ANALYSIS & PDE

Volume 10 No. 6 2017

msp

Analysis & PDE

msp.org/apde

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

The subscription price for 2017 is US $265/year for the electronic version, and $470/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 10, No. 6, 2017
dx.doi.org/10.2140/apde.2017.10.1285

LOCAL ENERGY DECAY AND SMOOTHING EFFECT FOR THE DAMPED SCHRÖDINGER EQUATION

MOEZ KHENISSI AND JULIEN ROYER

We prove the local energy decay and the global smoothing effect for the damped Schrödinger equation on $\mathbb{R}^d$. The self-adjoint part is a Laplacian associated to a long-range perturbation of the flat metric. The proofs are based on uniform resolvent estimates obtained by the dissipative Mourre method. All the results depend on the strength of the dissipation that we consider.

1. Introduction

Let $d \geqslant 3$. Our purpose in this paper is to study on $\mathbb{R}^d$ the local energy decay and the Kato smoothing effect for the damped Schrödinger equation

$$\begin{cases} -i\partial_t u + Pu - ia(x)\langle D\rangle^\alpha a(x)u = 0, \\ u(0) = u_0. \end{cases} \tag{1-1}$$

The operator P is a Laplacian in divergence form associated to a long-range perturbation of the usual flat metric (see (1-2) below). For the dissipative part we have denoted by $\langle\cdot\rangle$ the function $(1+|\cdot|^2)^{\frac{1}{2}}$ and by D the square root of the free Laplacian, so that $\langle D\rangle^\alpha$ stands for $(1-\Delta)^{\frac{\alpha}{2}}$. The parameter α belongs to $[0,2[$. The nonnegative-valued function a will be assumed to be of short range (see (1-3)), so that in terms of spacial decay, we have an absorption index $a(x)^2$ which decays at least like $\langle x\rangle^{-2-2\rho}$ for some $\rho > 0$.

It is known that the free Schrödinger equation ((1-1) with $P = -\Delta$ and $a = 0$) preserves the L^2-norm but satisfies the local energy decay: if u_0 is supported in the ball $B(R) = \{|x| \leqslant R\}$ of $\mathbb{R}^d$ for some $R > 0$ we have

$$\|e^{it\Delta}u_0\|_{L^2(B(R))} \leqslant C_R\langle t\rangle^{-\frac{d}{2}}\|u_0\|_{L^2(\mathbb{R}^d)}.$$

This means that the "mass" of the solution escapes at infinity. On the other hand, the Schrödinger equation has a regularizing effect:

$$\int_{\mathbb{R}} \|(1-\Delta)^{\frac{1}{4}}e^{it\Delta}u_0\|^2_{L^2(B(R))}\,dt \leqslant C_R\|u_0\|^2_{L^2(\mathbb{R}^d)}.$$

There are many papers dealing with these properties for more and more general Schrödinger equations. Concerning the local energy decay for a self-adjoint Schrödinger equation, we only refer to [Rauch 1978] for the Schrödinger operator with an exponentially decaying potential, to [Tsutsumi 1984] for the free Schrödinger equation on an exterior domain, and to [Bouclet 2011; Bony and Häfner 2012] for a Laplacian associated to a long-range perturbation of the flat metric. For all these papers, the local energy

MSC2010: 35B40, 35Q41, 35B65, 47A55, 47B44.
Keywords: local energy decay, smoothing effect, damped Schrödinger equation, resolvent estimates.

decays like $t^{-\frac{d}{2}}$ or like $t^{-\frac{d}{2}+\varepsilon}$ under a nontrapping assumption. There is also a huge literature for the closely related problem of the local energy decay for the wave equation (see [Lax et al. 1963; Ralston 1969; Morawetz et al. 1977; Burq 1998; Tataru 2013; Guillarmou et al. 2013]).

Concerning the smoothing effect we mention [Constantin and Saut 1988; Sjölin 1987] for the Laplacian on $\mathbb{R}^d$, [Ben-Artzi and Klainerman 1992] for the Schrödinger operator with a potential, [Burq et al. 2004] for the problem on an exterior domain and [Erdoğan et al. 2009] for the magnetic Schrödinger equation. We also refer to [Doi 1996; 2000; Burq 2004] for the necessity of the nontrapping condition.

In the dissipative context, the local energy decay for the damped Schrödinger equation in an exterior domain has been proved in [Aloui and Khenissi 2007]. In this context, the nontrapping condition can be replaced by the geometric control condition: there can be bounded classical trajectories but they have to go through the damping region (see [Rauch and Taylor 1974; Bardos et al. 1992] for the original geometric control condition, and [Aloui and Khenissi 2002] for the exterior geometric condition on an unbounded domain). Then the local energy decays like $t^{-\frac{d}{2}}$, as in the self-adjoint case under the nontrapping condition. A similar result has been obtained in [Aloui and Khenissi 2010] on an exterior domain with dissipation at the boundary, and in [Royer 2015] for the same problem on a wave guide (see also [D'Ancona and Racke 2012] for the undamped problem on a nonflat wave guide). In the latter case, the global energy decays exponentially and we have a smoothing effect in the unbounded directions. We also mention [Bortot and Cavalcanti 2014], where an exponential decay for the global energy is proved for the solution of the Schrödinger equation with a dissipation effective on a neighborhood of the infinity.

The dissipation by a potential ($\alpha = 0$ in our setting) is not strong enough to recover under the damping condition the same smoothing effect as under the nontrapping condition. However, it is known that this is the case for the so-called regularized Schrödinger equation ($\alpha = 1$). See [Aloui 2008a; 2008b] for the problem on a compact manifold and [Aloui et al. 2017] for the problem on an exterior domain. As in the self-adjoint case (see [Burq 2004]), we can recover a $H^{\frac{1}{2}-\varepsilon}$ smoothing effect if only a few classical trajectories fail to satisfy the assumption (see [Aloui et al. 2013]).

In these works, the problem is a compact perturbation of the free Schrödinger equation. Our purpose in this paper is to prove the local energy decay and the Kato smoothing effect for an asymptotically vanishing perturbation. In a similar context, the local energy decay has been studied for the dissipative wave equation in [Bouclet and Royer 2014].

We now describe more precisely the setting of our paper. We consider on $\mathbb{R}^d$ a long-range perturbation $G(x)$ of the identity: for some $\rho > 0$ there exist constants C_β for $\beta \in \mathbb{N}^d$ such that

$$\left|\partial^\beta (G(x) - I_d)\right| \leqslant C_\beta \langle x\rangle^{-\rho-|\beta|}. \tag{1-2}$$

Concerning the dissipative term, a is a smooth and nonnegative-valued function on $\mathbb{R}^d$. As already mentioned, it is of short range:

$$|\partial^\beta a(x)| \leqslant C_\beta \langle x\rangle^{-1-\rho-|\beta|}. \tag{1-3}$$

We will use the notation

$$B_\alpha = a(x)\langle D\rangle^\alpha a(x) \quad \text{and} \quad H = P - iB_\alpha. \tag{1-4}$$

We recall that $\alpha \in [0, 2[$ and we set

$$\tilde{\alpha} = \min(1, \alpha) \quad \text{and} \quad \kappa = \begin{cases} \frac{d}{2} & \text{if } d \text{ is even,} \\ \frac{d+1}{2} & \text{if } d \text{ is odd.} \end{cases} \tag{1-5}$$

Then $\tilde{\alpha} \in [0, 1]$ and $\kappa \geqslant 2$.

We will see that H is a maximal dissipative operator on L^2. In particular, for $u_0 \in \mathcal{D}(H) = H^2$ the problem (1-1) has a unique solution $t \mapsto e^{-itH} u_0$. The main purpose of this paper is to prove that this solution satisfies the local energy decay and the Kato smoothing effect as stated in the following two theorems:

Theorem 1.1 (local energy decay). *Let $\varepsilon > 0$. Let $\delta > \kappa + \frac{1}{2}$, $N \in \mathbb{N}$ and $\sigma \in [0, 2]$. Assume that*

(i) *there are no bounded geodesics* (*see the nontrapping condition* (1-8) *below*) *or*

(ii) *the bounded geodesics go through the damping region* (*see* (1-9)), $N\tilde{\alpha} + \sigma \geqslant 2$ *and* $\delta > N - \frac{1}{2}$.

Then there exists $C \geqslant 0$ such that for $u_0 \in H^{\sigma,\delta}$ and $t \geqslant 0$ we have

$$\| e^{-itH} u_0 \|_{L^{2,-\delta}} \leqslant C t^{-\frac{d}{2}+\varepsilon} \| u_0 \|_{H^{\sigma,\delta}}.$$

In this statement $L^{2,-\delta}$ denotes the weighted space $L^2(\langle x \rangle^{-2\delta}\, dx)$, while $\| u_0 \|_{H^{\sigma,\delta}}$ is the L^2-norm of $\langle x \rangle^{\delta} \langle D \rangle^{\sigma} u_0$.

We remark that we have to take $\sigma = 2$ in the second case if $\alpha = 0$. This means that we have a loss of two derivatives. If $\alpha > 0$ we can take $\sigma = 0$ (no loss of derivative) as long as we choose δ large enough (if $\alpha \geqslant 1$ then we can take $N = 2$, and in this case the condition $\delta > N - \frac{1}{2}$ is weaker than $\delta > \kappa + \frac{1}{2}$). Under the nontrapping condition we can always take $\sigma = 0$.

In this setting, we obtain a decay which is almost as good as in the free case. We recall that for such a P, this is the best decay known even in the particular case $a = 0$ (see [Bouclet 2011]).

Theorem 1.2 (global smoothing effect). *Assume that the damping condition* (1-9) *holds. Then there exists $C \geqslant 0$ such that for all $u_0 \in L^2$ we have*

$$\int_0^{+\infty} \| \langle x \rangle^{-1} \langle D \rangle^{\frac{\tilde{\alpha}}{2}} e^{-itH} u_0 \|_{L^2}^2 \, dt \leqslant C \| u_0 \|_{L^2}^2.$$

Moreover, under the nontrapping condition (1-8), *we can replace $\tilde{\alpha}$ by* 1.

The last statement says that, despite the non-self-adjointness of H, we recover the same gain of regularity as in the self-adjoint case under the nontrapping assumption. However, the main result is that if the damping is strong enough, we have the same result for a trapping metric under the usual geometric condition. For a weaker damping we cannot reach the optimal result, but we still have some gain of regularity. As for the local energy decay above and for the resolvent estimates below, we can consider a very strong damping ($\alpha > 1$), but this does not improve the results (even if we allow trapped trajectories, there still are trajectories going to infinity, and their contributions are not controlled by the damping).

The proofs of Theorems 1.1 and 1.2 are based on uniform resolvent estimates. According to Proposition 2.2 below, the operator H is maximal dissipative, so for all z in

$$\mathbb{C}_+ := \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$$

we can consider in $\mathcal{L}(L^2)$ (the space of bounded operators on L^2) the resolvent

$$R(z) = (H - z)^{-1}. \tag{1-6}$$

After a Fourier transform, the solution u of (1-1) can be written as the integral over frequencies $\operatorname{Re}(z)$ of this resolvent when $\operatorname{Im}(z)$ goes to 0 (see Section 6). Thus the problem will be reduced to proving uniform estimates for $R(z)$ and its derivatives for $\operatorname{Im}(z)$ small, and then to controlling the dependence of these estimates with respect to $\operatorname{Re}(z)$. Since the self-adjoint part P of H is a nonnegative operator, the estimates for $\operatorname{Re}(z) < 0$ are easy: for $n \in \mathbb{N}$ and $z \in \mathbb{C}_+$ with $\operatorname{Re}(z) \leqslant -c_0 < 0$ we have

$$\|R^{n+1}(z)\|_{\mathcal{L}(L^2)} \leqslant \frac{C}{|\operatorname{Re}(z)|^{n+1}}. \tag{1-7}$$

Thus we will focus on $z \in \mathbb{C}_+$ with $\operatorname{Re}(z) \geqslant -c_0$, where $0 < c_0 \ll 1$. As usual, the difficulties will arise for low frequencies ($\operatorname{Re}(z)$ close to 0) and high frequencies ($\operatorname{Re}(z) \gg 1$). We first state the uniform resolvent estimates for intermediate frequencies:

Theorem 1.3 (intermediate-frequency estimates). *Let K be a compact subset of $\mathbb{C} \setminus \{0\}$. Let $n \in \mathbb{N}$ and $\delta > n + \frac{1}{2}$. Then there exists $C \geqslant 0$ such that for all $z \in K \cap \mathbb{C}_+$ we have*

$$\|\langle x\rangle^{-\delta} R^{n+1}(z) \langle x\rangle^{-\delta}\|_{\mathcal{L}(L^2)} \leqslant C.$$

We remark that compared to the resolvent for the dissipative wave equation (see [Bouclet and Royer 2014]), the derivatives of the resolvent correspond to its powers:

$$R^{(n)}(z) = n!\, R^{n+1}(z).$$

This will significantly simplify the discussion.

It is known that even for the free Laplacian, the estimates of Theorem 1.3 fail to hold uniformly when z goes to 0 if n is too large. This explains the restriction in the rate of decay in Theorem 1.1. For low frequencies we prove the following result:

Theorem 1.4 (low-frequency estimates). *Let $\varepsilon > 0$. Let $n \in \mathbb{N}$ and let δ be such that*

$$\delta > \begin{cases} n + \frac{1}{2} & \text{if } 2n + 1 \geqslant d, \\ n + 1 & \text{if } 2n + 1 < d. \end{cases}$$

Then there exist $C \geqslant 0$ and a neighborhood $\mathcal{U}$ of 0 in $\mathbb{C}$ such that for all $z \in \mathcal{U} \cap \mathbb{C}_+$ we have

$$\|\langle x\rangle^{-\delta} R^{n+1}(z) \langle x\rangle^{-\delta}\|_{\mathcal{L}(L^2)} \leqslant C(1 + |z|^{\frac{d}{2} - \varepsilon - 1 - n}).$$

In the self-adjoint case we can improve the estimate for a single resolvent. More precisely we can replace the weight $\langle x\rangle^{-\delta}$ for $\delta > 1$ by $\langle x\rangle^{-1}$. See [Bouclet and Royer 2015]. This is particularly interesting for Theorem 1.2, which does not require estimates for the derivatives of the resolvent. This sharp resolvent estimate is also valid in our dissipative context:

Theorem 1.5 (sharp low-frequency estimate). *There exist $C \geqslant 0$ and a neighborhood $\mathcal{U}$ of 0 in $\mathbb{C}$ such that for all $z \in \mathcal{U} \cap \mathbb{C}_+$ we have*

$$\|\langle x\rangle^{-1} R(z)\langle x\rangle^{-1}\|_{\mathcal{L}(L^2)} \leqslant C.$$

The high-frequency properties of the problem are closely related to the corresponding classical problem. Here, the classical flow is the geodesic flow on $\mathbb{R}^{2d} \simeq T^*\mathbb{R}^d$ for the metric $G(x)^{-1}$. It is the Hamiltonian flow corresponding to the symbol

$$p(x,\xi) = \langle G(x)\xi, \xi\rangle.$$

We denote this flow by $\phi^t = (X(t), \Xi(t))$. Let

$$\Omega_b = \big\{w \in p^{-1}(\{1\}) : \sup_{t\in\mathbb{R}} |X(t,w)| < +\infty\big\}.$$

The assumptions used in the statements of Theorems 1.1 and 1.2 are the following. We say that the classical flow is nontrapping if there is no bounded geodesic:

$$\Omega_b = \varnothing. \tag{1-8}$$

We say that the damping condition on bounded geodesics (or geometric control condition) is satisfied if every bounded geodesic goes through the damping region $\{a(x) > 0\}$:

$$\forall w \in \Omega_b,\ \exists T \in \mathbb{R} \quad \text{such that} \quad a(X(T,w)) > 0. \tag{1-9}$$

Theorem 1.6 (high-frequency estimates). *Let $n \in \mathbb{N}$ and $\delta > n + \frac{1}{2}$.*

(i) *Assume that the nontrapping assumption* (1-8) *holds. Then there exists $C \geqslant 0$ such that for $z \in \mathbb{C}_+$ with* $\mathrm{Re}(z) \geqslant C$ *we have*

$$\|\langle x\rangle^{-\delta} R^{n+1}(z)\langle x\rangle^{-\delta}\|_{\mathcal{L}(L^2)} \leqslant C|z|^{-\frac{n+1}{2}}.$$

(ii) *Assume that the damping condition* (1-9) *holds. Then there exists $C \geqslant 0$ such that for $z \in \mathbb{C}_+$ with* $\mathrm{Re}(z) \geqslant C$ *we have*

$$\|\langle x\rangle^{-\delta} R^{n+1}(z)\langle x\rangle^{-\delta}\|_{\mathcal{L}(L^2)} \leqslant C|z|^{-\frac{(n+1)\tilde{\alpha}}{2}}$$

(*we recall that $\tilde{\alpha}$ was defined in* (1-5)).

To prove the uniform estimates of Theorems 1.3, 1.4 and 1.6 we use the commutators method of Mourre [1981] (see also [Amrein et al. 1996] for an overview of the subject). The method has been generalized to the dissipative setting in [Royer 2010], then in [Bouclet and Royer 2014] for the estimates of the derivatives of the resolvent and finally in [Royer 2016] for a dissipative perturbation in the sense of forms. Here the dissipative perturbation B_α is well defined as an operator on L^2 relatively bounded with respect to the self-adjoint part P. However, for $d \in \{3,4\}$ the rescaled version of the dissipative part which we are going to use for low frequencies will be uniformly bounded as an operator in $\mathcal{L}(H^1, H^{-1})$ but not in $\mathcal{L}(H^2, L^2)$, so we will have to see H as a dissipative perturbation of P in the sense of forms. See Remark 4.7.

Let us come back to the statement of Theorem 1.2. To prove this theorem we will use in particular the resolvent estimates of Theorem 1.6, which in turn rely on the damping assumption (1-9). These estimates and hence the smoothing effect we obtain are optimal (in the sense that they are as good as in the self-adjoint case with the nontrapping condition) when $\alpha \geqslant 1$. However, with a weaker dissipation ($\alpha < 1$) we can obtain (weaker) resolvent estimates and a (weaker) smoothing effect. Similarly, it is possible to prove high-frequency resolvent estimates weaker than those of Theorem 1.6 without the damping condition. We have already mentioned [Burq 2004] in the self-adjoint case and [Aloui et al. 2013] in the dissipative setting, where only a few hyperbolic classical trajectories deny the assumption $\big($in these cases the high-frequency resolvent estimates are of size $\ln|z|/\sqrt{|z|}$, which gives a gain of $\frac{1}{2}-\varepsilon$ derivative$\big)$. We do not prove resolvent estimates without damping condition in this paper, but we emphasize this fact with a more general version of Theorem 1.2 (for self-adjoint operators, we mention the result of [Thomann 2010], which gives a relation between the smoothing effect and the decay of the spectral projections).

Theorem 1.7. *Let $\gamma \in [0,2]$. Assume that there exists $C \geqslant 0$ such that for all $z \in \mathbb{C}_+$ we have*

$$\| \langle x \rangle^{-1} R(z) \langle x \rangle^{-1} \|_{\mathcal{L}(L^2)} \leqslant C \langle z \rangle^{-\frac{\gamma}{2}}. \tag{1-10}$$

Then for all $u_0 \in L^2$ we have

$$\int_0^{+\infty} \| \langle x \rangle^{-1} \langle D \rangle^{\frac{\gamma}{2}} e^{-itH} u_0 \|_{L^2}^2 \, dt \leqslant C \| u_0 \|_{L^2}^2.$$

It is classical in the self-adjoint setting to prove the smoothing effect from resolvent estimates by means of the theory of relatively smooth operators in the sense of Kato [1966] (see also [Reed and Simon 1978]). Other ideas have been used for dissipative operators (see [Aloui et al. 2013; 2017]). However, the theory of Kato can also be used in this context (see [Royer 2010; 2015]). We will follow this idea to prove Theorem 1.7 and hence Theorem 1.2.

This paper is organized as follows. In Section 2 we recall all the abstract properties we need concerning dissipative operators (including the statement of the Mourre method). In Section 3 we prove Theorem 1.3. In Section 4 we deal with low frequencies. We first prove Theorem 1.4 for a small perturbation of the free Laplacian in Section 4A and then in the general setting in Section 4B. Theorem 1.5 is proved in Section 4C. In Section 5 we prove Theorem 1.6 concerning the high-frequency resolvent estimates. Finally, we turn to the time-dependent problem: we prove Theorem 1.1 in Section 6 and Theorems 1.7 and 1.2 in Section 7.

2. Abstract properties for dissipative operators

In this section we recall some general properties about dissipative operators. In particular we give the version of the Mourre's method that we use in this paper.

Let $\mathcal{H}$ be a Hilbert space. An operator H with domain $\mathcal{D}(H)$ on $\mathcal{H}$ is said to be dissipative (respectively accretive) if

$$\forall \varphi \in \mathcal{D}(H), \quad \operatorname{Im}\langle H\varphi, \varphi \rangle_{\mathcal{H}} \leqslant 0 \quad (\text{respectively } \operatorname{Re}\langle H\varphi, \varphi \rangle_{\mathcal{H}} \geqslant 0).$$

Moreover H is said to be maximal dissipative (respectively maximal accretive) if it has no other dissipative (respectively accretive) extension than itself. Notice that H is (maximal) dissipative if and only if iH is (maximal) accretive. We recall that a dissipative operator H is maximal dissipative if and only if there exists $z \in \mathbb{C}_+$ such that the operator $(H-z)$ has a bounded inverse on $\mathcal{H}$. In this case any $z \in \mathbb{C}_+$ belongs to the resolvent set of H and

$$\|(H-z)^{-1}\|_{\mathcal{L}(\mathcal{H})} \leqslant \frac{1}{\operatorname{Im}(z)}. \tag{2-1}$$

According to the Hille–Yosida theorem this implies in particular that $-iH$ generates a contractions semigroup, and then for all $u_0 \in \mathcal{D}(H)$ the function $u : t \mapsto e^{-itH} u_0$ belongs to $C^0(\mathbb{R}_+, \mathcal{D}(H)) \cap C^1(\mathbb{R}_+, \mathcal{H})$ and is the unique solution for the problem

$$\begin{cases} -i\partial_t u + Hu = 0 & \forall t > 0, \\ u(0) = u_0. \end{cases}$$

Moreover we have

$$\forall t \geqslant 0, \quad \|u(t)\|_{\mathcal{H}} \leqslant \|u_0\|_{\mathcal{H}}.$$

Remark 2.1. Assume that H is both dissipative and accretive. Then it is maximal dissipative if and only if it is maximal accretive. Indeed both properties are equivalent to the fact that $(H-(-1+i))$ has a bounded inverse on $\mathcal{H}$. Moreover, for $z \in \mathbb{C}$ with $\operatorname{Im}(z) > 0$ or $\operatorname{Re}(z) < 0$ we have

$$\|(H-z)^{-1}\|_{\mathcal{L}(\mathcal{H})} \leqslant \frac{1}{\max(\operatorname{Im}(z), -\operatorname{Re}(z))}.$$

Proposition 2.2. *The operator H defined by* (1-4) *is maximal dissipative and maximal accretive on L^2.*

Proof. The operators P and B_α are self-adjoint and nonnegative on L^2, so $H = P - iB_\alpha$ is dissipative and accretive. Let $\varphi \in \mathcal{D}(P) = H^2$. By interpolation there exists $C \geqslant 0$ (which only depends on a and α) such that for any $\varepsilon > 0$

$$\|B_\alpha \varphi\|_{L^2} \leqslant C\|\varphi\|_{H^\alpha} \leqslant C\|\varphi\|_{H^2}^{\frac{\alpha}{2}} \|\varphi\|_{L^2}^{1-\frac{\alpha}{2}} \leqslant \tfrac{1}{2}\alpha\varepsilon C\|\varphi\|_{H^2} + C\big(1-\tfrac{1}{2}\alpha\big)\varepsilon^{-\frac{\alpha}{2-\alpha}}\|\varphi\|_{L^2}.$$

With $\varepsilon > 0$ small enough we obtain that the dissipative operator $-iB_\alpha$ is relatively bounded with respect to P with relative bound less than 1. According to [Royer 2010, Lemma 2.1], this proves that H is maximal dissipative in L^2. By Remark 2.1, H is also maximal accretive. □

According to Proposition 2.2, the estimate of Remark 2.1 holds for H in $\mathcal{L}(L^2)$. As already mentioned, the difficulties in Theorems 1.3, 1.4 and 1.6 come from the behavior of the resolvent $R(z)$ when the spectral parameter $z \in \mathbb{C}_+$ approaches the nonnegative real axis. For this we are going to use a dissipative version of the Mourre method, which we recall now.

Let q_0 be a quadratic form closed, densely defined, symmetric and bounded from below on $\mathcal{H}$. We set $\mathcal{K} = \mathcal{D}(q_0)$. Let q_Θ be another symmetric form on $\mathcal{H}$, nonnegative and q_0-bounded. Let $q = q_0 - iq_\Theta$ and let H be the corresponding maximal dissipative operator (see Proposition 2.2 in [Royer 2016]). We denote by $\widetilde{H} : \mathcal{K} \to \mathcal{K}^*$ the operator which satisfies $q(\varphi, \psi) = \langle \widetilde{H}\varphi, \psi \rangle_{\mathcal{K}^*,\mathcal{K}}$ for all $\varphi, \psi \in \mathcal{K}$. Similarly, we denote by $\widetilde{H}_0$ and Θ the operators in $\mathcal{L}(\mathcal{K}, \mathcal{K}^*)$ which correspond to the forms q_0 and q_Θ, respectively. By the Lax–Milgram theorem, the operator $(\widetilde{H} - z)$ has a bounded inverse in $\mathcal{L}(\mathcal{K}^*, \mathcal{K})$ for all $z \in \mathbb{C}_+$. Moreover for $\varphi \in \mathcal{H}$ we have $(H-z)^{-1}\varphi = (\widetilde{H} - z)^{-1}\varphi$.

Definition 2.3. Let A be a self-adjoint operator on $\mathcal{H}$ and $N \in \mathbb{N}^*$. We say that A is a conjugate operator (in the sense of forms) to H on the interval J, up to order N, and with bounds $\alpha_0 \in]0,1]$, $\beta \geqslant 0$ and $\Upsilon_N \geqslant 0$ if the following conditions are satisfied:

(i) The form domain $\mathcal{K}$ is left invariant by e^{-itA} for all $t \in \mathbb{R}$. We denote by $\mathcal{E}$ the domain of the generator of $e^{-itA}|_{\mathcal{K}}$.

(ii) The commutators $\Lambda^0 = [\widetilde{H}_0, iA]$ and $\Lambda_1 = [\widetilde{H}, iA]$, a priori defined as operators in $\mathcal{L}(\mathcal{E}, \mathcal{E}^*)$, extend to operators in $\mathcal{L}(\mathcal{K}, \mathcal{K}^*)$. Then for all $n \in [\![1, N]\!]$ the operator $[\Lambda_n, iA]$ defined (inductively) in $\mathcal{L}(\mathcal{E}, \mathcal{E}^*)$ extends to an operator in $\mathcal{L}(\mathcal{K}, \mathcal{K}^*)$, which we denote by Λ_{n+1}.

(iii) We have

$$\|\Lambda_1\| \leqslant \sqrt{\alpha_0}\,\Upsilon_N, \quad \|\Lambda_1 + \beta\Theta\|\,\|\Lambda^0\| \leqslant \alpha_0 \Upsilon_N, \quad \|[\Lambda_1, A]\| + \beta\|[\Theta, A]\| \leqslant \alpha_0 \Upsilon_N$$

and

$$\sum_{n=2}^{N+1} \|\Lambda_n\|_{\mathcal{L}(\mathcal{K},\mathcal{K}^*)} \leqslant \alpha_0 \Upsilon_N,$$

where all the norms are in $\mathcal{L}(\mathcal{K}, \mathcal{K}^*)$.

(iv) We have

$$\mathbb{1}_J(H_0)(\Lambda^0 + \beta\Theta)\mathbb{1}_J(H_0) \geqslant \alpha_0 \mathbb{1}_J(H_0). \tag{2-2}$$

Theorem 5.5 of [Royer 2016] in the particular case where all the inserted factors are equal to $\mathrm{Id}_{\mathcal{H}}$ gives the following abstract resolvent estimates:

Theorem 2.4. *Suppose the self-adjoint operator A is conjugate to the maximal dissipative operator H on J up to order $N \geqslant 2$ with bounds $(\alpha_0, \beta, \Upsilon_N)$. Let $n \in [\![1, N]\!]$. Let $I \subset \mathring{J}$ be a compact interval. Let $\delta > n - \frac{1}{2}$. Then there exists $c \geqslant 0$ which only depends on J, I, δ, β and Υ_N and such that for all $z \in \mathbb{C}_{I,+}$ we have*

$$\|\langle A\rangle^{-\delta}(H - z)^{-n}\langle A\rangle^{-\delta}\|_{\mathcal{L}(\mathcal{H})} \leqslant \frac{c}{\alpha_0^n}.$$

We finish this general section with the so-called quadratic estimates. The following result is a consequence of Proposition 4.4 in [Royer 2016]:

Proposition 2.5. *Let $T \in \mathcal{L}(\mathcal{K}, \mathcal{H})$ be such that $T^*T \leqslant q_\Theta$ in the sense of forms on $\mathcal{K}$. Let $Q \in \mathcal{L}(\mathcal{H}, \mathcal{K}^*)$. Then for all $z \in \mathbb{C}_+$ we have*

$$\|T(\widetilde{H} - z)^{-1} Q\|_{\mathcal{L}(\mathcal{H})} \leqslant \|Q^*(\widetilde{H} - z)^{-1} Q\|_{\mathcal{L}(\mathcal{H})}^{\frac{1}{2}}.$$

Applied with $Q = T^*$, this proposition gives the following particular case:

Corollary 2.6. *Let T be as in Proposition 2.5. Then for all $z \in \mathbb{C}_+$ we have*

$$\|T(\widetilde{H} - z)^{-1} T^*\|_{\mathcal{L}(\mathcal{H})} \leqslant 1.$$

We are going to use all these results with the forms $q_0 : \varphi \mapsto \langle P\varphi, \varphi\rangle$ and $q_\Theta : \varphi \mapsto \langle B_\alpha \varphi, \varphi\rangle$ defined on $\mathcal{K} = H^1(\mathbb{R}^d)$.

3. Intermediate-frequency estimates

In this section we prove Theorem 1.3. For this, we will apply Theorem 2.4 with the generator of dilations as the conjugate operator. Let

$$A = -\tfrac{1}{2} i (x \cdot \nabla + \nabla \cdot x) = -i\,(x \cdot \nabla) - \tfrac{1}{2} i d.$$

We recall in the following proposition the main properties of A that we are going to use in this paper:

Proposition 3.1. (i) *For $\theta \in \mathbb{R}$, $u \in \mathcal{S}$ and $x \in \mathbb{R}^d$ we have*

$$(e^{i\theta A} u)(x) = e^{\frac{d\theta}{2}} u(e^{\theta} x).$$

(ii) *For $j \in [\![1, d]\!]$ and $\gamma \in C^{\infty}(\mathbb{R}^d)$ we have on $\mathcal{S}$*

$$[\partial_j, iA] = \partial_j \quad \textit{and} \quad [\gamma, iA] = -(x \cdot \nabla)\gamma.$$

(iii) *For $p \in [1, +\infty]$, $\theta \in \mathbb{R}$ and $u \in \mathcal{S}$ we have*

$$\|e^{i\theta A} u\|_{L^p} = e^{\theta\left(\frac{d}{2} - \frac{d}{p}\right)} \|u\|_{L^p}.$$

Now we give a proof of Theorem 1.3:

Proof of Theorem 1.3. Let $E > 0$. We check that the generator of dilations A is a conjugate operator for H on a neighborhood J of E in the sense of Definition 2.3. The form domain of H is the Sobolev space $H^1(\mathbb{R}^d)$. According to Proposition 3.1, it is left invariant by the dilation e^{-itA} for any $t \in \mathbb{R}$. By pseudodifferential calculus we can see that the commutators $[P, iA]$, $[[P, iA], iA]$, $[B_\alpha, iA]$ and $[[B_\alpha, iA], iA]$ define operators in $\mathcal{L}(H^2, L^2)$, hence in $\mathcal{L}(L^2, H^{-2})$ by duality, and in $\mathcal{L}(H^1, H^{-1})$ by interpolation.[1] Finally, we use the usual trick for the main assumption. For $\sigma > 0$ we set $J_\sigma = [E-\sigma, E+\sigma]$. We have

$$\begin{aligned} \mathbb{1}_{J_\sigma}(P)[P, iA]\mathbb{1}_{J_\sigma}(P) &= \mathbb{1}_{J_\sigma}(P)\, 2P\, \mathbb{1}_{J_\sigma}(P) + W \mathbb{1}_{J_\sigma}(P) \\ &\geqslant 2(E-\sigma)\mathbb{1}_{J_\sigma}(P) + W \mathbb{1}_{J_\sigma}(P), \end{aligned}$$

where

$$W := \mathbb{1}_{J_\sigma}(P)\, \operatorname{div}\big((x \cdot \nabla) G(x)\big) \nabla$$

is a compact operator. Since $E > 0$ is not an eigenvalue of P (see [Koch and Tataru 2006]) the operator $\mathbb{1}_{J_\sigma}(P)$ goes strongly to 0 when σ goes to 0. Then for σ small enough we have

$$\mathbb{1}_{J_\sigma}(P)[P, iA]\mathbb{1}_{J_\sigma}(P) \geqslant E\, \mathbb{1}_{J_\sigma}(P).$$

Thus we can apply Theorem 2.4, which gives Theorem 1.3 for $\operatorname{Re}(z) \in J_\sigma$ and with weights $\langle A \rangle^{-\delta}$. By compactness of $K \subset \mathbb{C}^*$ and the easy estimate of Remark 2.1, we have a uniform estimate for all

[1] In fact we can also compute these commutators explicitly with Proposition 3.1, except for the commutators of $\langle D \rangle^\alpha$ with A: for this we can write $\langle D \rangle^\alpha = (1-\Delta)^2 \times (1-\Delta)^{-2} \langle -\Delta \rangle^{\frac{\alpha}{2}}$ and use the Helffer–Sjöstrand formula for the second factor (see [Dimassi and Sjöstrand 1999; Davies 1995]).

$z \in \mathbb{C}_+ \cap K$. It remains to replace $\langle A\rangle^{-\delta}$ by $\langle x\rangle^{-\delta}$. For this, we use the resolvent identity

$$R(z) = R(i) + (z-i)R(i)R(z) = R(i) + (z-i)R(z)R(i).$$

It gives in particular, for $\nu \geqslant 2$,

$$R^\nu(z) = R(i)\big(R^\nu(z) + 2(z-i)R^{\nu-1}(z) + (z-i)^2 R^\nu(z)\big)R(i).$$

With these equalities in hand, we can prove by induction on $m \in \mathbb{N}^*$ that $R^{n+1}(z)$ can be written as a sum of terms of the form $(z-i)^\beta R^{n+1+\beta}(i)$ with $\beta \in \mathbb{N}$ or

$$(z-i)^{2m-n-1+\nu} R^m(i) R^\nu(z) R^m(i),$$

where $\max(1, n+1-2m) \leqslant \nu \leqslant n+1$. For any $\beta \in \mathbb{N}$, we know $R^{n+1+\beta}(i)$ is uniformly bounded in $\mathcal{L}(L^2)$. On the other hand,

$$\big\|\langle x\rangle^{-\delta} R^m(i) R^\nu(z) R^m(i)\langle x\rangle^{-\delta}\big\| \leqslant \big\|\langle x\rangle^{-\delta} R^m(i)\langle A\rangle^{\delta}\big\| \big\|\langle A\rangle^{-\delta} R^\nu(z)\langle A\rangle^{-\delta}\big\| \big\|\langle A\rangle^{\delta} R^m(i)\langle x\rangle^{-\delta}\big\|.$$

The first and third factors are bounded by pseudodifferential calculus if m is large enough and the second has been estimated uniformly by the Mourre method. This concludes the proof of Theorem 1.3. ☐

4. Low-frequency estimates

In this section we prove Theorems 1.4 and 1.5. As in [Bouclet 2011; Bouclet and Royer 2014], the proof of Theorem 1.4 is based on a scaling argument for a small perturbation of the free Laplacian (see Section 4A), and then on a perturbation argument to deal with the general case (see Section 4B). Theorem 1.5 is proved in Section 4C.

Let $\chi \in C_0^\infty(\mathbb{R}^d)$ be equal to 1 on a neighborhood of 0. For $\eta \in \,]0,1]$ we set $\chi_\eta : x \mapsto \chi(\eta x)$. Then for $\eta_1 \in \,]0,1]$ we set $G_{\eta_1}(x) = \chi_{\eta_1}(x) I_d + (1-\chi_{\eta_1}(x))G(x)$,

$$P_{\eta_1} = -\operatorname{div} G_{\eta_1}(x)\nabla \quad \text{and} \quad P_{\eta_1,c} = P - P_{\eta_1} = -\operatorname{div}\big(\chi_{\eta_1}(x)(G(x)-I_d)\big)\nabla. \tag{4-1}$$

For the dissipative part we set

$$B^\alpha_{\eta_2} = a(1-\chi_{\eta_2})\langle D\rangle^\alpha a + a\chi_{\eta_2}\langle D\rangle^\alpha a(1-\chi_{\eta_2}) \tag{4-2}$$

and

$$B^\alpha_{\eta_2,c} = B_\alpha - B^\alpha_{\eta_2} = a\chi_{\eta_2}\langle D\rangle^\alpha a\chi_{\eta_2},$$

where $\eta_2 \in \,]0,1]$. Finally, for the full operator we define

$$H_{\bar\eta} = P_{\eta_1} - iB^\alpha_{\eta_2} \quad \text{and} \quad R_{\bar\eta}(z) = (H_{\bar\eta} - z)^{-1},$$

where $\bar\eta = (\eta_1, \eta_2) \in \,]0,1]^2$ and $z \in \mathbb{C}_+$. The fact that we can choose $\eta_1 \neq \eta_2$ will be important in the sequel (see Remark 4.11).

4A. ***Low-frequency estimates for a small perturbation of the Laplacian.*** In this paragraph we prove Theorem 1.4 with $R(z)$ replaced by $R_{\tilde{\eta}}(z)$. Then in Section 4B we will add the contributions of $P_{\eta_1,c}$ and $B^{\alpha}_{\eta_2,c}$.

The proof relies on a scaling argument. To this purpose we use for $z \in \mathbb{C}^*$ the operator

$$\Theta_z = \exp\big(\tfrac{1}{2} i \ln|z| A\big).$$

For $u \in \mathcal{S}$ and $x \in \mathbb{R}^d$ we have $(\Theta_z u)(x) = |z|^{\frac{d}{4}} u(|z|^{\frac{1}{2}} x)$. According to Proposition 3.1 we have for $p \in [1, +\infty]$

$$\|\Theta_z\|_{\mathcal{L}(L^p)} = |z|^{\frac{d}{4} - \frac{d}{2p}}. \tag{4-3}$$

For a function u on $\mathbb{R}^d$ and $z \in \mathbb{C}^*$ we denote by u_z the function

$$u_z : x \mapsto u\left(\frac{x}{\sqrt{|z|}}\right).$$

Compared to the scaling for the wave equation we are using the parameter $\sqrt{|z|}$ instead of $|z|$.

Now we introduce the rescaled versions of our operators:

$$H_{\tilde{\eta},z} = \frac{1}{|z|} \Theta_z^{-1} H_{\tilde{\eta}} \Theta_z = P_{\eta_1,z} - i B^{\alpha}_{\eta_2,z},$$

where $P_{\eta_1,z} = -\operatorname{div} G_{\eta_1,z}(x) \nabla$ and

$$B^{\alpha}_{\eta_2,z} = \frac{1}{|z|}((1-\chi_{\eta_2})a)_z (1 - |z|\Delta)^{\frac{\alpha}{2}} a_z + \frac{1}{|z|}(\chi_{\eta_2} a)_z (1 - |z|\Delta)^{\frac{\alpha}{2}} ((1-\chi_{\eta_2})a)_z.$$

Then for $\zeta \in \mathbb{C}_+$ we set $R_{\tilde{\eta},z}(\zeta) = (H_{\tilde{\eta},z} - \zeta)^{-1}$, so that with the notation $\hat{z} = z/|z|$ we have for $z \in \mathbb{C}_+$

$$R_{\tilde{\eta}}(z) = \frac{1}{|z|} \Theta_z R_{\tilde{\eta},z}(\hat{z}) \Theta_z^{-1}.$$

Our analysis of the rescaled operators is based on the fact that if a function ϕ decays like $\langle x\rangle^{-\nu-\frac{\rho}{2}}$ (recall that $\rho > 0$ is fixed by (1-2) and (1-3)) then the multiplication by the rescaled function ϕ_λ behaves like a differential operator of order ν for low frequencies, in the sense that it is of size λ^ν as an operator from H^s to $H^{s-\nu}$. Since this observation relies on the Sobolev embeddings, there is however a restriction on the choices of ν and s. For $\sigma \in \mathbb{R}$, let $\mathcal{S}^{-\sigma}(\mathbb{R}^d)$ be the set of functions $\phi \in C^\infty(\mathbb{R}^d)$ such that

$$|\partial^\beta \phi(x)| \lesssim \langle x\rangle^{-\sigma-|\beta|}.$$

For $\nu \geqslant 0$, $N \in \mathbb{N}$ and $\phi \in \mathcal{S}^{-\nu-\frac{\rho}{2}}(\mathbb{R}^d)$ we set

$$\|\phi\|_{\nu,N} = \sup_{|\beta| \leqslant \kappa+1} \sum_{0 \leqslant m \leqslant N} \sup_{x \in \mathbb{R}^d} \left| \langle x\rangle^{\nu+\frac{\rho}{2}+|\beta|} (\partial^\beta (x\cdot\nabla)^m \phi)(x) \right|.$$

We recall that the integer κ was defined in (1-5). The following result is Proposition 7.2 in [Bouclet and Royer 2014]:

Proposition 4.1. *Let $\nu \in \left[0, \frac{d}{2}\right[$ and $s \in \left]-\frac{d}{2}, \frac{d}{2}\right[$ be such that $s-\nu \in \left]-\frac{d}{2}, \frac{d}{2}\right[$. Then there exists $C \geqslant 0$ such that for $\phi \in \mathcal{S}^{-\nu-\frac{\rho}{2}}(\mathbb{R}^d)$, $u \in H^s$ and $\lambda > 0$ we have*

$$\|\phi_\lambda u\|_{\dot{H}^{s-\nu}} \leqslant C\lambda^\nu \|\phi\|_{\nu,0} \|u\|_{\dot{H}^s},$$
$$\|\phi_\lambda u\|_{H^{s-\nu}} \leqslant C\lambda^\nu \|\phi\|_{\nu,0} \|u\|_{H^s}.$$

The reason for replacing $G(x)$ by $G_{\eta_1}(x)$ and a by $a(1-\chi_{\eta_2})$ in the definition of $H_{\bar\eta}$ is that for all $N \in \mathbb{N}$ we have

$$\mathcal{N}_{\bar\eta,N} := \sum_{j,k=1}^{d} \|G_{\eta_1,j,k}(x) - \delta_{j,k}\|_{0,N} + \|(1-\chi_{\eta_2})a\|_{1,N}\big(\|a\|_{1,N} + \|\chi_{\eta_2}a\|_{1,N}\big) = \underset{\bar\eta\to 0}{O}(|\bar\eta|^{\rho/2}). \tag{4-4}$$

Thus this quantity is as small as we wish if we choose η_1 and η_2 small enough.

Given two operators T and S we set $\mathrm{ad}_T^0(S) = S$, $\mathrm{ad}_T(S) = \mathrm{ad}_T^1(S) = [S, T]$ and then, for $m \geqslant 2$, $\mathrm{ad}_T^m(S) = [\mathrm{ad}_T^{m-1}(S), S]$. For $\mu = (\mu_1, \dots, \mu_d) \in \mathbb{N}^d$ we set

$$\mathrm{ad}_x^\mu := \mathrm{ad}_{x_1}^{\mu_1} \cdots \mathrm{ad}_{x_d}^{\mu_d}.$$

At the beginning of the section we said that $H_{\bar\eta}$ has to be close to the free Laplacian. What we need precisely is the following result:

Proposition 4.2. *Let $\mu \in \mathbb{N}^d$, $m \in \mathbb{N}$, $\varepsilon_0 > 0$ and $s \in \mathbb{R}$. There exists $\eta_0 \in]0,1]$ such that for $\bar\eta = (\eta_1, \eta_2) \in]0, \eta_0]^2$ the following statements hold:*

(i) *If $s \in \left]-\frac{d}{2}, \frac{d}{2}\right[$ then for $z \in \mathbb{C}_+$ with $|z| \leqslant 1$ we have*

$$\|\mathrm{ad}_x^\mu \mathrm{ad}_A^m (P_{\eta_1,z}+)\|_{\mathcal{L}(H^{s+1},H^{s-1})} \leqslant \varepsilon_0.$$

(ii) *If $s \in \left]-\frac{d}{2}+1, \frac{d}{2}-1\right[$ then we also have*

$$\|\mathrm{ad}_x^\mu \mathrm{ad}_A^m B_{\eta_2,z}^\alpha\|_{\mathcal{L}(H^{s+1},H^{s-1})} \leqslant \varepsilon_0.$$

(iii) *For $u \in H^2$ we have*

$$\tfrac{1}{2}\|u\|_{\dot{H}^2} \leqslant \|P_{\eta_1}u\|_{L^2} \leqslant 2\|u\|_{\dot{H}^2}.$$

Proof. The first statement is the same as for the wave equation. See Proposition 7.6 in [Bouclet and Royer 2014]. In particular with $s = 1$, $|z| = 1$ and $\varepsilon_0 = \frac{1}{2}$ we obtain the last statement. It remains to prove (ii). Let $D_z = \sqrt{|z|}D$. We write

$$((1-\chi_{\eta_2})a)_z \langle D_z\rangle^\alpha a_z = ((1-\chi_{\eta_2})a)_z(-|z|\Delta + 1)\langle D_z\rangle^{\alpha-2} a_z.$$

Then $\mathrm{ad}_x^\mu \mathrm{ad}_A^m\big(((1-\chi_{\eta_2})a)_z \langle D_z\rangle^\alpha a_z\big)$ can be written as a sum of terms of the form

$$\mathrm{ad}_x^{\mu_1}\mathrm{ad}_A^{m_1}\big(((1-\chi_{\eta_2})a)_z\big)\,\mathrm{ad}_x^{\mu_2}\mathrm{ad}_A^{m_2}(-|z|\Delta+1)\,\mathrm{ad}_x^{\mu_3}\mathrm{ad}_A^{m_3}(\langle D_z\rangle^{\alpha-2})\,\mathrm{ad}_x^{\mu_4}\mathrm{ad}_A^{m_4}(a_z),$$

where $\mu_1, \mu_2, \mu_3, \mu_4 \in \mathbb{N}^d$ and $m_1, m_2, m_3, m_4 \in [\![0, m]\!]$ are such that $\mu_1 + \mu_2 + \mu_3 + \mu_4 = \mu$ and $m_1 + m_2 + m_3 + m_4 = m$. Let $\gamma \in [0,1]$. According to Proposition 4.1 we have for $z \in \mathbb{C}_+$

$$\|\mathrm{ad}_x^{\mu_1}\mathrm{ad}_A^{m_1}((1-\chi_{\eta_2})a)_z\|_{\mathcal{L}(H^{s-1+\gamma},H^{s-1})} \lesssim \eta_2^{1+\frac{\rho}{2}-\gamma}|z|^{\frac{\gamma}{2}} \tag{4-5}$$

and

$$\|\mathrm{ad}_x^{\mu_4}\mathrm{ad}_A^{m_4} a_z\|_{\mathcal{L}(H^{s+1},H^{s+1-\gamma})} \lesssim |z|^{\frac{\gamma}{2}}. \tag{4-6}$$

To estimate $\mathrm{ad}_x^{\mu_3}\mathrm{ad}_A^{m_3}(\langle D_z\rangle^{\alpha-2})$, we use the Helffer–Sjöstrand formula (see [Dimassi and Sjöstrand 1999; Davies 1995]). We can check that for $\zeta \in \mathbb{C}\setminus\mathbb{R}$ we have

$$\left\|\mathrm{ad}_x^{\mu_3}\mathrm{ad}_A^{m_3}(-|z|\Delta-\zeta)^{-1}\right\|_{\mathcal{L}(H^{s+1-\gamma})} \lesssim \frac{\langle\zeta\rangle^{|\mu_3|+m_3}}{|\mathrm{Im}(\zeta)|^{|\mu_3|+m_3+1}}.$$

Let $f:\tau\mapsto(\tau+1)^{\frac{\alpha-2}{2}}$. Let $\phi\in C_0^\infty(\mathbb{R},[0,1])$ be supported in $[-2,2]$ and equal to 1 on $[-1,1]$. For $M>|\mu_3|+m_3+1$ and $\zeta=x+iy$ we set

$$\tilde f_M(\zeta)=\phi\left(\frac{y}{\langle x\rangle}\right)\sum_{k=0}^{M} f^{(k)}(x)\frac{(iy)^k}{k!}.$$

We have

$$\left|\frac{\partial\tilde f_M}{\partial\bar\zeta}(\zeta)\right| \leqslant \mathbb{1}_{\{\langle x\rangle\leqslant|y|\leqslant 2\langle x\rangle\}}(\zeta)\langle x\rangle^{-1+\frac{\alpha-2}{2}}+\mathbb{1}_{\{|y|\leqslant 2\langle x\rangle\}}(\zeta)|y|^M\langle x\rangle^{-M-1+\frac{\alpha-2}{2}},$$

so we can write

$$(-|z|\Delta+1)^{\frac{\alpha-2}{2}}=\frac{1}{\pi}\int_{\zeta=x+iy\in\mathbb{C}}\frac{\partial\tilde f_M}{\partial\bar z}(\zeta)(-|z|\Delta-\zeta)^{-1}\,dx\,dy.$$

Then we can check that

$$\|\mathrm{ad}_x^{\mu_3}\mathrm{ad}_A^{m_3}\langle D_z\rangle^{\alpha-2}\|_{\mathcal{L}(H^{s+1-\gamma})}\lesssim 1. \tag{4-7}$$

It remains to estimate

$$\mathrm{ad}_x^{\mu_2}\mathrm{ad}_A^{m_2}(\langle D_z\rangle^2)=-|z|\mathrm{ad}_x^{\mu_2}\mathrm{ad}_A^{m_2}\Delta+\mathrm{ad}_x^{\mu_2}\mathrm{ad}_A^{m_2}(1). \tag{4-8}$$

We have $\||z|\mathrm{ad}_x^{\mu_2}\mathrm{ad}_A^{m_2}\Delta\|\lesssim|z|$ in $\mathcal{L}(H^{s+1},H^{s-1})$ so with (4-5), (4-6) and (4-7) applied with $\gamma=0$ we obtain in $\mathcal{L}(H^{s+1},H^{s-1})$

$$\left\|\mathrm{ad}_x^{\mu_1}\mathrm{ad}_A^{m_1}\big(((1-\chi_{\eta_2})a)_z\big)\,\mathrm{ad}_x^{\mu_2}\mathrm{ad}_A^{m_2}(-|z|\Delta)\,\mathrm{ad}_x^{\mu_3}\mathrm{ad}_A^{m_3}(\langle D_z\rangle^{\alpha-2})\,\mathrm{ad}_x^{\mu_4}\mathrm{ad}_A^{m_4}(a_z)\right\|\lesssim|z|\eta_2^{1+\frac{\rho}{2}}. \tag{4-9}$$

If $|\mu_2|=m_2=0$ we also have to consider the second term in (4-8). For this we apply (4-5), (4-6) and (4-7) with $\gamma=1$, which gives

$$\left\|\mathrm{ad}_x^{\mu_1}\mathrm{ad}_A^{m_1}\big(((1-\chi_{\eta_2})a)_z\big)\,\mathrm{ad}_x^{\mu_3}\mathrm{ad}_A^{m_3}(\langle D_z\rangle^{\alpha-2})\,\mathrm{ad}_x^{\mu_4}\mathrm{ad}_A^{m_4}(a_z)\right\|_{\mathcal{L}(H^{s+1},H^{s-1})}\lesssim|z|\eta_2^{\frac{\rho}{2}}.$$

Thus we have proved that $\mathrm{ad}_x^{\mu}\mathrm{ad}_A^{m}\big(((1-\chi_{\eta_2})a)_z\langle D_z\rangle^\alpha a_z\big)$ is of size $O(|z|\eta_2^{\frac{\rho}{2}})$ in $\mathcal{L}(H^{s+1},H^{s-1})$. We proceed similarly for $\mathrm{ad}_x^{\mu}\mathrm{ad}_A^{m}\big((\chi_{\eta_2}a)_z\langle D_z\rangle^\alpha((1-\chi_{\eta_2})a)_z\big)$, and the statement follows. ☐

Remark 4.3. If $d\geqslant 5$ we can replace P_{η_1} by $H_{\tilde\eta}$ in the last statement of Proposition 4.2. This is not the case for $d\in\{3,4\}$. This is due to the fact that $s=1$ does not belong to $\left]-\frac{d}{2}+1,\frac{d}{2}-1\right[$ and hence $B^\alpha_{\eta_2,z}$ is not small in $\mathcal{L}(\dot H^2,L^2)$ in these cases.

Proposition 4.4. *Let $\mu \in \mathbb{N}^d$, $m \in \mathbb{N}$ and $s \in \left]-\frac{d}{2}+1, \frac{d}{2}-1\right[$. There exists $\eta_0 \in]0,1]$ such that the operator*

$$\mathrm{ad}_x^{\mu} \mathrm{ad}_A^{m} R_{\bar\eta,z}(-1)$$

is bounded as an operator from H^{s-1} to H^{s+1} uniformly in $z \in \mathbb{C}_+$ with $|z| \leqslant 1$ and $\bar\eta = (\eta_1, \eta_2) \in]0, \eta_0]^2$.

Proof. The idea of the proof is the same as the proof of Proposition 7.9 in [Bouclet and Royer 2014]. We only have to be careful with the fact that the dissipative term has to be seen as an operator of order 2. However, with the smallness assumption on $a(1-\chi_{\eta_2})$, it is still a small perturbation of $-\Delta$, and we can proceed as for the wave resolvent. We also have to be careful with the restriction on s, which is stronger than for the wave equation. This is due to the analogous restriction in the second statement of Proposition 4.2. We omit the details. □

Proposition 4.5. (i) *Let $s \in \left[0, \frac{d}{2}\right[$, $\delta > s$ and $m \in \mathbb{N}$ be such that $m \geqslant s$. Then there exist $\eta_0 \in]0,1]$ and $C \geqslant 0$ such that for $z \in \mathbb{C}_+$ with $|z| \leqslant 1$ and $\bar\eta \in]0, \eta_0]^2$ we have*

$$\| \langle x\rangle^{-\delta} \Theta_z R^m_{\bar\eta,z}(-1) \Theta_z^{-1} \langle x\rangle^{-\delta} \|_{\mathcal{L}(L^2)} \leqslant C|z|^s.$$

(ii) *Let $s \in \left[0, \frac{d}{2}\right[$, $\delta > s$, and $m \in \mathbb{N}$ large enough $\left(\text{say } m \geqslant \delta + \frac{s}{2} + 1\right)$. Then there exist $\eta_0 \in]0,1]$ and $C \geqslant 0$ such that for $z \in \mathbb{C}_+$ with $|z| \leqslant 1$ and $\bar\eta \in]0, \eta_0]^2$ we have*

$$\| \langle x\rangle^{-\delta} \Theta_z R^m_{\bar\eta,z}(-1) \langle A\rangle^{\delta} \|_{\mathcal{L}(L^2)} \leqslant C|z|^{\frac{s}{2}},$$
$$\| \langle A\rangle^{\delta} R^m_{\bar\eta,z}(-1) \Theta_z^{-1} \langle x\rangle^{-\delta} \|_{\mathcal{L}(L^2)} \leqslant C|z|^{\frac{s}{2}}.$$

Proof. According to Proposition 4.4 the operator $R^m_{\bar\eta,z}(-1)$ is bounded in $\mathcal{L}(H^{-s}, H^s)$ uniformly in $z \in \mathbb{C}_+$ with $|z| \leqslant 1$ and $\bar\eta$ close to (0,0). On the other hand, according to the Sobolev embedding $H^s \subset L^p$ for $p = \frac{2d}{d-2s}$, the fact that $\langle x\rangle^{-\delta}$ belongs to $\mathcal{L}(L^p, L^2)$ and (4-3) we have

$$\| \langle x\rangle^{-\delta} \Theta_z \|_{\mathcal{L}(H^s, L^2)} \lesssim \|\Theta_z\|_{\mathcal{L}(L^p)} \lesssim |z|^{\frac{s}{2}}.$$

We similarly have

$$\| \Theta_z^{-1} \langle x\rangle^{-\delta} \|_{\mathcal{L}(L^2, H^{-s})} \lesssim |z|^{\frac{s}{2}},$$

and the first statement follows. For the second statement we use the same idea as in the proof of Proposition 7.11 in [Bouclet and Royer 2014]. We only prove the first estimate. For this we first remark that

$$\begin{aligned}\| \langle x\rangle^{-\delta} \Theta_z (1+|x|^{\delta}) \|_{\mathcal{L}(H^s, L^2)} &\leqslant \| \langle x\rangle^{-\delta} \Theta_z \|_{\mathcal{L}(H^s, L^2)} + \left\| \langle x\rangle^{-\delta} \Theta_z |x|^{\delta} \right\|_{\mathcal{L}(L^2)} \\ &\lesssim \|\Theta_z\|_{\mathcal{L}(L^p)} + |z|^{\frac{\delta}{2}} \left\| \langle x\rangle^{-\delta} |x|^{\delta} \Theta_z \right\|_{\mathcal{L}(L^2)} \lesssim |z|^{\frac{s}{2}},\end{aligned}$$

where, again, p stands for $\frac{2d}{d-2s}$. Then it remains to prove that for all $\delta \geqslant 0$ (we no longer need the assumption that $\delta > s$), $m \geqslant \delta + \frac{s}{2} + 1$ and $\mu \in \mathbb{N}^d$ the operator

$$\langle x\rangle^{-\delta} \mathrm{ad}_x^{\mu} \left(R^m_{\bar\eta,z}(-1) \langle A\rangle^{\delta} \right)$$

is bounded in $\mathcal{L}(L^2, H^s)$ uniformly in $z \in \mathbb{C}_+$. With $\mu = 0$ this will conclude the proof. By interpolation it is enough to consider the case where δ is an integer and $m \geqslant \delta + \frac{s}{2}$ (we do not mean to be sharp with this assumption). We proceed by induction. The statement for $\delta = 0$ is given by Proposition 4.4. Now let $\delta \in \mathbb{N}^*$. We have

$$R^{m}_{\bar\eta,z}(-1)A^{\delta} = \sum_{k=0}^{\delta} C^k_\delta \, R^{m-1}_{\bar\eta,z}(-1)A^{\delta-k}\operatorname{ad}^k_A(R_{\bar\eta,z}(-1)).$$

When $k \neq 0$ we can apply the inductive assumption to $R^{m-1}_{\bar\eta,z}(-1)A^{\delta-k}$. With Proposition 4.4 we obtain that the contributions of the corresponding terms are uniformly bounded in $\mathcal{L}(L^2, H^s)$ as expected. It remains to consider the term corresponding to $k = 0$. It is enough to consider

$$R^{m-1}_{\bar\eta,z}(-1)A^{\delta-1}x_j D_j R_{\bar\eta,z}(-1)$$

for some $j \in [\![1, d]\!]$. The operator $D_j R_{\bar\eta,z}(-1)$ and its commutators with powers of x are uniformly bounded operators on L^2, and

$$R^{m-1}_{\bar\eta,z}(-1)A^{\delta-1}x_j = x_j R^{m-1}_{\bar\eta,z}(-1)A^{\delta-1} + \operatorname{ad}_{x_j}\big(R^{m-1}_{\bar\eta,z}(-1)A^{\delta-1}\big).$$

We conclude with the inductive assumption. □

Proposition 4.6. *Let $k \in \mathbb{N}$ and $\delta > k + \frac{1}{2}$. Then there exist $\eta_0 \in]0, 1]$ and $C \geqslant 0$ such that for $z \in \mathbb{C}_+$ with $|z| \leqslant 1$ and $\bar\eta = (\eta_1, \eta_2) \in]0, \eta_0]^2$ we have*

$$\|\langle A\rangle^{-\delta} R^{k+1}_{\bar\eta,z}(\hat z)\langle A\rangle^{-\delta}\|_{\mathcal{L}(L^2)} \leqslant C.$$

Proof. The estimate is clear when $\hat z$ is outside some neighborhood of 1. For $\hat z$ close to 1 we apply Theorem 2.4 uniformly in z with A as a conjugate operator. We have already said that e^{-itA} leaves H^1 invariant for all $t \in \mathbb{R}$. The assumptions (ii) and (iii) of Definition 2.3 with $\alpha_0 = \frac{1}{2}$ and $\beta = 0$ are consequences of Proposition 4.2 applied with $s = 0$, $\mu = 0$ and $m \in \mathbb{N}^*$. For $m \in \{0, 1\}$, $z \in \mathbb{C}_+$ and $u \in \mathcal{S}$ we have

$$\begin{aligned}\big|\langle \operatorname{ad}^m_{iA}(P_{\eta_1,z})u, u\rangle_{L^2} - 2^m\langle -\Delta u, u\rangle_{L^2}\big| &\leqslant \sum_{j,k=1}^{d}\big|\langle (2 - x\cdot\nabla)^m (G_{\eta_1,z,j,k} - \delta_{j,k})D_j u, D_k u\rangle_{L^2}\big| \\ &\lesssim O(\eta_1^{\rho/2})\|\nabla u\|^2_{L^2},\end{aligned}$$

and hence

$$[P_{\eta_1,z}, iA] \geqslant (2 - O(\eta_1^{\rho/2}))(-\Delta) \geqslant (2 - O(\eta_1^{\rho/2}))P_{\eta_1,z}.$$

Let $J =]\frac{1}{2}, \frac{3}{2}[$. After conjugation by $\mathbb{1}_J(P_{\eta_1,z})$ we obtain that if η_0 is small enough then for all $\eta_1 \in]0, \eta_0]$ and $z \in \mathbb{C}_+$ we have

$$\mathbb{1}_J(P_{\eta_1,z})[P_{\eta_1,z}, iA]\mathbb{1}_J(P_{\eta_1,z}) \geqslant \tfrac{1}{2}\mathbb{1}_J(P_{\eta_1,z}).$$

Then Proposition 4.6 follows from Theorem 2.4. □

Remark 4.7. It is important to notice that we have estimated $[B^\alpha_{\eta_2}, iA]$ and $[[B^\alpha_{\eta_2}, iA], iA]$ in $\mathcal{L}(H^1, H^{-1})$ and not in $\mathcal{L}(H^2, L^2)$. By pseudodifferential calculus, these two commutators define operators in

$\mathcal{L}(H^2, L^2)$. But in low dimensions ($d \in \{3,4\}$) they can be estimated uniformly by Proposition 4.2 only in the sense of forms. This is why we need a form version of the dissipative Mourre method here.

Proposition 4.8. *Let $\varepsilon > 0$ and $n \in \mathbb{N}$. Let δ be as in the statement of Theorem 1.4. Then there exist $\eta_0 \in]0,1]$, $C \geqslant 0$ and a neighborhood $\mathcal{U}$ of 0 in $\mathbb{C}$ such that for $\bar{\eta} = (\eta_1, \eta_2) \in]0, \eta_0]^2$ and $\beta_l, \beta_r \in \mathbb{R}_+$ with $\beta_l + \beta_r \leqslant 2$ we have*

$$\left\| \langle x \rangle^{-\delta} \langle D \rangle^{\beta_l} R_{\bar{\eta}}^{n+1}(z) \langle D \rangle^{\beta_r} \langle x \rangle^{-\delta} \right\|_{\mathcal{L}(L^2)} \leqslant C\left(1 + |z|^{\frac{d}{2}-\varepsilon-1-n}\right).$$

Remark 4.9. Compared to the analogous result for the wave equation (see Theorem 1.3 in [Bouclet and Royer 2014]) there is no gain when we add a derivative. This is a consequence of the restriction on the Sobolev index s in Proposition 4.4, which is stronger than in Proposition 7.9 in [Bouclet and Royer 2014].

Proof of Proposition 4.8. First assume that $n \geqslant 1$. By the resolvent identity we have

$$\begin{aligned} \langle x \rangle^{-\delta} \langle D \rangle^{\beta_l} R_{\bar{\eta}}^{n+1}(z) \langle D \rangle^{\beta_r} \langle x \rangle^{-\delta} = {} & \langle x \rangle^{-\delta} \langle D \rangle^{\beta_l} R_{\bar{\eta}}(-1) \langle x \rangle^{\delta} \\ & \times \langle x \rangle^{-\delta} \big(R_{\bar{\eta}}^{n-1}(z) + 2(1+z) R_{\bar{\eta}}^{n}(z) + (1+z)^2 R^{n+1}(z) \big) \langle x \rangle^{-\delta} \\ & \times \langle x \rangle^{\delta} R_{\bar{\eta}}(-1) \langle D \rangle^{\beta_r} \langle x \rangle^{-\delta}. \end{aligned}$$

The first and last factors are bounded on L^2 uniformly in $\bar{\eta} \in]0,1]^2$ by pseudodifferential calculus, so it is enough to prove the statement without additional derivatives if $n \geqslant 1$. Since $\beta_l + \beta_r \leqslant 2$ we have a similar argument for $n = 0$.

We have

$$\langle x \rangle^{-\delta} R_{\bar{\eta}}^{n+1}(z) \langle x \rangle^{-\delta} = |z|^{-(n+1)} \langle x \rangle^{-\delta} \Theta_z R_{\bar{\eta},z}^{n+1}(\hat{z}) \Theta_z^{-1} \langle x \rangle^{-\delta}.$$

As in the proof of Theorem 1.3 in Section 3 we can prove by induction on $m \in \mathbb{N}^*$ that $R_{\bar{\eta},z}^{n+1}(\hat{z})$ can be written as a sum of terms of the form

$$(1+\hat{z})^{\beta} R_{\bar{\eta},z}^{n+1+\beta}(-1) \quad \text{or} \quad (1+\hat{z})^{2m-n-1+\nu} R_{\bar{\eta},z}^{m}(-1) R_{\bar{\eta},z}^{\nu}(\hat{z}) R_{\bar{\eta},z}^{m}(-1), \tag{4-10}$$

where $\max(1, n+1-2m) \leqslant \nu \leqslant n+1$ and $\beta \in \mathbb{N}$. Let $s = \min\big(n+1, \frac{d}{2} - \varepsilon\big)$. For $\beta \in \mathbb{N}$ we have $s \in \big[0, \frac{d}{2}\big[$, $n+1+\beta \geqslant s$ and $\delta > s$, so according to the first statement of Proposition 4.5 we have

$$|z|^{-(n+1)} \left\| \langle x \rangle^{-\delta} R_{\bar{\eta},z}^{n+1+\beta}(-1) \langle x \rangle^{-\delta} \right\|_{\mathcal{L}(L^2)} \lesssim |z|^{s-(n+1)} \lesssim 1 + |z|^{\frac{d}{2}-\varepsilon-n-1}.$$

Now we consider the contributions of terms of the second kind in (4-10). We can assume that m is large enough to apply the second statement of Proposition 4.5. We have $\delta > \nu - \frac{1}{2}$, so with Proposition 4.6 we get

$$\begin{aligned} & |z|^{-(n+1)} \left\| \langle x \rangle^{-\delta} \Theta_z R_{\bar{\eta},z}^{m}(-1) R_{\bar{\eta},z}^{\nu}(\hat{z}) R_{\bar{\eta},z}^{m}(-1) \Theta_z^{-1} \langle x \rangle^{-\delta} \right\| \\ & \quad \leqslant |z|^{-(n+1)} \left\| \langle x \rangle^{-\delta} \Theta_z R_{\bar{\eta},z}^{m}(-1) \langle A \rangle^{\delta} \right\| \left\| \langle A \rangle^{-\delta} R_{\bar{\eta},z}^{\nu}(\hat{z}) \langle A \rangle^{-\delta} \right\| \left\| \langle A \rangle^{\delta} R_{\bar{\eta},z}^{m}(-1) \Theta_z^{-1} \langle x \rangle^{-\delta} \right\| \\ & \quad \lesssim |z|^{s-(n+1)} \lesssim 1 + |z|^{\frac{d}{2}-\varepsilon-n-1}. \end{aligned}$$

∎

4B. ***Low-frequency estimates for a general perturbation of the Laplacian.*** In this paragraph we use the estimates on $R_{\bar\eta}(z)$ to prove the same estimates for $R(z)$. To this purpose we have to add the contribution of $P_{\eta_1,c}$ in the self-adjoint part and the contribution of $B^\alpha_{\eta_2,c}$ in the dissipative part.

For $\eta_0,\eta_2 \in]0,1]$ and $\bar\eta = (\eta_0,\eta_2)$ we set, for $\psi \in C_0^\infty(\mathbb{R})$,

$$S_{\psi,\bar\eta}(z) = P_{\eta_0,c} R_{\bar\eta}(z)\psi(P).$$

From Proposition 4.8 we obtain the following result:

Proposition 4.10. *Let $\varepsilon > 0$, $n \in \mathbb{N}$ and $M \in \mathbb{R}$. Let $\psi \in C_0^\infty(\mathbb{R})$. Let δ be as in the statement of Theorem 1.4. Let $\eta_0 \in]0,1]$ be given by Proposition 4.8. Then there exist $C \geqslant 0$ and a neighborhood $\mathcal{U}$ of 0 in $\mathbb{C}$ such that for $\eta_2 \in]0,\eta_0]$ and $z \in \mathcal{U} \cap \mathbb{C}_+$ we have*

$$\|\langle x\rangle^M S^{(n)}_{\psi,\bar\eta}(z)\langle x\rangle^{-\delta}\|_{\mathcal{L}(L^2)} \leqslant C(1+|z|^{\frac{d}{2}-\varepsilon-1-n}),$$

where $\bar\eta = (\eta_0,\eta_2)$.

Proof. The proposition is a consequence of Proposition 4.8, the boundedness of $\psi(P)$ in $L^{2,\delta}$ and the boundedness of $\langle x\rangle^M P_{\eta_0,c}(1-\Delta)^{-1}\langle x\rangle^\delta$. ☐

Remark 4.11. Until now we had not used the distinction between η_1 and η_2. However, the size of $\langle x\rangle^M P_{\eta_1,c}$ depends on η_1, so η_1 has to be fixed in order to obtain uniform estimates in Proposition 4.10 and in Proposition 4.12 below. On the other hand, we have to keep the possibility to take η_2 small. More precisely, the choice of the cut-off function ψ in Proposition 4.12 (and hence in the proof of Proposition 4.13) will depend on η_1, and then the choice of η_2 will in turn depend on ψ. This is why we could not simply take $\eta_1 = \eta_2$ in the definition of $H_{\bar\eta}$.

Proposition 4.12. *Let $\eta_0 \in]0,1]$ be given by Proposition 4.8. Let $\varepsilon_1 > 0$, $\sigma > 2$ and $M \geqslant 0$. Then there exist a bounded neighborhood $\mathcal{U}$ of 0 in $\mathbb{C}$, $\psi \in C_0^\infty(\mathbb{R})$ equal to 1 on a neighborhood of 0 and $\tilde\eta \in]0,\eta_0]$ such that for $\eta_2 \in]0,\tilde\eta]$ and $z \in \mathcal{U}\cap\mathbb{C}_+$ we have*

$$\|\langle x\rangle^M S_{\psi,\bar\eta}(z)\langle x\rangle^{-\sigma}\|_{\mathcal{L}(L^2)} \leqslant \varepsilon_1,$$

where $\bar\eta = (\eta_0,\eta_2)$.

Proof. According to the Hardy inequality we have for $u \in \mathcal{S}$

$$\|\langle x\rangle^M P_{\eta_0,c}u\|_{L^2} \lesssim \sum_{j,k=1}^d \big\|\langle x\rangle^M (D_j(\chi_{\eta_0}G_{j,k}))D_k u\big\|_{L^2} + \sum_{j,k=1}^d \big\|\langle x\rangle^M \chi_{\eta_0}(G_{j,k}-\delta_{j,k})D_jD_k u\big\|_{L^2} \lesssim \|u\|_{\dot H^2}.$$

According to the third statement of Proposition 4.2 we obtain for $\mu > 0$

$$\begin{aligned}
&\|\langle x\rangle^M P_{\eta_0,c}R_{\bar\eta}(i\mu)\psi(P)\langle x\rangle^{-\sigma}u\|_{L^2}\\
&\qquad\lesssim \|R_{\bar\eta}(i\mu)\psi(P)\langle x\rangle^{-\sigma}u\|_{\dot H^2}\\
&\qquad\lesssim \|P_{\eta_0}R_{\bar\eta}(i\mu)\psi(P)\langle x\rangle^{-\sigma}u\|_{L^2}\\
&\qquad\lesssim \|\psi(P)\langle x\rangle^{-\sigma}u\|_{L^2} + \mu\|R_{\bar\eta}(i\mu)\psi(P)\langle x\rangle^{-\sigma}u\|_{L^2} + \|B^\alpha_{\eta_2}R_{\bar\eta}(i\mu)\psi(P)\langle x\rangle^{-\sigma}u\|\\
&\qquad\lesssim \|\psi(P)\langle x\rangle^{-\sigma}u\|_{L^2} + O(\eta_2^{1+\rho})\big\|\langle x\rangle^{-1-\rho}(-\Delta+1)R_{\bar\eta}(i\mu)\psi(P)\langle x\rangle^{-\sigma}u\big\|_{L^2}.
\end{aligned}$$

The term with the factor μ is estimated by the analog of (2-1) for $H_{\bar\eta}$. For the term involving $B^{\alpha}_{\eta_2}$ we have used the fact that

$$\|B^{\alpha}_{\eta_2}(-\Delta+1)^{-1}\langle x\rangle^{1+\rho}\|_{\mathcal{L}(L^2)} = \underset{\eta_2\to 0}{O}(\eta_2^{1+\rho}).$$

Let $\psi_1 \in C_0^\infty(\mathbb{R})$ be equal to 1 on $[-1,1]$. For $\psi \in C_0^\infty(\mathbb{R})$ supported in $]-1,1[$ we have $\psi(P)\langle x\rangle^{-\sigma} = \psi(P)\psi_1(P)\langle x\rangle^{-\sigma}$. The operator $\psi_1(P)\langle x\rangle^{-\sigma}$ is compact. On the other hand, since 0 is not an eigenvalue of P, the operator $\psi(P)$ goes weakly to 0 when the support of ψ shrinks to $\{0\}$. Thus we can find ψ equal to 1 on a neighborhood of 0 such that for $\mu > 0$ and η_2 small enough we have

$$\|\langle x\rangle^{M} S_{\psi,\bar\eta}(i\mu)\langle x\rangle^{-\sigma}\|_{\mathcal{L}(L^2)} \leqslant \tfrac{1}{2}\varepsilon_1.$$

Now let $\tau \in \mathbb{R}$ and $\mu > 0$. We have

$$\begin{aligned}\big\|\langle x\rangle^{M} P_{\eta_0,c}\big(R_{\bar\eta}(\tau+i\mu) - R_{\bar\eta}(i\mu)\big)\psi(P)\langle x\rangle^{-\sigma}\big\| \\ \leqslant \big\|\langle x\rangle^{M} P_{\eta_0,c}(-\Delta+1)^{-1}\langle x\rangle^{\sigma}\big\| \times \int_0^\tau \big\|\langle x\rangle^{-\sigma}(-\Delta+1)R_{\bar\eta}^2(\theta+i\mu)\langle x\rangle^{-\sigma}\big\|\, d\theta.\end{aligned}$$

The first factor is bounded by pseudodifferential calculus, and the second factor is of size $O(|\tau|)$ by Proposition 4.8. Thus this norm is not greater that $\frac{1}{2}\varepsilon_1$ if τ is small enough, and the proposition is proved. □

For $z \in \mathbb{C}_+$ and $\eta_2 \in]0,1]$ we set

$$R_0(z) = (P-z)^{-1} \tag{4-11}$$

and

$$\widetilde{R}_{\eta_2}(z) = (P - iB^{\alpha}_{\eta_2} - z)^{-1}.$$

In the following proposition we prove the resolvent estimates for $\widetilde{R}_{\eta_2}(z)$. Then we will add the contribution of $B^{\alpha}_{\eta_2,c}$ in the dissipative part to conclude the proof of Theorem 1.4.

Proposition 4.13. *Let $\varepsilon > 0$ and $n \in \mathbb{N}$. Let δ be as in Theorem 1.4. Then there exist η_2, $C \geqslant 0$ and a neighborhood $\mathcal{U}$ of 0 in $\mathbb{C}$ such that for $z \in \mathcal{U}\cap\mathbb{C}_+$ and $\beta_l, \beta_r \in \mathbb{R}_+$ with $\beta_l+\beta_r \leqslant 2$ we have*

$$\big\|\langle x\rangle^{-\delta}\langle D\rangle^{\beta_l}\widetilde{R}^{n+1}_{\eta_2}(z)\langle D\rangle^{\beta_r}\langle x\rangle^{-\delta}\big\|_{\mathcal{L}(L^2)} \leqslant C(1+|z|^{\frac{d}{2}-\varepsilon-1-n}).$$

Proof. As for Proposition 4.8 we see that it is enough to consider the case $\beta_l = \beta_r = 0$. Let $\sigma = \max(\delta, 3)$. Let $\varepsilon_1 \in \big]0, \frac14\big]$ and consider $\psi \in C_0^\infty(\mathbb{R})$ as given by Proposition 4.12 for $M = \sigma$. We set $\mathcal{B}_\psi(z) = R_0(z)(1-\psi)(P)$. For any $\gamma \in \mathbb{R}$, this operator and its derivatives are uniformly bounded on $L^{2,\gamma}$ for $z \in \mathbb{C}_+$ close to 0. Let η_0 be given by Proposition 4.8. Given $\eta_2 \in]0,\eta_0]$ we write $\bar\eta$ for (η_0,η_2). We have

$$\widetilde{R}_{\eta_2}(z) = R_{\bar\eta}(z)\psi(P) - \widetilde{R}_{\eta_2}(z)S_{\psi,\bar\eta}(z) + \mathcal{B}_\psi(z) + i\,\widetilde{R}_{\eta_2}(z)B^{\alpha}_{\eta_2}\mathcal{B}_\psi(z),$$

and hence for $n \in \mathbb{N}$

$$\begin{aligned}\widetilde{R}^{(n)}_{\eta_2}(z) = R^{(n)}_{\bar\eta}(z)\psi(P) + \widetilde{R}^{(n)}_{\eta_2}(z)\big(-S_{\psi,\bar\eta}(z) + iB^{\alpha}_{\eta_2}\mathcal{B}_\psi(z)\big) + \mathcal{B}^{(n)}_\psi(z) \\ + i\sum_{j=0}^{n-1} C_j^n\,\widetilde{R}^{(j)}_{\eta_2}(z)\big(-S^{(n-j)}_{\psi,\bar\eta}(z) + iB^{\alpha}_{\eta_2}\mathcal{B}^{(n-j)}_\psi(z)\big).\end{aligned} \tag{4-12}$$

We prove by induction on $n \in \mathbb{N}$ that

$$\|\langle x\rangle^{-\delta} \tilde{R}_{\eta_2}^{(n)}(z)\langle x\rangle^{-\sigma}\| \lesssim 1+|z|^{\frac{d}{2}-\varepsilon-n-1}. \tag{4-13}$$

According to Propositions 4.8, 4.10 and 4.12, the fact that $\psi(P)$ is uniformly bounded on $L^{2,\sigma}$ and the inductive assumption for the sum in (4-12) (it vanishes if $n=0$), there exists $C \geqslant 0$ such that for $z \in \mathbb{C}_+$ close to 0 we have

$$\|\langle x\rangle^{-\delta} \tilde{R}_{\eta_2}^{(n)}(z)\langle x\rangle^{-\sigma}\|\big(1-\varepsilon_1-\|\langle x\rangle^{\sigma} B_{\eta_2}^{\alpha}\mathcal{B}_{\psi}(z)\langle x\rangle^{-\sigma}\|\big) \leqslant C(1+|z|^{\frac{d}{2}-\varepsilon-n-1}).$$

By pseudodifferential calculus we see that the norm of $\langle x\rangle^{\sigma} B_{\eta_2}^{\alpha}\mathcal{B}_{\psi}(z)\langle x\rangle^{-\sigma}$ goes to 0 when η_2 goes to 0. Thus if η_2 is small enough we have

$$1-\varepsilon_1-\|\langle x\rangle^{\sigma} B_{\eta_2}^{\alpha}\mathcal{B}_{\psi}(z)\langle x\rangle^{-\sigma}\| \geqslant \tfrac{1}{2},$$

which concludes the proof of (4-13). In order to replace σ by δ we use (4-12) again and, estimating the second term with (4-13) and Proposition 4.10 instead of Proposition 4.12 we obtain

$$\|\langle x\rangle^{-\delta} \tilde{R}_{\eta_2}^{(n)}(z)\langle x\rangle^{-\delta}\|\big(1-\|\langle x\rangle^{\delta} B_{\eta_2}^{\alpha}\mathcal{B}_{\psi}(z)\langle x\rangle^{-\delta}\|\big) \leqslant C(1+|z|^{\frac{d}{2}-\varepsilon-n-1}),$$

and we conclude similarly. □

In order to prove Theorem 1.4 it remains to add the dissipative part with compactly supported absorption index. We begin with a lemma:

Lemma 4.14. *Let $\mathcal{H}$ be a Hilbert space. Let $\mathcal{R}_0, \mathcal{R}_1 \in \mathcal{L}(\mathcal{H})$ and let $\mathcal{B}$ be such that*

$$\mathcal{R}_1 = \mathcal{R}_0 - \mathcal{R}_0\mathcal{B}\mathcal{R}_1 = \mathcal{R}_0 - \mathcal{R}_1\mathcal{B}\mathcal{R}_0.$$

Then for all $m \in \mathbb{N}$ we can write $\mathcal{R}_1^{m+1}$ as a linear combination of terms of the form

$$\mathcal{R}_0^{m_1+1}\mathcal{B}\mathcal{R}_{j_2}^{m_2+1}\mathcal{B}\cdots\mathcal{B}\mathcal{R}_{j_{k-1}}^{m_k+1}\mathcal{B}\mathcal{R}_0^{m_k+1}, \tag{4-14}$$

where $k \in \mathbb{N}^*$, $j_1, \dots, j_{k-1} \in \{0,1\}$ and $m_1, \dots, m_k \in \mathbb{N}$ are such that

$$\sum_{l=1}^{k} m_l \leqslant m \quad \text{and} \quad m_l = 0 \quad \text{if } j_l = 1.$$

Proof. Using both of the identities above involving $\mathcal{R}_1$ and $\mathcal{R}_0$ we obtain

$$\mathcal{R}_1(z) = \mathcal{R}_0(z) - \mathcal{R}_0(z)\mathcal{B}\mathcal{R}_0(z) + \mathcal{R}_0(z)\mathcal{B}\mathcal{R}_1(z)\mathcal{B}\mathcal{R}_0(z).$$

Then the result is proved by induction on m. □

Now we can finish the proof of Theorem 1.4:

Proof of Theorem 1.4. Let η_2 be given by Proposition 4.13. Let $T = \langle D\rangle^{\frac{\alpha}{2}} a\chi_{\eta_2} \in \mathcal{L}(H^1, L^2)$. We have $T^*T = B_{\eta_2,c}^{\alpha} \leqslant B_\alpha$, so according to Corollary 2.6 we have

$$\|TR(z)T^*\|_{\mathcal{L}(L^2)} \leqslant 1.$$

Let $M \geqslant 0$ and $T_M = \langle x\rangle^{-M}\langle D\rangle^{\frac{\alpha}{2}}$. We can write $B^{\alpha}_{\eta_2,c} = T_M^* \mathcal{B}_1 T = T^* \mathcal{B}_2 T_M = T_M^* \mathcal{B}_3 T_M$, where $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}_3$ are bounded on L^2. According to Lemma 4.14 applied with

$$\mathcal{R}_0 = \widetilde{R}_{\eta_2}(z), \quad \mathcal{R}_1 = R(z) \quad \text{and} \quad \mathcal{B} = B^{\alpha}_{\eta_2,c},$$

we can write $R^{m+1}(z)$ as a sum of terms of the form

$$\mathcal{T} = \mathcal{R}_0^{m_1+1}(z)\mathcal{B}\mathcal{R}_{j_2}^{m_2+1}(z)\mathcal{B}\cdots\mathcal{B}\mathcal{R}_{j_{k-1}}^{m_k+1}(z)\mathcal{B}\mathcal{R}_0^{m_k+1}(z), \tag{4-15}$$

where $k \in \mathbb{N}^*$, $j_1,\dots,j_{k-1} \in \{0,1\}$ and $m_1,\dots,m_k \in \mathbb{N}$ are such that $\sum_{l=1}^k m_l \leqslant m$ and $m_l = 0$ if $j_l = 1$. If $k \geqslant 3$ and M is large enough we obtain for such a term

$$\begin{aligned}
&\|\langle x\rangle^{-\delta}\mathcal{T}\langle x\rangle^{-\delta}\| \\
&\quad\lesssim \|\langle x\rangle^{-\delta}\widetilde{R}_{\eta_2}^{m_1+1}(z)T_M^*\| \times \prod_{\substack{l=2\\ j_l=0}}^{k-1} \|T_M \widetilde{R}_{\eta_2}^{m_l+1}(z)T_M^*\| \times \prod_{\substack{l=2\\ j_l=1}}^{k-1} \|TR(z)T^*\| \times \|T_M \widetilde{R}_{\eta_2}^{m_k+1}(z)\langle x\rangle^{-\delta}\| \\
&\quad\lesssim \prod_{l=1}^{k} (1+|z|^{\frac{d}{2}-1-m_l-\varepsilon}) \lesssim (1+|z|^{\frac{d}{2}-1-m-\varepsilon}).
\end{aligned}$$

The cases $k=1$ and $k=2$ are estimated similarly. This concludes the proof of Theorem 1.4. □

4C. *Sharp low-frequency resolvent estimate.* We finish this section with the proof of Theorem 1.5. The result follows from the self-adjoint analog by a simple perturbation argument, using the quadratic estimates and the spatial decay of the dissipative term:

Proof of Theorem 1.5. According to the resolvent identity, Proposition 2.5 and Theorem 1.1 in [Bouclet and Royer 2015] we have

$$\begin{aligned}
\big\|\langle x\rangle^{-1}R(z)\langle x\rangle^{-1}\big\| &= \big\|\langle x\rangle^{-1}R_0(z)\langle x\rangle^{-1}\big\| + \big\|\langle x\rangle^{-1}R_0(z)\sqrt{B_\alpha}\big\|\,\big\|\sqrt{B_\alpha}R(z)\langle x\rangle^{-1}\big\| \\
&\lesssim 1 + \big\|\langle x\rangle^{-1}R_0(z)\sqrt{B_\alpha}\big\|\,\big\|\langle x\rangle^{-1}R(z)\langle x\rangle^{-1}\big\|^{\frac{1}{2}}.
\end{aligned}$$

Moreover,

$$\big\|\langle x\rangle^{-1}R_0(z)\sqrt{B_\alpha}\big\| \leqslant \big\|\langle x\rangle^{-1}R_0(i)\sqrt{B_\alpha}\big\| + |z-i|\,\big\|\langle x\rangle^{-1}R_0(z)\langle x\rangle^{-1}\big\|\,\big\|\langle x\rangle R_0(i)\sqrt{B_\alpha}\big\| \lesssim 1.$$

For the norms involving $R_0(i)$ we have used the fact that $\langle x\rangle^{\sigma}R_0(i)\sqrt{B_\alpha}$ extends to a bounded operator since for $\sigma \leqslant 1$ and $u \in \mathcal{S}$ we have by pseudodifferential calculus

$$\|\sqrt{B_\alpha}R_0(i)\langle x\rangle^{\sigma}u\|_{L^2}^2 \leqslant \big\langle \langle x\rangle^{\sigma}R_0(-i)B_\alpha R_0(i)\langle x\rangle^{\sigma}u, u\big\rangle \lesssim \|u\|_{L^2}^2.$$

This gives

$$\|\langle x\rangle^{-1}R(z)\langle x\rangle^{-1}\| \lesssim 1 + \|\langle x\rangle^{-1}R(z)\langle x\rangle^{-1}\|^{\frac{1}{2}},$$

from which the conclusion follows. □

5. High-frequency estimates

In this section we prove Theorem 1.6. To this purpose we use semiclassical analysis (see for instance [Zworski 2012]). For $h > 0$ and $\zeta \in \mathbb{C}_+$ we set $H_h = h^2 H$, $P_h = h^2 P$ and $R_h(\zeta) = (H_h - \zeta)^{-1}$. Then for $n \in \mathbb{N}$, $z \in \mathbb{C}_+$ and $h = |z|^{-\frac{1}{2}}$ we have

$$R(z)^{n+1} = \frac{1}{|z|^{n+1}} R_h^{n+1}(\hat{z}) = h^{2(n+1)} R_h^{n+1}(\hat{z}) \tag{5-1}$$

(we recall that $\hat{z} = z/|z|$).

In order to prove uniform estimates for the resolvent $R_h(z)$ we use again the Mourre method. For high frequencies and in a dissipative context we follow [Royer 2010; Bouclet and Royer 2014]. Here we have to be careful with the form of the dissipative part $h^2 B_\alpha$.

