overline{B_{R_0}}$ for some $R_0 > 0$, then there exist smooth functions F_n such that, for any $R_1 > R_0$,*

$$u(x) = \frac{\mathrm{e}^{\mathrm{i}kr}}{r^{(d-1)/2}} \sum_{n=0}^{\infty} \frac{F_n(x/r)}{r^n} \quad \text{for } r := |x| \geq R_1 \tag{A.6}$$

where the sum in (A.6) (and all its derivatives) converges absolutely and uniformly.

Proof of Lemma A.3 By the definitions of $Q_{r,\alpha}(v)$ (A.4) and $\mathcal{M}_{r,\alpha}v$ (A.1),

$$\begin{aligned} Q_{r,\alpha}(u) \cdot \widehat{x} &= rk^{-2}\left(2\Re\left(u_r\overline{\left(u_r - \mathrm{i}ku + \frac{\alpha}{r}u\right)}\right) + k^2|u|^2 - |\nabla u|^2\right) \\ &= rk^{-2}\left(|u_r|^2 + 2\Re\left(u_r\overline{\left(-\mathrm{i}ku + \frac{\alpha}{r}u\right)}\right) + k^2|u|^2 - \big(|\nabla u|^2 - |u_r|^2\big)\right) \\ &= rk^{-2}\left(\frac{|\mathcal{M}_{r,\alpha}u|^2}{r^2} + \alpha^2\frac{|u|^2}{r^2} - \big(|\nabla u|^2 - |u_r|^2\big)\right), \end{aligned} \tag{A.7}$$

where we have again used the identity $2\Re(z_1\overline{z_2}) = |z_1+z_2|^2 - |z_1|^2 - |z_2|^2$. We now claim that the term in large brackets in (A.7) is $O(r^{-d-1})$; if this is true, then

$$\int_{\Gamma_{\partial B_{R_1}}} Q_{R_1,\alpha}(u) \cdot \widehat{x} = O\left(\frac{1}{R_1}\right) \text{ as } R_1 \to \infty,$$

and thus (A.5) follows. By the Atkinson–Wilcox expansion (A.6), $|u|^2 = O(r^{1-d})$ and $r^{-2}|\mathcal{M}_{r,\alpha}u|^2 = O(r^{-d-1})$. To prove the result, therefore, we only need to show that $|\nabla u|^2 - |u_r|^2 = O(r^{-d-1})$. The quantity $|\nabla u|^2 - |u_r|^2$ equals $|\nabla_S u|^2$ where ∇_S is the surface gradient on $|x| = r$, which satisfies $\nabla_S u = \nabla u - \widehat{x}u_r$. This differential operator is equal to $1/r$ multiplied by an operator acting only on $\widehat{x}$, i.e. the angular variables; thus, $|\nabla_S u|^2$ is $O(r^{-d-1})$ and the proof is complete. □

We now integrate (A.3) over $B_{R_1} \setminus B_R$, send $R_1 \to \infty$, and obtain an inequality involving the boundary term on ∂B_R.

Lemma A.5 *If u is an outgoing solution of $\mathcal{L}u = 0$ in $\mathbb{R}^d \setminus \overline{B_{R_0}}$, for some $R_0 > 0$, then, for $R > R_0$,*

$$\int_{\partial B_R} Q_{R,(d-1)/2}(u) \cdot \widehat{x} \le 0. \tag{A.8}$$

Proof We integrate the identity (A.3) over $B_{R_1} \setminus B_R$, where $R_1 > R$, with $v = u$ and then use the divergence theorem $\int_D \nabla \cdot F = \int_{\partial D} F$. The divergence theorem is valid for $F \in C^\infty(\overline{D})$ and D Lipschitz (see, e.g. [41, Theorem 3.34]); we can use it here since, by elliptic regularity, $u \in C^\infty(\overline{B_{R_1} \setminus B_R})$ (see, e.g. [41, Theorem 4.16]). This results in

$$\int_{\partial B_{R_1}} Q_{R_1,\alpha}(u) \cdot \widehat{x} - \int_{\partial B_R} Q_{R,\alpha}(u) \cdot \widehat{x} = \int_{B_{R_1} \setminus B_R} -\big(2\alpha - (d-1)\big)\big(|v|^2 - k^{-2}|\nabla v|^2\big) + k^{-2}\big(|\nabla v|^2 - |v_r|^2\big) + \big|k^{-1}v_r - \mathrm{i}v\big|^2. \tag{A.9}$$

Setting $\alpha = (d-1)/2$ eliminates the first term on the right-hand side of (A.9). Since $|v_r| \le |\nabla v|$, the remaining terms on the right-hand side of (A.9) are non-negative, and thus

$$\int_{\partial B_{R_1}} Q_{R_1,(d-1)/2}(u) \cdot \widehat{x} - \int_{\partial B_R} Q_{R,(d-1)/2}(u) \cdot \widehat{x} \ge 0.$$

Sending $R_1 \to \infty$ and using (A.5), we obtain the result (A.8). □

Proof of Theorem 2.2 The plan is to integrate the identity (A.2) over B_R with $v = u$, $\beta = R$, and $\alpha = (d-1)/2$, and then use the divergence theorem. We justify using the divergence theorem just as we did at the beginning of the proof of Lemma A.3 to find that if $v \in C^\infty(\overline{D})$ then

$$\int_{B_R} 2\Re\big(\overline{\mathcal{M}_{R,(d-1)/2}v}\,\mathcal{L}v\big) = \int_{\partial B_R} 2k^{-1}\Re\big(\overline{\mathcal{M}_{R,(d-1)/2}v}\,k^{-1}\nabla v\big) + \big(|v|^2 - k^{-2}|\nabla v|^2\big)x - \int_{B_R} \big(|v|^2 + k^{-2}|\nabla v|^2\big). \tag{A.10}$$

We now claim that (A.10) holds for $v \in H^2(B_R)$; this follows since $C^\infty(\overline{D})$ is dense in $H^2(D)$ [41, Page 77], and, by the trace theorem (see, e.g. [41, Theorem 3.37]), (A.10) is continuous in v with respect to the topology of $H^2(B_R)$. By Corollary 2.3, $u \in H^2(B_R)$, and thus (A.10) holds with $v = u$. Using in (A.10) the definition of Q (A.4), the fact that u is an outgoing solution of the Helmholtz equation (2.1), and Lemma A.5, we find that

$$-\int_{B_R} 2\Re\big(\overline{\mathcal{M}_{R,(d-1)/2}u}\,f\big) + \|u\|^2_{H^1_k(B_R)} = \int_{\partial B_R} Q_{R,(d-1)/2}(u) \cdot \widehat{x} \le 0.$$

Thus

$$\|u\|^2_{H^1_k(B_R)} \le 2\big\|\mathcal{M}_{R,(d-1)/2}u\big\|_{L^2(B_R)} \|f\|_{L^2(B_R)}\,. \tag{A.11}$$

By the inequality $|a+b|^2 \leq 2|a|^2 + 2|b|^2$, the fact that $|\mathrm{i}kr+\alpha|^2 = k^2r^2+\alpha^2$, and the bound $r \leq R$ on B_R,

$$\begin{aligned}\left\|\mathcal{M}_{R,(d-1)/2}u\right\|^2_{L^2(B_R)} &\leq 2R^2\,\|\nabla u\|^2_{L^2(B_R)} + 2\big((kR)^2+\alpha^2\big)\,\|u\|^2_{L^2(B_R)}\\ &\leq 2k^2R^2\left(1+\frac{\alpha^2}{k^2R^2}\right)\|u\|^2_{H^1_k(B_R)}.\end{aligned}$$