Let $\chi_\alpha \in C_0^\infty(\mathbb{R})$ be positive in a neighborhood of 1 and such that $0 \leqslant \chi_\alpha(r) \leqslant r^{\frac{\alpha}{2}}$ for all $r \in \mathbb{R}_+$. For $h \in]0, 1]$ we set

$$B_h^\alpha = a(x) \chi_\alpha(-h^2 \Delta) a(x).$$

Then we have

$$0 \leqslant h^{2-\tilde{\alpha}} B_h^\alpha \leqslant h^{2-\alpha} B_h^\alpha \leqslant h^2 a(x)(-\Delta)^{\frac{\alpha}{2}} a(x) \leqslant h^2 B_\alpha, \tag{5-2}$$

in the sense that for all $\varphi \in H^{\alpha/2}(\mathbb{R}^d)$ we have

$$0 \leqslant h^{2-\tilde{\alpha}} \langle B_h^\alpha \varphi, \varphi \rangle_{L^2(\mathbb{R}^d)} \leqslant h^2 \langle B_\alpha \varphi, \varphi \rangle_{L^2(\mathbb{R}^d)}. \tag{5-3}$$

The operator B_h^α is a bounded pseudodifferential operator on L^2. Its principal symbol is

$$b(x, \xi) = a(x)^2 \chi_\alpha(|\xi|^2).$$

The damping assumption (1-9) on bounded trajectories is satisfied with b instead of a:

$$\forall w \in \Omega_b,\ \exists T \in \mathbb{R} \quad \text{such that} \quad b(\phi^T(w)) > 0.$$

Set

$$f_0(x, \xi) = x \cdot \xi.$$

As in [Bouclet and Royer 2014] (see Proposition 8.1), we can prove that there exist an open neighborhood $\tilde{J}$ of 1, $f_c \in C_0^\infty(\mathbb{R}^{2n}, \mathbb{R})$, $\beta \geqslant 0$ and $c_0 > 0$ such that on $p^{-1}(\tilde{J})$ we have

$$\{p, f_0 + f_c\} + \beta b \geqslant 3c_0, \tag{5-4}$$

where $\{p, q\}$ is the Poisson bracket $\nabla_\xi p \cdot \nabla_x q - \nabla_x p \cdot \nabla_\xi q$. The fact that the symbol of the dissipative part depends on ξ does not change anything in the proof of this statement. We set

$$F_h = \mathrm{Op}_h^w(f_0 + f_c),$$

where Op_h^w is the Weyl quantization:

$$\mathrm{Op}_h^w(q)u(x) = \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{\frac{i}{h}\langle x-y, \xi\rangle} q\left(\frac{x+y}{2}, \xi\right) u(y)\, dy\, d\xi.$$

Let J be a neighborhood of 1 and a compact subset of $\widetilde{J}$. Let $\chi \in C_0^\infty(\widetilde{J}, [0,1])$ be equal to 1 on a neighborhood of J. After multiplication by $(\chi \circ p)^2$, the (easy) Gårding inequality (Theorem 4.26 in [Zworski 2012]) gives for $h > 0$ small enough

$$\operatorname{Op}_h\big((\chi \circ p)^2\{p, f_0 + f_c\} + \beta b(\chi \circ p)^2 + 3c_0(1 - (\chi \circ p)^2)\big) \geqslant 3c_0 - O(h) \geqslant 2c_0.$$

After multiplication by $h^{2-\tilde{\alpha}}$ we obtain

$$\chi(P_h)\big([P_h, ih^{1-\tilde{\alpha}} F_h] + \beta h^{2-\tilde{\alpha}} B_h^\alpha\big)\chi(P_h) + 3c_0(1-\chi^2)(P_h) \geqslant 2c_0 h^{2-\tilde{\alpha}} - O(h^{3-\tilde{\alpha}}).$$

After conjugation by $\mathbb{1}_J(P_h)$ we obtain for h small enough

$$\mathbb{1}_J(P_h)\big([P_h, ih^{1-\tilde{\alpha}} F_h] + \beta h^{2-\tilde{\alpha}} B_h^\alpha\big)\mathbb{1}_J(P_h) \geqslant c_0 h^{2-\tilde{\alpha}} \mathbb{1}_J(P_h).$$

According to (5-2) this finally gives

$$\mathbb{1}_J(P_h)\big([P_h, ih^{1-\tilde{\alpha}} F_h] + \beta h^2 B_\alpha\big)\mathbb{1}_J(P_h) \geqslant c_0 h^{2-\tilde{\alpha}} \mathbb{1}_J(P_h), \tag{5-5}$$

which is the main assumption of Definition 2.3 with βh^2 instead of β and $\alpha = c_0 h^{2-\tilde{\alpha}}$.

It remains to check the other assumptions of Definition 2.3. The first is proved as in [Bouclet and Royer 2014] (except that we look at the norm in the form domain H^1 instead of the domain H^2), and the commutator properties are proved using (standard) pseudodifferential calculus, considering h as a parameter (for the dissipative part we cannot use $h^{2-\alpha} B_h^\alpha$ as above, so we have to control directly the commutators of $h^2 B_\alpha$ with $h^{1-\tilde{\alpha}} F_h$).

Thus we have proved that for $h \in \,]0, h_0]$ the operator $h^{1-\tilde{\alpha}} F_h$ is a conjugate operator to H_h on a neighborhood J of 1 with lower bounds $h^{2-\tilde{\alpha}} c_0$ for some $c_0 > 0$. According to Theorem 2.4 we have proved the following result with $\langle F_h \rangle^{-\delta}$ instead of $\langle x \rangle^{-\delta}$:

Proposition 5.1. *Let $n \in \mathbb{N}$ and $\delta > n + \frac{1}{2}$. There exists a neighborhood J of 1, $h_0 > 0$ and $C \geqslant 0$ such that for all $\zeta \in \mathbb{C}_+$ with $\operatorname{Re}(\zeta) \in J$ we have*

$$\|\langle x \rangle^{-\delta} R_h^{n+1}(\zeta) \langle x \rangle^{-\delta}\|_{\mathcal{L}(L^2)} \leqslant \frac{C}{h^{(2-\tilde{\alpha})(n+1)}}.$$

In order to have the estimate with $\langle x \rangle^{-\delta}$ we proceed as usual (see the end of Section 3 for intermediate frequencies or [Royer 2010] in the semiclassical context). With (5-1) and Proposition 5.1 we obtain the second statement of Theorem 1.6. For the first statement, we observe that under the nontrapping condition we can proceed as above with $\beta = 0$ and with $\tilde{\alpha}$ replaced by 1 in (5-5).

6. Local energy decay

In this section we use Theorems 1.3, 1.4 and 1.6 to prove Theorem 1.1. Let $u_0 \in \mathcal{S}$. We denote by u the solution of (1-1). Let $\mu > 0$. For $t \in \mathbb{R}$ we set

$$u_\mu(t) = \mathbb{1}_{\mathbb{R}_+}(t) u(t) e^{-t\mu}.$$

Then for $\tau \in \mathbb{R}$ we set

$$\check{u}_\mu(\tau) = \int_{\mathbb{R}} e^{it\tau} u_\mu(t)\, dt = \int_0^{+\infty} e^{it(\tau+i\mu)} u(t)\, dt, \tag{6-1}$$

so that for all $n \in \mathbb{N}$ and $\tau \in \mathbb{R}$ we have

$$\check{u}_\mu^{(n)}(\tau) = \int_{\mathbb{R}} (it)^n e^{it\tau} u_\mu(t)\, dt. \tag{6-2}$$

We multiply (1-1) by $e^{it(\tau+i\mu)}$ and integrate over $\mathbb{R}_+$. This yields

$$(H - (\tau + i\mu))\check{u}_\mu(\tau) = -iu_0$$

and hence, for all $n \in \mathbb{N}$

$$\check{u}_\mu^{(n)}(\tau) = -i\, n!\, R^{n+1}(\tau + i\mu)u_0. \tag{6-3}$$

Lemma 6.1. *For all $n \in \mathbb{N}^*$ and $\mu > 0$ the map $\tau \mapsto R^{n+1}(\tau + i\mu)u_0$ belongs to $L^1(\mathbb{R}, L^2(\mathbb{R}^d))$.*

Proof. Let $\chi_0 \in C_0^\infty(\mathbb{R}, [0, 1])$ be equal to 1 on a neighborhood of 0. According to (1-7) the map $\tau \mapsto R^{n+1}(\tau + i\mu)u_0$ is bounded, so it is enough to prove that $\tau \mapsto (1-\chi_0)(\tau)R^{n+1}(\tau + i\mu)u_0$ belongs to $L^1(\mathbb{R})$. Let $z \in \mathbb{C}_+$. Using twice the identity

$$R(z) = \frac{R(z)(H+1) - 1}{z+1},$$

we get

$$R(z)u_0 = \frac{1}{(z+1)^2} R(z)(H+1)^2 u_0 - \frac{1}{(z+1)^2}(H+1)u_0 - \frac{1}{z+1}u_0.$$

The result follows after at least one differentiation with respect to z. □

This lemma does not provide any uniform estimate, but now we can take the Fourier transform of (6-2). With (6-3) this gives for all $t \geqslant 0$

$$(it)^n e^{-t\mu} u(t) = -\frac{i\, n!}{2\pi} \int_{\tau\in\mathbb{R}} e^{-it\tau} R^{n+1}(\tau + i\mu)u_0\, d\tau. \tag{6-4}$$

We consider $\chi_-, \chi_0, \chi \in C^\infty(\mathbb{R}, [0, 1])$ such that χ_- is supported in $]-\infty, 0[$, χ_0 is compactly supported and equal to 1 on a neighborhood of 0, χ is compactly supported in $]0, +\infty[$ and

$$\chi_- + \chi_0 + \sum_{j\in\mathbb{N}^*} \chi_j = 1 \quad \text{on } \mathbb{R},$$

where for $j \in \mathbb{N}^*$ and $\tau \in \mathbb{R}$ we have set $\chi_j(\tau) = \chi(\tau/2^{j-1})$. We set $\chi_+ = \sum_{j\in\mathbb{N}^*} \chi_j$. Starting from (6-4) applied with $n = \kappa - 1$ (κ was defined in (1-5)) we can write

$$u_\mu(t) = -\frac{i\, n!}{2\pi (it)^{\kappa-1}} \big(v_-(t) + v_0(t) + v_+(t)\big), \tag{6-5}$$

where for $* \in \{-, 0, +\}$ we have set

$$v_*(t) = \int_{\tau\in\mathbb{R}} \chi_*(\tau) e^{-it\tau} R^\kappa(\tau + i\mu)u_0\, d\tau. \tag{6-6}$$

To simplify the notation we forget the dependence on μ. From now on, all the quantities depend on $\mu > 0$ but the estimates are uniform in μ.

Proposition 6.2. *Let $k \in \mathbb{N}$. There exists $C \geqslant 0$ which does not depend on $u_0 \in \mathcal{S}$ such that for all $\mu > 0$ and $t \geqslant 0$ we have*

$$\|v_-(t)\|_{L^2} \leqslant C\,\langle t\rangle^{-k}\,\|u_0\|_{L^2}.$$

This implies that the corresponding contribution for $u(t)$ decays like any power of t in L^2.

Proof. After k partial integrations in (6-6) we get

$$(it)^k v_-(t) = \int_{\mathbb{R}} e^{-it\tau} \frac{d^k}{d\tau^k}\big(\chi_-(\tau)R(\tau+i\mu)^\kappa\big)u_0\,d\tau.$$

According to Remark 2.1 we have

$$\left\| \frac{d^k}{d\tau^k}\big(\chi_-(\tau)R(\tau+i\mu)^\kappa\big)\right\|_{\mathcal{L}(L^2)} \lesssim \langle\tau\rangle^{-(\kappa+k)},$$

and the result follows. □

We now deal with v_0. The following result is (a slightly modified version of) Lemma 4.3 in [Bouclet and Royer 2014]:

Lemma 6.3. *Let $\mathcal{H}$ be a Hilbert space. Let $f \in C^1(\mathbb{R}^*, \mathcal{H})$ be equal to 0 outside a compact subset of $\mathbb{R}$. Assume that for some $\gamma \in \,]0,1[$ and $M_f \geqslant 0$ we have*

$$\forall \tau \in \mathbb{R}^*, \quad \|f(\tau)\|_{\mathcal{H}} \leqslant M_f|\tau|^{-\gamma} \quad \textit{and} \quad \|f'(\tau)\|_{\mathcal{H}} \leqslant M_f|\tau|^{-1-\gamma}.$$

Let $\beta \in [0,1[$. Then there exists $C \geqslant 0$ which does not depend on f and such that for all $t \in \mathbb{R}$ we have

$$\|\hat{f}(t)\|_{\mathcal{H}} \leqslant C\,M_f\,\langle t\rangle^{\beta(\gamma-1)}.$$

Proof. Following the proof of [Bouclet and Royer 2014] we set $f_t(\tau) = \int_{-1}^{1}\phi(s)f\big(\tau-\frac{s}{t}\big)\,ds$, where $\phi \in C_0^\infty(]-1,1[,\mathbb{R})$ satisfies $\int_{\mathbb{R}}\phi = 1$ and we write for $|t| \geqslant 1$

$$|\hat{f}(t)| \leqslant \int_{|\tau|\leqslant t^{-\beta}} \|f(\tau)\|\,d\tau + \int_{|\tau|\geqslant t^{-\beta}} \|f(\tau)-f_t(\tau)\|\,d\tau + \left\|\int_{|\tau|\geqslant t^{-\beta}} e^{-it\tau} f_t(\tau)\,d\tau\right\|$$
$$\lesssim |t|^{-\beta(1-\gamma)} + |t|^{\gamma\beta-1} + \frac{1}{t}\left(\|f_t(t^{-\beta})\| + \|f_t(-t^{-\beta})\| + \left\|\int_{|\tau|\geqslant t^{-\beta}} e^{-it\tau} f_t'(\tau)\,d\tau\right\|\right) \lesssim |t|^{-\beta(1-\gamma)}.$$

We omit the details. □

Proposition 6.4. *Let $\varepsilon \in \big]0,\frac{1}{2}\big[$ and $\delta > \kappa + \frac{1}{2}$. Then there exists $C \geqslant 0$ which does not depend on $u_0 \in \mathcal{S}$ and such that for all $\mu > 0$ and $t \geqslant 0$ we have*

$$\|v_0(t)\|_{L^{2,-\delta}} \leqslant \langle t\rangle^{\kappa-1-\frac{d}{2}+\varepsilon}\|u_0\|_{L^{2,\delta}}.$$

Proof. According to Theorem 1.4 applied with $\frac{\varepsilon}{2}$ instead of ε and Theorem 1.3 there exists $C \geqslant 0$ (which does not depend on u_0) such that for $\mu > 0$, $\tau \in \mathbb{R}$ and $z = \tau + i\mu$ we have

$$\|\chi_0(\tau) R^\kappa(z) u_0\|_{L^{2,-\delta}} \leqslant C|z|^{\frac{d}{2}-\kappa-\frac{\varepsilon}{2}} \|u_0\|_{L^{2,\delta}},$$

$$\left\| \frac{d}{d\tau}\big(\chi_0(\tau) R^\kappa(z)\big) u_0 \right\|_{L^{2,-\delta}} \leqslant C|z|^{\frac{d}{2}-\kappa-1-\frac{\varepsilon}{2}} \|u_0\|_{L^{2,\delta}}.$$

Then the statement follows from Lemma 6.3 applied with $\beta \in]0,1[$ so close to 1 that

$$\beta\Big(\kappa - \frac{d}{2} - 1 + \frac{\varepsilon}{2}\Big) \leqslant \kappa - \frac{d}{2} - 1 + \varepsilon. \qquad \square$$

To finish the proof of Theorem 1.1 we have to estimate $v_+(t)$. As for $v_-(t)$ above, k partial integrations yield

$$\begin{aligned}(it)^k v_+(t) &= \int_{\mathbb{R}} e^{-it\tau} \sum_{j=1}^{k} C_k^j \chi_+^{(j)}(\tau) R^{\kappa+k-j}(\tau+i\mu) u_0 \, d\tau + \int_{\mathbb{R}} e^{-it\tau} \chi_+(\tau) R^{\kappa+k}(\tau+i\mu) u_0 \, d\tau \\ &=: v_{+,k}^0(t) + w_k(t).\end{aligned}$$

The following proposition proves that the contribution of $v_+(t)$ in (6-5) decays like any power of t. However, there may be a loss of two derivatives when $\alpha = 0$ if the nontrapping assumption does not hold. We apply the following result with $k \geqslant 1$ to conclude the proof of Theorem 1.1.

Proposition 6.5. *Let $k \in \mathbb{N}^*$ and $\delta > \kappa + k - \frac{1}{2}$. Let $\sigma \in [0,2]$.*

(i) *There exists $C \geqslant 0$ which does not depend on u_0 and such that for all $\mu > 0$ and $t \geqslant 1$ we have*

$$\|\langle x\rangle^{-\delta} v_{+,k}^0(t)\|_{L^2} \leqslant C\|u_0\|_{L^{2,\delta}}.$$

(ii) *Assume that the nontrapping assumption* (1-8) *holds or that we have the damping condition* (1-9) *together with $(\kappa+k)\tilde{\alpha} + \sigma \geqslant 2$. Then there exists $C \geqslant 0$ which does not depend on u_0 such that for all $\mu > 0$ and $t \geqslant 1$ we have*

$$\|\langle x\rangle^{-\delta} w_k(t)\|_{L^2} \leqslant C\|u_0\|_{H^{\sigma,\delta}}.$$

Proof. Statement (i) follows from Theorem 1.3 and the fact that $\chi_+^{(j)}$ is compactly supported in $]0,+\infty[$ for all $j \geqslant 1$. We turn to the proof of (ii).

• For $j \in \mathbb{N}^*$ we set

$$w_{k,j}(t) = \int_{\tau\in\mathbb{R}} \chi_j(\tau) e^{-it\tau} R^{\kappa+k}(\tau+i\mu) u_0 \, d\tau.$$

Let $\tilde{\chi} \in C_0^\infty(\mathbb{R}_+^*, [0,1])$ be equal to 1 on a neighborhood of $\operatorname{supp}\chi$. For $\tau \in \mathbb{R}$ and $j \in \mathbb{N}^*$ we set $\tilde{\chi}_j(\tau) = \tilde{\chi}(\tau/2^{j-1})$. Let

$$I_{k,j}(t) = \int_{\tau\in\mathbb{R}} \chi_j(\tau) e^{-it\tau} \langle x\rangle^{-\delta} R^{\kappa+k}(\tau+i\mu) \langle x\rangle^{-\delta} \, d\tau \in \mathcal{L}(L^2).$$

We have

$$\langle x\rangle^{-\delta} w_{k,j}(t) = w_{k,j}^1(t) + w_{k,j}^2(t) + w_{k,j}^3(t),$$

where

$$w_{k,j}^1(t) = \tilde{\chi}_j(P) I_{k,j}(t) \tilde{\chi}_j(P) \langle x \rangle^{\delta} u_0,$$
$$w_{k,j}^2(t) = (1-\tilde{\chi}_j)(P) I_{k,j}(t) \tilde{\chi}_j(P) \langle x \rangle^{\delta} u_0,$$
$$w_{k,j}^3(t) = I_{k,j}(t)(1-\tilde{\chi}_j)(P) \langle x \rangle^{\delta} u_0.$$

- By almost orthogonality, Theorem 1.6 and almost orthogonality again we have

$$\begin{aligned}
\Big\| \sum_{j\in\mathbb{N}^*} w_{k,j}^1(t) \Big\|^2 &\lesssim \sum_{j\in\mathbb{N}^*} \| w_{k,j}^1(t) \|^2 \\
&\lesssim \sup_{j\in\mathbb{N}^*} \left(\int_{\tau\in\mathbb{R}} \chi_j(\tau) \big\| \langle x \rangle^{-\delta} R^{\kappa+k}(\tau+i\mu) \langle x \rangle^{-\delta} \big\| \, d\tau \right)^2 \times \sum_{j\in\mathbb{N}^*} \big\| \tilde{\chi}_j(P) \langle P \rangle^{-\frac{\sigma}{2}} \langle P \rangle^{\frac{\sigma}{2}} \langle x \rangle^{\delta} u_0 \big\|^2 \\
&\lesssim \sup_{j\in\mathbb{N}^*} 2^{2j} 2^{-j(\kappa+k)\alpha} 2^{-j\sigma} \| \langle P \rangle^{\frac{\sigma}{2}} \langle x \rangle^{\delta} u_0 \|^2 \\
&\lesssim \| u_0 \|_{H^{\sigma,\delta}}^2.
\end{aligned}$$

It remains to prove that

$$\| w_{k,j}^2(t) \| + \| w_{k,j}^3(t) \| \lesssim 2^{-j} \| u_0 \|_{L^{2,\delta}}. \tag{6-7}$$

- For the contribution of $w_{k,j}^2(t)$ we prove that there exists $C \geqslant 0$ such that for $j \in \mathbb{N}^*$ and $\tau \in \operatorname{supp}(\chi_j)$ we have

$$\big\| (1-\tilde{\chi}_j)(P) \langle x \rangle^{-\delta} R^{\kappa+k}(\tau+i\mu) \langle x \rangle^{-\delta} \big\|_{\mathcal{L}(L^2)} \leqslant C 2^{-2j}. \tag{6-8}$$

For this, we prove by induction on $m \in \mathbb{N}^*$ and then on $\ell \in \mathbb{N}$ that for $\delta > m - \frac{1}{2}$ and $\phi \in C_0^\infty(\mathbb{R}_+^*, [0,1])$ equal to 1 on a neighborhood of $\operatorname{supp}(\chi)$ there exists $C \geqslant 0$ such that for all $j \in \mathbb{N}^*$, $z = \tau + i\mu$ with $\tau \in \operatorname{supp}(\chi)$ and $\mu > 0$

$$\big\| \langle x \rangle^{-\delta} (1-\phi_j)(P) R^m(z) \langle x \rangle^{-\delta} \big\| \leqslant C 2^{-j \min\left(m, \ell\left(1-\frac{\alpha}{2}\right)\right)}, \tag{6-9}$$

where for $j \in \mathbb{N}^*$ we have set $\phi_j = \phi(\cdot/2^{j-1})$. Let $m \in \mathbb{N}^*$. If $m \geqslant 2$, we assume that the estimate is proved up to order $m-1$ (for all $\ell \in \mathbb{N}$). Note that we will not use any inductive assumption on m for $m = 1$. Then we prove the estimate by induction on $\ell \in \mathbb{N}$. For $\ell = 0$ it follows from Theorem 1.6 and the boundedness of $(1-\phi_j)(P)$ in weighted spaces. Assume that (6-9) is proved up to order $\ell - 1$ for some $\ell \in \mathbb{N}^*$. Let $\tilde{\phi} \in C_0^\infty(\mathbb{R}_+^*)$ be equal to 1 on a neighborhood of $\operatorname{supp}(\chi)$ and such that $\phi = 1$ on a neighborhood of $\operatorname{supp}(\tilde{\phi})$. For $j \in \mathbb{N}^*$ we set $\tilde{\phi}_j = \tilde{\phi}(\cdot/2^{j-1})$. We recall that for $z \in \mathbb{C}_+$ we have set $R_0(z) = (P-z)^{-1}$. By the resolvent identity we have

$$R^m(z) = R_0(z) R^{m-1}(z) + i R_0(z) B_\alpha R^m(z). \tag{6-10}$$

If $m = 1$, the first term is just $R_0(z)$ and we have

$$\big\| \langle x \rangle^{-\delta} (1-\phi_j)(P) R_0(z) \langle x \rangle^{-\delta} \big\| \lesssim 2^{-j}.$$

If $m \geqslant 2$, the contribution of the first term is estimated as follows:

$$\begin{aligned}
&\big\| \langle x\rangle^{-\delta}(1-\phi_j)(P)R_0(z)R^{m-1}(z)\langle x\rangle^{-\delta}\big\| \\
&\qquad\leqslant \big\| \langle x\rangle^{-\delta}(1-\phi_j)(P)R_0(z)\tilde{\phi}_j(P)\langle x\rangle^{\delta}\big\| \big\| \langle x\rangle^{-\delta}R^{m-1}(z)\langle x\rangle^{-\delta}\big\| \\
&\qquad\leqslant \big\| \langle x\rangle^{-\delta}(1-\phi_j)(P)R_0(z)\langle x\rangle^{\delta}\big\| \big\| \langle x\rangle^{-\delta}(1-\tilde{\phi}_j(P))R^{m-1}(z)\langle x\rangle^{-\delta}\big\|.
\end{aligned}$$

Using Theorem 8.7 in [Dimassi and Sjöstrand 1999] about functions of a self-adjoint semiclassical pseudodifferential operator (with $h = 2^{-\frac{j-1}{2}}$) we see that for any $M \geqslant 0$ we have

$$\big\| \langle x\rangle^{-\delta}(1-\phi_j)(P)R_0(z)\langle x\rangle^{\delta}\tilde{\phi}_j(P)\big\|_{\mathcal{L}(L^2)} \lesssim 2^{-jM}.$$

We also have

$$\big\| \langle x\rangle^{-\delta}(1-\phi_j)(P)R_0(z)\langle x\rangle^{\delta}\big\| \lesssim 2^{-j}.$$

With Theorem 1.6 and the inductive assumption (for $\tilde{\phi}$ instead of ϕ) we obtain (6-9) with $R^m(z)$ replaced by $R_0(z)R^{m-1}(z)$. For the contribution of the second term in (6-10) we similarly write

$$\begin{aligned}
&\big\| \langle x\rangle^{-\delta}(1-\phi_j)(P)R_0(z)B_\alpha R^{m}(z)\langle x\rangle^{-\delta}\big\| \\
&\qquad\leqslant \big\| \langle x\rangle^{-\delta}(1-\phi_j)(P)R_0(z)B_\alpha\tilde{\phi}_j(P)\langle x\rangle^{\delta}\big\| \big\| \langle x\rangle^{-\delta}R^{m}(z)\langle x\rangle^{-\delta}\big\| \\
&\qquad\leqslant \big\| \langle x\rangle^{-\delta}(1-\phi_j)(P)R_0(z)B_\alpha\langle x\rangle^{\delta}\big\| \big\| \langle x\rangle^{-\delta}(1-\tilde{\phi}_j(P))R^{m}(z)\langle x\rangle^{-\delta}\big\|.
\end{aligned}$$

Here we only have

$$\big\| \langle x\rangle^{-\delta}(1-\phi_j)(P)R_0(z)B_\alpha\langle x\rangle^{\delta}\big\| \lesssim 2^{-j(1-\frac{\alpha}{2})},$$

but with the inductive assumption (on ℓ), we still can conclude. Thus, (6-9) is proved for all $m, \ell \in \mathbb{N}$. With $m = \kappa + k$ and ℓ large enough (we recall that $\alpha < 2$), this gives (6-8). After integration over $\tau \in \operatorname{supp}(\chi_j)$, this gives (6-7) for $w^2_{k,j}(t)$. The contribution of $w^3_{k,j}(t)$ is estimated similarly, and the proof is complete. ☐

7. Smoothing effect

In this section we prove Theorem 1.7. With Theorems 1.3, 1.5 and 1.6 it implies Theorem 1.2. For this we use a dissipative version of the theory of relatively smooth operators in the sense of Kato.

Proposition 7.1. *Under the assumption of Theorem 1.7 there exists $C \geqslant 0$ such that for all $z \in \mathbb{C}_+$ we have*

$$\|\langle x\rangle^{-1}\langle P\rangle^{\frac{\gamma}{4}}R(z)\langle P\rangle^{\frac{\gamma}{4}}\langle x\rangle^{-1}\|_{\mathcal{L}(L^2)} \leqslant C.$$

Proof. • Let K be a compact subset of $\mathbb{C}$. Using the resolvent identity

$$R(z) = R(i) + (z-i)R(i)^2 + (z-i)^2R(i)R(z)R(i),$$

we obtain for $z \in \mathbb{C}_+ \cap K$

$$\big\|\langle x\rangle^{-1}\langle P\rangle^{\frac{\gamma}{4}}R(z)\langle P\rangle^{\frac{\gamma}{4}}\langle x\rangle^{-1}\big\| \lesssim 1 + \big\|\langle x\rangle^{-1}\langle P\rangle^{\frac{\gamma}{4}}R(i)\langle x\rangle\big\| \big\|\langle x\rangle^{-1}R(z)\langle x\rangle^{-1}\big\| \big\|\langle x\rangle R(i)\langle P\rangle^{\frac{\gamma}{4}}\langle x\rangle^{-1}\big\|.$$

By pseudodifferential calculus the operators $\langle P\rangle^{\frac{\gamma}{4}} R(i)$ and $R(i)\langle P\rangle^{\frac{\gamma}{4}}$ are bounded on $L^{2,-1}$ and $L^{2,1}$, respectively. For the second factor on the right-hand side we use (1-10), and the conclusion follows for $z \in \mathbb{C}_+ \cap K$.

- It remains to prove the result for $|z| \gg 1$. Let $\chi \in C_0^\infty(\mathbb{R}, [0, 1])$ be supported on $[-3, 3]$ and equal to 1 on $[-2, 2]$. For $z \in \mathbb{C}_+$ we define $\chi_z : \lambda \mapsto \chi(\lambda/|z|)$. The operator $\varepsilon^{\frac{\gamma}{4}} \langle P\rangle^{\frac{\gamma}{4}} \langle \varepsilon P\rangle^{-\frac{\gamma}{4}}$ is a pseudodifferential operator whose symbol has bounded derivatives uniformly in $\varepsilon \in]0, 1]$, so the operator

$$|z|^{-\frac{\gamma}{4}} \langle x\rangle^{-1} \langle P\rangle^{\frac{\gamma}{4}} \left\langle \frac{P}{|z|}\right\rangle^{-\frac{\gamma}{4}} \langle x\rangle \tag{7-1}$$

extends to a bounded operator on L^2 uniformly in z with $|z| \geqslant 1$. The operator

$$\langle x\rangle^{-1} \left\langle \frac{P}{|z|}\right\rangle^{\frac{\gamma}{4}} \chi\left(\frac{P}{|z|}\right) \langle x\rangle \tag{7-2}$$

is also bounded on L^2 uniformly in z with $|z| \geqslant 1$, and we have similar estimates for the adjoint operators of (7-1) and (7-2). Thus

$$\begin{aligned}\left\|\langle x\rangle^{-1} \langle P\rangle^{\frac{\gamma}{4}} \chi_z(P) R(z) \chi_z(P) \langle P\rangle^{\frac{\gamma}{4}} \langle x\rangle^{-1}\right\| &\lesssim |z|^{\frac{\gamma}{2}} \left\|\langle x\rangle^{-1} \langle |z|^{-1} P\rangle^{\frac{\gamma}{4}} \chi_z(P) R(z) \chi_z(P) \langle |z|^{-1} P\rangle^{\frac{\gamma}{4}} \langle x\rangle^{-1}\right\| \\ &\lesssim |z|^{\frac{\gamma}{2}} \left\|\langle x\rangle^{-1} R(z) \langle x\rangle^{-1}\right\| \lesssim 1.\end{aligned}$$

- With $R_0(z) = (P - z)^{-1}$ we have the resolvent identity

$$R(z) = R_0(z) + i R(z) B_\alpha R_0(z).$$

We have

$$\begin{aligned}\left\|\langle x\rangle^{-1} \langle P\rangle^{\frac{\gamma}{4}} \chi_z(P) R_0(z) (1-\chi_z)(P) \langle P\rangle^{\frac{\gamma}{4}} \langle x\rangle^{-1}\right\| &\leqslant \left\|\langle x\rangle^{-1} \langle P\rangle^{\frac{\gamma}{4}} \chi_z(P)\right\| \left\|R_0(z)(1-\chi_z)(P) \langle P\rangle^{\frac{\gamma}{4}} \langle x\rangle^{-1}\right\| \\ &\lesssim \langle z\rangle^{\frac{\gamma}{4}} \langle z\rangle^{\frac{\gamma}{4}-1} \lesssim 1.\end{aligned}$$

We have estimated the first factor as above and the second by the spectral theorem. On the other hand, since the operator $\sqrt{B_\alpha} \langle P\rangle^{-\frac{1}{2}}$ is bounded we also have by Proposition 2.5

$$\begin{aligned}&\left\|\langle x\rangle^{-1} \langle P\rangle^{\frac{\gamma}{4}} \chi_z(P) R(z) B_\alpha R_0(z) (1-\chi_z)(P) \langle P\rangle^{\frac{\gamma}{4}} \langle x\rangle^{-1}\right\| \\ &\qquad \leqslant \left\|\langle x\rangle^{-1} \langle P\rangle^{\frac{\gamma}{4}} \chi_z(P) \langle x\rangle\right\| \left\|\langle x\rangle^{-1} R(z) \sqrt{B_\alpha}\right\| \left\|\langle P\rangle^{\frac{1}{2}} R_0(z) (1-\chi_z)(P) \langle P\rangle^{\frac{\gamma}{4}}\right\| \\ &\qquad \lesssim \langle z\rangle^{\frac{\gamma}{4}} \langle z\rangle^{-\frac{\gamma}{4}} \langle z\rangle^{\frac{1}{2}+\frac{\gamma}{4}-1} \lesssim 1.\end{aligned}$$

This proves that

$$\left\|\langle x\rangle^{-1} \langle P\rangle^{\frac{\gamma}{4}} \chi_z(P) R(z) (1-\chi_z)(P) \langle P\rangle^{\frac{\gamma}{4}} \langle x\rangle^{-1}\right\| \lesssim 1.$$

- The operator

$$\langle x\rangle^{-1} \langle P\rangle^{\frac{\gamma}{4}} (1-\chi_z)(P) R(z) \chi_z(P) \langle P\rangle^{\frac{\gamma}{4}} \langle x\rangle^{-1}$$

is estimated similarly. Finally for

$$\langle x\rangle^{-1} \langle P\rangle^{\frac{\gamma}{4}} (1-\chi_z)(P) R(z) (1-\chi_z)(P) \langle P\rangle^{\frac{\gamma}{4}} \langle x\rangle^{-1}$$

we only have to use twice the resolvent identity

$$R(z) = R_0(z) + iR_0(z)B_\alpha R_0(z) - R_0(z)B_\alpha R(z)B_\alpha R_0(z).$$

Then we apply the same idea as above, using Corollary 2.6 to estimate $\sqrt{B_\alpha}R(z)\sqrt{B_\alpha}$. □

Taking the adjoint in the estimate of Proposition 7.1 we obtain the same estimate with $R(z)$ replaced by $R(z)^* = (P + iB_\alpha - \bar{z})^{-1}$ (the same is true for the estimates of Theorems 1.3, 1.5 and 1.6). In particular we obtain the following result:

Corollary 7.2. *Then there exists $C \geqslant 0$ such that for all $z \in \mathbb{C}_+$ and $\varphi \in \mathcal{S}$ we have*

$$\left|\left\langle\left((H-z)^{-1} - (H^* - \bar{z})^{-1}\right)\langle P\rangle^{\frac{\gamma}{4}}\langle x\rangle^{-1}\varphi, \langle P\rangle^{\frac{\gamma}{4}}\langle x\rangle^{-1}\varphi\right\rangle_{L^2}\right| \leqslant C\|\varphi\|_{L^2}^2.$$

It is known that such an estimate on the resolvent implies Theorem 1.7. This comes from the dissipative version of the theory of relatively smooth operators. The self-adjoint theory can be found in [Reed and Simon 1978, §XIII.7]. The dissipative version uses the theory of self-adjoint dilations for a dissipative operator described in [Sz.-Nagy et al. 2010]. All this has been combined in Proposition 6.2 in [Royer 2016], from which Theorem 1.7 follows.

References

[Aloui 2008a] L. Aloui, "Smoothing effect for regularized Schrödinger equation on bounded domains", *Asymptot. Anal.* **59**:3-4 (2008), 179–193. MR Zbl

[Aloui 2008b] L. Aloui, "Smoothing effect for regularized Schrödinger equation on compact manifolds", *Collect. Math.* **59**:1 (2008), 53–62. MR Zbl

[Aloui and Khenissi 2002] L. Aloui and M. Khenissi, "Stabilisation pour l'équation des ondes dans un domaine extérieur", *Rev. Mat. Iberoamericana* **18**:1 (2002), 1–16. MR Zbl

[Aloui and Khenissi 2007] L. Aloui and M. Khenissi, "Stabilization of Schrödinger equation in exterior domains", *ESAIM Control Optim. Calc. Var.* **13**:3 (2007), 570–579. MR Zbl

[Aloui and Khenissi 2010] L. Aloui and M. Khenissi, "Boundary stabilization of the wave and Schrödinger equations in exterior domains", *Discrete Contin. Dyn. Syst.* **27**:3 (2010), 919–934. MR Zbl

[Aloui et al. 2013] L. Aloui, M. Khenissi, and G. Vodev, "Smoothing effect for the regularized Schrödinger equation with non-controlled orbits", *Comm. Partial Differential Equations* **38**:2 (2013), 265–275. MR Zbl

[Aloui et al. 2017] L. Aloui, M. Khenissi, and L. Robbiano, "The Kato smoothing effect for regularized Schrödinger equations in exterior domains", *Ann. Inst. H. Poincaré Anal. Non Linéaire* (online publication January 2017).

[Amrein et al. 1996] W. O. Amrein, A. Boutet de Monvel, and V. Georgescu, *C_0-groups, commutator methods and spectral theory of N-body Hamiltonians*, Progress in Mathematics **135**, Birkhäuser, Basel, 1996. MR Zbl

[Bardos et al. 1992] C. Bardos, G. Lebeau, and J. Rauch, "Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary", *SIAM J. Control Optim.* **30**:5 (1992), 1024–1065. MR Zbl

[Ben-Artzi and Klainerman 1992] M. Ben-Artzi and S. Klainerman, "Decay and regularity for the Schrödinger equation", *J. Anal. Math.* **58** (1992), 25–37. MR Zbl

[Bony and Häfner 2012] J.-F. Bony and D. Häfner, "Local energy decay for several evolution equations on asymptotically Euclidean manifolds", *Ann. Sci. Éc. Norm. Supér.* (4) **45**:2 (2012), 311–335. MR Zbl

[Bortot and Cavalcanti 2014] C. A. Bortot and M. M. Cavalcanti, "Asymptotic stability for the damped Schrödinger equation on noncompact Riemannian manifolds and exterior domains", *Comm. Partial Differential Equations* **39**:9 (2014), 1791–1820. MR Zbl

[Bouclet 2011] J.-M. Bouclet, “Low frequency estimates and local energy decay for asymptotically Euclidean Laplacians”, *Comm. Partial Differential Equations* **36**:7 (2011), 1239–1286. MR Zbl

[Bouclet and Royer 2014] J.-M. Bouclet and J. Royer, “Local energy decay for the damped wave equation”, *J. Funct. Anal.* **266**:7 (2014), 4538–4615. MR Zbl

[Bouclet and Royer 2015] J.-M. Bouclet and J. Royer, “Sharp low frequency resolvent estimates on asymptotically conical manifolds”, *Comm. Math. Phys.* **335**:2 (2015), 809–850. MR Zbl

[Burq 1998] N. Burq, “Décroissance de l’énergie locale de l’équation des ondes pour le problème extérieur et absence de résonance au voisinage du réel”, *Acta Math.* **180**:1 (1998), 1–29. MR Zbl

[Burq 2004] N. Burq, “Smoothing effect for Schrödinger boundary value problems”, *Duke Math. J.* **123**:2 (2004), 403–427. MR Zbl

[Burq et al. 2004] N. Burq, P. Gérard, and N. Tzvetkov, “On nonlinear Schrödinger equations in exterior domains”, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **21**:3 (2004), 295–318. MR Zbl

[Constantin and Saut 1988] P. Constantin and J.-C. Saut, “Local smoothing properties of dispersive equations”, *J. Amer. Math. Soc.* **1**:2 (1988), 413–439. MR Zbl

[D’Ancona and Racke 2012] P. D’Ancona and R. Racke, “Evolution equations on non-flat waveguides”, *Arch. Ration. Mech. Anal.* **206**:1 (2012), 81–110. MR Zbl

[Davies 1995] E. B. Davies, “The functional calculus”, *J. London Math. Soc.* (2) **52**:1 (1995), 166–176. MR Zbl

[Dimassi and Sjöstrand 1999] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semi-classical limit*, London Mathematical Society Lecture Note Series **268**, Cambridge University Press, 1999. MR Zbl

[Doi 1996] S.-i. Doi, “Smoothing effects of Schrödinger evolution groups on Riemannian manifolds”, *Duke Math. J.* **82**:3 (1996), 679–706. MR Zbl

[Doi 2000] S.-i. Doi, “Smoothing effects for Schrödinger evolution equation and global behavior of geodesic flow”, *Math. Ann.* **318**:2 (2000), 355–389. MR Zbl

[Erdoğan et al. 2009] M. B. Erdoğan, M. Goldberg, and W. Schlag, “Strichartz and smoothing estimates for Schrödinger operators with almost critical magnetic potentials in three and higher dimensions”, *Forum Math.* **21**:4 (2009), 687–722. MR Zbl

[Guillarmou et al. 2013] C. Guillarmou, A. Hassell, and A. Sikora, “Resolvent at low energy, III: The spectral measure”, *Trans. Amer. Math. Soc.* **365**:11 (2013), 6103–6148. MR Zbl

[Kato 1966] T. Kato, “Wave operators and similarity for some non-selfadjoint operators”, *Math. Ann.* **162** (1966), 258–279. MR Zbl

[Koch and Tataru 2006] H. Koch and D. Tataru, “Carleman estimates and absence of embedded eigenvalues”, *Comm. Math. Phys.* **267**:2 (2006), 419–449. MR Zbl

[Lax et al. 1963] P. D. Lax, C. S. Morawetz, and R. S. Phillips, “Exponential decay of solutions of the wave equation in the exterior of a star-shaped obstacle”, *Comm. Pure Appl. Math.* **16** (1963), 477–486. MR Zbl

[Morawetz et al. 1977] C. S. Morawetz, J. V. Ralston, and W. A. Strauss, “Decay of solutions of the wave equation outside nontrapping obstacles”, *Comm. Pure Appl. Math.* **30**:4 (1977), 447–508. MR Zbl

[Mourre 1981] E. Mourre, “Absence of singular continuous spectrum for certain selfadjoint operators”, *Comm. Math. Phys.* **78**:3 (1981), 391–408. MR Zbl

[Ralston 1969] J. V. Ralston, “Solutions of the wave equation with localized energy”, *Comm. Pure Appl. Math.* **22** (1969), 807–823. MR Zbl

[Rauch 1978] J. Rauch, “Local decay of scattering solutions to Schrödinger’s equation”, *Comm. Math. Phys.* **61**:2 (1978), 149–168. MR Zbl

[Rauch and Taylor 1974] J. Rauch and M. Taylor, “Exponential decay of solutions to hyperbolic equations in bounded domains”, *Indiana Univ. Math. J.* **24** (1974), 79–86. MR Zbl

[Reed and Simon 1978] M. Reed and B. Simon, *Methods of modern mathematical physics, IV: Analysis of operators*, Academic Press, New York, 1978. MR Zbl

[Royer 2010] J. Royer, "Limiting absorption principle for the dissipative Helmholtz equation", *Comm. Partial Differential Equations* **35**:8 (2010), 1458–1489. MR Zbl

[Royer 2015] J. Royer, "Exponential decay for the Schrödinger equation on a dissipative waveguide", *Ann. Henri Poincaré* **16**:8 (2015), 1807–1836. MR Zbl

[Royer 2016] J. Royer, "Mourre's commutators method for a dissipative form perturbation", *J. Operator Theory* **76**:2 (2016), 351–385. MR Zbl

[Sjölin 1987] P. Sjölin, "Regularity of solutions to the Schrödinger equation", *Duke Math. J.* **55**:3 (1987), 699–715. MR Zbl

[Sz.-Nagy et al. 2010] B. Sz.-Nagy, C. Foias, H. Bercovici, and L. Kérchy, *Harmonic analysis of operators on Hilbert space*, 2nd ed., Springer, 2010. MR Zbl

[Tataru 2013] D. Tataru, "Local decay of waves on asymptotically flat stationary space-times", *Amer. J. Math.* **135**:2 (2013), 361–401. MR Zbl

[Thomann 2010] L. Thomann, "A remark on the Schrödinger smoothing effect", *Asymptot. Anal.* **69**:1-2 (2010), 117–123. MR Zbl

[Tsutsumi 1984] Y. Tsutsumi, "Local energy decay of solutions to the free Schrödinger equation in exterior domains", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **31**:1 (1984), 97–108. MR Zbl

[Zworski 2012] M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics **138**, American Mathematical Society, Providence, RI, 2012. MR Zbl

Received 3 Jun 2015. Revised 14 Mar 2017. Accepted 24 Apr 2017.

MOEZ KHENISSI: moez.khenissi@fsg.rnu.tn
LAMMDA, École Supérieure des Sciences et de Technologie de Hammam Sousse, Université de Sousse, Rue Lamine El Abbessi, 4011 Hammam Sousse, Tunisia

JULIEN ROYER: julien.royer@math.univ-toulouse.fr
Institut de Mathématiques de Toulouse, 118, route de Narbonne, 31062 Toulouse Cédex 9, France

ANALYSIS AND PDE
Vol. 10, No. 6, 2017
dx.doi.org/10.2140/apde.2017.10.1317

A CLASS OF UNSTABLE FREE BOUNDARY PROBLEMS

SERENA DIPIERRO, ARAM KARAKHANYAN AND ENRICO VALDINOCI

We consider the free boundary problem arising from an energy functional which is the sum of a Dirichlet energy and a nonlinear function of either the classical or the fractional perimeter.

The main difference with the existing literature is that the total energy is here a nonlinear superposition of the either local or nonlocal surface tension effect with the elastic energy.

In sharp contrast with the linear case, the problem considered in this paper is unstable; namely a minimizer in a given domain is not necessarily a minimizer in a smaller domain.

We provide an explicit example for this instability. We also give a free boundary condition, which emphasizes the role played by the domain in the geometry of the free boundary. In addition, we provide density estimates for the free boundary and regularity results for the minimal solution.

As far as we know, this is the first case in which a nonlinear function of the perimeter is studied in this type of problem. Also, the results obtained in this nonlinear setting are new even in the case of the local perimeter, and indeed the instability feature is not a consequence of the possible nonlocality of the problem, but it is due to the nonlinear character of the energy functional.

1. Introduction

In this paper we consider a free boundary problem given by the superposition of a Dirichlet energy and an either classical or nonlocal perimeter functional. Differently from the existing literature, here we take into account the possibility that this energy superposition occurs in a nonlinear way; that is, the total energy functional is the sum of the Dirichlet energy plus a nonlinear function of the either local or nonlocal perimeter of the interface.

Unlike the cases already present in the literature, the nonlinear problem that we study may present a structural instability induced by the domain; namely a minimizer in a large domain may fail to be a minimizer in a small domain. This fact prevents the use of scaling arguments, which are frequently exploited in classical free boundary problems.

In this paper, after providing an explicit example of this type of structural instability, we describe the free boundary equation, which also underlines the striking role played by the total (either local or nonlocal) perimeter of the minimizing set in the domain, as modulated by the nonlinearity, in the local geometry of the interface. Then, we will present results concerning the Hölder regularity of the minimal solutions and the density of the interfaces in the one-phase problem.

MSC2010: 35R35.
Keywords: free boundary problems, regularity, nonlinear phenomena.

The mathematical setting in which we work is the following. Given an (open, Lipschitz and bounded) domain $\Omega \subset \mathbb{R}^n$ and $\sigma \in (0, 1]$, we use the notation $\operatorname{Per}_\sigma(E, \Omega)$ for the classical perimeter of E in Ω when $\sigma = 1$ (which will be often denoted as $\operatorname{Per}(E, \Omega)$, see, e.g., [Ambrosio et al. 2000; Maggi 2012]) and the fractional perimeter of E in Ω when $\sigma \in (0, 1)$ (see [Caffarelli et al. 2010]). More explicitly, if $\sigma \in (0, 1)$, we have

$$\operatorname{Per}_\sigma(E, \Omega) := L(E \cap \Omega, E^c) + L(E^c \cap \Omega, E \cap \Omega^c), \tag{1-1}$$

where, for any measurable subsets $A, B \subseteq \mathbb{R}^n$ with $A \cap B$ of measure zero, we set

$$L(A, B) := \iint_{A \times B} \frac{dx\,dy}{|x - y|^{n+\sigma}}.$$

As customary, we are using here the superscript c for complementary set; i.e., $E^c := \mathbb{R}^n \setminus E$.

The notation used for $\operatorname{Per}_\sigma$ when $\sigma = 1$ is inspired by the fact that $\operatorname{Per}_\sigma$, suitably rescaled, approaches the classical perimeter as $s \nearrow 1$; see, e.g., [Bourgain et al. 2001; Dávila 2002; Caffarelli and Valdinoci 2011; Ambrosio et al. 2011].

In our framework, the role played by the fractional perimeter is to allow long-range interaction to contribute to the energy arising from surface tension and phase segregation.

As a matter of fact, the fractional perimeter $\operatorname{Per}_\sigma$ naturally arises when one considers phase transition models with long-range particle interactions (see, e.g., [Savin and Valdinoci 2014]): roughly speaking, in this type of model, the remote interactions of the particles are sufficiently strong to persist even at a large scale, by possibly modifying the behavior of the phase separation.

The fractional perimeter $\operatorname{Per}_\sigma$ has also natural applications in motion by nonlocal mean curvatures, which in turn arises naturally in the study of cellular automata and in image digitization procedures (see, e.g., [Imbert 2009]).

It is also convenient[1] to fix $\Upsilon \in \left(0, \frac{1}{100}\right]$ and set

$$\Omega_\Upsilon := \bigcup_{p \in \Omega} B_\Upsilon(p) \quad \text{and} \quad \operatorname{Per}^\star_\sigma(E, \Omega) = \begin{cases} \operatorname{Per}(E, \Omega_\Upsilon) & \text{if } \sigma = 1, \\ \operatorname{Per}_\sigma(E, \Omega) & \text{if } \sigma \in (0, 1). \end{cases} \tag{1-2}$$

We consider a monotone nondecreasing and lower semicontinuous function $\Phi : [0, +\infty) \to [0, +\infty)$, with

$$\lim_{t \to +\infty} \Phi(t) = +\infty. \tag{1-3}$$

[1]The explicit value of Υ plays no major role, since it can be fixed by an "initial scaling" of the problem, but we decided to require it to be less than $\frac{1}{100}$ to emphasize, from the psychological point of view, that Ω_Υ can be thought as a small enlargement of Ω.

The reason we introduced such an Υ is that, in the classical case, the interfaces inside Ω do not see the contributions that may come along $\partial\Omega$, since Ω is taken to be open (conversely, in the nonlocal case, these contributions are always counted). By enlarging the domain Ω by a small quantity Υ, we are able to count also the contributions on $\partial\Omega$ and this, roughly speaking, boils down to computing the classical perimeter in the closure of Ω.

For any measurable function $u : \mathbb{R}^n \to \mathbb{R}$ such that $|\nabla u| \in L^2(\Omega)$ and any measurable subset $E \subseteq \mathbb{R}^n$ such that $u \geqslant 0$ a.e. in E and $u \leqslant 0$ a.e. in E^c, we consider the energy functional

$$\mathcal{E}_\Omega(u, E) := \int_\Omega |\nabla u(x)|^2\, dx + \Phi(\mathrm{Per}^\star_\sigma(E, \Omega)). \tag{1-4}$$

As usual, the notation ∇u stands for the distributional gradient.

When Φ is the identity, the functional in (1-4) provides a typical problem for (either local or nonlocal) free boundary problems; see [Athanasopoulos et al. 2001; Caffarelli et al. 2015].

The goal of this paper is to study the minimizers of the functional in (1-4). For this, we say that (u, E) is an admissible pair if

- $u : \mathbb{R}^n \to \mathbb{R}$ is a measurable function such that $u \in H^1(\Omega)$,
- $E \subseteq \mathbb{R}^n$ is a measurable set with $\mathrm{Per}^\star_\sigma(E, \Omega) < +\infty$, and
- $u \geqslant 0$ a.e. in E and $u \leqslant 0$ a.e. in E^c.

Then, we say that (u, E) is a minimal pair in Ω if

- (u, E) is an admissible pair,
- $\mathcal{E}_\Omega(u, E) < +\infty$, and
- for any admissible pair (v, F) such that $v - u \in H^1_0(\Omega)$ and $F \setminus \Omega = E \setminus \Omega$ up to sets of measure zero, we have
$$\mathcal{E}_\Omega(u, E) \leqslant \mathcal{E}_\Omega(v, F).$$

The existence[2] of minimal pairs for fixed domains and fixed conditions outside the domain follows from the direct methods in the calculus of variations (see Lemma 2.3 below for details).

A natural question in this framework is whether or not this minimization procedure is "stable" with respect to the choice of the domain, i.e., whether or not a minimal pair in a domain Ω is also a minimal pair in any subdomain $\Omega' \subset \Omega$. This stability property is indeed typical for "linear" free boundary problems, i.e., when Φ is the identity, see [Athanasopoulos et al. 2001; Caffarelli et al. 2015], and it often plays a crucial role in many arguments based on scaling and blow-up analysis.

In the "nonlinear" case, i.e., when Φ is not the identity, this stability property is lost, and we will provide a concrete example for that. In further detail, we consider the planar case of $\mathbb{R}^2$, we take coordinates $X := (x, y) \in \mathbb{R}^2$ and we set

$$\tilde{u}(x, y) := xy \tag{1-5}$$

and

$$\begin{aligned} \widetilde{E} &:= \{(x, y) \in \mathbb{R}^2 : xy > 0\} \\ &= \{(x, y) \in \mathbb{R}^2 : x > 0 \text{ and } y > 0\} \cup \{(x, y) \in \mathbb{R}^2 : x < 0 \text{ and } y < 0\}. \end{aligned} \tag{1-6}$$

[2] As a technical remark, we point out that the definition in (1-2) is useful to make sense of nontrivial versions of this minimization problem when $\sigma = 1$ and $u \geqslant 0$. Indeed, in this case, the setting in (1-2) "forces" the sets to interact with the boundary data. This expedient is not necessary when $\sigma = 0$ since, in this case, the nonlocal effect produces the nontrivial interactions.

In this setting, we show that:

Theorem 1.1 (an explicit counterexample). *There exists $K_o > 2$ such that the following statement is true. Let $n = 2$. Assume*
$$\Phi(t) = t^\gamma \quad \text{for any } t \in [0, 1]$$
for some
$$\gamma \in \Big(0, \frac{4}{2-\sigma}\Big),$$
and
$$\Phi(t) = 1 \quad \text{for any } t \in [2, K_o]. \tag{1-7}$$
Then, there exist $R_o > r_o > 0$ such that $(\tilde{u}, \widetilde{E})$ is a minimal pair in B_{R_o} and is not a minimal pair in B_r for any $r \in (0, r_o]$.

The heuristic idea underlying Theorem 1.1 is, roughly speaking, that the nonlinear energy term Φ weights differently the fractional perimeter with respect to the Dirichlet energy in different energy regimes, so it may favor a minimal pair (u, E) to be either "close to a harmonic function" in the u or "close to a fractional minimal surface" in the E, depending on the minimal energy level reached in a given domain.

It is worth stressing that, in other circumstances, rather surprising instability features in interface problems arise as a consequence of the fractional behavior of the energy; see, for instance, [Dipierro et al. 2017]. Differently from these cases, the unstable free boundaries presented in Theorem 1.1 are not caused by the existence of possibly nonlocal features, and indeed Theorem 1.1 holds true (and is new) even in the case of the local perimeter.

The instability phenomenon pointed out by Theorem 1.1 in a concrete case is also quite general, as it can be understood also in the light of the associated equation on the free boundary. Indeed, the free boundary equation takes into account a "global" term of the type $\Phi'(\mathrm{Per}^\star_\sigma(E, \Omega))$, which varies in dependence of the domain Ω. To clarify this point, we denote by H^E_σ the (either classical or fractional) mean curvature of ∂E (see [Caffarelli et al. 2010; Abatangelo and Valdinoci 2014] for the case $\sigma \in (0, 1)$). Namely, if $\sigma = 1$ the above notation stands for the classical mean curvature, while for $\sigma \in (0, 1)$, if $x \in \partial E$, we set
$$H^E_\sigma(x) := \limsup_{\delta \to 0} \int_{\mathbb{R}^n \setminus B_\delta(x)} \frac{\chi_{E^c}(y) - \chi_E(y)}{|x-y|^{n+\sigma}}\, dy.$$
In this setting, we have:

Theorem 1.2 (free boundary equation). *Let $\Phi \in C^{1,\alpha}(0, +\infty)$ for some $\alpha \in (0, 1)$. Assume (u, E) is a minimal pair in Ω. Assume*
$$(\partial E) \cap \Omega \text{ is of class } C^{1,\tau} \text{ with } \tau \in (\sigma, 1) \text{ when } \sigma \in (0, 1) \text{ and of class } C^2 \text{ when } \sigma = 1. \tag{1-8}$$
Suppose also
$$u > 0 \text{ in the interior of } E \cap \Omega, \quad u < 0 \text{ in the interior of } E^c \cap \Omega, \tag{1-9}$$
and
$$u \in C^1(\overline{\{u > 0\} \cap \Omega}) \cap C^1(\overline{\{u < 0\} \cap \Omega}). \tag{1-10}$$

Let also ν be the exterior normal of E, and for any $x \in (\partial E) \cap \Omega$ let

$$\partial^+_\nu u(x) := \lim_{t\to 0} \frac{u(x-t\nu)-u(x)}{t} \quad \text{and} \quad \partial^-_\nu u(x) := \lim_{t\to 0} \frac{u(x+t\nu)-u(x)}{t}. \tag{1-11}$$

Then, for any $x \in (\partial E) \cap \Omega$, we have

$$(\partial^+_\nu u(x))^2 - (\partial^-_\nu u(x))^2 = H^E_\sigma(x)\,\Phi'(\mathrm{Per}^\star_\sigma(E,\Omega)). \tag{1-12}$$

We remark that (1-12) has a simple geometric consequence when $\Phi' > 0$ and we consider the one-phase problem in which $u \geqslant 0$: indeed, in this case, we have $\partial^-_\nu u = 0$ and therefore formula (1-12) reduces to

$$(\partial^+_\nu u(x))^2 = H^E_\sigma(x)\,\Phi'(\mathrm{Per}^\star_\sigma(E,\Omega)).$$

In particular, we get that $H^E_\sigma \geqslant 0$; namely, in this case, the (either classical or fractional) mean curvature of the free boundary is nonnegative.

In order to better understand the structure of the solution and of the free boundary points, we now focus, for the sake of simplicity, on the one-phase case; i.e., we suppose that $u \geqslant 0$ to start with. In this setting, we investigate the Hölder regularity of the function u by obtaining uniform bounds and uniform growth conditions from the free boundary. For this, it is also convenient to introduce the auxiliary set

$$\mathcal{U}_0 := \big\{x \in \Omega : \text{there exists a sequence } x_k \in \Omega : x_k \to x \text{ with } u(x_k) \to 0 \text{ as } k \to +\infty\big\}. \tag{1-13}$$

Notice that $\{u = 0\}$ lies in $\mathcal{U}_0$ (just taking a constant sequence in the definition above). Also, if $u \geqslant 0$, then ∂E lies in $\mathcal{U}_0$ (since in this case u must vanish in the complement of E).

Of course, when u is continuous, such a set lies in the zero level set of u, but since we do not have this information a priori, it is useful to consider explicitly this set, and prove the following result:

Theorem 1.3 (growth from the free boundary). *Let R_o, $Q > 0$. Assume*

$$\Phi \textit{ is Lipschitz continuous in } [0, Q], \textit{ with Lipschitz constant bounded by } L_Q. \tag{1-14}$$

Assume (u, E) is a minimal pair in Ω, with $B_{R_o} \Subset \Omega$,

$$0 \in \mathcal{U}_0 \tag{1-15}$$

and $u \geqslant 0$ in $\mathbb{R}^n \setminus \Omega$. Suppose $R \in (0, R_o]$ and

$$\mathrm{Per}^\star_\sigma(E,\Omega) + R^{n-\sigma}\,\mathrm{Per}_\sigma(B_1,\mathbb{R}^n) \leqslant Q. \tag{1-16}$$

Then, there exists $C > 0$, possibly depending on R_o, n and σ such that, for any $x \in B_{R/2}$,

$$u(x) \leqslant C\sqrt{L_Q}\,|x|^{1-\sigma/2}.$$

We observe that condition (1-14) is always satisfied if Φ is globally Lipschitz, but the statement of Theorem 1.3 is more general, since it may take into account a locally Lipschitz Φ, provided that the domain is small enough to satisfy (1-16) (indeed, small domains satisfy this condition for locally Lipschitz Φ, as remarked in the forthcoming Lemma 2.8).

We also point out that (1-16) may be equivalently written

$$\mathrm{Per}^{\star}_{\sigma}(E,\Omega)+\mathrm{Per}_{\sigma}(B_R,\mathbb{R}^n)\leqslant Q. \tag{1-17}$$

One natural way to interpret (1-16), or (1-17), is that once $\mathrm{Per}^{\star}_{\sigma}(E,\Omega)$ is strictly less than Q (i.e., strictly less than the size of the interval in which Φ is Lipschitz), then (1-16), and thus (1-17), holds true as long as R is sufficiently small.

The growth result in Theorem 1.3 implies, as a byproduct, an interior Hölder regularity result:

Corollary 1.4. *Let $Q>0$ and assume Φ is Lipschitz continuous in $[0,Q]$, with Lipschitz constant bounded by L_Q.*

Assume (u,E) is a minimal pair in Ω, with $B_R\Subset\Omega$ and $u\geqslant 0$ in $\mathbb{R}^n\setminus\Omega$.

Suppose that $\mathrm{Per}^{\star}_{\sigma}(E,\Omega)+R^{n-\sigma}\,\mathrm{Per}_{\sigma}(B_1,\mathbb{R}^n)\leqslant Q$ and that $u\leqslant M$ on $\partial\Omega$.

Then $u\in C^{1-\sigma/2}(B_{R/4})$, with

$$\|u\|_{C^{1-\sigma/2}(B_{R/4})}\leqslant C\left(\sqrt{L_Q}+\frac{M}{R^{1-\sigma/2}}\right),$$

for some $C>0$, possibly depending on n and σ.

When Φ is linear, the result in Corollary 1.4 was obtained in Theorem 3.1 of [Athanasopoulos et al. 2001] if $\sigma=1$ and in Theorem 1.1 of [Caffarelli et al. 2015] if $\sigma\in(0,1)$. Differently than in our framework, in both papers mentioned above, scaling arguments are available, since scaling is compatible with the minimization procedure.

Now we investigate the structure of the free boundary points in terms of local densities of the phases. Indeed, we show that the free boundary points always have uniform density from outside E, according to the following result:

Theorem 1.5 (density estimate from the null side). *Assume (u,E) is a minimal pair in Ω, with $B_R\subseteq\Omega$, $0\in\partial E$ and $u\geqslant 0$ in $\mathbb{R}^n\setminus\Omega$. Set*

$$P=P(E,\Omega,R):=\mathrm{Per}^{\star}_{\sigma}(E,\Omega)+R^{n-\sigma}\,\mathrm{Per}_{\sigma}(B_1,\mathbb{R}^n) \tag{1-18}$$

and assume

$$\Phi \text{ is strictly increasing in the interval } (0,P). \tag{1-19}$$

Then there exists $\delta>0$, possibly depending on n and σ such that, for any $r\in\left(0,\tfrac12 R\right)$,

$$|B_r\setminus E|\geqslant\delta r^n.$$

We point out that condition (1-19) is always satisfied if Φ is strictly increasing in the whole of $[0,+\infty)$, but Theorem 1.5 is also general enough to take into consideration the case in which Φ is strictly increasing only in a subinterval, provided that the energy domain is sufficiently small to make the perimeter values lie in the strict monotonicity interval of Φ (as a matter of fact, the perimeter contributions in small domains are small, as we will point out in the forthcoming Lemma 2.8).

The investigation of the density properties of the free boundary is also completed by the following counterpart of Theorem 1.5, which proves the positive density of the set E:

Theorem 1.6 (density estimate from the positive side). *Let $Q > 0$ and assume*

$$\Phi \text{ is Lipschitz continuous in } [0, Q], \text{ with Lipschitz constant bounded by } L_Q, \tag{1-20}$$

and

$$\Phi' \geqslant c_o \quad \text{a.e. in } [0, Q] \tag{1-21}$$

for some $c_o > 0$.

Assume (u, E) is a minimal pair in Ω, with $B_R \Subset \Omega$, $0 \in \partial E$ and $u \geqslant 0$ in $\mathbb{R}^n \setminus \Omega$. Suppose

$$\operatorname{Per}^\star_\sigma(E, \Omega) + R^{n-\sigma} \operatorname{Per}_\sigma(B_1, \mathbb{R}^n) \leqslant Q. \tag{1-22}$$

Then there exists $\delta_* > 0$, possibly depending on n, σ, c_o and L_Q, such that, for any $r \in \left(0, \frac{1}{2}R\right)$,

$$|B_r \cap E| \geqslant \delta_* r^n.$$

More explicitly, such δ_* can be taken to be of the form

$$\delta_* := \delta_o \min\left\{1, \left(\frac{c_o}{L_Q}\right)^{n/\sigma}\right\} \tag{1-23}$$

for some $\delta_o > 0$, possibly depending on n and σ.

We remark that the results obtained in this paper are new even in the local case in which $\sigma = 1$. Also, we think it is an interesting point of this paper that all the cases $\sigma \in (0, 1)$ and $\sigma = 1$ are treated simultaneously in a unified fashion. The methods presented are also general enough to treat the case $\sigma = 0$, which would correspond to a volume term (see, e.g., [Maz'ya and Shaposhnikova 2002; Dipierro et al. 2013]). This case is in fact rich in results and so we will discuss it in detail in a forthcoming paper.

The rest of the paper is organized as follows. In Section 2 we show some preliminary properties of the minimal pair, such as existence, harmonicity and subharmonicity properties, and a comparison principle. We also prove a "locality" property for the (either classical or fractional) perimeter and provide a uniform bound on the (classical or fractional) perimeter of the set in the minimal pair.

Section 3 is devoted to the construction of the counterexample in Theorem 1.1. In Section 4 we provide the free boundary equation and prove Theorem 1.2.

Then we deal with the regularity of the function u in the minimal pair in the one-phase case, and we prove Theorem 1.3 and Corollary 1.4 in Sections 5 and 6, respectively. Finally, Sections 7 and 8 are devoted to the proofs of the density estimates from both sides provided by Theorems 1.5 and 1.6, respectively.

Since we hope that the paper may be of interest for different communities (such as scientists working in free boundary problems, variational methods, partial differential equations, geometric measure theory and fractional problems), we made an effort to give the details of the arguments involved in the proofs in a clear and widely accessible way.

2. Preliminaries

We start with a useful observation about the positivity sets of sequences of admissible pairs:

Lemma 2.1. *Let (u_j, E_j) be a sequence of admissible pairs. Assume $u_j \to u$ a.e. in $\mathbb{R}^n$ and $\chi_{E_j} \to \chi_E$ a.e. in $\mathbb{R}^n$ for some u and E. Then $u \geqslant 0$ a.e. in E and $u \leqslant 0$ a.e. in E^c.*

Proof. We show that $u \geqslant 0$ a.e. in E (the other claim being analogous). For this, we write $\mathbb{R}^n = X \cup Z$, with $|Z| = 0$ and such that for any $x \in X$ we have

$$\lim_{j\to+\infty} u_j(x) = u(x) \quad \text{and} \quad \lim_{j\to+\infty} \chi_{E_j}(x) = \chi_E(x).$$

Let now $x \in E \cap X$. Then

$$\lim_{j\to+\infty} \chi_{E_j}(x) = \chi_E(x) = 1,$$

and so there exists $j_x \in \mathbb{N}$ such that $\chi_{E_j}(x) \geqslant \frac{1}{2}$ for any $j \geqslant j_x$. Since the image of a characteristic function lies in $\{0, 1\}$, this implies $\chi_{E_j}(x) = 1$ for any $j \geqslant j_x$, and therefore $u_j(x) \geqslant 0$ for any $j \geqslant j_x$. Taking the limit, we obtain $u(x) \geqslant 0$. Since this is valid for any $x \in E \cap X$ and $E \cap X^c \subseteq Z$, which has null measure, we have obtained the desired result. □

Now we recall a useful auxiliary identity for the (classical or fractional) perimeter:

Lemma 2.2 ("clean cut" lemma). *Let $\Omega' \Subset \Omega$. Assume $\operatorname{Per}_\sigma(E, \Omega) < +\infty$ and $\operatorname{Per}_\sigma(F, \Omega) < +\infty$. Suppose also that*

$$E \setminus \overline{\Omega}' = F \setminus \overline{\Omega}'. \tag{2-1}$$

Then

$$\operatorname{Per}_\sigma(E, \Omega) - \operatorname{Per}_\sigma(F, \Omega) = \operatorname{Per}_\sigma(E, \overline{\Omega}') - \operatorname{Per}_\sigma(F, \overline{\Omega}'). \tag{2-2}$$

If in addition $\operatorname{Per}^\star_\sigma(E, \Omega) < +\infty$ and $\operatorname{Per}^\star_\sigma(F, \Omega) < +\infty$, then

$$\operatorname{Per}^\star_\sigma(E, \Omega) - \operatorname{Per}^\star_\sigma(F, \Omega) = \operatorname{Per}_\sigma(E, \overline{\Omega}') - \operatorname{Per}_\sigma(F, \overline{\Omega}'). \tag{2-3}$$

Proof. For completeness, we distinguish the cases $\sigma = 1$ and $\sigma \in (0, 1)$. If $\sigma = 1$, we write the perimeter of E in term of the Gauss–Green measure μ_E (see Remark 12.2 in [Maggi 2012]); namely

$$\operatorname{Per}(E, \Omega) = |\mu_E|(\Omega).$$

So we define

$$U := \Omega \setminus \overline{\Omega}'. \tag{2-4}$$

We remark that U is open and $\Omega = \overline{\Omega}' \cup U$, with disjoint union. Thus we obtain

$$\begin{aligned}
\operatorname{Per}(E, \Omega) &- \operatorname{Per}(F, \Omega) - \operatorname{Per}(E, \overline{\Omega}') + \operatorname{Per}(F, \overline{\Omega}') \\
&= |\mu_E|(\Omega) - |\mu_F|(\Omega) - |\mu_E|(\overline{\Omega}') + |\mu_F|(\overline{\Omega}') \\
&= |\mu_E|(\overline{\Omega}' \cup U) - |\mu_F|(\overline{\Omega}' \cup U) - |\mu_E|(\overline{\Omega}') + |\mu_F|(\overline{\Omega}') \\
&= |\mu_E|(\overline{\Omega}') + |\mu_E|(U) - |\mu_F|(\overline{\Omega}') - |\mu_F|(U) - |\mu_E|(\overline{\Omega}') + |\mu_F|(\overline{\Omega}') \\
&= |\mu_E|(U) - |\mu_F|(U) = \operatorname{Per}(E, U) - \operatorname{Per}(F, U).
\end{aligned} \tag{2-5}$$

Now we observe that

$$E \cap U = E \cap (\Omega \setminus \overline{\Omega}') = E \cap \Omega \cap (\overline{\Omega}')^c = (E \setminus \overline{\Omega}') \cap \Omega,$$

and a similar set identity holds for F. Thus, by (2-1), it follows that $E \cap U = F \cap U$. Therefore, by the locality of the classical perimeter (see, e.g., Proposition 3.38(c) in [Ambrosio et al. 2000]), we obtain

$$\mathrm{Per}(E, U) = \mathrm{Per}(F, U).$$

If one inserts this into (2-5), then one obtains (2-2) when $\sigma = 1$.

Now we deal with the case $\sigma \in (0, 1)$. For this we use (1-1) and (2-4) and we get

$$\begin{aligned}
\mathrm{Per}_\sigma(E, \Omega) - \mathrm{Per}_\sigma(E, \overline{\Omega}') &= L(E \cap \Omega, E^c) + L(E^c \cap \Omega, E \setminus \Omega) - L(E \cap \overline{\Omega}', E^c) - L(E^c \cap \overline{\Omega}', E \setminus \overline{\Omega}') \\
&= L(E \cap \overline{\Omega}', E^c) + L(E \cap U, E^c) + L(E^c \cap \overline{\Omega}', E \setminus \Omega) + L(E^c \cap U, E \setminus \Omega) \\
&\qquad - L(E \cap \overline{\Omega}', E^c) - L(E^c \cap \overline{\Omega}', E \setminus \Omega) - L(E^c \cap \overline{\Omega}', E \cap U) \\
&= L(E \cap U, E^c) + L(E^c \cap U, E \setminus \Omega) - L(E^c \cap \overline{\Omega}', E \cap U) \\
&= L(E \cap U, E^c \setminus \overline{\Omega}') + L(E^c \cap U, E \setminus \Omega),
\end{aligned}$$

and a similar formula holds for F replacing E. Now, from (2-1), we see that

$$E \cap U = F \cap U, \quad E^c \cap U = F^c \cap U, \quad E^c \setminus \overline{\Omega}' = F^c \setminus \overline{\Omega}' \quad \text{and} \quad E \setminus \Omega = F \setminus \Omega;$$

thus we obtain (2-2) when $\sigma \in (0, 1)$.

Now, to prove (2-3), we can focus on the case $\sigma = 1$ (since $\mathrm{Per}^\star_\sigma = \mathrm{Per}_\sigma$ when $\sigma \in (0, 1)$, in this case we return simply to (2-2)). To this end, we observe that $\Omega' \Subset \Omega_\Upsilon$ (recall formula (1-2)), so we can apply (2-2) to the sets Ω' and Ω_Υ and obtain, when $\sigma = 1$,

$$\mathrm{Per}^\star_\sigma(E, \Omega) - \mathrm{Per}^\star_\sigma(F, \Omega) = \mathrm{Per}(E, \Omega_\Upsilon) - \mathrm{Per}(F, \Omega_\Upsilon) = \mathrm{Per}(E, \overline{\Omega}') - \mathrm{Per}(F, \overline{\Omega}').$$

This completes the proof of (2-3). □

Now we state the basic existence result for the minimizers of the functional in (1-4):

Lemma 2.3 (existence of minimal pairs). *Fix an admissible pair (u_o, E_o) such that $\mathcal{E}_\Omega(u_o, E_o) < +\infty$. Then there exists a minimal pair (u, E) in Ω such that $u - u_o \in H^1_0(\Omega)$ and $E \setminus \Omega$ coincides with $E_o \setminus \Omega$ up to sets of measure zero.*

Proof. Let (u_j, E_j) be a minimizing sequence, namely

$$\lim_{j \to +\infty} \mathcal{E}_\Omega(u_j, E_j) = \inf_{X_\Omega(u_o, E_o)} \mathcal{E}_\Omega, \tag{2-6}$$

where $X_\Omega(u_o, E_o)$ denotes the family of all admissible pairs (v, F) in Ω such that $v - u_o \in H^1_0(\Omega)$ and $F \setminus \Omega$ coincides with $E_o \setminus \Omega$ up to sets of measure zero.

We stress that

$$\sup_{j \in \mathbb{N}} \Phi(\mathrm{Per}^\star_\sigma(E_j, \Omega)) < +\infty,$$

thanks to (2-6). By this and (1-3), we obtain

$$\sup_{j\in\mathbb{N}} \mathrm{Per}_\sigma(E_j,\Omega) < +\infty.$$

Using this and (2-6), by compactness (see, e.g., Corollary 3.49 in [Ambrosio et al. 2000] for the case $\sigma = 1$ or Theorem 7.1 in [Di Nezza et al. 2012] for the case $\sigma \in (0, 1)$), we obtain that, up to subsequences, u_j converges to some u weakly in $H^1(\Omega)$ and strongly in $L^2(\Omega)$, and χ_{E_j} converges to some χ_E strongly in $L^1(\Omega)$ as $j \to +\infty$. By Lemma 2.1, we have that (u, E) is an admissible pair, and so by construction

$$(u, E) \in X_\Omega(u_o, E_o). \tag{2-7}$$

Also, by the lower semicontinuity (or Fatou's lemma; see, e.g., Proposition 3.38(b) in [Ambrosio et al. 2000] for the case $\sigma = 1$) we have

$$\liminf_{j\to+\infty}\int_\Omega |\nabla u_j(x)|^2\,dx \geqslant \int_\Omega |\nabla u(x)|^2\,dx \quad \text{and} \quad \liminf_{j\to+\infty} \mathrm{Per}^\star_\sigma(E_j,\Omega) \geqslant \mathrm{Per}^\star_\sigma(E,\Omega),$$

and so, using also the monotonicity and the lower semicontinuity of Φ,

$$\liminf_{j\to+\infty} \Phi(\mathrm{Per}^\star_\sigma(E_j,\Omega)) \geqslant \Phi\big(\liminf_{j\to+\infty} \mathrm{Per}^\star_\sigma(E_j,\Omega)\big) \geqslant \Phi(\mathrm{Per}^\star_\sigma(E,\Omega)).$$

These inequalities and (2-6) give that

$$\mathcal{E}_\Omega(u, E) \leqslant \inf_{X_\Omega(u_o,E_o)} \mathcal{E}_\Omega,$$

and then equality holds in the formula above, thanks to (2-7). □

As it often happens in free boundary problems (see, e.g., [Alt and Caffarelli 1981; Athanasopoulos et al. 2001; Caffarelli et al. 2015]), the solutions are harmonic in the positivity or negativity sets. This happens also in our case, as clarified by the following observation:

Lemma 2.4. *Let (u, E) be a minimal pair in Ω. Let U be an open set. Assume that either* $\inf_U u > 0$ *or* $\sup_U u < 0$. *Then u is harmonic in U.*

Proof. The proof is standard, but we give the details to assist the reader. We suppose

$$\inf_U u > 0, \tag{2-8}$$

the other case being similar. Let $x_o \in U$. Since U is open, there exists $r > 0$ such that $B_r(x_o) \subset U$. Let $\psi \in C^\infty_0(B_{r/2}(x_o))$. Let also $u_\epsilon := u + \epsilon\psi$ and

$$m := \inf_{B_{r/2}(x_o)} u.$$

By (2-8), we know $m > 0$. Thus, if $\epsilon \in \mathbb{R}$, with $|\epsilon| < (1+\|\psi\|_{L^\infty(\mathbb{R}^n)})^{-1} m$, we have $u_\epsilon \geqslant u - \epsilon\|\psi\|_{L^\infty(\mathbb{R}^n)} \geqslant 0$ in $B_{r/2}(x_o)$. This and the fact that ψ vanishes outside $B_{r/2}(x_o)$ give that (u_ϵ, E) is an admissible pair. Thus, the minimality of (u, E) gives

$$0 \leqslant \mathcal{E}_\Omega(u_\epsilon, E) - \mathcal{E}_\Omega(u, E) = \int_\Omega \big(|\nabla u(x) + \epsilon\nabla\psi(x)|^2 - |\nabla u(x)|^2\big)\,dx,$$

from which the desired result easily follows. □

As often happens in free boundary problems, the minimizers satisfy the following subharmonicity property:

Lemma 2.5. *Let (u, E) be a minimal pair in Ω and $u^+ := \max\{u, 0\}$ and $u^- := u^+ - u = -\min\{u, 0\}$. Then both u^+ and u^- are subharmonic in Ω in the sense that*

$$\int_\Omega \nabla u^\pm(x) \cdot \nabla \psi(x)\, dx \leqslant 0$$

for any $\psi \in H^1_0(\Omega)$, with $\psi \geqslant 0$ a.e. in Ω.

Proof. The proof is a modification of the one in Lemma 2.7 in [Athanasopoulos et al. 2001], where this result was proved for the case in which Φ is the identity and $\sigma = 1$. We give the details to assist the reader. We argue for u^+, since a similar reasoning works for u^-. We define $v^\star$ to be the harmonic replacement of u^+ in Ω which vanishes in E^c, that is, the minimizer of the Dirichlet energy in Ω among all the functions v in $H^1(\Omega)$ such that $v - u^+ \in H^1_0(\Omega)$ and $v = 0$ a.e. in E^c. For the existence and the uniqueness of the harmonic replacement, see, e.g., Section 2 in [Athanasopoulos et al. 2001] or Lemma 2.1 in [Dipierro and Valdinoci 2015]. In particular, the uniqueness result gives that

$$\text{if } v \text{ in } H^1(\Omega) \text{ is such that } v - u^+ \in H^1_0(\Omega),\ v = 0 \text{ a.e. in } E^c \\ \text{and } \int_\Omega |\nabla v(x)|^2\, dx \leqslant \int_\Omega |\nabla v^\star(x)|^2\, dx, \text{ then } v = v^\star \text{ a.e. in } \mathbb{R}^n. \tag{2-9}$$

Moreover, by Lemma 2.3 in [Athanasopoulos et al. 2001], we have

$$v^\star \text{ is subharmonic.} \tag{2-10}$$

We also notice that $v^\star \geqslant 0$ by the classical maximum principle and therefore $(v^\star, E)$ is an admissible pair. Then, the minimality of (u, E) implies

$$0 \geqslant \mathcal{E}_\Omega(u, E) - \mathcal{E}_\Omega(v^\star, E) = \int_\Omega |\nabla u(x)|^2\, dx - \int_\Omega |\nabla v^\star(x)|^2\, dx \geqslant \int_\Omega |\nabla u^+(x)|^2\, dx - \int_\Omega |\nabla v^\star(x)|^2\, dx.$$

This implies that u^+ coincides with $v^\star$, thanks to (2-9), and so it is subharmonic, in light of (2-10). □

Remark 2.6. In light of Lemma 2.5, we have (see, e.g., Proposition 2.2 in [Giaquinta 1983]) that the map

$$R \to \frac{1}{|B_R|} \int_{B_R(p)} u^+(x)\, dx$$

is monotone nondecreasing; therefore, up to changing u^+ in a set of measure zero, we can (and implicitly do from now on) suppose

$$u(p) = \lim_{\epsilon \searrow 0} \frac{1}{|B_\epsilon|} \int_{B_\epsilon(p)} u^+(x)\, dx.$$

Another simple and interesting property of the solution is given by the following maximum principle:

Lemma 2.7. *Assume*

$$\Phi(0) < \Phi(t) \quad \textit{for any } t > 0. \tag{2-11}$$

Let (u, E) be a minimal pair in Ω and let $a \in \mathbb{R}$. If $u \leqslant a$ in Ω^c, then $u \leqslant a$ in the whole of $\mathbb{R}^n$.

Similarly, if $u \geqslant a$ in Ω^c, then $u \geqslant a$ in the whole of $\mathbb{R}^n$.

Proof. We suppose
$$u \geqslant a \quad \text{in } \Omega^c, \tag{2-12}$$
the other case being analogous.

We need to distinguish the cases $a \leqslant 0$ and $a > 0$.

If $a \leqslant 0$, we take $u^\star := \max\{u, a\}$. Notice that $(u^\star, E)$ is an admissible pair: indeed, a.e. in E we have $0 \leqslant u \leqslant u^\star$, while a.e. in E^c we have $u \leqslant 0$ and so $u^\star \leqslant 0$. Also, by (2-12), we have $u \geqslant a$ in Ω^c, and so $u^\star = u$ in Ω^c. As a consequence, the minimality of (u, E) gives
$$0 \leqslant \mathcal{E}_\Omega(u^\star, E) - \mathcal{E}_\Omega(u, E) = \int_\Omega \big(|\nabla u^\star(x)|^2 - |\nabla u(x)|^2\big)\, dx = -\int_{\Omega\cap\{u<a\}} |\nabla u(x)|^2\, dx,$$
which implies $u \geqslant a$, as desired.

Now suppose $a > 0$. We take $u^\sharp$ to be the minimizer of the Dirichlet energy in Ω with trace datum u along $\partial\Omega$ (and thus we set $u^\sharp := u$ outside Ω); then we have
$$\Gamma := \int_\Omega |\nabla u(x)|^2\, dx - \int_\Omega |\nabla u^\sharp(x)|^2\, dx \geqslant 0. \tag{2-13}$$
Moreover, by (2-12) and the classical maximum principle, we know
$$u^\sharp \geqslant a \quad \text{in the whole of } \mathbb{R}^n. \tag{2-14}$$
Thus, $u^\sharp > 0$ and so $(u^\sharp, \mathbb{R}^n)$ is an admissible pair. Accordingly, the minimality of (u, E) and (2-13) give
$$\begin{aligned} 0 &\leqslant \mathcal{E}_\Omega(u^\sharp, \mathbb{R}^n) - \mathcal{E}_\Omega(u, E) \\ &= \int_\Omega |\nabla u^\sharp(x)|^2\, dx + \Phi(0) - \int_\Omega |\nabla u(x)|^2\, dx - \Phi(\mathrm{Per}^\star_\sigma(E, \Omega)) \\ &= -\Gamma + \Phi(0) - \Phi(\mathrm{Per}^\star_\sigma(E, \Omega)). \end{aligned} \tag{2-15}$$
As a consequence,
$$\Phi(\mathrm{Per}^\star_\sigma(E, \Omega)) \leqslant -\Gamma + \Phi(0) \leqslant \Phi(0);$$
hence, exploiting (2-11), we see that $\mathrm{Per}^\star_\sigma(E, \Omega) = 0$. Plugging this information into (2-15), we obtain that $0 \leqslant -\Gamma$ and thus, recalling (2-13), we conclude that $\Gamma = 0$. By the uniqueness of the minimizer of the Dirichlet energy, this implies that $u^\sharp$ coincides with u. In light of this and of (2-14), we have $u = u^\sharp \geqslant a$, as desired. □

Now we give a uniform bound on the (classical or fractional) perimeter of the sets in the minimal pairs:

Lemma 2.8. *Suppose Ω is strictly star-shaped, i.e., $t\overline{\Omega} \subseteq \Omega$ for any $t \in (0, 1)$, and that*
$$\Phi \text{ is strictly monotone.} \tag{2-16}$$
Let (u, E) be a minimal pair in Ω. Assume $u \geqslant 0$. Then, for any $\Omega' \subseteq \Omega$, with Ω' open, Lipschitz and bounded, we have
$$\mathrm{Per}_\sigma(E, \Omega') \leqslant 2\,\mathrm{Per}_\sigma(\Omega', \mathbb{R}^n). \tag{2-17}$$

In particular, if $\Omega \supseteq B_R$, *then, for any* $r \in (0, R]$,

$$\mathrm{Per}_\sigma(E, B_r) \leqslant C r^{n-\sigma} \tag{2-18}$$

for some $C > 0$ *possibly depending on n and* σ.

Proof. We observe that (2-18) follows from (2-17) by taking $\Omega' := B_r$, so we focus on the proof of (2-17). For this, first we suppose that $\Omega' \Subset \Omega$ (the general case in which $\Omega' \subseteq \Omega$ will be considered at the end of the proof, by a limit procedure). Let $F := E \cup \Omega'$. Notice that $F \setminus \overline{\Omega}' = E \cup \Omega' \cap (\overline{\Omega}')^c = E \setminus \overline{\Omega}'$. Thus, by formula (2-3) in Lemma 2.2, we get

$$\mathrm{Per}^\star_\sigma(E, \Omega) - \mathrm{Per}^\star_\sigma(F, \Omega) = \mathrm{Per}_\sigma(E, \overline{\Omega}') - \mathrm{Per}_\sigma(F, \overline{\Omega}'). \tag{2-19}$$

Now, let v be the minimizer of the Dirichlet energy in Ω' with trace datum u along $\partial\Omega'$ (then take $v := u$ outside Ω'). Since $u \geqslant 0$, so is v. Hence, the pair (v, F) is admissible. Therefore, the minimality of (u, E) implies

$$\begin{aligned} 0 &\leqslant \mathcal{E}_\Omega(v, F) - \mathcal{E}_\Omega(u, E) \\ &= \int_{\Omega'} |\nabla v(x)|^2\,dx - \int_{\Omega'} |\nabla u(x)|^2\,dx + \Phi(\mathrm{Per}^\star_\sigma(F, \Omega)) - \Phi(\mathrm{Per}^\star_\sigma(E, \Omega)) \\ &\leqslant 0 + \Phi(\mathrm{Per}^\star_\sigma(F, \Omega)) - \Phi(\mathrm{Per}^\star_\sigma(E, \Omega)). \end{aligned}$$

Hence, by (2-16), we have $\mathrm{Per}^\star_\sigma(E, \Omega) \leqslant \mathrm{Per}^\star_\sigma(F, \Omega)$ and so, by (2-19),

$$\mathrm{Per}_\sigma(E, \overline{\Omega}') - \mathrm{Per}_\sigma(F, \overline{\Omega}') = \mathrm{Per}^\star_\sigma(E, \Omega) - \mathrm{Per}^\star_\sigma(F, \Omega) \leqslant 0. \tag{2-20}$$

In addition, we have

$$\mathrm{Per}_\sigma(F, \overline{\Omega}') = \mathrm{Per}_\sigma(E \cup \Omega', \overline{\Omega}') \leqslant 2\,\mathrm{Per}_\sigma(\Omega', \mathbb{R}^n),$$

where the last formula follows using (1-1) if $\sigma \in (0, 1)$ and, for instance, formula (16.12) in [Maggi 2012] when $\sigma = 1$.

The latter inequality and (2-20) give

$$\mathrm{Per}_\sigma(E, \Omega') \leqslant \mathrm{Per}_\sigma(E, \overline{\Omega}') \leqslant \mathrm{Per}_\sigma(F, \overline{\Omega}') \leqslant 2\,\mathrm{Per}_\sigma(\Omega', \mathbb{R}^n).$$

This proves the desired result when $\Omega' \Subset \Omega$. Let us now deal with the case $\Omega' \subseteq \Omega$. For this, we set $\Omega'_\epsilon := (1-\epsilon)\Omega'$. Since Ω is strictly star-shaped, we have $\overline{\Omega}'_\epsilon = (1-\epsilon)\overline{\Omega}' \subseteq (1-\epsilon)\overline{\Omega} \subseteq \Omega$ for any $\epsilon \in (0, 1)$, so we can use the result already proved and we get

$$\mathrm{Per}_\sigma(E, \Omega'_\epsilon) \leqslant 2\,\mathrm{Per}_\sigma(\Omega'_\epsilon, \mathbb{R}^n). \tag{2-21}$$

Moreover,

$$\mathrm{Per}_\sigma(\Omega'_\epsilon, \mathbb{R}^n) = (1-\epsilon)^{n-\sigma}\,\mathrm{Per}_\sigma(\Omega', \mathbb{R}^n). \tag{2-22}$$

Also, we claim that

$$\lim_{\epsilon \searrow 0} \mathrm{Per}_\sigma(E, \Omega'_\epsilon) = \mathrm{Per}_\sigma(E, \Omega'). \tag{2-23}$$

To prove it, we distinguish the cases $\sigma = 1$ and $\sigma \in (0, 1)$. If $\sigma = 1$, we use the representation of the perimeter of E in term of the Gauss–Green measure μ_E (see Remark 12.2 in [Maggi 2012]) and the monotone convergence theorem (applied to the monotone sequence of sets Ω'_ϵ, see, e.g., Theorem 1.26(a) in [Yeh 2006]): in this way, we have

$$\lim_{\epsilon\searrow 0}\operatorname{Per}(E,\Omega'_\epsilon)=\lim_{\epsilon\searrow 0}|\mu_E|(\Omega'_\epsilon)=|\mu_E|(\Omega')=\operatorname{Per}(E,\Omega').$$

This proves (2-23) when $\sigma=1$. If instead $\sigma\in(0,1)$, we first observe that $\operatorname{Per}_\sigma(E,\Omega'_\epsilon)\leqslant\operatorname{Per}_\sigma(E,\Omega')$ and then

$$\limsup_{\epsilon\searrow 0}\operatorname{Per}_\sigma(E,\Omega'_\epsilon)\leqslant\operatorname{Per}_\sigma(E,\Omega'). \tag{2-24}$$

Conversely, we use (1-1) to write

$$\begin{aligned}\operatorname{Per}_\sigma(E,\Omega'_\epsilon)&=L(E\cap\Omega'_\epsilon,E^c)+L(E^c\cap\Omega'_\epsilon,E\cap(\Omega'_\epsilon)^c)\\&\geqslant L(E\cap\Omega'_\epsilon,E^c)+L(E^c\cap\Omega'_\epsilon,E\cap(\Omega')^c).\end{aligned}$$

Consequently, by taking the limit of the inequality above and using Fatou's lemma,

$$\liminf_{\epsilon\searrow 0}\operatorname{Per}_\sigma(E,\Omega'_\epsilon)\geqslant L(E\cap\Omega',E^c)+L(E^c\cap\Omega',E\cap(\Omega')^c)=\operatorname{Per}_\sigma(E,\Omega').$$

This, together with (2-24), establishes (2-23).

Now, combining (2-21)–(2-23), we obtain (2-17) by taking a limit in ϵ. □

3. Proof of Theorem 1.1

Now we prove Theorem 1.1. The idea of the proof is that, on the one hand, for large balls, we obtain a large contribution of the perimeter, which makes the energy functional simply the Dirichlet energy plus a constant, due to the special form of Φ. On the other hand, for small balls, both the Dirichlet energy and the perimeter give a small contribution, and in this range the contribution of the perimeter becomes predominant. This dichotomy of the energy behavior makes the minimal pair change accordingly; namely, in large balls, harmonic functions are favored, somehow independently of their level sets, while, conversely, for small balls the sets which minimize the perimeter are favored, somehow independently on the Dirichlet energy of the function that they support. That is, in the end, the core of the counterexample is, roughly speaking, that being a minimal surface is something rather different than being the level set of a harmonic function.

Of course, some computations are needed to justify the above heuristic arguments and we present now all the details of the proof.

Estimates on $\operatorname{Per}_\sigma(E,B_R)$ from below. Here we obtain bounds from below for the (either classical or fractional) perimeter of a set E in B_R, once E is "suitably fixed" outside[3] the ball $B_R\subset\mathbb{R}^2$. For this scope, we recall the notation in (1-5) and (1-6), and we have:

[3]For simplicity, we state and prove all the results of this part only in $\mathbb{R}^2$, though some of the arguments would also be valid in higher dimensions.

Lemma 3.1. *Let $c_o > 0$. Let (u, E) be an admissible pair in $\mathbb{R}^2$. Assume $u - \tilde{u} \in H_0^1(B_1)$ and*

$$\int_{B_1} |\nabla u(X)|^2 \, dX \leqslant c_o.$$

Then there exists $c > 0$, possibly depending on c_o, such that

$$\operatorname{Per}_\sigma(E, B_1) \geqslant c. \tag{3-1}$$

Proof. We argue by contradiction. If the thesis in (3-1) were false, there would exist a sequence of admissible pairs (u_j, E_j) such that $u_j - \tilde{u} \in H_0^1(B_1)$,

$$\int_{B_1} |\nabla u_j(X)|^2 \, dX \leqslant c_o$$

and

$$\operatorname{Per}_\sigma(E_j, B_1) \leqslant \frac{1}{j}. \tag{3-2}$$

Thus, by compactness, (see, e.g., Corollary 3.49 in [Ambrosio et al. 2000] for the case $\sigma = 1$ or Theorem 7.1 in [Di Nezza et al. 2012] for the case $\sigma \in (0, 1)$), we conclude that, up to subsequences, u_j converges to some u_∞ weakly in $H^1(B_1)$ and strongly in $L^2(B_1)$, with

$$u_\infty - \tilde{u} \in H_0^1(B_1), \tag{3-3}$$

and χ_{E_j} converges to some χ_{E_∞} strongly in $L^1(B_1)$ as $j \to +\infty$. Accordingly, by the lower semicontinuity of the (either classical or fractional) perimeter (or by Fatou's lemma; see, e.g., Proposition 3.38(b) in [Ambrosio et al. 2000] for the case $\sigma = 1$) we deduce from (3-2) that

$$\operatorname{Per}_\sigma(E_\infty, B_1) = 0.$$

Hence, from the relative isoperimetric inequality (see, e.g., Lemma 2.5 in [Di Castro et al. 2015] when $\sigma \in (0, 1)$ and formula (12.46) in [Maggi 2012] when $\sigma = 1$),

$$\min\left\{|B_1 \cap E_\infty|^{(2-\sigma)/2}, |B_1 \setminus E_\infty|^{(2-\sigma)/2}\right\} \leqslant \widehat{C} \operatorname{Per}_\sigma(E_\infty, B_1) = 0$$

for some $\widehat{C} > 0$. Thus, we can suppose

$$|B_1 \cap E_\infty| = 0, \tag{3-4}$$

the case $|B_1 \setminus E_\infty| = 0$ being similar. Also, by virtue of Lemma 2.1, we have $u_\infty \geqslant 0$ a.e. in E_∞ and $u_\infty \leqslant 0$ a.e. in E_∞^c. Thus, by (3-4), we obtain that $u_\infty \leqslant 0$ a.e. in B_1. Looking at a neighborhood of ∂B_1 in the first quadrant, we obtain that this is in contradiction with (3-3), thus proving the desired result. □

By scaling Lemma 3.1, we obtain:

Lemma 3.2. *Let $c_o > 0$ and $R > 0$. Let (u, E) be an admissible pair in $\mathbb{R}^2$. Assume $u - \tilde{u} \in H_0^1(B_R)$ and*

$$\int_{B_R} |\nabla u(X)|^2 \, dX \leqslant c_o R^4. \tag{3-5}$$

Then there exists $c > 0$, possibly depending on c_o, such that

$$\operatorname{Per}_\sigma(E, B_R) \geqslant c R^{2-\sigma}.$$

Proof. We set

$$u_*(X) := R^{-2}u(RX) \quad \text{and} \quad E_* := \frac{E}{R} := \left\{\frac{X}{R} : X \in E\right\}.$$

Notice that $R^{-2}\tilde{u}(RX) = R^{-2}(Rx)(Ry) = \tilde{u}(X)$; therefore $u_* - \tilde{u} \in H_0^1(B_1)$. Also, (u_*, E_*) is an admissible pair. In addition,

$$\int_{B_1} |\nabla u_*(X)|^2 \, dX = R^{-2}\int_{B_1} |\nabla u(RX)|^2 \, dX = R^{-4}\int_{B_R} |\nabla u(Y)|^2 \, dY \leqslant c_o,$$

thanks to (3-5). As a consequence, we are in a position to apply Lemma 3.1 to the pair (u_*, E_*) and thus we obtain

$$c \leqslant \operatorname{Per}_\sigma(E_*, B_1) = \operatorname{Per}_\sigma\left(\frac{E}{R}, \frac{B_R}{R}\right) = \frac{1}{R^{2-\sigma}} \operatorname{Per}_\sigma(E, B_R),$$

as desired. □

Analysis of minimizers in large balls. Now we give a concrete example of a minimizer in $B_R \subset \mathbb{R}^2$ for R large enough. To this end, we consider a monotone nondecreasing and lower semicontinuous function $\widetilde{\Phi} : [0, +\infty) \to [0, +\infty)$, with

$$\widetilde{\Phi}(t) = 1 \quad \text{for any } t \in [2, +\infty). \tag{3-6}$$

We let

$$\widetilde{\mathcal{E}}_\Omega(u, E) := \int_\Omega |\nabla u(X)|^2 \, dX + \widetilde{\Phi}(\operatorname{Per}^\star_\sigma(E, \Omega)).$$

We remark that, in principle, the minimization procedure in Lemma 2.3 fails for this functional, since the coercivity assumption (1-3) is not satisfied by $\widetilde{\Phi}$. Nevertheless, we will be able to construct explicitly a minimizer for large balls of $\widetilde{\mathcal{E}}$. Then, we will modify $\widetilde{\Phi}$ at infinity and we will obtain from it a minimizer for a functional of the type in (1-4), with a coercive Φ. The details are as follows.

Proposition 3.3. *Let $n = 2$. Let $\tilde{u}$ and $\widetilde{E}$ be as in* (1-5) *and* (1-6).

Then, there exists $R_o > 0$, only depending on n and σ, such that if $R \geqslant R_o$ then

$$\widetilde{\mathcal{E}}_{B_R}(\tilde{u}, \widetilde{E}) \leqslant \widetilde{\mathcal{E}}_{B_R}(v, F) \tag{3-7}$$

for any admissible pair (v, F) such that $v - \tilde{u} \in H_0^1(B_R)$ and $F \setminus B_R = \widetilde{E} \setminus B_R$, up to sets of measure zero.

Proof. We observe that $\nabla\tilde{u}(x, y) = (y, x)$, and so

$$\int_{B_R} |\nabla\tilde{u}(X)|^2 \, dX = \int_{B_R} |X|^2 \, dX \leqslant C_1 R^4 \tag{3-8}$$

for some $C_1 > 0$. Moreover, since $\widetilde{E}$ is a cone, we have $\widetilde{E} = R\widetilde{E}$; thus

$$\operatorname{Per}_\sigma(\widetilde{E}, B_R) = \operatorname{Per}_\sigma(R\widetilde{E}, RB_1) = C_2 R^{2-\sigma}$$

for some $C_2 > 0$. In particular, if $R \geqslant (2/C_2)^{1/(2-\sigma)}$, we have

$$\operatorname{Per}^\star_\sigma(\widetilde{E}, B_R) \geqslant \operatorname{Per}_\sigma(\widetilde{E}, B_R) \geqslant 2,$$

and then, by (3-6),

$$\widetilde{\Phi}(\mathrm{Per}^{\star}_{\sigma}(\widetilde{E}, B_R)) = 1. \tag{3-9}$$

This and (3-8) imply that

$$\widetilde{\mathcal{E}}_{B_R}(\tilde{u}, \widetilde{E}) \leqslant C_1 R^4 + 1 \leqslant 2C_1 R^4 \tag{3-10}$$

if R is large enough.

Now suppose, by contradiction, that (3-7) is violated, i.e.,

$$\widetilde{\mathcal{E}}_{B_R}(\tilde{u}, \widetilde{E}) > \widetilde{\mathcal{E}}_{B_R}(v, F) \tag{3-11}$$

for some competitor (v, F). In particular, by (3-10),

$$\int_{B_R} |\nabla v(X)|^2 \, dX \leqslant \widetilde{\mathcal{E}}_{B_R}(v, F) \leqslant \widetilde{\mathcal{E}}_{B_R}(\tilde{u}, \widetilde{E}) \leqslant 2C_1 R^4. \tag{3-12}$$

This says that formula (3-5) is satisfied by the pair (v, F) with $c_o := 2C_1$, and so Lemma 3.2 gives

$$\mathrm{Per}^{\star}_{\sigma}(F, B_R) \geqslant \mathrm{Per}_{\sigma}(F, B_R) \geqslant c R^{2-\sigma}$$

for some $c > 0$. In particular, for large R, we have

$$\widetilde{\Phi}(\mathrm{Per}^{\star}_{\sigma}(F, B_R)) = 1$$

and therefore

$$\widetilde{\mathcal{E}}_{B_R}(v, F) = \int_{B_R} |\nabla v(X)|^2 \, dX + 1. \tag{3-13}$$

On the other hand, since $\tilde{u}$ is harmonic,

$$\int_{B_R} |\nabla v(X)|^2 \, dX \geqslant \int_{B_R} |\nabla \tilde{u}(X)|^2 \, dX;$$

hence (3-13) and (3-9) give

$$\widetilde{\mathcal{E}}_{B_R}(v, F) \geqslant \int_{B_R} |\nabla \tilde{u}(X)|^2 \, dX + 1 = \widetilde{\mathcal{E}}_{B_R}(\tilde{u}, \widetilde{E}).$$

This is in contradiction with (3-11) and so the desired result is established. □

Corollary 3.4. *Let $n = 2$. Let $\tilde{u}$ and $\widetilde{E}$ be as in* (1-5) *and* (1-6). *There exists $K_o > 2$ such that the following statement is true. Assume*

$$\Phi(t) = 1 \quad \textit{for any } t \in [2, K_o]. \tag{3-14}$$

Then, there exists $R_o > 0$ such that $(\tilde{u}, \widetilde{E})$ is a minimal pair in B_{R_o}.

Proof. We define

$$\widetilde{\Phi}(t) := \begin{cases} \Phi(t) & \text{if } t \in [0, 2], \\ 1 & \text{if } t \in (2, +\infty). \end{cases}$$

Then we are in the setting of Proposition 3.3 and we obtain that there exists $R_o > 0$, only depending on n and σ, such that $(\tilde{u}, \widetilde{E})$ is a minimal pair for $\widetilde{\mathcal{E}}_{B_{R_o}}$. So we define

$$K_o := \mathrm{Per}^\star_\sigma(\widetilde{E}, B_{R_o}) + 3.$$

Notice that K_o only depends on n and σ, since does R_o also, and $\tilde{u}$ and $\widetilde{E}$ are fixed.

To complete the proof of the desired claim, we need to show that $(\tilde{u}, \widetilde{E})$ is a minimal pair for $\mathcal{E}_{B_{R_o}}$, as long as (3-14) is satisfied. For this, we remark that, since Φ is monotone, we have $\Phi(t) \geqslant \Phi(2) = 1$ for any $t \geqslant 2$. As a consequence, we get $\Phi(t) \geqslant \widetilde{\Phi}(t)$ for any $t \geqslant 0$. Therefore, if (v, F) is a competitor for $(\tilde{u}, \widetilde{E})$, we deduce from (3-7) that

$$\widetilde{\mathcal{E}}_{B_{R_o}}(\tilde{u}, \widetilde{E}) \leqslant \widetilde{\mathcal{E}}_{B_{R_o}}(v, F) \leqslant \mathcal{E}_{B_{R_o}}(v, F). \tag{3-15}$$

On the other hand,

$$\mathrm{Per}^\star_\sigma(\widetilde{E}, B_{R_o}) \leqslant K_o. \tag{3-16}$$

Moreover, we have $\widetilde{\Phi}(t) = 1 = \Phi(t)$ if $t \in (2, K_o]$. Therefore, we get $\widetilde{\Phi} = \Phi$ in $[0, K_o]$ and thus, by (3-16),

$$\widetilde{\Phi}(\mathrm{Per}^\star_\sigma(\widetilde{E}, B_{R_o})) = \Phi(\mathrm{Per}^\star_\sigma(\widetilde{E}, B_{R_o})).$$

By plugging this into (3-15), we conclude that

$$\mathcal{E}_{B_{R_o}}(\tilde{u}, \widetilde{E}) = \widetilde{\mathcal{E}}_{B_{R_o}}(\tilde{u}, \widetilde{E}) \leqslant \mathcal{E}_{B_{R_o}}(v, F),$$

as desired. □

Estimates in small balls. Here, we show that the minimal pair constructed in Corollary 3.4 in large balls does not remain minimal in small balls.

Proposition 3.5. *Let $n = 2$. Assume*

$$\Phi(t) = t^\gamma \text{ for any } t \in [0, 1] \tag{3-17}$$

for some

$$\gamma \in \left(0, \frac{4}{2-\sigma}\right). \tag{3-18}$$

Let $\tilde{u}$ and $\widetilde{E}$ be as in (1-5) *and* (1-6).

Then there exists $r_o > 0$ such that if $r \in (0, r_o]$ then the pair $(\tilde{u}, \widetilde{E})$ is not minimal in B_r.

Proof. We suppose, by contradiction, that $(\tilde{u}, \widetilde{E})$ is minimal in B_r, with r sufficiently small.

We observe that $\widetilde{E}$ is not a minimizer of the perimeter in $\overline{B}_{1/2}$ (see [Savin and Valdinoci 2013] for the case $\sigma \in (0, 1)$). Therefore there exists a perturbation $E_\sharp$ of $\widetilde{E}$ inside $\overline{B}_{1/2}$ for which

$$\mathrm{Per}_\sigma(E_\sharp, \overline{B}_{1/2}) \leqslant \mathrm{Per}_\sigma(\widetilde{E}, \overline{B}_{1/2}) - a$$

for some (small, but fixed) $a > 0$. As a consequence, recalling Lemma 2.2,

$$\mathrm{Per}_\sigma(E_\sharp, B_1) - \mathrm{Per}_\sigma(\widetilde{E}, B_1) = \mathrm{Per}_\sigma(E_\sharp, \overline{B}_{1/2}) - \mathrm{Per}_\sigma(\widetilde{E}, \overline{B}_{1/2}) \leqslant -a. \tag{3-19}$$

Now we take $\psi \in C^\infty(\mathbb{R}^2, [0,1])$ such that $\psi(X)=0$ for any $X \in B_{3/4}$ and $\psi(X)=1$ for any $X \in B^c_{9/10}$. We define

$$u_\sharp(X) = u_\sharp(x,y) := \tilde{u}(X)\,\psi(X) = xy\,\psi(x,y).$$

We claim that

$$u_\sharp \geqslant 0 \quad \text{a.e. in } E_\sharp \qquad \text{and} \qquad u_\sharp \leqslant 0 \quad \text{a.e. in } E^c_\sharp. \tag{3-20}$$

To check this, we observe that $u_\sharp = 0$ in $B_{3/4}$, so it is enough to prove (3-20) for points outside $B_{3/4}$. Then, we also remark that $E_\sharp \setminus B_{3/4} = \widetilde{E} \setminus B_{3/4}$, and, as a consequence, we get that $\tilde{u} \geqslant 0$ a.e. in $E_\sharp \setminus B_{3/4}$ and $\tilde{u} \leqslant 0$ a.e. in $E^c_\sharp \setminus B_{3/4}$. Hence, since $\psi \geqslant 0$, we obtain that $u_\sharp \geqslant 0$ a.e. in $E_\sharp \setminus B_{3/4}$ and $u_\sharp \leqslant 0$ a.e. in $E^c_\sharp \setminus B_{3/4}$. These observations complete the proof of (3-20).

Now we define

$$u_r(X) := r^2 u_\sharp\Big(\frac{X}{r}\Big) = xy\,\psi\Big(\frac{X}{r}\Big) = \tilde{u}(X)\,\psi\Big(\frac{X}{r}\Big)$$

and

$$E_r := r E_\sharp.$$

From (3-20), we obtain that $u_r \geqslant 0$ a.e. in E_r and $u_r \leqslant 0$ a.e. in E^c_r, and thus (u_r, E_r) is an admissible pair.

Now we check that the data of (u_r, E_r) coincide with $(\tilde{u}, \widetilde{E})$ outside B_r. First of all, we have that $\psi = 1$ in $B^c_{9/10}$; thus, if $X \in B^c_{9r/10}$ we have $u_r(X) = \tilde{u}(X)$. This shows that

$$u_r - \tilde{u} \in H^1_0(B_r). \tag{3-21}$$

Moreover,

$$E_r \setminus B_r = \{X \in B^c_r : r^{-1}X \in E_\sharp\} = \{X = rY : Y \in E_\sharp \setminus B_1\} = \{X = rY : Y \in \widetilde{E} \setminus B_1\}.$$

Now, since $\widetilde{E}$ is a cone, we have $Y \in \widetilde{E}$ if and only if $rY \in \widetilde{E}$, and so, as a consequence,

$$E_r \setminus B_r = \{X = rY \in \widetilde{E} : Y \in B^c_1\} = \widetilde{E} \setminus B_r.$$

Using this and (3-21), we obtain that, if $(\tilde{u}, \widetilde{E})$ is minimal in B_r, then

$$\mathcal{E}_{B_r}(\tilde{u}, \widetilde{E}) \leqslant \mathcal{E}_{B_r}(u_r, E_r). \tag{3-22}$$

Now we remark that, since $\widetilde{E}$ is a cone,

$$\operatorname{Per}_\sigma(\widetilde{E}, B_r) = \operatorname{Per}_\sigma(r\widetilde{E}, rB_1) = r^{2-\sigma}\operatorname{Per}_\sigma(\widetilde{E}, B_1). \tag{3-23}$$

Now we define

$$\vartheta := \begin{cases} 4\Upsilon & \text{if } \sigma = 1, \\ 0 & \text{if } \sigma \in (0,1), \end{cases}$$

and we claim that

$$\operatorname{Per}^\star_\sigma(\widetilde{E}, B_r) = r^{2-\sigma}\operatorname{Per}_\sigma(\widetilde{E}, B_1) + \vartheta. \tag{3-24}$$

Indeed, if $\sigma \in (0,1)$, then (3-24) boils down to (3-23). If instead $\sigma = 1$, we use (3-23) in the following computation:

$$\operatorname{Per}^\star_\sigma(\widetilde{E}, B_r) = \operatorname{Per}(\widetilde{E}, B_{r+\Upsilon}) = \operatorname{Per}(\widetilde{E}, B_r) + \operatorname{Per}(\widetilde{E}, B_{r+\Upsilon} \setminus B_r) = r^{2-\sigma}\operatorname{Per}_\sigma(\widetilde{E}, B_1) + 4\Upsilon.$$

This proves (3-24).

From (3-24) we obtain that

$$\mathcal{E}_{B_r}(\tilde{u}, \widetilde{E}) \geqslant \Phi\big(r^{2-\sigma} \operatorname{Per}_\sigma(\widetilde{E}, B_1) + \vartheta\big). \tag{3-25}$$

On the other hand, recalling (3-19), we have

$$\operatorname{Per}_\sigma(E_r, B_r) = \operatorname{Per}_\sigma(r E_\sharp, B_r) = r^{2-\sigma} \operatorname{Per}_\sigma(E_\sharp, B_1) \leqslant r^{2-\sigma}(\operatorname{Per}_\sigma(\widetilde{E}, B_1) - a). \tag{3-26}$$

Now we claim that

$$\operatorname{Per}^\star_\sigma(E_r, B_r) \leqslant r^{2-\sigma}(\operatorname{Per}_\sigma(\widetilde{E}, B_1) - a) + \vartheta. \tag{3-27}$$

Indeed, if $\sigma \in (0, 1)$ then (3-27) reduces to (3-26). If instead $\sigma = 1$, we use the fact that E_r coincides with $\widetilde{E}$ outside B_r and (3-26) to see that

$$\operatorname{Per}^\star_\sigma(E_r, B_r) = \operatorname{Per}(E_r, B_{r+\Upsilon}) = \operatorname{Per}(E_r, B_r) + \operatorname{Per}(E_r, B_{r+\Upsilon} \setminus B_r) \leqslant r^{2-\sigma}(\operatorname{Per}_\sigma(\widetilde{E}, B_1) - a) + 4\Upsilon.$$

This establishes (3-27).

Then, the monotonicity of Φ and (3-27) give

$$\Phi(\operatorname{Per}^\star_\sigma(E_r, B_r)) \leqslant \Phi\big(r^{2-\sigma}(\operatorname{Per}_\sigma(\widetilde{E}, B_1) - a) + \vartheta\big) \tag{3-28}$$

Now we remark that

$$|\nabla u_r(X)| \leqslant |\nabla \tilde{u}(X)\, \psi(X/r)| + r^{-1}|\tilde{u}(X)\, \nabla\psi(X/r)| \leqslant |X| + Cr^{-1}|X|^2,$$

for some $C > 0$. As a consequence of this, and possibly renaming $C > 0$, we obtain

$$\int_{B_r} |\nabla u_r(X)|^2\, dX \leqslant C \int_{B_r} \big(|X|^2 + r^{-2}|X|^4\big)\, dX \leqslant Cr^4.$$

This and (3-28) give

$$\mathcal{E}_{B_r}(u_r, E_r) \leqslant Cr^4 + \Phi\big(r^{2-\sigma}(\operatorname{Per}_\sigma(\widetilde{E}, B_1) - a) + \vartheta\big).$$

Putting together this, (3-22) and (3-25), we conclude that

$$\Phi\big(r^{2-\sigma} \operatorname{Per}_\sigma(\widetilde{E}, B_1) + \vartheta\big) \leqslant Cr^4 + \Phi\big(r^{2-\sigma}(\operatorname{Per}_\sigma(\widetilde{E}, B_1) - a) + \vartheta\big).$$

Thus, if $r^{2-\sigma} \operatorname{Per}_\sigma(\widetilde{E}, B_1) \leqslant \frac{1}{2}$, and so $\operatorname{Per}_\sigma(\widetilde{E}, B_1) + \vartheta \leqslant 1$, we can use (3-17) and obtain

$$\big[r^{2-\sigma} \operatorname{Per}_\sigma(\widetilde{E}, B_1) + \vartheta\big]^\gamma \leqslant Cr^4 + \big[r^{2-\sigma}(\operatorname{Per}_\sigma(\widetilde{E}, B_1) - a) + \vartheta\big]^\gamma. \tag{3-29}$$

Now we distinguish the cases $\sigma \in (0, 1)$ and $\sigma = 1$. When $\sigma \in (0, 1)$, we have $\vartheta = 0$ and so (3-29) becomes

$$r^{(2-\sigma)\gamma}(\operatorname{Per}_\sigma(\widetilde{E}, B_1))^\gamma \leqslant Cr^4 + r^{(2-\sigma)\gamma}(\operatorname{Per}_\sigma(\widetilde{E}, B_1) - a)^\gamma.$$

So we multiply by $r^{(\sigma-2)\gamma}$ and we get

$$a_* := (\operatorname{Per}_\sigma(\widetilde{E}, B_1))^\gamma - (\operatorname{Per}_\sigma(\widetilde{E}, B_1) - a)^\gamma \leqslant Cr^{4+(\sigma-2)\gamma}.$$

Notice that $a_* > 0$ since $a > 0$, and therefore the latter inequality gives a contradiction if r is small enough, thanks to (3-18). This concludes the case in which $\sigma \in (0, 1)$.

If instead $\sigma = 1$, then we have $\vartheta > 0$ and so, for small t, we have

$$(t + \vartheta)^\gamma = \vartheta^\gamma + \gamma\vartheta^{\gamma-1} t + O(t^2).$$

Therefore, we infer from (3-29) that

$$\vartheta^\gamma + \gamma\vartheta^{\gamma-1} r^{2-\sigma} \operatorname{Per}_\sigma(\widetilde{E}, B_1) \leqslant \vartheta^\gamma + \gamma\vartheta^{\gamma-1} r^{2-\sigma}(\operatorname{Per}_\sigma(\widetilde{E}, B_1) - a) + O(r^{4-2\sigma}).$$

Hence we simplify some terms and we divide by $r^{2-\sigma}$ to obtain

$$a \leqslant O(r^{2-\sigma}),$$

which gives a contradiction for small $r > 0$. This completes the case $\sigma = 1$. □

Completion of the proof of Theorem 1.1. The claim in Theorem 1.1 now follows plainly by combining Corollary 3.4 and Proposition 3.5.

4. Proof of Theorem 1.2

The argument is a combination of a classical domain variation (see, e.g., [Alt and Caffarelli 1981]) with an expansion of the (classical or fractional) perimeter. Some similar perturbative methods appear, in the classical case, for instance, in [Garofalo and Lin 1986; Caffarelli et al. 2009]. Since the arguments involved here use both standard and nonstandard observations, we give all the details to assist the reader. First, we observe that

$$\text{the function } \Xi := (\partial_\nu^+ u(x))^2 - (\partial_\nu^- u(x))^2 - H_\sigma^E(x)\,\Phi'(\mathrm{Per}_\sigma^\star(E,\Omega)) \text{ belongs to } C(\partial E \cap \Omega), \tag{4-1}$$

thanks to (1-8), (1-10) and Proposition 6.3 in [Figalli et al. 2015] (to be used when $\sigma \in (0, 1)$).

Also, given a vector field $V \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ such that

$$V(x) = 0 \quad \text{for any } x \in \Omega^c, \tag{4-2}$$

for small $t \in \mathbb{R}$ we consider the ODE flow $y = y(t; x)$ given by the Cauchy problem

$$\begin{cases} \partial_t y(t; x) = V(y(t; x)), \\ y(0; x) = x. \end{cases} \tag{4-3}$$

We remark that, for small $t \in \mathbb{R}$,

$$y(t; x) = x + tV(y(t; x)) + o(t) = x + tV(x) + o(t). \tag{4-4}$$

Accordingly,

$$D_x y(t; x) = I + t\,DV(x) + o(t) = I + t\,DV(y(t; x)) + o(t), \tag{4-5}$$

where I denotes the n-dimensional identity matrix.

Also, the map $\mathbb{R}^n \ni x \mapsto y(t; x)$ is invertible for small t; i.e., we can consider the inverse diffeomorphism $x(t; y)$. In this way,

$$x(t; y(t; x)) = x \quad \text{and} \quad y(t; x(t; x)) = y. \tag{4-6}$$

By (4-4), we know

$$x(t; y) = y(t; x(t; y)) - tV(y(t; x(t; y))) + o(t) = y - tV(y) + o(t), \tag{4-7}$$

and therefore

$$D_y x(t; y) = I - t\,DV(y) + o(t).$$

In particular,

$$\det D_y x(t;\, y) = 1 - t \operatorname{div} V(y) + o(t). \tag{4-8}$$

Now, given a minimal pair (u, E) as in the statement of Theorem 1.2, we define

$$u_t(y) := u(x(t;\, y)).$$

We remark that the subscript t above does not represent a time derivative. By (4-6), we can write $u(x) = u_t(y(t;\, x))$ and thus, recalling (4-5),

$$\nabla u(x) = D_x y(t;\, x)\, \nabla u_t(y(t;\, x)) = \nabla u_t(y(t;\, x)) + t\, DV(y(t;\, x))\, \nabla u_t(y(t;\, x)) + o(t). \tag{4-9}$$

Also, we consider the image of the set E under the diffeomorphism $y(t;\, \cdot)$; i.e., we define

$$E_t := y(t;\, E).$$

We claim that

$$\text{the pair } (u_t, E_t) \text{ is admissible.} \tag{4-10}$$

To check this, let $y \in E_t$ (resp., $y \in E_t^c$). Then there exists

$$x \in E \quad (\text{resp. } x \in E^c) \tag{4-11}$$

such that $y = y(t;\, x)$. Then, by (4-6), we have

$$x(t;\, y) = x(t;\, y(t;\, x)) = x.$$

This identity and (4-11) imply

$$0 \leqslant u(x) = u(x(t;\, y)) = u_t(y) \quad (\text{resp. } 0 \geqslant u_t(y)).$$

From this, we obtain (4-10).

In addition, we recall that

$$y(t;\, x) = x \quad \text{for any } x \in \Omega^c, \tag{4-12}$$

thanks to (4-2) and (4-3). Therefore, we have

$$y(t;\, \Omega) = \Omega. \tag{4-13}$$

Moreover, as a consequence of (4-12) and of (4-10), and using the minimality of (u, E), we have

$$0 \leqslant \mathcal{E}_\Omega(u_t, E_t) - \mathcal{E}_\Omega(u, E). \tag{4-14}$$

Now we compute the first order in t of the right-hand side of (4-14). For this scope, using, for instance, formula (6.3) (when $\sigma = 1$) or formula (6.12) (when $\sigma \in (0, 1)$) in [Figalli et al. 2015], and recalling that V vanishes outside Ω, one obtains that

$$\operatorname{Per}^\star_\sigma(E_t, \Omega) = \operatorname{Per}^\star_\sigma(E, \Omega) + t \int_{(\partial E)\cap\Omega} H^E_\sigma(x)\, V(x) \cdot \nu(x)\, d\mathcal{H}^{n-1}(x) + o(t). \tag{4-15}$$

Above, we denote by ν the exterior normal of E and by $\mathcal{H}^{n-1}$ the $(n-1)$-dimensional Hausdorff measure.

From (4-15), we obtain that

$$\begin{aligned}\Phi(\mathrm{Per}^\star_\sigma(E_t,\Omega)) &= \Phi\left(\mathrm{Per}^\star_\sigma(E,\Omega)+t\int_{(\partial E)\cap\Omega} H^E_\sigma(x)\,V(x)\cdot\nu(x)\,d\mathcal{H}^{n-1}(x)+o(t)\right)\\ &= \Phi(\mathrm{Per}^\star_\sigma(E,\Omega))+t\,\Phi'(\mathrm{Per}^\star_\sigma(E,\Omega))\int_{(\partial E)\cap\Omega} H^E_\sigma(x)\,V(x)\cdot\nu(x)\,d\mathcal{H}^{n-1}(x)+o(t). \qquad (4\text{-}16)\end{aligned}$$

Moreover, by (4-9),

$$|\nabla u(x)|^2 = |\nabla u_t(y(t;x))|^2 + 2t\,\nabla u_t(y(t;x))\cdot\big(DV(y(t;x))\,\nabla u_t(y(t;x))\big)+o(t).$$

Now we integrate this equation in x over Ω and we use the change of variable $y := y(t;x)$. In this way, recalling (4-8) and (4-13), we see that

$$\begin{aligned}\int_\Omega |\nabla u(x)|^2\,dx &= \int_\Omega \big[|\nabla u_t(y(t;x))|^2 + 2t\,\nabla u_t(y(t;x))\cdot\big(DV(y(t;x))\,\nabla u_t(y(t;x))\big)\big]\,dx + o(t)\\ &= \int_\Omega \big[|\nabla u_t(y)|^2 + 2t\,\nabla u_t(y)\cdot(DV(y)\,\nabla u_t(y))\big]\,|\det D_y x(t;y)|\,dy + o(t)\\ &= \int_\Omega \big[|\nabla u_t(y)|^2 + 2t\,\nabla u_t(y)\cdot(DV(y)\,\nabla u_t(y))\big]\,[1-t\,\mathrm{div}\,V(y)]\,dy + o(t)\\ &= \int_\Omega \big[|\nabla u_t(y)|^2 + 2t\,\nabla u_t(y)\cdot(DV(y)\,\nabla u_t(y)) - t\,|\nabla u_t(y)|^2\,\mathrm{div}\,V(y)\big]\,dy + o(t).\end{aligned}$$

We write this formula as

$$\begin{aligned}&\int_\Omega |\nabla u_t(y)|^2\,dy\\ &\quad= \int_\Omega |\nabla u(x)|^2\,dx + t\int_\Omega \big[|\nabla u_t(y)|^2\,\mathrm{div}\,V(y) - 2\,\nabla u_t(y)\cdot(DV(y)\,\nabla u_t(y))\big]\,dy + o(t). \qquad (4\text{-}17)\end{aligned}$$

Also, by (4-9),

$$\nabla u(x) = \nabla u_t(y(t;x)) + O(t),$$

and so, evaluating this expression at $x := x(t;y)$ and using (4-7), we get

$$\nabla u_t(y) = \nabla u_t(y(t;x(t;y))) = \nabla u(x(t;y)) + O(t) = \nabla u(y) + O(t).$$

We can substitute this into (4-17), thus obtaining

$$\begin{aligned}&\int_\Omega |\nabla u_t(y)|^2\,dy\\ &\quad= \int_\Omega |\nabla u(x)|^2\,dx + t\int_\Omega \big[|\nabla u(y)|^2\,\mathrm{div}\,V(y) - 2\,\nabla u(y)\cdot(DV(y)\,\nabla u(y))\big]\,dy + o(t). \qquad (4\text{-}18)\end{aligned}$$

Now we define $\Omega_1 := \Omega\cap\{u>0\}$ and $\Omega_2 := \Omega\cap\{u<0\}$. Notice that $\Delta u = 0$ in Ω_1 and in Ω_2, thanks to Lemma 2.4. Accordingly, in both Ω_1 and Ω_2 we have

$$\mathrm{div}(|\nabla u|^2\,V) = |\nabla u|^2\,\mathrm{div}\,V + 2V\cdot(D^2u\,\nabla u) \qquad (4\text{-}19)$$

and

$$\mathrm{div}((V\cdot\nabla u)\nabla u) = \nabla(V\cdot\nabla u)\cdot\nabla u = \nabla u\cdot(DV\,\nabla u) + V\cdot(D^2u\,\nabla u). \qquad (4\text{-}20)$$

So, we take the quantity in (4-19) and we subtract twice the quantity in (4-20); in this way we see that, in both Ω_1 and Ω_2,

$$\begin{aligned}\operatorname{div}(|\nabla u|^2\, V) - 2\operatorname{div}((V\cdot\nabla u)\nabla u) &= |\nabla u|^2 \operatorname{div} V + 2V\cdot(D^2u\,\nabla u) - 2\big[\nabla u\cdot(DV\nabla u) + V\cdot(D^2u\,\nabla u)\big]\\ &= |\nabla u|^2 \operatorname{div} V - 2\nabla u\cdot(DV\nabla u).\end{aligned}$$

We remark that the last expression is exactly the quantity appearing in one integrand of (4-18); therefore we can write (4-18) as

$$\begin{aligned}&\int_\Omega |\nabla u_t(y)|^2\,dy\\ &\quad = \int_\Omega |\nabla u(x)|^2\,dx + t\sum_{i\in\{1,2\}}\int_{\Omega_i}\big[\operatorname{div}\big(|\nabla u(y)|^2\,V(y)\big) - 2\operatorname{div}\big((V(y)\cdot\nabla u(y))\nabla u(y)\big)\big]\,dy + o(t). \qquad (4\text{-}21)\end{aligned}$$

Now we recall (1-9) and we notice that the exterior normal ν_1 of Ω_1 coincides with ν, while the exterior normal ν_2 of Ω_2 coincides with $-\nu$. Furthermore, by (1-11), we see that $\nu_1 = -\nabla u/|\nabla u| = -\nabla u/|\partial_\nu^+ u|$ coming from Ω_1 and $\nu_2 = \nabla u/|\nabla u| = \nabla u/|\partial_\nu^- u|$ coming from Ω_2. Accordingly, coming from Ω_1, we have

$$\partial_{\nu_1} u = \nu_1\cdot\nabla u = -\frac{\nabla u}{|\nabla u|}\cdot\nabla u = -|\partial_\nu^+ u|.$$

Similarly, coming from Ω_2,

$$\partial_{\nu_2} u = \nu_2\cdot\nabla u = \frac{\nabla u}{|\nabla u|}\cdot\nabla u = |\partial_\nu^- u|.$$

Therefore, coming from Ω_1,

$$\nabla u\,\partial_{\nu_1} u = -|\nabla u|\,\partial_{\nu_1} u\,\nu_1 = |\partial_\nu^+ u|^2\,\nu,$$

and coming from Ω_2,

$$\nabla u\,\partial_{\nu_2} u = |\nabla u|\,\partial_{\nu_2} u\,\nu_2 = -|\partial_\nu^- u|^2\,\nu.$$

Consequently, coming from Ω_1 we have

$$|\nabla u|^2\,V\cdot\nu_1 - 2(V\cdot\nabla u)\partial_{\nu_1} u = |\partial_\nu^+ u|^2 V\cdot\nu - 2(V\cdot\nu)\,|\partial_\nu^+ u|^2 = -|\partial_\nu^+ u|^2 V\cdot\nu,$$

while, coming from Ω_2,

$$|\nabla u|^2\,V\cdot\nu_2 - 2(V\cdot\nabla u)\partial_{\nu_2} u = -|\partial_\nu^- u|^2 V\cdot\nu + 2(V\cdot\nu)\,|\partial_\nu^- u|^2 = |\partial_\nu^- u|^2 V\cdot\nu.$$

Hence, if we apply the divergence theorem in (4-21), we obtain

$$\begin{aligned}&\int_\Omega |\nabla u_t(y)|^2\,dy - \int_\Omega |\nabla u(x)|^2\,dx\\ &\quad = t\sum_{i\in\{1,2\}}\int_{\partial\Omega_i}\big[|\nabla u(y)|^2\,V(y)\cdot\nu_i(y) - 2(V(y)\cdot\nabla u(y))\partial_{\nu_i} u(y)\big]\,d\mathcal{H}^{n-1}(y) + o(t)\\ &\quad = -t\int_{(\partial E)\cap\Omega} |\partial_\nu^+ u(y)|^2\,V(y)\cdot\nu(y)\,d\mathcal{H}^{n-1}(y) + t\int_{(\partial E)\cap\Omega} |\partial_\nu^- u(y)|^2\,V(y)\cdot\nu(y)\,d\mathcal{H}^{n-1}(y) + o(t). \qquad (4\text{-}22)\end{aligned}$$

Using this and (4-16), and also recalling the definition in (4-1), we conclude that

$$\begin{aligned}\mathcal{E}_\Omega(u_t, E_t) - \mathcal{E}_\Omega(u, E) &= \int_\Omega |\nabla u_t(y)|^2\,dy - \int_\Omega |\nabla u(x)|^2\,dx + \Phi(\operatorname{Per}^\star_\sigma(E_t,\Omega)) - \Phi(\operatorname{Per}^\star_\sigma(E,\Omega))\\ &= t\int_{(\partial E)\cap\Omega}\big(|\partial_\nu^- u(y)|^2 - |\partial_\nu^+ u(y)|^2\big)\,V(y)\cdot\nu(y)\,d\mathcal{H}^{n-1}(y)\\ &\qquad + t\,\Phi'(\operatorname{Per}^\star_\sigma(E,\Omega))\int_{(\partial E)\cap\Omega} H^E_\sigma(x)\,V(x)\cdot\nu(x)\,d\mathcal{H}^{n-1}(x) + o(t)\\ &= -t\int_{(\partial E)\cap\Omega}\Xi(x)\,V(x)\cdot\nu(x)\,d\mathcal{H}^{n-1}(x) + o(t).\end{aligned}$$

This and (4-14) imply

$$\int_{(\partial E)\cap\Omega}\Xi(x)\,V(x)\cdot\nu(x)\,d\mathcal{H}^{n-1}(x) = 0.$$

Since V is arbitrary, the latter identity and (4-1) imply that Ξ vanishes in the whole of $\partial E\cap\Omega$, which completes the proof of Theorem 1.2.

5. Proof of Theorem 1.3

Energy of the harmonic replacement of a minimal solution. We start with a computation on the harmonic replacement:

Lemma 5.1. *Assume that* (1-14) *holds true. Let* (u, E) *be a minimal pair in* Ω, *with* $u \geqslant 0$ *a.e. in* Ω^c *and* $B_{R_o} \Subset \Omega$. *Let* $R \in (0, R_o]$ *and* u_R *be the function minimizing the Dirichlet energy in* B_R *among all the functions* v *such that* $v - u \in H^1_0(B_R)$. *Then*

$$\int_{B_R} |\nabla u(x) - \nabla u_R(x)|^2\,dx \leqslant C\,L_Q\,R^{n-\sigma}$$

for some $C > 0$, *possibly depending on* R_o, n *and* σ, *and* L_Q *is the one introduced in* (1-14).

Proof. We observe that $u \geqslant 0$ a.e. in $\mathbb{R}^n$, thanks to Lemma 2.7. Hence $u_R \geqslant 0$ a.e., by the classical maximum principle, and therefore, taking $u_R := u$ in B_R^c, we see that $(u_R, E\cup B_R)$ is an admissible pair, and an admissible competitor against (u, E). Therefore, by the minimality of (u, E),

$$\begin{aligned}0 &\leqslant \mathcal{E}_\Omega(u_R, E\cup B_R) - \mathcal{E}_\Omega(u, E)\\ &= \int_{B_R}\big(|\nabla u_R(x)|^2 - |\nabla u(x)|^2\big)\,dx + \Phi(\operatorname{Per}^\star_\sigma(E\cup B_R,\Omega)) - \Phi(\operatorname{Per}^\star_\sigma(E,\Omega)).\end{aligned}\tag{5-1}$$

Now we use the subadditivity of the (either classical or fractional) perimeter (see, e.g., Proposition 3.38(d) in [Ambrosio et al. 2000] when $\sigma = 1$ and formula (3.1) in [Dipierro et al. 2013] when $\sigma \in (0, 1)$) and we remark that

$$\begin{aligned}\operatorname{Per}^\star_\sigma(E\cup B_R,\Omega) &\leqslant \operatorname{Per}^\star_\sigma(E,\Omega) + \operatorname{Per}^\star_\sigma(B_R,\Omega) \leqslant \operatorname{Per}^\star_\sigma(E,\Omega) + \operatorname{Per}_\sigma(B_R,\mathbb{R}^n)\\ &= \operatorname{Per}^\star_\sigma(E,\Omega) + R^{n-\sigma}\operatorname{Per}_\sigma(B_1,\mathbb{R}^n) \leqslant Q,\end{aligned}\tag{5-2}$$

in light of (1-16).

Now we claim that

$$\Phi(\mathrm{Per}^\star_\sigma(E\cup B_R,\Omega))-\Phi(\mathrm{Per}^\star_\sigma(E,\Omega))\leqslant C\,L_Q\,R^{n-\sigma}. \tag{5-3}$$

To prove it, we observe that if $\mathrm{Per}^\star_\sigma(E\cup B_R,\Omega)\leqslant \mathrm{Per}^\star_\sigma(E,\Omega)$ then, by the monotonicity of Φ it follows that $\Phi(\mathrm{Per}^\star_\sigma(E\cup B_R,\Omega))\leqslant \Phi(\mathrm{Per}^\star_\sigma(E,\Omega))$, which implies (5-3). Therefore, we can assume that $\mathrm{Per}^\star_\sigma(E\cup B_R,\Omega)>\mathrm{Per}^\star_\sigma(E,\Omega)$. Then, by (1-14), which can be utilized here in view of (5-2), and using again the subadditivity of the (either classical or fractional) perimeter,

$$\begin{aligned}\Phi(\mathrm{Per}^\star_\sigma(E\cup B_R,\Omega))-\Phi(\mathrm{Per}^\star_\sigma(E,\Omega))&\leqslant L_Q\left|\mathrm{Per}^\star_\sigma(E\cup B_R,\Omega)-\mathrm{Per}^\star_\sigma(E,\Omega)\right|\\&\leqslant L_Q\,\mathrm{Per}^\star_\sigma(B_R,\Omega)\leqslant L_Q\,\mathrm{Per}_\sigma(B_R,\mathbb{R}^n)\leqslant C\,L_Q\,R^{n-\sigma}.\end{aligned}$$

This proves (5-3).

By (5-3) and (5-1) we obtain

$$\begin{aligned}C\,L_Q\,R^{n-\sigma}&\geqslant\int_{B_R}\big(|\nabla u(x)|^2-|\nabla u_R(x)|^2\big)\,dx\\&=\int_{B_R}(\nabla u(x)+\nabla u_R(x))\cdot(\nabla u(x)-\nabla u_R(x))\,dx\\&=\int_{B_R}\big(\nabla u(x)-\nabla u_R(x)+2\nabla u_R(x)\big)\cdot(\nabla u(x)-\nabla u_R(x))\,dx\\&=\int_{B_R}|\nabla u(x)-\nabla u_R(x)|^2\,dx+2\int_{B_R}\nabla u_R(x)\cdot(\nabla u(x)-\nabla u_R(x))\,dx\\&=\int_{B_R}|\nabla u(x)-\nabla u_R(x)|^2\,dx,\end{aligned}$$

where the last equality follows from the fact that u_R is harmonic in B_R. The desired result is thus established. □

Remark 5.2. From Lemma 5.1 it follows that the gradient of the minimizers locally belongs to the Campanato space $\mathcal{L}^{p,\lambda}$, with $p:=2$ and $\lambda:=n-\sigma$, and thus to the Morrey space $L^{2,n-\sigma}$. This and the Poincaré inequality would give that the minimizers belong to the Campanato space $\mathcal{L}^{2,n+2-\sigma}$, and thus to the Hölder space of continuous functions with exponents $\frac{1}{2}((n+2-\sigma)-n)=1-\frac{1}{2}\sigma$. In any case, in the forthcoming Section 6 we will provide an alternate approach to continuity results.

Estimate on the average of minimal solutions. Now we estimate the average in balls for minimal solutions:

Lemma 5.3. *Assume that* (1-14) *holds true. Let* (u,E) *be a minimal pair in* Ω, *with* $u\geqslant 0$ *a.e. in* Ω^c *and* $B_{R_o}(p)\Subset\Omega$. *Assume* $R\in(0,R_o]$ *and* $p\in\mathcal{U}_0$. *Then*

$$\frac{1}{|B_R(p)|}\int_{B_R(p)}u(x)\,dx\leqslant C\sqrt{L_Q}\,R^{1-\sigma/2}$$

for some $C>0$, *possibly depending on* R_o, n *and* σ, *and* L_Q *is the one introduced in* (1-14).

Proof. By (1-13), we can take a sequence p_k with

$$\lim_{k\to+\infty} u(p_k) = 0. \tag{5-4}$$

For any $r \in (0, R]$ and for any $k \in \mathbb{N}$, we define

$$\psi(r) := r^{-n} \int_{B_r(p)} u(x)\, dx \quad \text{and} \quad \psi_k(r) := r^{-n} \int_{B_r(p_k)} u(x)\, dx.$$

We observe that

$$\lim_{k\to+\infty} \psi_k(r) = \psi(r). \tag{5-5}$$

To check this, we let $\bar{R} > R_o$, with $B_{\bar{R}}(p) \Subset \Omega$ and we consider a continuous approximation of u in $L^1(B_{\bar{R}}(p))$. That is, we take continuous functions u_ϵ such that

$$\lim_{\epsilon\searrow 0} \int_{B_{\bar{R}}(p)} |u(x) - u_\epsilon(x)|\, dx = 0. \tag{5-6}$$

For large k, we have $B_r(p_k) \subseteq B_{\bar{R}}(p)$, and so

$$\begin{aligned}
r^n\, |\psi_k(r) - \psi(r)| &= \left| \int_{B_r(p_k)} u(x)\, dx - \int_{B_r(p)} u(x)\, dx \right| \\
&\leqslant \left| \int_{B_r(p_k)} u_\epsilon(x)\, dx - \int_{B_r(p)} u_\epsilon(x)\, dx \right| + 2 \int_{B_{\bar{R}}(p)} |u(x) - u_\epsilon(x)|\, dx \\
&= \left| \int_{B_r} \big(u_\epsilon(x + p_k) - u_\epsilon(x + p)\big)\, dx \right| + 2 \int_{B_{\bar{R}}(p)} |u(x) - u_\epsilon(x)|\, dx.
\end{aligned}$$

Hence, taking the limit in k and using the dominated convergence theorem, we get

$$\lim_{k\to+\infty} r^n\, |\psi_k(r) - \psi(r)| \leqslant 2 \int_{B_{\bar{R}}(p)} |u(x) - u_\epsilon(x)|\, dx.$$

Then, we take the limit in ϵ and we obtain (5-5) from (5-6), as desired.

Now, we recall that $u \geqslant 0$ a.e. in $\mathbb{R}^n$, thanks to Lemma 2.7. Thus, by Remark 2.6,

$$\psi_k(0) := \lim_{r\searrow 0} \psi_k(r) = u(p_k). \tag{5-7}$$

Furthermore, using polar coordinates,

$$\begin{aligned}
\psi_k'(r) &= \frac{d}{dr} \int_{B_1} u(p_k + ry)\, dy = \int_{B_1} \nabla u(p_k + ry) \cdot y\, dy \\
&= \int_0^1 \left[t^n \int_{S^{n-1}} \nabla u(p_k + rt\omega) \cdot \omega\, d\mathcal{H}^{n-1}(\omega) \right] dt = \int_0^1 \left[t^n \int_{\partial B_1} \partial_\nu u(p_k + rt\omega)\, d\mathcal{H}^{n-1}(\omega) \right] dt,
\end{aligned} \tag{5-8}$$

where ν is the exterior normal of B_1.

Now, for a fixed $k \in \mathbb{N}$, we use the notation of Lemma 5.1 for the harmonic replacement u_r in $B_r(p_k) \Subset \Omega$. For $\rho \in (0, r]$, we define $v_r(x) := u_r(p_k + \rho x)$ and we observe that, for any $x \in B_1$, we have $\Delta v_r(x) =$

$\rho^2 \Delta u_r(p_k + \rho x) = 0$, and so

$$0 = \int_{B_1} \Delta v_r(x)\,dx = \int_{\partial B_1} \partial_\nu v_r(\omega)\,d\mathcal{H}^{n-1}(\omega) = \rho \int_{\partial B_1} \partial_\nu u_r(p_k + \rho\omega)\,d\mathcal{H}^{n-1}(\omega).$$

We take $\rho := rt$ and we insert this into (5-8). In this way, we obtain

$$\psi_k'(r) = \int_0^1 \left[t^n \int_{\partial B_1} \big(\partial_\nu u(p_k + rt\omega) - \partial_\nu u_r(p_k + rt\omega)\big)\,d\mathcal{H}^{n-1}(\omega) \right] dt.$$

That is, switching from polar to Cartesian coordinates and making the change of variable $y := p_k + rx$,

$$\psi_k'(r) = \int_{B_1} x \cdot \big(\nabla u(p_k + rx) - \nabla u_r(p_k + rx)\big)\,dx = r^{-(n+1)} \int_{B_r(p_k)} (y - p_k) \cdot (\nabla u(y) - \nabla u_r(y))\,dy.$$

Hence, using the Hölder inequality and Lemma 5.1,

$$\psi_k'(r) \leqslant r^{-n} \int_{B_r(p_k)} |\nabla u(y) - \nabla u_r(y)|\,dy \leqslant C r^{-n/2} \left(\int_{B_r(p_k)} |\nabla u(y) - \nabla u_r(y)|^2\,dy \right)^{1/2} \leqslant C \sqrt{L_Q}\, r^{-\sigma/2}$$

for some $C > 0$. This and (5-7) give

$$\psi_k(R) - u(p_k) = \psi_k(R) - \psi_k(0) = \int_0^R \psi_k'(r)\,dr \leqslant C \sqrt{L_Q} \int_0^R r^{-\sigma/2} \leqslant C \sqrt{L_Q}\, R^{1-\sigma/2},$$

up to renaming constants. Hence, making use of (5-4) and (5-5), we find that

$$\psi(R) \leqslant C \sqrt{L_Q}\, R^{1-\sigma/2},$$

which is the desired claim. □

Completion of the proof of Theorem 1.3. We recall that $u \geqslant 0$ a.e. in $\mathbb{R}^n$, thanks to Lemma 2.7. In particular, u is subharmonic, thanks to Lemma 2.5, and thus

$$\frac{1}{|B_\rho|} \int_{B_\rho(x)} u(y)\,dy \geqslant u(x) \tag{5-9}$$

for small $\rho > 0$. Now we take $x \in \Omega$, with $|x|$ suitably small, and we define $R := |x|$. Notice that $B_R(x) \subseteq B_{2R}$ and therefore, since $u \geqslant 0$,

$$\int_{B_R(x)} u(y)\,dy \leqslant \int_{B_{2R}} u(y)\,dy. \tag{5-10}$$

In addition, by applying Lemma 5.3 in B_{2R}, we find that

$$\frac{1}{R^n} \int_{B_{2R}} u(y)\,dy \leqslant C \sqrt{L_Q}\, R^{1-\sigma/2}.$$

As a result, exploiting (5-9) and (5-10),

$$u(x) \leqslant \frac{C}{R^n} \int_{B_R(x)} u(y)\,dy \leqslant \frac{C}{R^n} \int_{B_{2R}} u(y)\,dy \leqslant C \sqrt{L_Q}\, R^{1-\sigma/2} = C \sqrt{L_Q}\, |x|^{1-\sigma/2},$$

up to renaming constants. This proves Theorem 1.3.

6. Proof of Corollary 1.4

To prove Corollary 1.4, it is useful to point out a strengthening of Lemma 2.4 in which one replaces the condition on the infimum with a pointwise condition (this refinement is possible by virtue of Theorem 1.3):

Lemma 6.1. *Let the assumptions of Corollary 1.4 hold true. Let (u, E) be a minimal pair in Ω, with $u \geqslant 0$. Let $U \Subset \Omega$ be an open set with $u > 0$ in U. Then u is harmonic in U.*

Proof. Let $U' \Subset U$ be open. The claim is proved if we show that u is harmonic in U'. To this aim, we claim that

$$\inf_{U'} u > 0. \tag{6-1}$$

We argue for a contradiction, assuming that this infimum is equal to 0. Then, recalling (1-13), we have that there exists $x_\star \in \overline{U}' \cap \mathcal{U}_0$. In particular, since $x_\star \in \overline{U}' \subset U$, we know that

$$u(x_\star) > 0. \tag{6-2}$$

On the other hand, by Theorem 1.3, for small y,

$$u(x_\star + y) \leqslant C\sqrt{L_Q}\,|y|^{1-\sigma/2}.$$

As a result, recalling Remark 2.6,

$$u(x_\star) = u^+(x_\star) = \lim_{\epsilon \searrow 0} \frac{1}{|B_\epsilon|} \int_{B_\epsilon} u^+(x_\star + y)\,dy \leqslant C\sqrt{L_Q} \lim_{\epsilon \searrow 0} \frac{1}{|B_\epsilon|} \int_{B_\epsilon} |y|^{1-\sigma/2}\,dy = 0.$$

This is in contradiction with (6-2) and so we have proved (6-1).

Then, in light of (6-1), we fall under the assumptions of Lemma 2.4, which in turn implies the desired claim. ☐

First we recall that $u \geqslant 0$ a.e. in $\mathbb{R}^n$, thanks to Lemma 2.7. Also we know that u is subharmonic in Ω (recall Lemma 2.5) and therefore, by the classical maximum principle,

$$u(x) \leqslant M \tag{6-3}$$

for any $x \in \Omega$. Also, we may suppose that

$$\text{there exists } q_o \in B_{3R/10} \text{ such that } u(q_o) = 0. \tag{6-4}$$

Indeed, if this does not hold, then u is harmonic in $B_{3R/10}$, due to Lemma 6.1, and thus

$$\sup_{B_{R/4}} |\nabla u| \leqslant \frac{C}{R} \sup_{B_{3R/10}} u \leqslant \frac{CM}{R}$$

for some $C > 0$, where we also used (6-3) in the latter inequality. This implies

$$|u(x) - u(y)| \leqslant \frac{CM}{R}|x - y| \leqslant \frac{CM}{R^{1-\sigma/2}}|x - y|^{1-\sigma/2},$$

which gives the desired result in this case.

Hence, from now on, we can suppose that (6-4) holds true. We fix $x \neq y \in B_{R/4}$ and we define $d(x)$ to be the distance from x to the set $\{u = 0\}$; we define $d(y)$ analogously. By (6-4), we know that $d(x)$, $d(y) \in \left[0, \frac{3}{5}R\right]$. We distinguish two cases:

Case 1: $|x-y| \geqslant \frac{1}{2}\max\{d(x),\, d(y)\}$.

Case 2: $|x-y| < \frac{1}{2}\max\{d(x),\, d(y)\}$.

First, we deal with Case 1. In this case, we use Theorem 1.3 and we have

$$|u(x)| \leqslant C\sqrt{L_Q}\,(d(x))^{1-\sigma/2} \quad \text{and} \quad |u(y)| \leqslant C\sqrt{L_Q}\,(d(y))^{1-\sigma/2}.$$

As a consequence,

$$|u(x)-u(y)| \leqslant |u(x)|+|u(y)| \leqslant C\sqrt{L_Q}\left((d(x))^{1-\sigma/2}+(d(y))^{1-\sigma/2}\right).$$

Then, the assumption of Case 1 implies

$$|u(x)-u(y)| \leqslant C\sqrt{L_Q}\,|x-y|^{1-\sigma/2},$$

up to renaming constants, which gives the desired result in this case.

Now we consider Case 2. In this case, up to exchanging x and y, we have

$$0 \leqslant 2|x-y| < d(x) = \max\{d(x),\, d(y)\} \tag{6-5}$$

and $u > 0$ in $B_{d(x)}(x)$. Then, by Lemma 6.1, we know that u is harmonic in $B_{d(x)}(x)$ and thus

$$\sup_{B_{9d(x)/10}(x)} |\nabla u| \leqslant \frac{C}{d(x)} \sup_{B_{d(x)}(x)} u \tag{6-6}$$

for some $C > 0$.

Now, we prove

$$\sup_{B_{d(x)}(x)} u \leqslant C\sqrt{L_Q}\,(d(x))^{1-\sigma/2} \tag{6-7}$$

for some $C > 0$. For this, take $\eta \in B_{d(x)}(x)$. By construction, there exists $\zeta \in \overline{B_{d(x)}(x)}$ such that $u(\zeta) = 0$. Accordingly, we have $|\eta-\zeta| \leqslant |\eta-x|+|x-\zeta| \leqslant 2d(x)$, and then, by Theorem 1.3,

$$u(\eta) \leqslant C\sqrt{L_Q}\,|\eta-\zeta|^{1-\sigma/2} \leqslant C\sqrt{L_Q}\,(d(x))^{1-\sigma/2},$$

up to renaming $C > 0$, and this establishes (6-7).

Thus, exploiting (6-6) and (6-7), and possibly renaming constants, we obtain that

$$\sup_{B_{9d(x)/10}(x)} |\nabla u| \leqslant C\sqrt{L_Q}\,(d(x))^{-\sigma/2}.$$

Notice now that $y \in B_{d(x)/2}(x) \subset B_{9d(x)/10}(x)$, thanks to (6-5); therefore

$$|u(x)-u(y)| \leqslant C\sqrt{L_Q}\,(d(x))^{-\sigma/2}\,|x-y| \leqslant C\sqrt{L_Q}\,|x-y|^{1-\sigma/2},$$

up to renaming constants. This establishes the desired result also in Case 2 and so the proof of Corollary 1.4 is now completed.

7. Proof of Theorem 1.5

The proof is based on a measure theoretic argument that was used, in different forms, in [Caffarelli et al. 2015; Dipierro and Valdinoci 2016], but unlike the proof in the existing literature, we cannot use here the scaling properties of the functional: namely, the existing proofs can always reduce to the unit ball, since the rescaled minimal pair is a minimal pair for the rescaled functional, whereas this procedure fails in our case (as stressed for instance by Theorem 1.1). For this reason, we need to perform a measure-theoretic argument which works at every scale. To this end, for any $r \in (0, R)$ we define

$$V(r) := |B_r \setminus E| \quad \text{and} \quad a(r) := \mathcal{H}^{n-1}((\partial B_r) \setminus E)$$

and we observe that

$$V(r) = \int_0^r a(t)\,dt; \tag{7-1}$$

see, e.g., formula (13.3) in [Maggi 2012].

The proof of Theorem 1.5 is by contradiction: we suppose that, for some $r_o \in \left(0, \frac{1}{2}R\right)$, we have

$$V(r_o) = |B_{r_o} \setminus E| \leqslant \delta r_o^n \tag{7-2}$$

and we derive a contradiction if $\delta > 0$ is sufficiently small. We recall that $u \geqslant 0$ a.e. in $\mathbb{R}^n$, due to Lemma 2.7, and we define

$$A := B_r \setminus E.$$

We observe that $(u, E \cup A)$ is admissible, since $(E \cup A)^c = E^c \cap A^c \subseteq E^c$. Then, by the minimality of (u, E), we obtain

$$0 \leqslant \mathcal{E}_\Omega(u, E \cup A) - \mathcal{E}_\Omega(u, E) = \Phi(\mathrm{Per}^\star_\sigma(E \cup A, \Omega)) - \Phi(\mathrm{Per}^\star_\sigma(E, \Omega)). \tag{7-3}$$

Now, by the subadditivity of the (either classical or fractional) perimeter (see, e.g., Proposition 3.38(d) in [Ambrosio et al. 2000] when $\sigma = 1$ and formula (3.1) in [Dipierro et al. 2013] when $\sigma \in (0, 1)$), we have

$$\begin{aligned}\mathrm{Per}^\star_\sigma(E \cup A, \Omega) &= \mathrm{Per}^\star_\sigma(E \cup B_r, \Omega) \leqslant \mathrm{Per}^\star_\sigma(E, \Omega) + \mathrm{Per}^\star_\sigma(B_r, \Omega)\\ &\leqslant \mathrm{Per}^\star_\sigma(E, \Omega) + \mathrm{Per}_\sigma(B_r, \mathbb{R}^n) \leqslant \mathrm{Per}^\star_\sigma(E, \Omega) + R^{n-\sigma}\,\mathrm{Per}_\sigma(B_1, \mathbb{R}^n).\end{aligned}$$

Then, both $\mathrm{Per}^\star_\sigma(E, \Omega)$ and $\mathrm{Per}^\star_\sigma(E \cup A, \Omega)$ are bounded by P, as defined in (1-18), and so they lie in the invertibility range of Φ, as prescribed by (1-19). This observation and (7-3) imply

$$\mathrm{Per}^\star_\sigma(E, \Omega) \leqslant \mathrm{Per}^\star_\sigma(E \cup A, \Omega). \tag{7-4}$$

Now we claim that

$$\mathrm{Per}_\sigma(E, \Omega) \leqslant \mathrm{Per}_\sigma(E \cup A, \Omega). \tag{7-5}$$

Indeed, if $\sigma \in (0, 1)$, then (7-5) is simply (7-4). If instead $\sigma = 1$, we notice that $E \setminus \bar{B}_r = (E \cup A) \setminus \bar{B}_r$ and so we use (2-2), (2-3) and (7-4) to obtain

$$\begin{aligned}0 &\leqslant \mathrm{Per}^\star_\sigma(E \cup A, \Omega) - \mathrm{Per}^\star_\sigma(E, \Omega)\\ &= \mathrm{Per}_\sigma(E \cup A, \bar{B}_r) - \mathrm{Per}_\sigma(E, \bar{B}_r) = \mathrm{Per}_\sigma(E \cup A, \Omega) - \mathrm{Per}_\sigma(E, \Omega),\end{aligned}$$

which establishes (7-5).

Now we use the (either classical or fractional) isoperimetric inequality in the whole of $\mathbb{R}^n$ (see, e.g., Theorem 3.46 in [Ambrosio et al. 2000] when $\sigma = 1$, and [Frank et al. 2008], or Corollary 25 in [Caffarelli and Valdinoci 2011] when $\sigma \in (0, 1)$); in this way, we have

$$(V(r))^{(n-\sigma)/n} = |B_r \setminus E|^{(n-\sigma)/n} = |A|^{(n-\sigma)/n} \leqslant C \operatorname{Per}_\sigma(A, \mathbb{R}^n) \tag{7-6}$$

for some $C > 0$.

Now we claim that, for a.e. $r \in (0, R)$,

$$\operatorname{Per}_\sigma(A, \mathbb{R}^n) \leqslant \begin{cases} C\, a(r) & \text{if } \sigma = 1, \\ C \int_0^r a(\rho)(r-\rho)^{-\sigma}\, d\rho & \text{if } \sigma \in (0, 1) \end{cases} \tag{7-7}$$

for some $C > 0$ (up to renaming C). First we prove (7-7) when $\sigma = 1$. For this, we write the perimeter of E in term of the Gauss–Green measure μ_E (see Remark 12.2 in [Maggi 2012]), we use the additivity of the measures on disjoint sets and we obtain

$$\begin{aligned} \operatorname{Per}(E, B_r) + \operatorname{Per}(E, \Omega \setminus \bar{B}_r) &= |\mu_E|(B_r) + |\mu_E|(\Omega \setminus \bar{B}_r) \\ &\leqslant |\mu_E|(B_r) + |\mu_E|(\Omega \setminus B_r) = |\mu_E|(\Omega) = \operatorname{Per}(E, \Omega). \end{aligned} \tag{7-8}$$

Now we prove that, for a.e. $r \in (0, R)$, we have

$$\mathcal{H}^{n-1}((\partial B_r) \setminus E) = \operatorname{Per}(B_r \setminus E, \Omega) - \operatorname{Per}(E, B_r). \tag{7-9}$$

For this scope, we make use of the property of the Gauss–Green measure with respect to the intersection with balls (see formula (15.14) in Lemma 15.12 of [Maggi 2012], applied here to the complement of E). In this way, we see

$$\begin{aligned} \mathcal{H}^{n-1}((\partial B_r) \setminus E) &= \mathcal{H}^{n-1}((\partial B_r) \cap E^c \cap \Omega) = \mathcal{H}^{n-1}\big|_{E^c \cap (\partial B_r)}(\Omega) \\ &= |\mu_{E^c \cap B_r}|(\Omega) - |\mu_{E^c}|\big|_{B_r}(\Omega) \\ &= \operatorname{Per}(E^c \cap B_r, \Omega) - |\mu_{E^c}|(B_r \cap \Omega) \\ &= \operatorname{Per}(E^c \cap B_r, \Omega) - |\mu_{E^c}|(B_r) \\ &= \operatorname{Per}(E^c \cap B_r, \Omega) - \operatorname{Per}(E^c, B_r). \end{aligned}$$

From this and the fact that $\operatorname{Per}(E^c, B_r) = \operatorname{Per}(E, B_r)$ (see, for instance, Proposition 3.38(d) in [Ambrosio et al. 2000]), we obtain that (7-9) holds true.

Now we claim that, for a.e. $r \in (0, R)$, we have

$$\operatorname{Per}(E \cup B_r, \bar{B}_r) = \mathcal{H}^{n-1}((\partial B_r) \setminus E). \tag{7-10}$$

Since it is not easy to find a complete reference for such formula in the literature, we try to give here an exhaustive proof. To this end, given a set F and $t \in [0, 1]$, we denote by $F^{(t)}$ the set of points of density t of F (see, e.g., Example 5.17 in [Maggi 2012]), that is,

$$F^{(t)} := \left\{ x \in \mathbb{R}^n : \lim_{r \to 0} \frac{|F \cap B_r(x)|}{|B_r|} = t \right\}.$$

With this notation, we observe that $B_r^{(0)} = \mathbb{R}^n \setminus \bar{B}_r$, and thus

$$B_r^{(0)} \cap \bar{B}_r = \varnothing. \tag{7-11}$$

We denote by ∂^* the reduced boundary of a set of locally finite perimeter (see, e.g., formula (15.1) in [Maggi 2012]); we recall that for any $x \in \partial^* E$ one can define the measure-theoretic outer unit normal to E, which we denote by ν_E. We also recall that, by De Giorgi's structure theorem (see, e.g., formula (15.10) in [Maggi 2012]),

$$|\mu_E| = \mathcal{H}^{n-1}\big|_{\partial^* E}. \tag{7-12}$$

We also set

$$N_r := \{x \in (\partial^* E) \cap (\partial B_r) : \nu_E = \nu_{B_r}\}.$$

We claim that, for a.e. $r \in (0, R)$,

$$\mathcal{H}^{n-1}(N_r) = 0. \tag{7-13}$$

To check this, for any $k \in \mathbb{N}$ we define

$$\beta_k := \left\{ r \in (0, R) : \mathcal{H}^{n-1}(N_r) \geqslant \frac{1}{k} \right\}.$$

Then, if $r \in \beta_k$, by (7-12) we have

$$|\mu_E|(\partial B_r) = \mathcal{H}^{n-1}\big|_{\partial^* E}(\partial B_r) = \mathcal{H}^{n-1}((\partial^* E) \cap (\partial B_r)) \geqslant \mathcal{H}^{n-1}(N_r) \geqslant \frac{1}{k}.$$

As a consequence, if $r_1, \dots, r_j \in \beta_k$ and $r \in (0, R)$, we obtain

$$\operatorname{Per}(E, B_R) = |\mu_E|(B_R) \geqslant |\mu_E| \left(\bigcup_{i=1}^{j} (\partial B_{r_i}) \right) = \sum_{i=1}^{j} |\mu_E|(\partial B_{r_i}) \geqslant \frac{j}{k},$$

that is, $j \leqslant k \operatorname{Per}(E, B_R)$.

This says that β_k has a finite (indeed less than $k \operatorname{Per}(E, B_R)$) number of elements. Thus the following set is countable (and so is of measure zero):

$$\bigcup_{k=1}^{+\infty} \beta_k = \{r \in (0, R) : \mathcal{H}^{n-1}(N_r) > 0\} = \{r \in (0, R) : \text{(7-13) does not hold}\}.$$

This proves (7-13).

Now we use the known formula about the perimeter of the union. For instance, exploiting formula (16.12) of [Maggi 2012] (used here with $F = B_r$ and $G := \bar{B}_r$) we have

$$\operatorname{Per}(E \cup B_r, \bar{B}_r) = \operatorname{Per}(E, B_r^{(0)} \cap \bar{B}_r) + \operatorname{Per}(B_r, E^{(0)} \cap \bar{B}_r) + \mathcal{H}^{n-1}(N_r \cap \bar{B}_r).$$

In particular, using (7-11) and (7-13), we obtain

$$\operatorname{Per}(E \cup B_r, \bar{B}_r) = \operatorname{Per}(B_r, E^{(0)} \cap \bar{B}_r) \tag{7-14}$$

for a.e. $r \in (0, R)$. On the other hand, B_r is a smooth set and so (see, e.g., Example 12.6 in [Maggi 2012]) we have

$$\mathrm{Per}(B_r, E^{(0)} \cap \bar{B}_r) = \mathcal{H}^{n-1}(E^{(0)} \cap \bar{B}_r \cap (\partial B_r)) = \mathcal{H}^{n-1}(E^{(0)} \cap (\partial B_r)),$$

and so (7-14) becomes

$$\mathrm{Per}(E \cup B_r, \bar{B}_r) = \mathcal{H}^{n-1}(E^{(0)} \cap (\partial B_r)). \tag{7-15}$$

Now we set

$$S := (E^{(0)} \setminus E^c) \cup (E^c \setminus E^{(0)})$$

and we remark that $|S| = 0$ (see, e.g., formula (5.19) in [Maggi 2012]). Then, also $|S \cap B_r| = 0$. Therefore (see, e.g., Remark 12.4 in [Maggi 2012]) we get that $\mathrm{Per}(S, \mathbb{R}^n) = 0 = \mathrm{Per}(S \cap B_r, \mathbb{R}^n)$ and then (see, e.g., formula (15.15) in [Maggi 2012]) for a.e. $r \in (0, R)$ we obtain

$$\mathcal{H}^{n-1}(S \cap (\partial B_r)) = \mathrm{Per}(S \cap B_r, \mathbb{R}^n) - \mathrm{Per}(S, B_r) = 0,$$

and so, as a consequence,

$$\mathcal{H}^{n-1}(E^{(0)} \cap (\partial B_r)) = \mathcal{H}^{n-1}(E^c \cap (\partial B_r)).$$

Now we combine this and (7-15) and we finally complete the proof of (7-10).

Now we show that, for a.e. $r \in (0, R)$,

$$\mathrm{Per}(E \cup B_r, \Omega) - \mathrm{Per}(E, \Omega \setminus \bar{B}_r) = \mathrm{Per}(B_r \setminus E, \Omega) - \mathrm{Per}(E, B_r). \tag{7-16}$$

To prove this, we notice that $(E \cup B_r) \setminus \bar{B}_r = E \setminus \bar{B}_r$, and so we use Lemma 2.2 to see

$$\mathrm{Per}(E \cup B_r, \Omega) - \mathrm{Per}(E, \Omega) = \mathrm{Per}(E \cup B_r, \bar{B}_r) - \mathrm{Per}(E, \bar{B}_r).$$

As a consequence,

$$\begin{aligned}\mathrm{Per}(E \cup B_r, \Omega) - \mathrm{Per}(E, \Omega \setminus \bar{B}_r) &= \mathrm{Per}(E \cup B_r, \bar{B}_r) - \mathrm{Per}(E, \bar{B}_r) + \mathrm{Per}(E, \Omega) - \mathrm{Per}(E, \Omega \setminus \bar{B}_r)\\ &= \mathrm{Per}(E \cup B_r, \bar{B}_r) - |\mu_E|(\bar{B}_r) + |\mu_E|(\Omega) - |\mu_E|(\Omega \setminus \bar{B}_r)\\ &= \mathrm{Per}(E \cup B_r, \bar{B}_r),\end{aligned}$$

thanks to the additivity of the Gauss–Green measure μ_E. Then, we use (7-10) and we obtain

$$\mathrm{Per}(E \cup B_r, \Omega) - \mathrm{Per}(E, \Omega \setminus \bar{B}_r) = \mathcal{H}^{n-1}((\partial B_r) \setminus E).$$

Then, we exploit (7-9) and we complete the proof of (7-16).

Now we observe that, using (7-9) and (7-16), we obtain, for a.e. $r \in (0, R)$,

$$\mathrm{Per}(E \cup B_r, \Omega) = \mathrm{Per}(E, \Omega \setminus \bar{B}_r) + \mathcal{H}^{n-1}((\partial B_r) \setminus E). \tag{7-17}$$

Now, putting together (7-8) and (7-17), and noticing that $E \cup B_r = E \cup A$, we have

$$\begin{aligned}\mathrm{Per}(E, B_r) &\leqslant \mathrm{Per}(E, \Omega) - \mathrm{Per}(E, \Omega \setminus \bar{B}_r)\\ &= \mathrm{Per}(E, \Omega) - \mathrm{Per}(E \cup B_r, \Omega) + \mathcal{H}^{n-1}((\partial B_r) \setminus E)\\ &= \mathrm{Per}(E, \Omega) - \mathrm{Per}(E \cup A, \Omega) + \mathcal{H}^{n-1}((\partial B_r) \setminus E).\end{aligned}$$

Therefore, recalling (7-5) (used here with $\sigma = 1$), we conclude that

$$\operatorname{Per}(E, B_r) \leqslant \mathcal{H}^{n-1}((\partial B_r) \setminus E). \tag{7-18}$$

Now we take $r' \in (r, R)$ and we observe that $B_r \Subset B_{r'} \Subset \Omega$. Also, we see that $A \setminus \bar{B}_{r'} = \varnothing$; thus, by Lemma 2.2 (applied here with $F := \varnothing$),

$$\operatorname{Per}(A, \mathbb{R}^n) = \operatorname{Per}(A, \bar{B}_{r'}) \leqslant \operatorname{Per}(A, \Omega) = \operatorname{Per}(B_r \setminus E, \Omega).$$

As a consequence of this and of (7-16), we obtain

$$\operatorname{Per}(A, \mathbb{R}^n) \leqslant \operatorname{Per}(E \cup B_r, \Omega) - \operatorname{Per}(E, \Omega \setminus \bar{B}_r) + \operatorname{Per}(E, B_r).$$

Hence, in light of (7-17) and (7-18),

$$\operatorname{Per}(A, \mathbb{R}^n) \leqslant 2\mathcal{H}^{n-1}((\partial B_r) \setminus E) = 2a(r).$$

This completes the proof of (7-7) when $\sigma = 1$.

When $\sigma \in (0, 1)$, to prove (7-7) we use a modification of the argument contained in formulas (5.8)–(5.12) in [Dipierro and Valdinoci 2016]. We first observe that

$$\operatorname{Per}_\sigma(E, \Omega) - \operatorname{Per}_\sigma(E \cup A, \Omega) = L(A, E) - L(A, (E \cup A)^c).$$

As a consequence,

$$\begin{aligned}\operatorname{Per}_\sigma(A, \mathbb{R}^n) &= L(A, A^c) = L(A, E) + L(A, (E \cup A)^c)\\ &= 2L(A, (E \cup A)^c) + \operatorname{Per}_\sigma(E, \Omega) - \operatorname{Per}_\sigma(E \cup A, \Omega).\end{aligned}$$

This and (7-5) give

$$\operatorname{Per}_\sigma(A, \mathbb{R}^n) \leqslant 2L(A, (E \cup A)^c) \leqslant 2L(A, B_r^c). \tag{7-19}$$

Now we recall that $A \subseteq B_r$ and so, using the change of coordinates $\zeta := x - y$, we obtain

$$\begin{aligned}L(A, B_r^c) &= \int_{A \times B_r^c} \frac{dx\,dy}{|x - y|^{n+\sigma}} \leqslant \int_{\{(x,\zeta)\in A\times\mathbb{R}^n : |\zeta| \geqslant r - |x|\}} \frac{dx\,d\zeta}{|\zeta|^{n+\sigma}}\\ &\leqslant C \int_A \left[\int_{r-|x|}^{+\infty} \frac{\rho^{n-1}\,d\rho}{\rho^{n+\sigma}} \right] dx \leqslant C \int_A \frac{dx}{(r - |x|)^\sigma}.\end{aligned} \tag{7-20}$$

Now we use the coarea formula (see, e.g., Theorem 2 on page 117 of [Evans and Gariepy 1992], applied here in codimension 1 to the functions $f(x) = |x|$ and $g(x) := \chi_A(x)/(r - |x|)^\sigma$), and we deduce that

$$\begin{aligned}\int_A \frac{dx}{(r - |x|)^\sigma} &= \int_{\mathbb{R}} \left[\int_{\partial B_t} \frac{\chi_A(x)}{(r - |x|)^\sigma}\, d\mathcal{H}^{n-1}(x) \right] dt\\ &= \int_0^r \left[\int_{\partial B_t} \frac{\chi_{E^c}(x)}{(r - t)^\sigma}\, d\mathcal{H}^{n-1}(x) \right] dt = \int_0^r \frac{\mathcal{H}^{n-1}(E^c \cap (\partial B_t))}{(r - t)^\sigma}\, dt = \int_0^r \frac{a(t)}{(r - t)^\sigma}\, dt.\end{aligned}$$

This and (7-20) imply

$$L(A, B_r^c) \leqslant C \int_0^r \frac{a(t)}{(r - t)^\sigma}\, dt.$$

Inserting this into (7-19) we get

$$\operatorname{Per}_\sigma(A, \mathbb{R}^n) \leqslant C \int_0^r \frac{a(t)}{(r-t)^\sigma}\, dt,$$

which gives the desired claim in (7-7) when $\sigma \in (0, 1)$.

Using (7-6) and (7-7), and possibly renaming constants, we conclude that, for a.e. $r \in (0, R)$,

$$(V(r))^{(n-\sigma)/n} \leqslant \begin{cases} C\, a(r) & \text{if } \sigma = 1, \\ C \int_0^r a(\rho)(r-\rho)^{-\sigma}\, d\rho & \text{if } \sigma \in (0, 1). \end{cases} \tag{7-21}$$

Our next goal is to show that, for any $t \in \left[\frac{1}{4}, \frac{1}{2}\right]$, we have

$$\int_{r_o/4}^{tr_o} (V(r))^{(n-\sigma)/n}\, dr \leqslant C t^{1-\sigma} r_o^{1-\sigma}\, V(tr_o) \tag{7-22}$$

for some $C > 0$. To prove this, we integrate (7-21) in $r \in \left[\frac{1}{4} r_o, tr_o\right]$. Then, when $\sigma = 1$, we obtain (7-22) directly from (7-1). If instead $\sigma \in (0, 1)$, we obtain

$$\begin{aligned} \int_{r_o/4}^{tr_o} (V(r))^{(n-\sigma)/n}\, dr &\leqslant C \int_{r_o/4}^{tr_o} \left[\int_0^r a(\rho)(r-\rho)^{-\sigma}\, d\rho \right] dr \\ &\leqslant C \int_0^{tr_o} \left[\int_\rho^{tr_o} a(\rho)(r-\rho)^{-\sigma}\, dr \right] d\rho = \frac{C}{1-\sigma} \int_0^{tr_o} a(\rho)(tr_o - \rho)^{1-\sigma}\, d\rho \\ &\leqslant \frac{C}{1-\sigma} \int_0^{tr_o} a(\rho)(tr_o)^{1-\sigma}\, d\rho = \frac{C\,(tr_o)^{1-\sigma}}{1-\sigma}\, V(tr_o), \end{aligned}$$

where we used (7-1) in the last identity. This completes the proof of (7-22), up to renaming the constants.

Now we define $t_k := \frac{1}{4} + \frac{1}{2^k}$ for any $k \geqslant 2$. Let also $w_k := r_o^{-n} V(t_k r_o)$. Notice that $t_{k+1} \geqslant \frac{1}{4}$. Then we use (7-22) with $t := t_k$ and we obtain

$$C t_k^{1-\sigma} r_o^{1-\sigma}\, V(t_k r_o) \geqslant \int_{r_o/4}^{t_k r_o} (V(r))^{(n-\sigma)/n}\, dr \geqslant \int_{t_{k+1} r_o}^{t_k r_o} (V(r))^{(n-\sigma)/n}\, dr.$$

Thus, since $V(\cdot)$ is monotone,

$$C t_k^{1-\sigma} r_o^{1-\sigma}\, V(t_k r_o) \geqslant (t_k r_o - t_{k+1} r_o)(V(t_{k+1} r_o))^{(n-\sigma)/n} = \frac{r_o}{2^{k+1}}\, (V(t_{k+1} r_o))^{(n-\sigma)/n}.$$

This can be written as

$$w_{k+1}^{(n-\sigma)/n} = r_o^{\sigma-n}\, (V(t_{k+1} r_o))^{(n-\sigma)/n} \leqslant 2^{k+1} C\, t_k^{1-\sigma} r_o^{-n} V(t_k r_o) = 2^{k+1} C\, t_k^{1-\sigma} w_k.$$

Consequently, using that $t_k \leqslant 1$ and possibly renaming $C > 0$, we obtain

$$w_{k+1}^{(n-\sigma)/n} \leqslant C^k w_k. \tag{7-23}$$

Also, we have $t_2 = \frac{1}{2}$ and thus

$$w_2 = r_o^{-n} V\left(\tfrac{1}{2} r_o\right) \leqslant r_o^{-n} V(r_o) \leqslant \delta,$$

in view of (7-2). Then, if $\delta > 0$ is sufficiently small, we have $w_k \to 0$ as $k \to +\infty$ (see, e.g., formula (8.18) in [Dipierro et al. 2014] for explicit bounds). This and the fact that $t_k \geqslant \frac{1}{4}$ say that

$$0 = \lim_{k\to+\infty} r_o^{-n} V(t_k r_o) = \lim_{k\to+\infty} r_o^{-n} |B_{t_k r_o} \setminus E| \geqslant r_o^{-n} |B_{r_o/4} \setminus E|.$$

Hence, we have $|B_{r_o/4} \setminus E| = 0$, in contradiction with the assumption that $0 \in \partial E$ (in the measure-theoretic sense). The proof of Theorem 1.5 is thus complete.

8. Proof of Theorem 1.6

By Lemma 2.7, we have

$$u \geqslant 0 \quad \text{a.e. in } \mathbb{R}^n. \tag{8-1}$$

For any $r \in (0, R)$ we define

$$V(r) := |B_r \cap E| \quad \text{and} \quad a(r) := \mathcal{H}^{n-1}((\partial B_r) \cap E),$$

and we observe that

$$V(r) = \int_0^r a(t)\, dt; \tag{8-2}$$

see, e.g., formula (13.3) in [Maggi 2012].

The proof of Theorem 1.6 is obtained by a contradiction argument. Namely, we suppose that, for some $r_o \in \left(0, \frac{1}{2}R\right)$ we have

$$V(r_o) = |B_{r_o} \cap E| \leqslant \delta_* r_o^n, \tag{8-3}$$

and we derive a contradiction if $\delta_* > 0$ is sufficiently small.