Using this in (A.11), and recalling that $\alpha = (d-1)/2$, we find that $\|u\|_{H^1_k(B_R)} \leq C\,\|f\|_{L^2(B_R)}$ with C the right-hand side of (2.7). $\square$

Acknowledgements The idea for this paper came out of the "Nachdiplom" lecture course I taught at ETH Zürich in Fall 2021; it is a pleasure to thank the Institute for Mathematical Research (FIM) and the Seminar for Applied Mathematics at ETH for their hospitality during that time. In particular, the proof of Theorem 5.1 presented in this paper was the result of Ralf Hiptmair asking me about the minimum technicalities needed for the proof in [32] of the analogous splitting in the variable-coefficient case. I also thank Martin Averseng for many useful discussions in his role as the teaching assistant for this course. Finally, I thank Alastair Spence (University of Bath) and Jared Wunsch (Northwestern University) for their insightful comments on a draft of this paper.

Funding I acknowledge support from EPSRC grant EP/R005591/1.

Declarations

Conflict of interest The author declares no competing interests.

References

1. Aziz, A.K., Kellogg, R.B., Stephens, A.B.: A two point boundary value problem with a rapidly oscillating solution. Numer. Math. **53**(1), 107–121 (1988)
2. Babuška, I.M., Sauter, S.A.: Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers? SIAM Rev., 451–484 (2000)
3. Barucq, H., Chaumont-Frelet, T., Gout, C.: Stability analysis of heterogeneous Helmholtz problems and finite element solution based on propagation media approximation. Math. Comp. **86**(307), 2129–2157 (2017). https://doi.org/10.1090/mcom/3165
4. Bernkopf, M., Chaumont-Frelet, T., Melenk, J.M.: Wavenumber-explicit stability and convergence analysis of *hp* finite element discretizations of Helmholtz problems in piecewise smooth media. arXiv:2209.03601 (2022)
5. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, 3rd edn. Texts in Applied Mathematics, vol. 15. Springer (2008)
6. Brown, D.L., Gallistl, D., Peterseim, D.: Multiscale Petrov-Galerkin method for high-frequency heterogeneous Helmholtz equations. In: Meshfree Methods for Partial Differential Equations VIII, pp. 85–115. Springer (2017)