We let $A := B_r \cap E$. Let also $\tilde{v}$ be the minimizer of the Dirichlet energy in B_{r_o} among all the possible candidates $v : \mathbb{R}^n \to \mathbb{R}$ such that $v = u$ outside B_{r_o}, $v - u \in H_0^1(B_{r_o})$ and $v = 0$ a.e. in $E^c \cup A$ (for the existence and the uniqueness of such harmonic replacement see, e.g., page 481 in [Athanasopoulos et al. 2001]). By (8-1) and Lemma 2.3 in [Athanasopoulos et al. 2001] we have

$$\tilde{v} \geqslant 0 \quad \text{a.e. in } \mathbb{R}^n. \tag{8-4}$$

Now we set $F := E \setminus A$. We observe that $\tilde{v} = 0$ a.e. in $F^c = E^c \cup A$ by construction. This and (8-4) give that $(\tilde{v}, F)$ is an admissible pair, and recall also that $\tilde{v} - u \in H_0^1(B_{r_o}) \subseteq H_0^1(\Omega)$. Hence, the minimality of (u, E) gives

$$0 \leqslant \mathcal{E}_\Omega(\tilde{v}, F) - \mathcal{E}_\Omega(u, E) = \int_\Omega |\nabla \tilde{v}(x)|^2\, dx - \int_\Omega |\nabla u(x)|^2\, dx + \Phi(\mathrm{Per}^\star_\sigma(F, \Omega)) - \Phi(\mathrm{Per}^\star_\sigma(E, \Omega)).$$

Using this and the fact that $\tilde{v}$ and u coincide outside B_{r_o}, we obtain

$$\Phi(\mathrm{Per}^\star_\sigma(E, \Omega)) - \Phi(\mathrm{Per}^\star_\sigma(F, \Omega)) \leqslant \int_{B_{r_o}} |\nabla \tilde{v}(x)|^2\, dx - \int_{B_{r_o}} |\nabla u(x)|^2\, dx. \tag{8-5}$$

Now we take $\tilde{w}$ to be the minimizer of the Dirichlet energy in B_{r_o} among all the functions $w:\mathbb{R}^n\to\mathbb{R}$ such that $w=u$ outside B_{r_o}, $w-u\in H^1_0(B_{r_o})$ and $w=0$ a.e. in E^c. We remark that u is a competitor with such $\tilde{w}$ and therefore

$$\int_{B_{r_o}} |\nabla\tilde{w}(x)|^2\,dx \leqslant \int_{B_{r_o}} |\nabla u(x)|^2\,dx.$$

Plugging this into (8-5), we deduce that

$$\Phi(\mathrm{Per}^\star_\sigma(E,\Omega))-\Phi(\mathrm{Per}^\star_\sigma(F,\Omega)) \leqslant \int_{B_{r_o}} |\nabla\tilde{v}(x)|^2\,dx - \int_{B_{r_o}} |\nabla\tilde{w}(x)|^2\,dx.$$

This and Lemma 2.3 in [Caffarelli et al. 2015] imply

$$\Phi(\mathrm{Per}^\star_\sigma(E,\Omega))-\Phi(\mathrm{Per}^\star_\sigma(F,\Omega)) \leqslant C\,{r_o}^{-2}\,|A|\,\|\tilde{w}\|^2_{L^\infty(B_{r_o})}. \tag{8-6}$$

Since, by Lemma 2.3 in [Athanasopoulos et al. 2001], we know that $\tilde{w}\geqslant 0$ a.e. in $\mathbb{R}^n$ and is subharmonic, we have that w in B_{r_o} takes its maximum along ∂B_{r_o}, where it coincides with u. Hence

$$\|\tilde{w}\|_{L^\infty(B_{r_o})} \leqslant \sup_{\partial B_{r_o}} u. \tag{8-7}$$

Now we observe that condition (1-20) allows us to use Theorem 1.3, which gives

$$\sup_{\partial B_{r_o}} u \leqslant C\,\sqrt{L_Q}\,r_o^{1-\sigma/2}$$

for some $C>0$. Hence (8-7) gives

$$\|\tilde{w}\|_{L^\infty(B_{r_o})} \leqslant C\,\sqrt{L_Q}\,{r_o}^{1-\sigma/2}.$$

Thus, recalling (8-6), and possibly renaming constants, we conclude that

$$\Phi(\mathrm{Per}^\star_\sigma(E,\Omega))-\Phi(\mathrm{Per}^\star_\sigma(F,\Omega)) \leqslant C\,{r_o}^{-\sigma}\,|A|\,L_Q. \tag{8-8}$$

Now we claim that

$$\mathrm{Per}_\sigma(E,\Omega)-\mathrm{Per}_\sigma(F,\Omega) \leqslant C\,c_o^{-1}\,{r_o}^{-\sigma}\,|A|\,L_Q, \tag{8-9}$$

where $c_o>0$ is the one introduced in (1-21). To check this, we may suppose that $\lambda_1:=\mathrm{Per}_\sigma(E,\Omega)>\mathrm{Per}_\sigma(F,\Omega)=:\lambda_2$, otherwise we are done. Then, by (1-22), both λ_1 and λ_2 belong to $[0,Q]$; therefore we can make use of (1-21) and obtain

$$\begin{aligned}\Phi(\mathrm{Per}^\star_\sigma(E,\Omega))-\Phi(\mathrm{Per}^\star_\sigma(F,\Omega)) &= \Phi(\lambda_1)-\Phi(\lambda_2)\\ &= \int_{\lambda_2}^{\lambda_1}\Phi'(t)\,dt \geqslant c_o(\lambda_1-\lambda_2) = c_o\big(\mathrm{Per}^\star_\sigma(E,\Omega)-\mathrm{Per}^\star_\sigma(F,\Omega)\big),\end{aligned}$$

and then it follows from (8-8) that

$$\mathrm{Per}^\star_\sigma(E,\Omega)-\mathrm{Per}^\star_\sigma(F,\Omega) \leqslant C\,c_o^{-1}\,{r_o}^{-\sigma}\,|A|\,L_Q. \tag{8-10}$$

Now we observe that $E \setminus \bar{B}_r = F \setminus \bar{B}_r$; therefore, using (2-2) and (2-3), we see that

$$\mathrm{Per}^\star_\sigma(E, \Omega) - \mathrm{Per}^\star_\sigma(F, \Omega) = \mathrm{Per}_\sigma(E, \bar{B}_r) - \mathrm{Per}_\sigma(F, \bar{B}_r) = \mathrm{Per}_\sigma(E, \Omega) - \mathrm{Per}_\sigma(F, \Omega).$$

Putting together this and (8-10) we obtain (8-9).

Now we show that, for a.e. $r \in (0, r_o)$,

$$\mathrm{Per}_\sigma(A, \mathbb{R}^n) \leqslant \begin{cases} C\big(a(r) + c_o^{-1} r_o{}^{-\sigma} |A| L_Q\big) & \text{if } \sigma = 1, \\ C\big(\int_0^r a(\rho)(r-\rho)^{-\sigma}\, d\rho + c_o^{-1} r_o{}^{-\sigma} |A| L_Q\big) & \text{if } \sigma \in (0,1). \end{cases} \tag{8-11}$$

To prove (8-11) we distinguish the cases $\sigma = 1$ and $\sigma \in (0, 1)$. If $\sigma = 1$, we notice that $A \setminus \bar{B}_r = (B_r \cap E) \setminus \bar{B}_r = \varnothing$; hence, by Lemma 2.2, we have

$$\mathrm{Per}(A, \mathbb{R}^n) = \mathrm{Per}(A, \bar{B}_r) = \mathrm{Per}(E \cap B_r, \bar{B}_r).$$

Hence we use the formula for the perimeter associated with the intersection with balls (see, e.g., (15.14) in Lemma 15.12 of [Maggi 2012]) and we obtain

$$\begin{aligned} \mathrm{Per}(A, \mathbb{R}^n) &= |\mu_{E \cap B_r}|(\bar{B}_r) = \mathcal{H}^{n-1}\big|_{E \cap (\partial B_r)}(\bar{B}_r) + |\mu_E|\big|_{B_r}(\bar{B}_r) \\ &= \mathcal{H}^{n-1}(E \cap (\partial B_r) \cap \bar{B}_r) + \mathrm{Per}(E, B_r \cap \bar{B}_r) \\ &= \mathcal{H}^{n-1}(E \cap (\partial B_r)) + \mathrm{Per}(E, B_r). \end{aligned} \tag{8-12}$$

On the other hand, we have $(E \setminus B_r)^c = E^c \cup B_r$; hence (see, e.g., formula (16.11) in [Maggi 2012]) we obtain that $\mathrm{Per}(E \setminus B_r, \bar{B}_r) = \mathrm{Per}(E^c \cup B_r, \bar{B}_r)$ for a.e. $r \in (0, r_o)$. Hence, by Lemma 2.2,

$$\begin{aligned} \mathrm{Per}(E, \Omega) - \mathrm{Per}(F, \Omega) &= \mathrm{Per}(E, \bar{B}_r) - \mathrm{Per}(F, \bar{B}_r) \\ &= \mathrm{Per}(E, \bar{B}_r) - \mathrm{Per}(E \setminus B_r, \bar{B}_r) = \mathrm{Per}(E, \bar{B}_r) - \mathrm{Per}(E^c \cup B_r, \bar{B}_r) \end{aligned} \tag{8-13}$$

for a.e. $r \in (0, r_o)$. Moreover (see, e.g., formula (7-10), applied here to the complementary set), we have

$$\mathrm{Per}(E^c \cup B_r, \bar{B}_r) = \mathcal{H}^{n-1}((\partial B_r) \cap E),$$

so we can write (8-13) as

$$\mathrm{Per}(E, \Omega) - \mathrm{Per}(F, \Omega) = \mathrm{Per}(E, \bar{B}_r) - \mathcal{H}^{n-1}((\partial B_r) \cap E).$$

In particular

$$\mathrm{Per}(E, B_r) \leqslant \mathrm{Per}(E, \bar{B}_r) = \mathrm{Per}(E, \Omega) - \mathrm{Per}(F, \Omega) + \mathcal{H}^{n-1}((\partial B_r) \cap E).$$

Then we insert this information into (8-12) and we obtain

$$\mathrm{Per}(A, \mathbb{R}^n) \leqslant 2\mathcal{H}^{n-1}(E \cap (\partial B_r)) + \mathrm{Per}(E, \Omega) - \mathrm{Per}(F, \Omega).$$

Now we recall (8-9), which completes the proof of (8-11) when $\sigma = 1$, and focus on the case $\sigma \in (0, 1)$. For this, we use (1-1) and we see that

$$\mathrm{Per}_\sigma(E, \Omega) - \mathrm{Per}_\sigma(F, \Omega) = \mathrm{Per}_\sigma(E, \Omega) - \mathrm{Per}_\sigma(E \setminus A, \Omega) = L(A, E^c) - L(A, E \setminus A).$$

Therefore

$$\mathrm{Per}_\sigma(A,\mathbb{R}^n)=L(A,A^c)=L(A,E^c)+L(A,E\setminus A)=\mathrm{Per}_\sigma(E,\Omega)-\mathrm{Per}_\sigma(F,\Omega)+2L(A,E\setminus A). \quad (8\text{-}14)$$

Now we use the fact that $A\subseteq B_r$ and the change of coordinates $\zeta:=x-y$ to write

$$\begin{aligned} L(A,E\setminus A)\leqslant L(A,B_r^c)&=\int_{A\times B_r^c}\frac{dx\,dy}{|x-y|^{n+\sigma}}\leqslant\int_{\{(x,\zeta)\in A\times\mathbb{R}^n:|\zeta|\geqslant r-|x|\}}\frac{dx\,d\zeta}{|\zeta|^{n+\sigma}}\\ &\leqslant C\int_A\left[\int_{r-|x|}^{+\infty}\frac{\rho^{n-1}\,d\rho}{\rho^{n+\sigma}}\right]dx\leqslant C\int_A\frac{dx}{(r-|x|)^\sigma}. \end{aligned} \quad (8\text{-}15)$$

Now we observe that, by the coarea formula (see, e.g., Theorem 2 on page 117 of [Evans and Gariepy 1992], applied here in codimension 1 to the functions $f(x)=|x|$ and $g(x):=\chi_A(x)/(r-|x|)^\sigma$),

$$\begin{aligned}\int_A\frac{dx}{(r-|x|)^\sigma}&=\int_{\mathbb{R}}\left[\int_{\partial B_t}\frac{\chi_A(x)}{(r-|x|)^\sigma}\,d\mathcal{H}^{n-1}(x)\right]dt\\ &=\int_0^r\left[\int_{\partial B_t}\frac{\chi_E(x)}{(r-t)^\sigma}\,d\mathcal{H}^{n-1}(x)\right]dt=\int_0^r\frac{\mathcal{H}^{n-1}(E\cap(\partial B_t))}{(r-t)^\sigma}\,dt=\int_0^r\frac{a(t)}{(r-t)^\sigma}\,dt.\end{aligned}$$

This and (8-15) give

$$L(A,E\setminus A)\leqslant C\int_0^r\frac{a(t)}{(r-t)^\sigma}\,dt.$$

So we substitute this and (8-9) into (8-14) and we complete the proof of (8-11) when $\sigma\in(0,1)$.

Now we recall that $|A|=V(r)$ and we use the (either classical or fractional) isoperimetric inequality in the whole of $\mathbb{R}^n$ (see, e.g., Theorem 3.46 in [Ambrosio et al. 2000] when $\sigma=1$, and [Frank et al. 2008], or Corollary 25 in [Caffarelli and Valdinoci 2011] when $\sigma\in(0,1)$) and we deduce from (8-11) that, for a.e. $r\in(0,r_o)$,

$$(V(r))^{(n-\sigma)/n}=|A|^{(n-\sigma)/n}\leqslant\begin{cases}C\big(a(r)+c_o^{-1}r_o{}^{-\sigma}\,V(r)\,L_Q\big) & \text{if } \sigma=1,\\ C\big(\int_0^r a(\rho)(r-\rho)^{-\sigma}\,d\rho+c_o^{-1}r_o{}^{-\sigma}\,V(r)\,L_Q\big) & \text{if } \sigma\in(0,1),\end{cases} \quad (8\text{-}16)$$

up to renaming $C>0$. Now we recall (8-3) and we notice that, if $r\in(0,r_o)$,

$$c_o^{-1}r_o{}^{-\sigma}\,V(r)\,L_Q\leqslant c_o^{-1}r_o{}^{-\sigma}(V(r))^{(n-\sigma)/n}(V(r_o))^{\sigma/n}L_Q\leqslant\delta_*^{\sigma/n}c_o^{-1}(V(r))^{(n-\sigma)/n}L_Q.$$

This means that, if $\delta_*>0$ is small enough, or more precisely if

$$\delta_*^{\sigma/n}c_o^{-1}L_Q\leqslant\frac{1}{2C}, \quad (8\text{-}17)$$

we can reabsorb[4] one term in the left-hand side of (8-16): in this way, possibly renaming constants, we obtain that, for a.e. $r\in(0,r_o)$,

$$(V(r))^{(n-\sigma)/n}\leqslant\begin{cases}C\,a(r) & \text{if } \sigma=1,\\ C\int_0^r a(\rho)(r-\rho)^{-\sigma}\,d\rho & \text{if } \sigma\in(0,1).\end{cases}$$

[4]It is interesting to point out that the possibility of absorbing the term $C\,c_o^{-1}\,r_o{}^{-\sigma}\,V(r)\,L_Q$ into the left-hand side of (8-16) crucially depends on the fact that the power produced by the (either classical or fractional) isoperimetric inequality and the one given by the growth result in Theorem 1.3 match together in the appropriate way.

This implies that, for any $t \in \left[\frac{1}{4}, \frac{1}{2}\right]$, we have

$$\int_{r_o/4}^{tr_o} (V(r))^{(n-\sigma)/n}\, dr \leqslant C t^{1-\sigma} r_o^{1-\sigma} V(tr_o) \tag{8-18}$$

for some $C > 0$. Indeed, the proof of (8-18) is obtained in the same way as that of (7-22) (the only difference is that here one has to use (8-2) in lieu of (7-1)). Then, one defines $t_k := \frac{1}{4} + \frac{1}{2^k}$ and $w_k := r_o^{-n} V(t_k r_o)$ and observes that

$$w_{k+1}^{(n-\sigma)/n} \leqslant C^k w_k. \tag{8-19}$$

Indeed, (8-19) can be obtained as in the proof of (7-23) (but using here (8-18) instead of (7-22)). Furthermore

$$w_2 = r_o^{-n} V\left(\tfrac{1}{2} r_o\right) \leqslant \delta_*,$$

thanks to (8-3). This says that

$$\text{if } \delta_* > 0 \text{ is sufficiently small (with respect to a universal constant),} \tag{8-20}$$

then $w_k \to 0$ as $k \to +\infty$ (see formula (8.18) in [Dipierro et al. 2014] for explicit bounds). Thus

$$0 = \lim_{k \to +\infty} r_o^{-n} V(t_k r_o) = \lim_{k \to +\infty} r_o^{-n} |B_{t_k r_o} \cap E| \geqslant r_o^{-n} |B_{r_o/4} \cap E|.$$

This is in contradiction with the assumption that $0 \in \partial E$ (in the measure-theoretic sense) and so the proof of Theorem 1.6 is finished. We stress that the explicit condition in (1-23) comes from (8-17) and (8-20).

Acknowledgements

It is a pleasure to thank Francesco Maggi for an interesting discussion and the School of Mathematics of the University of Edinburgh for the warm hospitality.

This work has been supported by EPSRC grant EP/K024566/1 (Monotonicity Formula Methods for Nonlinear PDEs), Humboldt Foundation, ERC grant 277749 (EPSILON: Elliptic Pde's and Symmetry of Interfaces and Layers for Odd Nonlinearities) and PRIN grant 2012 (Critical Point Theory and Perturbative Methods for Nonlinear Differential Equations).

References

[Abatangelo and Valdinoci 2014] N. Abatangelo and E. Valdinoci, "A notion of nonlocal curvature", *Numer. Funct. Anal. Optim.* **35**:7-9 (2014), 793–815. MR Zbl

[Alt and Caffarelli 1981] H. W. Alt and L. A. Caffarelli, "Existence and regularity for a minimum problem with free boundary", *J. Reine Angew. Math.* **325** (1981), 105–144. MR Zbl

[Ambrosio et al. 2000] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, New York, 2000. MR Zbl

[Ambrosio et al. 2011] L. Ambrosio, G. De Philippis, and L. Martinazzi, "Gamma-convergence of nonlocal perimeter functionals", *Manuscripta Math.* **134**:3-4 (2011), 377–403. MR Zbl

[Athanasopoulos et al. 2001] I. Athanasopoulos, L. A. Caffarelli, C. Kenig, and S. Salsa, "An area-Dirichlet integral minimization problem", *Comm. Pure Appl. Math.* **54**:4 (2001), 479–499. MR Zbl

[Bourgain et al. 2001] J. Bourgain, H. Brezis, and P. Mironescu, "Another look at Sobolev spaces", pp. 439–455 in *Optimal control and partial differential equations*, edited by J. L. Menaldi et al., IOS, Amsterdam, 2001. MR Zbl

[Caffarelli and Valdinoci 2011] L. Caffarelli and E. Valdinoci, "Uniform estimates and limiting arguments for nonlocal minimal surfaces", *Calc. Var. Partial Differential Equations* **41**:1-2 (2011), 203–240. MR Zbl

[Caffarelli et al. 2009] L. A. Caffarelli, A. L. Karakhanyan, and F.-H. Lin, "The geometry of solutions to a segregation problem for nondivergence systems", *J. Fixed Point Theory Appl.* **5**:2 (2009), 319–351. MR Zbl

[Caffarelli et al. 2010] L. Caffarelli, J.-M. Roquejoffre, and O. Savin, "Nonlocal minimal surfaces", *Comm. Pure Appl. Math.* **63**:9 (2010), 1111–1144. MR Zbl

[Caffarelli et al. 2015] L. Caffarelli, O. Savin, and E. Valdinoci, "Minimization of a fractional perimeter-Dirichlet integral functional", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **32**:4 (2015), 901–924. MR Zbl

[Dávila 2002] J. Dávila, "On an open question about functions of bounded variation", *Calc. Var. Partial Differential Equations* **15**:4 (2002), 519–527. MR Zbl

[Di Castro et al. 2015] A. Di Castro, M. Novaga, B. Ruffini, and E. Valdinoci, "Nonlocal quantitative isoperimetric inequalities", *Calc. Var. Partial Differential Equations* **54**:3 (2015), 2421–2464. MR Zbl

[Di Nezza et al. 2012] E. Di Nezza, G. Palatucci, and E. Valdinoci, "Hitchhiker's guide to the fractional Sobolev spaces", *Bull. Sci. Math.* **136**:5 (2012), 521–573. MR Zbl

[Dipierro and Valdinoci 2015] S. Dipierro and E. Valdinoci, "On a fractional harmonic replacement", *Discrete Contin. Dyn. Syst.* **35**:8 (2015), 3377–3392. MR Zbl

[Dipierro and Valdinoci 2016] S. Dipierro and E. Valdinoci, "Continuity and density results for a one-phase nonlocal free boundary problem", *Ann. Inst. H. Poincaré Anal. Non Linéaire* (online publication November 2016).

[Dipierro et al. 2013] S. Dipierro, A. Figalli, G. Palatucci, and E. Valdinoci, "Asymptotics of the s-perimeter as $s \searrow 0$", *Discrete Contin. Dyn. Syst.* **33**:7 (2013), 2777–2790. MR Zbl

[Dipierro et al. 2014] S. Dipierro, A. Figalli, and E. Valdinoci, "Strongly nonlocal dislocation dynamics in crystals", *Comm. Partial Differential Equations* **39**:12 (2014), 2351–2387. MR Zbl

[Dipierro et al. 2017] S. Dipierro, O. Savin, and E. Valdinoci, "Boundary behavior of nonlocal minimal surfaces", *J. Funct. Anal.* **272**:5 (2017), 1791–1851. MR Zbl

[Evans and Gariepy 1992] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992. MR Zbl

[Figalli et al. 2015] A. Figalli, N. Fusco, F. Maggi, V. Millot, and M. Morini, "Isoperimetry and stability properties of balls with respect to nonlocal energies", *Comm. Math. Phys.* **336**:1 (2015), 441–507. MR Zbl

[Frank et al. 2008] R. L. Frank, E. H. Lieb, and R. Seiringer, "Hardy–Lieb–Thirring inequalities for fractional Schrödinger operators", *J. Amer. Math. Soc.* **21**:4 (2008), 925–950. MR Zbl

[Garofalo and Lin 1986] N. Garofalo and F.-H. Lin, "Monotonicity properties of variational integrals, A_p weights and unique continuation", *Indiana Univ. Math. J.* **35**:2 (1986), 245–268. MR Zbl

[Giaquinta 1983] M. Giaquinta, *Multiple integrals in the calculus of variations and nonlinear elliptic systems*, Annals of Mathematics Studies **105**, Princeton University Press, 1983. MR Zbl

[Imbert 2009] C. Imbert, "Level set approach for fractional mean curvature flows", *Interfaces Free Bound.* **11**:1 (2009), 153–176. MR Zbl

[Maggi 2012] F. Maggi, *Sets of finite perimeter and geometric variational problems: an introduction to geometric measure theory*, Cambridge Studies in Advanced Mathematics **135**, Cambridge University Press, 2012. MR Zbl

[Maz'ya and Shaposhnikova 2002] V. Maz'ya and T. Shaposhnikova, "On the Bourgain, Brezis, and Mironescu theorem concerning limiting embeddings of fractional Sobolev spaces", *J. Funct. Anal.* **195**:2 (2002), 230–238. MR Zbl

[Savin and Valdinoci 2013] O. Savin and E. Valdinoci, "Regularity of nonlocal minimal cones in dimension 2", *Calc. Var. Partial Differential Equations* **48**:1-2 (2013), 33–39. MR Zbl

[Savin and Valdinoci 2014] O. Savin and E. Valdinoci, "Density estimates for a variational model driven by the Gagliardo norm", *J. Math. Pures Appl.* (9) **101**:1 (2014), 1–26. MR Zbl

[Yeh 2006] J. Yeh, *Real analysis: theory of measure and integration*, 2nd ed., World Scientific, Hackensack, NJ, 2006. MR Zbl

Received 11 Dec 2015. Accepted 9 May 2017.

SERENA DIPIERRO: s.dipierro@unimelb.edu.au
School of Mathematics and Statistics, University of Melbourne, 813 Swanston Street, Parkville VIC 3010, Australia

and

Dipartimento di Matematica, Università degli studi di Milano, Via Saldini 50, 20133 Milan, Italy

ARAM KARAKHANYAN: aram.karakhanyan@ed.ac.uk
Maxwell Institute for Mathematical Sciences and School of Mathematics, University of Edinburgh, James Clerk Maxwell Building, Peter Guthrie Tait Road, Edinburgh EH9 3FD, United Kingdom

ENRICO VALDINOCI: enrico@math.utexas.edu
School of Mathematics and Statistics, University of Melbourne, 813 Swanston Street, Parkville VIC 3010, Australia

and

Istituto di Matematica Applicata e Tecnologie Informatiche, Consiglio Nazionale delle Ricerche, Via Ferrata 1, 27100 Pavia, Italy

and

Dipartimento di Matematica, Università degli studi di Milano, Via Saldini 50, 20133 Milan, Italy

ANALYSIS AND PDE
Vol. 10, No. 6, 2017
dx.doi.org/10.2140/apde.2017.10.1361

GLOBAL WELL-POSEDNESS OF THE MHD EQUATIONS IN A HOMOGENEOUS MAGNETIC FIELD

DONGYI WEI AND ZHIFEI ZHANG

We study the MHD equations with small viscosity and resistivity coefficients, which may be different. This is a typical setting in high temperature plasmas. It was proved that the MHD equations are globally well-posed if the initial velocity is close to 0 and the initial magnetic field is close to a homogeneous magnetic field in the weighted Hölder spaces. The main novelty is that the closeness is independent of the dissipation coefficients.

1. Introduction

We consider the incompressible magnetohydrodynamics (MHD) equations in $[0, T) \times \Omega$, with $\Omega \subseteq \mathbb{R}^d$,

$$\begin{cases} \partial_t v - \nu\, \Delta v + v \cdot \nabla v + \nabla p = b \cdot \nabla b, \\ \partial_t b - \mu\, \Delta b + v \cdot \nabla b = b \cdot \nabla v, \\ \operatorname{div} v = \operatorname{div} b = 0, \end{cases} \tag{1-1}$$

where v denotes the velocity field and b denotes the magnetic field, $\nu \geq 0$ is the viscosity coefficient, and $\mu \geq 0$ is the resistivity coefficient. If $\nu = \mu = 0$, (1-1) consists of the so-called ideal MHD equations; if $\nu > 0$ and $b = 0$, (1-1) is reduced to the Navier–Stokes equations. We refer to [Sermange and Temam 1983] for a mathematical introduction to the MHD equations.

It is well known that the 2-dimensional MHD equations with full viscosities (i.e., $\nu > 0$ and $\mu > 0$) have a global smooth solution. In the general case, the question of whether a smooth solution of the MHD equations develops a singularity in finite time is basically open [Sermange and Temam 1983; Cordoba and Fefferman 2001]. Recently, Cao and Wu [2011] studied the global regularity of the 2-dimensional MHD equations with partial dissipation and magnetic diffusion. We refer to [Cao et al. 2013; Chemin et al. 2016; Fefferman et al. 2014; He et al. 2014; Jiu et al. 2015; Lei 2015] for more relevant results.

In this paper, we are concerned with the global well-posedness of the MHD equations in a homogeneous magnetic field B_0. Recently, there have been a lot of works [Abidi and Zhang 2016; Lin et al. 2015; Ren et al. 2014; 2016; Zhang 2014] devoted to the case without resistivity (i.e, $\nu > 0$ and $\mu = 0$). Roughly speaking, it was proved that the MHD equations are globally well-posed and the solution decays in time if the initial velocity field is close to 0 and the initial magnetic field is close to B_0. These results especially justify the numerical observation [Califano and Chiuderi 1999]: the energy of the MHD equations is dissipated at a rate independent of the ohmic resistivity.

MSC2010: 76W05.
Keywords: MHD equations, global well-posedness, Hölder spaces.

In high temperature plasmas, both the viscosity coefficient ν and resistivity coefficient μ are usually very small [Califano and Chiuderi 1999]. Up to now, the heating mechanism of the solar corona is still an unsolved problem in physics [Priest et al. 1998], so it is very interesting to investigate the long-time dynamics of the MHD equations in the case when the dissipation coefficients are very small.

For simplicity, let us first look at the case $\mu = \nu$. Following [Bardos et al. 1988], we rewrite the system (1-1) in terms of the Elsässer variables

$$Z_+ = v + b, \quad Z_- = v - b.$$

Then the ideal MHD equations (1-1) can be written as

$$\begin{cases} \partial_t Z_+ + Z_- \cdot \nabla Z_+ = \nu\, \Delta Z_+ - \nabla p, \\ \partial_t Z_- + Z_+ \cdot \nabla Z_- = \nu\, \Delta Z_- - \nabla p, \\ \operatorname{div} Z_+ = \operatorname{div} Z_- = 0. \end{cases} \tag{1-2}$$

We introduce the fluctuations

$$z_+ = Z_+ - B_0, \quad z_- = Z_- + B_0.$$

Then the system (1-2) can be reformulated as

$$\begin{cases} \partial_t z_+ + Z_- \cdot \nabla z_+ = \nu\, \Delta z_+ - \nabla p, \\ \partial_t z_- + Z_+ \cdot \nabla z_- = \nu\, \Delta z_- - \nabla p, \\ \operatorname{div} z_+ = \operatorname{div} z_- = 0. \end{cases} \tag{1-3}$$

In the case of $\Omega = \mathbb{R}^d$ and $\nu = 0$, Bardos, Sulem and Sulem [Bardos et al. 1988] proved that for large time, the solution $z_\pm$ of (1-3) tends to linear Alfvén waves:

$$\partial_t w_\pm \mp B_0 \cdot \nabla w_\pm = 0.$$

Cai and Lei [2016] and He, Xu and Yu [He et al. 2016] studied the global well-posedness of (1-1) for any $\nu \geq 0$ and $\Omega = \mathbb{R}^3$. The result in [Cai and Lei 2016] also includes the case $\Omega = \mathbb{R}^2$. These works are based on an important observation: the nonlinear terms $z_- \cdot \nabla z_+$ and $z_+ \cdot \nabla z_-$ can be essentially neglected after a long time since $z_\pm$ are transported along the opposite direction. To justify this observation, the key point is to make weighted estimates for the fluctuations $z_\pm$. Due to the nonlocal pressure, the choice of weight function is very delicate. On the other hand, the viscosity gives rise to more technical troubles compared with the ideal case.

From the physical point of view, it is more natural to consider the MHD equations in a domain with boundary. One frequently used domain in physics is a slab bounded by two hyperplanes, i.e., $\Omega = \mathbb{R}^{d-1} \times [0, 1]$. More importantly, although both ν and μ are very small, they should be different in the real case. However, the proof in [Cai and Lei 2016; He et al. 2016] strongly relies on the facts that Ω is a whole space and $\nu = \mu$. In particular, the formulation (1-3) plays a crucial role.

The main goal of this paper is to prove the global well-posedness of (1-1) in the physical case when Ω is a slab and $\nu \neq \mu$. In this case, we need to impose suitable boundary conditions on $z_\pm$. Let $z_\pm$ be a function of (t, x, y), $(x, y) \in \Omega$. In the case when $\nu = \mu = 0$, we impose the nonpenetrating boundary condition

$$z_{\pm}^{d} = 0 \quad \text{on } y = 0, 1. \tag{1-4}$$

In the case when $\nu > 0$ and $\mu > 0$, we impose the Navier-slip boundary condition

$$z_{\pm}^{d} = 0, \qquad \partial_d z_{\pm}^{i} = 0, \quad i = 1, \dots, d-1, \qquad \text{on } y = 0, 1. \tag{1-5}$$

To deal with the boundary case, our idea is to use the symmetric extension and solve the MHD equations in the framework of Hölder spaces $C^{1,\alpha}$ for $0 < \alpha < 1$. In the ideal case, we give a representation formula of the pressure by using the symmetric extension. Although the extended solution does not have the same regularity as the original one under the nonpenetrating boundary condition, we have the important observation that ∇p still lies in $C^{1,\alpha}$ based on the representation formula. In the viscous case, we can reduce the slab domain to $\Omega = \mathbb{R}^{d-1} \times \mathbb{T}$ by using the symmetric extension, because the extended solution still keeps the $C^{1,\alpha}$ regularity under the Navier-slip boundary condition.

The most challenging task comes from the case $\nu \neq \mu$. To handle this case, we need to introduce some new ideas. First of all, we introduce a key decomposition: let $\mu_1 = \frac{1}{2}(\nu + \mu)$, $\mu_2 = \frac{1}{2}(\nu - \mu)$, and we have the decompositions $z_+ = z_+^{(1)} + z_+^{(2)}$ and $z_- = z_-^{(1)} + z_-^{(2)}$ such that

$$\begin{cases} \partial_t z_+^{(1)} + Z_- \cdot \nabla z_+^{(1)} = \mu_1 \, \Delta z_+^{(1)} - \nabla p_+^{(1)}, \\ \partial_t z_-^{(1)} + Z_+ \cdot \nabla z_-^{(1)} = \mu_1 \, \Delta z_-^{(1)} - \nabla p_-^{(1)}, \\ \partial_t z_+^{(2)} + Z_- \cdot \nabla z_+^{(2)} = \mu_1 \, \Delta z_+^{(2)} + \mu_2 \, \Delta z_- - \nabla p_+^{(2)}, \\ \partial_t z_-^{(2)} + Z_+ \cdot \nabla z_-^{(2)} = \mu_1 \, \Delta z_-^{(2)} + \mu_2 \, \Delta z_+ - \nabla p_-^{(2)}. \end{cases}$$

The next task is to establish a closed uniform estimate for the fluctuations $z_{\pm}^{(1)}$ and $z_{\pm}^{(2)}$ with respect to μ_1 and t. For this, we need the following key ingredients:

- The construction of the weighted Hölder spaces for the solution. Due to the appearance of the extra problematic terms $\Delta z_{\pm}$, we have to work in spaces with different regularity and weight for the solution $z_{\pm}^{(1)}$, $z_{\pm}^{(2)}$. Such inconsistencies give rise to the essential difficulties. In particular, the choice of the weight is very delicate. In [Bardos et al. 1988; Cai and Lei 2016; He et al. 2016], the weight has decay in all directions. For the slap domain, the weight is only allowed to decay in partial directions. Again, the weight has to be compatible with the estimate of the nonlocal pressure.
- Uniform estimates of the transport equation in the weighted Hölder spaces, which are very crucial to control the growth of the Lagrangian map.
- Uniform estimates for the parabolic equation with variable coefficients in the suitable weighted Hölder spaces. This is the most important step.
- Boundedness of the Riesz transform and its commutator in the weighted Hölder spaces, which is essentially used to handle the nonlocal pressure. To our knowledge, these results are new and may be independent of interest. The proof is highly nontrivial.

In this work, we require that μ_2/μ_1 is small. However, this cannot be handled as a perturbation of the case $\mu_2 = 0$ except when $|\mu_2| \leq \mu_1^{\alpha}$ for some $\alpha > 1$. In this case, the smallness of $z_{\pm}^{(2)}$ is not easily observed. If we directly use the energy method, we can only prove that $\|z_\pm^{(2)}(t)\|_{L^2} = O(|\mu_2|/\mu_1)$ for fixed $z_\pm(0)$. However, we can show that $|z_\pm^{(2)}(t)|_{0,\alpha} = O(\mu_1)$ for $t \sim 1/\mu_1$ and fixed $z_\pm(0)$.

In this paper, we consider the MHD equations in a homogeneous magnetic field. In the real case (for example, solar corona), it is more natural to consider the MHD equations in an inhomogeneous magnetic field. An important question is to consider the decay of Alfvén waves in an inhomogeneous magnetic field $B_0(y) = (b_1(y), b_2(y), 0)$. This is similar to the situation of Landau damping.

2. The weighted Hölder spaces and symmetric extension

Weighted Hölder spaces. Let $\Omega \subseteq \mathbb{R}^d$ be a domain and $\alpha \in (0, 1]$. We denote by $C^{k,\alpha}(\Omega)$, $(k = 0, 1)$ the Hölder space equipped with the norm

$$|u|_{0,\alpha;\Omega} := |u|_{0;\Omega} + [u]_{\alpha;\Omega}, \quad |u|_{1,\alpha;\Omega} := |u|_{0;\Omega} + |\nabla u|_{0,\alpha;\Omega},$$

where

$$|u|_{0;\Omega} = \sup_{X\in\Omega} |u(X)|, \quad [u]_{\alpha;\Omega} = \sup_{X,Y\in\Omega} \frac{|u(X) - u(Y)|}{|X - Y|^\alpha}.$$

Let $h(X) \in C(\mathbb{R}^d)$ be a positive bounded function. We introduce the weighted $C^{k,\alpha}$ norms

$$|u|_{0,\alpha;h,\Omega} := |u|_{0;h,\Omega} + [u]_{\alpha;h,\Omega}, \quad |u|_{1,\alpha;h,\Omega} := |u|_{0;h,\Omega} + |\nabla u|_{0,\alpha;h,\Omega},$$

where

$$|u|_{0;h,\Omega} = \left|\frac{u}{h}\right|_{0;\Omega}, \quad [u]_{\alpha;h,\Omega} = \sup_{X,Y\in\Omega} \frac{|u(X) - u(Y)|}{(h(X) + h(Y))|X - Y|^\alpha}.$$

We say that $u \in C_h^{k,\alpha}(\Omega)$ if $|u|_{k,\alpha;h,\Omega} < +\infty$. We also introduce

$$|u|_{k,\alpha;h,\Omega,T} := \sup_{0\le t\le T} |u(t)|_{k,\alpha;h(t),\Omega}.$$

When $\Omega = \mathbb{R}^d$, we will omit the subscript Ω in the norm of Hölder spaces.

The following two lemmas can be proved by using the definition of Hölder norm.

Lemma 2.1. *Let h, h_1, h_2 be the weight functions such that there exists a constant c_0 such that*

$$0 < c_0 h(X) \le h(Y) \quad \textit{for any } X, Y \in \mathbb{R}^d, \ |X - Y| \le 2. \tag{2-1}$$

Then there exists a constant C depending only on c_0 such that, for $k = 0, 1$,

$$|u|_{0,\alpha;h,\Omega} \le C\big(|u|_{0;h,\Omega} + |\nabla u|_{0;h,\Omega}\big),$$

$$|uw|_{k,\alpha;h_1h_2,\Omega} \le C|u|_{k,\alpha;h_1,\Omega}|w|_{k,\alpha;h_2,\Omega},$$

$$\left|\int_t^s u(r)\,dr\right|_{k,\alpha;\int_t^s h(r)\,dr,\Omega} \le \sup_{t\le r\le s} |u(r)|_{k,\alpha;h(r),\Omega}.$$

Lemma 2.2. *Let Φ be a map from Ω to Ω with $\nabla\Phi \in C^{0,\alpha}(\Omega)$. It holds that*

$$|u \circ \Phi|_{0,\alpha;h\circ\Phi,\Omega} \le |u|_{0,\alpha;h,\Omega} \max(|\nabla\Phi|^{\alpha}_{0;\Omega}, 1),$$
$$|u \circ \Phi|_{1,\alpha;h\circ\Phi,\Omega} \le |u|_{1,\alpha;h,\Omega} \max(|\nabla\Phi|^{\alpha}_{0;\Omega}, 1) \max(|\nabla\Phi|_{0,\alpha;\Omega}, 1).$$

Here and in what follows, $|\nabla\Phi|$ *denotes the matrix norm defined by*

$$|A| := \sup_{|X|=1} |AX|. \tag{2-2}$$

To deal with the viscous case, we introduce the following scaled weighted Hölder space. Let $\alpha \in (0, 1)$, $R \ge 0$ and define

$$|u|_{0,\alpha;h,R} := |u|_{0;h} + R^{\alpha}[u]_{\alpha;h},$$
$$|u|_{1,\alpha;h,R} := |u|_{0,\alpha;h} + \max(R, R^{1-\alpha})|\nabla u|_{0,\alpha;h,R}.$$

For these kinds of weighted spaces, we have analogues of Lemmas 2.1 and 2.2. For example, if $h(X)$ satisfies

$$0 < c_0 h(X) \le h(Y) \quad \text{for any } X, Y \in \mathbb{R}^d,\ |X - Y| \le 2R, \tag{2-3}$$

then for $R \ge 1$, we have

$$|u|_{0;h} + R|\nabla u|_{0,\alpha;h,R} \le |u|_{1,\alpha;h,R} \le |u|_{0,\alpha;h,R} + R|\nabla u|_{0,\alpha;h,R} \le C\big(|u|_{0;h} + R|\nabla u|_{0,\alpha;h,R}\big).$$

Here C is a constant depending only on c_0. In the following, we will fix $\alpha \in (0, 1)$.

Lemma 2.3. *Let $\gamma > 0$ and $h(X) > 0$. Then there exists a constant C independent of h, γ, t such that*

$$\begin{aligned}\left|\int_0^t u(s)\,ds\right|_{1,\alpha;h,\sqrt{k+\gamma t}} &\le C\gamma^{-1} \sup_{0<s<t}\Big((\gamma s)^{\frac12}(\gamma(t-s))^{\frac12}|u(s)|_{0,\alpha;h} + \varphi_\alpha(\sqrt{k+\gamma s})(\gamma(t-s))^{1-\frac{\alpha}{2}}|\nabla u(s)|_{0;h} \\ &\qquad + \varphi_\alpha(\sqrt{k+\gamma s})(\gamma(t-s))^{\frac{3-\alpha}{2}}[\nabla u(s)]_{1;h}\Big),\end{aligned}$$

where $\varphi_\alpha(R) = \max(R, R^{1+\alpha})$.

Proof. We denote by $C\gamma^{-1}A$ the right-hand side of the inequality. Then we have

$$\left|\int_0^t u(s)\,ds\right|_{0,\alpha;h} \le \int_0^t |u(s)|_{0,\alpha;h}\,ds \le \int_0^t (\gamma s)^{-\frac12}(\gamma(t-s))^{-\frac12}\,ds A \le C\gamma^{-1}A,$$
$$\begin{aligned}\left|\nabla\int_0^t u(s)\,ds\right|_{0;h} \le \int_0^t |\nabla u(s)|_{0;h}\,ds &\le \int_0^t \varphi_\alpha(\sqrt{k+\gamma s})^{-1}(\gamma(t-s))^{-1+\frac{\alpha}{2}}\,ds A \\ &\le C\gamma^{-1}\min\big((k+\gamma t)^{-\frac12}, (k+\gamma t)^{-\frac{1-\alpha}{2}}\big)A.\end{aligned}$$

For any $X, Y \in \mathbb{R}^d$, we have

$$|\nabla u(s,X) - \nabla u(s,Y)| \le |X-Y|(h(X)+h(Y))[\nabla u(s)]_{1;h},$$
$$|\nabla u(s,X) - \nabla u(s,Y)| \le |\nabla u(s,X)| + |\nabla u(s,Y)| \le (h(X)+h(Y))|\nabla u(s)|_{0;h}.$$

This gives

$$\begin{aligned}|\nabla u(s,X)-\nabla u(s,Y)| &\le \min\big((\gamma(t-s))^{\frac12},|X-Y|\big)(h(X)+h(Y))\big([\nabla u(s)]_{1;h}+(\gamma(t-s))^{-\frac12}|\nabla u(s)|_{0;h}\big)\\ &\le \min\big((\gamma(t-s))^{\frac12},|X-Y|\big)(h(X)+h(Y))\varphi_\alpha(\sqrt{k+\gamma s})^{-1}(\gamma(t-s))^{-\frac{3-\alpha}{2}}A.\end{aligned}$$

Therefore,

$$\begin{aligned}&\left|\nabla\int_0^t u(s)\,ds(X)-\nabla\int_0^t u(s)\,ds(Y)\right|\\ &\quad\le\int_0^t|\nabla u(s,X)-\nabla u(s,Y)|\,ds\\ &\quad\le\int_0^t \min\big((\gamma(t-s))^{\frac12},|X-Y|\big)(h(X)+h(Y))\varphi_\alpha(\sqrt{k+\gamma s})^{-1}(\gamma(t-s))^{-\frac{3-\alpha}{2}}A\,ds\\ &\quad\le C(h(X)+h(Y))A\bigg(\min\big((\gamma t)^{\frac12},|X-Y|\big)\int_0^{\frac t2}\varphi_\alpha(\sqrt{k+\gamma s})^{-1}\,ds(\gamma t)^{-\frac{3-\alpha}{2}}\\ &\qquad\qquad+\int_{\frac t2}^t\min\big((\gamma(t-s))^{\frac12},|X-Y|\big)(\gamma(t-s))^{-\frac{3-\alpha}{2}}\,ds\varphi_\alpha(\sqrt{k+\gamma t})^{-1}\bigg)\\ &\quad\le C(h(X)+h(Y))A\big((\gamma t)^{\frac{1-\alpha}{2}}|X-Y|^\alpha t\varphi_\alpha(\sqrt{k+\gamma t})^{-1}(\gamma t)^{-\frac{3-\alpha}{2}}+\gamma^{-1}|X-Y|^\alpha\varphi_\alpha(\sqrt{k+\gamma t})^{-1}\big)\\ &\quad\le C\gamma^{-1}(h(X)+h(Y))A|X-Y|^\alpha\varphi_\alpha(\sqrt{k+\gamma t})^{-1}.\end{aligned}$$

Hence, we deduce our result. □

Lemma 2.4. *Let Φ be a map from $\mathbb{R}^d$ to $\mathbb{R}^d$ with $\nabla\Phi\in C^{0,\alpha}(\mathbb{R}^d)$. It holds that*

$$\begin{aligned}|u\circ\Phi|_{0,\alpha;h\circ\Phi,R}&\le|u|_{0,\alpha;h,R}\max(|\nabla\Phi|_0^\alpha,1),\\ |u\circ\Phi|_{1,\alpha;h\circ\Phi,R}&\le|u|_{1,\alpha;h,R}\max(|\nabla\Phi|_0^\alpha,1)\max(|\nabla\Phi|_{0,\alpha;1,R},1).\end{aligned}$$

Symmetric extension. Let $\Omega=\mathbb{R}^{d-1}\times[0,1]$ be a strip and $X=(x,y),\ x\in\mathbb{R}^{d-1},\ y\in[0,1]$ be a point in Ω.

Let T_e be an even extension from $C(\Omega)$ to $C(\mathbb{R}^d)$ defined by

$$T_ef(x,2n+y)=T_ef(x,2n-y)=f(x,y)$$

for $x\in\mathbb{R}^{d-1},\ y\in[0,1]$, $n\in\mathbb{Z}$. Let T_o be an odd extension from $C_0(\Omega)=\{u\in C(\Omega):u=0\text{ on }\partial\Omega\}$ to $C(\mathbb{R}^d)$ defined by

$$T_of(x,2n-y)=-f(x,y),\quad T_of(x,2n+y)=f(x,y)$$

for $x\in\mathbb{R}^{d-1},\ y\in[0,1],\ n\in\mathbb{Z}$.

Lemma 2.5. *It holds that*

$$|T_ef|_{0,\alpha}=|f|_{0,\alpha,\Omega},\quad |f|_{0,\alpha;\Omega}\le|T_of|_{0,\alpha}\le2|f|_{0,\alpha;\Omega}.$$

The same result holds for the weighted Hölder norm $|\cdot|_{0,\alpha;h}$ if the weight function $h(X)$ depends only on x.

Proof. First of all, it is obvious that

$$|f|_{0,\alpha;\Omega} \le |T_e f|_{0,\alpha}, \quad |f|_{0,\alpha;\Omega} \le |T_o f|_{0,\alpha},$$

and the same is true for the weighted Hölder norm $|\cdot|_{0,\alpha;h}$. We define

$$\begin{aligned} \rho_0(y) &= \inf_{n\in\mathbb{Z}} |y-2n| \in [0,1] \quad && \text{for } y \in \mathbb{R}, \\ \rho(X) &= (x, \rho_0(y)) \in \Omega && \text{for } X = (x,y) \in \mathbb{R}^d, \end{aligned}$$

and let

$$\Omega_+ = \bigcup_{n\in\mathbb{Z}} \mathbb{R}^{d-1} \times [2n, 2n+1], \quad \Omega_- = \bigcup_{n\in\mathbb{Z}} \mathbb{R}^{d-1} \times [2n-1, 2n].$$

Then it is easy to see that

$$\begin{gathered} T_e f = f \circ \rho, \\ T_o f = f \circ \rho \quad \text{in } \Omega_+, \quad T_o f = -f \circ \rho \quad \text{in } \Omega_-, \\ |\rho_0(y) - \rho_0(y')| \le |y-y'|, \quad |\rho(X) - \rho(Y)| \le |X-Y|, \end{gathered}$$

from which, it follows that

$$\begin{gathered} |T_e f|_{0,\alpha} \le |f|_{0,\alpha;\Omega}, \quad |T_e f|_{0,\alpha;h} \le |f|_{0,\alpha;h,\Omega}, \\ |T_o f|_0 \le |f|_{0;\Omega}, \qquad |T_o f|_{0;h} \le |f|_{0;h,\Omega}. \end{gathered}$$

Given $X = (x,y),\ Y = (x', y') \in \mathbb{R}^d$ with $y \le y'$, if $X, Y \in \Omega_+$ or $X, Y \in \Omega_-$, then

$$\begin{aligned} |T_o f(X) - T_o f(Y)| &= |f \circ \rho(X) - f \circ \rho(Y)| \\ &\le |f|_{0,\alpha;h,\Omega} \big(h \circ \rho(X) + h \circ \rho(Y)\big) |\rho(X) - \rho(Y)|^\alpha \\ &\le |f|_{0,\alpha;h,\Omega} (h(X) + h(Y)) |X-Y|^\alpha. \end{aligned}$$

Here we used $h \circ \rho(X) = h(X)$. Otherwise, there exists $y_1, y_2 \in \mathbb{Z}$ so that $y_1 - 1 \le y \le y_1 \le y_2 \le y' \le y_2 + 1$. Let $X' = (x, y_1),\ Y' = (x', y_2)$. Then for $f \in C_0(\Omega)$, we have

$$\begin{aligned} |T_o f(X)| = |f \circ \rho(X)| &= |f \circ \rho(X) - f \circ \rho(X')| \\ &\le |f|_{0,\alpha;h,\Omega} \big(h \circ \rho(X) + h \circ \rho(X')\big) |\rho(X) - \rho(X')|^\alpha \\ &\le 2|f|_{0,\alpha;h,\Omega} h(X) |X - X'|^\alpha. \end{aligned}$$

Similarly, we have

$$|T_o f(Y)| \le 2|f|_{0,\alpha;h,\Omega} h(Y) |Y - Y'|^\alpha.$$

Then, using $|X - X'| + |Y - Y'| \le |X - Y|$, we get

$$|T_o f(X) - T_o f(Y)| \le 2|f|_{0,\alpha;h,\Omega} (h(X) + h(Y)) |X-Y|^\alpha.$$

This shows $[T_o f]_{\alpha;h} \le 2[f]_{\alpha;h,\Omega}$. Similarly, $[T_o f]_\alpha \le 2[f]_{\alpha;\Omega}$. □

3. Global well-posedness for the ideal MHD equations

This section is devoted to the proof of the global well-posedness of the ideal MHD equations in $\mathbb{R}^{d-1}\times[0,1]$ with the boundary condition (1-4). Recall that in terms of the Elsässer variables $z_\pm = Z_\pm \pm B_0$, the ideal MHD equations take

$$\begin{cases} \partial_t z_+ + Z_-\cdot\nabla z_+ = -\nabla p, \\ \partial_t z_- + Z_+\cdot\nabla z_- = -\nabla p, \\ \operatorname{div} z_+ = \operatorname{div} z_- = 0, \\ z_\pm^d(t,x,y) = 0 \quad \text{on } y = 0, 1. \end{cases} \tag{3-1}$$

Without loss of generality, we take the background magnetic field $B_0 = (1, 0, \dots, 0)$.

Main result. Let $f(x,y) = f_0(x_1)$, where $f_0 \in C^1(\mathbb{R})$ is chosen so that $|f_0'| < f_0 < 1$ and for some $C_1^* > 0$,

$$\begin{aligned} \delta(T) \triangleq \sup_{Y\in\mathbb{R}^d} \int_{-T}^{T} f(Y + 2B_0 t)\,dt \le C_1^* \quad & \text{for any } T > 0, \\ \int_{\mathbb{R}^d} \frac{f(Y)}{1+|X-Y|^{d+1}}\,dY \le C_1^* f(X) \quad & \text{for any } X \in \mathbb{R}^d, \\ f(X) \le 2f(Y) \quad & \text{for any } |X-Y| \le 2. \end{aligned} \tag{3-2}$$

In fact, $f_0(r) = (C_0 + r^2)^{-\frac{\delta+1}{2}}$ satisfies the above conditions for some $C_0 > 1$ and $0 < \delta < 1$.

Now we introduce the weight function $f_\pm(t,X)$ given by

$$f_\pm(t,X) \triangleq f(X \pm B_0 t),$$

which satisfies (2-1) with a uniform constant c_0 independent of t. Let

$$M_\pm(t) \triangleq \sup_{|s|\le t} |z_\pm(s)|_{1,\alpha;f_\pm(s),\Omega}.$$

The main result of this section is stated as follows.

Theorem 3.1. *Let $\alpha \in (0,1)$. There exists $\varepsilon > 0$ such that if $M_\pm(0) \le \varepsilon$, then there exists a global in time unique solution $(z_+, z_-) \in L^\infty\big(0,+\infty; C^{1,\alpha}(\Omega)\big)$, with the pressure p determined by* (3-10), *to the ideal MHD equations* (3-1), *which satisfies*

$$M_\pm(t) \le C\varepsilon \quad \textit{for any } t \in [0,+\infty).$$

Remark 3.2. Since $M_\pm(0) \sim |z_\pm(0)\langle x_1\rangle^{1+\delta}|_{1,\alpha;\Omega}$ if $f_0(r) = (C_0 + r^2)^{-\frac{\delta+1}{2}}$, the initial data decays at infinity only in one direction. This is very crucial for the global well-posedness in the slap domain, especially in $\mathbb{R}\times[0,1]$.

We conclude this subsection by introducing some properties of weighted functions. Let

$$g(t,X) \triangleq \int_{\mathbb{R}^d} \frac{f(Y+B_0 t) f(Y - B_0 t)}{1+|X-Y|^{d+1}}\,dY.$$

We have the following important facts.

Lemma 3.3. *There exists a constant $C > 0$ such that for any $X \in \mathbb{R}^d$, $t \in \mathbb{R}$,*

$$\begin{aligned} &f(X+B_0t)f(X-B_0t) \le Cg(t,X), \\ &g(t,X) \le C\bigl(1+|X-Y|\bigr)^{d+1}g(t,Y), \\ &\int_{-T}^{T} g(t, X \pm B_0t)\,dt \le C\delta(T)f(X). \end{aligned}$$

Proof. Thanks to $f(Y) \ge f(X)/2$ for $|X-Y| < 2$, we get

$$\begin{aligned} g(t,X) &\ge \int_{B(X,2)} \frac{f(Y+B_0t)f(Y-B_0t)}{1+|X-Y|^{d+1}}\,dY \ge \frac{1}{4}\int_{B(X,2)} \frac{f(X+B_0t)f(X-B_0t)}{1+|X-Y|^{d+1}}\,dY \\ &\ge C^{-1}f(X+B_0t)f(X-B_0t), \end{aligned}$$

which gives the first inequality.

Using the inequality

$$\frac{1}{1+|X-Z|^{d+1}} \le C\frac{1+|X-Y|^{d+1}}{1+|Y-Z|^{d+1}},$$

we infer

$$\begin{aligned} g(t,X) &= \int_{\mathbb{R}^d} \frac{f(Z+B_0t)f(Z-B_0t)}{1+|X-Z|^{d+1}}\,dZ \le C\int_{\mathbb{R}^d} \frac{f(Z+B_0t)f(Z-B_0t)}{1+|Y-Z|^{d+1}}(1+|X-Y|^{d+1})\,dY \\ &= C\bigl(1+|X-Y|^{d+1}\bigr)g(t,Y), \end{aligned}$$

which gives the second inequality.

Make a change of variable

$$g(t,X+B_0t) = \int_{\mathbb{R}^d} \frac{f(Y+B_0t)f(Y-B_0t)}{1+|X+B_0t-Y|^{d+1}}\,dY = \int_{\mathbb{R}^d} \frac{f(Y+2B_0t)f(Y)}{1+|X-Y|^{d+1}}\,dY,$$

which along with (3-2) gives

$$\int_{-T}^{T} g(t,X+B_0t) = \int_{\mathbb{R}^d} \frac{\int_{-T}^{T} f(Y+2B_0t)f(Y)\,dt}{1+|X-Y|^{d+1}}\,dY \le C\int_{\mathbb{R}^d} \frac{\delta(T)f(Y)}{1+|X-Y|^{d+1}}\,dY \le C\delta(T)f(X).$$

Similarly, we have

$$\int_{-T}^{T} g(t,X-B_0t) \le C\delta(T)f(X). \qquad \square$$

Weighted $C^{1,\alpha}$ estimate for the transport equation. Let $Z \in C^1\bigl([0,T]\times\Omega\bigr)$ be a vector field with $Z^d = 0$ on $\partial\Omega$. We introduce the characteristic associated with Z:

$$\frac{d}{dt}\Phi(s,t,X) = Z(t,\Phi(s,t,X)), \quad \Phi(s,s,X) = X. \tag{3-3}$$

Then $\Phi(s,t,X) \in C^1\bigl([0,T]\times[0,T]\times\Omega\bigr)$ is a diffeomorphism from Ω to Ω and $\partial\Omega$ to $\partial\Omega$ having the property

$$\Phi(r,t)\circ\Phi(s,r) = \Phi(s,t), \quad \Phi(s,s) = \mathrm{Id}.$$

Lemma 3.4. *If $Z(t,X)$ satisfies the extra condition*

$$|\nabla Z|_{0,\alpha;h,\Omega,T}\int_{t_0}^{T} h(t,\Phi(T,t,X))\,dt \le A_0 \quad \textit{for any } X\in\Omega, \tag{3-4}$$

then it holds that for $0\le t_0\le t\le s<T$,

$$\begin{aligned}|\nabla\Phi(s,t)-\mathrm{Id}|_{0;\Omega} &\le e^{A_0}-1,\\ |\nabla\Phi(s,t)|_{0;\Omega} &\le e^{A_0},\\ [\nabla\Phi(s,t)]_{\alpha;\Omega} &\le 2A_0e^{(2+\alpha)A_0}.\end{aligned}$$

Proof. Thanks to the definition of $\Phi(s,t)$, we have

$$\begin{gathered}\partial_t\nabla\Phi(s,t)=\nabla\Phi(s,t)\big((\nabla Z(t))\circ\Phi(s,t)\big),\\ \Phi(s,s)=\mathrm{Id},\quad \nabla\Phi(s,s)=\mathrm{Id},\\ |\nabla\Phi(s,t)|\le|\nabla\Phi(s,t)-\mathrm{Id}|+1.\end{gathered}$$

Here $|\nabla\Phi(s,t)|$ is the matrix norm defined by (2-2). Therefore,

$$\begin{aligned}\big|\nabla\Phi(s,t)-\mathrm{Id}\big| &\le \int_t^s |\partial_r\nabla\Phi(s,r)|\,dr\\ &\le \int_t^s |\nabla\Phi(s,r)|\big|(\nabla Z(r))\circ\Phi(s,r)\big|\,dr\\ &\le \int_t^s \big|(\nabla Z(r))\circ\Phi(s,r)\big|\,dr+\int_t^s\big|\nabla\Phi(s,r)-\mathrm{Id}\big|\big|(\nabla Z(r))\circ\Phi(s,r)\big|\,dr,\end{aligned}$$

which implies

$$\big|\nabla\Phi(s,t)-\mathrm{Id}\big|\le\exp\bigg(\int_t^s\big|(\nabla Z(r))\circ\Phi(s,r)\big|\,dr\bigg)-1.$$

Thanks to

$$\big|(\nabla Z(r))\circ\Phi(s,r)\big|\le|\nabla Z|_{0,\alpha;h,\Omega,T}\,h(r)\circ\Phi(s,r),$$

we get by (3-4) that

$$\begin{aligned}\int_t^s\big|(\nabla Z(r))\circ\Phi(s,r)(X)\big|\,dr &\le |\nabla Z|_{0,\alpha;h,\Omega,T}\int_t^s h(r)\circ\Phi(s,r)(X)\,dr\\ &= |\nabla Z|_{0,\alpha;h,\Omega,T}\int_t^s h\big(r,\Phi(T,r,\Phi(s,T)(X))\big)\,dr\le A_0.\end{aligned}$$

Thus, we conclude that

$$\begin{gathered}|\nabla\Phi(s,t)-\mathrm{Id}|_{0;\Omega}\le e^{A_0}-1,\\ |\nabla\Phi(s,t)|_{0;\Omega}\le e^{A_0},\\ |\Phi(s,t,X)-\Phi(s,t,Y)|\le|\nabla\Phi(s,t)|_{0;\Omega}|X-Y|\le e^{A_0}|X-Y|.\end{gathered}$$

Notice that

$$\begin{aligned}\big|\nabla\Phi(s,t,X)-\nabla\Phi(s,t,Y)\big| \le \int_t^s \big|\nabla\Phi(s,r,X)-\nabla\Phi(s,r,Y)\big|\big|(\nabla Z(r))\circ\Phi(s,r,X)\big|\,dr \\ + \int_t^s \big|\nabla\Phi(s,r,Y)\big|\big|(\nabla Z(r))\circ\Phi(s,r,X)-(\nabla Z(r))\circ\Phi(s,r,Y)\big|\,dr.\end{aligned}$$

From this and Gronwall's inequality, we infer

$$\begin{aligned}&\big|\nabla\Phi(s,t,X)-\nabla\Phi(s,t,Y)\big| \\ &\le \int_t^s \big|\nabla\Phi(s,r,Y)\big|\big|(\nabla Z(r))\circ\Phi(s,r,X)-(\nabla Z(r))\circ\Phi(s,r,Y)\big|\,dr \exp\Big(\int_t^s \big|(\nabla Z(r))\circ\Phi(s,r,X)\big|\,dr\Big) \\ &\le \int_t^s \big|\nabla\Phi(s,r,Y)\big||\nabla Z|_{0,\alpha;h,\Omega,T}\big(h(r,\Phi(s,r,X))+h(r,\Phi(s,r,Y))\big)\big|\Phi(s,r,X)-\Phi(s,r,Y)\big|^\alpha\,dr\,e^{A_0} \\ &\le \int_t^s e^{A_0}|\nabla Z|_{0,\alpha;h,\Omega,T}\big(h(r,\Phi(s,r,X))+h(r,\Phi(s,r,Y))\big)e^{\alpha A_0}|X-Y|^\alpha\,dr\,e^{A_0} \\ &= e^{(2+\alpha)A_0}|X-Y|^\alpha|\nabla Z|_{0,\alpha;h,\Omega,T}\int_t^s \big(h(r,\Phi(s,r,X))+h(r,\Phi(s,r,Y))\big)\,dr \\ &\le 2A_0e^{(2+\alpha)A_0}|X-Y|^\alpha,\end{aligned}$$

which shows the last inequality of the lemma. □

Next we consider the transport equation

$$\partial_t u + Z\cdot\nabla u = F, \quad u(0,X)=u_0(X). \tag{3-5}$$

Using the characteristic, the solution $u(t,X)$ is given by

$$u(t,X)=u_0(\Phi(t,0,X))+\int_0^t F(s,\Phi(t,s,X))\,ds. \tag{3-6}$$

Lemma 3.5. *If Z satisfies* (3-4), *then we have*

$$\begin{aligned}|u(t)|_{0,\alpha;\Omega} \le e^{\alpha A_0}\Big(|u_0|_{0,\alpha;\Omega}+\int_0^t |F(s)|_{0,\alpha;\Omega}\,ds\Big), \\ |\mathrm{div}\ u(t)|_{0;\Omega} \le |\mathrm{div}\ u_0|_{0;\Omega}+\int_0^t \big|\big(\mathrm{tr}(\nabla Z\,\nabla u)-\mathrm{div}\,F\big)(s)\big|_{0;\Omega}\,ds.\end{aligned}$$

Proof. Using (3-6) and Lemmas 2.2 and 3.4, we get

$$\begin{aligned}|u(t)|_{0,\alpha;\Omega} &\le |u_0\circ\Phi(t,0)|_{0,\alpha;\Omega}+\int_0^t |F(s)\circ\Phi(t,s)|_{0,\alpha;\Omega}\,ds \\ &\le |u_0|_{0,\alpha;\Omega}\max\big(|\nabla\Phi(t,0)|^\alpha_{0;\Omega},1\big)+\int_0^t |F(s)|_{0,\alpha;\Omega}\max\big(|\nabla\Phi(t,s)|^\alpha_{0;\Omega},1\big)\,ds \\ &\le e^{\alpha A_0}\Big(|u_0|_{0,\alpha;\Omega}+\int_0^t |F(s)|_{0,\alpha;\Omega}\,ds\Big).\end{aligned}$$

Taking the divergence of (3-5), we obtain

$$\partial_t \operatorname{div} u + Z\cdot\nabla \operatorname{div} u + \operatorname{tr}(\nabla Z\,\nabla u) = \operatorname{div} F,\quad u(0,X)=u_0(X).$$

So, we have

$$\operatorname{div} u(t) = \operatorname{div} u_0\circ\Phi(t,0) + \int_0^t \big(\operatorname{div} F - \operatorname{tr}(\nabla Z\,\nabla u)\big)(s)\circ\Phi(t,s)\,ds,$$

and then the second inequality follows easily. □

Proposition 3.6. *If* $|Z+B_0|_{1,\alpha;f_-,\Omega,T}\,\delta(T)<1$, *then we have*

$$|u|_{1,\alpha;f_+,\Omega,T} \le C\big(|u_0|_{1,\alpha;f,\Omega} + \delta(T)|F|_{1,\alpha;g,\Omega,T}\big).$$

If $|Z-B_0|_{1,\alpha;f_+,\Omega,T}\,\delta(T)<1$, *then we have*

$$|u|_{1,\alpha;f_-,\Omega,T} \le C\big(|u_0|_{1,\alpha;f,\Omega} + \delta(T)|F|_{1,\alpha;g,\Omega,T}\big).$$

Here C is a constant independent of T.

Proof. We only prove the first inequality; the proof of the second one is similar. Let us claim

$$\big|\Phi(s,t,X)+B_0(t-s)-X\big|<2\quad\text{for } 0\le t\le s\le T. \tag{3-7}$$

Otherwise, there exists $t\in[0,s]$ such that $\big|\Phi(s,t,X)+B_0(t-s)-X\big|=2$ and $\big|\Phi(s,r,X)+B_0(r-s)-X\big|\le 2$ for $r\in[t,s]$. Thus,

$$\begin{aligned}\big|\Phi(s,t,X)+B_0(t-s)-X\big| &\le \int_t^s |\partial_r\Phi(s,r,X)+B_0|\,dr\\ &= \int_t^s |Z(r,\Phi(s,r,X))+B_0|\,dr\\ &\le \int_t^s |Z+B_0|_{1,\alpha;f_-,\Omega,T}\,f_-(r,\Phi(s,r,X))\,dr\\ &= |Z+B_0|_{1,\alpha;f_-,\Omega,T}\int_t^s f(\Phi(s,r,X)-B_0r)\,dr,\end{aligned}$$

while, by (3-2),

$$\int_t^s f(\Phi(s,r,X)-B_0r)\,dr \le 2\int_t^s f(X-B_0(r-s)-B_0r)\,dr\le 2\delta(T).$$

This shows

$$\big|\Phi(s,t,X)+B_0(t-s)-X\big|\le 2|Z+B_0|_{1,\alpha;f_-,\Omega,T}\,\delta(T)<2,$$

which is a contradiction; hence (3-7) is true.

Now we verify (3-4) for $h=f_-$ and $A_0=2$. Indeed, by (3-2) and (3-7),

$$\int_0^T f_-(t,\Phi(T,t,X))\,dt = \int_0^T f(\Phi(T,t,X)-B_0t)\,dt \le 2\int_0^T f(X-B_0(t-T)-B_0t)\,dt\le 2\delta(T),$$

which implies (3-4). Then we infer from Lemma 3.4 that

$$|\nabla\Phi(t,s)|_{0,\alpha;\Omega}\le C. \tag{3-8}$$

It follows from Lemma 3.3 and (3-7) that

$$\int_0^t g(r,\Phi(t,r,X))\,dr\le C\int_0^t g(r,X-B_0(r-t))\,dr\le C\delta(T)f(X+B_0t),$$

which implies

$$|u(t)|_{1,\alpha;f_+(t),\Omega}\le|u_0\circ\Phi(t,0)|_{1,\alpha;f_+(t),\Omega}+C\delta(T)\sup_{0\le s\le t}|F(s)\circ\Phi(t,s)|_{0,\alpha;g(s)\circ\Phi(t,s),\Omega}.$$

Using the fact $f(\Phi(t,0,X))\le 2f(X-B_0(0-t))=2f_+(t,X)$, we get

$$|u_0\circ\Phi(t,0)|_{1,\alpha;f_+(t),\Omega}\le 2|u_0\circ\Phi(t,0)|_{1,\alpha;f\circ\Phi(t,0),\Omega}.$$

Then by Lemma 2.2 and (3-8), we obtain

$$\begin{aligned}&|u(t)|_{1,\alpha;f_+(t),\Omega}\\&\quad\le C\big(|u_0\circ\Phi(t,0)|_{1,\alpha;f,\Omega}+\delta(T)\sup_{0\le s\le t}|F(s)|_{1,\alpha;g(s),\Omega}\big)\max\big(|\nabla\Phi(t,s)|^{\alpha}_{0;\Omega},1\big)\max\big(|\nabla\Phi(t,s)|_{0,\alpha;\Omega},1\big)\\&\quad\le C|u_0|_{1,\alpha;f,\Omega}+C\delta(T)\sup_{0\le s\le t}|F(s)|_{1,\alpha;g(s),\Omega}.\end{aligned}$$

This shows the first inequality of the lemma. □

Representation formula of the pressure. In this subsection, we give a representation formula of the pressure by using the symmetric extension.

Let (v,b,p) be a smooth solution of (1-1) in $[0,T]\times\Omega$ with the boundary condition (1-4). We make the following symmetric extension for the solution:

$$\bar v=Tv:=\big(T_ev^1,\dots,T_ev^{d-1},T_ov^d\big),\quad \bar b=Tb,\quad \bar p=T_ep.$$

Then $(\bar v,\bar b,\bar p)$ satisfies (1-1) in $[0,T]\times\mathbb{R}^d$ in the weak sense. Although the solution after the symmetric extension does not have the same smoothness as the original one, we have the following important observation.

Lemma 3.7. *Let h be a weight satisfying* (2-1). *Let $u=(u^1,\dots,u^d)$, $w=(w^1,\dots,w^d)\in C_h^{1,\alpha}(\Omega)$ be two vector fields with $u^d=w^d=0$ on $\partial\Omega$. Let $\bar u=Tu$ and $\bar w=Tw$. Then it holds that for $i,j=1,\dots,d$,*

$$\begin{gathered}|\partial_i\bar u^j\,\partial_j\bar w^i|_{0,\alpha;h}+|\partial_i\bar u^i\,\partial_j\bar w^j|_{0,\alpha;h}\le C|\nabla u|_{0,\alpha;h,\Omega}|\nabla w|_{0,\alpha;h,\Omega},\\|\bar u^j\,\partial_j\bar w^i|_{0,\alpha;h}+|\bar u^i\,\partial_j\bar w^j|_{0,\alpha;h}\le C|u|_{0,\alpha;h,\Omega}|\nabla w|_{0,\alpha;h,\Omega}.\end{gathered}$$

Proof. It is easy to verify that

$$\begin{aligned} \partial_i \bar u^j\, \partial_j \bar w^i &= T_e(\partial_i u^j\, \partial_j w^i), & \partial_i \bar u^i\, \partial_j \bar w^j &= T_e(\partial_i u^i\, \partial_j w^j),\\ \bar u^j\, \partial_j \bar w^i &= T_e(u^j\, \partial_j w^i), & \bar u^i\, \partial_j \bar w^j &= T_e(u^i\, \partial_j w^j) \quad \text{for } i = 1,\dots,d-1,\\ \bar u^j\, \partial_j \bar w^d &= T_o(u^j\, \partial_j w^d), & \bar u^d\, \partial_j \bar w^j &= T_0(u^d\, \partial_j w^j). \end{aligned}$$

Then the lemma follows easily from Lemma 2.5. □

Taking the divergence of the first equation of (1-1), we get

$$-\Delta \bar p = \partial_i (\bar v^j \partial_j \bar v^i - \bar b^j \partial_j \bar b^i).$$

Formally, we have

$$\nabla \bar p(t,X) = \nabla \int_{\mathbb{R}^d} N(X-Y)\partial_i (\bar v^j \partial_j \bar v^i - \bar b^j \partial_j \bar b^i)(t,Y)\, dY,$$

where $N(X)$ is the Newton potential. In terms of the Elsässer variables $\bar z_\pm(t,X)$, we have

$$\nabla \bar p(t,X) = \nabla \int_{\mathbb{R}^d} N(X-Y)\, \partial_i (\bar z_+^j \partial_j \bar z_-^i)(t,Y)\, dY.$$

However, this integral does not make sense for $\partial_i (\bar z_+^j \partial_j \bar z_-^i) \in C^{0,\alpha}$. To overcome this trouble, we introduce a smooth cut-off function $\theta(r)$ such that

$$\theta(r) = \begin{cases} 1 & \text{for } |r| \le 1,\\ 0 & \text{for } |r| \ge 2. \end{cases} \tag{3-9}$$

Integrating by parts, we can split $\nabla \bar p(t,X)$ as

$$\begin{aligned} -\nabla \bar p(t,X) = \int_{\mathbb{R}^d} \nabla N(X-Y)(\partial_i \bar z_+^j \partial_j \bar z_-^i)(t,Y)\, dY \\ + \int_{\mathbb{R}^d} \partial_i \partial_j \big(\nabla N(X-Y)(1-\theta(|X-Y|))\big)(\bar z_+^j \bar z_-^i)(t,Y)\, dY. \end{aligned} \tag{3-10}$$

It is easy to check that this representation makes sense for $\bar z_\pm \in W^{1,\infty}(\mathbb{R}^d)$.

We define

$$\begin{aligned} T_1 u &\triangleq \int_{\mathbb{R}^d} \nabla N(X-Y)\theta(|X-Y|)u(Y)\, dY,\\ T_{ij} w &\triangleq \int_{\mathbb{R}^d} \partial_i \partial_j \big(\nabla N(X-Y)(1-\theta(|X-Y|))\big) w(Y)\, dY. \end{aligned} \tag{3-11}$$

Let $u, w \in C^{1,\alpha}(\Omega)$ be two vector fields with $u^d = w^d = 0$ on $\partial\Omega$. Let $\bar u = Tu$ and $\bar w = Tw$ be the symmetric extension. We define

$$I(u,w) \triangleq T_1(\partial_i \bar u^j \partial_j \bar w^i - \partial_j \bar u^j \partial_i \bar w^i) + T_{ij}(\bar u^i \bar w^j). \tag{3-12}$$

Here and in what follows, the repeated index denotes the summation. Thanks to

$$\partial_i \bar u^j \partial_j \bar w^i - \partial_j \bar u^j \partial_i \bar w^i = \partial_i (\bar u^j \partial_j \bar w^i - \bar u^i \partial_j \bar w^j), \tag{3-13}$$

we infer from Lemmas A.1 and 3.7 that

$$|I(u,w)|_{0,\alpha;\Omega} \le C|u|_{0,\alpha;\Omega}|w|_{1,\alpha;\Omega}. \tag{3-14}$$

Using Lemma A.2 and (3-13), we calculate

$$\begin{aligned}
&\operatorname{div} I(u,w) + (\partial_i u^j\, \partial_j w^i - \partial_i u^i\, \partial_j w^j) \\
&\quad = \int_{\mathbb{R}^d} \nabla N(X-Y)\cdot\nabla\theta(|X-Y|)(\partial_i \bar{u}^j\, \partial_j \bar{w}^i - \partial_i \bar{u}^i\, \partial_j \bar{w}^j)(Y)\,dY \\
&\qquad\qquad - \int_{\mathbb{R}^d} \partial_i\partial_j\big(\nabla N(X-Y)\cdot\nabla\theta(|X-Y|)\big)(\bar{u}^j\bar{w}^i)(Y)\,dY \\
&\quad = \int_{\mathbb{R}^d} \partial_i\big(\nabla N(X-Y)\cdot\nabla\theta(|X-Y|)\big)\big(-\bar{u}^j\partial_j\bar{w}^i + \bar{u}^i\,\partial_j\bar{w}^j + \partial_j(\bar{u}^j\bar{w}^i)\big)(Y)\,dY \\
&\quad = \int_{\mathbb{R}^d} \partial_i\big(\nabla N(X-Y)\cdot\nabla\theta(|X-Y|)\big)(\bar{u}^i \operatorname{div}\bar{w} + \bar{w}^i \operatorname{div}\bar{u})(Y)\,dY,
\end{aligned}$$

which implies

$$\big|\operatorname{div} I(u,w) - (\partial_i u^j\, \partial_j w^i - \partial_i u^i\, \partial_j w^j)\big|_{0;\Omega} \le C\big(|u|_{0;\Omega}|\operatorname{div} w|_{0;\Omega} + |w|_{0,\Omega}|\operatorname{div} u|_{0;\Omega}\big). \tag{3-15}$$

In the case of $\mathbb{R}^d$, the pressure $p(t,X)$ can also be expressed as

$$-\nabla p(t,X) = I(z_+, z_-), \tag{3-16}$$

where

$$\begin{aligned}
I(u,w) \triangleq &\int_{\mathbb{R}^d} \nabla N(X-Y)\theta(|X-Y|)(\partial_i u^j\, \partial_j v^i)(Y)\,dY \\
&\qquad + \int_{\mathbb{R}^d} \partial_i\partial_j\big(\nabla N(X-Y)(1-\theta(|X-Y|))\big)(u^j v^i)(Y)\,dY.
\end{aligned} \tag{3-17}$$

Notice that the representation formula (3-16) is independent of the choice of θ in $I(u,w)$.

Proof of Theorem 3.1. Since we cannot find a well-posedness theory for the ideal MHD equations in the weighted Hölder spaces, we will present a complete proof of Theorem 3.1. In fact, we find that the proof of the existence part is very nontrivial.

Using the representation of the pressure (3-10), we rewrite the system (3-1) as

$$\begin{cases} \partial_t z_+ + Z_-\cdot\nabla z_+ = -I(z_+, z_-), \\ \partial_t z_- + Z_+\cdot\nabla z_- = -I(z_+, z_-), \\ z_+(0,X) = z_{+0}(X), \quad z_-(0,X) = z_{-0}(X). \end{cases} \tag{3-18}$$

Let $T > 0$ be determined later and

$$A_1 = |z_{+0}|_{1,\alpha;f,\Omega} + |z_{-0}|_{1,\alpha;f,\Omega}.$$

When A_1 is sufficiently small, T can be taken to be $+\infty$. The system (3-18) is solved by the following iteration scheme:

$$z_+^{(0)} = z_-^{(0)} = 0, \quad Z_+^{(n)} = z_+^{(n)} + B_0, \quad Z_-^{(n)} = z_-^{(n)} - B_0.$$

Let us inductively assume that $z_\pm^{(n)}$ satisfies

$$|z_+^{(n)}|_{1,\alpha;f_+,\Omega,T}\le 2C_1A_1,\quad |z_-^{(n)}|_{1,\alpha;f_-,\Omega,T}\le 2C_1A_1,$$

where C_1 is the constant in Proposition 3.6.

Take $T>0$ so that $4C_1A_1\delta(T)<1$. Then we have

$$|z_+^{(n)}|_{1,\alpha;f_+,\Omega,T}\,\delta(T)<\tfrac{1}{2},\quad |z_-^{(n)}|_{1,\alpha;f_-,\Omega,T}\,\delta(T)<\tfrac{1}{2}. \tag{3-19}$$

Now, the solution $z_+^{(n+1)}$, $z_-^{(n+1)}$ is determined by

$$\begin{cases}\partial_t z_+^{(n+1)}+Z_-^{(n)}\cdot\nabla z_+^{(n+1)}=-I(z_+^{(n)},z_-^{(n)}),\\ \partial_t z_-^{(n+1)}+Z_+^{(n)}\cdot\nabla z_-^{(n+1)}=-I(z_+^{(n)},z_-^{(n)}),\\ z_+^{(n+1)}(0,X)=z_{+0}(X),\quad z_-^{(n+1)}(0,X)=z_{-0}(X).\end{cases}$$

It follows from Proposition 3.6 that

$$\begin{aligned}|z_+^{(n+1)}|_{1,\alpha;f_+,\Omega,T}&\le C_1\big(|z_{+0}|_{1,\alpha;f,\Omega}+\delta(T)|z_+^{(n)}|_{1,\alpha;f_+,\Omega,T}|z_-^{(n)}|_{1,\alpha;f_-,\Omega,T}\big),\\ |z_-^{(n+1)}|_{1,\alpha;f_-,\Omega,T}&\le C_1\big(|z_{-0}|_{1,\alpha;f,\Omega}+\delta(T)|z_+^{(n)}|_{1,\alpha;f_+,\Omega,T}|z_-^{(n)}|_{1,\alpha;f_-,\Omega,T}\big).\end{aligned}$$

Here we used

$$|I(u,w)|_{1,\alpha;g,\Omega}\le C|\partial_i\bar u^j\partial_j\bar w^i-\partial_j\bar u^j\partial_i\bar w^i|_{0,\alpha;h}+C|\bar u\bar w|_{0;h}\le C|u|_{1,\alpha;h,\Omega}|w|_{1,\alpha;h,\Omega},$$

which follows from Lemma A.1 with $h(t,X)=f_+f_-(t,X)$ and Lemma 3.7.

Due to (3-19), we obtain

$$|z_+^{(n+1)}|_{1,\alpha;f_+,\Omega,T}\le 2C_1A_1,\quad |z_-^{(n+1)}|_{1,\alpha;f_-,\Omega,T}\le 2C_1A_1.$$

In particular, we show that for any n,

$$|z_+^{(n)}|_{1,\alpha;f_+,\Omega,T}\le C,\quad |z_-^{(n)}|_{1,\alpha;f_-,\Omega,T}\le C.$$

Next, we show that $\{z_\pm^{(n)}\}_{n\ge 0}$ are Cauchy sequences in $C^{0,\alpha}(\Omega)$. Indeed, we have

$$\begin{aligned}\partial_t(z_+^{(n+1)}-z_+^{(n)})+Z_-^{(n)}\cdot\nabla(z_+^{(n+1)}-z_+^{(n)})+(z_-^{(n)}-z_-^{(n-1)})\cdot\nabla z_+^{(n)}&\\ +I(z_+^{(n)}-z_+^{(n-1)},z_-^{(n)})+I(z_+^{(n-1)},z_-^{(n)}-z_-^{(n-1)})&=0,\\ \partial_t(z_-^{(n+1)}-z_-^{(n)})+Z_+^{(n)}\cdot\nabla(z_-^{(n+1)}-z_-^{(n)})+(z_+^{(n)}-z_+^{(n-1)})\cdot\nabla z_-^{(n)}&\\ +I(z_+^{(n)}-z_+^{(n-1)},z_-^{(n)})+I(z_+^{(n-1)},z_-^{(n)}-z_-^{(n-1)})&=0,\end{aligned}$$

$$(z_+^{(n+1)}-z_+^{(n)})(0,X)=0,\quad (z_-^{(n+1)}-z_-^{(n)})(0,X)=0.$$

Then it follows from Lemma 3.5 and (3-14) that

$$\begin{aligned}\big|(z_+^{(n+1)} - z_+^{(n)})(t)\big|_{0,\alpha;\Omega} &\le C\int_0^t \big|(z_-^{(n)} - z_-^{(n-1)})(s)\big|_{0,\alpha;\Omega}\,|\nabla z_+^{(n)}(s)|_{0,\alpha;\Omega}\,ds \\ &\quad + C\int_0^t \big|(z_+^{(n)} - z_+^{(n-1)})(s)\big|_{0,\alpha;\Omega}\,|z_-^{(n)}(s)|_{1,\alpha;\Omega}\,ds \\ &\quad + C\int_0^t \big|(z_-^{(n)} - z_-^{(n-1)})(s)\big|_{0,\alpha;\Omega}\,|z_+^{(n-1)}(s)|_{1,\alpha;\Omega}\,ds \\ &\le C_2\int_0^t \Big(\big|(z_+^{(n)} - z_+^{(n-1)})(s)\big|_{0,\alpha;\Omega} + \big|(z_-^{(n)} - z_-^{(n-1)})(s)\big|_{0,\alpha;\Omega}\Big)\,ds.\end{aligned}$$

Similarly, we have

$$\big|(z_-^{(n+1)} - z_-^{(n)})(t)\big|_{0,\alpha;\Omega} \le C_2\int_0^t \Big(\big|(z_+^{(n)} - z_+^{(n-1)})(s)\big|_{0,\alpha;\Omega} + \big|(z_-^{(n)} - z_-^{(n-1)})(s)\big|_{0,\alpha;\Omega}\Big)\,ds.$$

This implies that

$$\big|(z_+^{(n+1)} - z_+^{(n)})(t)\big|_{0,\alpha;\Omega} + \big|(z_-^{(n+1)} - z_-^{(n)})(t)\big|_{0,\alpha;\Omega} \le C(2C_2t)^n/n!.$$

Therefore, $z_+^{(n)}, z_-^{(n)}$ converge to some z_+, z_- uniformly in $[0,t]\times\Omega$ for any $0<t<T$. As $z_+^{(n)}, z_-^{(n)}$ are uniformly bounded in $C^{1,\alpha}$, we have $z_+, z_- \in C^{1,\alpha}$. Then $\nabla z_+^{(n)}, \nabla z_-^{(n)}$ converge to $\nabla z_+, \nabla z_-$ uniformly in $[0,t]\times\Omega$ for any $0<t<T$. Using the equations of $z_+^{(n+1)}, z_-^{(n+1)}$, we have $\partial_t z_+^{(n)},\ \partial_t z_-^{(n)}$ also converge uniformly in $[0,t]\times\Omega$ for any $0<t<T$. Thus, $z_+,\ z_- \in C^1([0,t]\times\overline{\Omega})$ satisfies (3-18) and $z_+^d = z_-^d = 0$ on $\partial\Omega$.

Finally, it remains to prove that if $\operatorname{div} z_{+0} = \operatorname{div} z_{-0} = 0$, then $\operatorname{div} z_+ = \operatorname{div} z_- = 0$. It follows from Lemma 3.5 and (3-15) that

$$\begin{aligned}&|\operatorname{div} z_+(t)|_{0;\Omega} \\ &\quad\le \int_0^t \big|\big(\partial_i z_+^j\,\partial_j z_-^i - \operatorname{div} I(z_+, z_-)\big)(s)\big|_{0;\Omega}\,ds \\ &\quad\le C\int_0^t \big(|\operatorname{div} z_+(s)|_{0;\Omega}\,|\operatorname{div} z_-(s)|_{0;\Omega} + |z_+(s)|_{0;\Omega}\,|\operatorname{div} z_-(s)|_{0;\Omega} + |\operatorname{div} z_+(s)|_{0;\Omega}\,|z_-(s)|_{0;\Omega}\big)\,ds \\ &\quad\le C\int_0^t \big(|\operatorname{div} z_+(s)|_{0;\Omega} + |\operatorname{div} z_-(s)|_{0;\Omega}\big)\,ds.\end{aligned}$$

Similarly,

$$|\operatorname{div} z_-(t)|_{0;\Omega} \le C\int_0^t \big(|\operatorname{div} z_+(s)|_{0;\Omega} + |\operatorname{div} z_-(s)|_{0;\Omega}\big)\,ds.$$

This implies that $\operatorname{div} z_+ = \operatorname{div} z_- = 0$.

Let us remark that $I(z_+, z_-)$ can be expressed as ∇p. Indeed, we can find $\theta_1, \theta_2 \in C^\infty(0,+\infty)$ such that $\theta_1'(r) = -\theta(r)N(r)$ and $\theta_2'(r) = (\theta(r)-1)N(r)$. Let $\theta_{ij}(X) = \partial_i\partial_j\theta_2(|X|)$ and

$$\begin{aligned}&I_*(u,w)(x) \\ &\quad= \int_{\mathbb{R}^d} \theta_1(|X-Y|)(\partial_i u^j\,\partial_j w^i - \partial_j u^j\,\partial_i w^i)(Y)\,dY + \int_{\mathbb{R}^d} \big(\theta_{ij}(X-Y) - \theta_{i,j}(-Y)\big)(u^j w^i)(Y)\,dY.\end{aligned}$$

Then we have $\nabla I_*(u, v) = I(u, v)$. Therefore, we can take $p = I_*(\bar{z}_+, \bar{z}_-)$, which satisfies $|p| \le C \ln(2 + |x|)$. This completes the proof of Theorem 3.1.

4. Global well-posedness for the viscous MHD equations

In this section, we study the global well-posedness for the viscous MHD equations in the slab domain $\Omega = \mathbb{R}^{d-1} \times [0, 1]$ with the Navier-slip boundary condition. Because we can reduce the slab domain $\Omega = \mathbb{R}^{d-1} \times [0, 1]$ to $\mathbb{R}^{d-1} \times \mathbb{T}$ by using the symmetric extension, we will consider more general domain $\Omega = \mathbb{R}^k \times \mathbb{T}^{d-k}$ for $2 \le k \le d$. The case $k = 1$ is more difficult and will be dealt in the future work.

In fact, $\Omega = \mathbb{R}^k \times \mathbb{T}^{d-k}$ is a special case of $\mathbb{R}^d$ periodic in $d-k$ directions $e_1, \dots, e_{d-k}$. We will assume that $e_1, \dots, e_{d-k}, B_0$ are linearly independent.

New formulation. Let $\mu_1 = \frac{1}{2}(\nu + \mu)$ and $\mu_2 = \frac{1}{2}(\nu - \mu)$. In terms of the Elsässer variables $Z_\pm = v \pm b$, the MHD equations (1-1) read

$$\begin{cases} \partial_t z_+ + Z_- \cdot \nabla z_+ = \mu_1 \Delta z_+ + \mu_2 \Delta z_- - \nabla p, \\ \partial_t z_- + Z_+ \cdot \nabla z_- = \mu_1 \Delta z_- + \mu_2 \Delta z_+ - \nabla p, \\ \operatorname{div} z_+ = \operatorname{div} z_- = 0, \end{cases} \tag{4-1}$$

where $z_\pm = Z_+ \pm B_0$. In the case of $\nu = \mu$ (thus, $\mu_2 = 0$), the formulation (4-1) plays a crucial role in the proof of [Cai and Lei 2016; He et al. 2016]. To deal with the case of $\nu \ne \mu$, we need to introduce the key decomposition

$$z_+ = z_+^{(1)} + z_+^{(2)}, \quad z_- = z_-^{(1)} + z_-^{(2)},$$

where $z_\pm^{(1)}$ and $z_\pm^{(2)}$ are determined by

$$\begin{cases} \partial_t z_+^{(1)} + Z_- \cdot \nabla z_+^{(1)} = \mu_1 \Delta z_+^{(1)} - \nabla p_+^{(1)}, \\ \partial_t z_-^{(1)} + Z_+ \cdot \nabla z_-^{(1)} = \mu_1 \Delta z_-^{(1)} - \nabla p_-^{(1)}, \\ \operatorname{div} z_+^{(1)} = \operatorname{div} z_-^{(1)} = 0, \\ z_+^{(1)}(0) = z_+(0), \ z_-^{(1)}(0) = z_-(0), \end{cases} \tag{4-2}$$

and

$$\begin{cases} \partial_t z_+^{(2)} + Z_- \cdot \nabla z_+^{(2)} = \mu_1 \Delta z_+^{(2)} + \mu_2 \Delta z_- - \nabla p_+^{(2)}, \\ \partial_t z_-^{(2)} + Z_+ \cdot \nabla z_-^{(2)} = \mu_1 \Delta z_-^{(2)} + \mu_2 \Delta z_+ - \nabla p_-^{(2)}, \\ \operatorname{div} z_+^{(2)} = \operatorname{div} z_-^{(2)} = 0, \\ z_+^{(2)}(0) = z_-^{(2)}(0) = 0. \end{cases} \tag{4-3}$$

To estimate $z_\pm^{(1)}$, we rewrite (4-2) as

$$\begin{cases} \partial_t z_+^{(1)} + Z_-^{(1)} \cdot \nabla z_+^{(1)} = \mu_1 \Delta z_+^{(1)} - z_-^{(2)} \cdot \nabla z_+^{(1)} - I(z_-^{(2)}, z_+^{(1)}) - I(z_-^{(1)}, z_+^{(1)}), \\ \partial_t z_-^{(1)} + Z_+^{(1)} \cdot \nabla z_-^{(1)} = \mu_1 \Delta z_-^{(1)} - z_+^{(2)} \cdot \nabla z_-^{(1)} - I(z_+^{(2)}, z_-^{(1)}) - I(z_+^{(1)}, z_-^{(1)}), \end{cases} \tag{4-4}$$

where $I(u, w)$ is defined by (3-17). We also need to use the equation of $J_{\pm}^{(1)} = \operatorname{curl} z_{\pm}^{(1)}$, which is given by

$$\begin{cases} \partial_t J_+^{(1)} + Z_-^{(1)} \cdot \nabla J_+^{(1)} + \nabla z_-^{(1)} \wedge \nabla z_+^{(1)} + \operatorname{curl}(z_-^{(2)} \cdot \nabla z_+^{(1)}) = \mu\, \Delta J_+^{(1)}, \\ \partial_t J_-^{(1)} + Z_+^{(1)} \cdot \nabla J_-^{(1)} + \nabla z_+^{(1)} \wedge \nabla z_-^{(1)} + \operatorname{curl}(z_+^{(2)} \cdot \nabla z_-^{(1)}) = \mu\, \Delta J_-^{(1)}. \end{cases} \tag{4-5}$$

Here $A \wedge B = (AB) - (AB)^T$ is understood as matrix multiplication.

To estimate $z_{\pm}^{(2)}$, we need to introduce another formulation in terms of the stream function $\psi_{\pm}^{(2)} = \Delta^{-1} \operatorname{curl} z_{\pm}^{(2)}$, which satisfies

$$\begin{cases} \partial_t \psi_+^{(2)} + \Delta^{-1} \operatorname{curl}(Z_- \cdot \nabla z_+^{(2)}) = \mu_1\, \Delta \psi_+^{(2)} + \mu_2 J_-, \\ \partial_t \psi_-^{(2)} + \Delta^{-1} \operatorname{curl}(Z_+ \cdot \nabla z_-^{(2)}) = \mu_1\, \Delta \psi_-^{(2)} + \mu_2 J_+, \end{cases}$$

where

$$J_{\pm} = \operatorname{curl} z_{\pm} = J_{\pm}^{(1)} + \operatorname{curl} z_{\pm}^{(2)}. \tag{4-6}$$

We introduce

$$\mathrm{II}_1(u, w) \triangleq \Delta^{-1} \operatorname{curl}\ \operatorname{div}(u \otimes w),$$
$$\mathrm{II}_2(u, w) \triangleq \Delta^{-1} \operatorname{curl}(u \cdot \nabla w) - u \cdot \nabla\, \Delta^{-1} \operatorname{curl} w.$$

So, we get

$$\Delta^{-1} \operatorname{curl}(Z_- \cdot \nabla z_+^{(2)}) = Z_-^{(1)} \cdot \nabla \psi_+^{(2)} + \mathrm{II}_1(z_-^{(2)}, z_+^{(2)}) + \mathrm{II}_2(z_-^{(1)}, z_+^{(2)}).$$

Then we deduce that

$$\begin{cases} \partial_t \psi_+^{(2)} + Z_-^{(1)} \cdot \nabla \psi_+^{(2)} + \mathrm{II}_2(z_-^{(1)}, z_+^{(2)}) + \mathrm{II}_1(z_-^{(2)}, z_+^{(2)}) = \mu_1\, \Delta \psi_+^{(2)} + \mu_2 J_-, \\ \partial_t \psi_-^{(2)} + Z_+^{(1)} \cdot \nabla \psi_-^{(2)} + \mathrm{II}_2(z_+^{(1)}, z_-^{(2)}) + \mathrm{II}_1(z_+^{(2)}, z_-^{(2)}) = \mu_1\, \Delta \psi_-^{(2)} + \mu_2 J_+. \end{cases} \tag{4-7}$$

A direct calculation shows

$$-\big(\Delta^{-1} \operatorname{curl}(u \cdot \nabla w)\big)^{jk} = \Delta^{-1}\big(\partial_k \partial_i (u^i w^j) - \partial_j \partial_i (u^i w^k)\big) = -R_k R_i (u^i w^j) + R_j R_i (u^i w^k),$$
$$-\big(u \cdot \nabla(\Delta^{-1} \operatorname{curl} w)\big)^{jk} = u^i\, \partial_i \Delta^{-1}(\partial_k w^j - \partial_j w^k) = u^i(-R_i R_k w^j + R_i R_j w^k),$$

where R_i is the Riesz transform defined by $R_i = \partial_i (-\Delta)^{-\frac{1}{2}}$. This gives

$$\mathrm{II}_2(u, w)^{jk} = [u^i, R_i R_j] w^k - [u^i, R_i R_k] w^j. \tag{4-8}$$

Weighted $C^{1,\alpha}$ estimates for the parabolic equation. We consider the parabolic equation with variable coefficients

$$\partial_t u - \gamma\, \partial_i (a_{ij}\, \partial_j u) + F_1 + F_2 + \partial_i G^i = 0, \tag{4-9}$$

where $\gamma > 0$ and the coefficients $a_{ij}(t, X)$ satisfy

$$\sup_{t \in [0,T]} \big(|a_{ij}(t) - \delta_{ij}|_0 + (1 + \gamma t)^{\frac{\alpha}{2}} [a_{ij}(t)]_\alpha\big) \le \varepsilon_0 \tag{4-10}$$

for some $\alpha \in (0, 1)$, $\varepsilon_0 > 0$ and $T > 0$.

Let $f(t,X)$ and $h(t,X)$ be two weight functions satisfying (2-1) with a uniform constant c_0 independent of t and

$$\int_0^t H(2\gamma(t-s))h(s,X)\,ds \le c_0^{-1} f(t,X), \quad H(2\gamma(t-s))f(s,X) \le c_0^{-1} f(t,X) \tag{4-11}$$

for all $0 \le s < t \le T,\ X \in \mathbb{R}^d$, where

$$H(t)\varphi(X) = \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d} e^{-\frac{|X-Y|^2}{4t}} \varphi(Y)\,dY.$$

Let $\delta > 0$. We introduce

$$\begin{aligned}\Lambda_1(T,F_1,F_2,G,f,h) \triangleq \sup_{0<t\le T} \Big(&|F_1(t)|_{1,\alpha;h(t),(1+\gamma t)^{1/2}} + \gamma^{-1}((\gamma t)^{\frac12} + (\gamma t)^{1+\frac{\delta}{2}})|F_2(t)|_{0,\alpha;f(t)} \\ &+ \gamma^{-1}(1+\gamma t)^{\frac12}|G(t)|_{0,\alpha;f(t),(1+\gamma t)^{1/2}}\Big),\end{aligned}$$

and

$$\begin{aligned}\Lambda_0(T,F_1,F_2,G,f,h) \triangleq \sup_{0<t\le T} \Big(&|F_1(t)|_{1,\alpha;h(t),(\gamma t)^{1/2}} + \gamma^{-1}((\gamma t)^{1-\frac{\alpha}{2}} + (\gamma t)^{1+\frac{\delta}{2}})|F_2(t)|_{0,\alpha;f(t)} \\ &+ \gamma^{-1}((\gamma t)^{\frac12} + (\gamma t)^{\frac{1-\alpha}{2}})|G(t)|_{0,\alpha;f(t),(\gamma t)^{1/2}}\Big).\end{aligned}$$

Proposition 4.1. *There exist $\varepsilon_0 > 0$ and $C > 0$ independent of γ and T such that*

$$\sup_{0<t\le T} |u(t)|_{1,\alpha;f(t),(1+\gamma t)^{1/2}} \le C\big(|u(0)|_{1,\alpha;f(0),1} + \Lambda_1(T,F_1,F_2,G,f,h)\big),$$
$$\sup_{0<t\le T} |u(t)|_{1,\alpha;f(t),(\gamma t)^{1/2}} \le C\big(|u(0)|_{0,\alpha;f(0)} + \Lambda_0(T,F_1,F_2,G,f,h)\big).$$

Proof. Let us first consider the case $a_{ij} = \delta_{ij}$. Then we get

$$u(t) = H(\gamma t)u(0) + \int_0^t \big(H(\gamma(t-s))(F_1(s)+F_2(s)) + \partial_i H(\gamma(t-s))G^i(s)\big)\,ds.$$

Using $H(2\gamma t)f(0) \le c_0^{-1} f(0,X)$, we get by Lemma A.4 that

$$|H(\gamma t)u(0)|_{1,\alpha;f(t),(1+\gamma t)^{1/2}} \le C|H(\gamma t)u(0)|_{1,\alpha;H(2\gamma t)f(0),(1+\gamma t)^{1/2}} \le C|u(0)|_{1,\alpha;f(0),1},$$
$$|H(\gamma t)u(0)|_{1,\alpha;f(t),(\gamma t)^{1/2}} \le C|H(\gamma t)u(0)|_{1,\alpha;H(2\gamma t)f(0),(\gamma t)^{1/2}} \le C|u(0)|_{0,\alpha;f(0)}.$$

By (4-11) and Lemma A.4, we have

$$\begin{aligned}\left|\int_0^t H(\gamma(t-s))F_1(s)\,ds\right|_{1,\alpha;f(t),(k+\gamma t)^{1/2}} &\le C \sup_{0<s<t} |H(\gamma(t-s))F_1(s)|_{1,\alpha;H(2\gamma(t-s))h(s),(k+\gamma t)^{1/2}} \\ &= C \sup_{0<s<t} |H(\gamma(t-s))F_1(s)|_{1,\alpha;H(2\gamma(t-s))h(s),(k+\gamma s+\gamma(t-s))^{1/2}} \\ &\le C \sup_{0<s<t} |F_1(s)|_{1,\alpha;h(s),(k+\gamma s)^{1/2}},\end{aligned}$$

and by Lemma A.4,

$$\left|\int_0^t H(\gamma(t-s))F_2(s)\,ds\right|_{1,\alpha;f(t),(k+\gamma t)^{1/2}} \le C\int_0^t |H(\gamma(t-s))F_2(s)|_{1,\alpha;H(2\gamma(t-s))f(s),(k+\gamma t)^{1/2}}\,ds$$
$$\le C\int_0^t \frac{\varphi_\alpha(\sqrt{k+\gamma t})}{\varphi_\alpha(\sqrt{\gamma(t-s)})}|F_2(s)|_{0,\alpha;f(s)}\,ds$$

for $k=0,1$. Recall that $\varphi_\alpha(R)=\max(R,R^{1+\alpha})$ for $k=0,1$,

$$\int_0^t \varphi_\alpha(\sqrt{\gamma(t-s)})^{-1}\min\big((\gamma s)^{-1+\frac{\alpha+k(1-\alpha)}{2}},(\gamma s)^{-1-\frac{\delta}{2}}\big)\,ds \le \int_0^t (\gamma(t-s))^{-\frac{1}{2}}(\gamma s)^{-1+\frac{\alpha+k(1-\alpha)}{2}}\,ds$$
$$\le C\gamma^{-1}(\gamma t)^{-\frac{(1-k)(1-\alpha)}{2}},$$

and

$$\begin{aligned}\int_0^t &\varphi_\alpha(\sqrt{\gamma(t-s)})^{-1}\min\big((\gamma s)^{-1+\frac{\alpha+k(1-\alpha)}{2}},(\gamma s)^{-1-\frac{\delta}{2}}\big)\,ds\\ &\le \int_0^t (\gamma(t-s))^{-\frac{1+\alpha}{2}}\min\big((\gamma s)^{-1+\frac{\alpha+k(1-\alpha)}{2}},(\gamma s)^{-1-\frac{\delta}{2}}\big)\,ds\\ &\le C\int_0^{\frac{t}{2}} (\gamma t)^{-\frac{1+\alpha}{2}}\min\big((\gamma s)^{-1+\frac{\alpha+k(1-\alpha)}{2}},(\gamma s)^{-1-\frac{\delta}{2}}\big)\,ds + \int_{\frac{t}{2}}^t (\gamma(t-s))^{-\frac{1+\alpha}{2}}(\gamma t)^{-1}\,ds\\ &\le C\gamma^{-1}(\gamma t)^{-\frac{1+\alpha}{2}}.\end{aligned}$$

Thus, we have

$$\begin{aligned}\int_0^t &\frac{\varphi_\alpha(\sqrt{k+\gamma t})}{\varphi_\alpha(\sqrt{\gamma(t-s)})}\min\big((\gamma s)^{-1+\frac{\alpha+k(1-\alpha)}{2}},(\gamma s)^{-1-\frac{\delta}{2}}\big)\,ds\\ &\le C\gamma^{-1}\max\big((k+\gamma t)^{\frac{1}{2}},(k+\gamma t)^{\frac{1+\alpha}{2}}\big)\min\big((\gamma t)^{-\frac{(1-k)(1-\alpha)}{2}},(\gamma t)^{-\frac{1+\alpha}{2}}\big)\le C\gamma^{-1}.\end{aligned}$$

Therefore, we deduce that for $k=0,1$ and $j=1,2$,

$$\left|\int_0^t H(\gamma(t-s))F_j(s)\,ds\right|_{1,\alpha;f(t),(k+\gamma t)^{1/2}} \le C\Lambda_k(T,F_1,F_2,G,f,h).$$

It follows from Lemmas 2.3 and A.3 that for $k=0,1$,

$$\begin{aligned}&\left|\int_0^t \partial_i H(\gamma(t-s))G^i(s))\,ds\right|_{1,\alpha;f(t),(k+\gamma t)^{1/2}}\\ &\le C\gamma^{-1}\sup_{0<s<t}\Big((\gamma s)^{\frac{1}{2}}(\gamma(t-s))^{\frac{1}{2}}\big|\partial_i H(\gamma(t-s))G^i(s)\big|_{0,\alpha;f(t)}\\ &\qquad+\varphi_\alpha(\sqrt{k+\gamma s})(\gamma(t-s))^{1-\frac{\alpha}{2}}\big|\nabla\partial_i H(\gamma(t-s))G^i(s)\big|_{0;f(t)}\\ &\qquad\qquad+\varphi_\alpha(\sqrt{k+\gamma s})(\gamma(t-s))^{\frac{3-\alpha}{2}}[\nabla\partial_i H(\gamma(t-s))G^i(s)]_{1;f(t)}\Big)\\ &\le C\gamma^{-1}\sup_{0<s<t}\big((\gamma s)^{\frac{1}{2}}|G(s)|_{0,\alpha;f(s)}+\varphi_\alpha(\sqrt{k+\gamma s})[G(s)]_{\alpha;f(s)}\big)\end{aligned}$$

$$\le C\gamma^{-1}\sup_{0<s<t}\big((k+\gamma s)^{\frac12}+(k+\gamma s)^{\frac{1-\alpha}{2}}\big)|G(s)|_{0,\alpha;f(s),(k+\gamma s)^{1/2}}$$
$$\le C\Lambda_k(T,F_1,F_2,G,f,h).$$

Summing up, we conclude the proof for the case $a_{ij}=\delta_{ij}$.

To deal with the general case, we rewrite (4-9) as

$$\partial_t u-\gamma\,\Delta u+F_1+F_2+\partial_i\widehat{G}^i=0,$$

where $\widehat{G}^i=G^i-\gamma(a_{ij}-\delta_{ij})\,\partial_j u$. Thus, we have

$$\sup_{0<t\le T}|u(t)|_{1,\alpha;f(t),(k+\gamma t)^{1/2}}\le C\big(|u(0)|_{1,\alpha;f(0),k}+\Lambda_k(T,F_1,F_2,\widehat{G},f,h)\big)$$

for $k=0,1$, where

$$\begin{aligned}&\Lambda_k(T,F_1,F_2,\widehat{G},f,h)\\&\qquad\le\Lambda_k(T,F_1,F_2,G,f,h)+\sup_{0\le t\le T}\sup_i\big((k+\gamma t)^{\frac12}+(\gamma t)^{\frac{1-\alpha}{2}}\big)\big|(a_{ij}-\delta_{ij})\,\partial_j u(t)\big|_{0,\alpha;f(t),(k+\gamma t)^{1/2}},\end{aligned}$$

and by (4-10),

$$\begin{aligned}\big|(a_{ij}-\delta_{ij})\,\partial_j u(t)\big|_{0,\alpha;f(t),(k+\gamma t)^{1/2}}&\le C|a_{ij}(t)-\delta_{ij}|_{0,\alpha;1,(k+\gamma t)^{1/2}}|\partial_j u(t)|_{0,\alpha;f(t),(k+\gamma t)^{1/2}}\\&\le C\varepsilon_0|\nabla u(t)|_{0,\alpha;f(t),(k+\gamma t)^{1/2}}\\&\le C\varepsilon_0\min\big((k+\gamma t)^{-\frac12},(k+\gamma t)^{-\frac{1-\alpha}{2}}\big)|u(t)|_{1,\alpha;f(t),(k+\gamma t)^{1/2}}.\end{aligned}$$

This shows that

$$\begin{aligned}&\sup_{0<t\le T}|u(t)|_{1,\alpha;f(t),(k+\gamma t)^{1/2}}\\&\qquad\le C\big(|u(0)|_{1,\alpha;f(0)}+\Lambda_k(T,F_1,F_2,G,f,h)+\varepsilon_0\sup_{0\le t\le T}|u(t)|_{1,\alpha;f(t),(k+\gamma t)^{1/2}}\big),\end{aligned}$$

which gives the desired result by taking ε_0 such that $C\varepsilon_0\le\frac12$. □

Weighted $C^{1,\alpha}$ estimates for the transport-diffusion equation. We consider the transport-diffusion equation with general form

$$\partial_t u+Z\cdot\nabla u-\gamma\,\Delta u+F_1+F_2+\partial_i G^i=0,\quad u(0,X)=u_0(X).\tag{4-12}$$

Given the divergence-free vector field $Z(t,X)\in C^1([0,T]\times\mathbb{R}^d)$ and $s\in[0,T]$, we define

$$\frac{d}{dt}\Phi(s,t,X)=Z(t,\Phi(s,t,X)),\quad \Phi(s,s,X)=X.$$

We denote by $D\Phi$ and $\nabla\Phi$ the matrix with the convention

$$(D\Phi)_{ij}=\partial_j\Phi^i,\quad (\nabla\Phi)_{ij}=\partial_i\Phi^j.$$

That is, $(D\Phi)=(\nabla\Phi)^T$. We introduce

$$b=(D\Phi)^{-1},\quad a=(D\Phi)^{-1}(\nabla\Phi)^{-1},\quad a_{ij}=b_{ik}b_{kj}.$$

For $v(t,X)$ defined in $[0,T]\times\mathbb{R}^d$, we define

$$v^*(t,X)\triangleq v(t,\Phi(s,t,X)).$$

Using the formulas

$$(\operatorname{div}G)\circ\Phi=\operatorname{div}\big((D\Phi)^{-1}G\circ\Phi\big),\quad (\Delta u)\circ\Phi=\operatorname{div}\big((D\Phi)^{-1}(\nabla\Phi)^{-1}\,\nabla u\circ\Phi\big),$$

we can transform (4-12) into the form

$$\partial_t u^*(t)-\gamma\partial_i(a_{ij}\partial_j u^*(t))+F_1^*+F_2^*+\partial_i G_*^i=0, \tag{4-13}$$

where $G_*^i=b_{ij}(G^*)^j$.

We introduce the weight functions $f(t,X)$, $\hat f(t,X)$, $h(t,X)$, which satisfy (2-1) with a uniform constant c_0 and

$$\begin{gathered}\int_0^t H(2\gamma(t-s))h_\pm(s,X)\,ds\le c_0^{-1}\hat f(t,X)\quad\text{for all } 0\le t\le T,\ X\in\mathbb{R}^d,\\ \int_0^T f_\pm(t,X\pm B_0t)\,dt=\int_0^T f(t,X\pm 2B_0t)\,dt\le c_0^{-1},\\ H(2\gamma(t-s))\hat f(s,X)\le c_0^{-1}\hat f(t,X)\quad\text{for all } 0\le s\le t\le T,\ X\in\mathbb{R}^d,\end{gathered} \tag{4-14}$$

where we set

$$f_\pm(t,X)=U(\pm t)f(t,X),\quad U(t)f(s,X)=f(s,X+B_0t).$$

Proposition 4.2. *There exist $\varepsilon_1>0$ and $C>0$ independent of γ and T such that if*

$$|Z(t)+B_0|_{1,\alpha;f_-(t),(1+\gamma t)^{1/2}}<\varepsilon_1$$

and (4-14) *holds for the minus sign, then it holds that for $k=0,1$,*

$$\sup_{0\le t\le T}|u(t)|_{1,\alpha;\hat f_+(t),(k+\gamma t)^{1/2}}\le C\big(|u_0|_{1,\alpha;\hat f(0),k}+\Lambda_k(T,F_1,F_2,G,\hat f_+,h)\big).$$

Similarly, if

$$|Z(t)-B_0|_{1,\alpha;f_+(t),(1+\gamma t)^{1/2}}<\varepsilon_1,$$

and (4-14) *holds for the plus sign, then it holds that for $k=0,1$,*

$$\sup_{0\le t\le T}|u(t)|_{1,\alpha;\hat f_-(t),(k+\gamma t)^{1/2}}\le C\big(|u_0|_{1,\alpha;\hat f(0),k}+\Lambda_k(T,F_1,F_2,G,\hat f_-,h)\big).$$

Proof. We only consider the case $|Z(t)+B_0|_{1,\alpha;f_-(t),(1+\gamma t)^{1/2}}<\varepsilon_1$. In this case, similar to (3-7), we have

$$|\Phi(s,t,X)+B_0(t-s)-X|<2\quad\text{for } 0\le t\le s\le T.$$

Then we get by (2-1) and (4-14) that

$$\sup_{t\le s\le T}|\nabla Z(s)|_{0,\alpha;f_-(s)}\int_0^T f_-\big(s,\Phi(T,s,X)\big)\,ds$$
$$\le \varepsilon_1(1+\gamma t)^{-\frac12}c_0^{-1}\int_0^T f_-\big(s,X-B_0(s-T)\big)\,ds\le \varepsilon_1(1+\gamma t)^{-\frac12}c_0^{-1}, \tag{4-15}$$

and by (2-1),

$$U(s-t)h(t)=h(t,X+B_0(s-t))\ge c_0h(t)\circ\Phi(s,t), \tag{4-16}$$
$$U(s)\hat f(t)=U(s-t)\hat f_+(t)\ge c_0\hat f_+(t)\circ\Phi(s,t). \tag{4-17}$$

Now we fix $s\ge 0$ and assume $0\le t\le s\le T$. With (4-15), we infer from Lemma 3.4 that

$$|\nabla\Phi(s,t)-\mathrm{Id}|_{0,\alpha}\le C\varepsilon_1(1+\gamma t)^{-\frac12}. \tag{4-18}$$

This implies that

$$|a_{ij}(t)-\delta_{ij}|_{0,\alpha}\le C\varepsilon_1(1+\gamma t)^{-\frac12},\quad |b_{ij}(t)|_{0,\alpha;1,(1+\gamma t)^{1/2}}\le C.$$

Using (4-14), it is easy to verify that

$$H(2\gamma(t-\tau))U(s)\hat f(\tau,X)=U(s)H(2\gamma(t-\tau))\hat f(\tau,X)\le c_0^{-1}U(s)\hat f(t),$$

and

$$\int_0^t H(2\gamma(t-\tau))U(s-\tau)h(\tau,X)\,d\tau=\int_0^t H(2\gamma(t-\tau))U(s)h_-(\tau,X)\,d\tau\le c_0^{-1}U(s)\hat f(t).$$

Therefore, if we take $\varepsilon_1>0$ so that $C\varepsilon_1\le\varepsilon_0$, then we can apply Proposition 4.1 to obtain

$$\sup_{0<t\le s}|u^*(t)|_{1,\alpha;U(s)\hat f(t),(k+\gamma t)^{1/2}}$$
$$\le C\big(|u_0\circ\Phi(s,0)|_{1,\alpha;U(s)\hat f(0),k}+\Lambda_k(s,F_1^*,F_2^*,G_*,U(s)\hat f,U(s-\cdot)h)\big).$$

Thanks to (4-18), we get by Lemma 2.4, (4-16) and (4-17) that

$$|u_0\circ\Phi(s,0)|_{1,\alpha;U(s)\hat f(0),k}\le C|u_0\circ\Phi(s,0)|_{1,\alpha;\hat f(0)\circ\Phi(s,0),k}\le C|u_0|_{1,\alpha;\hat f(0),k},$$
$$|F_2^*(t)|_{0,\alpha;U(s)\hat f(t)}\le C|F_2^*(t)|_{0,\alpha;\hat f_+(t)\circ\Phi(s,t)}\le C|F_2(t)|_{0,\alpha;\hat f_+(t)},$$
$$|F_1^*(t)|_{1,\alpha;U(s-t)h(t),(k+\gamma t)^{1/2}}\le C|F_1^*(t)|_{1,\alpha;h(t)\circ\Phi(s,t),(k+\gamma t)^{1/2}}\le C|F_1(t)|_{1,\alpha;h(t),(k+\gamma t)^{1/2}},$$

and

$$\begin{aligned}|G_*(t)|_{0,\alpha;U(s)\hat f(t),(k+\gamma t)^{1/2}}&\le C|G_*(t)|_{0,\alpha;\hat f_+(t)\circ\Phi(s,t),(k+\gamma t)^{1/2}}\\&\le C|b(t)|_{0,\alpha;1,(1+\gamma t)^{1/2}}|G(t)\circ\Phi(s,t)|_{0,\alpha;\hat f_+(t)\circ\Phi(s,t),(k+\gamma t)^{1/2}}\\&\le C|G(t)|_{0,\alpha;\hat f_+(t),(k+\gamma t)^{1/2}}.\end{aligned}$$

This proves

$$\Lambda_k\big(s,F_1^*,F_2^*,G_*,U(s)\hat f,U(s-t)h\big)\le C\Lambda_k(s,F_1,F_2,G,\hat f_+,h).$$

Therefore, we conclude

$$\sup_{0<t\le s} |u^*(t)|_{1,\alpha;U(s)\hat f(t),(k+\gamma t)^{1/2}} \le C\big(|u_0|_{1,\alpha;\hat f(0),k} + \Lambda_k(s,F_1,F_2,G,\hat f_+,h)\big).$$

Thanks to $u^*(s)=u(s)$ and $U(s)\hat f(s)=\hat f_+(s)$, we have

$$|u(s)|_{1,\alpha;\hat f_+(s),(k+\gamma s)^{1/2}} \le C\big(|u_0|_{1,\alpha;\hat f(0),k} + \Lambda_k(s,F_1,F_2,G,\hat f_+,h)\big)$$

for all $0<s\le T$. The case $s=0$ is trivial. □

Main result. Let us first introduce the weight functions

$$f(t)=H(1+2\mu_1 t)\varphi_1, \quad f_1(t)=H(1+2\mu_1 t)\varphi_0,$$

where if $B_0=(1,0,\dots,0)$, we may take

$$\varphi_1(X)=|x_1^2+x_2^2|^{-\frac{1+\delta}{2}}, \quad \varphi_0(X)=|x_2|^{-\delta} \tag{4-19}$$

for some $0<\delta<\frac{1}{2}$. Let

$$g(t,X)\triangleq\int_{\mathbb{R}^d}\frac{f_+(t,Y)f_-(t,Y)}{1+|X-Y|^{d+1}}\,dY.$$

We introduce

$$M_\pm(t)\triangleq\sup_{0\le\tau\le t}\big(|z_\pm^{(1)}(\tau)|_{1,\alpha;f_\pm(\tau),(1+\mu_1\tau)^{1/2}}+|J_\pm^{(1)}(\tau)|_{1,\alpha;f_\pm(\tau),(\mu_1\tau)^{1/2}}+\mu_1^{-1}|\psi_\pm^{(2)}(\tau)|_{1,\alpha;f_1(\tau),(\mu_1\tau)^{1/2}}\big).$$

The main result of this section is stated as follows.

Theorem 4.3. *Let $\alpha\in(0,1)$. There exists $\varepsilon_2>0$ such that if $M_\pm(0)+|\mu_2|/\mu_1\le\varepsilon\le\varepsilon_2$, then there exists a global in time unique solution $(z_+,z_-)\in L^\infty((0,+\infty)\times\Omega)$, with the pressure p determined by* (3-16), *to the viscous MHD equations* (4-1) *satisfying*

$$M_\pm(t)\le C\varepsilon \quad \textit{for any } t\in[0,+\infty).$$

Remark 4.4. Since $M_\pm(0)\sim|z_\pm(0)\langle(x_1,x_2)\rangle^{1+\delta}|_{1,\alpha}$, the initial data decays at infinity only in two directions. This is crucial for the global well-posedness in domains like $\mathbb{R}^2$ and $\mathbb{R}^2\times[0,1]$.

Remark 4.5. The condition $|\mu_2|\le\varepsilon\mu_1$ is crucial to our proof. Although μ_2/μ_1 is small, the smallness is independent of μ_1. It remains open whether one can prove a similar result for any $\mu>0$, $\nu>0$.

Remark 4.6. In numerical simulation, μ_2 is usually taken to be zero, although it is unreasonable in physics. However, our result provides a theoretical base for the validity of such a choice, because our result shows that a small discrepancy between the dissipation coefficients does not change the dynamics of the system.

To proceed, we need to verify that the weight functions introduced here satisfy some key properties, (2-3) and (4-14).

With the choice of (4-19), it is easy to check that for $k=0,1$,

$$C^{-1}R^d\min(\varphi_k(X),R^{-k-\delta})\le\int_{B(X,R)}\varphi_k(Y)\,dY\le CR^d\min(\varphi_k(X),R^{-k-\delta}),$$

$$\int_{\mathbb{R}}\varphi_1(X+B_0t)\,dt\le C\varphi_0(X),$$

which imply

$$C^{-1}\min\bigl(\varphi_1(X),(1+\mu_1t)^{-\frac{1+\delta}{2}}\bigr)\le f(t,X)\le C\min\bigl(\varphi_1(X),(1+\mu_1t)^{-\frac{1+\delta}{2}}\bigr),\tag{4-20}$$

$$C^{-1}\min\bigl(\varphi_0(X),(1+\mu_1t)^{-\frac{\delta}{2}}\bigr)\le f_1(t,X)\le C\min\bigl(\varphi_0(X),(1+\mu_1t)^{-\frac{\delta}{2}}\bigr),\tag{4-21}$$

$$\int_{\mathbb{R}}f(t,X+B_0s)\,ds\le Cf_1(t,X).\tag{4-22}$$

Therefore,

$$\int_{B(X,R)}f_1(t,Y)\,dY\le CR^d\min\bigl(R^{-\delta},(1+\mu_1t)^{-\frac{1+\delta}{2}}\bigr)\tag{4-23}$$

and

$$\int_{B(X,R)}h(Y)\,dY\le Ch(X),\tag{4-24}$$

which is true for $h=1$, $f(t)$, $f_1(t)$, and $f_\pm(t)$ by translation. Thus,

$$\int_{\mathbb{R}^d}\frac{f_\pm(t,Y)\,dY}{R^{d+1}+|X-Y|^{d+1}}\le CR^{-1}f_\pm(t,X).\tag{4-25}$$

Lemma 4.7. (1) *The weight functions $f(t,X)$, $f_1(t,X)$, $g(t,X)$ satisfy* (2-3) *with $R=(1+\mu_1t)^{\frac12}$ and a uniform constant c_0 independent of t.*

(2) *Property* (4-14) *with $\gamma=\mu_1$ holds true for $(\hat f,h)=(f,g)$ and $(\hat f,h)=(f_1,f_-)$ for the minus sign or $(\hat f,h)=(f_1,f_+)$ for the plus sign.*

Proof. We deduce from (4-20) and (4-21) that $f(t)$ and $f_1(t)$ satisfy (2-3) with $R=(1+\mu_1t)^{\frac12}$. So do $f_\pm(t)$ and $f_+(t)f_-(t)$, and thus $g(t)$. This also implies

$$g(t,X)\ge C^{-1}f_+(t,X)f_-(t,X).$$

By definition, we have

$$H(2\mu_1(t-s))f(s,X)=f(t,X),\quad H(2\mu_1(t-s))f_1(s,X)=f_1(t,X),$$

which give the third inequality of (4-14).

By

$$\int_0^Tf_\pm(t,X\pm B_0t)\,dt=\int_0^Tf(t,X\pm2B_0t)\,dt\le Cf_1(t,X)\le C,$$

we get the second inequality of (4-14).

Thanks to

$$\begin{aligned}\int_0^t H(2\mu_1(t-s))f_-(s,X)\,ds &= \int_0^t H(2\mu_1(t-s))U(-2s)f(s,X)\,ds\\ &= \int_0^t f(t,X-2B_0s)\,ds \le Cf_1(t,X),\end{aligned}$$

we get the first inequality of (4-14) with minus sign for $(\hat{f},h)=(f_1,f_-)$. Similarly, the first inequality of (4-14) with plus sign for $(\hat{f},h)=(f_1,f_+)$ is true.

Notice that

$$\begin{aligned}H(2\mu_1(t-s))g_\pm(s,X) &= \int_{\mathbb{R}^d}\frac{H(2\mu_1(t-s))(f_+(s)f_-(s))(Y\pm B_0s)}{1+|X-Y|^{d+1}}\,dY\\ &= \int_{\mathbb{R}^d}\frac{H(2\mu_1(t-s))\big(f(s)U(\pm 2s)f(s)\big)(Y)}{1+|X-Y|^{d+1}}\,dY.\end{aligned}$$

By (4-20), we have

$$f(t,X)\le C\left(1+\frac{|Y-X|}{\sqrt{1+\mu_1 t}}\right)^{1+\delta} f(t,Y),$$

which gives

$$f(s)U(\pm 2s)f(s)(X)\le C\left(1+\frac{|Y-X|}{\sqrt{1+\mu_1 s}}\right)^{2+2\delta} f(s)U(\pm 2s)f(s)(Y).$$

Therefore, for $t/2\le s<t$,

$$\begin{aligned}&H(2\mu_1(t-s))\big(f(s)U(\pm 2s)f(s)\big)(Y)\\ &\qquad= \int_{\mathbb{R}^d} K\big(2\mu_1(t-s),X-Y\big)\big(f(s)U(\pm 2s)f(s)\big)(X)\,dX\\ &\qquad\le C\int_{\mathbb{R}^d} K\big(2\mu_1(t-s),X-Y\big)\left(1+\frac{|Y-X|}{\sqrt{1+\mu_1 s}}\right)^{2+2\delta} f(s)U(\pm 2s)f(s)(Y)\,dX\\ &\qquad\le Cf(s)U(\pm 2s)f(s))(Y)\le Cf(t)U(\pm 2s)f(0))(Y),\end{aligned}$$

and for $0\le s\le t/2$,

$$H(2\mu_1(t-s))\big(f(s)U(\pm 2s)f(s)\big)\le CH(2\mu_1 t)\big(f(s)U(\pm 2s)f(s)\big)\le CH(2\mu_1 t)\big(f(0)U(\pm 2s)f(0)\big).$$

Therefore,

$$\begin{aligned}&\int_0^t H(2\mu_1(t-s))\big(f(s)U(\pm 2s)f(s)\big)\,ds\\ &\qquad\le C\int_0^{\frac{t}{2}} H(2\mu_1 t)\big(f(0)U(\pm 2s)f(0)\big)\,ds + C\int_{\frac{t}{2}}^t \big(f(t)U(\pm 2s)f(0)\big)\,ds\\ &\qquad\le CH(2\mu_1 t)(f(0)f_1(0)) + Cf(t)f_1(0)\le Cf(t).\end{aligned}$$

This shows that

$$\int_0^t H(2\mu_1(t-s))g_\pm(s,X)\,ds \le C\int_{\mathbb{R}^d}\frac{f(t,Y)}{1+|X-Y|^{d+1}}\,dY \le Cf(t,X),$$

which gives the first inequality of (4-14) for $(\hat{f},h)=(f,g)$. $\square$

Proof of Theorem 4.3. The following lemma gives the relation between the Hölder norms of $z_\pm^{(i)}(t)$, $i=1,2$, and $M_\pm(t)$.

Lemma 4.8. *It holds that*

$$|z_\pm^{(2)}(t)|_{0,\alpha;f_1(t),(\mu_1 t)^{1/2}} \le C\mu_1 \min\big((\mu_1 t)^{-\frac12},(\mu_1 t)^{-\frac{1-\alpha}{2}}\big)M_\pm(t),$$
$$|z_\pm^{(2)}(t)|_{0,\alpha;1,(1+\mu_1 t)^{1/2}} \le C\mu_1 \min\big((\mu_1 t)^{-\frac12},(\mu_1 t)^{-\frac{1+\delta}{2}}\big)M_\pm(t),$$
$$|\nabla z_\pm^{(1)}|_{1,\alpha;f_\pm(t),(\mu_1 t)^{1/2}} \le CM_\pm(t).$$

Proof. As $z_\pm^{(2)}=\operatorname{div}\psi_\pm^{(2)}$, we have

$$\begin{aligned}|z_\pm^{(2)}(t)|_{0,\alpha;f_1(t),(\mu_1 t)^{1/2}} &\le C|\nabla\psi_\pm^{(2)}(t)|_{0,\alpha;f_1(t),(\mu_1 t)^{1/2}}\\ &\le C|\psi_\pm^{(2)}(t)|_{1,\alpha;f_1(t),(\mu_1 t)^{1/2}}\min\big((\mu_1 t)^{-\frac12},(\mu_1 t)^{-\frac{1-\alpha}{2}}\big)\\ &\le C\mu_1\min\big((\mu_1 t)^{-\frac12},(\mu_1 t)^{-\frac{1-\alpha}{2}}\big)M_\pm(t),\end{aligned}$$

which along with (4-21) gives

$$\begin{aligned}|z_\pm^{(2)}(t)|_{0,\alpha;1,(1+\mu_1 t)^{1/2}} &\le |z_\pm^{(2)}(t)|_{0,\alpha;f_1(t),(\mu_1 t)^{1/2}}\Big(1+\frac{1}{\mu_1 t}\Big)^{\frac{\alpha}{2}}|f_1(t)|_0\\ &\le C\mu_1\min\big((\mu_1 t)^{-\frac12},(\mu_1 t)^{-\frac{1-\alpha}{2}}\big)M_\pm(t)\Big(1+\frac{1}{\mu_1 t}\Big)^{\frac{\alpha}{2}}(1+\mu_1 t)^{-\frac{\delta}{2}}\\ &\le C\mu_1\min\big((\mu_1 t)^{-\frac12},(\mu_1 t)^{-\frac{1+\delta}{2}}\big)M_\pm(t).\end{aligned}$$

Obviously, we have

$$|\nabla z_\pm^{(1)}|_{0,\alpha;f_\pm(t)} \le |z_\pm^{(1)}|_{1,\alpha;f_\pm(t),(1+\mu_1 t)^{1/2}} \le M_\pm(s).$$

Thanks to $\Delta z_\pm^{(1)}=\operatorname{div} J_\pm^{(1)}$, we have

$$\begin{aligned}|\Delta z_\pm^{(1)}|_{0,\alpha;f_\pm(t),(\mu_1 t)^{1/2}} &\le C|J_\pm^{(1)}|_{1,\alpha;f_\pm(t),(\mu_1 t)^{1/2}}\min\big((\mu_1 t)^{-\frac12},(\mu_1 t)^{-\frac{1-\alpha}{2}}\big)\\ &\le CM_\pm(s)\min\big((\mu_1 t)^{-\frac12},(\mu_1 t)^{-\frac{1-\alpha}{2}}\big).\end{aligned}$$

Notice that by Lemma 4.7,

$$f_\pm(t,X)\le Cf_\pm(t,Y)\quad \text{if } |X-Y|\le(1+\mu_1 t)^{\frac12}.$$

Then we infer from Lemma A.10 that

$$\begin{aligned}|\nabla^2 z_\pm^{(1)}|_{0,\alpha;f_\pm(t),(\mu_1 t)^{1/2}} &\le C\big(|\nabla z_\pm^{(1)}|_{0,\alpha;f_\pm(t)}\min\big((\mu_1 t)^{-\frac12},(\mu_1 t)^{-\frac{1-\alpha}{2}}\big)+|\Delta z_\pm^{(1)}|_{0,\alpha;f_\pm(t),(\mu_1 t)^{1/2}}\big)\\ &\le CM_\pm(s)\min\big((\mu_1 t)^{-\frac12},(\mu_1 t)^{-\frac{1-\alpha}{2}}\big).\end{aligned}$$

This proves the third inequality. $\square$

Proof of Theorem 4.3. For fixed $\nu > 0$ and $\mu > 0$, the local well-posedness of the MHD equations in the weighted Hölder spaces can be proved by using the semigroup method and the estimates of the heat operator in the weighted Hölder spaces (see the subsection on page 1394). Here we omit the details. The local well-posedness of the linear equations (4-2)–(4-7) in the weighted Hölder spaces is also true.

The proof of global well-posedness is based on a continuity argument. Let us first assume

$$M_{\pm}(s) < \varepsilon_1 \tag{4-26}$$

for $\varepsilon_1 > 0$ given by Proposition 4.2. This in particular gives

$$|Z_{\pm}^{(1)}(t) \pm B_0|_{1,\alpha;f_{\mp}(t),(1+\mu_1 t)^{1/2}} < \varepsilon_1.$$

Our next goal is to show that

$$M_+(s) \le C\big(M_+(0) + (M_+(s) + |\mu_2|/\mu_1)M_-(s)\big), \tag{4-27}$$

$$M_-(s) \le C\big(M_-(0) + (M_-(s) + |\mu_2|/\mu_1)M_+(s)\big). \tag{4-28}$$

With the above estimates, we can deduce our result if ε_2 is taken small enough that

$$CM_{\pm}(0) \le C\varepsilon_2 < \tfrac{1}{2}\varepsilon_1, \quad C^2\varepsilon_2 < \tfrac{1}{2}.$$

This condition on ε_2 implies that if $M_{\pm}(s) < \varepsilon_1$ then $M_{\pm}(s) \le 2CM_{\pm}(0) < \varepsilon_1$.

The proof of (4-27) and (4-28) is split into three steps.

Step 1: $C^{1,\alpha}$ estimate for $z_{\pm}^{(1)}$. For the system (4-4), we apply Proposition 4.2 to obtain

$$\begin{aligned}\sup_{0\le t\le s} &|z_+^{(1)}(t)|_{1,\alpha;f_+(t),(1+\mu_1 t)^{1/2}} \\ &\le C\big(|z_+(0)|_{1,\alpha;f(0),1} + \Lambda_1\big(s, I(z_-^{(1)}, z_+^{(1)}), z_-^{(2)}\cdot\nabla z_+^{(1)} + I(z_-^{(2)}, z_+^{(1)}), 0, f_+, g\big)\big).\end{aligned}$$

By (A-5), we have

$$\begin{aligned}\big|I(z_-^{(1)}(t), z_+^{(1)}(t))\big|_{1,\alpha;g(t),(1+\mu_1 t)^{1/2}} &\le C|z_-^{(1)}(t)|_{1,\alpha;f_-(t),(1+\mu_1 t)^{1/2}}|z_+^{(1)}(t)|_{1,\alpha;f_+(t),(1+\mu_1 t)^{1/2}} \\ &\le CM_+(s)M_-(s).\end{aligned}$$

By (A-6) and Lemma 4.8,

$$\begin{aligned}\big|z_-^{(2)}\cdot\nabla z_+^{(1)}(t) &+ I(z_-^{(2)}(t), z_+^{(1)}(t))\big|_{0,\alpha;f_+(t)} \\ &\le C|z_-^{(2)}(t)|_{0,\alpha;1,(1+\mu_1 t)^{1/2}}|z_+^{(1)}(t)|_{1,\alpha;f_+(t),(1+\mu_1 t)^{1/2}}(1+\mu_1 t)^{-\frac{1}{2}} \\ &\le C\mu_1 M_-(s)\min\big((\mu_1 t)^{-\frac{1}{2}}, (\mu_1 t)^{-\frac{1+\delta}{2}}\big)M_+(s)(1+\mu_1 t)^{-\frac{1}{2}} \\ &\le C\mu_1 M_+(s)M_-(s)\min\big((\mu_1 t)^{-\frac{1}{2}}, (\mu_1 t)^{-1-\frac{\delta}{2}}\big),\end{aligned}$$

and obviously,

$$|z_+(0)|_{1,\alpha;f(0),1} \le M_+(0).$$

Therefore, we obtain

$$\sup_{0\le t\le s} |z_+^{(1)}(t)|_{1,\alpha;f_+(t),(1+\mu_1 t)^{1/2}} \le C\big(M_+(0)+M_+(s)M_-(s)\big).$$

Similarly, we have

$$\sup_{0\le t\le s} |z_-^{(1)}(t)|_{1,\alpha;f_-(t),(1+\mu_1 t)^{1/2}} \le C\big(M_-(0)+M_+(s)M_-(s)\big).$$

Step 2: $C^{1,\alpha}$ estimate for $J_\pm^{(1)}$. For the system (4-5), we apply Proposition 4.2 to obtain

$$\sup_{0\le t\le s} |J_+^{(1)}(t)|_{1,\alpha;f_+(t),(\mu_1 t)^{1/2}} \le C\Big(|J_+(0)|_{0,\alpha;\hat f(0)} + \Lambda_0\big(s,\nabla z_-^{(1)}\wedge\nabla z_+^{(1)},0,z_-^{(2)}\cdot\nabla z_+^{(1)},f_+,g\big)\Big).$$

Thanks to the choice of weight functions, we have

$$f_-(t,X)f_+(t,X)\le Cg(t,X).$$

Then by Lemma 4.8 and the analogue of Lemma 2.1, we have

$$\begin{aligned}
\big|\nabla z_-^{(1)}\wedge\nabla z_+^{(1)}(t)\big|_{1,\alpha;g(t),(\mu_1 t)^{1/2}} &\le C|\nabla z_-^{(1)}|_{1,\alpha;f_-(t),(\mu_1 t)^{1/2}}|\nabla z_+^{(1)}(t)|_{1,\alpha;f_+(t),(\mu_1 t)^{1/2}}\\
&\le CM_+(s)M_-(s),\\
\big|z_-^{(2)}\cdot\nabla z_+^{(1)}(t)\big|_{0,\alpha;f_+(t),(\mu_1 t)^{1/2}} &\le C|z_-^{(2)}(t)|_{0,\alpha;1,(\mu_1 t)^{1/2}}|\nabla z_+^{(1)}(t)|_{0,\alpha;f_+(t),(\mu_1 t)^{1/2}}\\
&\le C\mu_1\min\big((\mu_1 t)^{-\frac{1-\alpha}{2}},(\mu_1 t)^{-\frac12}\big)M_-(s)M_+(s),
\end{aligned}$$

and $|J_+(0)|_{0,\alpha;f(0)}\le M_+(0)$. Therefore, we obtain

$$\sup_{0\le t\le s} |J_+^{(1)}(t)|_{1,\alpha;f_+(t),(\mu_1 t)^{1/2}} \le C\big(M_+(0)+M_+(s)M_-(s)\big).$$

Similarly, we have

$$\sup_{0\le t\le s} |J_-^{(1)}(t)|_{1,\alpha;f_-(t),(\mu_1 t)^{1/2}} \le C\big(M_-(0)+M_+(s)M_-(s)\big).$$

Step 3: $C^{1,\alpha}$ estimate for $\psi_\pm^{(2)}$. For the system (4-7), we apply Proposition 4.2 to obtain

$$\sup_{0\le t\le s} |\psi_+^{(2)}(t)|_{1,\alpha;f_1(t),(\mu_1 t)^{1/2}} \le C\Lambda_0\big(s,\mathrm{II}_2(z_-^{(1)},z_+^{(2)})-\mu_2 J_-^{(1)},\mathrm{II}_1(z_-^{(2)},z_+^{(2)}),\mu_2 z_-^{(2)},f_1,f_-\big),$$

where we used the facts that $\psi_\pm^{(2)}(0)=0$ and $f_{1\pm}=f_1$, and the decomposition of $J_\pm$ in (4-6). We get by Proposition A.6 and Lemma 4.8 that

$$\begin{aligned}
&\big|\mathrm{II}_2(z_-^{(1)}(t),z_+^{(2)}(t))-\mu_2 J_-^{(1)}(t)\big|_{1,\alpha;f_-(t),(\mu_1 t)^{1/2}}\\
&\qquad\le C|z_-^{(1)}(t)|_{1,\alpha;f_-(t),(\mu_1 t)^{1/2}}|\nabla\psi_+^{(2)}(t)|_{1,\alpha;f_1(t),(\mu_1 t)^{1/2}}+|\mu_2||J_-^{(1)}(t)|_{1,\alpha;f(t),(\mu_1 t)^{1/2}}\\
&\qquad\le C\mu_1 M_+(s)M_-(s)+|\mu_2|M_-(s),
\end{aligned}$$

and

$$\begin{aligned}\big|\mathrm{II}_1(z_-^{(2)}(t), z_+^{(2)}(t))\big|_{0,\alpha;f_1(t)} &\\ &\le C|z_-^{(2)}(t)|_{0,\alpha;f_1(t),(\mu_1 t)^{1/2}}|z_-^{(2)}(t)|_{0,\alpha;f_1(t),(\mu_1 t)^{1/2}}(1+\mu_1 t)^{-\frac{\delta}{2}}(1+(\mu_1 t)^{-\frac{\alpha}{2}})\\ &\le C\mu_1^2 \min\big((\mu_1 t)^{-1+\frac{\alpha}{2}}, (\mu_1 t)^{-1-\frac{\delta}{2}}\big)M_-(s)M_+(s),\end{aligned}$$

and

$$|\mu_2 z_-^{(2)}(t)|_{0,\alpha;f_1(t),(\gamma t)^{1/2}} \le C\mu_1|\mu_2| \min\big((\mu_1 t)^{-\frac{1}{2}}, (\mu_1 t)^{-\frac{1-\alpha}{2}}\big)M_-(s).$$

This shows that

$$\sup_{0\le t\le s} |\psi_+^{(2)}(t)|_{1,\alpha;f_1(t),(\mu_1 t)^{1/2}} \le C\big(\mu_1 M_+(s)+|\mu_2|\big)M_-(s).$$

Similarly, we have

$$\sup_{0\le t\le s} |\psi_-^{(2)}(t)|_{1,\alpha;f_1(t),(\mu_1 t)^{1/2}} \le C\big(\mu_1 M_-(s)+|\mu_2|\big)M_+(s).$$

Summing up the estimates in Steps 1–3, we conclude (4-27) and (4-28). □

Appendix

Weighted $C^{1,\alpha}$ estimate for the integral operator. Recall that

$$T_1 u \triangleq \int_{\mathbb{R}^d} \nabla N(X-Y)\theta(|X-Y|)u(Y)\,dY, \quad T_{ij} w \triangleq \int_{\mathbb{R}^d} \partial_i\partial_j\big(\nabla N(X-Y)(1-\theta(|X-Y|))\big)w(Y)\,dY,$$

where the cut-off function θ is given by (3-9).

Lemma A.1. *Let $u, w \in C_h^{0,\alpha}(\mathbb{R}^d)$, with the weight h satisfying* (2-1). *Then there exists a constant $C > 0$ depending only on c_0 such that*

$$|T_1 u|_{1,\alpha;h} \le C|u|_{0,\alpha;h}, \quad |T_{ij} w|_{1,\alpha;g} \le C|w|_{0;h},$$

where

$$g(X) = \int_{\mathbb{R}^d} \frac{h(Y)}{1+|X-Y|^{d+1}}\,dy.$$

In particular, we have

$$|T_1 u + T_{ij} w|_{1,\alpha;g} \le C\big(|u|_{0,\alpha;h} + |w|_{0,h}\big).$$

Proof. Thanks to

$$\big|\nabla^k \partial_i\partial_j\big(\nabla N(X-Y)\cdot(1-\theta(|x-y|))\big)\big| \le \frac{C}{1+|x-y|^{d+1}}, \quad k = 0, 1, 2,$$

and $h(X) \le Cg(X)$, we get

$$|\nabla^k T_{ij} w(X)| \le Cg(X)\left|\frac{w}{h}\right|_0,$$

which in particular implies

$$|T_{ij} w|_{1,\alpha;g} \le C\left|\frac{w}{h}\right|_0. \tag{A-1}$$

To deal with $T_1 u$, we decompose it as

$$T_1 u = \sum_{k=0}^{+\infty} B_k(u),$$

where

$$B_k(u) = \int_{\mathbb{R}^d} \varphi_k(X-Y)u(Y)\,dY, \quad \varphi_k(X) = \nabla N(X)\cdot\big(\theta(2^k|X|) - \theta(2^{k+1}|X|)\big).$$

To proceed, we need to use the simple facts

$$\int_{\mathbb{R}^d} |\varphi_k(X)|\,dX \le C2^{-k}, \quad \int_{\mathbb{R}^d} |\nabla\varphi_k(X)||X|^\alpha\,dX \le C2^{-k\alpha}, \quad \int_{\mathbb{R}^d} |\nabla^2\varphi_k(X)||X|^\alpha\,dX \le C2^{k(1-\alpha)},$$
$$\varphi_k(X) = 0 \quad \text{for } |X| > 2,\ k \ge 0.$$

Then we have

$$|B_k(u)(X)| \le \int_{\mathbb{R}^d} |\varphi_k(X-Y)||h(Y)|\,dY\left|\frac{u}{h}\right|_0 \le C2^{-k}h(X)\left|\frac{u}{h}\right|_0. \tag{A-2}$$

Notice that

$$\nabla B_k(u)(X) = \int_{\mathbb{R}^d} \nabla\varphi_k(X-Y)(u(Y)-u(X))\,dY,$$

from which, we deduce

$$|\nabla B_k(u)(X)| \le \int_{\mathbb{R}^d} |\nabla\varphi_k(X-Y)||X-Y|^\alpha(h(X)+h(Y))\,dY|u|_{0,\alpha;h} \le C2^{-k\alpha}h(X)|u|_{0,\alpha;h}. \tag{A-3}$$

Similarly, we have

$$|\nabla^2 B_k(u)(X)| \le C2^{k(1-\alpha)}h(X)|u|_{0,\alpha;h}. \tag{A-4}$$

It follows from (A-2) and (A-3) that

$$\sum_{k=0}^{+\infty} |B_k(u)(X)| \le \sum_{k=0}^{+\infty} C2^{-k}h(X)\left|\frac{u}{h}\right|_0 \le Ch(X)\left|\frac{u}{h}\right|_0,$$
$$\sum_{k=0}^{+\infty} |\nabla B_k(u)(X)| \le \sum_{k=0}^{+\infty} C2^{-k\alpha}h(X)|u|_{0,\alpha} \le Ch(X)|u|_{0,\alpha;h}.$$

It follows from (A-3) and (A-4) that

$$\big|\nabla B_k(u)(X) - \nabla B_k(u)(Y)\big| \le C2^{-k\alpha}(h(X)+h(Y))|u|_{0,\alpha;h}\min(1, 2^k|X-Y|),$$

which gives

$$\begin{aligned}\left|\sum_{k=0}^{+\infty} \nabla\big(B_k(u)(X) - B_k(u)(Y)\big)\right| &\le C(h(X)+h(Y))|u|_{0,\alpha;h}\sum_{k=0}^{+\infty} 2^{-k\alpha}\min(1, 2^k|X-Y|)\\ &\le C\big(h(X)+h(Y)\big)|u|_{0,\alpha;h}|X-Y|^\alpha.\end{aligned}$$

Now we can conclude that

$$|T_1 u|_{1,\alpha;h} \le \Big|\sum_{k=0}^{+\infty} B_k(u)\Big|_{1,\alpha;h} \le C|u|_{0,\alpha;h}. \qquad \square$$

Lemma A.2. *It holds that*

$$\begin{aligned}&\operatorname{div}(T_1 u + T_{ij} w^{ij}) + u\\ &\quad= \int_{\mathbb{R}^d} \nabla N(X-Y)\cdot\nabla\theta(|X-Y|)u(Y)\,dY - \int_{\mathbb{R}^d} \partial_i\partial_j\big(\nabla N(X-Y)\cdot\nabla\theta(|X-Y|)\big)w^{ij}(Y)\,dY.\end{aligned}$$

Proof. With the notations in Lemma A.1, a direct calculation gives

$$\begin{aligned}\operatorname{div}\ T_{ij}(w^{ij}) &= -\int_{\mathbb{R}^d} \partial_i\partial_j\big(\nabla N(X-Y)\cdot\nabla\theta(|X-Y|)\big)w^{ij}(Y)\,dY,\\ \operatorname{div} B_k(u) &= \int_{\mathbb{R}^d} \operatorname{div}\varphi_k(X-Y)u(Y)\,dY,\end{aligned}$$

where

$$\begin{aligned}\operatorname{div}\varphi_k(X) &= \nabla N(X)\cdot\nabla\big(\theta(2^k|X|)-\theta(2^{k+1}|X|)\big) = \varphi_k^*(X)-\varphi_{k+1}^*(X),\\ \varphi_k^*(X) &= \nabla N(X)\cdot\nabla\theta(2^k|X|) = -c_d\frac{2^k\theta'(2^k|X|)}{|X|^{d-1}} \ge 0.\end{aligned}$$

Therefore,

$$\begin{aligned}\operatorname{div}\sum_{k=0}^{N} B_k(u)+u &= \int_{\mathbb{R}^d}\big(\varphi_0^*(X-Y)-\varphi_{N+1}^*(X-Y)\big)u(Y)\,dY+u(X)\\ &= \int_{\mathbb{R}^d}\varphi_0^*(|X-Y|)u(Y)\,dY-\int_{\mathbb{R}^d}\varphi_{N+1}^*(X-Y)(u(Y)-u(X))\,dY \triangleq I_0^*-I_{N+1}^*,\end{aligned}$$

where we used $\int_{\mathbb{R}^d}\varphi_k^*(X)\,dX = 1$. Now,

$$|I_{N+1}^*| \le [u]_\alpha\int_{\mathbb{R}^d}\varphi_{N+1}^*(X-Y)|X-Y|^\alpha\,dY = C[u]_\alpha 2^{-N\alpha}\to 0$$

as $N\to+\infty$. This proves the lemma. $\square$

We also introduce

$$\begin{aligned}T_1(u,R) &\triangleq \int_{\mathbb{R}^d}\nabla N(X-Y)\theta(|X-Y|/R)u(Y)\,dY,\\ T_{ij}(w,R) &\triangleq \int_{\mathbb{R}^d}\partial_i\partial_j\big(\nabla N(X-Y)(1-\theta(|X-Y|/R))\big)w(Y)\,dY,\end{aligned}$$

where $N(X)$ is the Newton potential. Let $R\ge 1$. If $h(X)\le C_0h(Y)$ for $|X-Y|\le 2R$, then we can deduce by following the proof of Lemma A.1 that

$$|T_1(u,R)|_{1,\alpha;g,R}+|T_{ij}(w,R)|_{1,\alpha;g,R} \le C\big(R^2|u|_{0,\alpha;h,R}+|w|_{0;h}\big),$$

where

$$g(X)=\int_{\mathbb{R}^d}\frac{h(Y)}{R^{d+1}+|X-Y|^{d+1}}\,dy.$$

Due to (4-25), we also have

$$R^{-1}|T_1(u,R)|_{1,\alpha;f_\pm(t),R}+|T_{ij}(w,R)|_{0,\alpha;f_\pm(t),R}\le C\big(|u|_{0,\alpha;f_\pm(t),R}+R^{-1}|w|_{0;f_\pm(t)}\big)$$

for $R=\sqrt{1+\mu_1 t}$.

In particular, we have

$$\begin{aligned}&|I(u,w)|_{1,\alpha;g(t),(1+\mu_1t)^{1/2}}\\&\quad\le C\big((1+\mu_1t)|\nabla u|_{0,\alpha;f_+(t),(1+\mu_1t)^{1/2}}|\nabla w|_{0,\alpha;f_-(t),(1+\mu_1t)^{1/2}}+|u|_{0;f_+(t)}|w|_{0;f_-(t)}\big)\\&\quad\le C|u|_{1,\alpha;f_+(t),(1+\mu_1t)^{1/2}}|w|_{1,\alpha;f_-(t),(1+\mu_1t)^{1/2}},\end{aligned}\tag{A-5}$$

where g, $f_\pm$ are defined as in the subsection on page 1385.

For $\operatorname{div} u=\operatorname{div} w=0$, we have

$$I(u,w)\triangleq T_1(\partial_iu^j\partial_jw^i,R)+T_{ij}(u^iw^j,R)=\partial_iT_1(u^j\partial_jw^i,R)+T_{ij}(u^iw^j,R).$$

Therefore, we deduce

$$|I(u,w)|_{0,\alpha;f_\pm(t)}\le C|u|_{0,\alpha;1,(1+\gamma t)^{1/2}}|w|_{1,\alpha;f_\pm(t),(1+\gamma t)^{1/2}}(1+\gamma t)^{-\frac12}.\tag{A-6}$$

Weighted Hölder estimates for the heat operator. Let $H(t)$ be the heat operator given by

$$H(t)f(X):=\frac{1}{(4\pi t)^{d/2}}\int_{\mathbb{R}^d}e^{-\frac{|X-Y|^2}{4t}}f(Y)\,dY=\int_{\mathbb{R}^d}K(t,X-Y)f(Y)\,dY,$$

where $K(t,X)=(4\pi t)^{-\frac d2}e^{-\frac{|X|^2}{4t}}$. Let $\alpha\ge0$ and $k\in\mathbb{N}$. It is easy to verify the properties

$$\begin{gathered}|\nabla^kK(t,X)|\le Ct^{-\frac k2}K(2t,X),\\|\nabla^kK(t,X)||X'|^\alpha\le Ct^{-\frac{k-\alpha}{2}}K(2t,X),\\|\nabla^kK(t,X)-\nabla^kK(t,Y)|\le Ct^{-\frac{k+1}{2}}K(2t,X)|X-Y|,\\|\nabla^kK(t,X)-\nabla^kK(t,Y)||X'|^\alpha\le Ct^{-\frac{k+1-\alpha}{2}}K(2t,X)|X-Y|\end{gathered}\tag{A-7}$$

for any $X',Y\in B(X,\sqrt t)$. Here C is a constant independent of t.

We introduce the seminorm

$$[u]_{1;h}:=\sup_{X,Y\in\mathbb{R}^d}\frac{|u(X)-u(Y)|}{(h(X)+h(Y))|X-Y|}.$$

Then it is easy to check that

$$[u]_{\alpha;h}\le[u]_{1;h}^\alpha|u|_{0;h}^{1-\alpha},\quad|\nabla u|_{0;h}\le2[u]_{1;h}.\tag{A-8}$$

Lemma A.3. *Let $u \in C_h^{0,\alpha}(\mathbb{R}^d)$, with $0 < h < C_0$ and $\alpha \in (0,1)$. Then there exists a constant $C > 0$ depending only on d, α, k such that, for $k \in \mathbb{N}$,*

$$|\nabla^k H(t)u|_{0;H(2t)h} \le Ct^{-\frac{k}{2}}|u|_{0;h}, \quad [\nabla^k H(t)u]_{1;H(2t)h} \le Ct^{-\frac{k+1}{2}}|u|_{0;h},$$
$$[\nabla^k H(t)u]_{\alpha;H(2t)h} \le Ct^{-\frac{k}{2}}[u]_{\alpha;h}, \quad [\nabla^k H(t)u]_{1;H(2t)h} \le Ct^{-\frac{k+1-\alpha}{2}}[u]_{\alpha;h}.$$

Proof. Thanks to (A-7), we have

$$\begin{aligned}|\nabla^k H(t)u(X)| &= \left|\int_{\mathbb{R}^d} \nabla^k K(t, X-Y)u(Y)\,dY\right| \\ &\le \int_{\mathbb{R}^d} |\nabla^k K(t, X-Y)||u(Y)|\,dY \\ &\le Ct^{-\frac{k}{2}} \int_{\mathbb{R}^d} K(2t, X-Y)h(Y)\,dY|u|_{0;h} \\ &\le Ct^{-\frac{k}{2}} H(2t)h(X)|u|_{0;h},\end{aligned}$$

which gives the first inequality.

If $|X-Y| < \sqrt{t}$, then we get by (A-7) that

$$\begin{aligned}\left|\nabla^k H(t)u(X) - \nabla^k H(t)u(Y)\right| &= \left|\int_{\mathbb{R}^d} (\nabla^k K(t, X-X') - \nabla^k K(t, Y-X'))u(X')\,dX'\right| \\ &\le \int_{\mathbb{R}^d} \left|\nabla^k K(t, X-X') - \nabla^k K(t, Y-X')\right||u(X')|\,dX' \\ &\le Ct^{-\frac{k+1}{2}}|X-Y| \int_{\mathbb{R}^d} K(2t, X-X')h(X')\,dX'|u|_{0;h} \\ &\le Ct^{-\frac{k+1}{2}}|X-Y|H(2t)h(X)|u|_{0;h},\end{aligned}$$

and if $|X-Y| \ge \sqrt{t}$, then

$$\begin{aligned}\left|\nabla^k H(t)u(X) - \nabla^k H(t)u(Y)\right| &\le |\nabla^k H(t)u(X)| + |\nabla^k H(t)u(Y)| \\ &\le Ct^{-\frac{k}{2}} H(2t)h(X)|u|_{0;h} + Ct^{-\frac{k}{2}} H(2t)h(Y)|u|_{0;h} \\ &\le Ct^{-\frac{k+1}{2}}|X-Y|\big(H(2t)h(X) + H(2t)h(Y)\big)|u|_{0;h},\end{aligned}$$

which imply the second inequality.

For any $X, Y \in \mathbb{R}^d$, we have

$$\begin{aligned}\left|\nabla^k H(t)u(X) - \nabla^k H(t)u(Y)\right| &= \left|\int_{\mathbb{R}^d} \nabla^k K(t, X')u(X-X')\,dX' - \int_{\mathbb{R}^d} \nabla^k K(t, X')u(Y-X')\,dX'\right| \\ &\le \int_{\mathbb{R}^d} |\nabla^k K(t, X')|\left|u(X-X') - u(Y-X')\right|dX' \\ &\le Ct^{-\frac{k}{2}} \int_{\mathbb{R}^d} K(2t, X')\big(h(X-X') + h(Y-X')\big)\,dX'|X-Y|^\alpha[u]_{\alpha;h} \\ &\le Ct^{-\frac{k}{2}}\big(H(2t)h(X) + H(2t)h(Y)\big)|X-Y|^\alpha[u]_{\alpha;h},\end{aligned}$$

which gives the third inequality.

For any $X, Y \in \mathbb{R}^d$, if $|X-Y| < \sqrt{t}$, we take $Y' \in B(X, \sqrt{t})$ so that

$$h(Y') \int_{B(X,\sqrt{t})} K(2t, X-X')\, dX' \le \int_{B(X,\sqrt{t})} K(2t, X-X') h(X')\, dX' \le H(2t) h(X),$$

which gives $h(Y') \le C H(2t) h(X)$. Then we deduce, for $|X-Y| < \sqrt{t}$,

$$\begin{aligned}
&|\nabla^k H(t)u(X) - \nabla^k H(t)u(Y)| \\
&\qquad = \left| \int_{\mathbb{R}^d} \big(\nabla^k K(t, X-X') - \nabla^k K(t, Y-X')\big)\big(u(X') - u(Y')\big)\, dX' \right| \\
&\qquad \le \int_{\mathbb{R}^d} \big|\nabla^k K(t, X-X') - \nabla^k K(t, Y-X')\big| \big|u(X') - u(Y')\big|\, dX' \\
&\qquad \le \int_{\mathbb{R}^d} \big|\nabla^k K(t, X-X') - \nabla^k K(t, Y-X')\big| |X'-Y'|^\alpha \big(h(X') + h(Y')\big)\, dX' [u]_{\alpha;h} \\
&\qquad \le C t^{-\frac{k+1-\alpha}{2}} |X-Y| \int_{\mathbb{R}^d} K(2t, X-X')\big(h(X') + h(Y')\big)\, dX' [u]_{\alpha;h} \\
&\qquad \le C t^{-\frac{k+1-\alpha}{2}} |X-Y| \big(H(2t)h(X) + h(Y')\big) [u]_{\alpha;h} \\
&\qquad \le C t^{-\frac{k+1-\alpha}{2}} |X-Y| H(2t) h(X) [u]_{\alpha;h}.
\end{aligned}$$

While, if $|X-Y| \ge \sqrt{t}$, then

$$\begin{aligned}
\big|\nabla^k H(t)u(X) - \nabla^k H(t)u(Y)\big| &\le C t^{-\frac{k}{2}} \big(H(2t)h(X) + H(2t)h(Y)\big)|X-Y|^\alpha [u]_{\alpha;h} \\
&\le C t^{-\frac{k+1-\alpha}{2}} \big(H(2t)h(X) + H(2t)h(Y)\big)|X-Y| [u]_{\alpha;h}.
\end{aligned}$$

This proves the fourth inequality. ☐

Lemma A.4. *Let $\gamma > 0$, $k \ge 0$, and $u \in C_h^{0,\alpha}(\mathbb{R}^d)$, with $0 < h < C_0$. Let $R \ge \sqrt{t} > 0$. Then there exists a constant $C > 0$ depending only on d, α such that*

$$|H(t)u|_{1,\alpha;H(2t)h,\sqrt{k+t}} \le C |u|_{1,\alpha;h,\sqrt{k}}, \quad |H(t)u|_{1,\alpha;H(2t)h,R} \le C \frac{\varphi_\alpha(R)}{\varphi_\alpha(\sqrt{t})} |u|_{0,\alpha;h},$$

where $\varphi_\alpha(R) = \max(R, R^{1+\alpha})$.

Proof. By Lemma A.3 and (A-8), we have

$$\begin{gathered}
|H(t)u|_{0;H(2t)h} \le C|u|_{0;h}, \quad [H(t)u]_{\alpha;H(2t)h} \le C[u]_{\alpha;h}, \quad |H(t)u|_{0,\alpha;H(2t)h} \le C|u|_{0,\alpha;h}, \\
|\nabla H(t)u|_{0;H(2t)h} \le \min\big(C t^{-\frac{1}{2}} |u|_{0;h}, C t^{-\frac{1-\alpha}{2}} [u]_{\alpha;h}\big) \le C \min(t^{-\frac{1}{2}}, t^{-\frac{1-\alpha}{2}}) |u|_{0,\alpha;h}, \\
[\nabla H(t)u]_{\alpha;H(2t)h} \le \min\big(C t^{-\frac{1+\alpha}{2}} |u|_{0;h}, C t^{-\frac{1}{2}} [u]_{\alpha;h}\big) \le C \min(t^{-\frac{1+\alpha}{2}}, t^{-\frac{1}{2}}) |u|_{0,\alpha;h}.
\end{gathered}$$

Due to $\nabla H(t)u = H(t)\,\nabla u$, we have

$$|\nabla H(t)u|_{0;H(2t)h} \le C|\nabla u|_{0;h}, \quad [\nabla H(t)u]_{\alpha;H(2t)h} \le C[\nabla u]_{\alpha;h}.$$

Therefore,

$$\begin{aligned}|H(t)u|_{1,\alpha;H(2t)h,\sqrt{k+t}} &= |H(t)u|_{0,\alpha;H(2t)h} + \max((k+t)^{\frac{1-\alpha}{2}}, (k+t)^{\frac{1}{2}})|\nabla H(t)u|_{0;H(2t)h} \\ &\quad + \max((k+t)^{\frac{1}{2}}, (k+t)^{\frac{1+\alpha}{2}})[\nabla H(t)u]_{\alpha;H(2t)h} \\ &\le C|u|_{0,\alpha;h} + \max(k^{\frac{1-\alpha}{2}}, k^{\frac{1}{2}})|\nabla H(t)u|_{0;H(2t)h} \\ &\quad + \max(t^{\frac{1-\alpha}{2}}, t^{\frac{1}{2}})|\nabla H(t)u|_{0;H(2t)h} + \max(k^{\frac{1}{2}}, k^{\frac{1+\alpha}{2}})[\nabla H(t)u]_{\alpha;H(2t)h} \\ &\qquad + \max(t^{\frac{1}{2}}, t^{\frac{1+\alpha}{2}})[\nabla H(t)u]_{\alpha;H(2t)h} \\ &\le C|u|_{0,\alpha;h} + C\max(k^{\frac{1-\alpha}{2}}, k^{\frac{1}{2}})|\nabla u|_{0;h} + C|u|_{0,\alpha;h} \\ &\quad + C\max(k^{\frac{1}{2}}, k^{\frac{1+\alpha}{2}})[\nabla u]_{\alpha;h} + C|u|_{0,\alpha;h} \\ &\le C|u|_{1,\alpha;h,\sqrt{k}},\end{aligned}$$

which gives the first inequality. Also,

$$\begin{aligned}|H(t)u|_{1,\alpha;H(2t)h,R} &= |H(t)u|_{0,\alpha;H(2t)h} + \max(R^{1-\alpha}, R)\big(|\nabla H(t)u|_{0;H(2t)h} + R^{\alpha}[\nabla H(t)u]_{\alpha;H(2t)h}\big) \\ &\le C|u|_{0,\alpha;h} + \max(R, R^{1+\alpha})\big(t^{-\frac{\alpha}{2}}|\nabla H(t)u|_{0;H(2t)h} + [\nabla H(t)u]_{\alpha;H(2t)h}\big) \\ &\le C|u|_{0,\alpha;h} + C\varphi_\alpha(R)\min(t^{-\frac{1+\alpha}{2}}, t^{-\frac{1}{2}})|u|_{0,\alpha;h} \\ &\le C\frac{\varphi_\alpha(R)}{\varphi_\alpha(\sqrt{t})}|u|_{0,\alpha;h},\end{aligned}$$

which gives the second inequality. □

Riesz transform in the weighted Hölder spaces. Throughout this subsection, we take f, f_1, $f_\pm$ to be as in the subsection on page 1385. We need the following property for the weight functions.

Lemma A.5. *For $h = 1$, $f_1(t)$, $f(t)$, and $f_\pm(t)$, we have*

$$R^{-d}\int_{B(X,R)} h(Y) f_1(t,Y)\,dY \le Ch(X)\min(R^{-\delta}, (1+\mu_1 t)^{-\frac{\delta}{2}}). \tag{A-9}$$

Proof. The case of $h = 1$ follows from (4-23). We define

$$\rho_1(X) = |x_2|, \quad \rho_2(X) = |(x_1, x_2)| \quad \text{for } X = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Then by (4-21), for $h = f_1(t)$ if $\rho_1(X) \ge 2R$ or $\rho_1(X) \le 2\sqrt{1+\mu_1 t}$, we have

$$h(Y) \le Ch(X) \quad \text{for } |Y - X| \le R,$$

which gives,

$$R^{-d}\int_{B(X,R)} h(Y) f_1(t,Y)\,dY \le CR^{-d}\int_{B(X,R)} h(X) f_1(t,Y)\,dY \le Ch(X)\min(R^{-\delta}, (1+\mu_1 t)^{-\frac{\delta}{2}}).$$

Using (4-20), the above inequality holds for $h = f(t)$ if $\rho_2(X) \ge 2R$ or $\rho_2(X) \le 2\sqrt{1+\mu_1 t}$.

For the case $h = f_1(t)$, if $2\sqrt{1+\mu_1 t} \le \rho_1(X) \le 2R$, then by (4-21),

$$h(X) \ge C^{-1}\varphi_1(X) \ge C^{-1}R^{-\delta}, \quad h(Y) f_1(t,Y) \le C\varphi_1(Y)^2 = C\rho_1(Y)^{-2\delta},$$

which imply

$$R^{-d}\int_{B(X,R)} h(Y)f_1(t,Y)\,dY \le CR^{-d}\int_{B(X,R)} \rho_1(Y)^{-2\delta}\,dY \le CR^{-2\delta} \le Ch(X)R^{-\delta}.$$

For the case $h=f(t)$, if $2\sqrt{1+\mu_1 t}\le \rho_2(X)\le 2R$, then by (4-20),

$$h(X)\ge C^{-1}\varphi_2(X)\ge C^{-1}R^{-1-\delta},\quad h(Y)f_1(t,Y)\le C\varphi_1(Y)\varphi_2(Y)=C|y_1|^{-\frac12-\delta}|y_2|^{-\frac12-\delta},$$

which imply

$$R^{-d}\int_{B(X,R)} h(Y)f_1(t,Y)\,dY \le CR^{-1-2\delta}\le Ch(X)R^{-\delta}.$$

Thus (A-9) is true for $h=f_1(t), f(t)$. The case $h=f_\pm(t)$ follows from the case $h=f(t)$ by translation. □

Proposition A.6. *It holds that*

$$\big|[u,R_iR_j]\,\partial_k w\big|_{1,\alpha;f_\pm(t),(\mu_1 t)^{1/2}} \le C|u|_{1,\alpha;f_\pm(t),(1+\mu_1 t)^{1/2}}|w|_{1,\alpha;f_1(t),(\mu_1 t)^{1/2}},$$

$$|R_iR_j(uw)|_{0,\alpha;f_1(t)} \le C(1+\mu_1 t)^{-\frac{\delta}{2}}(1+(\mu_1 t)^{-\frac{\alpha}{2}})|u|_{0,\alpha;f_1(t),(\mu_1 t)^{1/2}}|w|_{0,\alpha;f_1(t),(\mu_1 t)^{1/2}}.$$

The proof of the proposition is very complicated. Let us begin with some reductions. For fixed i, j,

$$R_iR_jw(X)+\frac{\delta_{ij}}{d}w(X) = -\text{p.v.}\int_{\mathbb{R}^d}\partial_i\partial_j N(X-Y)w(Y)\,dY \triangleq \sum_{n=-\infty}^{\infty} R^n_{ij}(w),$$

where

$$R^n_{ij}(u) = -\int_{\mathbb{R}^d}\varphi_n(X-Y)u(Y)\,dY,$$

with $\varphi_n(X)=\partial_i\partial_j N(X)\big(\theta(2^n|X|)-\theta(2^{n+1}|X|)\big)$. Therefore,

$$[u,R_iR_j]\,\partial_k w = \sum_{n=-\infty}^{\infty}[u,R^n_{ij}]\,\partial_k w. \tag{A-10}$$

Lemma A.7. *For $h=1, f_1(t), f(t)$ and $f_\pm(t)$, it holds that*

$$|R_iR_j(u)|_{0,\alpha;h,(1+\mu_1 t)^{1/2}} \le C(1+\mu_1 t)^{-\frac{\delta}{2}}|u|_{0,\alpha;hf_1(t),(1+\mu_1 t)^{1/2}}.$$

Proof. Notice that

$$\int_{R^d}\varphi_n(X)\,dX=0,\quad \operatorname{supp}\varphi_n\subset B(0,2^{1-n})\setminus B(0,2^{-1-n}),\ \ |\nabla^l\varphi_n|\le C2^{n(d+l)},\ l=0,1,2.$$

We deduce from Lemma A.5 that

$$\begin{aligned}|R^n_{ij}(u)(X)| &\le \int_{\mathbb{R}^d}|\varphi_n(X-Y)|h(Y)f_1(t,Y)\,dY|u|_{0;hf_1(t)}\\ &\le C2^{nd}\int_{B(X,2^{1-n})}h(Y)f_1(t,Y)\,dY|u|_{0;hf_1(t)} \le C2^{n\delta}h(X)|u|_{0;hf_1(t)}.\end{aligned}$$

For $X\in\mathbb{R}^d$,

$$R^n_{ij}(u)(X) = -\int_{\mathbb{R}^d}\varphi_n(X-Y)(u(Y)-u(X))\,dY,$$

which along with (4-24) gives

$$\begin{aligned}|R^n_{ij}(u)(X)| &\le \int_{\mathbb{R}^d} |\varphi_n(X-Y)|(h(X)+h(Y))|X-Y|^\alpha \, dY\,[u]_{\alpha;h} \\ &\le C2^{n(d-\alpha)}\int_{B(X,2^{1-n})}(h(X)+h(Y))\,dY\,[u]_{\alpha;h} \le C2^{-n\alpha}h(X)[u]_{\alpha;h}.\end{aligned}$$

By (4-21), we have

$$[u]_{\alpha;h} \le C(1+\mu_1 t)^{-\frac{\delta}{2}}[u]_{\alpha;hf_1(t)} \le C(1+\mu_1 t)^{-\frac{\alpha+\delta}{2}}|u|_{0,\alpha;hf_1(t),(1+\mu_1 t)^{1/2}}.$$

Thus, we can conclude

$$\begin{aligned}|R_iR_j(u)(X)| &\le \sum_{n=-\infty}^{\infty}|R^n_{ij}(u)(X)| \\ &\le \sum_{n=-\infty}^{\infty} C\min\bigl(2^{n\delta},2^{-n\alpha}(1+\mu_1 t)^{-\frac{\alpha+\delta}{2}}\bigr)h(X)|u|_{0,\alpha;hf_1(t),(1+\mu_1 t)^{1/2}} \\ &\le C(1+\mu_1 t)^{-\frac{\delta}{2}}h(X)|u|_{0,\alpha;hf_1(t),(1+\mu_1 t)^{1/2}}.\end{aligned}$$

For any $X, X' \in \mathbb{R}^d$, with $|X-X'|\le 2^{-n}$,

$$\begin{aligned}\bigl|R^n_{ij}(u)(X)-R^n_{ij}(u)(X')\bigr| &\le \int_{\mathbb{R}^d}\bigl|\varphi_n(X-Y)-\varphi_n(X'-Y)\bigr|(h(X)+h(Y))|X-Y|^\alpha\,dY\,[u]_{\alpha;h} \\ &\le C2^{n(d+1-\alpha)}|X-X'|\int_{B(X,2^{1-n})}(h(X)+h(Y))\,dY\,[u]_{\alpha;h} \\ &\le C2^{n(1-\alpha)}|X-X'|h(X)[u]_{\alpha;h},\end{aligned}$$

which gives, for any $X, X' \in \mathbb{R}^d$,

$$\bigl|R^n_{ij}(u)(X)-R^n_{ij}(u)(X')\bigr| \le C2^{-n\alpha}\min(1,2^n|X-X'|)(h(X)+h(X'))[u]_{\alpha;h}.$$

Then we have

$$\begin{aligned}\bigl|R_iR_j(u)(X)-R_iR_j(u)(X')\bigr| &\le \sum_{n=-\infty}^{\infty}\bigl|R^n_{ij}(u)(X)-R^n_{ij}(u)(X')\bigr| \\ &\le \sum_{n=-\infty}^{\infty}C2^{-n\alpha}\min(1,2^n|X-X'|)(h(X)+h(X'))[u]_{\alpha;h} \\ &\le C|X-X'|^\alpha(h(X)+h(X'))[u]_{\alpha;h},\end{aligned}$$

which implies $[R_iR_ju]_{\alpha;h}\le C[u]_{\alpha;h}$. Thus,

$$\begin{aligned}|R_iR_j(u)|_{0,\alpha;h,(1+\mu_1 t)^{1/2}} &= |R_iR_j(u)|_{0;h}+(1+\mu_1 t)^{\frac{\alpha}{2}}[R_iR_ju]_{\alpha;h} \\ &\le C(1+\mu_1 t)^{-\frac{\delta}{2}}|u|_{0,\alpha;hf_1(t),(1+\mu_1 t)^{1/2}}+(1+\mu_1 t)^{\frac{\alpha}{2}}[u]_{\alpha;h} \\ &\le C(1+\mu_1 t)^{-\frac{\delta}{2}}|u|_{0,\alpha;hf_1(t),(1+\mu_1 t)^{1/2}},\end{aligned}$$

which gives our result. □

Lemma A.8. *For $l = 0, 1$, it holds that*

$$\left|\nabla^l [u, R^n_{ij}]\, \partial_k w(X)\right| \le C 2^{n(l-\alpha)} |\nabla u|_{0;B(X,2^{1-n})} [w]_\alpha.$$

Proof. Thanks to

$$\begin{aligned}[u, R^n_{ij}]\, \partial_k w(X) &= \int_{\mathbb{R}^d} \varphi_n(X-Y)(u(Y)-u(X))\, \partial_k w(Y)\, dY \\ &= \int_{\mathbb{R}^d} \partial_k \varphi_n(X-Y)(u(Y)-u(X)) w(Y) dY - \int_{\mathbb{R}^d} \varphi_n(X-Y)\, \partial_k u(Y) w(Y)\, dY, \quad \text{(A-11)}\end{aligned}$$

we deduce that

$$\begin{aligned}\left|[u, R^n_{ij}]\, \partial_k w(X)\right| &\le \int_{\mathbb{R}^d} |\partial_k \varphi_n(X-Y)|\, |X-Y|\, dY\, |\nabla u|_{0;B(X,2^{1-n})} |w|_{0;B(X,2^{1-n})} \\ &\qquad + \int_{\mathbb{R}^d} |\varphi_n(X-Y)|\, dY\, |\nabla u|_{0;B(X,2^{1-n})} |w|_{0;B(X,2^{1-n})} \\ &\le C |\nabla u|_{0;B(X,2^{1-n})} |w|_{0;B(X,2^{1-n})},\end{aligned}$$

Thanks to

$$\begin{aligned}\nabla[u, R^n_{ij}]\, \partial_k w(X) &= \int_{\mathbb{R}^d} \nabla \partial_k \varphi_n(X-Y)(u(Y)-u(X)) w(Y)\, dY \\ &\quad - \nabla u(X) \int_{\mathbb{R}^d} \partial_k \varphi_n(X-Y) w(Y) dY - \int_{\mathbb{R}^d} \nabla \varphi_n(X-Y)\, \partial_k u(Y) w(Y)\, dY, \quad \text{(A-12)}\end{aligned}$$

we can similarly deduce that

$$\left|\nabla[u, R^n_{ij}]\, \partial_k w(X)\right| \le C 2^n |\nabla u|_{0;B(X,2^{1-n})} |w|_{0;B(X,2^{1-n})}.$$

As $[u, R^n_{ij}]\, \partial_k w = [u, R^n_{ij}]\, \partial_k (w - w(X))$, we have, for $l = 0, 1$,

$$\left|\nabla^l [u, R^n_{ij}]\, \partial_k w(X)\right| \le C 2^{nl} |\nabla u|_{0;B(X,2^{1-n})} |w - w(X)|_{0;B(X,2^{1-n})} \le C 2^{n(l-\alpha)} |\nabla u|_{0;B(X,2^{1-n})} [w]_\alpha.$$

□

Lemma A.9. *If $|u|_{1,\alpha;h,(1+\mu_1 t)^{1/2}} = |w|_{1,\alpha;f_1(t),(\mu_1 t)^{1/2}} = 1$ for $h = 1$, $f_1(t)$, $f(t)$ and $f_\pm(t)$, then we have*

$$\begin{aligned}\left|[u, R^n_{ij}]\, \partial_k w(X)\right| &\le C h(X) \min\left(2^{n\delta}(1+\mu_1 t)^{-\frac{1}{2}}, 2^{-n\alpha}(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}\right), \\ \left|\partial_l [u, R^n_{ij}]\, \partial_k w(X)\right| &\le C h(X)(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}} \min\left(2^{n(1-\alpha)}, 2^{-n\alpha}(\mu_1 t)^{-\frac{1}{2}}\right).\end{aligned}$$

Proof. As $|u|_{1,\alpha;h,(1+\mu_1 t)^{1/2}} = |w|_{1,\alpha;f_1(t),(\mu_1 t)^{1/2}} = 1$, we have

$$|u(X)| \le h(X), \quad |\nabla u(X)| \le h(X)(1+\mu_1 t)^{-\frac{1}{2}}, \quad |w(X)| \le f_1(t, X).$$

Using $f_1(t, X) \le C(1+\mu_1 t)^{-\frac{\delta}{2}}$, we also have

$$\begin{aligned}&|w|_0 \le C(1+\mu_1 t)^{-\frac{\delta}{2}}, \quad [w]_\alpha \le C(1+\mu_1 t)^{-\frac{\delta}{2}}, \quad |\nabla w|_0 \le C(1+\mu_1 t)^{-\frac{\delta}{2}}(\mu_1 t)^{-\frac{1}{2}}, \\ &[\nabla w]_\alpha \le C(1+\mu_1 t)^{-\frac{\delta}{2}} \min\left((\mu_1 t)^{-\frac{1}{2}}, (\mu_1 t)^{-\frac{1+\alpha}{2}}\right) \le C(1+\mu_1 t)^{-\frac{\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}},\end{aligned}$$

and

$$[w]_\alpha \le C|w|_0^{1-\alpha}|\nabla w|_0^\alpha \le C(1+\mu_1 t)^{-\frac{\delta}{2}}(\mu_1 t)^{-\frac{\alpha}{2}}.$$

Therefore

$$[w]_\alpha \le C(1+\mu_1 t)^{-\frac{\delta}{2}}\min(1,(\mu_1 t)^{-\frac{\alpha}{2}}) \le C(1+\mu_1 t)^{-\frac{\delta+\alpha}{2}}.$$

Then we deduce from (A-11) and Lemma A.5 that, for $2^{-n} \ge \sqrt{1+\mu_1 t}$,

$$\begin{aligned}
&\big|[u,R_{ij}^n]\,\partial_k w(X)\big| \\
&\le \int_{\mathbb{R}^d} |\partial_k\varphi_n(X-Y)|(h(X)+h(Y))\,f_1(t,Y)\,dY + (1+\mu_1 t)^{-\frac{1}{2}}\int_{\mathbb{R}^d}|\varphi_n(X-Y)|h(Y)\,f_1(t,Y)\,dY \\
&\le C2^{n(d+1)}\int_{B(X,2^{1-n})}(h(X)+h(Y))\,f_1(t,Y)\,dY + C(1+\mu_1 t)^{-\frac{1}{2}}2^{nd}\int_{B(X,2^{1-n})}h(Y)\,f_1(t,Y)\,dY \\
&\le C2^{n(1+\delta)}h(X) + C(1+\mu_1 t)^{-\frac{1}{2}}2^{n\delta}h(X) \\
&\le C(1+\mu_1 t)^{-\frac{1}{2}}2^{n\delta}h(X).
\end{aligned}$$

For $2^{-n} \le \sqrt{1+\mu_1 t}$, we have

$$|\nabla u|_{0;B(X,2^{1-n})} \le |\nabla u|_{0;h,B(X,2^{1-n})}|h|_{0;B(X,2^{1-n})} \le Ch(X)(1+\mu_1 t)^{-\frac{1}{2}},$$

where we used the fact that h satisfies (2-3) with $R=\sqrt{1+\mu_1 t}$. Similarly, we have

$$[\nabla u]_{\alpha;B(X,2^{1-n})} \le [\nabla u]_{\alpha;h,B(X,2^{1-n})}|h|_{0;B(X,2^{1-n})} \le Ch(X)(1+\mu_1 t)^{-\frac{1+\alpha}{2}}.$$

Then we get by Lemma A.8 that

$$\big|[u,R_{ij}^n]\,\partial_k w(X)\big| \le C2^{n\alpha}h(X)|\nabla u|_{0;B(X,2^{1-n})}[w]_\alpha \le C2^{n\alpha}h(X)(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}},$$

which gives the first inequality of the lemma.