7. Cao, H., Wu, H.: IPCDGM and multiscale IPDPGM for the Helmholtz problem with large wave number. J. Comput. Appl. Math. **369**, 112590 (2020)
8. Chandler-Wilde, S.N., Monk, P.: Wave-number-explicit bounds in time-harmonic scattering. SIAM J. Math. Anal. **39**(5), 1428–1455 (2008)
9. Chaumont-Frelet, T.: On high order methods for the heterogeneous Helmholtz equation. Comput. Math. Appl. **72**(9), 2203–2225 (2016)
10. Chaumont-Frelet, T., Valentin, F.: A multiscale hybrid-mixed method for the Helmholtz equation in heterogeneous domains. SIAM J. Num. Anal. **58**(2), 1029–1067 (2020)
11. Chaumont-Frelet, T., Nicaise, S.: Wavenumber explicit convergence analysis for finite element discretizations of general wave propagation problem. IMA J. Numer. Anal. **40**(2), 1503–1543 (2020)
12. Ciarlet, P.G.: Basic error estimates for elliptic problems. In: Handbook of Numerical Analysis, Vol. II. Handb. Numer. Anal., II, pp. 17–351, North-Holland (1991)
13. Colton, D.L., Kress, R.: Integral Equation Methods in Scattering Theory, p. 271. Wiley, New York (1983)
14. Du, Y., Wu, H.: Preasymptotic error analysis of higher order FEM and CIP-FEM for Helmholtz equation with high wave number. SIAM J. Numer. Anal. **53**(2), 782–804 (2015)
15. Dyatlov, S., Zworski, M.: Mathematical Theory of Scattering Resonances. AMS (2019)
16. Esterhazy, S., Melenk, J.M.: On stability of discretizations of the Helmholtz equation. In: Graham, I.G., Hou, T.Y., Lakkis, O., Scheichl, R. (eds.) Numerical Analysis of Multiscale Problems. Lecture Notes in Computational Science and Engineering, vol. 83, pp. 285–324. Springer (2012)
17. Feng, X., Wu, H.: Discontinuous Galerkin methods for the Helmholtz equation with large wave number. SIAM J. Numer. Anal. **47**(4), 2872–2896 (2009)
18. Feng, X., Wu, H.: *hp*-Discontinuous Galerkin methods for the Helmholtz equation with large wave number. Math. Comput. **80**(276), 1997–2024 (2011)
19. Gallistl, D., Chaumont-Frelet, T., Nicaise, S., Tomezyk, J.: Wavenumber explicit convergence analysis for finite element discretizations of time-harmonic wave propagation problems with perfectly matched layers. Comm. Math. Sci. **20**(1), 1–52 (2022)
20. Gallistl, D., Peterseim, D.: Stable multiscale Petrov–Galerkin finite element method for high frequency acoustic scattering. Comput. Methods Appl. Mech. Eng. **295**, 1–17 (2015)
21. Galkowski, J., Lafontaine, D., Spence, E.A., Wunsch, J.: Decompositions of high-frequency Helmholtz solutions via functional calculus, and application to the finite element method. SIAM J. Math. Anal. to appear. arXiv:2102.13081 (2021)
22. Galkowski, J., Lafontaine, D., Spence, E.A., Wunsch, J.: The *hp*-FEM applied to the Helmholtz equation with PML truncation does not suffer from the pollution effect. arXiv:2207.05542 (2022)
23. Galkowski, J.: Lower bounds for piecewise polynomial approximations of oscillatory functions. arXiv:2211.04757 (2022)
24. Galkowski, J., Spence, E.A.: Sharp preasymptotic error bounds for the Helmholtz *h*-FEM. arXiv:2301.03574 (2023)
25. Galkowski, J., Spence, E.A., Wunsch, J.: Optimal constants in nontrapping resolvent estimates. Pure Appl. Anal. **2**(1), 157–202 (2020)
26. Galkowski, J., Spence, E.A.: Does the Helmholtz boundary element method suffer from the pollution effect? SIAM Review to appear (2023)
27. Graham, I.G., Sauter, S.: Stability and finite element error analysis for the Helmholtz equation with variable coefficients. Math. Comp. **89**(321), 105–138 (2020)
28. Graham, I.G., Pembery, O.R., Spence, E.A.: The Helmholtz equation in heterogeneous media: a priori bounds, well-posedness, and resonances. J. Differ. Equ. **266**(6), 2869–2923 (2019)
29. Ihlenburg, F., Babuška, I.: Dispersion analysis and error estimation of Galerkin finite element methods for the Helmholtz equation. Int. J. Numer. Meth. Eng. **38**(22), 3745–3774 (1995)
30. Ihlenburg, F., Babuska, I.: Finite element solution of the Helmholtz equation with high wave number part II: the *hp* version of the FEM. SIAM J. Numer. Anal. **34**(1), 315–358 (1997)
31. Ihlenburg, F., Babuška, I.: Finite element solution of the Helmholtz equation with high wave number Part I: The h-version of the FEM. Comp. Math. Appl. **30**(9), 9–37 (1995)
32. Lafontaine, D., Spence, E.A., Wunsch, J.: Wavenumber-explicit convergence of the *hp*-FEM for the full-space heterogeneous Helmholtz equation with smooth coefficients. Comp. Math. Appl. **113**, 59–69 (2022)
33. Li, Y., Wu, H.: FEM and CIP-FEM for Helmholtz equation with high wave number and perfectly matched layer truncation. SIAM J. Numer. Anal. **57**(1), 96–126 (2019)