Similarly, by (A-12) and Lemmas A.5 and A.8, we can deduce

$$\begin{aligned}
|\partial_l[u,R_{ij}^n]\,\partial_k w(X)| &\le Ch(X)2^n\min\big(2^{n\delta}(1+\mu_1 t)^{-\frac{1}{2}},2^{-n\alpha}(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}\big) \\
&\le Ch(X)2^{n(1-\alpha)}(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\partial_l[u,R_{ij}^n]\,\partial_k w(X) &= \int_{\mathbb{R}^d}\partial_l\varphi_n(X-Y)(u(Y)-u(X))\,\partial_k w(Y)\,dY - \partial_l u(X)\int_{\mathbb{R}^d}\varphi_n(X-Y)\,\partial_k w(Y)\,dY \\
&\triangleq [u,\partial_l R_{ij}^n]\,\partial_k w(X) + \partial_l u(X)R_{ij}^n\,\partial_k w(X).
\end{aligned}$$

From the proof of Lemma A.7, we can see that

$$|R_{ij}^n\,\partial_k w(X)| \le C2^{-n\alpha}[\partial_k w]_\alpha \le C2^{-n\alpha}(1+\mu_1 t)^{-\frac{\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}}.$$

By (4-24), we deduce that, for $2^{-n} \geq \sqrt{1+\mu_1 t}$,

$$\begin{aligned}\big|[u, \partial_l R_{ij}^n]\, \partial_k w(X)\big| &\leq \int_{\mathbb{R}^d} |\partial_l \varphi_n(X-Y)| \big(|u(Y)|+|u(X)|\big)\, |\partial_k w(Y)|\, dY \\ &\leq C 2^{n(d+1)}(1+\mu_1 t)^{-\frac{\delta}{2}}(\mu_1 t)^{-\frac{1}{2}} \int_{B(X,2^{1-n})} (h(Y)+h(X))\, dY \\ &\leq C 2^{n}(1+\mu_1 t)^{-\frac{\delta}{2}}(\mu_1 t)^{-\frac{1}{2}} h(X) \\ &\leq C 2^{-n\alpha}(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}} h(X).\end{aligned}$$

For $2^{-n} \leq \sqrt{1+\mu_1 t}$, using the formula

$$\begin{aligned}[u, \partial_l R_{ij}^n]\, \partial_k w(X) = \int_{\mathbb{R}^d} \partial_l \varphi_n(X-Y)(u(Y)-u(X))(\partial_k w(Y)-\partial_k w(X))\, dY \\ +\partial_k w(X) \int_{\mathbb{R}^d} \varphi_n(X-Y)(\partial_l u(Y)-\partial_l u(X))\, dY,\end{aligned}$$

we deduce that

$$\begin{aligned}\big|[u, \partial_l R_{ij}^n]\, \partial_k w(X)\big| &\leq \int_{\mathbb{R}^d} |\partial_l \varphi_n(X-Y)|\, |X-Y|^{1+\alpha}\, dY\, |\nabla u|_{0;B(X,2^{1-n})} [\partial_k w]_\alpha \\ &\qquad + |\partial_k w|_0 \int_{\mathbb{R}^d} \varphi_n(X-Y) |X-Y|^{\alpha}\, dY\, [\nabla u]_{\alpha;B(X,2^{1-n})} \\ &\leq C 2^{-n\alpha} h(X)(1+\mu_1 t)^{-\frac{1}{2}}(1+\mu_1 t)^{-\frac{\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}} \\ &\qquad + C(1+\mu_1 t)^{-\frac{\delta}{2}}(\mu_1 t)^{-\frac{1}{2}} 2^{-n\alpha} h(X)(1+\mu_1 t)^{-\frac{1+\alpha}{2}} \\ &\leq C 2^{-n\alpha} h(X)(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}}.\end{aligned}$$

This shows that

$$\begin{aligned}&\big|\partial_l [u, R_{ij}^n]\, \partial_k w(X)\big| \\ &\quad\leq \big|[u, \partial_l R_{ij}^n]\, \partial_k w(X)\big| + \big|\partial_l u(X) R_{ij}^n\, \partial_k w(X)\big| \\ &\quad\leq C 2^{-n\alpha} h(X)(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}} + C h(X)(1+\mu_1 t)^{-\frac{1}{2}} 2^{-n\alpha}(1+\mu_1 t)^{-\frac{\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}} \\ &\quad\leq C 2^{-n\alpha} h(X)(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}},\end{aligned}$$

which gives the second inequality of the lemma. □

Using the formula

$$\begin{aligned}&\partial_m [u, \partial_l R_{ij}^n]\, \partial_k w(X) \\ &\quad= \int_{\mathbb{R}^d} \partial_m \partial_l \varphi_n(X-Y)(u(Y)-u(X))\, \partial_k w(Y)\, dY - \partial_m u(X) \int_{\mathbb{R}^d} \partial_l \varphi_n(X-Y)\, \partial_k w(Y)\, dY,\end{aligned}$$

we can also deduce that

$$\big|\partial_m [u, \partial_l R_{ij}^n]\, \partial_k w(X)\big| \leq C 2^{n(1-\alpha)} h(X)(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}}. \tag{A-13}$$

Now we are in position to prove Proposition A.6.

Proof. We get by Lemma A.7 with $h = f_1(t)$ that

$$\begin{aligned}|R_iR_j(uw)|_{0,\alpha;f_1(t)} &\le C(1+\mu_1 t)^{-\frac{\delta}{2}}|uw|_{0,\alpha;f_1(t)^2,(1+\mu_1 t)^{1/2}}\\ &\le C(1+\mu_1 t)^{-\frac{\delta}{2}}(1+(\mu_1 t)^{-\frac{\alpha}{2}})|uw|_{0,\alpha;f_1(t)^2,(\mu_1 t)^{1/2}}\\ &\le C(1+\mu_1 t)^{-\frac{\delta}{2}}(1+(\mu_1 t)^{-\frac{\alpha}{2}})|u|_{0,\alpha;f_1(t),(\mu_1 t)^{1/2}}|w|_{0,\alpha;f_1(t),(\mu_1 t)^{1/2}},\end{aligned}$$

which gives the second inequality of the proposition.

For the first inequality, without loss of generality, we can assume

$$|u|_{1,\alpha;h,(1+\mu_1 t)^{1/2}} = |w|_{1,\alpha;f_1(t),(\mu_1 t)^{1/2}} = 1,$$

where $h = f_\pm(t)$.

First of all, by Lemma A.9, we have

$$\begin{aligned}\big|[u,R_iR_j]\,\partial_k w(X)\big| &\le \sum_{n=-\infty}^{\infty}\big|[u,R_{ij}^n]\,\partial_k w(X)\big|\\ &\le C\sum_{n=-\infty}^{\infty} h(X)\min\big(2^{n\delta}(1+\mu_1 t)^{-\frac12}, 2^{-n\alpha}(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}\big) \le Ch(X)(1+\mu_1 t)^{-\frac{1+\delta}{2}},\end{aligned}$$

and

$$\begin{aligned}\big|\partial_l[u,R_iR_j]\,\partial_k w(X)\big| &\le \sum_{n=-\infty}^{\infty}\big|\partial_l[u,R_{ij}^n]\,\partial_k w(X)\big|\\ &\le C\sum_{n=-\infty}^{\infty} h(X)(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}\min\big(2^{n(1-\alpha)}, 2^{-n\alpha}(\mu_1 t)^{-\frac12}\big)\\ &\le Ch(X)(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1-\alpha}{2}}.\end{aligned}$$

Now we consider $X, Y \in \mathbb{R}^d$, with $|X-Y| \le \sqrt{1+\mu_1 t}$. It follows from Lemma A.9 that

$$\big|[u,R_{ij}^n]\,\partial_k w(X) - [u,R_{ij}^n]\,\partial_k w(Y)\big| \le Ch(X)2^{-n\alpha}(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}\min(1, 2^n|X-Y|),$$

where we used the fact that h satisfies (2-3) with $R = \sqrt{1+\mu_1 t}$. Therefore,

$$\begin{aligned}\big|[u,R_iR_j]\,\partial_k w(X) - [u,R_iR_j]\,\partial_k w(Y)\big| &\le \sum_{n=-\infty}^{\infty}\big|[u,R_{ij}^n]\,\partial_k w(X) - [u,R_{ij}^n]\,\partial_k w(Y)\big|\\ &\le C\sum_{n=-\infty}^{\infty} h(X)2^{-n\alpha}(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}\min(1, 2^n|X-Y|)\\ &\le Ch(X)(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}|X-Y|^\alpha.\end{aligned}$$

We write

$$\partial_l[u,R_iR_j]\,\partial_k w = [u,\partial_l R_iR_j]\,\partial_k w + \partial_l u\, R_iR_j\,\partial_k w,$$

where

$$[u,\partial_l R_iR_j]\,\partial_k w = \sum_{n=-\infty}^{\infty}[u,\partial_l R_{ij}^n]\,\partial_k w.$$

We get by Lemma A.9 and (A-13) that

$$\big|[u,\partial_l R_{ij}^n]\,\partial_k w(X) - [u,\partial_l R_{ij}^n]\,\partial_k w(Y)\big| \le Ch(X)2^{-n\alpha}(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac12}\min(1, 2^n|X-Y|),$$

which gives

$$
\begin{aligned}
&\big|[u,\partial_l R_i R_j]\,\partial_k w(X)-[u,\partial_l R_i R_j]\,\partial_k w(Y)\big|\\
&\qquad\le \sum_{n=-\infty}^{\infty}\big|[u,\partial_l R_{ij}^n]\,\partial_k w(X)-[u,\partial_l R_{ij}^n]\,\partial_k w(Y)\big|\\
&\qquad\le C\sum_{n=-\infty}^{\infty} h(X)2^{-n\alpha}(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}}\min(1,2^n|X-Y|)\\
&\qquad\le Ch(X)(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}}|X-Y|^{\alpha}.
\end{aligned}
$$

By

$$
\begin{aligned}
|\partial_k w|_{0,\alpha;f_1(t),(1+\mu_1 t)^{1/2}} &\le (1+(\mu_1 t)^{-\frac{\alpha}{2}})|\partial_k w|_{0,\alpha;f_1(t),(\mu_1 t)^{1/2}}\\
&\le (1+(\mu_1 t)^{-\frac{\alpha}{2}})\min\big((\mu_1 t)^{-\frac{1-\alpha}{2}},(\mu_1 t)^{-\frac{1}{2}}\big)\le 2(\mu_1 t)^{-\frac{1}{2}},
\end{aligned}
$$

we infer from Lemma A.7 that

$$
\begin{aligned}
|\partial_l u\ R_i R_j\ \partial_k w|_{0,\alpha;h,(1+\mu_1 t)^{1/2}} &\le C|\partial_l u|_{0,\alpha;h,(1+\mu_1 t)^{1/2}}|R_i R_j\ \partial_k w|_{0,\alpha;1,(1+\mu_1 t)^{1/2}}\\
&\le C(1+\mu_1 t)^{-\frac{1}{2}}(1+\mu_1 t)^{-\frac{\delta}{2}}|\partial_k w|_{0,\alpha;h,(1+\mu_1 t)^{1/2}}\\
&\le C(1+\mu_1 t)^{-\frac{1+\delta}{2}}(\mu_1 t)^{-\frac{1}{2}},
\end{aligned}
$$

and

$$[\partial_l u\ R_i R_j\ \partial_k w]_{\alpha;h}\le C(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}}.$$

This shows that

$$\big|\partial_l[u,R_i R_j]\,\partial_k w(X)-\partial_l[u,R_i R_j]\,\partial_k w(Y)\big|\le Ch(X)(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}}|X-Y|^{\alpha}.$$

For the case $X,Y\in\mathbb{R}^d$, with $|X-Y|\ge\sqrt{1+\mu_1 t}$, we have

$$
\begin{aligned}
\big|[u,R_i R_j]\,\partial_k w(X)-[u,R_i R_j]\,\partial_k w(Y)\big| &\le C(h(X)+h(Y))(1+\mu_1 t)^{-\frac{1+\delta}{2}}\\
&\le C(h(X)+h(Y))(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}|X-Y|^{\alpha},
\end{aligned}
$$

and

$$
\begin{aligned}
\big|\partial_l[u,R_i R_j]\,\partial_k w(X)-\partial_l[u,R_i R_j]\,\partial_k w(Y)\big| &\le C(h(X)+h(Y))(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1-\alpha}{2}}\\
&\le C(h(X)+h(Y))(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}}|X-Y|^{\alpha}.
\end{aligned}
$$

In summary, we conclude

$$
\begin{aligned}
&\big|[u,R_i R_j]\,\partial_k w\big|_{1,\alpha;h,(\mu_1 t)^{1/2}}\\
&=\big|[u,R_i R_j]\,\partial_k w\big|_{0;h}+\big[[u,R_i R_j]\,\partial_k w\big]_{\alpha;h}\\
&\qquad+\max((\mu_1 t)^{\frac{1-\alpha}{2}},(\mu_1 t)^{\frac{1}{2}})\Big(\big|\nabla[u,R_i R_j]\,\partial_k w\big|_{0;h}+(\mu_1 t)^{\frac{\alpha}{2}}\big[\nabla[u,R_i R_j]\,\partial_k w\big]_{\alpha;h}\Big)\\
&\le C(1+\mu_1 t)^{-\frac{1+\delta}{2}}+C(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}\\
&\quad+C\max\big((\mu_1 t)^{\frac{1-\alpha}{2}},(\mu_1 t)^{\frac{1}{2}}\big)\big((1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1-\alpha}{2}}+(\mu_1 t)^{\frac{\alpha}{2}}(1+\mu_1 t)^{-\frac{1+\delta+\alpha}{2}}(\mu_1 t)^{-\frac{1}{2}}\big)\le C,
\end{aligned}
$$

which gives the first inequality of the proposition. $\square$

Weighted Schauder estimate. Let $h(X)$ be a positive bounded weight satisfying

$$h(X) \le C_0 h(Y) \quad \text{for } |X-Y| \le 2R,\ R>0.$$

Lemma A.10. *Let $u \in C_h^{2,\alpha}(\mathbb{R}^d)$. Then we have*

$$|\nabla^2 u|_{0,\alpha;h,R} \le C\big(|\nabla u|_{0,\alpha;h} \min(R^{-1+\alpha}, R^{-1}) + |\Delta u|_{0,\alpha;h,R}\big).$$

Here C is a constant depending only on C_0.

Proof. Fix $X \in \mathbb{R}^d$ and consider the function $w(Y) = u(Y) - u(X) - (Y-X)\cdot\nabla u(X)$. So,

$$\nabla^2 w = \nabla^2 u, \quad \Delta w = \Delta u, \quad |\Delta u|_{0,\alpha;B(X,2R),R} \le 2C_0 h(X)|\Delta u|_{0,\alpha;h,R},$$

where

$$|u|_{0,\alpha;B(X,2R),R} \triangleq |u|_{0;B(X,2R)} + R^\alpha [u]_{\alpha;B(X,2R)}.$$

As $\nabla w(Y) = \nabla u(Y) - \nabla u(X)$, we have for $|X-Y| \le 2R$,

$$|\nabla w(Y)| = |\nabla u(Y) - \nabla u(X)| \le (h(X)+h(Y))|X-Y|^\alpha |\nabla u|_{0,\alpha;h} \le 4C_0 h(X) R^\alpha |\nabla u|_{0,\alpha;h},$$

$$|\nabla w(Y)| \le |\nabla u(Y)| + |\nabla u(X)| \le (h(X)+h(Y))|\nabla u|_{0,\alpha;h} \le 2C_0 h(X)|\nabla u|_{0,\alpha;h}.$$

This shows that

$$|\nabla w|_{0;B(X,2R)} \le 4C_0 h(X) \min(R^\alpha, 1)|\nabla u|_{0,\alpha;h},$$

from which and $w(X)=0$, we infer

$$|w|_{0;B(X,2R)} \le 2R|\nabla w|_{0;B(X,2R)} \le 8C_0 h(X)\min(R^{1+\alpha}, R)|\nabla u|_{0,\alpha;h}.$$

Then by the (scaled) Schauder estimate, we obtain

$$\begin{aligned} |\nabla^2 w|_{0,\alpha;B(X,R),R} &\le C\big(R^{-2}|w|_{0;B(X,2R)} + |\Delta w|_{0,\alpha;B(X,2R),R}\big) \\ &\le Ch(X)\big(\min(R^{-1+\alpha}, R^{-1})|\nabla u|_{0,\alpha;h} + |\Delta u|_{0,\alpha;h,R}\big) \triangleq Ch(X)A, \end{aligned}$$

which in particular shows

$$|\nabla^2 u(X)| = |\nabla^2 w(X)| \le |\nabla^2 w|_{0,\alpha;B(X,R),R} \le Ch(X)A.$$

On the other hand, if $|Y-X| < R$, then

$$\big|\nabla^2 u(X) - \nabla^2 u(Y)\big| \le |X-Y|^\alpha R^{-\alpha} |\nabla^2 w|_{0,\alpha;B(X,R),R} \le Ch(X)A|X-Y|^\alpha R^{-\alpha},$$

and if $|Y-X| \ge R$, then

$$\big|\nabla^2 u(X) - \nabla^2 u(Y)\big| \le |\nabla^2 u(X)| + |\nabla^2 u(Y)| \le Ch(X)A + Ch(Y)A \le C(h(X)+h(Y))A|X-Y|^\alpha R^{-\alpha}.$$

This gives

$$|\nabla^2 u|_{0,\alpha;h,R} = |\nabla^2 u|_{0;h} + R^\alpha [\nabla^2 u]_{\alpha;h} \le CA. \qquad \square$$

Acknowledgment

Zhang is partially supported by NSF of China under under grant 11425103.

References

[Abidi and Zhang 2016] H. Abidi and P. Zhang, "On the global solution of a 3-D MHD system with initial data near equilibrium", *Comm. Pure Appl. Math* (online publication May 2016).

[Bardos et al. 1988] C. Bardos, C. Sulem, and P.-L. Sulem, "Longtime dynamics of a conductive fluid in the presence of a strong magnetic field", *Trans. Amer. Math. Soc.* **305**:1 (1988), 175–191. MR Zbl

[Cai and Lei 2016] Y. Cai and Z. Lei, "Global well-posedness of the incompressible magnetohydrodynamics", preprint, 2016. arXiv

[Califano and Chiuderi 1999] F. Califano and C. Chiuderi, "Resistivity-independent dissipation of magnetrodydrodynamic waves in an inhomogeneous plasma", *Phys. Rev. E* **60**:4 (1999), 4701–4707.

[Cao and Wu 2011] C. Cao and J. Wu, "Global regularity for the 2D MHD equations with mixed partial dissipation and magnetic diffusion", *Adv. Math.* **226**:2 (2011), 1803–1822. MR Zbl

[Cao et al. 2013] C. Cao, D. Regmi, and J. Wu, "The 2D MHD equations with horizontal dissipation and horizontal magnetic diffusion", *J. Differential Equations* **254**:7 (2013), 2661–2681. MR Zbl

[Chemin et al. 2016] J.-Y. Chemin, D. S. McCormick, J. C. Robinson, and J. L. Rodrigo, "Local existence for the non-resistive MHD equations in Besov spaces", *Adv. Math.* **286** (2016), 1–31. MR Zbl

[Cordoba and Fefferman 2001] D. Cordoba and C. Fefferman, "Behavior of several two-dimensional fluid equations in singular scenarios", *Proc. Natl. Acad. Sci. USA* **98**:8 (2001), 4311–4312. MR Zbl

[Fefferman et al. 2014] C. L. Fefferman, D. S. McCormick, J. C. Robinson, and J. L. Rodrigo, "Higher order commutator estimates and local existence for the non-resistive MHD equations and related models", *J. Funct. Anal.* **267**:4 (2014), 1035–1056. MR Zbl

[He et al. 2014] C. He, X. Huang, and Y. Wang, "On some new global existence results for 3D magnetohydrodynamic equations", *Nonlinearity* **27**:2 (2014), 343–352. MR Zbl

[He et al. 2016] L.-B. He, L. Xu, and P. Yu, "On global dynamics of three dimensional magnetohydrodynamics: nonlinear stability of Alfvén waves", preprint, 2016. arXiv

[Jiu et al. 2015] Q. Jiu, D. Niu, J. Wu, X. Xu, and H. Yu, "The 2D magnetohydrodynamic equations with magnetic diffusion", *Nonlinearity* **28**:11 (2015), 3935–3955. MR Zbl

[Lei 2015] Z. Lei, "On axially symmetric incompressible magnetohydrodynamics in three dimensions", *J. Differential Equations* **259**:7 (2015), 3202–3215. MR Zbl

[Lin et al. 2015] F. Lin, L. Xu, and P. Zhang, "Global small solutions of 2-D incompressible MHD system", *J. Differential Equations* **259**:10 (2015), 5440–5485. MR Zbl

[Priest et al. 1998] E. R. Priest, C. R. Foley, J. Heyvaerts, T. D. Arber, J. L. Culhane, and L. W. Acton, "Nature of the heating mechanism for the diffuse solar corona", *Nature* **393**:6685 (1998), 545–547.

[Ren et al. 2014] X. Ren, J. Wu, Z. Xiang, and Z. Zhang, "Global existence and decay of smooth solution for the 2-D MHD equations without magnetic diffusion", *J. Funct. Anal.* **267**:2 (2014), 503–541. MR Zbl

[Ren et al. 2016] X. Ren, Z. Xiang, and Z. Zhang, "Global well-posedness for the 2D MHD equations without magnetic diffusion in a strip domain", *Nonlinearity* **29**:4 (2016), 1257–1291. MR Zbl

[Sermange and Temam 1983] M. Sermange and R. Temam, "Some mathematical questions related to the MHD equations", *Comm. Pure Appl. Math.* **36**:5 (1983), 635–664. MR Zbl

[Zhang 2014] T. Zhang, "An elementary proof of the global existence and uniqueness theorem to 2-D incompressible non-resistive MHD system", preprint, 2014. arXiv

Received 20 Sep 2016. Revised 31 Mar 2017. Accepted 9 May 2017.

DONGYI WEI: jnwdyi@163.com
School of Mathematical Sciences, Peking University, Beijing, 100871, China

ZHIFEI ZHANG: zfzhang@math.pku.edu.cn
School of Mathematical Sciences, Peking University, Beijing, 100871, China

ANALYSIS AND PDE
Vol. 10, No. 6, 2017
dx.doi.org/10.2140/apde.2017.10.1407

NONNEGATIVE KERNELS AND 1-RECTIFIABILITY IN THE HEISENBERG GROUP

VASILEIOS CHOUSIONIS AND SEAN LI

Let E be a 1-regular subset of the Heisenberg group $\mathbb{H}$. We prove that there exists a -1-homogeneous kernel K_1 such that if E is contained in a 1-regular curve, the corresponding singular integral is bounded in $L^2(E)$. Conversely, we prove that there exists another -1-homogeneous kernel K_2 such that the $L^2(E)$-boundedness of its corresponding singular integral implies that E is contained in a 1-regular curve. These are the first non-Euclidean examples of kernels with such properties. Both K_1 and K_2 are weighted versions of the Riesz kernel corresponding to the vertical component of $\mathbb{H}$. Unlike the Euclidean case, where all known kernels related to rectifiability are antisymmetric, the kernels K_1 and K_2 are even and nonnegative.

1. Introduction

One of the standard topics in classical harmonic analysis is the study of singular integral operators (SIOs) of the form

$$Tf(x) = \int \frac{\Omega(x-y)}{|x-y|^n} f(y)\, d\mathcal{L}^n(y),$$

where Ω is a 0-homogeneous function and $\mathcal{L}^n$ is the Lebesgue measure in $\mathbb{R}^n$; see, e.g., [Stein 1993]. A considerable amount of research has been devoted to such SIOs, and nowadays they are well understood. On the other hand if the singular integral is defined on lower-dimensional measures, the situation is much more complicated even when one considers the simplest of kernels.

As an example the reader should think of the Cauchy transform

$$C_E f(z) = \int_E \frac{f(w)}{z-w}\, d\mathcal{H}^1(w), \qquad E \subset \mathbb{C},$$

where $\mathcal{H}^1$ denotes the 1-dimensional Hausdorff measure in the complex plane. Two questions arise naturally. For which sets E is C_E bounded in $L^2(E)$? And, if C_E is bounded in $L^2(E)$, what does this imply about E? Here $L^2(E)$-boundedness means that there exists a constant $C > 0$ such that the truncated operator

$$C_E^{\varepsilon} f(z) = \int_{E \setminus B(z,\varepsilon)} \frac{f(w)}{z-w}\, d\mathcal{H}^1(w)$$

Chousionis was supported by the Academy of Finland through the grant *Geometric harmonic analysis*, grant number 267047. Li is supported by NSF grant DMS-1600804.
MSC2010: primary 28A75; secondary 28C10, 35R03.
Keywords: Heisenberg group, rectifiability, singular integrals.

satisfies $\|C_E^\varepsilon f\|_{L^2(\mathcal{H}^1|_E)} \le C\|f\|_{L^2(\mathcal{H}^1|_E)}$ for all $f \in L^2(\mathcal{H}^1|_E)$. It turns out that the $L^2(E)$-boundedness of the Cauchy transform depends crucially on the geometric structure of E.

The problem of exploring this relation has a long history and it is deeply related to rectifiability and analytic capacity; we refer to the recent book of Tolsa [2014] for an extensive treatment. One of the landmarks in the field was the characterization of the 1-regular sets E on which the Cauchy transform is bounded in $L^2(E)$. Recall that an $\mathcal{H}^1$-measurable set E is 1-*Ahlfors-regular*, if there exists a constant $1 \le C < \infty$ such that

$$C^{-1} r \le \mathcal{H}^1(B(x,r) \cap E) \le Cr$$

for all $x \in E$, and $0 < r \le \operatorname{diam} E$. It turns out that if E is 1-regular, the Cauchy transform C_E is bounded in $L^2(E)$ if and only if E is contained in a 1-regular curve. The sufficient condition is due to David [1988] and it even holds for more general smooth antisymmetric kernels. The necessary condition is due to Mattila, Melnikov and Verdera [Mattila et al. 1996]. It is a remarkable fact that their proof depends crucially on a special subtle positivity property of the Cauchy kernel related to an old notion of curvature named after Menger; see, e.g., [Melnikov and Verdera 1995; Mattila et al. 1996]. We also note that the above characterization also holds for the SIOs associated to the coordinate parts of the Cauchy kernel.

Very few things are known for the action of SIOs associated with other -1-homogeneous, 1-dimensional Calderón–Zygmund kernels (see Section 2 for the exact definition) on 1-regular sets in the complex plane. Call a kernel "good" if its associated SIO is bounded on $L^2(E)$ if and only if E is contained in a 1-regular curve. It is noteworthy that all known good or bad kernels are related to the kernels

$$k_n(z) = \frac{x^{2n-1}}{|z|^{2n}}, \quad z = (x, y) \in \mathbb{C} \setminus \{0\},\ n \in \mathbb{N}.$$

Observe that k_1 is a good kernel as it is the x-coordinate of the Cauchy kernel; see [Mattila et al. 1996]. It was shown in [Chousionis et al. 2012] that the kernels k_n, $n > 1$, are good as well, and these were the first nontrivial examples of good kernels not directly related to the Cauchy kernel. Now let

$$\kappa_t(z) = k_2(z) + t \cdot k_1(z), \quad t \in \mathbb{R}.$$

It follows by [Chousionis et al. 2012] and [Mattila et al. 1996] that κ_t is good for $t > 0$. Recently Chunaev [2016] showed that κ_t is good for $t \le -2$ and Chunaev, Mateu and Tolsa [Chunaev et al. 2016] proved that κ_t is good for $t \in (-2, -\sqrt{2})$. For $t = -1$ and $t = -\frac{3}{4}$ there exist intricate examples of sets E, due to Huovinen [2001] and Jaye and Nazarov [2013] respectively, which show that the $L^2(E)$-boundedness of the SIO associated to κ_{-1} and $\kappa_{-3/4}$ does not imply rectifiability for E. Therefore the kernels $\kappa_{-1}(z) = xy^2/|z|^4$ and $\kappa_{-3/4}(x, y) = (x^3 - 3xy^2)/|z|^4$ are bad kernels.

Notice that all the kernels mentioned so far are odd and this is very reasonable. Consider, for example, a 1-dimensional Calderón–Zygmund kernel $k : \mathbb{R} \times \mathbb{R} \setminus \{x = y\} \to \mathbb{R}^+$ which is not locally integrable along the diagonal. Take, for example, $k(x, y) = |x - y|^{-1}$. Then $\int_I k(x, y)\, dy = \infty$ for all open intervals $I \subset \mathbb{R}$. It becomes evident that defining a SIO which makes sense on lines and other "nice" 1-dimensional objects depends crucially on the cancellation properties of the kernel. Surprisingly in the Heisenberg group $\mathbb{H}$ the situation is very different.

The *Heisenberg group* $\mathbb{H}$ is $\mathbb{R}^3$ endowed with the group law

$$p \cdot q = \big(x + x',\ y + y',\ z + z' + \tfrac{1}{2}(xy' - yx')\big) \tag{1-1}$$

for $p = (x, y, t)$, $q = (x', y', t') \in \mathbb{R}^3$. We use the following metric on $\mathbb{H}$:

$$d_{\mathbb{H}} : \mathbb{H} \times \mathbb{H} \to [0, \infty), \quad d_{\mathbb{H}}(p, q) := N(q^{-1} \cdot p),$$

where $N : \mathbb{H} \to [0, \infty)$ is the *Korányi norm* in $\mathbb{H}$,

$$N(x, y, z) := ((x^2 + y^2)^2 + z^2)^{1/4}.$$

We also let

$$NH(x, y, z) = |z|^{1/2},$$

where NH stands for nonhorizontal. Note that

$$d_{\mathbb{H}}(x, y) = (|\pi(x) - \pi(y)|^4 + NH(x^{-1}y)^4)^{1/4}.$$

We also remark that the metric $d_{\mathbb{H}}$ is homogeneous with respect to the dilations

$$\delta_r : \mathbb{H} \to \mathbb{H}, \quad \delta_r((x, y, z)) = (rx, ry, r^2 z), \quad (r > 0).$$

Finally let $\Omega : \mathbb{H} \setminus \{0\} \to [0, \infty)$,

$$\Omega(p) = \frac{NH(p)}{N(p)} \tag{1-2}$$

and notice that Ω is 0-homogeneous, as $\Omega(\delta_r(p)) = \Omega(p)$ for all $r > 0$. One can also define the dilations for $r < 0$ for which the metric is still 1-homogeneous (although with absolute value).

In our first main theorem we prove that, in contrast to the Euclidean case, there exists a nonnegative, -1 homogeneous, Calderón–Zygmund kernel which is bounded in $L^2(E)$ for every 1-regular set E which is contained in a 1-regular curve. We warn the reader that from now on $\mathcal{H}^1$ will denote the 1-dimensional Hausdorff measure in $(\mathbb{H}, d_{\mathbb{H}})$.

Theorem 1.1. *Let $K_1 : \mathbb{H} \setminus \{0\} \to [0, \infty)$ be defined by*

$$K_1(p) = \frac{\Omega(p)^8}{N(p)},$$

and let E be a 1*-regular set which is contained in a* 1*-regular curve. Then the corresponding truncated singular integrals*

$$T_1^{\varepsilon} f(p) = \int_{E \setminus B_{\mathbb{H}}(p,\varepsilon)} K_1(q^{-1} \cdot p) f(q)\, d\mathcal{H}^1(q)$$

are uniformly bounded in $L^2(E)$.

There are abundant examples of 1-regular sets in $\mathbb{H}$ which are not contained in 1-regular curves. For example, one can consider suitable 1-dimensional Cantor sets in the vertical axis, $T = \{(0, 0, z) : z \in \mathbb{R}\}$, which is 2-dimensional.

We define the principal value of f at p to be

$$\text{p.v.}\, T_1 f(p) = \lim_{\varepsilon \to 0} T_1^{\varepsilon}(f)(p),$$

when the limit exists. Because the kernel is positive, we will be able to use Theorem 1.1 to easily show that the principal value operator is bounded in L^2.

Corollary 1.2. *If $f \in L^2(E)$, then* p.v. $T_1 f(x)$ *exists almost everywhere and is in $L^2(E)$. Moreover, we have that there exists a constant $C > 0$ such that*

$$\|\text{p.v.}\, T_1 f\|_{L^2(E)} \le C \|f\|_{L^2(E)} \quad \forall f \in L^2(E).$$

Let us quickly give an intuition behind why one would expect a positive kernel like $NH(x)^m/N(x)^{m+1}$ to be bounded on Lipschitz curves. Rademacher's theorem says that Lipschitz curves in $\mathbb{R}^n$ infinitesimally resemble affine lines, and antisymmetric kernels cancel on affine lines. This is essentially what controls the singularity. In the Heisenberg setting, a Rademacher-type theorem by Pansu [1989] says that Lipschitz curves infinitesimally resemble *horizontal* lines and NH is 0 on horizontal lines. Thus, we again have control over the singularity.

Some heuristic motivation comes from the fact that the positive Riesz kernel $|z|/(x^2 + y^2 + z^2)^{3/2}$ defines a SIO which is trivially bounded in $\mathbb{R}^3$ for curves in the xy-plane. In this case, however, the boundedness of this SIO tells us nothing about the regularity of the xy-curve. An analogous concern in the Heisenberg group would be whether the boundedness of kernels of the form $NH(z)^p/N(z)^{p+1}$ implies anything about the regularity of the sets if the vertical direction is "orthogonal" to Lipschitz curves. While we do not know if the boundedness of the kernel of Theorem 1.1 says anything about regularity, our next result shows that there exists some p for which these vertical Riesz kernels do:

Theorem 1.3. *Let $K_2 : \mathbb{H} \setminus \{0\} \to [0, \infty)$ be defined by*

$$K_2(p) = \frac{\Omega(p)^2}{N(p)},$$

and let E be a 1-regular set. If the corresponding truncated singular integrals

$$T_2^{\varepsilon} f\,(p) = \int_{E \setminus B_{\mathbb{H}}(p,\varepsilon)} K_2(q^{-1} \cdot p) f(q)\, d\mathcal{H}^1(q)$$

are uniformly bounded in $L^2(E)$ then E is contained in a 1-regular curve.

One can interpret this statement as saying that the vertical fluctuations of a 1-regular set $E \subset \mathbb{H}$ (that is, $K_i(p^{-1} \cdot q)$ when $p, q \in E$) contain enough information to determine that it lies on a 1-regular curve.

The following question arises naturally from Theorems 1.1 and 1.3. *Does there exist some $m \in \mathbb{N}$ such that any 1-regular set E is contained in some 1-regular curve if and only if the operators*

$$T^{\varepsilon} f\,(p) = \int_E \frac{\Omega(q^{-1} \cdot p)^m}{N(q^{-1} \cdot p)} f(q)\, d\mathcal{H}^1(q)$$

are uniformly bounded in $L^2(E)$? The methods developed in this paper do not allow us to answer this question, partly because our proof for Theorem 1.1 seems to require a large power for $\Omega(p)$. This is essential because we are using a positive kernel and so are not able to use antisymmetry to gain additional control from the bilinearity, as is commonly used in these types of arguments; for example, see Section 6.2 of [Tolsa 2009]. The proof of Theorem 1.3 uses delicate estimates regarding the Korányi norm and is also not likely to be improved without a major change in the proof structure.

A motivation for the geometric study of SIOs in $\mathbb{R}^n$ is their significance in PDE and potential theory. In particular the d-dimensional Riesz transforms (the SIOs associated to the kernels $x/|x|^{d+1}$) for $d = 1$ and $d = n-1$ play a crucial role in the geometric characterization of *removable sets* for bounded analytic functions and Lipschitz harmonic functions. Landmark contributions by David [1998], David and Mattila [2000], and Nazarov, Tolsa and Volberg [Nazarov et al. 2014a; 2014b] established that these removable sets coincide with the purely $(n-1)$-unrectifiable sets in $\mathbb{R}^n$, i.e., the sets which intersect every $(n-1)$-dimensional Lipschitz graph in a set of vanishing $(n-1)$-dimensional Hausdorff measure. For an excellent review of the topic and its connections to nonhomogeneous harmonic analysis, we refer the reader to the survey [Volberg and Èiderman 2013].

The same motivation exists in several noncommutative Lie groups as well. For example, the problem of characterizing removable sets for Lipschitz harmonic functions has a natural analogue in Carnot groups. In that case the harmonic functions are, by definition, the solutions to the sub-Laplacian equation. It was shown in [Chousionis and Mattila 2014] that in the case of the Heisenberg group, the dimension threshold for such removable sets is $\dim \mathbb{H} - 1 = 3$, where $\dim \mathbb{H}$ denotes the Hausdorff dimension of $\mathbb{H}$. See also [Chousionis et al. 2015] for an extension of the previous result to all Carnot groups. As in the Euclidean case, one has to handle a SIO whose kernel is the horizontal gradient of the fundamental solution of the sub-Laplacian. For example, in $\mathbb{H}$, such a kernel can be explicitly written as

$$K(p) := \left(\frac{x(x^2+y^2)+yz}{((x^2+y^2)^2+z^2)^{3/2}}, \frac{y(x^2+y^2)-xz}{((x^2+y^2)^2+z^2)^{3/2}} \right)$$

for $p = (x, y, z) \in \mathbb{H}$. Currently we know very little about the action of this kernel on 3-dimensional subsets of $\mathbb{H}$. Nevertheless it has motivated research on SIOs on lower-dimensional subsets of $\mathbb{H}$, e.g., [Chousionis and Mattila 2011] and the present paper, as well as the very recent study of quantitative rectifiability in $\mathbb{H}$; see [Chousionis et al. 2016].

2. Preliminaries

Although we have already defined a metric on $\mathbb{H}$, we will also need the fact that there exists a natural path metric on $\mathbb{H}$. Notice that the Heisenberg group is a Lie group with respect to the group operation defined in (1-1), and the Lie algebra of the left invariant vector fields in $\mathbb{H}$ is generated by the vector fields

$$X := \partial_x + y\partial_z, \quad Y := \partial_y - x\partial_z, \quad T := \partial_z.$$

The vector fields X and Y define the *horizontal subbundle* $H\mathbb{H}$ of the tangent vector bundle of $\mathbb{R}^3$. For every point $p \in \mathbb{H}$ we will denote the horizontal fiber by $H_p\mathbb{H}$. Every such horizontal fiber is endowed with the left invariant scalar product $\langle \cdot, \cdot \rangle_p$ and the corresponding norm $|\cdot|_p$ that make the vector fields X, Y, T orthonormal.

Definition 2.1. An absolutely continuous curve $\gamma : [a, b] \to \mathbb{H}$ will be called *horizontal* with respect to the vector fields X, Y if

$$\dot{\gamma}(t) \in H_{\gamma(t)}\mathbb{H} \quad \text{for a.e. } t \in [a, b].$$

Definition 2.2. The *Carnot–Carathéodory distance* of $p, q \in \mathbb{H}$ is

$$d_{cc}(p, q) = \inf \int_a^b |\dot{\gamma}(t)|_{\gamma(t)}\, dt,$$

where the infimum is taken over all horizontal curves $\gamma : [a, b] \to \mathbb{H}$ such that $\gamma(a) = p$ and $\gamma(b) = q$.

By Chow's theorem, the above set of curves joining p and q is not empty and hence d_{cc} defines a metric on $\mathbb{H}$. Furthermore the infimum in the definition can be replaced by a minimum. See [Bonfiglioli et al. 2007] for more details.

Remark 2.3. It follows by results of Pansu [1982a; 1982b] that any 1-regular curve is horizontal; hence the reader should keep in mind that our two main theorems (Theorems 1.1 and 1.3) essentially involve subsets of horizontal curves.

A point $p \in \mathbb{H}$ is called *horizontal* if p lies on the xy-plane. We can now define an important family of curves in the Heisenberg group.

Definition 2.4. Let $p, q \in \mathbb{H}$ such that q is horizontal. The subsets of the form

$$\{p \cdot \delta_r(q) : r \in \mathbb{R}\}$$

are called *horizontal lines*.

Observe that horizontal lines are horizontal curves with constant tangent vector. Thus, in the horizontal line above, the element q can be thought of as defining a "horizontal direction" for the line.

Note also that the horizontal lines going through a specified point in $\mathbb{H}$ span only two dimensions instead of three as in $\mathbb{R}^3$. This is a significant difference between Heisenberg and Euclidean geometry.

It is easy to see that the homomorphic projection $\pi : \mathbb{H} \to \mathbb{R}^2$ defined by

$$\pi(x, y, z) = (x, y)$$

is 1-Lipshitz. We will also use the map $\tilde{\pi} : \mathbb{H} \to \mathbb{H}$ defined by

$$\tilde{\pi}(x, y, z) = (x, y, 0).$$

We stress that $\tilde{\pi}$ is *not* a homomorphism.

Definition 2.5 (horizontal interpolation). For $p, q \in \mathbb{H}$,

$$\overline{pq} = \{p \cdot \delta_r \tilde{\pi}(p^{-1} \cdot q) : r \in [0, 1]\}.$$

Note that $\overline{pq}$ is a horizontal segment starting from p traveling in the horizontal direction of $p^{-1} \cdot q$.

Definition 2.6. Let (X, d) be a metric space. We say that

$$k(\cdot,\cdot): X \times X \setminus \{x = y\} \to \mathbb{R}$$

is an n-dimensional *Calderón–Zygmund* (CZ)-kernel if there exist constants $c > 0$ and η, with $0 < \eta \leq 1$, such that for all $x, y \in X$, with $x \neq y$,

(1) $|k(x, y)| \leq c/d(x, y)^n$,

(2) $|k(x, y) - k(x', y)| + |k(y, x) - k(y, x')| \leq cd(x, x')^\eta / d(x, y)^{n+\eta}$ if $d(x, x') \leq d(x, y)/2$.

For the next lemma, recall the definition (1-2) of the functions Ω.

Lemma 2.7. *Fix $m \in \mathbb{N}$, and let $k : \mathbb{H} \times \mathbb{H} \setminus \{x = y\} \to \mathbb{R}$ be defined as*

$$k(p, q) = \frac{\Omega(q^{-1} \cdot p)^m}{N(q^{-1} \cdot p)}.$$

Then k is a 1-dimensional CZ-kernel.

Proof. We need to verify (1) and (2) from Definition 2.6. Notice that (1) is immediate because by the definition of the Korányi norm, $NH(p) \leq N(p)$ for all $p \in \mathbb{H}$. For (2) we will use the fact that the function

$$f(p) = \frac{\Omega(p)^m}{N(p)}, \quad p \in \mathbb{H} \setminus \{0\},$$

is C^1 away from the origin and it is also -1-homogeneous, that is,

$$f(\delta_r(p)) = \frac{1}{r} f(p)$$

for all $r > 0$ and $p \in \mathbb{H} \setminus \{0\}$. Hence by [Folland and Stein 1982, Proposition 1.7] there exists some constant $C > 0$ such that for all $P, Q \in \mathbb{H}$ with $N(Q) \leq N(P)/2$,

$$|f(P \cdot Q) - f(P)| \leq C \frac{N(Q)}{N(P)^2}.$$

Hence if $p, p', q \in \mathbb{H}$ such that $d_{\mathbb{H}}(p, p') \leq d_{\mathbb{H}}(p, q)/2$,

$$\begin{aligned} |k(p, q) - k(p', q)| &= |f(q^{-1} \cdot p) - f(q^{-1} \cdot p')| \\ &= |f(q^{-1} \cdot p) - f(q^{-1} \cdot p \cdot p^{-1} \cdot p')| \leq C \frac{N(p'^{-1} \cdot p)}{N(q^{-1} \cdot p)^2} = C \frac{d_{\mathbb{H}}(p', p)}{d_{\mathbb{H}}(p, q)^2}. \end{aligned} \tag{2-1}$$

Since k is symmetric, from (2-1) we deduce that k also satisfies (2) of Definition 2.6. ☐

In the sequel, we will use the notation $a \lesssim b$ or $a \gtrsim b$ to mean that there exists a universal constant C so that $a \leq Cb$ or $a \geq Cb$. This universal constant can change from instance to instance. We let $a \asymp b$ mean both $a \lesssim b$ and $b \lesssim a$. Given another fixed quantity α, we let $a \lesssim_\alpha b$ and $b \lesssim_\alpha a$ mean that the quantity C can depend only on α.

3. Necessary conditions

In order to simplify notation, in the two following sections we will denote $d := d_{\mathbb{H}}$, $B(p,r) := B_{\mathbb{H}}(p,r)$ and $ab := a \cdot b$ for $a, b \in \mathbb{H}$.

Let $E \subset \mathbb{H}$ such that $\mu = \mathcal{H}^1|_E$ satisfies the 1-regularity condition

$$\xi r \leq \mu(B(x,r)) \leq \xi^{-1} r \quad \forall x \in E,\ r > 0,$$

for some $\xi < 1$. We now recall the construction of *David cubes* [1991]. David cubes can be constructed on any regular set of a geometrically doubling metric space. In particular, for the set E, we obtain a constant $c > 0$ and a family of partitions Δ_j of E, $j \in \mathbb{Z}$, with the following properties:

(D1) If $k \leq j$, $Q \in \Delta_j$ and $Q' \in \Delta_k$, then either $Q \cap Q' = \varnothing$ or $Q \subset Q'$.

(D2) If $Q \in \Delta_j$, then $\operatorname{diam} Q \leq 2^{-j}$.

(D3) Every set $Q \in \Delta_j$ contains a set of the form $B(p_Q, c2^{-j}) \cap E$ for some $p_Q \in Q$.

The sets in $\Delta := \bigcup \Delta_j$ are called David cubes, or *dyadic cubes*, of E. Notice that $\operatorname{diam} Q \asymp 2^{-j}$ if $Q \in \Delta_j$. For a cube $S \in \Delta$, we define

$$\Delta(S) := \{Q \in \Delta : Q \subseteq S\}.$$

Given a cube $Q \in \Delta$ and $\lambda \geq 1$, we define

$$\lambda Q := \{x \in E : d(x, Q) \leq (\lambda - 1) \operatorname{diam} Q\}.$$

It follows from (D1), (D2), and the 1-regularity of E that $\mu(Q) \sim 2^{-j}$ for $Q \in \Delta_j$.

Define the positive symmetric -1-homogeneous kernel K by

$$K(p) = \frac{\Omega^8(p)}{N(p)} = \frac{NH(p)^8}{N(p)^9}.$$

For any $\varepsilon > 0$, we can define the truncated operator as before:

$$T_1^\varepsilon f(x) = \int_{d(y,x)>\varepsilon} K(y^{-1}x) f(x)\, d\mu(y).$$

Proof of Theorem 1.1. Our goal is to show that when E lies on a rectifiable curve, there exists a uniform bound $C < \infty$ that can depend on ξ such that

$$\|T_1^\varepsilon \chi_S\|_{L^2(S)}^2 \leq C\mu(S) \quad \forall S \in \Delta,\ \forall \varepsilon > 0. \tag{3-1}$$

We then apply the $T(1)$ theorem for homogeneous spaces — see, e.g., [Deng and Han 2009; David 1991] — to deduce the uniform L^2-boundedness of T_1^ε for all $\varepsilon > 0$. We may suppose E is a 1-regular rectifiable curve, as taking a subset can only decrease the L^2-bound of $T_1^\varepsilon \chi_S$.

From now on we assume the 1-regular set E actually lies on a rectifiable curve. For $x \in E$ and $r > 0$, we define

$$\beta_E(x,r) = \inf_L \sup_{z \in E \cap B(x,r)} \frac{d(z,L)}{r},$$

where the infimum is taken over all horizontal lines.

Proposition 3.1. *There exists a constant $C \geq 1$ depending only ξ so that for any $S \in \Delta$, we have*

$$\sum_{Q \in \Delta(S)} \beta(10Q)^4 \mu(Q) \leq C\mu(S). \tag{3-2}$$

Proof. This essentially follows from Theorem I of [Li and Schul 2016b], which says that there exists some universal constant $C > 0$ such that

$$\int_{\mathbb{H}} \int_0^\infty \beta_E(B(x,t))^4 \, \frac{dt}{t^4} \, d\mathcal{H}^4(x) \leq C\mathcal{H}^1(E)$$

when E is simply a horizontal curve. When E is in addition 1-regular, it is a standard argument to use the Ahlfors regularity to bound this integral from below by a constant multiple — which can depend on ξ — of the left-hand side of (3-2) (after intersecting E with S). In fact, one can easily show that the integral and sum are comparable up to multiplicative constants.

One then gets

$$\sum_{Q \in \Delta(S)} \beta(10Q)^4 \mu(Q) \leq C\mathcal{H}^1(E \cap S) \lesssim_\xi \mu(S),$$

where we again used 1-regularity of E in the final inequality. □

We now fix $S \in \Delta$ a cube.

Now define a positive, even Lipschitz function $\psi : \mathbb{R} \to \mathbb{R}$ such that $\chi_{B(0,1/2)} \leq \psi \leq \chi_{B(0,2)}$. We define

$$\psi_j : \mathbb{H} \to \mathbb{R}, \quad z \mapsto \psi(2^j N(z)),$$

and $\phi_j := \psi_j - \psi_{j+1}$. Thus, ϕ_j is supported on the annulus $B(0, 2^{1-j}) \backslash B(0, 2^{-2-j})$ in $\mathbb{H}$ and we have

$$\chi_{\mathbb{H}\backslash B(0,2^{-n+1})} \leq \sum_{n \leq N} \phi_n \leq \chi_{\mathbb{H}\backslash B(0,2^{-n-2})}. \tag{3-3}$$

For each $j \in \mathbb{Z}$, we can define $K_{(j)} = \phi_j \cdot K$ and also

$$T_{(j)}\chi_S(x) = \int_S K_{(j)}(y^{-1}x) \, d\mu(y).$$

Define $S_N = \sum_{n \leq N} T_{(n)}$. As the kernel K is positive, we can easily get the following *pointwise* estimates for any positive function f from (3-3):

$$0 \leq T_1^\varepsilon f \leq S_{n+1} f \quad \forall \varepsilon \geq 2^{-n}.$$

Thus, to show uniform bound (3-1), it suffices to show that there exists $C < \infty$ depending possibly on ξ such that

$$\|S_n \chi_S\|^2_{L^2(S)} \leq C\mu(S) \quad \forall S \in \Delta, \ \forall n \in \mathbb{Z}.$$

We now fix $S \in \Delta_\ell$.

We will need the following lemma.

Lemma 3.2 [Li and Schul 2016a, Lemma 3.3]. *For every $a, b \in \mathbb{H}$ and horizontal line $L \subset \mathbb{H}$, we have*

$$\max\{d(a, L), d(b, L)\} \geq \frac{1}{16} \frac{NH(a^{-1}b)^2}{d(a, b)}. \tag{3-4}$$

Lemma 3.3. *For any $j \in \mathbb{Z}$ and $x \in E$, we have*

$$T_{(j)}1(x) \lesssim_\xi \beta_E(x, 2^{1-j})^4. \tag{3-5}$$

Proof. Define the annulus $A = E \cap A(x, 2^{-2-j}, 2^{1-j})$. Then

$$T_{(j)}1(x) \leq \int_E \phi_j(y^{-1}x)K(y^{-1}x)\,d\mu(y) \leq 2^{j+2}\int_A \frac{NH(y^{-1}x)^8}{N(y^{-1}x)^8}\,d\mu(y) \lesssim_\xi \sup_{y\in A}\frac{NH(y^{-1}x)^8}{d(x,y)^8}.$$

It suffices to show

$$\frac{NH(y^{-1}x)^8}{d(x,y)^8} \leq 8^4\beta_E(B(x,2^{1-j}))^4$$

when $y \in A$. This follows easily from (3-4). Indeed, as $y \in A$, we have $d(x, y) \geq 2^{-j-2}$. We can then find a horizontal line so that

$$\beta_{\{x,y\}}(B(x,2^{1-j})) = \frac{\max\{d(x,L),d(y,L)\}}{2^{1-j}} \geq \frac{\max\{d(x,L),d(y,L)\}}{8d(x,y)} \overset{(3\text{-}4)}{\geq} \frac{NH(x^{-1}y)^2}{128d(x,y)^2}.$$

The statement now follows as $\beta_E(B(x,2^{1-j})) \geq \beta_{\{x,y\}}(B(x,2^{1-j}))$. □

We now have the following easy corollary.

Corollary 3.4. *Let $R \in \Delta_j$. Then for any $\alpha > 0$, we have*

$$\int_R T_{(j)}1(x)^\alpha\,d\mu(x) \lesssim_\xi \beta_E(10R)^{4\alpha}\mu(R). \tag{3-6}$$

Remark 3.5. We may replace the constant 1 function in (3-5) and (3-6) with any positive function $f \leq 1$ (such as $f = \chi_S$ for some $S \in \Delta$). This is again because the kernel of T_j is positive and so respects the partial ordering of positive functions.

For any $Q \in \Delta$, we can also define

$$T_Q\chi_S := \chi_Q T_{(j(Q))}\chi_S.$$

Thus, we have

$$S_n\chi_S = \sum_{j\leq n} T_{(j)}\chi_S = \sum_{j\leq n}\sum_{Q\in\Delta_j} T_Q\chi_S.$$

and so

$$\|S_n\chi_S\|^2_{L^2(S)} = \sum_{j\leq n}\|T_{(j)}\chi_S\|^2_{L^2(S)} + 2\sum_{j<k\leq n}\langle T_{(j)}\chi_S, T_{(k)}\chi_S\rangle, \tag{3-7}$$

where the inner product $\langle\cdot,\cdot\rangle$ is integration on S. We will bound the two terms on the right-hand side separately.

Let $S^* \in \Delta_{\ell-2}$ be such that $S \subset S^*$. It follows from (D1) that S^* is unique for S. It follows from the ϕ_j factor and the fact that cubes of Δ_ℓ have diameter at most $2^{-\ell}$ that $T_{(j)}\chi_S(x) = 0$ for $x \in S \in \Delta_\ell$ whenever $j < \ell - 2$. Thus, as $S \in \Delta_\ell$, we have

$$\sum_{j\leq n}\|T_{(j)}\chi_S\|^2_{L^2(S)} \leq \sum_{\ell-2\leq j\leq n}\sum_{Q\in\Delta_j, Q\subseteq S}\int_Q T_{(j)}\chi_S(x)^2\,d\mu(x) \overset{(3\text{-}6)}{\lesssim_\xi} \sum_{Q\in\Delta(S^*)}\beta(10Q)^8\mu(Q). \tag{3-8}$$

We now have to bound the off-diagonal terms of (3-7). We have

$$\begin{aligned}\sum_{j\geq \ell-2}\sum_{j<k\leq n}\int_S T_{(j)}\chi_S(x)\cdot T_{(k)}(x)\chi_S\,d\mu(x) &\overset{(3\text{-}5)}{\lesssim_\xi} \sum_{j\geq \ell-2}\sum_{Q\in\Delta_j(S)}\beta(10Q)^4\sum_{k>j}\int_Q T_{(k)}\chi_S\,d\mu(x)\\ &\overset{(3\text{-}6)}{\lesssim_\xi} \sum_{Q\in\Delta(S^*)}\beta(10Q)^4\sum_{R\in\Delta(Q)}\beta(10R)^4\mu(R)\\ &\overset{(3\text{-}2)}{\lesssim_\xi} C\sum_{Q\in\Delta(S^*)}\beta(10Q)^4\mu(Q).\end{aligned}\tag{3-9}$$

Note that the constants hidden in the $\lesssim$ of (3-8) and (3-9) do not depend on S or n.

Altogether, we have

$$\|S_n\chi_S\|^2_{L^2(S)} \overset{(3\text{-}7)\text{–}(3\text{-}9)}{\lesssim_\xi} \sum_{Q\in\Delta(S^*)}\beta(10Q)^4\mu(Q) \overset{(3\text{-}2)}{\lesssim_\xi} \mu(S^*)\lesssim_{\xi,c}\mu(S),$$

where we used properties (D2), (D3), and 1-regularity of E in the last inequality. □

We now demonstrate how using a positive kernel leads to an easy proof of Corollary 1.2.

Proof of Corollary 1.2. First suppose that $f\in L^2(E)$ is a nonnegative function. Then as the kernel K_1 is positive, we have for fixed $p\in E$ that $T_1^\varepsilon f(p)$ is a monotonically increasing sequence as $\varepsilon\to 0$ and so

$$\text{p.v. } T_1 f(p) := \lim_{\varepsilon\to 0} T_1^\varepsilon f(p)$$

is a well-defined function, although it be infinity. By Theorem 1.1, we get that there exists some $C>0$ such that

$$\sup_{\varepsilon>0}\int (T_1^\varepsilon f)^2\,d\mu \leq C\int f^2\,d\mu.$$

Thus, by Fatou's lemma, we get

$$\int (\text{p.v. } T_1 f)^2\,d\mu \leq \liminf_{\varepsilon\to 0}\int (T_1^\varepsilon f)^2 \leq C\int f^2\,d\mu.$$

This then proves the corollary for nonnegative functions.

Now let $f\in L^2(E)$ be a real-valued function. We have the decomposition $f=f^+-f^-$, where $f^+=\max\{f,0\}$ and $f^-=\max\{-f,0\}$. Then

$$\max(\|f^+\|_{L^2(E)},\|f^-\|_{L^2(E)})\leq \|f\|_{L^2(E)}$$

and so we get that the principal values of f^+ and f^- under T_1 are controlled by $C\|f\|_{L^2(E)}$. Thus, the principal values have to be finite almost everywhere and so we get p.v. $T_1 f=$ p.v. $T_1 f^+ -$ p.v. $T_1 f^-$ as $L^2(E)$ functions. Additionally, we get

$$\|\text{p.v. } T_1 f\|_{L^2(E)} \leq \|\text{p.v. } T_1 f^+\|_{L^2(E)} + \|\text{p.v. } T_1 f^-\|_{L^2(E)} \leq 2C\|f\|_{L^2(E)}.$$

This proves the entire corollary. □

4. Sufficient conditions

We will need the following “triangle inequality” for this section.

Lemma 4.1 (NH^2 triangle inequality). *Let $a, b, c \in \mathbb{H}$ and let A be the (unsigned) area of the triangle in $\mathbb{R}^2$ with vertices $\pi(a), \pi(b), \pi(c)$. For the four quantities*

$$A, \quad NH(a^{-1}b)^2, \quad NH(b^{-1}c)^2, \quad NH(c^{-1}a)^2,$$

any one of these numbers is less than the sum of the other three.

Proof. Let us first show A is less than the sum of the other three quantities. Since everything is invariant under left translation, we may suppose $c = (0, 0, 0)$, $a = (x, y, t)$, and $b = (x', y', t')$. Then $NH(c^{-1}a)^2 = |t|$ and $NH(b^{-1}c)^2 = |t'|$ and we have

$$A = \tfrac{1}{2}|x'y - xy'| \le \left|\tfrac{1}{2}x'y - xy' - t + t'\right| + |t'| + |t| \le NH(a^{-1}b)^2 + NH(b^{-1}c)^2 + NH(c^{-1}a)^2.$$

We now show that $NH(a^{-1}b)^2$ is less than the sum of the other three quantities. We will keep the same normalization as the last case:

$$NH(a^{-1}b)^2 = \left|\tfrac{1}{2}x'y - xy' - t + t'\right| \le \tfrac{1}{2}|x'y - xy'| + |t'| + |t| \le A + NH(b^{-1}c)^2 + NH(c^{-1}a)^2. \quad \square$$

For $r < R$ and $x \in \mathbb{H}$, we can define the annulus

$$A(x, r, R) := \{y \in \mathbb{H} : d(x, y) \in (r, R)\}.$$

For three points p_1, p_2, p_3 in $\mathbb{H}$, we define

$$\partial(p_1, p_2, p_3) = \min_{\sigma \in S_3}\bigl\{d(p_{\sigma(1)}, p_{\sigma(2)}) + d(p_{\sigma(2)}, p_{\sigma(3)}) - d(p_{\sigma(1)}, p_{\sigma(3)})\bigr\}.$$

For $\alpha \in (0, 1)$, $r > 0$, and a metric space X, we let $\Sigma_X(\alpha, r)$ denote the triples of points $(p_1, p_2, p_3) \in X$ such that

$$\alpha r \le d(p_i, p_j) \le r \quad \forall i \ne j.$$

We also let $\Sigma_X(\alpha) = \bigcup_{r>0} \Sigma_X(\alpha, r)$. For notational convenience, we will drop the X subscript when we want $X = E$, where E is the 1-regular set of the hypothesis of Theorem 1.3.

Lemma 4.2. *Let $(p_1, p_2, p_3) \in \Sigma(\alpha, r)$. If for some $\varepsilon \in (0, 1/2)$ we have*

$$NH(p_i^{-1}p_j) \le \varepsilon d(p_i, p_j), \tag{4-1}$$

then the point $\pi(p_i) \in \mathbb{R}^2$ is contained in the strip around the line $\overline{\pi(p_{i+1}), \pi(p_{i+2})}$ of width $16\alpha^{-1}\varepsilon^2 r$.

Proof. We will view $\overline{\pi(p_2), \pi(p_3)}$ as the base of a triangle with top vertex $\pi(p_1)$. It suffices to bound the height. We let A denote the area of the triangle.

Suppose $A \ge 4\varepsilon^2 r^2$. We have by the NH^2 triangle inequality that

$$NH(p_2^{-1}p_3)^2 \ge A - NH(p_1^{-1}p_2)^2 - NH(p_1^{-1}p_3)^2 \overset{(4\text{-}1)}{\ge} 2\varepsilon^2 r^2.$$

This is a contradiction of (4-1).

Thus, we may assume $A \leq 4\varepsilon^2 r^2$. But if $NH(p_2^{-1}p_3) \leq d(p_2, p_3)/2$, then $|\pi(p_2) - \pi(p_3)| \geq d(p_2, p_3)/2 \geq \alpha r/2$. Thus, the height of the triangle is less than

$$\frac{2A}{|\pi(p_2) - \pi(p_3)|} \leq \frac{16}{\alpha}\varepsilon^2 r. \qquad \square$$

Given $u, v, w \in \mathbb{H}$, we denote the largest and second largest quantities of

$$\left\{\frac{NH(u^{-1}v)}{d(u, v)}, \frac{NH(v^{-1}w)}{d(v, w)}, \frac{NH(u^{-1}w)}{d(u, w)}\right\}$$

by $\gamma_1(u, v, w)$ and $\gamma_2(u, v, w)$, respectively.

Lemma 4.3. *For all $\alpha > 0$, there exists a constant $c_1 > 0$ such that if $(p_1, p_2, p_3) \in \Sigma(\alpha, r)$, then*

$$\partial(p_1, p_2, p_3) \leq c_1 \gamma_1(p_1, p_2, p_3)^4 r.$$

Proof. Let $\gamma = \gamma_1(p_1, p_2, p_3)$, and we may suppose without loss of generality that

$$\partial(p_1, p_2, p_3) = d(p_1, p_2) + d(p_2, p_3) - d(p_1, p_3).$$

Suppose first that $\gamma < c$ for some $c > 0$ to be determined soon. Then

$$NH(p_i^{-1}p_j) \leq \gamma d(p_i, p_j) < c d(p_i, p_j) \quad \forall i \neq j, \tag{4-2}$$

and so

$$|\pi(p_i) - \pi(p_j)| = \bigl(d(p_i, p_j)^4 - NH(p_i^{-1}p_j)^4\bigr)^{1/4} \geq (1 - c^4)^{1/4} d(p_i, p_j).$$

By taking c small enough, we get that $(\pi(p_1), \pi(p_2), \pi(p_3)) \in \Sigma_{\mathbb{R}^2}(\alpha/2)$ and, by Taylor expansion of the Korányi norm, that

$$d(p_i, p_j) \leq |\pi(p_i) - \pi(p_j)| + \frac{NH(p_i^{-1}p_j)^4}{|\pi(p_i) - \pi(p_j)|^3} \leq |\pi(p_i) - \pi(p_j)| + (1 - c^4)^{-3/4}\gamma^4 r,$$

and so

$$\partial(p_1, p_2, p_3) \leq |\pi(p_1) - \pi(p_2)| + |\pi(p_2) - \pi(p_3)| - |\pi(p_1) - \pi(p_3)| + 2(1 - c^4)^{-3/4}\gamma^4 r. \tag{4-3}$$

As $(\pi(p_1), \pi(p_2), \pi(p_3)) \in \Sigma_{\mathbb{R}^2}(\alpha/2)$, we get by a Taylor approximation of the Euclidean metric that

$$|\pi(p_1) - \pi(p_2)| + |\pi(p_2) - \pi(p_3)| - |\pi(p_1) - \pi(p_3)| \lesssim_\alpha \frac{h^2}{r}, \tag{4-4}$$

where h is the height of the triangle in $\mathbb{R}^2$ defined by $\pi(p_i)$ with base $\overline{\pi(p_1), \pi(p_3)}$. From (4-1) and (4-2), we have

$$h \leq 16\alpha^{-1}\gamma^2 r. \tag{4-5}$$

The result now follows from (4-3)–(4-5).

Now suppose $\gamma \geq c$. As $\partial(p_1, p_2, p_3) \leq 3r$, the lemma trivially follows. $\square$

We let $E \subset \mathbb{H}$ be a set with $\mu = \mathcal{H}^1|_E$ satisfying the estimate

$$\xi r \le \mu(B(x,r)) \le \xi^{-1} r \quad \forall x \in E,\ r > 0,$$

where $\xi \le 1$.

Lemma 4.4. *Let $E \subset \mathbb{H}$ be a 1-regular set and $\alpha \in (0,1)$. There exists $c_2 \ge 1$ depending on α and ξ such that if $(p_1, p_2, p_3) \in \Sigma(\alpha, r)$, then one of the following is true*:

(1) $\gamma_1(p_1,p_2,p_3) \le c_2\gamma_2(p_1,p_2,p_3)$.

(2) *After a possible reindexing of p_i, there exists a set $V \subseteq E \cap B(p_1, \alpha r/10)$ with $\mu(V) \ge r/c_2$ such that for every $x \in V$ we have*

$$\gamma_1(p_1,p_2,p_3) \le c_2\gamma_2(x,p_2,p_3)$$

and $(x,p_2,p_3) \in \Sigma(c_2^{-1})$.

(3) *After a possible reindexing of p_i, there exist sets $W_1, W_2 \subseteq E \cap B(p_1, \alpha r/5)$ with $\mu(W_1), \mu(W_2) \ge r/c_2$ such that for all $(x,y) \in W_1 \times W_2$ we have*

$$\gamma_1(p_1,p_2,p_3) \le c_2\gamma_2(p_1,x,y)$$

and $(p_1,x,y) \in \Sigma(c_2^{-1}, r)$.

Proof. Throughout this proof, we will give a finite series of lower bounds for c_2. The final c_2 will then just be the maximum of these lower bounds. For simplicity of notation, let $\gamma_i = \gamma_i(p_1,p_2,p_3)$. We may of course suppose that $\gamma_2 \le c\gamma_1$ for some small $c > 0$ depending on α and ξ to be determined, as otherwise condition (1) would be satisfied. Without loss of generality, we can assume that $\gamma_1 = NH(p_2^{-1}p_3)/d(p_2,p_3)$. Let A denote the area of the triangle in $\mathbb{R}^2$ with vertices $\pi(p_i)$. Then we have from the NH^2 triangle inequality that

$$NH(p_2^{-1}p_3)^2 \le NH(p_1^{-1}p_2)^2 + NH(p_1^{-1}p_3)^2 + A,$$

and so if we set $c < \alpha/2$ (while still allowing ourselves to take c smaller) then

$$A \ge \tfrac{1}{2}\alpha^2\gamma_1^2 r^2. \tag{4-6}$$

Fix $\lambda \in (0,1)$ depending only ξ so that

$$\mu(A(x,\lambda\ell,\ell)) \ge \tfrac{1}{2}\xi\ell \quad \forall x \in E,\ \ell > 0.$$

Suppose now $A(p_1, \lambda\alpha r/10, \alpha r/10)$ contains a subset S of μ-measure at least $\xi\alpha r/40$ so that

$$\frac{NH(x^{-1}p_1)}{d(x,p_1)} < c\gamma_1 \quad \forall x \in S. \tag{4-7}$$

If there is a further subset $V \subseteq S$ with $\mu(V) \ge \xi\alpha r/80$ such that $NH(x^{-1}p_2) \ge c\gamma_1 d(x,p_2)$ for each $x \in V$, then we are done as we've satisfied condition (2) for large enough c_2 if we keep p_2, p_3 and draw x from V.

Thus, suppose there is a subset $V \subseteq S$ with $\mu(V) \ge \xi\alpha r/80$ and

$$\frac{NH(x^{-1}p_2)}{d(x,p_2)} < c\gamma_1 \quad \forall x \in V. \tag{4-8}$$

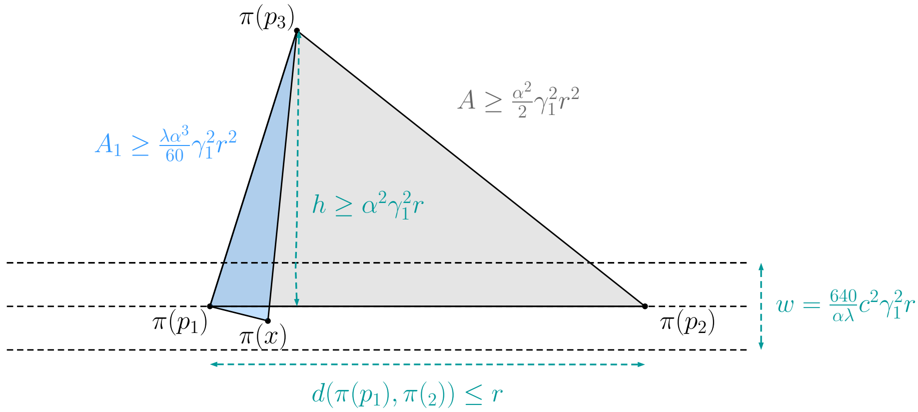

Figure 1. A denotes the area of the triangle determined by $\pi(p_i)$, $i = 1, 2, 3$, and A_1 denotes the area of the triangle determined by $\pi(p_1)$, $\pi(p_3)$ and $\pi(x)$.

Recalling

$$d(x, p_1) \in \left[\tfrac{1}{10}\lambda\alpha r, \tfrac{1}{10}\alpha r\right], \quad d(x, p_2) \in \left[\tfrac{1}{2}r, 2r\right], \quad \forall x \in V \subseteq A\big(p_1, \tfrac{1}{10}\lambda\alpha r, \tfrac{1}{10}\alpha r\big), \tag{4-9}$$

from (4-7), (4-8), and Lemma 4.2, for every $x \in V$ we get that $\pi(x)$ lies in the strip around $\overline{\pi(p_1), \pi(p_2)}$ of width

$$w = \frac{640}{\lambda\alpha}c^2\gamma_1^2 r. \tag{4-10}$$

As $NH(x^{-1}p_1) < c\gamma_1 d(x, p_1)$, we easily get (supposing c is small enough) that

$$|\pi(x) - \pi(p_1)| \geq \tfrac{1}{2}d(x, p_1) \overset{(4\text{-}9)}{\geq} \tfrac{1}{20}\lambda\alpha r. \tag{4-11}$$

As $d(p_1, p_2) \leq r$, we get that the height of the triangle given by $\pi(p_i)$ with base $\overline{\pi(p_1), \pi(p_2)}$ is then at least

$$h \geq \frac{2A}{d(p_1, p_2)} \overset{(4\text{-}6)}{\geq} \alpha^2\gamma_1^2 r.$$

Let A_1 denote the area of the triangle determined by $\pi(p_1)$, $\pi(x)$, $\pi(p_3)$. By (4-10), we have that w is at most some constant multiple (depending on α and λ) of $c^2 h$. Thus, if we choose c small enough to get $\pi(x)$ sufficiently close to the line $\overline{\pi(p_1), \pi(p_2)}$ compared to h, we get

$$A_1 \geq \tfrac{1}{3}h|\pi(p_1) - \pi(x)| \overset{(4\text{-}11)}{\geq} \tfrac{1}{60}\lambda\alpha^3\gamma_1^2 r^2.$$

See Figure 1 for an illustration of these triangles.

Now using the NH^2 triangle inequality, we get

$$\tfrac{1}{60}\alpha^3\lambda\gamma_1^2 r^2 \leq A_1 \leq NH(x^{-1}p_1)^2 + NH(p_1^{-1}p_3)^2 + NH(x^{-1}p_3)^2 \overset{(4\text{-}7),(4\text{-}9)}{\leq} 2c^2\gamma_1^2 r^2 + NH(x^{-1}p_3)^2.$$

Thus, if we choose c small enough compared to α and λ once and for all, we get

$$NH(x^{-1}p_3) \geq \tfrac{1}{10}\sqrt{\alpha^3\lambda}\gamma_1 r \geq \tfrac{1}{20}\sqrt{\alpha^3\lambda}\gamma_1 d(x, p_3).$$

Now we can satisfy condition (2) for sufficiently large c_2 by keeping p_2, p_3 and drawing x from V.

Thus, we may suppose that $E \cap A(p_1, \lambda\alpha r/10, \alpha r/10)$ contains a subset S so that $\mu(S) \geq \xi\alpha r/40$ and

$$NH(z^{-1}p_1) \geq c\gamma_1 d(z, p_1) \quad \forall z \in S.$$

Using the 1-regularity of E, an elementary, although tedious, packing argument shows that there exist $\eta, \tau < \lambda\alpha/100$ depending only on α and ξ and points $x', y' \in E \cap A(p_1, \lambda\alpha r/10, \alpha r/10)$ such that $d(x', y') \geq 10\tau r$ and

$$\min\bigl\{\mu(S \cap B(x', \tau r)),\ \mu(S \cap B(y', \tau r))\bigr\} \geq \eta r.$$

Note by the triangle inequality that we get

$$B(x', \tau r), B(y', \tau r) \subseteq A\bigl(p_1, \tfrac{1}{20}\lambda\alpha r, \tfrac{1}{5}\alpha r\bigr).$$

Thus, after setting c_2 large enough, we've satisfied condition (3) with $W_1 = S \cap B(x', \tau r)$ and $W_2 = S \cap B(y', \tau r)$, which would completely finish the proof of the lemma. We will present a quick sketch of the packing argument and leave the details to the reader.

Find a maximal τr-separated net $\mathcal{N}$ of $E \cap B(p_1, \alpha r)$ for $\tau > 0$ to be determined. By 1-regularity, we have $\#\mathcal{N} \gtrsim \alpha/\tau$. First use the 1-regularity of E to find $M \geq 1$ such that any subset $S \subseteq \mathcal{N}$ for which $\#S \geq M$ must contain $x', y' \in S$ so that $d(x', y') \geq 10\tau r$. Now $\{B(x, \tau r) : x \in \mathcal{N}\}$ is a covering of $B(p_1, \alpha r/10)$. Define $\mathcal{B} = \{B(x, \tau r) : x \in \mathcal{N}, \mu(S \cap B(x, r)) \geq \eta r\}$. By choosing η small enough relative to $\alpha\tau$, we can use the 1-regularity of E and the fact that $\mu(S) \gtrsim \alpha r$ to get that $\#\mathcal{B} \gtrsim \alpha\mathcal{N} \gtrsim \alpha^2/\tau$ (with no dependence on η). Now simply choose τ small enough so that $\#\mathcal{B} \geq M$. One then finds two balls $B(x', \tau r), B(y', \tau r) \in \mathcal{B}$ such that $d(x', y') \geq 10\tau r$, which finishes the sketch. $\square$

For $x, y \in E$, we let

$$\Sigma(\alpha, r; x) := \{(y, z) \in E^2 : (x, y, z) \in \Sigma(\alpha, r)\}, \quad \Sigma(\alpha; x, y) := \{z \in E : (x, y, z) \in \Sigma(\alpha)\}.$$

One easily has that there exists some constant $c_3 \geq 1$ depending on ξ such that

$$\frac{1}{c_3} r^2 \leq \mu \times \mu(\Sigma(\alpha, r; x)) \leq c_3 r^2, \quad \frac{1}{c_3} d(x, y) \leq \mu(\Sigma(\alpha; x, y)) \leq c_3 d(x, y).$$

For simplicity, we will adopt the convention that the integral $\int_A f(x)\,dx$ means $\int_A f(x)\,d\mu(x)$ when $A \subseteq E$. Recall that for three points p_1, p_2, p_3 in a metric space X, the Menger curvature $c(p_1, p_2, p_3) \in \mathbb{R}$ is defined as

$$c(p_1, p_2, p_3) = \frac{1}{R},$$

where R is the radius of the circle in $\mathbb{R}^2$ passing through a triangle defined by the vertices $p_1', p_2', p_3' \in \mathbb{R}^2$, where $d(p_i, p_j) = |p_i' - p_j'|$.

Proposition 4.5. *For any $\alpha > 0$, there exists $c_4 \geq 1$ such that*

$$\iiint_{\Sigma(\alpha)} c(x, y, z)^2\,dx\,dy\,dz \leq c_4 \iiint_{\Sigma(c_4^{-1})} \frac{\gamma_1(x, y, z)^2\,\gamma_2(x, y, z)^2}{\operatorname{diam}(\{x, y, z\})^2}\,dx\,dy\,dz. \tag{4-12}$$

Proof. We have by [Hahlomaa 2005] that there exists some $\tau > 0$ depending on α such that if $(x, y, z) \in \Sigma(\alpha)$, then

$$c(x, y, z)^2 \leq \tau \operatorname{diam}(\{x, y, z\})^{-3}\partial(x, y, z). \tag{4-13}$$

By Lemma 4.3, we have that there exists $c_1 > 0$ such that

$$\iiint_{\Sigma(\alpha)} \operatorname{diam}(\{x,y,z\})^{-3}\partial(x,y,z)\,dx\,dy\,dz \le c_1 \iiint_{\Sigma(\alpha)} \frac{\gamma_1(x,y,z)^4}{\operatorname{diam}(\{x,y,z\})^2}\,dx\,dy\,dz. \tag{4-14}$$

We now decompose $\Sigma(\alpha)$ into three pieces. For $i = 1, 2, 3$, let $S_i \subseteq \Sigma(\alpha)$ denote the triples of points for which condition (i) of Lemma 4.4 holds for some $r > 0$ (that can depend on the triple of points). Note $\Sigma(\alpha) \subseteq S_1 \cup S_2 \cup S_3$, but this decomposition need not be disjoint.

It will be convenient to define the functions

$$f(x,y,z) := \frac{\gamma_1(x,y,z)^4}{\operatorname{diam}(\{x,y,z\})^2}, \quad g(x,y,z) := \frac{\gamma_1(x,y,z)^2\,\gamma_2(x,y,z)^2}{\operatorname{diam}(\{x,y,z\})^2}.$$

We trivially have that

$$\iiint_{S_1} f(x,y,z)\,dx\,dy\,dz \le c_2^2 \iiint_{S_1} g(x,y,z)\,dx\,dy\,dz. \tag{4-15}$$

When we write a triple of points $(x,y,z) \in S_2$, we will always assume y, z play the role of p_2, p_3 in condition (2). Now let $(x,y,z) \in S_2 \cap \Sigma(\alpha)$. We then have that there exists a subset with $\mu(V) \ge r/c_2$,

$$f(x,y,z) \le c_2 g(u,y,z) \quad \forall u \in V.$$

We then have

$$f(x,y,z) \le c_2 \frac{1}{\mu(V)} \int_V g(u,y,z)\,du.$$

We also have $(u,y,z) \in \Sigma(c_2^{-1})$ for all $u \in V$ and so

$$\int_{\Sigma(\alpha;y,z)} f(x,y,z)\,dx \le c_2 \frac{\mu(\Sigma(\alpha;y,z))}{\mu(V)} \int_V g(u,y,z)\,du \le c_2^2 c_3 \int_{\Sigma(c_2^{-1};y,z)} g(u,y,z)\,du.$$

Now we have

$$\begin{aligned}\iiint_{S_2} f(x,y,z)\,dx\,dy\,dz &= \iiint_{\Sigma(\alpha)} \mathbf{1}_{S_2} f(x,y,z)\,dx\,dy\,dz \\ &\le \int_E \int_E \int_{\Sigma(\alpha;y,z)} \mathbf{1}_{S_2} f(x,y,z)\,dx\,dy\,dz \\ &\le c_2^2 c_3 \int_E \int_E \int_{\Sigma(c_2^{-1};y,z)} g(x,y,z)\,dx\,dy\,dz \\ &\le 6c_2^2 c_3 \iiint_{\Sigma(c_2^{-1})} g(x,y,z)\,dx\,dy\,dz. \end{aligned} \tag{4-16}$$

For S_3, we will write the points (x,y,z) with the understanding that z plays the role of p_1 in condition (3). Now let $(x,y,z) \in S_3 \cap \Sigma(\alpha/2, r)$. In a way similar to that above, we can use the properties of the conclusion of property (3) to get that

$$f(x,y,z) \le c_2^2 c_3 \iint_{\Sigma(c_2^{-1},r;z)} g(u,v,z)\,du\,dv.$$

It is elementary to see that if $(x, y, z) \in \Sigma(\alpha)$, then

$$\int_0^\infty \mathbf{1}_{\{r:(x,y)\in\Sigma(\alpha/2,r;z)\}} \frac{dr}{r} \asymp_\alpha 1.$$

Here, we need the extra factor of $\frac{1}{2}$ in case (x, y, z) achieves tightness in the $\Sigma(\alpha)$ condition. We can now decompose the integral:

$$\begin{aligned}
\iiint_{S_3} f(x, y, z)\, dx\, dy\, dz &\lesssim_\alpha \iiint_{S_3} f(x, y, z) \int_0^\infty \mathbf{1}_{\{r:(x,y)\in\Sigma(\alpha/2,r;z)\}} \frac{dr}{r}\, dx\, dy\, dz \\
&\le \int_E \int_0^\infty \iint_{\{(x,y)\in\Sigma(\alpha/2,r;z):(x,y,z)\in S_3\}} f(x, y, z)\, dx\, dy \frac{dr}{r}\, dz \\
&\le c_2^2 c_3 \int_E \int_0^\infty \iint_{\Sigma(c_2^{-1},r;z)} g(u, v, z)\, du\, dv \frac{dr}{r}\, dz \\
&\lesssim_\alpha \int_{\Sigma(c_2^{-1})} g(x, y, z) \int_0^\infty \mathbf{1}_{\{r:(u,v)\in\Sigma(c_2^{-1},r;z)\}} \frac{dr}{r}\, du\, dv\, dz \\
&\lesssim \iiint_{\Sigma(c_2^{-1})} g(x, y, z)\, dx\, dy\, dz. \qquad (4\text{-}17)
\end{aligned}$$

In the second and penultimate inequalities, we used Fubini. We then get the conclusion from (4-13)–(4-16) and (4-17). □

Proof of Theorem 1.3. By a result of Hahlomaa [2007, p. 123], it suffices to show that for some $\alpha > 0$,

$$\iiint_{\Sigma(\alpha)\cap B(p,R)^3} c^2(y_1, y_2, y_3)\, dy_1\, dy_2\, dy_3 \lesssim R \quad \forall p \in E,\ R > 0. \qquad (4\text{-}18)$$

Hence by (4-12), it is enough to prove that for some $\alpha > 0$,

$$\iiint_{\Sigma(\alpha)\cap B(p,R)^3} \frac{\gamma_1(y_1, y_2, y_3)^2 \gamma_2(y_1, y_2, y_3)^2}{\operatorname{diam}(\{y_1, y_2, y_3\})^2}\, dy_1\, dy_2\, dy_3 \lesssim R \quad \forall p \in E,\ R > 0. \qquad (4\text{-}19)$$

By our assumption, for all $\varepsilon > 0$ and every $f \in L^2(E)$,

$$\|T_2^\varepsilon f\|_{L^2(E)} \lesssim \|f\|_{L^2(E)}. \qquad (4\text{-}20)$$

Let $p \in E$ and $R > 0$. Applying (4-20) to $f = \chi_{B(p,R)}$, we get that there exists some $C \ge 0$ such that for every $\varepsilon > 0$,

$$\int_{E\cap B(p,R)} \int_{E\cap B(p,r)\cap B(y_1,\varepsilon)^c} \frac{NH(y_1^{-1}y_2)^2}{d(y_1, y_2)^3}\, dy_2 \int_{E\cap B(p,r)\cap B(y_1,\varepsilon)^c} \frac{NH(y_1^{-1}y_3)^2}{d(y_1, y_3)^3}\, dy_3\, dy_1 \le CR,$$

$$U_\varepsilon = \big\{(y_1, y_2, y_3) \in \Sigma(\alpha) \cap B(p, R)^3 : d(y_1, y_2) > \varepsilon,\ d(y_1, y_3) > \varepsilon\big\},$$

$$V_\varepsilon = \big\{(y_1, y_2, y_3) \in \Sigma(\alpha) \cap B(p, R)^3 : d(y_1, y_2) > \varepsilon,\ d(y_1, y_3) > \varepsilon,\ d(y_2, y_3) > \varepsilon\big\}.$$

It then easily follows from Fubini (remember that all the terms in the integrand are positive) that

$$\iiint_{U_\varepsilon} \frac{NH(y_1^{-1}y_2)^2\, NH(y_1^{-1}y_3)^2}{\operatorname{diam}(\{y_1, y_2, y_3\})^6}\, dy_1\, dy_2\, dy_3 \le CR. \tag{4-21}$$

Therefore,

$$CR \ge \iiint_{V_\varepsilon} \frac{NH(y_1^{-1}y_2)^2\, NH(y_1^{-1}y_3)^2}{\operatorname{diam}(\{y_1, y_2, y_3\})^6}\, dy_1\, dy_2\, dy_3 + \iiint_{U_\varepsilon\setminus V_\varepsilon} \frac{NH(y_1^{-1}y_2)^2\, NH(y_1^{-1}y_3)^2}{\operatorname{diam}(\{y_1, y_2, y_3\})^6}\, dy_1\, dy_2\, dy_3. \tag{4-22}$$

Using the upper regularity of μ, it is not difficult to show that

$$\iiint_{U_\varepsilon\setminus V_\varepsilon} \frac{NH(y_1^{-1}y_2)^2\, NH(y_1^{-1}y_3)^2}{\operatorname{diam}(\{y_1, y_2, y_3\})^6}\, dy_1\, dy_2\, dy_3 \lesssim_\xi R. \tag{4-23}$$

Using (4-21)–(4-23) and letting $\varepsilon \to 0$ we deduce that

$$\iiint_{\Sigma(\alpha)\cap B(p,R)^3} \frac{NH(y_1^{-1}y_2)^2\, NH(y_1^{-1}y_3)^2}{\operatorname{diam}(\{y_1, y_2, y_3\})^6}\, dy_1\, dy_2\, dy_3 \le CR.$$

By permuting variables, we get

$$\iiint_{\Sigma(\alpha)\cap B(p,R)^3} \sum_{\sigma\in S_3} \frac{NH(y_{\sigma(1)}^{-1}y_{\sigma(2)})^2\, NH(y_{\sigma(1)}^{-1}y_{\sigma(3)})^2}{\operatorname{diam}(\{y_1, y, y_3\})^6}\, dy_1\, dy_2\, dy_3 \le 6CR. \tag{4-24}$$

If $(y_1, y_2, y_3) \in \Sigma(\alpha)$, then it follows easily that

$$\begin{aligned}\frac{\gamma_1(y_1, y_2, y_3)^2\, \gamma_2(y_1, y_2, y_3)^2}{\operatorname{diam}(\{y_1, y_2, y_3\})^2} &\lesssim \max_{\sigma\in S_3} \frac{NH(y_{\sigma(1)}^{-1}y_{\sigma(2)})^2\, NH(y_{\sigma(1)}^{-1}y_{\sigma(3)})^2}{\operatorname{diam}(\{y_1, y_2, y_3\})^6} \\ &\le \sum_{\sigma\in S_3} \frac{NH(y_{\sigma(1)}^{-1}y_{\sigma(2)})^2\, NH(y_{\sigma(1)}^{-1}y_{\sigma(3)})^2}{\operatorname{diam}(\{y_1, y_2, y_3\})^6},\end{aligned} \tag{4-25}$$

where the constant multiple implicit in the first inequality depends on α. We then get (4-19) from (4-24) and (4-25). □

5. Norm independence

In this short section we will show that Theorems 1.1 and 1.3 do not depend on the Korányi metric.

Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two homogeneous norms on $\mathbb{H}$ and denote by

$$d_i(p, q) = \|q^{-1}\cdot p\|_i$$

the induced metrics for $i = 1, 2$. We will also denote by $B_i(p, r)$ the balls with respect to the metric d_i for $i = 1, 2$. It is well known — see, e.g., [Bonfiglioli et al. 2007, Proposition 5.1.4] — that all homogeneous norms in a Carnot group are globally equivalent. In particular there exists some $L \ge 0$ such that

$$L^{-1}\|p\|_2 \le \|p\|_1 \le L\|p\|_2 \quad \text{for } p \in \mathbb{H}.$$

Let $s > 0$ and define $k_1, k_2 : \mathbb{H} \setminus \{0\} \to (0, +\infty)$ by

$$k_1(p) = \frac{|z|^s}{\|p\|_1^{2s+1}} \quad \text{and} \quad k_2(p) = \frac{|z|^s}{\|p\|_2^{2s+1}},$$

where $p = (x, y, z) \in \mathbb{H} \setminus \{0\}$. As in the proof of Lemma 2.7 one can show that the kernels $k_i,\ i = 1, 2$, are CZ kernels. Note also that

$$L^{-s-1}\, k_2(p) \le k_1(p) \le L^{s+1}\, k_2(p).$$

Let μ be a 1-regular measure on $\mathbb{H}$ and define the truncated singular integrals

$$S_1^{\varepsilon} f(p) = \int_{B_1(p,\varepsilon)^c} k_1(q^{-1}\cdot p) f(q)\, d\mu(q) \quad \text{and} \quad S_2^{\varepsilon} f(p) = \int_{B_2(p,\varepsilon)^c} k_2(q^{-1}\cdot p) f(q)\, d\mu(q)$$

for $f \in L^2(\mu)$ and $\varepsilon > 0$.

Proposition 5.1. *The operator S_1 is bounded in $L^2(\mu)$ if and only if the operator S_2 is bounded in $L^2(\mu)$.*

Proof. It suffices to show that if S_2 is bounded in $L^2(\mu)$ then S_1 is bounded in $L^2(\mu)$. We define the following auxiliary truncated singular integral for $\varepsilon > 0$ and $f \in L^2(\mu)$:

$$\widetilde{S}_2^{\varepsilon} f(p) = \int_{B_2(p,\varepsilon)^c} k_1(q^{-1}\cdot p) f(q)\, d\mu(q).$$

Let Q be any David cube associated to μ, as in the beginning of Section 3. Then

$$\begin{aligned}\|\widetilde{S}_2^{\varepsilon}\chi_Q\|^2_{L^2(\mu)} &= \int\Big(\int_{Q\cap B_2(p,\varepsilon)^c} k_1(q^{-1}\cdot p)\, d\mu(q)\Big)^2 d\mu(p)\\ &\le L^{2(s+1)}\int\Big(\int_{Q\cap B_2(p,\varepsilon)^c} k_2(q^{-1}\cdot p)\, d\mu(q)\Big)^2 d\mu(p) \le L^{2(s+1)}\|S_2\chi_Q\|^2_{L^2(\mu)} \lesssim \mu(Q)\end{aligned}$$

because S_2 is bounded in $L^2(\mu)$. Hence by the $T(1)$ theorem for homogeneous spaces — see, e.g., [Deng and Han 2009; David 1991] — we deduce that $\widetilde{S}_2$ is bounded in $L^2(\mu)$.

For $f \in L^2(\mu),\ \varepsilon > 0$, and $p \in \mathbb{H}$, we have

$$\begin{aligned}|S_1^{\varepsilon} f(p) - \widetilde{S}_2^{\varepsilon} f(p)| &= \Big|\int_{B_1(p,\varepsilon)^c} k_1(q^{-1}\cdot p) f(q)\, d\mu(q) - \int_{B_2(p,\varepsilon)^c} k_1(q^{-1}\cdot p) f(q)\, d\mu(q)\Big|\\ &\lesssim \int_{B_1(p,\varepsilon)\setminus B_2(p,\varepsilon)} \frac{|f(q)|}{d_1(p,q)}\, d\mu(q) + \int_{B_2(p,\varepsilon)\setminus B_1(p,\varepsilon)} \frac{|f(q)|}{d_1(p,q)}\, d\mu(q).\end{aligned}$$

Note that

$$\begin{aligned}\int_{B_1(p,\varepsilon)\setminus B_2(p,\varepsilon)} \frac{|f(q)|}{d_1(p,q)}\, d\mu(q) &\le \int_{\{q:\varepsilon/L\le d_1(p,q)<\varepsilon\}} \frac{|f(q)|}{d_1(p,q)}\, d\mu(q)\\ &\le \frac{L}{\varepsilon}\int_{B_1(p,\varepsilon)} |f(y)|\, d\mu(q) \approx \frac{1}{\mu(B_1(p,\varepsilon))}\int_{B_1(p,\varepsilon)} |f(y)|\, d\mu(q) \le M_{\mu}^1 f(p),\end{aligned}$$

where M^1_μ denotes the Hardy–Littlewood maximal function with respect to d_1 and μ. Similarly,

$$\int_{B_2(p,\varepsilon)\setminus B_1(p,\varepsilon)} \frac{|f(q)|}{d_1(p,q)}\,d\mu(q) \lesssim M^1_\mu f(p),$$

and we have shown that

$$|S^\varepsilon_1 f(p) - \widetilde{S}^\varepsilon_2 f(p)| \lesssim M^1_\mu f(p).$$

Hence the proposition follows because we already showed that $\widetilde{S}_2$ is bounded in $L^2(\mu)$ and it is also well known that the maximal operator M^1_μ is bounded in $L^2(\mu)$. $\square$

In particular, as a corollary to Theorems 1.1 and 1.3 and Proposition 5.1, we obtain that Theorems 1.1 and 1.3 hold respectively for the kernels

$$K'_1(p) = \frac{|z|^4}{d_{cc}(p,0)^9} \quad \text{and} \quad K'_2(p) = \frac{|z|}{d_{cc}(p,0)^3},$$

where, recalling Definition 2.2, d_{cc} stands for the Carnot–Carathéodory distance.

Acknowledgement

We would like to thank Bruce Kleiner for useful comments and questions, which led to the material of Section 5.

References

[Bonfiglioli et al. 2007] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians*, Springer, 2007. MR Zbl

[Chousionis and Mattila 2011] V. Chousionis and P. Mattila, "Singular integrals on Ahlfors–David regular subsets of the Heisenberg group", *J. Geom. Anal.* **21**:1 (2011), 56–77. MR Zbl

[Chousionis and Mattila 2014] V. Chousionis and P. Mattila, "Singular integrals on self-similar sets and removability for Lipschitz harmonic functions in Heisenberg groups", *J. Reine Angew. Math.* **691** (2014), 29–60. MR Zbl

[Chousionis et al. 2012] V. Chousionis, J. Mateu, L. Prat, and X. Tolsa, "Calderón–Zygmund kernels and rectifiability in the plane", *Adv. Math.* **231**:1 (2012), 535–568. MR Zbl

[Chousionis et al. 2015] V. Chousionis, V. Magnani, and J. T. Tyson, "Removable sets for Lipschitz harmonic functions on Carnot groups", *Calc. Var. Partial Differential Equations* **53**:3-4 (2015), 755–780. MR Zbl

[Chousionis et al. 2016] V. Chousionis, K. Fässler, and T. Orponen, "Intrinsic Lipschitz graphs and vertical β-numbers in the Heisenberg group", preprint, 2016. arXiv

[Chunaev 2016] P. Chunaev, "A new family of singular integral operators whose L^2-boundedness implies rectifiability", preprint, 2016. arXiv

[Chunaev et al. 2016] P. Chunaev, J. Mateu, and X. Tolsa, "Singular integrals unsuitable for the curvature method whose L^2-boundedness still implies rectifiability", preprint, 2016. To appear in *J. Anal. Math.* arXiv

[David 1988] G. David, "Morceaux de graphes lipschitziens et intégrales singulières sur une surface", *Rev. Mat. Iberoamericana* **4**:1 (1988), 73–114. MR Zbl

[David 1991] G. David, *Wavelets and singular integrals on curves and surfaces*, Lecture Notes in Mathematics **1465**, Springer, 1991. MR Zbl

[David 1998] G. David, "Unrectifiable 1-sets have vanishing analytic capacity", *Rev. Mat. Iberoamericana* **14**:2 (1998), 369–479. MR Zbl

[David and Mattila 2000] G. David and P. Mattila, "Removable sets for Lipschitz harmonic functions in the plane", *Rev. Mat. Iberoamericana* **16**:1 (2000), 137–215. MR Zbl

[Deng and Han 2009] D. Deng and Y. Han, *Harmonic analysis on spaces of homogeneous type*, Lecture Notes in Mathematics **1966**, Springer, 2009. MR Zbl

[Folland and Stein 1982] G. B. Folland and E. M. Stein, *Hardy spaces on homogeneous groups*, Mathematical Notes **28**, Princeton University Press, 1982. MR Zbl

[Hahlomaa 2005] I. Hahlomaa, "Menger curvature and Lipschitz parametrizations in metric spaces", *Fund. Math.* **185**:2 (2005), 143–169. MR Zbl

[Hahlomaa 2007] I. Hahlomaa, "Curvature integral and Lipschitz parametrization in 1-regular metric spaces", *Ann. Acad. Sci. Fenn. Math.* **32**:1 (2007), 99–123. MR Zbl

[Huovinen 2001] P. Huovinen, "A nicely behaved singular integral on a purely unrectifiable set", *Proc. Amer. Math. Soc.* **129**:11 (2001), 3345–3351. MR Zbl

[Jaye and Nazarov 2013] B. Jaye and F. Nazarov, "Three revolutions in the kernel are worse than one", preprint, 2013. arXiv

[Li and Schul 2016a] S. Li and R. Schul, "The traveling salesman problem in the Heisenberg group: upper bounding curvature", *Trans. Amer. Math. Soc.* **368**:7 (2016), 4585–4620. MR Zbl

[Li and Schul 2016b] S. Li and R. Schul, "An upper bound for the length of a traveling salesman path in the Heisenberg group", *Rev. Mat. Iberoam.* **32**:2 (2016), 391–417. MR Zbl

[Mattila et al. 1996] P. Mattila, M. S. Melnikov, and J. Verdera, "The Cauchy integral, analytic capacity, and uniform rectifiability", *Ann. of Math.* (2) **144**:1 (1996), 127–136. MR Zbl

[Melnikov and Verdera 1995] M. S. Melnikov and J. Verdera, "A geometric proof of the L^2 boundedness of the Cauchy integral on Lipschitz graphs", *Internat. Math. Res. Notices* **1995**:7 (1995), 325–331. MR Zbl

[Nazarov et al. 2014a] F. Nazarov, X. Tolsa, and A. Volberg, "On the uniform rectifiability of AD-regular measures with bounded Riesz transform operator: the case of codimension 1", *Acta Math.* **213**:2 (2014), 237–321. MR Zbl

[Nazarov et al. 2014b] F. Nazarov, X. Tolsa, and A. Volberg, "The Riesz transform, rectifiability, and removability for Lipschitz harmonic functions", *Publ. Mat.* **58**:2 (2014), 517–532. MR Zbl

[Pansu 1982a] P. Pansu, *Géométrie du groupe d' Heisenberg*, thèse de 3ème cycle, Université Paris VII, 1982, available at https://www.math.u-psud.fr/~pansu/pansu_These_1982.pdf.

[Pansu 1982b] P. Pansu, "Une inégalité isopérimétrique sur le groupe de Heisenberg", *C. R. Acad. Sci. Paris Sér. I Math.* **295**:2 (1982), 127–130. MR Zbl

[Pansu 1989] P. Pansu, "Métriques de Carnot–Carathéodory et quasiisométries des espaces symétriques de rang un", *Ann. of Math.* (2) **129**:1 (1989), 1–60. MR Zbl

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

[Tolsa 2009] X. Tolsa, "Uniform rectifiability, Calderón–Zygmund operators with odd kernel, and quasiorthogonality", *Proc. Lond. Math. Soc.* (3) **98**:2 (2009), 393–426. MR Zbl

[Tolsa 2014] X. Tolsa, *Analytic capacity, the Cauchy transform, and non-homogeneous Calderón–Zygmund theory*, Progress in Mathematics **307**, Springer, 2014. MR Zbl

[Volberg and Èiderman 2013] A. L. Volberg and V. Y. Èiderman, "Nonhomogeneous harmonic analysis: 16 years of development", *Uspekhi Mat. Nauk* **68**:6(414) (2013), 3–58. In Russian; translated in *Russian Math. Surveys* **68**:6 (2013), 973–1026. MR Zbl

Received 21 Oct 2016. Revised 7 Apr 2017. Accepted 9 May 2017.

VASILEIOS CHOUSIONIS: vasileios.chousionis@uconn.edu
Department of Mathematics, University of Connecticut, Storrs, CT 06269, United States

SEAN LI: seanli@math.uchicago.edu
Department of Mathematics, University of Chicago, Chicago, IL 60637, United States

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 10, No. 6, 2017
dx.doi.org/10.2140/apde.2017.10.1429

BERGMAN KERNEL AND HYPERCONVEXITY INDEX

BO-YONG CHEN

Dedicated to Professor John Erik Fornaess on the occasion of his 70th birthday

Let $\Omega \subset \mathbb{C}^n$ be a bounded domain with the hyperconvexity index $\alpha(\Omega) > 0$. Let ϱ be the relative extremal function of a fixed closed ball in Ω, and set $\mu := |\varrho|(1+|\log|\varrho||)^{-1}$ and $\nu := |\varrho|(1+|\log|\varrho||)^n$. We obtain the following estimates for the Bergman kernel. (1) For every $0 < \alpha < \alpha(\Omega)$ and $2 \le p < 2 + 2\alpha(\Omega)/(2n - \alpha(\Omega))$, there exists a constant $C > 0$ such that $\int_\Omega |K_\Omega(\cdot, w)/\sqrt{K_\Omega(w)}|^p \le C|\mu(w)|^{-(p-2)n/\alpha}$ for all $w \in \Omega$. (2) For every $0 < r < 1$, there exists a constant $C > 0$ such that $|K_\Omega(z, w)|^2/(K_\Omega(z)K_\Omega(w)) \le C(\min\{\nu(z)/\mu(w), \nu(w)/\mu(z)\})^r$ for all $z, w \in \Omega$. Various applications of these estimates are given.

1. Introduction

A domain $\Omega \subset \mathbb{C}^n$ is called *hyperconvex* if there exists a negative continuous plurisubharmonic (psh) function ρ on Ω such that $\{\rho < c\} \Subset \Omega$ for any $c < 0$. The class of hyperconvex domains is very wide; e.g., every bounded pseudoconvex domain with Lipschitz boundary is hyperconvex [Demailly 1987]. Although hyperconvex domains already admit a rich function theory (see, e.g., [Ohsawa 1993; Błocki and Pflug 1998; Herbort 1999; Poletsky and Stessin 2008]), it is not enough to get quantitative results unless one imposes certain growth conditions on the bounded exhaustion function ρ (compare [Berndtsson and Charpentier 2000; Błocki 2005; Diederich and Ohsawa 1995]).

A meaningful condition is $-\rho \le C\delta^\alpha$ for some constants $\alpha, C > 0$, where δ denotes the boundary distance. Let $\alpha(\Omega)$ be the supremum of all α. We call it the *hyperconvexity index* of Ω. From the fundamental work of Diederich and Fornaess [1977], we know that if Ω is a bounded pseudoconvex domain with C^2-boundary then there exists a continuous negative psh function ρ on Ω such that $C^{-1}\delta^\eta \le -\rho \le C\delta^\eta$ for some constants $\eta, C > 0$. The supremum $\eta(\Omega)$ of all η is called the *Diederich–Fornaess index* of Ω (see, e.g., [Adachi and Brinkschulte 2015; Fu and Shaw 2016; Harrington 2008]). Clearly, $\alpha(\Omega) \ge \eta(\Omega)$. Recently, Harrington [2008] showed that if Ω is a bounded pseudoconvex domain with Lipschitz boundary then $\eta(\Omega) > 0$.

On the other hand, there are plenty of domains with very irregular boundaries such that $\alpha(\Omega) > 0$, while it is difficult to verify $\eta(\Omega) > 0$. For instance, Koebe's distortion theorem implies $\alpha(\Omega) \ge \frac{1}{2}$ if $\Omega \subsetneq \mathbb{C}$ is a simply connected domain [Carleson and Gamelin 1993, Chapter 1, Theorem 4.4]. Recently, Carleson and Totik [2004] and Totik [2006] obtained various Wiener-type criteria for planar domains with positive

Chen was supported by Grant IDH1411041 from Fudan University.
MSC2010: primary 32A25; secondary 32U35.
Keywords: Bergman kernel, hyperconvexity index.

hyperconvexity indices. In particular, if $\partial\Omega$ is uniformly perfect in the sense of Pommerenke [1979], then $\alpha(\Omega) > 0$ [Carleson and Totik 2004, Theorem 1.7]. Moreover, for domains like $\Omega = \mathbb{C} \setminus E$, where E is a compact set in $\mathbb{R}$ (e.g., Cantor-type sets), the connection between the metric properties of E and the precise value of $\alpha(\Omega)$ (especially the optimal case $\alpha(\Omega) = \frac{1}{2}$) was studied in detail in [Carleson and Totik 2004; Totik 2006]. In the Appendix of this paper, we will provide more examples of higher-dimensional domains with positive hyperconvexity indices. The Teichmüller space of a compact Riemann surface with genus ≥ 2 which is boundedly embedded in $\mathbb{C}^{3g-3}$ probably has a positive hyperconvexity index.

For a domain $\Omega \subset \mathbb{C}^n$, let ϱ be the *relative extremal function* of a (fixed) closed ball $\bar{B} \subset \Omega$; i.e.,

$$\varrho(z) := \varrho_{\bar{B}}(z) := \sup\{u(z) : u \in \mathrm{PSH}^-(\Omega),\ u|_{\bar{B}} \leq -1\},$$

where $\mathrm{PSH}^-(\Omega)$ denotes the set of negative psh functions on Ω. It is known that ϱ is continuous on $\overline{\Omega}$ if Ω is a bounded hyperconvex domain [Błocki 2002, Proposition 3.1.3(vii)]. Furthermore, it is easy to show that if $\alpha(\Omega) > 0$ then for every $0 < \alpha < \alpha(\Omega)$ there exists a constant $C > 0$ such that $-\varrho \leq C\delta^{\alpha}$.

The goal of this paper is to present some off-diagonal estimates of the Bergman kernel on domains with positive hyperconvexity indices, in terms of ϱ. Usually, off-diagonal behavior of the Bergman kernel is more sensitive to the geometry of a domain than on-diagonal behavior (compare to [Barrett 1992]).

Let $K_\Omega(z, w)$ be the Bergman kernel of Ω. It is well-known that $K_\Omega(\cdot\,, w) \in L^2(\Omega)$ for all $w \in \Omega$. Thus, it is natural to ask the following:

Problem. For which Ω and $p > 2$ does one have $K_\Omega(\cdot\,, w) \in L^p(\Omega)$ for all $w \in \Omega$?

For the sake of convenience, we set

$$\beta(\Omega) = \sup\{\beta \geq 2 : K_\Omega(\cdot\,, w) \in L^\beta(\Omega) \text{ for all } w \in \Omega\}.$$

We call it the *integrability index* of the Bergman kernel. From the well-known works of Kerzman, Catlin and Bell, we know that $\beta(\Omega) = \infty$ if Ω is a bounded pseudoconvex domain of finite D'Angelo type. On the other hand, it is not difficult to see from the work of Barrett [1992] that there exist unbounded Diederich–Fornaess worm domains with $\beta(\Omega)$ arbitrarily close to 2 (see, e.g., [Krantz and Peloso 2008, Lemma 7.5]). Thus, it is meaningful to show the following:

Theorem 1.1. *If $\Omega \subset \mathbb{C}^n$ is pseudoconvex, then $\beta(\Omega) \geq 2 + 2\alpha(\Omega)/(2n - \alpha(\Omega))$. Furthermore, if Ω is a bounded domain with $\alpha(\Omega) > 0$, then for every $0 < \alpha < \alpha(\Omega)$ and $2 \leq p < 2 + 2\alpha(\Omega)/(2n - \alpha(\Omega))$, there exists a constant $C > 0$ such that*

$$\int_\Omega \left|K_\Omega(\cdot\,, w)/\sqrt{K_\Omega(w)}\right|^p \leq C|\mu(w)|^{-(p-2)n/\alpha}, \quad w \in \Omega, \tag{1-1}$$

where $K_\Omega(w) = K_\Omega(w, w)$ and $\mu := |\varrho|(1 + |\log|\varrho||)^{-1}$.

The lower bound for $\beta(\Omega)$ can be improved substantially when $n = 1$:

Theorem 1.2. *If Ω is a domain in $\mathbb{C}$, then $\beta(\Omega) \geq 2 + \alpha(\Omega)/(1 - \alpha(\Omega))$.*

In particular, we obtain the known fact that if $\Omega \subsetneq \mathbb{C}$ is a simply connected domain then $\beta(\Omega) \geq 3$. A famous conjecture of Brennan [1978] suggests that the bound may be improved to $\beta(\Omega) \geq 4$; an equivalent statement is that, if $f : \Omega \to \mathbb{D}$ is a conformal mapping where $\mathbb{D}$ is the unit disc, then $f' \in L^p(\Omega)$ for all $p < 4$. There has been extensive research on this conjecture (see [Bertilsson 1998; Carleson and Jones 1992; Carleson and Makarov 1994; Pommerenke 1992], etc.).

Nevertheless, Theorem 1.2 is best understood in view of the following:

Proposition 1.3. *Let $E \subset \mathbb{C}$ be a compact set satisfying* $\operatorname{Cap}(E) > 0$ *and* $\dim_H(E) < 1$, *where* Cap *and* $\dim_H$ *denote the logarithmic capacity and the Hausdorff dimension, respectively. Set $\Omega := \mathbb{C} \setminus E$. Then $\beta(\Omega) \le 2 + \dim_H(E)/(1 - \dim_H(E))$.*

Example. There exists a Cantor-type set E with $\dim_H(E) = 0$ and $\operatorname{Cap}(E) > 0$ [Carleson 1967, §4, Theorem 5]. Thus, $\beta(\mathbb{C} \setminus E) = 2$ in view of Proposition 1.3.

Example. Andrievskii [2005] constructed a compact set $E \subset \mathbb{R}$ with $\dim_H(E) = \frac{1}{2}$ and $\alpha(\mathbb{C} \setminus E) = \frac{1}{2}$. It follows from Theorem 1.2 and Proposition 1.3 that $\beta(\mathbb{C} \setminus E) = 3$.

Problem. Is there a bounded domain $\Omega \subset \mathbb{C}$ with $\beta(\Omega) = 2$?

Theorems 1.1 and 1.2 shed some light on the study of the Bergman space

$$A^p(\Omega) = \left\{ f \in \mathbb{O}(\Omega) : \int_\Omega |f|^p < \infty \right\}$$

for domains with positive hyperconvexity indices. For instance, we can show that $A^p(\Omega) \cap A^2(\Omega)$ lies dense in $A^2(\Omega)$ for suitable $p > 2$ and the reproducing property of $K_\Omega(z, w)$ holds in $A^p(\Omega)$ for suitable $p < 2$ (see Section 4). A related problem is to study whether the Bergman projection can be extended to a bounded projection $L^p(\Omega) \to A^p(\Omega)$ for all p in some nonempty open interval around 2. For flat Hartogs triangles, a complete answer was recently given by Edholm and McNeal [2016]. For more information on this matter, we refer the reader to the review article of Lanzani [2015] and the references therein.

Set

$$K_{\Omega,p}(z) := \sup\{|f(z)| : f \in A^p(\Omega),\ \|f\|_{L^p(\Omega)} \le 1\}.$$

Using $f := (K_\Omega(\cdot\,, z)/\sqrt{K_\Omega(z)})/\|K_\Omega(\cdot\,, z)/\sqrt{K_\Omega(z)}\|_{L^p(\Omega)}$ as a candidate, we conclude from estimate (1-1):

Corollary 1.4. *Let $\Omega \subset \mathbb{C}^n$ be a bounded domain with $\alpha(\Omega) > 0$. For every $p < 2 + 2\alpha(\Omega)/(2n - \alpha(\Omega))$,*

$$K_{\Omega,p}(z) \ge C_{\alpha,p} \sqrt{K_\Omega(z)} |\mu(z)|^{(p-2)n/(p\alpha)}.$$

Remark. If Ω is a bounded pseudoconvex domain with C^2-boundary, then $K_\Omega(z) \ge C\delta(z)^{-2}$ in view of the Ohsawa–Takegoshi extension theorem [1987]. On the other hand, Hopf's lemma implies $|\varrho| \ge C\delta$. Thus,

$$K_{\Omega,p}(z) \ge C_{\alpha,p} \delta(z)^{-(1-(p-2)n/(p\alpha))} |\log \delta(z)|^{-(p-2)n/(p\alpha)}$$

as $z \to \partial\Omega$. Notice also that $(p-2)n/(p\alpha) < \frac{1}{2}$ if and only if $p < 2 + 2\alpha(\Omega)/(2n - \alpha(\Omega))$.

We would like to mention an interesting connection between the problem on page 1430 and the regularity problem of biholomorphic maps. The starting point is the following:

Theorem 1.5 [Lempert 1986, Theorem 6.2]. *Let $\Omega_1 \subset \mathbb{C}^n$ be a bounded domain with C^2-boundary such that its Bergman projection P_{Ω_1} maps $C_0^\infty(\Omega_1)$ into $L^p(\Omega_1)$ for some $p > 2$. Let $\Omega_2 \subset \mathbb{C}^n$ be a bounded domain with real-analytic boundary. Then any biholomorphic map $F : \Omega_1 \to \Omega_2$ extends to a Hölder-continuous map $\overline{\Omega}_1 \to \overline{\Omega}_2$.*

Notice that if Ω is a domain with $\int_\Omega |K_\Omega(\cdot\,, w)|^p$ locally uniformly bounded in w for some $p \ge 1$, then for any $\phi \in C_0^\infty(\Omega)$,

$$|P_\Omega(\phi)(z)|^p \le \int_{\zeta \in \operatorname{supp}\phi} |K_\Omega(\zeta, z)|^p \|\phi\|^p_{L^q(\Omega)}, \quad 1/p + 1/q = 1,$$

so that

$$\int_{z\in\Omega} |P_\Omega(\phi)(z)|^p \le \|\phi\|^p_{L^q(\Omega)} \int_{\zeta \in \operatorname{supp}\phi} \int_{z\in\Omega} |K_\Omega(z, \zeta)|^p < \infty, \tag{1-2}$$

i.e., P_Ω maps $C_0^\infty(\Omega)$ into $L^p(\Omega)$. Thus, we have:

Corollary 1.6. *Let $\Omega_1 \subset \mathbb{C}^n$ be a bounded domain with C^2-boundary such that the integral $\int_\Omega |K_\Omega(\cdot\,, w)|^p$ is locally uniformly bounded in w for some $p > 2$. Let $\Omega_2 \subset \mathbb{C}^n$ be a bounded domain with real-analytic boundary. Then any biholomorphic map $F : \Omega_1 \to \Omega_2$ extends to a Hölder-continuous map $\overline{\Omega}_1 \to \overline{\Omega}_2$.*

In particular, it follows from Corollary 1.6 and Theorem 1.1 that any biholomorphic map between a bounded *pseudoconvex* domain with C^2-boundary and a bounded domain with real-analytic boundary extends to a Hölder-continuous map between their closures, which was first proved in [Diederich and Fornaess 1979]. On the other hand, Barrett [1984] constructed a *nonpseudoconvex* bounded smooth domain $\Omega \subset \mathbb{C}^2$ such that P_Ω fails to map $C_0^\infty(\Omega)$ into $L^p(\Omega)$ for any $p > 2$ so that $\int_\Omega |K_\Omega(\cdot\,, w)|^p$ can not be locally uniformly bounded in w. However, it is still expected that if Ω is a bounded domain with *real-analytic* boundary then there exists $p > 2$ such that $\int_\Omega |K_\Omega(\cdot\,, w)|^p$ is locally uniformly bounded in w.

With the help of an elegant technique due to Błocki [2005] (see also [Herbort 2000] for prior related techniques) on estimating the pluricomplex Green function, we may prove the following:

Theorem 1.7. *Let $\Omega \subset \mathbb{C}^n$ be a bounded domain with $\alpha(\Omega) > 0$. For every $0 < r < 1$, there exists a constant $C > 0$ such that*

$$\mathcal{B}_\Omega(z, w) := \frac{|K_\Omega(z, w)|^2}{K_\Omega(z) K_\Omega(w)} \le C\left(\min\left\{\frac{\nu(z)}{\mu(w)}, \frac{\nu(w)}{\mu(z)}\right\}\right)^r, \quad z, w \in \Omega, \tag{1-3}$$

where $\mu := |\varrho|/(1 + |\log|\varrho||)$ and $\nu := |\varrho|(1 + |\log|\varrho||)^n$.

We call $\mathcal{B}_\Omega(z, w)$ the normalized Bergman kernel of Ω. There is a long list of papers about pointwise estimates of the *weighted* normalized Bergman kernel $\mathcal{B}_{\Omega,\varphi}(z, w) := |K_{\Omega,\varphi}(z, w)|^2/(K_{\Omega,\varphi}(z)K_{\Omega,\varphi}(w))$ when Ω is $\mathbb{C}^n$ or a compact algebraic manifold, after a seminal paper of Christ [1991] (see [Delin 1998; Lindholm 2001; Ma and Marinescu 2007; Christ 2013; Zelditch 2016], etc.). Quantitative measurements of positivity of $i\partial\bar\partial\varphi$ play a crucial role in these works.

The basic difference between $\mathcal{B}_\Omega(z, w)$ and $\mathcal{B}_{\Omega,\varphi}(z, w)$ is that the former is always a *biholomorphic invariant*. Skwarczyński [1980] showed that

$$d_S(z, w) := \left(1 - \sqrt{\mathcal{B}_\Omega(z, w)}\right)^{1/2}$$

gives an invariant distance on a bounded domain Ω. The relationship between d_S and the Bergman distance d_B is

$$d_B(z, w) \geq \sqrt{2} d_S(z, w) \tag{1-4}$$

(see, e.g., [Jarnicki and Pflug 1993, Corollary 6.4.7]). By Theorem 1.7 and (1-4), we may prove the following:

Corollary 1.8. *If Ω is a bounded domain with $\alpha(\Omega) > 0$, then for fixed $z_0 \in \Omega$, there exists a constant $C > 0$ such that*

$$d_B(z_0, z) \geq C \frac{|\log \delta(z)|}{\log|\log \delta(z)|}, \tag{1-5}$$

provided z sufficiently close to $\partial\Omega$.

Błocki [2005] first proved (1-5) for any bounded domain which admits a continuous negative psh function ρ with $C_1\delta^\alpha \leq -\rho \leq C_2\delta^\alpha$ for some constants $C_1, C_2, \alpha > 0$ (e.g., Ω is a pseudoconvex domain with Lipschitz boundary [Harrington 2008]). Diederich and Ohsawa [1995] proved earlier that the weaker inequality

$$d_B(z_0, z) \geq C \log|\log \delta(z)|$$

holds for more general bounded domains admitting a continuous negative psh function ρ with $C_1\delta^{1/\alpha} \leq -\rho \leq C_2\delta^\alpha$ for some constants $C_1, C_2, \alpha > 0$.

In order to study isometric embedding of Kähler manifolds, Calabi [1953] introduced the notion "diastasis". Marcel Berger [1996] wrote, "It seems to me that the notion of diastasis should make a comeback [...]. For example, it would be interesting to compare the diastasis with the various types of Kobayashi metrics (when they exist)."

Notice that the diastasis $D_B(z, w)$ with respect to the Bergman metric is $-\log \mathcal{B}_\Omega(z, w)$.

Corollary 1.9. *If Ω is a bounded domain with $\alpha(\Omega) > 0$, then for fixed $z_0 \in \Omega$, there exists a constant $C > 0$ such that*

$$D_B(z_0, z) \geq C d_K(z_0, z), \tag{1-6}$$

where d_K denotes the Kobayashi distance.

Problem. Does one have $d_B(z_0, z) \geq C d_K(z_0, z)$ for bounded domains with $\alpha(\Omega) > 0$?

A domain $\Omega \subset \mathbb{C}^n$ is called *weighted circular* if there exists an n-tuple $(a_1, \dots, a_n)$ of positive numbers such that $z \in \Omega$ implies $(e^{ia_1\theta} z_1, \dots, e^{ia_n\theta} z_n) \in \Omega$ for any $\theta \in \mathbb{R}$. As a final consequence of Theorem 1.7, we obtain:

Corollary 1.10. *Let $\Omega_1 \subset \mathbb{C}^n$ be a bounded domain with $\alpha(\Omega_1) > 0$. Let $\Omega_2 \subset \mathbb{C}^n$ be a bounded weighted circular domain which contains the origin. Let $0 < \alpha < \alpha(\Omega_1)$ be given. Then for any biholomorphic map $F : \Omega_1 \to \Omega_2$, there is a constant $C > 0$ such that*

$$\delta_2(F(z)) \leq C \delta_1(z)^{\alpha/(2n)}, \quad z \in \Omega_1. \tag{1-7}$$

Here δ_1 and δ_2 denote the boundary distances of Ω_1 and Ω_2, respectively.

Remark. Inequalities like (1-7) are crucial in the study of the regularity problem of biholomorphic maps (see, e.g., [Diederich and Fornaess 1979; Lempert 1986]).

2. L^2 boundary decay estimates of the Bergman kernel

Proposition 2.1. *Let $\Omega \subset \mathbb{C}^n$ be a pseudoconvex domain. Let ρ be a negative continuous psh function on Ω. Set*

$$\Omega_t = \{z \in \Omega : -\rho(z) > t\}, \quad t > 0.$$

Let $a > 0$ be given. For every $0 < r < 1$, there exist constants $\varepsilon_r, C_r > 0$ such that

$$\int_{-\rho\le\varepsilon} |K_\Omega(\cdot\,, w)|^2 \le C_r K_{\Omega_a}(w)(\varepsilon/a)^r \tag{2-1}$$

for all $w \in \Omega_a$ and $\varepsilon \le \varepsilon_r a$.

The proof of the proposition is essentially the same as for Proposition 6.1 in [Chen 2016]. For the sake of completeness, we include a proof here. The key ingredient is the following weighted estimate of the L^2-minimal solution of the $\bar\partial$-equation due to Berndtsson.

Theorem 2.2 [Chen 2016, Corollary 2.3]. *Let Ω be a bounded pseudoconvex domain in $\mathbb{C}^n$ and $\varphi \in \mathrm{PSH}(\Omega)$. Let ψ be a continuous psh function on Ω which satisfies $ri\partial\bar\partial\psi \ge i\partial\psi\wedge\bar\partial\psi$ as currents for some $0 < r < 1$. Suppose v is a $\bar\partial$-closed $(0, 1)$-form on Ω such that $\int_\Omega |v|^2 e^{-\varphi} < \infty$. Then the $L^2(\Omega, \varphi)$-minimal solution of $\bar\partial u = v$ satisfies*

$$\int_\Omega |u|^2 e^{-\psi-\varphi} \le \frac{1}{1-r}\int_\Omega |v|^2_{i\partial\bar\partial\psi} e^{-\psi-\varphi}. \tag{2-2}$$

Here $|v|^2_{i\partial\bar\partial\psi}$ should be understood as the infimum of nonnegative locally bounded functions H satisfying $i\bar v \wedge v \le Hi\partial\bar\partial\psi$ as currents.

Proof of Proposition 2.1. Assume first that Ω is bounded. Let $\kappa : \mathbb{R} \to [0, 1]$ be a smooth cut-off function such that $\kappa|_{(-\infty,1]} = 1$, $\kappa|_{[3/2,\infty)} = 0$ and $|\kappa'| \le 2$. We then have

$$\int_{-\rho\le\varepsilon} |K_\Omega(\cdot\,, w)|^2 \le \int_\Omega \kappa(-\rho/\varepsilon)|K_\Omega(\cdot\,, w)|^2.$$

By the well-known property of the Bergman projection, we obtain

$$\int_\Omega \kappa(-\rho/\varepsilon)K_\Omega(\cdot\,, w)\cdot\overline{K_\Omega(\cdot\,,\zeta)} = \kappa(-\rho(\zeta)/\varepsilon)K_\Omega(\zeta, w) - u(\zeta), \quad \zeta \in \Omega,$$

where u is the $L^2(\Omega)$-minimal solution of the equation

$$\bar\partial u = \bar\partial(\kappa(-\rho/\varepsilon)K_\Omega(\cdot\,, w)) =: v.$$

Since $\kappa(-\rho(w)/\varepsilon) = 0$ provided $\frac{3}{2}\varepsilon \le a$ (i.e., $\varepsilon \le 2a/3$),

$$\int_{-\rho\le\varepsilon} |K_\Omega(\cdot\,, w)|^2 \le -u(w). \tag{2-3}$$

Set

$$\psi = -r\log(-\rho), \quad 0 < r < 1.$$

Clearly, ψ is psh and satisfies $ri\partial\bar\partial\psi \ge i\partial\psi\wedge\bar\partial\psi$ so that

$$i\bar v\wedge v \le C_0 r^{-1}|\kappa'(-\rho/\varepsilon)|^2|K_\Omega(\cdot,w)|^2 i\partial\bar\partial\psi$$

for some numerical constant $C_0 > 0$. Thus, by Theorem 2.2,

$$\begin{aligned}\int_\Omega |u|^2e^{-\psi} &\le C_r\int_{\varepsilon\le-\rho\le(3/2)\varepsilon}|K_\Omega(\cdot,w)|^2e^{-\psi}\\ &\le C_r\varepsilon^r\int_{-\rho\le(3/2)\varepsilon}|K_\Omega(\cdot,w)|^2.\end{aligned}$$

Since $e^{-\psi}\ge a^r$ on Ω_a and u is holomorphic there, it follows that

$$\begin{aligned}|u(w)|^2 &\le K_{\Omega_a}(w)\int_{\Omega_a}|u|^2\\ &\le K_{\Omega_a}(w)a^{-r}\int_\Omega |u|^2e^{-\psi}\\ &\le C_rK_{\Omega_a}(w)(\varepsilon/a)^r\int_{-\rho\le(3/2)\varepsilon}|K_\Omega(\cdot,w)|^2.\end{aligned}$$

Thus, by (2-3),

$$\int_{-\rho\le\varepsilon}|K_\Omega(\cdot,w)|^2 \le C_rK_{\Omega_a}(w)^{1/2}(\varepsilon/a)^{r/2}\biggl(\int_{-\rho\le(3/2)\varepsilon}|K_\Omega(\cdot,w)|^2\biggr)^{1/2}.$$

Notice that

$$\int_{-\rho\le(3/2)\varepsilon}|K_\Omega(\cdot,w)|^2 \le \int_\Omega|K_\Omega(\cdot,w)|^2 = K_\Omega(w)\le K_{\Omega_a}(w)$$

provided $\frac{3}{2}\varepsilon\le a$. Thus,

$$\int_{-\rho\le\varepsilon}|K_\Omega(\cdot,w)|^2\le C_rK_{\Omega_a}(w)(\varepsilon/a)^{r/2}.$$

Replacing ε by $\frac{3}{2}\varepsilon$ in the argument above, we obtain

$$\int_{-\rho\le(3/2)\varepsilon}|K_\Omega(\cdot,w)|^2\le C_rK_{\Omega_a}(w)(3/2)^{r/2}(\varepsilon/a)^{r/2}$$

provided $(\frac{3}{2})^2\varepsilon\le a$. Thus, we may improve the upper bound by

$$\int_{-\rho\le\varepsilon}|K_\Omega(\cdot,w)|^2\le C_rK_{\Omega_a}(w)(\varepsilon/a)^{r/2+r/4}.$$

By induction, we conclude that, for every $k\in\mathbb{Z}^+$,

$$\int_{-\rho\le\varepsilon}|K_\Omega(\cdot,w)|^2\le C_{r,k}K_{\Omega_a}(w)(\varepsilon/a)^{r/2+r/4+\cdots+r/2^k}$$

provided $(\frac{3}{2})^k\varepsilon \le a$. Since $r/2+r/4+\cdots+r/2^k \to 1$ as $k\to\infty$ and $r\to 1$, we get the desired estimate under the assumption that Ω is bounded.

In general, Ω may be exhausted by an increasing sequence $\{\Omega_j\}$ of bounded pseudoconvex domains. From the argument above, we know that

$$\int_{\Omega_j\cap\{-\rho\le\varepsilon\}} |K_{\Omega_j}(\cdot\,,w)|^2 \le C_r K_{\Omega_j\cap\Omega_a}(w)(\varepsilon/a)^r$$

holds for all $j \gg 1$. Since $\Omega_j \uparrow \Omega$, it is well-known that $K_{\Omega_j}(\cdot\,,w)\to K_\Omega(\cdot\,,w)$ locally uniformly in Ω and $K_{\Omega_j\cap\Omega_a}(w)\to K_{\Omega_a}(w)$. It follows from Fatou's lemma that

$$\begin{aligned}\int_{-\rho\le\varepsilon} |K_\Omega(\cdot\,,w)|^2 &= \liminf_{j\to\infty}\int_{\Omega_j\cap\{-\rho\le\varepsilon\}} |K_{\Omega_j}(\cdot\,,w)|^2\\ &\le C_r K_{\Omega_a}(w)(\varepsilon/a)^r.\end{aligned}$$

□

Remark. One of the referees kindly suggested an alternative proof as follows. Berndtsson and Charpentier [2000] showed that, if $\int_\Omega |f|^2|\rho|^{-r}<\infty$ for some $0<r<1$, then

$$\int_\Omega |P_\Omega(f)|^2|\rho|^{-r} \le C_r\int_\Omega |f|^2|\rho|^{-r}<\infty$$

where $P_\Omega(f)(z):=\int_\Omega K_\Omega(z,\cdot\,)f(\cdot\,)$ is the Bergman projection. If one applies $f=\chi_{\Omega_a}K_{\Omega_a}(\cdot\,,w)$ where χ_{Ω_a} denotes the characteristic function on Ω_a, then $K_\Omega(z,w)=P_\Omega(f)(z)$ and

$$\int_\Omega |K_\Omega(\cdot\,,w)|^2|\rho|^{-r} \le C_r\int_{\Omega_a} |K_{\Omega_a}(\cdot\,,w)|^2|\rho|^{-r},$$

from which the estimate (2-1) immediately follows.

Let ϱ be the relative extremal function of a (fixed) closed ball $\bar B\subset\Omega$. We have:

Proposition 2.3. *Let $\Omega\subset\mathbb{C}^n$ be a bounded domain with $\alpha(\Omega)>0$. For every $0<r<1$, there exist constants $\varepsilon_r, C_r>0$ such that*

$$\int_{-\varrho\le\varepsilon} |K_\Omega(\cdot\,,w)|^2/K_\Omega(w) \le C_r(\varepsilon/\mu(w))^r \tag{2-4}$$

for all $\varepsilon\le\varepsilon_r\mu(w)$, where $\mu=|\varrho|(1+|\log|\varrho||)^{-1}$.

In order to prove this proposition, we need an elementary estimate of the pluricomplex Green function. Recall that the *pluricomplex Green function* $g_\Omega(z,w)$ of a domain $\Omega\subset\mathbb{C}^n$ is defined as

$$g_\Omega(z,w)=\sup\{u(z): u\in \mathrm{PSH}^-(\Omega),\ u(z)\le\log|z-w|+O(1)\text{ near }w\}.$$

We first show the following quasi-Hölder-continuity of ϱ.

Lemma 2.4. *Let $\Omega\subset\mathbb{C}^n$ be a bounded domain with $\alpha(\Omega)>0$. For every $r>1$ and $0<\alpha<\alpha(\Omega)$, there exists a constant $C>0$ such that*

$$\varrho(z_2)\ge r\varrho(z_1)-C|z_1-z_2|^\alpha,\quad z_1,z_2\in\Omega. \tag{2-5}$$

Proof. Choose $\rho \in C(\Omega) \cap \mathrm{PSH}^-(\Omega)$ with $-\rho \le C_\alpha \delta^\alpha$. Clearly

$$\varrho(z) \ge \frac{\rho(z)}{\inf_{\bar{B}} |\rho|} \ge -C_\alpha \delta^\alpha.$$

To get (2-5), we employ a well-known technique of Walsh [1968] as follows. Set $\varepsilon := |z_1 - z_2|$, $\Omega' := \Omega - (z_1 - z_2)$ and

$$u(z) = \begin{cases} \varrho(z) & \text{if } z \in \Omega \setminus \Omega', \\ \max\{\varrho(z), r\varrho(z + z_1 - z_2) - C\varepsilon^\alpha\} & \text{if } z \in \Omega \cap \Omega'. \end{cases}$$

We claim that $u \in \mathrm{PSH}^-(\Omega)$ provided $C \gg 1$. Indeed, if $z \in \Omega \cap \partial\Omega'$, then $\delta(z) \le \varepsilon$ so that

$$\varrho(z) \ge -C_\alpha \delta(z)^\alpha \ge -C_\alpha \varepsilon^\alpha \ge r\varrho(z + z_1 - z_2) - C_\alpha \varepsilon^\alpha.$$

Moreover, if $\varepsilon \le \varepsilon_r \ll 1$, then $\varrho(z + z_1 - z_2) \le -1/r$ for $z \in \bar{B}$ since ϱ is continuous on $\overline{\Omega}$. Thus, $u|_{\bar{B}} \le -1$. Since $z_2 = z_1 - (z_1 - z_2) \in \Omega \cap \Omega'$, it follows that

$$\varrho(z_2) \ge u(z_2) \ge r\varrho(z_1) - C_\alpha \varepsilon^\alpha.$$

If $\varepsilon = |z_1 - z_2| > \varepsilon_r$, then (2-5) trivially holds. □

Remark. It is not known whether ϱ is Hölder-continuous on $\overline{\Omega}$. The answer is positive if $n = 1$ [Carleson and Gamelin 1993, p. 138].

Proposition 2.5. *Let $\Omega \subset \mathbb{C}^n$ be a bounded domain with $\alpha(\Omega) > 0$. There exists a constant $C \gg 1$ such that*

$$\{g_\Omega(\cdot\,, w) < -1\} \subset \{\varrho < -C^{-1}\mu(w)\}, \quad w \in \Omega. \tag{2-6}$$

Proof. Fix $0 < \alpha < \alpha(\Omega)$. We have $-\varrho \le C_\alpha \delta^\alpha$ for some constant $C_\alpha > 0$. Clearly, it suffices to consider the case when $|\varrho(w)| \le \frac{1}{2}$. Applying Lemma 2.4 with $r = \frac{3}{2}$, we see that if $\varrho(z) = \varrho(w)/2$ then

$$C_1 |z - w|^\alpha \ge \tfrac{3}{2}\varrho(z) - \varrho(w) = -\tfrac{1}{4}\varrho(w)$$

so that

$$\log \frac{|z - w|}{R} \ge \frac{1}{\alpha} \log|\varrho(w)|/(4C_1) - \log R \ge C_2 \log|\varrho(w)|$$

for some constant $C_2 \gg 1$. It follows that

$$\psi(z) := \begin{cases} \log|z - w|/R & \text{if } \varrho(z) \le \varrho(w)/2, \\ \max\{\log|z - w|/R, 2C_2(\varrho(w)^{-1} \log|\varrho(w)|)\varrho(z)\} & \text{otherwise} \end{cases}$$

is a well-defined negative psh function on Ω with a logarithmic pole at w, and if $\varrho(z) \ge \varrho(w)/2$, then

$$g_\Omega(z, w) \ge \psi(z) \ge 2C_2(\varrho(w)^{-1} \log|\varrho(w)|)\varrho(z). \tag{2-7}$$

Thus,

$$\{g_\Omega(\cdot\,, w) < -1\} \cap \{\varrho \ge \varrho(w)/2\} \subset \{\varrho < -C^{-1}\mu(w)\}$$

provided $C \gg 1$. Since $\{\varrho < \varrho(w)/2\} \subset \{\varrho < -C^{-1}\mu(w)\}$ if $C \gg 1$, we conclude the proof. □

Proof of Proposition 2.3. Set $A_w := \{g_\Omega(\cdot\,, w) < -1\}$. It is known from [Herbort 1999] or [Chen 1999] that

$$K_{A_w}(w) \le C_n K_\Omega(w). \tag{2-8}$$

By Proposition 2.5,

$$A_w \subset \Omega_{a(w)} := \{\varrho < -a(w)\} \tag{2-9}$$

where $a(w) := C^{-1}\mu(w)$ with $C \gg 1$. If we choose $\rho = \varrho$ in Proposition 2.1, it follows that, for every $\varepsilon \le \varepsilon_r a(w)$,

$$\begin{aligned}\int_{-\varrho\le\varepsilon} |K_\Omega(\cdot\,, w)|^2 &\le C_r K_{\Omega_{a(w)}}(w)(\varepsilon/a(w))^r \\ &\le C_{n,r} K_\Omega(w)(\varepsilon/a(w))^r \end{aligned}\tag{2-10}$$

in view of (2-8) and (2-9). □

3. L^p-integrability of the Bergman kernel

Proof of Theorem 1.1. Without loss of generality, we may assume $\alpha(\Omega) > 0$. For every $0 < \alpha < \alpha(\Omega)$, we may choose $\rho \in \mathrm{PSH}^-(\Omega)$ such that

$$-\rho \le C_\alpha \delta^\alpha$$

for some constant $C_\alpha > 0$. Let S be a compact set in Ω, and let $w \in S$. By virtue of Proposition 2.1, we conclude that, for every $0 < r < 1$,

$$\int_{-\rho\le\varepsilon} |K_\Omega(\cdot\,, w)|^2 \le C\varepsilon^r$$

where $C = C(n, r, \alpha, S) > 0$. Since $\{\delta \le \varepsilon\} \subset \{-\rho \le C_\alpha \varepsilon^\alpha\}$, it follows that

$$\int_{\delta\le\varepsilon} |K_\Omega(\cdot\,, w)|^2 \le C\varepsilon^{r\alpha}.$$

Since $|\delta(\zeta) - \delta(z)| \le |\zeta - z|$, we have $B(z, \delta(z)) \subset \{\delta \le 2\delta(z)\}$. By the mean value inequality, we get

$$|K_\Omega(z, w)|^2 \le C_n \delta(z)^{-2n} \int_{\delta\le 2\delta(z)} |K_\Omega(\cdot\,, w)|^2 \le C\delta(z)^{r\alpha-2n}. \tag{3-1}$$

Thus, for every $\tau > 0$,

$$\begin{aligned}\int_\Omega |K_\Omega(\cdot\,, w)|^{2+\tau} &= \int_{\delta>1/2} |K_\Omega(\cdot\,, w)|^{2+\tau} + \sum_{k=1}^\infty \int_{2^{-k-1}<\delta\le 2^{-k}} |K_\Omega(\cdot\,, w)|^{2+\tau} \\ &\le C2^{n\tau} \int_\Omega |K_\Omega(\cdot\,, w)|^2 + C\sum_{k=1}^\infty 2^{(k+1)\tau(n-r\alpha/2)} \int_{\delta\le 2^{-k}} |K_\Omega(\cdot\,, w)|^2 \\ &\le C + C2^{\tau(n-r\alpha/2)} \sum_{k=1}^\infty 2^{-k(r\alpha+\tau(r\alpha/2-n))} \\ &< \infty \end{aligned}$$

provided $\tau < 2r\alpha/(2n - r\alpha)$. Since r and α can be arbitrarily close to 1 and $\alpha(\Omega)$, respectively, we conclude the proof of the first statement.

Since $\{\delta \le \varepsilon\} \subset \{-\varrho \le C_\alpha \varepsilon^\alpha\}$, it follows from Proposition 2.3 that

$$\int_{\delta \le \varepsilon} |K_\Omega(\cdot\,, w)|^2 / K_\Omega(w) \le C_{\alpha,r}(\varepsilon^\alpha/\mu(w))^r \tag{3-2}$$

provided $\varepsilon^\alpha/\mu(w) \le \varepsilon_r \ll 1$. For every $z \in \Omega$,

$$|K_\Omega(z, w)|^2 / K_\Omega(w) \le K_\Omega(z) \le C_n \delta(z)^{-2n}, \tag{3-3}$$

and if $(2\delta(z))^\alpha \le \varepsilon_r \mu(w)$,

$$\begin{aligned} |K_\Omega(z, w)|^2 &\le C_n \delta(z)^{-2n} \int_{\delta \le 2\delta(z)} |K_\Omega(\cdot\,, w)|^2 \\ &\le C_{\alpha,r} K_\Omega(w) \mu(w)^{-r} \delta(z)^{\alpha r - 2n}. \end{aligned} \tag{3-4}$$

For every $\tau < 2r\alpha/(2n - r\alpha)$, we conclude from (3-3) that

$$\begin{aligned} \int_{2\delta \ge (\varepsilon_r \mu(w))^{1/\alpha}} |K_\Omega(\cdot\,, w)|^{2+\tau} &\le C_n K_\Omega(w)^{\tau/2} \int_{2\delta \ge (\varepsilon_r \mu(w))^{1/\alpha}} |K_\Omega(\cdot\,, w)|^2 \delta^{-n\tau} \\ &\le C_{\alpha,r} \frac{K_\Omega(w)^{\tau/2}}{\mu(w)^{n\tau/\alpha}} \int_\Omega |K_\Omega(\cdot\,, w)|^2 \\ &\le C_{\alpha,r} \frac{K_\Omega(w)^{1+\tau/2}}{\mu(w)^{n\tau/\alpha}}. \end{aligned} \tag{3-5}$$

Now choose $k_w \in \mathbb{Z}^+$ such that $(\varepsilon_r \mu(w))^{1/\alpha} \in (2^{-k_w - 1}, 2^{-k_w}]$ (it suffices to consider the case when $\mu(w)$ is sufficiently small). We then have

$$\begin{aligned} \int_{2\delta < (\varepsilon_r \mu(w))^{1/\alpha}} |K_\Omega(\cdot\,, w)|^{2+\tau} &\le \sum_{k=k_w}^\infty \int_{2^{-k-1} < \delta \le 2^{-k}} |K_\Omega(\cdot\,, w)|^{2+\tau} \\ &\le C_{\alpha,r,\tau} \frac{K_\Omega(w)^{\tau/2}}{\mu(w)^{\tau r/2}} \sum_{k=k_w}^\infty 2^{k\tau(n - r\alpha/2)} \int_{\delta \le 2^{-k}} |K_\Omega(\cdot\,, w)|^2 \quad \text{(by (3-4))} \\ &\le C_{\alpha,r,\tau} \frac{K_\Omega(w)^{1+\tau/2}}{\mu(w)^{r(1+\tau/2)}} \sum_{k=k_w}^\infty 2^{-k(r\alpha + \tau(r\alpha/2 - n))} \quad \text{(by (3-2))} \\ &\le C_{\alpha,r,\tau} \frac{K_\Omega(w)^{1+\tau/2}}{\mu(w)^{r(1+\tau/2)}} \mu(w)^{(r\alpha + \tau(r\alpha/2 - n))/\alpha} \\ &\le C_{\alpha,r,\tau} \frac{K_\Omega(w)^{1+\tau/2}}{\mu(w)^{\tau n/\alpha}}. \end{aligned} \tag{3-6}$$

By (3-5) and (3-6), (1-1) immediately follows. □

Proof of Theorem 1.2. It suffices to use the following lemma instead of (3-1) in the proof of the first statement in Theorem 1.1. □

Lemma 3.1. *Let Ω be a domain in $\mathbb{C}$. For every compact set $S \subset \Omega$ and $\alpha < \alpha(\Omega)$, there exists a constant $C > 0$ such that*

$$|K_\Omega(z, w)| \le C\delta(z)^{\alpha-1}, \quad z \in \Omega, \ w \in S.$$

Proof. Let $g_\Omega(z, w)$ be the (negative) Green function on Ω. Let $\Delta(c, r)$ be the disc with center c and radius r. Fix $w \in S$ and $z \in \Omega$ for a moment. Clearly, it suffices to consider the case when $\delta(z) \le \delta(w)/4$. Since $g_\Omega(\xi, \zeta)$ is harmonic in $\xi \in \Delta(z, \delta(z))$ and $\zeta \in \Delta(w, \delta(w)/2)$, respectively, we conclude from Poisson's formula that

$$\begin{aligned} g_\Omega(\xi, \zeta) = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} & g_\Omega(z + \tfrac{1}{2}\delta(z)e^{i\theta}, w + \tfrac{1}{2}\delta(w)e^{i\vartheta}) \\ & \times \frac{\frac{1}{4}\delta(z)^2 - |\xi - z|^2}{\left|\frac{1}{2}\delta(z)e^{i\theta} - (\xi - z)\right|^2} \frac{\frac{1}{4}\delta(w)^2 - |\zeta - w|^2}{\left|\frac{1}{2}\delta(w)e^{i\vartheta} - (\zeta - w)\right|^2} \, d\theta \, d\vartheta \end{aligned}$$

where $\xi \in \Delta(z, \delta(z)/4)$ and $\zeta \in \Delta(w, \delta(w)/4)$. By the extremal property of g_Ω, it is easy to verify that $-g_\Omega \le C\delta(z)^\alpha$ on $\partial\Delta(z, \delta(z)/2) \times \partial\Delta(w, \delta(w)/2)$. Thus,

$$\left| \frac{\partial^2 g_\Omega(\xi, \zeta)}{\partial\xi \, \partial\bar{\zeta}} \right| \le C\delta(z)^{\alpha-1}.$$

Using the formula $K_\Omega(\xi, \zeta) = \dfrac{2}{\pi} \dfrac{\partial^2 g_\Omega(\xi, \zeta)}{\partial\xi \, \partial\bar{\zeta}}$ from [Schiffer 1946], the assertion immediately follows. □

In order to prove Proposition 1.3, we need the following:

Theorem 3.2 [Carleson 1967, §6, Theorem 1]. *Let $\Omega = \mathbb{C} \setminus E$ where $E \subset \mathbb{C}$ is a compact set. Then*

(1) $A^2(\Omega) \ne \{0\}$ *if and only if* $\mathrm{Cap}(E) > 0$, *and*

(2) $A^p(\Omega) = \{0\}$ *if* $\Lambda_{2-q}(E) < \infty$, $2 < p < \infty$ *and* $1/p + 1/q = 1$. *Here $\Lambda_s(E)$ denotes the s-dimensional Hausdorff measure of E.*

Remark. Let $\Omega \subset \mathbb{C}$ be a domain and E a closed polar set in Ω. It is well-known that E is removable for negative harmonic functions so that $g_{\Omega\setminus E}(z, w) = g_\Omega(z, w)$ for $z, w \in \Omega \setminus E$. Thus, $K_{\Omega\setminus E}(z, w) = K_\Omega(z, w)$ in view of Schiffer's formula. By the reproducing property of the Bergman kernel, we immediately get the known fact that $A^2(\Omega \setminus E) = A^2(\Omega)$.

Proof of Proposition 1.3. Suppose on the contrary $\beta(\Omega) > 2 + \dim_H(E)/(1 - \dim_H(E))$. Fix

$$\beta(\Omega) > p > 2 + \frac{\dim_H(E)}{1 - \dim_H(E)},$$

and let q be the conjugate exponent of p, i.e., $1/p + 1/q = 1$. We then have $K_\Omega(\cdot, w) \in A^p(\Omega)$ for fixed w. Since

$$\dim_H(E) = \sup\{s : \Lambda_s(E) = \infty\}$$

and $2 - q > \dim_H(E)$, it follows that $\Lambda_{2-q}(E) < \infty$ so that $K_\Omega(\cdot, w) = 0$ in view of Theorem 3.2(2). On the other hand, $\mathrm{Cap}(E) > 0$, so $K_\Omega(\cdot, w) \ne 0$ in view of Theorem 3.2(1), which is absurd. □

Theorem 1.2 implies $\beta(\Omega) \to \infty$ as $\alpha(\Omega) \to 1$ for planar domains (notice that $\alpha(\Omega) = 1$ when $\Omega \subset \mathbb{C}$ is convex or $\partial\Omega$ is C^1). It is also known that $\beta(\Omega) = \infty$ if Ω is a bounded smooth convex domain in $\mathbb{C}^n$ [Boas and Straube 1991]. Thus, it is reasonable to make the following:

Conjecture 3.3. *If $\Omega \subset \mathbb{C}^n$ is convex, then $\beta(\Omega) = \infty$.*

4. Applications of L^p-integrability of the Bergman kernel

We first study density of $A^p(\Omega) \cap A^2(\Omega)$ in $A^2(\Omega)$.

Proposition 4.1. *Let Ω be a pseudoconvex domain in $\mathbb{C}^n$. For every $1 \le p < 2 + 2\alpha(\Omega)/(2n - \alpha(\Omega))$, $A^p(\Omega) \cap A^2(\Omega)$ lies dense in $A^2(\Omega)$.*

Proof. Choose a sequence of functions $\chi_j \in C_0^\infty(\Omega)$ such that $0 \le \chi_j \le 1$ and the sequence of sets $\{\chi_j = 1\}$ exhausts Ω. Given $f \in A^2(\Omega)$, we set $f_j = P_\Omega(\chi_j f)$. Clearly, $f_j \in A^p(\Omega) \cap A^2(\Omega)$ in view of Theorem 1.1 and (1-2). Moreover,

$$\|f_j - f\|_{L^2(\Omega)} = \|P_\Omega((\chi_j - 1)f)\|_{L^2(\Omega)} \le \|(\chi_j - 1)f\|_{L^2(\Omega)} \to 0. \qquad \square$$

Similarly, we may prove the following:

Proposition 4.2. *Let Ω be a domain in $\mathbb{C}$. For every $1 \le p < 2 + \alpha(\Omega)/(1 - \alpha(\Omega))$, $A^p(\Omega) \cap A^2(\Omega)$ lies dense in $A^2(\Omega)$.*

Next we study the reproducing property of the Bergman kernel in $A^p(\Omega)$.

Proposition 4.3. *Let Ω be a bounded domain in $\mathbb{C}$ with $\alpha(\Omega) > 0$. If $p > 2 - \alpha(\Omega)$, then $f = P_\Omega(f)$ for all $f \in A^p(\Omega)$.*

Proof. Suppose $f \in A^p(\Omega)$ with $p > 2 - \alpha(\Omega)$. Let q be the conjugate exponent of p. Since $q < 2 + \alpha(\Omega)/(1 - \alpha(\Omega))$, the integral $\int_\Omega f(\cdot)K_\Omega(z, \cdot)$ is well-defined in view of Theorem 1.2. Clearly, it suffices to consider the case $p < 2$. By Theorem 1 of [Hedberg 1972], we may find a sequence $f_j \in \mathcal{O}(\overline{\Omega}) \subset A^2(\Omega) \subset A^p(\Omega)$ such that $\|f_j - f\|_{L^p(\Omega)} \to 0$. It follows that, for every $z \in \Omega$,

$$f(z) = \lim_{j\to\infty} f_j(z) = \lim_{j\to\infty} \int_\Omega f_j(\cdot)K_\Omega(z, \cdot) = \int_\Omega f(\cdot)K_\Omega(z, \cdot)$$

since $K_\Omega(z, \cdot) \in L^q(\Omega)$. $\square$

For a bounded domain $\Omega \subset \mathbb{C}^n$, the *Berezin transform* T_Ω of Ω is defined as

$$T_\Omega(f)(z) = \int_\Omega f(\cdot)\frac{|K_\Omega(\cdot, z)|^2}{K_\Omega(z)}, \quad z \in \Omega, \ f \in L^\infty(\Omega).$$

Clearly, one has $f = T_\Omega(f)$ for all $f \in A^\infty(\Omega)$.

Corollary 4.4. *Let Ω be a bounded domain in $\mathbb{C}$ with $\alpha(\Omega) > 0$. If $p > 2/\alpha(\Omega) - 1$, then $f = T_\Omega(f)$ for all $f \in A^p(\Omega)$.*

Proof. Set $p' = 2p/(p+1)$. It follows from Hölder's inequality that

$$\begin{aligned}\int_\Omega |f K_\Omega(\cdot\,,z)|^{p'} &\le \left(\int_\Omega |f|^{p'/(2-p')}\right)^{2-p'} \left(\int_\Omega |K_\Omega(\cdot\,,z)|^{p'/(p'-1)}\right)^{p'-1} \\ &= \left(\int_\Omega |f|^p\right)^{2-p'} \left(\int_\Omega |K_\Omega(\cdot\,,z)|^{p'/(p'-1)}\right)^{p'-1} \\ &< \infty\end{aligned}$$

since $p' > 2-\alpha(\Omega)$ and $p'/(p'-1) < 2+\alpha(\Omega)/(1-\alpha(\Omega))$. Thus, $h := f K_\Omega(\cdot\,,z)/K_\Omega(z) \in A^{p'}(\Omega)$ for fixed $z \in \Omega$ so that

$$f(z) = h(z) = \int_\Omega h(\cdot) K_\Omega(z,\cdot) = \int_\Omega f(\cdot) \frac{|K_\Omega(\cdot\,,z)|^2}{K_\Omega(z)}. \qquad \square$$

For higher-dimensional cases, we can only prove the following:

Proposition 4.5. *Let Ω be a bounded pseudoconvex domain in $\mathbb{C}^n$. Suppose there exists a negative psh exhaustion function ρ on Ω such that, for suitable constants $C, \alpha > 0$,*

$$|\rho(z) - \rho(w)| \le C|z-w|^\alpha, \quad z, w \in \Omega.$$

For every $p > 4n/(2n+\alpha)$, one has $f = P_\Omega(f)$ for all $f \in A^p(\Omega)$.

Proof. Set $\Omega_t = \{-\rho > t\}$, $t \ge 0$, and $\rho_t := \rho + t$. For every $z \in \Omega_t$, we choose $z^* \in \partial\Omega_t$ such that $|z - z^*| = \delta_t(z) := d(z, \partial\Omega_t)$. We then have

$$|\rho_t(z)| = |\rho_t(z) - \rho_t(z^*)| \le C|z-z^*|^\alpha = C\delta_t(z)^\alpha$$

where C is a constant independent of t. By a similar argument as the proof of Theorem 1.1, we may show that, for fixed $w \in \Omega$,

$$\int_{\Omega_t} |K_{\Omega_t}(\cdot\,, w)|^q \le C = C(q,w) < \infty$$

holds uniformly in $t \ll 1$ for every $q < 2 + 2\alpha/(2n-\alpha)$. Let $2 > p > 4n/(2n+\alpha)$ and $f \in A^p(\Omega)$. Fix $z \in \Omega$ for a moment. For every $t \ll 1$, we have $z \in \Omega_t$ and

$$f(z) = \int_{\Omega_t} f(\cdot) K_{\Omega_t}(z,\cdot). \tag{4-1}$$

Notice that

$$\begin{aligned}&\left|\int_\Omega f(\cdot)K_\Omega(z,\cdot) - \int_{\Omega_t} f(\cdot)K_{\Omega_t}(z,\cdot)\right| \\ &\qquad \le \int_{\Omega_t} |f||K_\Omega(z,\cdot) - K_{\Omega_t}(z,\cdot)| + \int_{\Omega\setminus\Omega_t} |f||K_\Omega(z,\cdot)| \\ &\qquad \le \|f\|_{L^p(\Omega)} \|K_\Omega(z,\cdot) - K_{\Omega_t}(z,\cdot)\|_{L^q(\Omega_t)} + \|f\|_{L^p(\Omega\setminus\Omega_t)} \|K_\Omega(z,\cdot)\|_{L^q(\Omega)}\end{aligned} \tag{4-2}$$

where $1/p+1/q=1$ (which implies $q<2+2\alpha/(2n-\alpha)$). Take $0<\gamma\ll 1$ so that $(q-\gamma)/(1-\gamma/2)<2+2\alpha/(2n-\alpha)$. We then have

$$\begin{aligned}\int_{\Omega_t}&|K_\Omega(z,\cdot)-K_{\Omega_t}(z,\cdot)|^q\\&=\int_{\Omega_t}|K_\Omega(z,\cdot)-K_{\Omega_t}(z,\cdot)|^\gamma|K_\Omega(z,\cdot)-K_{\Omega_t}(z,\cdot)|^{q-\gamma}\\&\le\left(\int_{\Omega_t}|K_\Omega(z,\cdot)-K_{\Omega_t}(z,\cdot)|^2\right)^{\gamma/2}\left(\int_{\Omega_t}|K_\Omega(z,\cdot)-K_{\Omega_t}(z,\cdot)|^{(q-\gamma)/(1-\gamma/2)}\right)^{1-\gamma/2}\end{aligned}$$

in view of Hölder's inequality. Since

$$\begin{aligned}\int_{\Omega_t}|K_\Omega(z,\cdot)-K_{\Omega_t}(z,\cdot)|^2&=\int_{\Omega_t}|K_\Omega(z,\cdot)|^2+\int_{\Omega_t}|K_{\Omega_t}(z,\cdot)|^2-2\,\mathrm{Re}\int_{\Omega_t}K_\Omega(z,\cdot)K_{\Omega_t}(\cdot,z)\\&\le K_{\Omega_t}(z)-K_\Omega(z)\\&\to 0\quad(t\to 0)\end{aligned}$$

and

$$\begin{aligned}\int_{\Omega_t}&|K_\Omega(z,\cdot)-K_{\Omega_t}(z,\cdot)|^{(q-\gamma)/(1-\gamma/2)}\\&\le 2^{(q-\gamma)/(1-\gamma/2)}\left(\int_\Omega|K_\Omega(z,\cdot)|^{(q-\gamma)/(1-\gamma/2)}+\int_{\Omega_t}|K_{\Omega_t}(z,\cdot)|^{(q-\gamma)/(1-\gamma/2)}\right)\\&\le C,\end{aligned}$$

it follows from (4-1) and (4-2) that $f=P_\Omega(f)$. □

Similarly, we have:

Corollary 4.6. *If $p>2n/\alpha$, then $f=T_\Omega(f)$ for all $f\in A^p(\Omega)$.*

5. Estimate of the pluricomplex Green function

The goal of this section is to show the following:

Proposition 5.1. *Let $\Omega\subset\mathbb{C}^n$ be a bounded domain with $\alpha(\Omega)>0$. There exists a constant $C\gg 1$ such that*

$$\{g_\Omega(\cdot,w)<-1\}\subset\{\varrho>-C\nu(w)\},\quad w\in\Omega,\tag{5-1}$$

where $\nu=|\varrho|(1+|\log|\varrho||)^n$.

We will follow the argument of Błocki [2005] with necessary modifications. The key observation is the following:

Lemma 5.2 [Błocki 2005]. *Let $\Omega\subset\mathbb{C}^n$ be a bounded hyperconvex domain. Suppose ζ and w are two points in Ω such that the closed balls $\bar B(\zeta,\varepsilon)$, $\bar B(w,\varepsilon)\subset\mathbb{C}^n$ and $\bar B(\zeta,\varepsilon)\cap\bar B(w,\varepsilon)=\varnothing$. Then there exists $\tilde\zeta\in\bar B(\zeta,\varepsilon)$ such that*

$$|g_\Omega(\tilde\zeta,w)|^n\le n!\,(\log R/\varepsilon)^{n-1}|g_\Omega(w,\zeta)|\tag{5-2}$$

where $R:=\mathrm{diam}(\Omega)$.

For the sake of completeness, we include a proof here, which relies heavily on the following fundamental results.

Theorem 5.3 [Demailly 1987]. *Let Ω be a bounded hyperconvex domain in $\mathbb{C}^n$.*

(1) *For every $w \in \Omega$, one has $(dd^c g_\Omega(\cdot, w))^n = (2\pi)^n \delta_w$, where δ_w denotes the Dirac measure at w.*

(2) *For every $\zeta \in \Omega$ and $\eta > 0$, one has $\int_\Omega (dd^c \max\{g_\Omega(\cdot, \zeta), -\eta\})^n = (2\pi)^n$.*

Theorem 5.4 ([Błocki 1993]; see also [Błocki 2002]). *Let Ω be a bounded domain in $\mathbb{C}^n$. Assume that $u, v \in \mathrm{PSH}^- \cap L^\infty(\Omega)$ are nonpositive psh functions such that $u = 0$ on $\partial\Omega$. Then*

$$\int_\Omega |u|^n (dd^c v)^n \le n!\, \|v\|_\infty^{n-1} \int_\Omega |v| (dd^c u)^n. \tag{5-3}$$

Proof of Lemma 5.2. Let $\eta = \log R/\varepsilon$. Since $g_\Omega(z, \zeta) \ge \log|z - \zeta|/R$, it follows that

$$\{g_\Omega(\cdot, \zeta) = -\eta\} \subset \bar{B}(\zeta, \varepsilon).$$

First applying Theorem 5.4 with $u = \max\{g_\Omega(\cdot, w), -t\}$ and $v = \max\{g_\Omega(\cdot, \zeta), -\eta\}$ and then letting $t \to +\infty$, we obtain

$$\int_\Omega |g_\Omega(\cdot, w)|^n (dd^c \max\{g_\Omega(\cdot, \zeta), -\eta\})^n \le n!\, (2\pi)^n \eta^{n-1} |g_\Omega(w, \zeta)|$$

in view of Theorem 5.3(1). Since $\bar{B}(\zeta, \varepsilon) \cap \bar{B}(w, \varepsilon) = \varnothing$, it follows that $g_\Omega(\cdot, w)$ is continuous on $\bar{B}(\zeta, \varepsilon)$ so that there exists $\tilde{\zeta} \in \bar{B}(\zeta, \varepsilon)$ such that

$$|g_\Omega(\tilde{\zeta}, w)| = \min_{\bar{B}(\zeta, \varepsilon)} |g_\Omega(\cdot, w)|.$$

Since the measure $(dd^c \max\{g_\Omega(\cdot, \zeta), -\eta\})^n$ is supported on $\{g_\Omega(\cdot, \zeta) = -\eta\}$ with total mass $(2\pi)^n$, we immediately get (5-2). □

Proof of Proposition 5.1. Clearly, it suffices to consider the case when w is sufficiently close to $\partial\Omega$. Fix $\zeta \in \Omega$ with $\varrho(\zeta) \le 2\varrho(w)$ for a moment. Set $\varepsilon := |\varrho(w)|^{2/\alpha}$. Since $\varepsilon \le C_\alpha^{2/\alpha} \delta(w)^2$, we see that $\bar{B}(w, \varepsilon) \subset \Omega$ provided $\delta(w) \le \varepsilon_\alpha \ll 1$. For every $z \in \Omega$ with $\delta(z) \le \varepsilon$, we have

$$|\varrho(z)| \le C_\alpha \delta(z)^\alpha \le C_\alpha \varepsilon^\alpha = C_\alpha |\varrho(w)|^2 \quad (\le |\varrho(w)|/2) \tag{5-4}$$

provided $\delta(w) \le \varepsilon_\alpha \ll 1$. It follows from (2-7) and (5-4) that for every $\tau > 0$ there exists $\varepsilon_\tau \ll \varepsilon_\alpha$ such that

$$\sup_{\delta \le \varepsilon} |g_\Omega(\cdot, w)| \le \tau \tag{5-5}$$

provided $\delta(w) \le \varepsilon_\tau$. Since

$$C_\alpha \delta(\zeta)^\alpha \ge -\varrho(\zeta) \ge -2\varrho(w) = 2\varepsilon^{\alpha/2}$$

and Lemma 2.4 yields

$$C_1 |\zeta - w|^\alpha \ge \tfrac{3}{2}\varrho(w) - \varrho(\zeta) \ge -\tfrac{1}{2}\varrho(w) = \tfrac{1}{2}\varepsilon^{\alpha/2},$$

it follows that if $\delta(w) \le \varepsilon_\tau \ll 1$ then $\bar{B}(\zeta, \varepsilon) \subset \Omega$ and

$$\bar{B}(\zeta, \varepsilon) \cap \bar{B}(w, \varepsilon) = \varnothing. \tag{5-6}$$

By Lemma 5.2, there exists $\tilde{\zeta} \in \bar{B}(\zeta, \varepsilon)$ such that (5-2) holds.

Now set

$$\Psi(z) := \sup\{u(z) : u \in \mathrm{PSH}^-(\Omega),\ u|_{\bar{B}(w,\varepsilon)} \le -1\}.$$

We claim that

$$g_\Omega(z, w) \ge \log R/\varepsilon \Psi(z), \quad z \in \Omega \setminus B(w, \varepsilon), \qquad g_\Omega(z, w) \le \log \delta(w)/\varepsilon \Psi(z), \quad z \in \Omega. \tag{5-7}$$

To see this, first notice that

$$\log \frac{|z-w|}{R} \le g_\Omega(z, w) \le \log \frac{|z-w|}{\delta(w)}, \quad z \in \Omega. \tag{5-8}$$

Since

$$u(z) = \begin{cases} \log|z-w|/R & \text{if } z \in B(w, \varepsilon), \\ \max\{\log|z-w|/R, \log R/\varepsilon \Psi(z)\} & \text{if } z \in \Omega \setminus B(w, \varepsilon) \end{cases}$$

is a negative psh function on Ω with a logarithmic pole at w, it follows that

$$g_\Omega(z, w) \ge \log R/\varepsilon \Psi(z), \quad z \in \Omega \setminus B(w, \varepsilon).$$

Since (5-8) implies $g_\Omega(\,\cdot\,, w)|_{\bar{B}(w,\varepsilon)} \le \log \varepsilon/\delta(w)$, we have

$$\Psi(z) \ge \frac{g_\Omega(z, w)}{\log \delta(w)/\varepsilon}, \quad z \in \Omega.$$

By (5-5) and (5-7), we obtain

$$\sup_{\delta \le \varepsilon} |\Psi| \le \frac{\tau}{\log \delta(w)/\varepsilon}. \tag{5-9}$$

Set $\widetilde{\Omega} = \Omega - (\tilde{\zeta} - \zeta)$ and

$$v(z) = \begin{cases} \Psi(z) & \text{if } z \in \Omega \setminus \widetilde{\Omega}, \\ \max\{\Psi(z), \Psi(z + \tilde{\zeta} - \zeta) - \tau/(\log \delta(w)/\varepsilon)\} & \text{if } z \in \Omega \cap \widetilde{\Omega}. \end{cases}$$

Since $\Omega \cap \partial\widetilde{\Omega} \subset \{\delta \le \varepsilon\}$, it follows from (5-9) that $v \in \mathrm{PSH}^-(\Omega)$. Since

$$\Psi(z) \le \frac{\log|z-w|/\delta(w)}{\log R/\varepsilon}, \quad z \in \Omega \setminus B(w, \varepsilon),$$

in view of (5-8) and (5-7), and $z + \tilde{\zeta} - \zeta \in \bar{B}(w, 2\varepsilon)$ if $z \in \bar{B}(w, \varepsilon)$, it follows from the maximal principle that

$$v|_{\bar{B}(w,\varepsilon)} \le -\frac{\log \delta(w)/(2\varepsilon)}{\log R/\varepsilon}.$$

Thus,

$$\Psi(\tilde{\zeta}) - \frac{\tau}{\log \delta(w)/\varepsilon} \le v(\zeta) \le \frac{\log \delta(w)/(2\varepsilon)}{\log R/\varepsilon} \Psi(\zeta).$$

Combining with (5-6) and (5-7), we obtain

$$g_\Omega(\zeta, w) \geq \frac{(\log R/\varepsilon)^2}{\log \delta(w)/\varepsilon \cdot \log \delta(w)/(2\varepsilon)}(g_\Omega(\tilde{\zeta}, w) - \tau) \geq C_3(g_\Omega(\tilde{\zeta}, w) - \tau)$$

since $\delta(w) \geq |\varrho(w)/C_\alpha|^{1/\alpha} = \sqrt{\varepsilon}/C_\alpha^{1/\alpha}$. If we choose $\tau = 1/(2C_3)$, then

$$\begin{aligned} g_\Omega(\zeta, w) &\geq -C_3(n!)^{1/n}(\log R/\varepsilon)^{1-1/n}|g_\Omega(w,\zeta)|^{1/n} - \tfrac{1}{2} \quad \text{(by (5-2))} \\ &\geq -C_4|\log|\varrho(w)||^{1-1/n}\frac{|\varrho(w)\log|\varrho(\zeta)||^{1/n}}{|\varrho(\zeta)|^{1/n}} - \tfrac{1}{2} \quad \text{(by (2-7))} \\ &\geq -C_5\frac{|\varrho(w)|^{1/n}|\log|\varrho(w)||}{|\varrho(\zeta)|^{1/n}} - \tfrac{1}{2} \end{aligned}$$

since $\varrho(\zeta) \leq 2\varrho(w)$. Thus,

$$\{g_\Omega(\cdot\,, w) < -1\} \cap \{\varrho \leq 2\varrho(w)\} \subset \{\varrho > -C\nu(w)\}$$

provided $C \gg 1$. Since $\{\varrho > 2\varrho(w)\} \subset \{\varrho > -C\nu(w)\}$ if $C \gg 1$, we conclude the proof. □

6. Pointwise estimate of the normalized Bergman kernel and applications

Proof of Theorem 1.7. By Proposition 2.3, we know that for every $0 < r < 1$ there exist constants $\varepsilon_r, C_r > 0$ such that

$$\int_{-\varrho \leq \varepsilon} |K_\Omega(\cdot\,, w)|^2/K_\Omega(w) \leq C_r(\varepsilon/\mu(w))^r$$

for all $\varepsilon \leq \varepsilon_r\mu(w)$. Fix $z \in \Omega$ with $b(z) := C\nu(z) \leq \varepsilon_r\mu(w)$ for a moment, where C is the constant in (5-1). Let $\chi : \mathbb{R} \to [0, 1]$ be a smooth function satisfying $\chi|_{(0,\infty)} = 0$ and $\chi|_{(-\infty,-\log 2)} = 1$. We proceed with the proof in a similar way as [Chen 1999]. Notice that $g_\Omega(\cdot\,, z)$ is a continuous negative psh function on $\Omega \setminus \{z\}$ which satisfies

$$-i\partial\bar{\partial}\log(-g_\Omega(\cdot\,, z)) \geq i\partial\log(-g_\Omega(\cdot\,, z)) \wedge \bar{\partial}\log(-g_\Omega(\cdot\,, z))$$

as currents. By virtue of the Donnelly–Fefferman estimate [1983] (see also [Berndtsson and Charpentier 2000]), there exists a solution of the equation

$$\bar{\partial}u = K_\Omega(\cdot\,, w)\bar{\partial}\chi(-\log(-g_\Omega(\cdot\,, z)))$$

such that

$$\begin{aligned} \int_\Omega |u|^2 e^{-2ng_\Omega(\cdot\,,z)} &\leq C_0\int_\Omega |K_\Omega(\cdot\,, w)|^2 |\bar{\partial}\chi(-\log(-g_\Omega(\cdot\,, z)))|^2_{-i\partial\bar{\partial}\log(-g_\Omega(\cdot\,,z))} e^{-2ng_\Omega(\cdot\,,z)} \\ &\leq C_n\int_{\varrho > -b(z)} |K_\Omega(\cdot\,, w)|^2 \quad \text{(by (5-1))} \\ &\leq C_{n,r}K_\Omega(w)(\nu(z)/\mu(w))^r. \end{aligned}$$

Set

$$f := K_\Omega(\cdot\,, w)\chi(-\log(-g_\Omega(\cdot\,, z))) - u.$$

Clearly, we have $f \in \mathcal{O}(\Omega)$. Since $g_\Omega(\zeta, z) = \log|\zeta - z| + O(1)$ as $\zeta \to z$ and u is holomorphic in a neighborhood of z, it follows that $u(z) = 0$, i.e., $f(z) = K_\Omega(z, w)$. Moreover,

$$\begin{aligned}\int_\Omega |f|^2 &\le 2\int_{\varrho > -b(z)} |K_\Omega(\cdot\,, w)|^2 + 2\int_\Omega |u|^2 \\ &\le C_{n,r} K_\Omega(w)(\nu(z)/\mu(w))^r\end{aligned}$$

since $g_\Omega(\cdot\,, z) < 0$. Thus, we get

$$K_\Omega(z) \ge \frac{|f(z)|^2}{\|f\|^2_{L^2(\Omega)}} \ge C_{n,r}^{-1}\frac{|K_\Omega(z, w)|^2}{K_\Omega(w)}(\mu(w)/\nu(z))^r,$$

and

$$\mathcal{B}_\Omega(z, w) \le C_{n,r}(\nu(z)/\mu(w))^r.$$

If $b(z) > \varepsilon_r \mu(w)$, then the inequality above trivially holds since $|K_\Omega(z, w)|^2/(K_\Omega(z)K_\Omega(w)) \le 1$. By symmetry of $\mathcal{B}_\Omega$, the assertion immediately follows. □

Remark. It would be interesting to get pointwise estimates for $|S_\Omega(z, w)|^2/(S_\Omega(z)S_\Omega(w))$, where S_Ω is the Szegö kernel (compare to [Chen and Fu 2011]).

Proof of Corollary 1.8. Let $z \in \Omega$ be an arbitrarily fixed point which is sufficiently close to $\partial\Omega$. By the Hopf–Rinow theorem, there exists a Bergman geodesic γ jointing z_0 to z, for ds_B^2 is complete on Ω. We may choose a finite number of points $\{z_k\}_{k=1}^m \subset \gamma$ with the order

$$z_0 \to z_1 \to z_2 \to \cdots \to z_m \to z,$$

where

$$|\varrho(z_{k+1})|(1 + |\log|\varrho(z_{k+1})||)^{n+2} = |\varrho(z_k)|$$

and

$$|\varrho(z)|(1 + |\log|\varrho(z)||)^{n+2} \ge |\varrho(z_m)|.$$

Since

$$\begin{aligned}\frac{\nu(z_{k+1})}{\mu(z_k)} &= \frac{|\varrho(z_{k+1})|}{|\varrho(z_k)|}(1 + |\log|\varrho(z_{k+1})||)^n(1 + |\log|\varrho(z_k)||) \\ &\le \frac{|\varrho(z_{k+1})|}{|\varrho(z_k)|}(1 + |\log|\varrho(z_{k+1})||)^{n+1} \\ &= (1 + |\log|\varrho(z_{k+1})||)^{-1},\end{aligned}$$

it follows from Theorem 1.7 that there exists $k_0 \in \mathbb{Z}^+$ such that $\mathcal{B}_\Omega(z_k, z_{k+1}) \le \frac{1}{4}$ for all $k \ge k_0$. By (1-4),

$$d_B(z_k, z_{k+1}) \ge 1.$$

Notice that

$$\begin{aligned}|\varrho(z_{k_0})| &= |\varrho(z_{k_0+1})||\log|\varrho(z_{k_0+1})||^{n+2} \\ &\le |\varrho(z_{k_0+2})||\log|\varrho(z_{k_0+2})||^{2(n+2)} \\ &\le \cdots \le |\varrho(z_m)||\log|\varrho(z_m)||^{(m-k_0)(n+2)}.\end{aligned}$$

Thus,

$$m - k_0 \ge \text{const.}\frac{|\log|\varrho(z_m)||}{\log|\log|\varrho(z_m)||} \ge \text{const.}\frac{|\log|\varrho(z)||}{\log|\log|\varrho(z)||}$$

so that

$$\begin{aligned} d_B(z, z_0) &\ge \sum_{k=k_0}^{m-1} d_B(z_k, z_{k+1}) \ge m - k_0 - 1 \\ &\ge \text{const.}\frac{|\log|\varrho(z)||}{|\log|\log|\varrho(z)|||} \\ &\ge \text{const.}\frac{|\log\delta(z)|}{\log|\log\delta(z)|} \end{aligned}$$

since $|\varrho(z)| \le C_\alpha \delta^\alpha$ for any $\alpha < \alpha(\Omega)$. □

Proof of Corollary 1.9. For every $0 < \alpha < \alpha(\Omega)$, we have $-\varrho \le C_\alpha \delta^\alpha$. Theorem 1.7 then yields

$$D_B(z_0, z) \ge \alpha|\log\delta(z)|$$

as $z \to \partial\Omega$. Thus, it suffices to show

$$d_K(z, z_0) \le C|\log\delta(z)| \tag{6-1}$$

as $z \to \partial\Omega$. To see this, let F_K be the Kobayashi–Royden metric. Since F_K is decreasing under holomorphic mappings, we conclude that $F_K(z; X)$ is dominated by the KR metric of the ball $B(z, \delta(z))$. Thus, $F_K(z; X) \le C|X|/\delta(z)$, from which (6-1) immediately follows (compare to the proof of Proposition 7.3 in [Chen 2016]). □

In order to prove Corollary 1.10, we need the following elementary fact.

Lemma 6.1. *If $\Omega \subset \mathbb{C}^n$ is a bounded weighted circular domain which contains the origin, then $K_\Omega(z, 0) = K_\Omega(0)$ for any $z \in \Omega$.*

Proof. For fixed $\theta \in \mathbb{R}$, we set $F_\theta(z) := (e^{ia_1\theta}z_1, \dots, e^{ia_n\theta}z_n)$. By the transform formula of the Bergman kernel,

$$K_\Omega(F_\theta(z), 0) = K_\Omega(z, 0), \quad z \in \Omega.$$

It follows that, for any n-tuple $(m_1, \dots, m_n)$ of nonnegative integers,

$$e^{i(a_1m_1+\cdots+a_nm_n)\theta}\frac{\partial^{m_1+\cdots+m_n}K_\Omega(z,0)}{\partial z_1^{m_1}\cdots\partial z_n^{m_n}}\bigg|_{z=0} = \frac{\partial^{m_1+\cdots+m_n}K_\Omega(z,0)}{\partial z_1^{m_1}\cdots\partial z_n^{m_n}}\bigg|_{z=0} \quad \text{for all } \theta \in \mathbb{R}$$

so that $\frac{\partial^{m_1+\cdots+m_n}K_\Omega(z,0)}{\partial z_1^{m_1}\cdots\partial z_n^{m_n}}\Big|_{z=0} = 0$ if not all m_j are zero. Taylor's expansion of $K_\Omega(z, 0)$ at $z = 0$ and the identity theorem of holomorphic functions yield $K_\Omega(z, 0) = K_\Omega(0)$ for any $z \in \Omega$. □

Proof of Corollary 1.10. By Lemma 6.1,

$$\mathcal{B}_{\Omega_2}(F(z), 0) = K_{\Omega_2}(0)K_{\Omega_2}(F(z))^{-1} \ge C^{-1}\delta_2(F(z))^{2n}.$$

On the other hand, Theorem 1.7 implies

$$\mathcal{B}_{\Omega_1}(z, F^{-1}(0)) \le C_\alpha \delta_1(z)^\alpha.$$

Since $\mathcal{B}_{\Omega_2}(F(z), 0) = \mathcal{B}_{\Omega_1}(z, F^{-1}(0))$, we conclude the proof. □

Appendix: Examples of domains with positive hyperconvexity indices

We start with the following almost trivial fact.