34. Melenk, J.M., Sauter, S.: Convergence analysis for finite element discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions. Math. Comp **79**(272), 1871–1914 (2010)
35. Melenk, J.M.: On generalized finite element methods. PhD thesis, The University of Maryland (1995)
36. Melenk, J.M., Sauter, S.: Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation. SIAM J. Numer. Anal. **49**, 1210–1243 (2011)
37. Melenk, J.M., Parsania, A., Sauter, S.: General DG-methods for highly indefinite Helmholtz problems. J. Sci. Comput. **57**(3), 536–581 (2013)
38. Ma, C., Alber, C., Scheichl, R.: Wavenumber explicit convergence of a multiscale GFEM for heterogeneous Helmholtz problems. arXiv:2112.10544 (2021)
39. Melenk, J.M., Sauter, S.A.: Wavenumber-explicit *hp*-FEM analysis for Maxwell's equations with transparent boundary conditions. Found. Comp. Math. **21**(1), 125–241 (2021)
40. Melenk, J.M., Sauter, S.A.: Wavenumber-explicit *hp*-FEM analysis for Maxwell's equations with impedance boundary conditions. arXiv:2201.02602 (2022)
41. McLean, W.C.H.: Strongly Elliptic Systems and Boundary Integral Equations. CUP (2000)
42. Morawetz, C.S., Ludwig, D.: An inequality for the reduced wave operator and the justification of geometrical optics. Comm. Pure Appl. Math. **21**, 187–203 (1968)
43. Morawetz, C.S.: Decay for solutions of the exterior problem for the wave equation. Commun. Pure Appl. Math. **28**(2), 229–264 (1975)
44. Nicaise, S., Tomezyk, J.: Convergence analysis of a *hp*-finite element approximation of the time-harmonic Maxwell equations with impedance boundary conditions in domains with an analytic boundary. Numer. Methods Partial Differ. Equ. **36**(6), 1868–1903 (2020)
45. Ohlberger, M., Verfurth, B.: A new heterogeneous multiscale method for the Helmholtz equation with high contrast. Multiscale Model. Simul. **16**(1), 385–411 (2018)
46. Peterseim, D.: Eliminating the pollution effect in Helmholtz problems by local subscale correction. Math. Comput. **86**(305), 1005–1036 (2017)
47. Sayas, F.-J., Brown, T.S., Hassell, M.E.: Variational Techniques for Elliptic Partial Differential Equations: Theoretical Tools and Advanced Applications. CRC Press, Boca Raton (2019)
48. Sauter, S.A.: A refined finite element convergence theory for highly indefinite Helmholtz problems. Computing **78**(2), 101–115 (2006)
49. Schatz, A.H.: An observation concerning Ritz-Galerkin methods with indefinite bilinear forms. Math. Comput. **28**(128), 959–962 (1974)
50. Shannon, C.E.: Communication in the presence of noise. Proc. IRE **37**(1), 10–21 (1949)
51. Spence, E.A., Kamotski, I.V., Smyshlyaev, V.P.: Coercivity of combined boundary integral equations in high-frequency scattering. Comm. Pure Appl. Math **68**(9), 1587–1639 (2015)
52. Wu, H.: Pre-asymptotic error analysis of CIP-FEM and FEM for the Helmholtz equation with high wave number. Part I: linear version. IMA J. Numer. Anal. **34**(3), 1266–1288 (2014)
53. Whittaker, E.T.: On the functions which are represented by the expansions of the interpolation-theory. Proc. R. Soc. Edinb. **35**, 181–194 (1915)
54. Zhu, B., Wu, H.: Preasymptotic error analysis of the HDG method for Helmholtz equation with large wave number. J. Sci. Comput. **87**(2), 1–34 (2021)
55. Zhu, L., Wu, H.: Preasymptotic error analysis of CIP-FEM and FEM for Helmholtz equation with high wave number. Part II: *hp* version. SIAM J. Numer. Anal. **51**(3), 1828–1852 (2013)