Proposition A.1. *Let Ω_1 and Ω_2 be two bounded domains in $\mathbb{C}^n$ such that there exists a biholomorphic map $F : \Omega_1 \to \Omega_2$ which extends to a Hölder-continuous map $\overline{\Omega}_1 \to \overline{\Omega}_2$. If $\alpha(\Omega_2) > 0$, then $\alpha(\Omega_1) > 0$.*

Proof. Let δ_1 and δ_2 denote the boundary distances of Ω_1 and Ω_2, respectively. Choose $\rho_2 \in \mathrm{PSH}^- \cap C(\Omega_2)$ such that $-\rho_2 \le C\delta_2^\alpha$ for some $C, \alpha > 0$. Set $\rho_1 := \rho_2 \circ F$. Clearly, $\rho_1 \in \mathrm{PSH}^- \cap C(\Omega_1)$. For fixed $z \in \Omega_1$, we choose $z^* \in \partial\Omega_1$ so that $|z - z^*| = \delta_1(z)$. Since $F(z^*) \in \partial\Omega_2$, it follows that

$$\begin{aligned} -\rho_1(z) &\le C\delta_2(F(z))^\alpha = C(\delta_2(F(z)) - \delta_2(F(z^*)))^\alpha \\ &\le C|F(z) - F(z^*)|^\alpha \le C|z - z^*|^{\gamma\alpha} \\ &\le C\delta_1(z)^{\gamma\alpha}, \end{aligned}$$

where γ is the order of Hölder continuity of F on $\overline{\Omega}_1$. □

Example. Let $D \subset \mathbb{C}$ be a bounded Jordan domain which admits a uniformly Hölder-continuous conformal map f onto the unit disc Δ (e.g., a quasidisc with a fractal boundary). Set $F(z_1, \dots, z_n) := (f(z_1), \dots, f(z_n))$. Clearly, F is a biholomorphic map between D^n and Δ^n which extends to a Hölder-continuous map between their closures. Let

$$\Omega_2 := \{z \in \mathbb{C}^n : |z_1|^{a_1} + \cdots + |z_n|^{a_n} < 1\},$$

where $a_j > 0$. Clearly, we have $\alpha(\Omega_2) > 0$. By Proposition A.1, we conclude that the domain $\Omega_1 := F^{-1}(\Omega_2)$ satisfies $\alpha(\Omega_1) > 0$. Notice that some parts of $\partial\Omega_1$ might be highly irregular.

A domain $\Omega \subset \mathbb{C}^n$ is called $\mathbb{C}$-convex if $\Omega \cap L$ is a simply connected domain in L for every affine complex line L. Clearly, every convex domain is $\mathbb{C}$-convex.

Proposition A.2. *If $\Omega \subset \mathbb{C}^n$ is a bounded $\mathbb{C}$-convex domain, then $\alpha(\Omega) \ge \frac{1}{2}$.*

Proof. Let $w \in \Omega$ be an arbitrarily fixed point. Let w^* be a point on $\partial\Omega$ satisfying $\delta(w) = |w - w^*|$. Let L be the complex line determined by w and w^*. Since every $\mathbb{C}$-convex domain is linearly convex [Hörmander 1994, Theorem 4.6.8], it follows that there exists an affine complex hyperplane $H \subset \mathbb{C}^n \setminus \Omega$ with $w^* \in H$. Since $|w - w^*| = \delta(w)$, H has to be *orthogonal* to L. Let π_L denote the natural projection $\mathbb{C}^n \to L$. Notice that $\pi_L(\Omega)$ is a bounded simply connected domain in L in view of [Hörmander 1994, Proposition 4.6.7]. By Proposition 7.3 in [Chen 2016], there exists a negative continuous function ρ_L on $\pi_L(\Omega)$ with

$$(\delta_L/\delta_L(z_L^0))^2 \le -\rho_L \le (\delta_L/\delta_L(z_L^0))^{1/2},$$

where δ_L denotes the boundary distance of $\pi_L(\Omega)$ and $z_L^0 \in \pi_L(\Omega)$ satisfies $\delta_L(z_L^0) = \sup_{\pi_L(\Omega)} \delta_L$. Fix a point $z^0 \in \Omega$. We have

$$\delta_L(z_L^0) \ge \delta_L(\pi_L(z^0)) \ge \delta(z^0).$$

Set

$$\varrho_{z_0}(z) = \sup\{u(z) : u \in \mathrm{PSH}^-(\Omega),\ u(z^0) \le -1\}.$$

Clearly, $\varrho_{z_0} \in \mathrm{PSH}^-(\Omega)$. Since $\Omega \subset \pi_L^{-1}(\pi_L(\Omega))$, it follows that $\pi_L^*(\rho_L) \in \mathrm{PSH}^-(\Omega)$. Since $\pi_L^*(\delta_L)(w) = \delta(w)$ and

$$\pi_L^*(\rho_L)(z^0) = \rho_L(\pi_L(z^0)) \le -(\delta_L(\pi_L(z^0))/\delta_L(z_L^0))^2,$$

then

$$\begin{aligned}\varrho_{z_0}(w) &\ge (\delta_L(z_L^0)/\delta_L(\pi_L(z^0)))^2 \pi_L^*(\rho_L)(w) \\ &\ge -(\delta_L(z_L^0)^{3/2}/\delta_L(\pi_L(z^0))^2)\delta(w)^{1/2} \\ &\ge -(R^{3/2}/\delta(z^0)^2)\delta(w)^{1/2},\end{aligned}$$

where $R = \mathrm{diam}(\Omega)$. Thus, $\alpha(\Omega) \ge \frac{1}{2}$. □

Remark. After the first version of this paper was finished, the author was kindly informed by Nikolai Nikolov that Proposition A.2 follows also from Proposition 3(ii) of [Nikolov and Trybuła 2015].

Complex dynamics also provides interesting examples of domains with $\alpha(\Omega) > 0$. Let $q(z) = \sum_{j=0}^d a_j z^j$ be a complex polynomial of degree $d \ge 2$. Let q^n denote the n-iterates of q. The attracting basin at ∞ of q is defined by

$$F_\infty := \{z \in \overline{\mathbb{C}} : q^n(z) \to \infty \text{ as } n \to \infty\},$$

which is a domain in $\overline{\mathbb{C}}$ with $q(F_\infty) = F_\infty$. The Julia set of q is defined by $J := \partial F_\infty$. It is known that J is always uniformly perfect. Thus, $\alpha(F_\infty) > 0$.

We say that q is *hyperbolic* if there exist constants $C > 0$ and $\gamma > 1$ such that

$$\inf_J |(q^n)'| \ge C\gamma^n \quad \text{for all } n \ge 1.$$

Consider a holomorphic family $\{q_\lambda\}$ of hyperbolic polynomials of constant degree $d \ge 2$ over the unit disc Δ. Let F_∞^λ denote the attracting basin at ∞ of q_λ, and let $J_\lambda := \partial F_\infty^\lambda$. Let Ω_r denote the total space of F_∞^λ over the disc $\Delta_r := \{z \in \mathbb{C} : |z| < r\}$, where $0 < r \le 1$, that is

$$\Omega_r = \{(\lambda, w) : \lambda \in \Delta_r,\ w \in F_\infty^\lambda\}.$$

Proposition A.3. *For every* $0 < r < 1$, Ω_r *is a bounded domain in* $\mathbb{C}^2$ *with* $\alpha(\Omega_r) > 0$.

Proof. We first show that Ω_r is a domain. Mañé, Sad and Sullivan [Mañé et al. 1983] showed that there exists a family of maps $\{f_\lambda\}_{\lambda\in\Delta}$ such that

(1) $f_\lambda : J_0 \to J_\lambda$ is a homeomorphism for each $\lambda \in \Delta$,
(2) $f_0 = \mathrm{id}|_{J_0}$,
(3) $f(\lambda, z) := f_\lambda(z)$ is holomorphic on Δ for each $z \in J_0$ and
(4) $q_\lambda = f_\lambda \circ q_0 \circ f_\lambda^{-1}$ on J_λ, for each $\lambda \in \Delta$.

In other words, properties (1)–(3) say that $\{f_\lambda\}_{\lambda\in\Delta}$ gives a *holomorphic motion* of J_0. By a result of Slodkowski [1991], $\{f_\lambda\}_{\lambda\in\Delta}$ may be extended to a holomorphic motion $\{\tilde{f}_\lambda\}_{\lambda\in\Delta}$ of $\overline{\mathbb{C}}$ such that

(a) $\tilde{f}_\lambda : \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ is a quasiconformal map of dilatation $\le (1+|\lambda|)/(1-|\lambda|)$, for each $\lambda \in \Delta$,
(b) $\tilde{f}_\lambda : F_\infty^0 \to F_\infty^\lambda$ is a homeomorphism for each $\lambda \in \Delta$ and
(c) $\tilde{f}(\lambda, z) := \tilde{f}_\lambda(z)$ is jointly Hölder-continuous in (λ, z).

It follows immediately that Ω_r is a domain in $\mathbb{C}^n$ for each $r \le 1$. Let δ_λ and δ denote the boundary distances of F_∞^λ and Ω_1, respectively. We claim that for every $0 < r < 1$ there exists $\gamma > 0$ such that

$$\delta_\lambda(w) \le C\delta(\lambda, w)^\gamma, \quad \lambda \in \Delta_r,\ w \in F_\infty^\lambda. \tag{A-1}$$

To see this, choose $(\lambda', w_{\lambda'})$ where $w_{\lambda'} \in J_{\lambda'}$, such that

$$\delta(\lambda, w) = \sqrt{|\lambda - \lambda'|^2 + |w - w_{\lambda'}|^2}.$$

Write $w_{\lambda'} = \tilde{f}(\lambda', z_0)$ where $z_0 \in J_0$. Since $\tilde{f}(\lambda, z_0) \in J_\lambda$, it follows that

$$\begin{aligned}\delta_\lambda(w) &\le |w - \tilde{f}(\lambda, z_0)| \le |w - w_{\lambda'}| + |\tilde{f}(\lambda', z_0) - \tilde{f}(\lambda, z_0)| \\ &\le |w - w_{\lambda'}| + C|\lambda - \lambda'|^\gamma \\ &\le \delta(\lambda, w) + C\delta(\lambda, w)^\gamma \\ &\le C'\delta(\lambda, w)^\gamma,\end{aligned}$$

where γ is the order of Hölder continuity of $\tilde{f}$ on Ω_r.

Recall that the Green function $g_\lambda(w) := g_{F_\infty^\lambda}(w, \infty)$ at ∞ of F_∞^λ satisfies

$$g_\lambda(w) = \lim_{n\to\infty} d^{-n} \log|q_\lambda^n(w)|, \quad w \in F_\infty^\lambda, \tag{A-2}$$

where the convergence is uniform on compact subsets of F_∞^λ [Ransford 1995, Corollary 6.5.4]. Actually the proof of that result shows that the convergence is also uniform on compact subsets of Ω_1. Since $\log|q_\lambda^n(w)|$ is psh in (λ, w), so is $g(\lambda, w) := g_\lambda(w)$. By (A-1) it suffices to verify that for every $0 < r < 1$ there are positive constants C and α such that $-g_\lambda(w) \le C\delta_\lambda(w)^\alpha$ for each $\lambda \in \Delta_r$ and $w \in F_\infty^\lambda$. This can be verified similarly to the proof of Theorem 3.2 in [Carleson and Gamelin 1993]. □

Conjecture A.4. *Let $D \subset \mathbb{C}$ be a domain with $\alpha(D) > 0$. Let $\{f_\lambda\}_{\lambda\in\Delta}$ be a holomorphic motion of D. Let*

$$\Omega_r := \{(\lambda, w) : \lambda \in \Delta_r,\ w \in f_\lambda(D)\}.$$

One has $\alpha(\Omega_r) > 0$ for each $r < 1$.

Acknowledgements

It is my pleasure to thank the valuable comments from the referees, Professor Nikolai Nikolov and Dr. Xieping Wang.

References

[Adachi and Brinkschulte 2015] M. Adachi and J. Brinkschulte, "A global estimate for the Diederich–Fornaess index of weakly pseudoconvex domains", *Nagoya Math. J.* **220** (2015), 67–80.

[Andrievskii 2005] V. V. Andrievskii, "On sparse sets with Green function of the highest smoothness", *Comput. Methods Funct. Theory* **5**:2 (2005), 301–322.

[Barrett 1984] D. E. Barrett, "Irregularity of the Bergman projection on a smooth bounded domain in $\mathbf{C}^2$", *Ann. of Math.* (2) **119**:2 (1984), 431–436.

[Barrett 1992] D. E. Barrett, “Behavior of the Bergman projection on the Diederich–Fornæss worm”, *Acta Math.* **168**:1–2 (1992), 1–10.

[Berger 1996] M. Berger, “Encounter with a geometer: Eugenio Calabi”, pp. 20–60 in *Manifolds and geometry* (Pisa, 1993), edited by P. de Bartolomeis et al., Sympos. Math. **36**, Cambridge University, 1996.

[Berndtsson and Charpentier 2000] B. Berndtsson and P. Charpentier, “A Sobolev mapping property of the Bergman kernel”, *Math. Z.* **235**:1 (2000), 1–10.

[Bertilsson 1998] D. Bertilsson, “Coefficient estimates for negative powers of the derivative of univalent functions”, *Ark. Mat.* **36**:2 (1998), 255–273.

[Błocki 1993] Z. Błocki, “Estimates for the complex Monge–Ampère operator”, *Bull. Polish Acad. Sci. Math.* **41**:2 (1993), 151–157.

[Błocki 2002] Z. Błocki, “The complex Monge–Ampère operator in pluripotential theory”, lecture notes, Jagiellonian University, Kraków, 2002, Available at http://gamma.im.uj.edu.pl/~blocki/publ/ln/wykl.pdf.

[Błocki 2005] Z. Błocki, “The Bergman metric and the pluricomplex Green function”, *Trans. Amer. Math. Soc.* **357**:7 (2005), 2613–2625.

[Błocki and Pflug 1998] Z. Błocki and P. Pflug, “Hyperconvexity and Bergman completeness”, *Nagoya Math. J.* **151** (1998), 221–225.

[Boas and Straube 1991] H. P. Boas and E. J. Straube, “Sobolev estimates for the $\bar\partial$-Neumann operator on domains in $\mathbf{C}^n$ admitting a defining function that is plurisubharmonic on the boundary”, *Math. Z.* **206**:1 (1991), 81–88.

[Brennan 1978] J. E. Brennan, “The integrability of the derivative in conformal mapping”, *J. London Math. Soc.* (2) **18**:2 (1978), 261–272.

[Calabi 1953] E. Calabi, “Isometric imbedding of complex manifolds”, *Ann. of Math.* (2) **58** (1953), 1–23.

[Carleson 1967] L. Carleson, *Selected problems on exceptional sets*, Van Nostrand Math. Stud. **13**, Van Nostrand, Princeton, 1967.

[Carleson and Gamelin 1993] L. Carleson and T. W. Gamelin, *Complex dynamics*, Springer, New York, 1993.

[Carleson and Jones 1992] L. Carleson and P. W. Jones, “On coefficient problems for univalent functions and conformal dimension”, *Duke Math. J.* **66**:2 (1992), 169–206.

[Carleson and Makarov 1994] L. Carleson and N. G. Makarov, “Some results connected with Brennan’s conjecture”, *Ark. Mat.* **32**:1 (1994), 33–62.

[Carleson and Totik 2004] L. Carleson and V. Totik, “Hölder continuity of Green’s functions”, *Acta Sci. Math.* (*Szeged*) **70**:3–4 (2004), 557–608.

[Chen 1999] B.-Y. Chen, “Completeness of the Bergman metric on non-smooth pseudoconvex domains”, *Ann. Polon. Math.* **71**:3 (1999), 241–251.

[Chen 2016] B.-Y. Chen, “Parameter dependence of the Bergman kernels”, *Adv. Math.* **299** (2016), 108–138.

[Chen and Fu 2011] B.-Y. Chen and S. Fu, “Comparison of the Bergman and Szegö kernels”, *Adv. Math.* **228**:4 (2011), 2366–2384.

[Christ 1991] M. Christ, “On the $\bar\partial$ equation in weighted L^2 norms in $\mathbb{C}^1$”, *J. Geom. Anal.* **1**:3 (1991), 193–230.

[Christ 2013] M. Christ, “Upper bounds for Bergman kernels associated to positive line bundles with smooth Hermitian metrics”, preprint, 2013. arXiv

[Delin 1998] H. Delin, “Pointwise estimates for the weighted Bergman projection kernel in $\mathbb{C}^n$, using a weighted L^2 estimate for the $\bar\partial$ equation”, *Ann. Inst. Fourier* (*Grenoble*) **48**:4 (1998), 967–997.

[Demailly 1987] J.-P. Demailly, “Mesures de Monge–Ampère et mesures pluriharmoniques”, *Math. Z.* **194**:4 (1987), 519–564.

[Diederich and Fornaess 1977] K. Diederich and J. E. Fornaess, “Pseudoconvex domains: bounded strictly plurisubharmonic exhaustion functions”, *Invent. Math.* **39**:2 (1977), 129–141.

[Diederich and Fornaess 1979] K. Diederich and J. E. Fornaess, “Proper holomorphic maps onto pseudoconvex domains with real-analytic boundary”, *Ann. of Math.* (2) **110**:3 (1979), 575–592.

[Diederich and Ohsawa 1995] K. Diederich and T. Ohsawa, "An estimate for the Bergman distance on pseudoconvex domains", *Ann. of Math.* (2) **141**:1 (1995), 181–190.

[Donnelly and Fefferman 1983] H. Donnelly and C. Fefferman, "L^2-cohomology and index theorem for the Bergman metric", *Ann. of Math.* (2) **118**:3 (1983), 593–618.

[Edholm and McNeal 2016] L. D. Edholm and J. D. McNeal, "The Bergman projection on fat Hartogs triangles: L^p boundedness", *Proc. Amer. Math. Soc.* **144**:5 (2016), 2185–2196.

[Fu and Shaw 2016] S. Fu and M.-C. Shaw, "The Diederich–Fornæss exponent and non-existence of Stein domains with Levi-flat boundaries", *J. Geom. Anal.* **26**:1 (2016), 220–230.

[Harrington 2008] P. S. Harrington, "The order of plurisubharmonicity on pseudoconvex domains with Lipschitz boundaries", *Math. Res. Lett.* **15**:3 (2008), 485–490.

[Hedberg 1972] L. I. Hedberg, "Approximation in the mean by analytic functions", *Trans. Amer. Math. Soc.* **163** (1972), 157–171.

[Herbort 1999] G. Herbort, "The Bergman metric on hyperconvex domains", *Math. Z.* **232**:1 (1999), 183–196.

[Herbort 2000] G. Herbort, "The pluricomplex Green function on pseudoconvex domains with a smooth boundary", *Internat. J. Math.* **11**:4 (2000), 509–522.

[Hörmander 1994] L. Hörmander, *Notions of convexity*, Progr. Math. **127**, Birkhäuser, Boston, 1994.

[Jarnicki and Pflug 1993] M. Jarnicki and P. Pflug, *Invariant distances and metrics in complex analysis*, De Gruyter Expos. Math. **9**, De Gruyter, Berlin, 1993.

[Krantz and Peloso 2008] S. G. Krantz and M. M. Peloso, "Analysis and geometry on worm domains", *J. Geom. Anal.* **18**:2 (2008), 478–510.

[Lanzani 2015] L. Lanzani, "Harmonic analysis techniques in several complex variables", reprint, 2015. arXiv

[Lempert 1986] L. Lempert, "On the boundary behavior of holomorphic mappings", pp. 193–215 in *Contributions to several complex variables*, edited by A. Howard and P.-M. Wong, Asp. Math. **E9**, Vieweg, Braunschweig, 1986.

[Lindholm 2001] N. Lindholm, "Sampling in weighted L^p spaces of entire functions in $\mathbb{C}^n$ and estimates of the Bergman kernel", *J. Funct. Anal.* **182**:2 (2001), 390–426.

[Ma and Marinescu 2007] X. Ma and G. Marinescu, *Holomorphic Morse inequalities and Bergman kernels*, Progr. Math. **254**, Birkhäuser, Basel, 2007.

[Mañé et al. 1983] R. Mañé, P. Sad, and D. Sullivan, "On the dynamics of rational maps", *Ann. Sci. École Norm. Sup.* (4) **16**:2 (1983), 193–217.

[Nikolov and Trybuła 2015] N. Nikolov and M. Trybuła, "The Kobayashi balls of ($\mathbb{C}$-)convex domains", *Monatsh. Math.* **177**:4 (2015), 627–635.

[Ohsawa 1993] T. Ohsawa, "On the Bergman kernel of hyperconvex domains", *Nagoya Math. J.* **129** (1993), 43–52.

[Ohsawa and Takegoshi 1987] T. Ohsawa and K. Takegoshi, "On the extension of L^2 holomorphic functions", *Math. Z.* **195**:2 (1987), 197–204.

[Poletsky and Stessin 2008] E. A. Poletsky and M. I. Stessin, "Hardy and Bergman spaces on hyperconvex domains and their composition operators", *Indiana Univ. Math. J.* **57**:5 (2008), 2153–2201.

[Pommerenke 1979] C. Pommerenke, "Uniformly perfect sets and the Poincaré metric", *Arch. Math.* (*Basel*) **32**:2 (1979), 192–199.

[Pommerenke 1992] C. Pommerenke, *Boundary behaviour of conformal maps*, Grundlehren math. Wissenschaften **299**, Springer, Berlin, 1992.

[Ransford 1995] T. Ransford, *Potential theory in the complex plane*, London Math. Soc. Student Texts **28**, Cambridge University, 1995.

[Schiffer 1946] M. Schiffer, "The kernel function of an orthonormal system", *Duke Math. J.* **13** (1946), 529–540.

[Skwarczyński 1980] M. Skwarczyński, *Biholomorphic invariants related to the Bergman function*, Dissertationes Math. (Rozprawy Mat.) **173**, Instytut Matematyczny Polskiej Akademii Nauk, Warsaw, 1980.

[Slodkowski 1991] Z. Slodkowski, "Holomorphic motions and polynomial hulls", *Proc. Amer. Math. Soc.* **111**:2 (1991), 347–355.

[Totik 2006] V. Totik, *Metric properties of harmonic measures*, Mem. Amer. Math. Soc. **867**, 2006.

[Walsh 1968] J. B. Walsh, “Continuity of envelopes of plurisubharmonic functions”, *J. Math. Mech.* **18** (1968), 143–148.

[Zelditch 2016] S. Zelditch, “Off-diagonal decay of toric Bergman kernels”, *Lett. Math. Phys.* **106**:12 (2016), 1849–1864.

Received 11 Nov 2016. Revised 27 Feb 2017. Accepted 24 Apr 2017.

BO-YONG CHEN: boychen@fudan.edu.cn
School of Mathematical Sciences, Fudan University, 220 Handan Road, Shanghai 200433, China

ANALYSIS AND PDE
Vol. 10, No. 6, 2017
dx.doi.org/10.2140/apde.2017.10.1455

STRUCTURE OF SETS WHICH ARE WELL APPROXIMATED BY ZERO SETS OF HARMONIC POLYNOMIALS

MATTHEW BADGER, MAX ENGELSTEIN AND TATIANA TORO

The zero sets of harmonic polynomials play a crucial role in the study of the free boundary regularity problem for harmonic measure. In order to understand the fine structure of these free boundaries, a detailed study of the singular points of these zero sets is required. In this paper we study how "degree-k points" sit inside zero sets of harmonic polynomials in $\mathbb{R}^n$ of degree d (for all $n \geq 2$ and $1 \leq k \leq d$) and inside sets that admit arbitrarily good local approximations by zero sets of harmonic polynomials. We obtain a general structure theorem for the latter type of sets, including sharp Hausdorff and Minkowski dimension estimates on the singular set of degree-k points ($k \geq 2$) without proving uniqueness of blowups or aid of PDE methods such as monotonicity formulas. In addition, we show that in the presence of a certain topological separation condition, the sharp dimension estimates improve and depend on the parity of k. An application is given to the two-phase free boundary regularity problem for harmonic measure below the continuous threshold introduced by Kenig and Toro.

1. Introduction

In this paper, we study the geometry of sets that admit arbitrarily good local approximations by zero sets of harmonic polynomials. As our conditions are reminiscent of those introduced by Reifenberg [1960], we often refer to these sets as Reifenberg-type sets which are well approximated by zero sets of harmonic

Badger was partially supported by NSF grant DMS 1500382. Engelstein was partially supported by an NSF Graduate Research Fellowship, NSF DGE 1144082. Toro was partially supported by NSF grant DMS 1361823, and the Robert R. & Elaine F. Phelps Professorship in Mathematics.
MSC2010: primary 33C55, 49J52; secondary 28A75, 31A15, 35R35.
Keywords: Reifenberg-type sets, harmonic polynomials, Łojasiewicz-type inequalities, singular set, Hausdorff dimension , Minkowski dimension, two-phase free boundary problems, harmonic measure, NTA domains.

polynomials. This class of sets plays a crucial role in the study of a two-phase free boundary problem for harmonic measure with weak initial regularity, examined first by Kenig and Toro [2006] and subsequently by Kenig, Preiss and Toro [Kenig et al. 2009], Badger [2011; 2013], Badger and Lewis [2015], and Engelstein [2016]. Our results are partly motivated by several open questions about the structure and size of the singular set in the free boundary, which we answer definitively below. In particular, we obtain sharp bounds on the upper Minkowski and Hausdorff dimensions of the singular set, which depend on the degree of blowups of the boundary. It is important to remark that this is one of those rare instances in which a singular set of a nonvariational problem can be well understood. Often, in this type of question, the lack of a monotonicity formula is a serious obstacle. A remarkable feature of the proof is that Łojasiewicz-type inequalities for harmonic polynomials are used to establish a relationship between the terms in the Taylor expansion of a harmonic polynomial at a given point in its zero set and the extent to which this zero set can be approximated by the zero set of a lower-order harmonic polynomial (see Sections 3 and 4). In a broader context, this paper also complements the recent investigations by Cheeger, Naber, and Valtorta [Cheeger et al. 2015] and Naber and Valtorta [2014] into volume estimates for the critical sets of harmonic functions and solutions to certain second-order elliptic operators with Lipschitz coefficients. Detailed descriptions of these past works and new results appear below, after we introduce some requisite notation.

For all $n \geq 2$ and $d \geq 1$, let $\mathcal{H}_{n,d}$ denote the collection of all zero sets Σ_p of nonconstant harmonic polynomials $p : \mathbb{R}^n \to \mathbb{R}$ of degree at most d such that $0 \in \Sigma_p$ (i.e., $p(0) = 0$). For every nonempty set $A \subseteq \mathbb{R}^n$, location $x \in A$, and scale $r > 0$, we introduce the *bilateral approximation number* $\Theta_A^{\mathcal{H}_{n,d}}(x, r)$, which, roughly speaking, records how well A looks like some zero set of a harmonic polynomial of degree at most d in the open ball $B(x, r) = \{y \in \mathbb{R}^n : |y - x| < r\}$:

$$\Theta_A^{\mathcal{H}_{n,d}}(x,r) = \frac{1}{r} \inf_{\Sigma_p \in \mathcal{H}_{n,d}} \max\Big\{ \sup_{a \in A \cap B(x,r)} \operatorname{dist}(a, x+\Sigma_p), \sup_{z \in (x+\Sigma_p) \cap B(x,r)} \operatorname{dist}(z, A)\Big\} \in [0,1]. \tag{1-1}$$

When $\Theta_A^{\mathcal{H}_{n,d}}(x, r) = 0$, the closure, $\bar{A}$, of A coincides with the zero set of some harmonic polynomial of degree at most d in $B(x, r)$. At the other extreme, when $\Theta_A^{\mathcal{H}_{n,d}}(x, r) \sim 1$, the set A stays "far away" in $B(x, r)$ from every zero set of a nonconstant harmonic polynomial of degree at most d containing x. We observe that the approximation numbers are scale invariant in the sense that $\Theta_{\lambda A}^{\mathcal{H}_{n,d}}(\lambda x, \lambda r) = \Theta_A^{\mathcal{H}_{n,d}}(x, r)$ for all $\lambda > 0$. A point x in a nonempty set A is called an $\mathcal{H}_{n,d}$ *point* of A if $\lim_{r \to 0} \Theta_A^{\mathcal{H}_{n,d}}(x, r) = 0$.

For all $n \geq 2$ and $k \geq 1$, let $\mathcal{F}_{n,k}$ denote the collection of all zero sets of homogeneous harmonic polynomials $p : \mathbb{R}^n \to \mathbb{R}$ of degree k. We note that

$$\mathcal{F}_{n,k} \subseteq \mathcal{H}_{n,d} \quad \text{whenever } 1 \leq k \leq d.$$

For every nonempty set $A \subseteq \mathbb{R}^n$, $x \in A$, and $r > 0$, the bilateral approximation number $\Theta_A^{\mathcal{F}_{n,k}}(x, r)$ is defined analogously to $\Theta_A^{\mathcal{H}_{n,d}}(x, r)$ except that the zero set Σ_p in the infimum ranges over $\mathcal{F}_{n,k}$ instead of $\mathcal{H}_{n,d}$. A point x in a nonempty set A is called an $\mathcal{F}_{n,k}$ *point* of A if $\lim_{r \to 0} \Theta_A^{\mathcal{F}_{n,k}}(x, r) = 0$. This means that infinitesimally at x, A looks like the zero set of a homogeneous harmonic polynomial of degree k.

We say that a nonempty set $A \subseteq \mathbb{R}^n$ is *locally bilaterally well approximated by* $\mathcal{H}_{n,d}$ if for all $\varepsilon > 0$ and for all compact sets $K \subseteq A$ there exists $r_{\varepsilon,K} > 0$ such that $\Theta_A^{\mathcal{H}_{n,d}}(x, r) \leq \varepsilon$ for all $x \in K$ and $0 < r \leq r_{\varepsilon,K}$.

If $k = 1$, then $\mathcal{H}_{n,1} = \mathcal{F}_{n,1} = G(n, n-1)$ is the collection of codimension-1 hyperplanes through the origin, and sets A that are locally bilaterally well approximated by $\mathcal{H}_{n,1}$ are also called *Reifenberg flat sets with vanishing constant* or *Reifenberg vanishing sets* (e.g., see [David et al. 2001]). Our initial result is the following structure theorem for sets that are locally bilaterally well approximated by $\mathcal{H}_{n,d}$.

Theorem 1.1. *Let $n \geq 2$ and $d \geq 2$. If $A \subseteq \mathbb{R}^n$ is locally bilaterally well approximated by $\mathcal{H}_{n,d}$, then we can write A as a disjoint union,*

$$A = A_1 \cup \cdots \cup A_d \quad (i \neq j \implies A_i \cap A_j = \varnothing),$$

with the following properties:

(i) *For all $1 \leq k \leq d$, we have $x \in A_k$ if and only if x is an $\mathcal{F}_{n,k}$ point of A.*

(ii) *For all $1 \leq k \leq d$, the set $U_k := A_1 \cup \cdots \cup A_k$ is relatively open in A.*

(iii) *For all $1 \leq k \leq d$, the set U_k is locally bilaterally well approximated by $\mathcal{H}_{n,k}$.*

(iv) *For all $2 \leq k \leq d$, the set A is locally bilaterally well approximated along A_k by $\mathcal{F}_{n,k}$; i.e., $\limsup_{r\downarrow 0} \sup_{x\in K} \Theta_A^{\mathcal{F}_{n,k}}(x, r) = 0$ for every compact set $K \subseteq A_k$.*

(v) *For all $1 \leq l < k \leq d$, the set U_l is relatively open in U_k and $A_{l+1} \cup \cdots \cup A_k$ is relatively closed in U_k.*

(vi) *The set A_1 is relatively dense in A; i.e., $\bar{A}_1 \cap A = A$.*

If, in addition, A is closed and nonempty, then

(vii) *A has upper Minkowski dimension and Hausdorff dimension $n-1$; and,*

(viii) *$A \setminus A_1 = A_2 \cup \cdots \cup A_d$ has upper Minkowski dimension at most $n-2$.*

Remark 1.2. If $\Sigma_p \in \mathcal{H}_{n,d}$, then Σ_p is locally bilaterally well approximated by $\mathcal{H}_{n,d}$, simply because $\Theta_{\Sigma_p}^{\mathcal{H}_{n,d}}(x, r) = 0$ for all $x \in \Sigma_p$ and $r > 0$. Since $A = \Sigma_p$ corresponding to $p(x_1, \ldots, x_n) = x_1 x_2$ has $A_2 = \{0\}^2 \times \mathbb{R}^{n-2}$, we see that the dimension bounds on $A \setminus A_1$ in Theorem 1.1 hold by example, and thus, are generically the best possible.

Remark 1.3. Note that A_1 is nonempty if A is nonempty by (vi), A_1 is locally closed if A is closed by (ii), and A_1 is locally Reifenberg flat with vanishing constant by (iii). Therefore, by Reifenberg's topological disk theorem (e.g., see [Reifenberg 1960] or [David and Toro 2012]), A_1 admits local bi-Hölder parametrizations by open subsets of $\mathbb{R}^{n-1}$ with bi-Hölder exponents arbitrarily close to 1 provided that A is closed and nonempty. However, we emphasize that while A_1 always has Hausdorff dimension $n-1$ under these conditions, A_1 may potentially have locally infinite $(n-1)$-dimensional Hausdorff measure or may even be purely unrectifiable (e.g., see [David and Toro 1999]).

The proof of Theorem 1.1 uses a general structure theorem for Reifenberg-type sets, developed in [Badger and Lewis 2015], as well as uniform Minkowski content estimates for the zero and singular sets of harmonic polynomials from [Naber and Valtorta 2014]. A Reifenberg-type set is a set $A \subseteq \mathbb{R}^n$ that admits uniform local bilateral approximations by sets in a cone $\mathcal{S}$ of model sets in $\mathbb{R}^n$. In the present setting, the role of the model sets $\mathcal{S}$ is played by $\mathcal{H}_{n,d}$. For background on the theory of local set approximation and a summary of results from [Badger and Lewis 2015], we refer the reader to Appendix A. The core geometric result at the heart of Theorem 1.1 is the following property of zero sets of harmonic polynomials: $\mathcal{H}_{n,k}$ points can be detected in zero sets of harmonic polynomials of degree d $(1 \le k \le d)$ by finding a *single*, sufficiently good approximation at a coarse scale. The precise statement is as follows.

Theorem 1.4. *For all $n \ge 2$ and $1 \le k < d$, there exists a constant $\delta_{n,d,k} > 0$, depending only on n, d, and k, such that for any harmonic polynomial $p : \mathbb{R}^n \to \mathbb{R}$ of degree d and, for any $x \in \Sigma_p$,*

$$\begin{aligned} \partial^\alpha p(x) = 0 \quad \text{for all } |\alpha| \le k \quad &\iff \quad \Theta^{\mathcal{H}_{n,k}}_{\Sigma_p}(x,r) \ge \delta_{n,d,k} \quad \text{for all } r > 0, \\ \partial^\alpha p(x) \ne 0 \quad \text{for some } |\alpha| \le k \quad &\iff \quad \Theta^{\mathcal{H}_{n,k}}_{\Sigma_p}(x,r) < \delta_{n,d,k} \quad \text{for some } r > 0. \end{aligned}$$

Moreover, there exists a constant $C_{n,d,k} > 1$ depending only on n, d, and k such that

$$\Theta^{\mathcal{H}_{n,k}}_{\Sigma_p}(x,r) < \delta_{n,d,k} \quad \text{for some } r > 0 \quad \Longrightarrow \quad \Theta^{\mathcal{H}_{n,k}}_{\Sigma_p}(x,sr) < C_{n,d,k}\, s^{1/k} \quad \text{for all } s \in (0,1). \tag{1-2}$$

In particular, applying (1-2) with $\Sigma_p \in \mathcal{H}_{n,d}$ and $x = 0$, we obtain the following property.

Corollary 1.5. *In the language of Definition A.12, $\mathcal{H}_{n,k}$ points are detectable in $\mathcal{H}_{n,d}$.*

Remark 1.6. The reader may recognize (1-2) as an "improvement-type lemma", which is often obtained as a consequence of a monotonicity formula or a blow-up argument. Here this improvement result states that at every $\mathcal{H}_{n,k}$ point in the zero set Σ_p of a harmonic polynomial of degree $d > k$, the zero set Σ_p resembles the zero set of a harmonic polynomial of degree at most k at scale r *with increasing certainty* as $r \downarrow 0$. In fact, (1-2) yields a precise rate of convergence for the approximation number $\Theta^{\mathcal{H}_{n,k}}_{\Sigma_p}(x,sr)$ as s goes to 0 provided $\Theta^{\mathcal{H}_{n,k}}_{\Sigma_p}(x,r)$ is small enough. However, we would like to emphasize that the proof of Theorem 1.1 does not require monotone convergence nor a definite rate of convergence of the blowups $(A-x)/r$ of the set A as $r \downarrow 0$. Rather, the proof of Theorem 1.1 relies only on the fact that the *pseudotangents* $T = \lim_{i\to\infty}(A-x_i)/t_i$ of A at x (along sequences $x_i \to x$ in A and $t_i \downarrow 0$) satisfy (1-2). The authors expect that both this improvement-type lemma as well as the way in which it is applied in the proof of Theorem 1.1 should be useful in other situations where questions about the structure and size of sets with singularities arise.

In the special case when $k = 1$, Theorem 1.4 first appeared in [Badger 2013, Theorem 1.4]. The proof of the general case, given in Sections 2–4 below, follows the same guidelines, but requires more sophisticated estimates. In particular, in Section 3, we establish uniform growth and size estimates for harmonic polynomials of bounded degree. Of some note, we prove that harmonic polynomials of bounded degree satisfy a Łojasiewicz-type inequality with uniform constants (see Theorem 3.1). These estimates are essential to show that the approximability $\Theta^{\mathcal{H}_{n,k}}_{\Sigma_p}(x,r)$ of a zero set $\Sigma_p \in \mathcal{H}_{n,d}$ is controlled from above by the relative size $\hat{\zeta}_k(p,x,r)$ of the terms of degree at most k appearing in the Taylor expansion for p at x (see Definition 2.3 and Lemma 4.1).

Applied to harmonic polynomials of degree at most d, [Naber and Valtorta 2014, Theorem A.3] says that

$$\operatorname{Vol}\left(\left\{x \in B\left(0, \tfrac{1}{2}\right) : \operatorname{dist}(x, \Sigma_p) \le r\right\}\right) \le (C(n)d)^d\, r \quad \text{for all } \Sigma_p \in \mathcal{H}_{n,d}, \tag{1-3}$$

and [Naber and Valtorta 2014, Theorem 3.37] says that

$$\mathrm{Vol}\left(\left\{x \in B\left(0, \tfrac{1}{2}\right) : \mathrm{dist}(x, S_p) \le r\right\}\right) \le C(n)^{d^2} r^2 \quad \text{for all } S_p \in \mathcal{SH}_{n,d}, \tag{1-4}$$

where $\mathcal{SH}_{n,d} = \{S_p = \Sigma_p \cap |Dp|^{-1}(0) : \Sigma_p \in \mathcal{H}_{n,d},\ 0 \in S_p\}$ denotes the collection of singular sets of nonconstant harmonic polynomials in $\mathbb{R}^n$ of degree at most d that include the origin. The latter estimate is a refinement of [Cheeger et al. 2015], which gave bounds on the volume of the r-neighborhood of the singular set of the form $C(n, d, \varepsilon) r^{2-\varepsilon}$ for all $\varepsilon > 0$. The results of Cheeger, Naber, and Valtorta [Cheeger et al. 2015] and Naber and Valtorta [2014] apply to solutions of a class of second-order elliptic operators with Lipschitz coefficients; we refer the reader to the original papers for the precise class. Estimates (1-3) and (1-4) imply that the zero sets and the singular sets of harmonic polynomials have locally finite $(n-1)$- and $(n-2)$-dimensional Hausdorff measure, respectively. They transfer to the dimension estimates in Theorem 1.1 for sets that are locally bilaterally well approximated by $\mathcal{H}_{n,d}$ using [Badger and Lewis 2015]. See the proof of Theorem 1.1 in Section 5 for details.

Although the singular set of a harmonic polynomial in $\mathbb{R}^n$ generically has dimension at most $n-2$, additional topological restrictions on the zero set may lead to better bounds. In the plane, for example, the zero set of a homogeneous harmonic polynomial of degree k is precisely the union of k lines through the origin, arranged in an equiangular pattern. Hence $\mathbb{R}^2 \setminus \Sigma_p$ has precisely two connected components for $\Sigma_p \in \mathcal{F}_{2,k}$ if and only if $k = 1$, and consequently, the singular set is empty for any harmonic polynomial whose zero set separates $\mathbb{R}^2$ into two connected components. When $n = 3$, Lewy [1977] proved that if $\mathbb{R}^3 \setminus \Sigma_p$ has precisely two connected components for $\Sigma_p \in \mathcal{F}_{3,k}$, then k is necessarily odd. Moreover, Lewy proved the existence of $\Sigma_p \in \mathcal{F}_{3,k}$ that separate $\mathbb{R}^3$ into two connected components for all odd $k \ge 3$; an explicit example due to Szulkin [1978] is $\Sigma_p \in \mathcal{F}_{3,3}$, where

$$p(x, y, z) = x^3 - 3xy^2 + z^3 - \tfrac{3}{2}(x^2 + y^2)z.$$

Starting with $n = 4$, zero sets of even-degree homogeneous harmonic polynomials can also separate $\mathbb{R}^n$ into two components, as shown, e.g., by Lemma 1.7, which we prove in Section 6.

Lemma 1.7. *Let $k \ge 2$, even or odd, and let $q : \mathbb{R}^2 \to \mathbb{R}$ be a homogeneous harmonic polynomial of degree k. For any pair of constants $a, b \ne 0$, consider the homogeneous harmonic polynomial $p : \mathbb{R}^4 \to \mathbb{R}$ of degree k given by*

$$p(x_1, y_1, x_2, y_2) = a\, q(x_1, y_1) + b\, q(x_2, y_2).$$

The zero set Σ_p of p separates $\mathbb{R}^4$ into two components.

Motivated by these examples, it is natural to ask whether it is possible to improve the dimension bounds on the singular set $A \setminus A_1 = A_2 \cup \cdots \cup A_d$ in Theorem 1.1 under additional topological restrictions on A. In this direction, we prove the following result in Section 6 below.

Theorem 1.8. *Let $n \ge 2$ and $d \ge 2$. Let $A \subseteq \mathbb{R}^n$ be a closed set that is locally bilaterally well approximated by $\mathcal{H}_{n,d}$. If $\mathbb{R}^n \setminus A = \Omega^+ \cup \Omega^-$ is a union of complimentary NTA domains Ω^+ and Ω^-, then*

(i) *$A \setminus A_1 = A_2 \cup \cdots \cup A_d$ has upper Minkowski dimension at most $n - 3$;*

(ii) *the "even singular set" $A_2 \cup A_4 \cup A_6 \cup \cdots$ has Hausdorff dimension at most $n - 4$.*

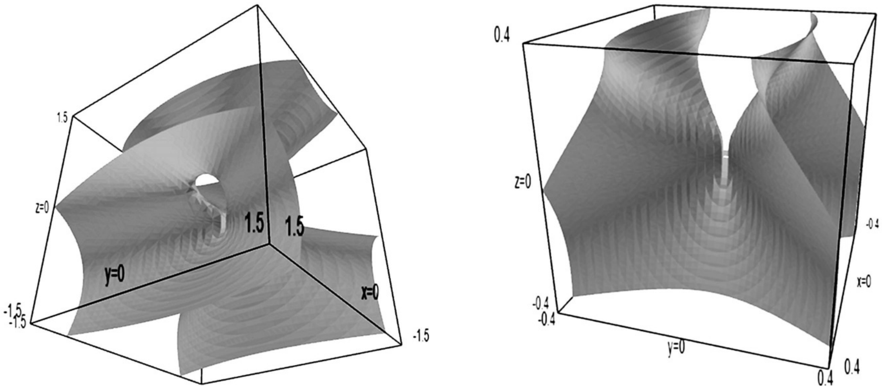

Figure 1. Select views of Σ_p, $p(x, y, z) = x^2 - y^2 + z^3 - 3x^2 z$, which separates $\mathbb{R}^3$ into two components *and* has a cusp at the origin.

NTA domains, or *nontangentially accessible domains*, were introduced by Jerison and Kenig [1982] to study the boundary behavior of harmonic functions in dimensions three and above. We defer their definition to Section 6. However, let us mention in particular that NTA domains satisfy a quantitative strengthening of path connectedness called the *Harnack chain condition*. This property guarantees that A appearing in Theorem 1.8 may be locally bilaterally well approximated by zero sets Σ_p of harmonic polynomials such that $\mathbb{R}^n \setminus \Sigma_p$ has two connected components. Without the Harnack chain condition, this property may fail, as in the following example by Logunov and Malinnikova [2015].

Example 1.9. Consider the harmonic polynomial $p(x, y, z) = x^2 - y^2 + z^3 - 3x^2 z$ from [Logunov and Malinnikova 2015, Example 5.1]. In that paper, they also show that $\mathbb{R}^n \setminus \Sigma_p = \Omega^+ \cup \Omega^-$ is the union of two domains, but remark that Ω^+ and Ω^- fail the Harnack chain condition, and thus, Ω^+ and Ω^- are not NTA domains (see Figure 1). Using Lemma 4.3 below, it can be shown that Σ_p has a unique tangent set at the origin (see Definition A.5 in Appendix A), given by Σ_q, where $q(x, y, z) = x^2 - y^2$. Note that Σ_q divides $\mathbb{R}^3$ into four components. However, if the set Σ_p is locally bilaterally well approximated by some closed class $\mathcal{S} \subseteq \mathcal{H}_{n,d}$, then $\Sigma_q \in \mathcal{S}$ by Theorem A.11.

Remark 1.10. It can be shown that $\mathbb{R}^n \setminus \Sigma_p = \Omega^+ \cup \Omega^-$ is a union of complementary NTA domains and Σ_p is smooth except at the origin when $p(x, y, z)$ is Szulkin's polynomial or when $p(x_1, y_1, x_2, y_2)$ is any polynomial from Lemma 1.7. Thus, the upper bounds given in Theorem 1.8 are generically the best possible. The reason that we obtain an upper Minkowski dimension bound on the full singular set $A \setminus A_1$, but only obtain a Hausdorff dimension bound on the even singular set $A_2 \cup A_4 \cup \cdots$ is that the former is always closed when A is closed, but we only know that the latter is F_σ when A is closed (see the proof of Theorem 1.8).

The improved dimension bounds on $A \setminus A_1$ in Theorem 1.8 require a refinement of (1-4) for $\Sigma_p \in \mathcal{H}_{n,d}$ that separate $\mathbb{R}^n$ into complementary NTA domains, whose existence was postulated in [Badger and Lewis 2015, Remark 9.5]. Using the quantitative stratification machinery introduced in [Cheeger et al. 2015], we demonstrate that near its singular points, a zero set $\Sigma_p \in \mathcal{H}_{n,d}$ with the separation property does not resemble $\Sigma_h \times \mathbb{R}^{n-2}$ for any $\Sigma_h \in \mathcal{F}_{2,k}$, $2 \le k \le d$. This leads us to a version of (1-4) with right-hand side $C(n, d, \varepsilon) r^{3-\varepsilon}$ for all $\varepsilon > 0$ and thence to $\overline{\dim}_M A \setminus A_1 \le n-3$ using [Badger and Lewis 2015]. In addition, we show that at "even-degree" singular points, a zero set Σ_p with the separation property, does not resemble $\Sigma_h \times \mathbb{R}^{n-3}$ for any $\Sigma_h \in \mathcal{F}_{3,2k}$, $2 \le 2k \le d$. This leads us to the bound $\dim_H \Gamma_2 \cup \Gamma_4 \cup \cdots \le n-4$. See the proof of Theorem 1.8 in Section 6 for details.

In the last section of the paper, Section 7, we specialize Theorems 1.1 and 1.8 to the setting of two-phase free boundary problems for harmonic measure mentioned above, which motivated our investigation. This includes the case that $A = \partial\Omega$ is the boundary of a 2-sided NTA domain $\Omega \subset \mathbb{R}^n$ whose interior harmonic measure ω^+ and exterior harmonic measure ω^- are mutually absolutely continuous and have Radon–Nikodym derivative $f = d\omega^-/d\omega^+$ satisfying $\log f \in C(\partial\Omega)$ or $\log f \in \mathrm{VMO}(d\omega^+)$.

2. Relative size of the low-order part of a polynomial

Given a polynomial $p(x) = \sum_{|\alpha| \le d} c_\alpha x^\alpha$ in $\mathbb{R}^n$, define the *height* by $H(p) = \max_{|\alpha|\le d} |c_\alpha|$; i.e., the height of p is the maximum in absolute value of the coefficients of p. The following lemma is an instance of the equivalence of norms on finite-dimensional vector spaces.

Lemma 2.1. *$H(p) \approx \|p\|_{L^\infty(B(0,1))}$ for every polynomial $p : \mathbb{R}^n \to \mathbb{R}$ of degree at most d, where the implicit constants depend only on n and d.*

Below we will need the following easy consequence of Lemma 2.1.

Corollary 2.2. *If $p \equiv p_d + \cdots + p_0$, where each $p_i : \mathbb{R}^n \to \mathbb{R}$ is zero or a homogeneous polynomial of degree i, then $\|p\|_{L^\infty(B(0,1))} \approx \sum_{i=0}^d H(p_i)$, where the implicit constants depend only on n and d.*

Proof. On one hand,

$$\|p\|_{L^\infty(B(0,1))} \le \sum_{i=0}^{d} \|p_i\|_{L^\infty(B(0,1))} \lesssim \sum_{i=0}^{d} H(p_i)$$

by Lemma 2.1 (applied $d+1$ times). On the other hand, the assumption that each p_i is zero or homogeneous of degree i ensures that $H(p) = \max_i H(p_i)$. Hence

$$\sum_{i=0}^{d} H(p_i) \le (d+1) H(p) \lesssim \|p\|_{L^\infty(B(0,1))}$$

by Lemma 2.1, again. □

By Taylor's theorem, for any polynomial $p : \mathbb{R}^n \to \mathbb{R}$ of degree $d \ge 1$ and for any $x \in \mathbb{R}^n$, we can write

$$p(x+y) = p_d^{(x)}(y) + p_{d-1}^{(x)}(y) + \cdots + p_0^{(x)}(y) \quad \text{for all } y \in \mathbb{R}^n, \tag{2-1}$$

where each term $p_i^{(x)} : \mathbb{R}^n \to \mathbb{R}$ is an i-homogeneous polynomial, i.e.,

$$p_i^{(x)}(ry) = r^i p_i^{(x)}(y) \quad \text{for all } y \in \mathbb{R}^n \text{ and } r > 0. \tag{2-2}$$

Definition 2.3. Let $p : \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree $d \ge 1$ and let $x \in \mathbb{R}^n$. For all $0 \le k < d$ and $r > 0$, define $\hat{\zeta}_k(p, x, r)$ by

$$\hat{\zeta}_k(p, x, r) = \max_{k<j\le d} \frac{\|p_j^{(x)}\|_{L^\infty(B(0,r))}}{\left\|\sum_{i=0}^k p_i^{(x)}\right\|_{L^\infty(B(0,r))}} \in [0, \infty].$$

Remark 2.4. The function $\hat{\zeta}_k(p, x, r)$ is a variant of the function $\zeta_k(p, x, r)$ appearing in [Badger 2013, Definition 2.1] and defined by

$$\zeta_k(p, x, r) = \max_{j\ne k} \frac{\|p_j^{(x)}\|_{L^\infty(B(0,r))}}{\|p_k^{(x)}\|_{L^\infty(B(0,r))}}.$$

The latter measured the relative size of the degree-k part of a polynomial compared to its parts of degree $j \ne k$, while the former measures the relative size of the low-order part of a polynomial, consisting of all terms of degree at most k, compared to its parts of degree $j > k$. We note that $\hat{\zeta}_1(p, x, r)$ and $\zeta_1(p, x, r)$ coincide whenever $x \in \Sigma_p$, the zero set of p.

The next lemma generalizes [Badger 2013, Lemma 2.10], which stated $\zeta_1(p, x, sr) \le s\zeta_1(p, x, r)$ for all $s \in (0, 1)$, for all polynomials $p : \mathbb{R}^n \to \mathbb{R}$, for all $x \in \Sigma_p$, and for all $r > 0$.

Lemma 2.5 (change of scales lemma). *For all polynomials $p : \mathbb{R}^n \to \mathbb{R}$ of degree $d \ge 1$, for all $0 \le k < d$, for all $x \in \mathbb{R}^n$ and for all $r > 0$,*

$$s^d\, \hat{\zeta}_k(p, x, r) \lesssim \hat{\zeta}_k(p, x, sr) \lesssim s\, \hat{\zeta}_k(p, x, r) \quad \textit{for all } s \in (0, 1),$$

where the implicit constants depends only on n and d.

Proof. Let $p : \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree $d \ge 1$, let $x \in \mathbb{R}^n$, and let $0 \le k < d$. Write $\tilde{p} = p_k^{(x)} + \cdots + p_0^{(x)}$ for the low-order part of p at x. Then, by repeated use of Corollary 2.2 and the i-homogeneity of each $p_i^{(x)}$, we have that for all $r > 0$ and $s \in (0, 1)$,

$$\begin{aligned}\|\tilde{p}\|_{L^\infty(B(0,sr))} &= \left\|\sum_{i=0}^k p_i^{(x)}(sr\,\cdot)\right\|_{L^\infty(B(0,1))} \gtrsim \sum_{i=0}^k H(p_i^{(x)}(sr\,\cdot)) \gtrsim \sum_{i=0}^k s^i H(p_i^{(x)}(r\,\cdot)) \\ &\gtrsim s^k \sum_{i=0}^k H(p^{(x)}(r\,\cdot)) \gtrsim s^k \left\|\sum_{i=0}^k p^{(x)}(r\,\cdot)\right\|_{L^\infty(B(0,1))} \gtrsim s^k \|\tilde{p}\|_{L^\infty(B(0,r))},\end{aligned} \tag{2-3}$$

where the implicit constants depend on only n and k. It immediately follows that

$$\hat{\zeta}_k(p, x, sr) = \max_{k<j\le d} \frac{\|p_j^{(x)}\|_{L^\infty(B(0,sr))}}{\|\tilde{p}\|_{L^\infty(B(0,sr))}} \lesssim \max_{k<j\le d} s^{j-k} \frac{\|p_j^{(x)}\|_{L^\infty(B(0,r))}}{\|\tilde{p}\|_{L^\infty(B(0,r))}} \lesssim s\, \hat{\zeta}_k(p, x, r),$$

where the implied constant depends only on n and k, and therefore, may be chosen to only depend on n and d. The other inequality follows similarly and is left to the reader. □

We end with a statement about the joint continuity of $\hat{\zeta}_k(p, x, r)$. Lemma 2.7 follows from elementary considerations; for some sample details, the reader may consult the proof of an analogous statement for $\zeta_k(p, x, r)$ in [Badger 2013, Lemma 2.8].

Definition 2.6. A sequence of polynomials $(p^i)_{i=1}^{\infty}$ in $\mathbb{R}^n$ converges *in coefficients* to a polynomial p in $\mathbb{R}^n$ if $d = \max_i \deg p^i < \infty$ and $H(p - p^i) \to 0$ as $i \to \infty$.

Lemma 2.7. *For every $k \geq 0$, the function $\hat{\zeta}_k(p, x, r)$ is jointly continuous in p, x, and r. That is,*

$$\hat{\zeta}_k(p^i, x_i, r_i) \to \hat{\zeta}_k(p, x, r)$$

whenever $\deg p > k$, $p^i \to p$ *in coefficients*, $x_i \to x \in \mathbb{R}^n$, *and* $r_i \to r \in (0, \infty)$.

3. Growth estimates for harmonic polynomials

We need several estimates on the growth of nonconstant harmonic polynomials of degree at most k. The main result of this section is the following uniform Łojasiewicz inequality for harmonic polynomials of bounded degree.

Theorem 3.1 (Łojasiewicz inequality for harmonic polynomials). *For all $n \geq 2$ and $k \geq 1$, there exists a constant $c = c(n, k) > 0$ with the following property. If $p : \mathbb{R}^n \to \mathbb{R}$ is a nonconstant harmonic polynomial of degree at most k and $x_0 \in \Sigma_p$, then*

$$|p(z)| \geq c\|p\|_{L^\infty(B(x_0,1))} \operatorname{dist}(z, \Sigma_p)^k \quad \textit{for all } z \in B\big(x_0, \tfrac{1}{2}\big). \tag{3-1}$$

Remark 3.2. Łojasiewicz [1959] proved the remarkable result that if f is a real analytic function on $\mathbb{R}^n$ and $x_0 \in \Sigma_f$ (the zero set of f), then there exist constants $C, \varepsilon, m > 0$ such that

$$|f(z)| \geq C \operatorname{dist}(z, \Sigma_f)^m \quad \text{for all } z \in B(x_0, \varepsilon).$$

The smallest possible m is called the *Łojasiewicz exponent* of f at x_0. It is perhaps a surprising fact that the Łojasiewicz exponent of a polynomial can exceed the degree of the polynomial. Bounding the Łojasiewicz exponent from above is a difficult problem in algebraic geometric; see, e.g., [Kollár 1999; Phạm 2012]. The content of Theorem 3.1 over the general form of the Łojasiewicz inequality is the tight bound on the Łojasiewicz exponent and uniformity of the constant c in (3-1) across all harmonic polynomials of bounded degree.

The key tools that we use in this section are Almgren's frequency formula and Harnack's inequality for positive harmonic functions. Let us now recall the definition of the former.

Definition 3.3. Let $f \in H^1_{\mathrm{loc}}(\mathbb{R}^n)$ and let

$$x_0 \in \Sigma_f = \{x \in \mathbb{R}^n : f(x) = 0\}.$$

For all $r > 0$, define the quantities $H(r, x_0, f)$ and $D(r, x_0, f)$ by

$$H(r, x_0, f) = \int_{\partial B(x_0,r)} f^2 \, d\sigma \quad \text{and} \quad D(r, x_0, f) = \int_{B(x_0,r)} |\nabla f|^2 \, dx.$$

Then the *frequency function* $N(r, x_0, f)$ is defined by

$$N(r, x_0, f) = \frac{r D(r, x_0, f)}{H(r, x_0, f)} \quad \text{for all } r > 0.$$

Almgren [1979] introduced the frequency function. It is a simple matter to show that for any harmonic polynomial p, we have $N(r, x_0, p) \le \deg p$. When f is any harmonic function, not necessarily a polynomial, Almgren proved that $N(r, x_0, f)$ is absolutely continuous in r and monotonically decreasing as $r \downarrow 0$, and moreover, $\lim_{r\downarrow 0} N(r, x_0, f)$ is the order to which f vanishes at x_0. It can also be verified that

$$\frac{d}{dr}\log\left(\frac{H(r, x_0, f)}{r^{n-1}}\right) = 2\frac{N(r, x_0, f)}{r}. \tag{3-2}$$

Integrating (3-2) and invoking the monotonicity of $N(r, x_0, f)$ in r, one can prove the following doubling property. For a proof of Lemma 3.4, see, e.g., [Han 2007, Corollary 1.5]; the result is stated there with $x_0 = 0$ and $R = 1$, but the general case readily follows by observing that $N(R, x_0, f) = N(1, 0, g)$, where $g(x) = f(x_0 + Rx)/R$.

Lemma 3.4. *If f is a harmonic function on $B(x_0, R)$, then for all $r \in \left(0, \tfrac{1}{2}R\right)$,*

$$⨍_{B(x_0,2r)} f^2\,dx \le 2^{2N(R,x_0,f)-1} ⨍_{B(x_0,r)} f^2\,dx. \tag{3-3}$$

Corollary 3.5. *For all $n \ge 2$ and $k \ge 1$, there exists a constant $C > 0$ such that if $p : \mathbb{R}^n \to \mathbb{R}$ is a harmonic polynomial of degree at most k, $x_0 \in \mathbb{R}^n$, and $r > 0$, then*

$$⨍_{B(x_0,2r)} p^2\,dx \le C ⨍_{B(x_0,r)} p^2\,dx \quad \text{and} \quad \sup_{B(x_0,r)} p^2 \le 2^n C ⨍_{B(x_0,r)} p^2\,dx. \tag{3-4}$$

Proof. The first inequality in (3-4) is an immediate consequence of Lemma 3.4 and the well known fact that $N(r, x_0, p) \le \deg p$ for every harmonic polynomial p.

To establish the second inequality in (3-4), first note that $B(z, r) \subseteq B(x_0, 2r)$ for all $z \in B(x_0, r)$. By the mean value property of harmonic functions and the first inequality,

$$p(z)^2 = \left(⨍_{B(z,r)} p\,dx\right)^2 \le ⨍_{B(z,r)} p^2\,dx \le 2^n ⨍_{B(x_0,2r)} p^2\,dx \le 2^n C ⨍_{B(x_0,r)} p^2\,dx.$$

This establishes (3-4). □

Next, as an application of Corollary 3.5 and Harnack's inequality, we show that $p(z)$ is relatively large when z is far enough away from Σ_p.

Lemma 3.6. *For all $n \ge 2$ and $k \ge 1$, there exists a constant $c > 0$ such that if $p : \mathbb{R}^n \to \mathbb{R}$ is a harmonic polynomial of degree at most k, $z \in \mathbb{R}^n$, and $x_0 \in \Sigma_p$ is any point such that $\rho := \operatorname{dist}(z, \Sigma_p) = |z - x_0|$, then*

$$|p(z)| \ge c \sup_{B(x_0,\rho)} |p|. \tag{3-5}$$

Proof. Let $n \ge 2$ and $k \ge 1$ be given, and let $p : \mathbb{R}^n \to \mathbb{R}$ be a harmonic polynomial of degree at most k. Since the conclusion is trivial for all $z \in \Sigma_p$, we may assume $z \in \mathbb{R}^n \setminus \Sigma_p$. Without loss of generality, we may further assume that p is positive in $B(z, \rho)$, where $\rho = \operatorname{dist}(z, \Sigma_p)$. By Harnack's inequality for positive harmonic functions (e.g., see [Axler et al. 2001, Theorem 3.4]), there exists a constant $A = A(n) > 0$ such that

$$p(z)^2 \geq A \sup_{B(z,\rho/2)} p^2 \geq A \fint_{B(z,\rho/2)} p^2\,dx.$$

Pick $x_0 \in \Sigma_p$ such that $\rho = |z - x_0|$ and note that $B(z, 2\rho) \supseteq B(x_0, \rho)$. Hence, by two applications of the first inequality in Corollary 3.5 and then by the second inequality,

$$\fint_{B(z,\rho/2)} p^2\,dx \geq C^2 \fint_{B(z,2\rho)} p^2\,dx \geq 2^{-n} C^2 \fint_{B(x_0,\rho)} p^2\,dx \geq 4^{-n} C \sup_{B(x_0,\rho)} p^2.$$

Combining the displayed equations, we conclude that (3-5) holds with $c = 2^{-n}\sqrt{AC}$. □

We can now obtain the Łojasiewicz inequality for harmonic polynomials (Theorem 3.1) by combining Lemma 3.6 with the estimate (2-3) from the proof of Lemma 2.5.

Proof of Theorem 3.1. Let $n \geq 2$ and $k \geq 1$ be given. Suppose that $p : \mathbb{R}^n \to \mathbb{R}$ is a nonconstant harmonic polynomial of degree at most k, and without loss of generality, assume that $0 \in \Sigma_p$ (the origin will play the role of x_0 in the statement of the theorem). Fix $z \in B\big(0, \frac{1}{2}\big)$ and choose $x_0 \in \Sigma_p$ to be any point such that $\rho := |z - x_0| = \operatorname{dist}(z, \Sigma_p)$. Note that $\rho < \frac{1}{2}$, since $0 \in \Sigma_p$ and $z \in B\big(0, \frac{1}{2}\big)$. On one hand, by Lemma 3.6,

$$|p(z)| \gtrsim \sup_{B(x_0,\rho)} |p|.$$

On the other hand, applying (2-3) with $r = 2$ and $s = \frac{1}{2}\rho$ (this is fine as $s < 1$),

$$\sup_{B(x_0,\rho)} |p| \gtrsim \rho^k \sup_{B(x_0,2)} |p| \geq \rho^k \|p\|_{L^\infty(B(0,1))}.$$

Here all implicit constants depend on at most n and k. The inequality (3-1) immediately follows by combining the displayed equations (and recalling the definition of ρ). □

As we work separately with the sets $\{p > 0\}$ and $\{p < 0\}$ below, it is important for us to know that $\sup p^+$ and $\sup p^-$ are comparable in any ball centered on Σ_p.

Lemma 3.7. *For all $n \geq 2$ and $k \geq 1$, there exists a constant $C > 1$ such that if $p : \mathbb{R}^n \to \mathbb{R}$ is a nonconstant harmonic polynomial of degree at most k, then*

$$C^{-1} \sup_{B(x_0,r)} p^+ \leq \sup_{B(x_0,r)} p^- \leq C \sup_{B(x_0,r)} p^+ \quad \text{for all } x_0 \in \Sigma_p \text{ and } r > 0. \tag{3-6}$$

Proof. Let $M^\pm = \sup_{B(x_0,r)} p^\pm$, and assume without loss of generality that $M^+ \geq M^-$. The argument now splits into two cases.

Case I. Assume that $\sup_{B(x_0,r/2)} |p| = \sup_{B(x_0,r/2)} p^-$. Then by the estimate (2-3) in the proof of Lemma 2.5,

$$M^- \geq \sup_{B(x_0,r/2)} p^- = \sup_{B(x_0,r/2)} |p| \gtrsim \sup_{B(x_0,r)} |p| = M^+,$$

where the implicit constant depends only on n and k.

Case II. Assume that $\sup_{B(x_0,r/2)} |p| = \sup_{B(x_0,r/2)} p^+$. Note that $p+2M^-$ is a positive harmonic function in $B(x_0, r)$. Thus, by Harnack's inequality,

$$2M^- = p(x_0) + 2M^- \geq a \sup_{B(x_0,r/2)} (p+2M^-) = a \sup_{B(x_0,r/2)} (p^+ + 2M^-), \tag{3-7}$$

where $a = a(n) > 0$. We now argue as in Case I. By (2-3),

$$\sup_{B(x_0,r/2)} p^+ = \sup_{B(x_0,r/2)} |p| \gtrsim \sup_{B(x_0,r)} |p| = M^+,$$

where the implicit constant depends only on n and k. Combining the displayed equations, we conclude that $M^- \gtrsim M^+$. □

Finally, we record a technical observation that will be needed in Section 6.

Lemma 3.8. *Let $n \geq 2$ and let $k \geq 1$. If $p : \mathbb{R}^n \to \mathbb{R}$ is a harmonic polynomial of degree at most k, then $\|p\|_{L^2(B(0,1))} \sim_{n,k} \|p\|_{L^2(\partial B(0,1))}$.*

Proof. The fact that $\|p\|_{L^2(\partial B(0,1))}$ is a norm on the space of harmonic polynomials follows from the maximum principle for harmonic functions. Thus, the equivalence of $\|p\|_{L^2(B(0,1))}$ and $\|p\|_{L^2(\partial B(0,1))}$ for harmonic polynomials of bounded degree follows from the equivalence of norms on finite-dimensional vector spaces. □

4. $\mathcal{H}_{n,k}$ points are detectable in $\mathcal{H}_{n,d}$

The next lemma shows that $\hat{\zeta}_k$ (see Definition 2.3 above) controls how close $\Sigma_p \in \mathcal{H}_{n,d}$ is to the zero set of a harmonic polynomial of degree at most k; cf. [Badger 2013, Lemma 4.1]. For the definition of the bilateral approximation number $\Theta_{\Sigma_p}^{\mathcal{H}_{n,k}}(x, r)$, we refer the reader to the Introduction; see (1-1).

Lemma 4.1. *For all $n \geq 2$ and $d \geq 2$, there exists $0 < C < \infty$ such that for every harmonic polynomial $p : \mathbb{R}^n \to \mathbb{R}$ of degree d and for every $1 \leq k < d$,*

$$\Theta_{\Sigma_p}^{\mathcal{H}_{n,k}}(x, r) \leq C\, \hat{\zeta}_k(p, x, r)^{1/k} \quad \textit{for all } x \in \Sigma_p \textit{ and } r > 0. \tag{4-1}$$

Proof. Let $p : \mathbb{R}^n \to \mathbb{R}$ be a harmonic polynomial of degree $d \geq 2$, let $1 \leq k < d$, and let $x \in \Sigma_p$. Write $p(\cdot + x) = p_d^{(x)} + \cdots + p_{k+1}^{(x)} + p_k^{(x)} + \cdots + p_1^{(x)}$, where each $p_i^{(x)} : \mathbb{R}^n \to \mathbb{R}$ is an i-homogeneous polynomial in y with coefficients depending on x. We remark that $x + \Sigma_{p(\cdot+x)} = \Sigma_p$. Now, since p is harmonic, each term $p_i^{(x)}$ is harmonic, as well. Set $\tilde{p} = p_k^{(x)} + \cdots + p_1^{(x)}$, the low-order part of p at x, and note that $\tilde{p}(0) = 0$. If $\tilde{p} \equiv 0$, then $\hat{\zeta}_k(p, x, r) = \infty$ for all $r > 0$ and (4-1) holds trivially. Thus, we may assume that $\tilde{p} \not\equiv 0$, in which case $\Sigma_{\tilde{p}} \in \mathcal{H}_{n,k}$. To prove (4-1), we shall prove a slightly stronger pair of inequalities,

$$r^{-1} \sup_{a \in \Sigma_p \cap B(x,r)} \operatorname{dist}\big(a, (x + \Sigma_{\tilde{p}}) \cap \overline{B(x, r)}\big) \leq C_1\, \hat{\zeta}_k(p, x, r)^{1/k} \tag{4-2}$$

and

$$r^{-1} \sup_{w \in (x+\Sigma_{\tilde{p}}) \cap B(x,r)} \operatorname{dist}(w, \Sigma_p) \leq C_2\, \hat{\zeta}_k(p, x, 2r)^{1/k} \tag{4-3}$$

for some constants C_1 and C_2 that depend only on n, d, and k, and therefore, may be chosen to depend only on n and d. With the help of Lemma 2.5, (4-1) follows immediately from (4-2) and (4-3).

Suppose $\tilde{p}(z)\neq 0$ for some $z\in B(0,r)$ and choose $y\in\Sigma_{\tilde{p}}\cap\overline{B(0,r)}$ such that $\rho:=\operatorname{dist}(z,\Sigma_{\tilde{p}}\cap\overline{B(0,r)})=|z-y|$. We note that $\rho\le r$, since $\tilde{p}(0)=0$, and $B(0,r)\subseteq B(y,2r)$. Hence, by Lemma 3.6,

$$|\tilde{p}(z)|\ge c\|\tilde{p}\|_{L^\infty(B(y,\rho))}\overset{(2\text{-}3)}{\ge} c\left(\frac{\rho}{2r}\right)^k\|\tilde{p}\|_{L^\infty(B(y,2r))}\ge c\left(\frac{\rho}{r}\right)^k\|\tilde{p}\|_{L^\infty(B(0,r))},$$

where at each occurrence c denotes a positive constant determined by n and k. Thus,

$$|p(z+x)|\ge|\tilde{p}(z)|-\sum_{j=k+1}^{d}\|p_j^{(x)}\|_{L^\infty(B(0,r))}\ge c_1\left(\frac{\rho}{r}\right)^k\|\tilde{p}\|_{L^\infty(B(0,r))}-(d-k)\hat{\zeta}_k(p,x,r)\|\tilde{p}\|_{L^\infty(B(0,r))},$$

where $c_1>0$ is a constant depending only on n and k. It follows that $|p(z+x)|>0$ whenever $z\in B(0,r)$ and $\operatorname{dist}(z,\Sigma_{\tilde{p}}\cap\overline{B(0,r)})=\rho>C_1\hat{\zeta}_k(p,x,r)^{1/k}r$, where

$$C_1=\left(\frac{d-k}{c_1}\right)^{1/k}.$$

Consequently, for any $a=z+x\in\Sigma_p\cap B(x,r)$, we have

$$\operatorname{dist}\big(a,(x+\Sigma_{\tilde{p}})\cap\overline{B(x,r)}\big)=\operatorname{dist}\big(z,\Sigma_{\tilde{p}}\cap\overline{B(0,r)}\big)\le C_1\hat{\zeta}_k(p,x,r)^{1/k}r.$$

This establishes (4-2).

Next, suppose that $w\in(x+\Sigma_{\tilde{p}})\cap B(x,r)$, say $w=x+z$ for some $z\in\Sigma_{\tilde{p}}\cap B(0,r)$. Let $\delta<r$ be a fixed scale, to be chosen below. Because $\tilde{p}$ is harmonic, we can locate points $z_\delta^{\pm}\in\partial B(z,\delta)$ such that

$$\tilde{p}(z_\delta^+)=\max_{z'\in\overline{B(z,\delta)}}\tilde{p}(z')>0\quad\text{and}\quad\tilde{p}(z_\delta^-)=\min_{z'\in\overline{B(z,\delta)}}\tilde{p}(z')<0.$$

Thus, by Lemma 3.7,

$$\pm\tilde{p}(z_\delta^{\pm})=|\tilde{p}(z_\delta^{\pm})|\ge c\|\tilde{p}\|_{L^\infty(B(z,\delta))}\overset{(2\text{-}3)}{\ge} c\left(\frac{\delta}{3r}\right)^k\|\tilde{p}\|_{L^\infty(B(z,3r))}\ge c\left(\frac{\delta}{r}\right)^k\|\tilde{p}\|_{L^\infty(B(0,2r))},$$

where at each occurrence $c>0$ depends only on n and k. We conclude that

$$\begin{aligned}\pm p(z_\delta^{\pm}+x)&\ge\pm\tilde{p}(z_\delta^{\pm})-\sum_{j=k+1}^{d}\|p_j^{(x)}\|_{L^\infty(B(0,2r))}\\&\ge c_2\left(\frac{\delta}{r}\right)^k\|\tilde{p}\|_{L^\infty(B(0,2r))}-(d-k)\hat{\zeta}_k(p,x,2r)\|\tilde{p}\|_{L^\infty(B(0,r))}>0\end{aligned}$$

provided that $\delta>C_2\hat{\zeta}_k(p,x,2r)^{1/k}r$, where $C_2=[(d-k)/c_2]^{1/k}$. But we also required $\delta<r$ above. To continue, there are two cases. On one hand, if $C_2\tilde{\zeta}_k(p,x,2r)^{1/k}\ge 1$, then $\Theta_{\Sigma_p}^{\mathcal{H}_{n,k}}(x,r)\le 1\le C_2\tilde{\zeta}_k(p,x,2r)^{1/k}$ holds trivially. On the other hand, suppose that $C_2\tilde{\zeta}_k(p,x,2r)^{1/k}<1$. In this case, pick any $\delta\in(C_2\tilde{\zeta}_k(p,x,2r)^{1/k}r,r)$. Then the estimate above gives $\pm p(z_\delta^{\pm}+x)>0$. In particular, the straight line segment ℓ that connects $z_\delta^+ +x$ to $z_\delta^- +x$ inside $\overline{B(z+x,\delta)}$ must intersect $\Sigma_p\cap\overline{B(z+x,\delta)}$ by the intermediate value theorem and the convexity of ball. Hence $\operatorname{dist}(w,\Sigma_p)=\operatorname{dist}(z+x,\Sigma_p)\le\delta$. Therefore, letting $\delta\downarrow C_2\tilde{\zeta}_k(p,x,2r)^{1/k}$, we obtain (4-3). $\square$

Remark 4.2. In the proof of Lemma 4.1, the harmonicity of p was only used to establish the harmonicity of $\tilde{p}$. Thus, the argument actually yields that $\Theta_{\Sigma_p}^{\mathcal{H}_{n,k}}(x, r) \lesssim_{n,d} \hat{\zeta}_k(p, x, r)$ for all $x \in \Sigma_p$ and for all $r > 0$, whenever $p : \mathbb{R}^n \to \mathbb{R}$ is a polynomial of degree $d > k$ such that $\tilde{p} = p_k^{(x)} + \cdots + p_1^{(x)}$ is harmonic.

The following useful fact facilitates normal families arguments with sequences in $\mathcal{H}_{n,d}$. It is ultimately a consequence of the mean value property of harmonic functions.

Lemma 4.3. *Suppose that $\Sigma_{p_1}, \Sigma_{p_2}, \ldots \in \mathcal{H}_{n,d}$. If $p_i \to p$ in coefficients and $H(p) \neq 0$, then $\Sigma_p \in \mathcal{H}_{n,d}$ and $\Sigma_{p_i} \to \Sigma_p$ in the Attouch–Wets topology (see Appendix A).*

Proof. Suppose that, for each $i \geq 1$, the function $p_i : \mathbb{R}^n \to \mathbb{R}$ is a harmonic polynomial of degree at most d such that $p_i(0) = 0$. Assume that $p_i \to p$ in coefficients and $H(p) \neq 0$. Then $p : \mathbb{R}^n \to \mathbb{R}$ is also a harmonic polynomial of degree at most d such that $p(0) = 0$, because $p_i \to p$ uniformly on compact subsets of $\mathbb{R}^n$, and p is nonconstant, because $H(p) \neq 0$. Hence $\Sigma_p \in \mathcal{H}_{n,d}$. It remains to show that $\Sigma_{p_i} \to \Sigma_p$ in the Attouch–Wets topology, which is metrizable. Thus, it suffices to prove that every subsequence $(\Sigma_{p_{ij}})_{j=1}^{\infty}$ of $(\Sigma_{p_i})_{i=1}^{\infty}$ has a further subsequence $(\Sigma_{p_{ijk}})_{k=1}^{\infty}$ such that $\Sigma_{p_{ijk}} \to \Sigma_p$ in the Attouch–Wets topology.

Fix an arbitrary subsequence $(\Sigma_{p_{ij}})_{j=1}^{\infty}$ of $(\Sigma_{p_i})_{i=1}^{\infty}$. Since $0 \in \Sigma_{p_{ij}}$ for all $j \geq 1$ and the set of closed sets in $\mathbb{R}^n$ containing the origin is sequentially compact, there exists a closed set $F \subseteq \mathbb{R}^n$ containing 0 and a subsequence $(\Sigma_{p_{ijk}})_{k=1}^{\infty}$ of $(\Sigma_{p_{ij}})_{j=1}^{\infty}$ such that $\Sigma_{p_{ijk}} \to F$. We claim that $F = \Sigma_p$. Indeed, on one hand, for any $y \in F$ there exists a sequence $y_k \in \Sigma_{p_{ijk}}$ such that $y_k \to y$; but $p(y) = \lim_{k\to\infty} p_{ijk}(y_k) = \lim_{k\to\infty} 0 = 0$, since $y_k \in \Sigma_{p_{ijk}}$, $p_{ijk} \to p$ uniformly on compact sets, and $y_k \to y$. Hence $y \in \Sigma_p$ for all $y \in F$. That is, $F \subseteq \Sigma_p$. On the other hand, suppose $z \in \Sigma_p$. Since $p(z) = 0$, but $p \not\equiv 0$, for all $r \in (0, 1)$ we can locate points $z_r^{\pm} \in B(z, r)$ such that $p(z_r^+) > 0$ and $p(z_r^-) < 0$ by the mean value theorem for harmonic functions. Because $p_{ijk} \to p$ pointwise, it follows that

$$p_{ijk}(z_r^+) > 0 \quad \text{and} \quad p_{ijk}(z_r^-) < 0$$

for all sufficiently large k depending on r. In particular, by the intermediate value theorem, the straight line segment connecting z_r^+ to z_r^- inside $B(z, r)$ must intersect $\Sigma_{p_{ijk}} \cap B(z, r)$ for all sufficiently large k depending on r. Hence $\operatorname{dist}(z, \Sigma_{p_{ijk}} \cap B(z, 1)) \to 0$ as $k \to \infty$. Ergo, since $\Sigma_{p_{ijk}} \to F$ in the Attouch–Wets topology,

$$\operatorname{dist}(z, F) \leq \liminf_{k\to\infty} \big(\operatorname{dist}(z, \Sigma_{p_{ijk}} \cap B(z, 1)) + \operatorname{ex}(\Sigma_{p_{ijk}} \cap B(z, 1), F)\big) = 0.$$

That is, $z \in F$ for all $z \in \Sigma_p$. Therefore, $\Sigma_p \subseteq F$, and the conclusion follows. $\square$

Corollary 4.4. *For all $n \geq 2$ and $1 \leq k \leq d$, the sets $\mathcal{H}_{n,d}$ and $\mathcal{F}_{n,k}$ are closed subsets of $\mathfrak{C}(0)$ with the Attouch–Wets topology.*

Proof. Suppose $\Sigma_{p_i} \in \mathcal{H}_{n,d}$ for all $i \geq 1$ and $\Sigma_{p_i} \to F$ for some closed set F in $\mathbb{R}^n$. Replacing each p_i by $p_i/H(p_i)$, which leaves Σ_{p_i} unchanged, we may assume $H(p_i) = 1$ for all $i \geq 1$. Hence we can find a polynomial p and a subsequence $(p_{ij})_{j=1}^{\infty}$ of $(p_i)_{i=1}^{\infty}$ such that $p_{ij} \to p$ in coefficients and $H(p) = 1$. Thus, by Lemma 4.3, $\Sigma_p \in \mathcal{H}_{n,d}$ and $\Sigma_{p_{ij}} \to \Sigma_p$. Therefore, $F = \lim_{i\to\infty} \Sigma_{p_i} = \lim_{j\to\infty} \Sigma_{p_{ij}} = \Sigma_p \in \mathcal{H}_{n,d}$. We conclude that $\mathcal{H}_{n,d}$ is closed. Finally, $\mathcal{F}_{n,k}$ is closed by the additional observation that p is homogeneous of degree k whenever p_{ij} is homogeneous of degree k for all j. $\square$

Remark 4.5. For any $\Sigma_p \in \mathcal{H}_{n,d}$ and $\lambda > 0$, the dilate $\lambda\Sigma_p$ is equal to Σ_q, where $q : \mathbb{R}^n \to \mathbb{R}$ is given by $q(x) = p(x/\lambda)$ for all $x \in \mathbb{R}^n$. Since p is a nonconstant polynomial of degree at most d such that $p(0) = 0$, so is q. Also, q is k-homogeneous whenever p is k-homogeneous. Finally, since p is harmonic on $\mathbb{R}^n$, the mean value theorem gives

$$\fint_{B(y,r)} q(x)\,dx = \fint_{B(y,r)} p(x/\lambda)\,dx = \fint_{B(y/\lambda,r/\lambda)} p(x)\,dx = p(y/\lambda) = q(y)$$

for all $y \in \mathbb{R}^n$ and $r > 0$. Thus, since q is continuous, it is also harmonic by the mean value theorem. This shows that $\lambda\Sigma_p \in \mathcal{H}_{n,d}$ for all $\Sigma_p \in \mathcal{H}_{n,d}$ and $\lambda > 0$. Likewise, $\lambda\Sigma_p \in \mathcal{F}_{n,k}$ for all $\Sigma_p \in \mathcal{F}_{n,k}$ and $\lambda > 0$. In other words, $\mathcal{H}_{n,d}$ and $\mathcal{F}_{n,k}$ are cones. Therefore, $\mathcal{H}_{n,d}$ and $\mathcal{F}_{n,k}$ are local approximation classes in the sense of Definition A.7(i). A similar argument shows that $\mathcal{H}_{n,d}$ is translation invariant in the sense that $\Sigma_p - x \in \mathcal{H}_{n,d}$ for all $\Sigma_p \in \mathcal{H}_{n,d}$ and $x \in \Sigma_p$.

The next lemma captures a weak rigidity property of real-valued harmonic functions: the zero set of a real-valued harmonic function determines the relative arrangement of its positive and negative components.

Lemma 4.6. *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ be harmonic functions, and let Σ_f and Σ_g denote the zero sets of f and g, respectively. If $\Sigma_f = \Sigma_g$, then f and g take the same or the opposite sign simultaneously on every connected component of $\mathbb{R}^n \setminus \Sigma_f = \mathbb{R}^n \setminus \Sigma_g$.*

Proof. Since the conclusion is trivial if f is identically zero, we may assume, in addition to the hypothesis, that f is not identically zero. According to [Logunov and Malinnikova 2015, Theorem 1.1], if u and v are harmonic functions defined on a domain $\Omega \subseteq \mathbb{R}^n$ whose zero sets satisfy $\Sigma_v \subseteq \Sigma_u$, then there exists a real-analytic function α in Ω such that $u = \alpha v$. Invoking this fact twice, we obtain that $f = \alpha g = \alpha\beta f$, where α and β are real analytic functions on $\mathbb{R}^n$. Since f is not identically zero, it follows that $\alpha\beta = 1$ on $\mathbb{R}^n$. In particular, $\operatorname{sign}(\alpha) = \pm 1$ on $\mathbb{R}^n$. Therefore, $\operatorname{sign}(f) = \operatorname{sign}(\alpha)\operatorname{sign}(g) = \pm\operatorname{sign}(g)$ on $\mathbb{R}^n$. $\square$

The following lemma indicates that zero sets of homogeneous harmonic polynomials of different degrees are uniformly separated on balls centered at the origin. This answers affirmatively a question posed in [Badger 2013, Remark 4.12].

Lemma 4.7. *For all $n \geq 2$ and $1 \leq j < k$, there exists a constant $\varepsilon > 0$ such that for all $\Sigma_p \in \mathcal{F}_{n,k}$ and $\Sigma_q \in \mathcal{F}_{n,j}$,*

$$\widetilde{\mathrm{D}}^{0,r}[\Sigma_p, \Sigma_q] = \frac{1}{r}\max\Big\{\sup_{x\in\Sigma_p\cap B(0,r)} \operatorname{dist}(x, \Sigma_q), \sup_{y\in\Sigma_q\cap B(0,r)} \operatorname{dist}(y, \Sigma_p)\Big\} \geq \varepsilon \quad \textit{for all } r > 0.$$

Proof. Note that $\lambda\Sigma_p = \Sigma_p$ and $\lambda\Sigma_q = \Sigma_q$ for all $\lambda > 0$ whenever $\Sigma_p \in \mathcal{F}_{n,k}$ and $\Sigma_q \in \mathcal{F}_{n,j}$. Hence $\widetilde{\mathrm{D}}^{0,r}[\Sigma_p, \Sigma_q] = \widetilde{\mathrm{D}}^{0,1}[r^{-1}\Sigma_p, r^{-1}\Sigma_q] = \widetilde{\mathrm{D}}^{0,1}[\Sigma_p, \Sigma_q]$ for all $r > 0$, whenever $n \geq 2$, $1 \leq j < k$, $\Sigma_p \in \mathcal{F}_{n,k}$, and $\Sigma_q \in \mathcal{F}_{n,j}$. Thus, it suffices to prove the claim with $r = 1$.

Assume to the contrary that for some $n \geq 2$ and $1 \leq j < k$ we can find sequences $p_1, p_2, \ldots \in \mathcal{F}_{n,k}$ and $q_1, q_2, \ldots \in \mathcal{F}_{n,j}$ such that

$$\widetilde{\mathrm{D}}^{0,1}[\Sigma_{p_i}, \Sigma_{q_i}] \leq \frac{1}{i} \quad \text{for all } i \geq 1. \tag{4-4}$$

By Corollary 4.4, passing to subsequences (which we relabel), we may assume that there exist $\Sigma_p \in \mathcal{F}_{n,k}$ and $\Sigma_q \in \mathcal{F}_{n,j}$ such that $\Sigma_{p_i} \to \Sigma_p$ and $\Sigma_{q_i} \to \Sigma_q$. Moreover, replacing each p_i and q_i by $p_i/H(p_i)$ and $q_i/H(q_i)$, respectively, and passing to further subsequences (which we again relabel), we may assume that $p_i \to p$ in coefficients and $q_i \to q$ in coefficients, where p and q are homogeneous harmonic polynomials of degrees k and j, respectively. By two applications of the weak quasitriangle inequality (see Appendix A),

$$\begin{aligned}\widetilde{\mathrm{D}}^{0,1/4}[\Sigma_p, \Sigma_q] &\le 2\widetilde{\mathrm{D}}^{0,1/2}[\Sigma_p, \Sigma_{p_i}] + 2\widetilde{\mathrm{D}}^{0,1/2}[\Sigma_{p_i}, \Sigma_q] \\ &\le 2\widetilde{\mathrm{D}}^{0,1/2}[\Sigma_p, \Sigma_{p_i}] + 4\widetilde{\mathrm{D}}^{0,1}[\Sigma_{p_i}, \Sigma_{q_i}] + 4\widetilde{\mathrm{D}}^{0,1}[\Sigma_{q_i}, \Sigma_q]. \end{aligned} \tag{4-5}$$

Letting $i \to \infty$, the first term vanishes since $\Sigma_{p_i} \to \Sigma_p$, the second term vanishes by (4-4), and the third term vanishes since $\Sigma_{q_i} \to \Sigma_q$. Hence $\widetilde{\mathrm{D}}^{0,1/4}[\Sigma_p, \Sigma_q] = 0$, which implies $\Sigma_p \cap B\big(0, \tfrac{1}{4}\big) = \Sigma_q \cap B\big(0, \tfrac{1}{4}\big)$. But Σ_p and Σ_q are cones, so in fact $\Sigma_p = \Sigma_q$. By Lemma 4.6, the functions p and q take the same or the opposite sign simultaneously on every connected component of $\mathbb{R}^n \setminus \Sigma_p = \mathbb{R}^n \setminus \Sigma_q$. Hence either $p(x)q(x) \ge 0$ for all $x \in \mathbb{R}^n$ or $p(x)q(x) \le 0$ for all $x \in \mathbb{R}^n$. It follows that either $\int_{S^{n-1}} pq\, d\sigma > 0$ or $\int_{S^{n-1}} pq\, d\sigma < 0$. This contradicts the fact that homogeneous harmonic polynomials of different degrees are orthogonal in $L^2(S^{n-1})$ (e.g., see [Axler et al. 2001, Proposition 5.9]). □

We now show that $\hat{\zeta}_k$ cannot grow arbitrarily large as $\Theta^{\mathcal{H}_{n,k}}_{\Sigma_p}$ becomes arbitrarily small; cf. [Badger 2013, Proposition 4.8].

Lemma 4.8. *For all $n \ge 2$ and $1 \le k < d$ there is $\delta_{n,d,k} > 0$ with the following property. If $p : \mathbb{R}^n \to \mathbb{R}$ is a harmonic polynomial of degree d and $\Theta^{\mathcal{H}_{n,k}}_{\Sigma_p}(x, r) < \delta_{n,d,k}$ for some $x \in \Sigma_p$ and $r > 0$, then $\hat{\zeta}_k(p, x, r) < \delta^{-1}_{n,d,k}$.*

Proof. Let $n \ge 2$ and $1 \le k < d$ be given. Suppose in order to reach a contradiction that for all $j \ge 1$ there exists a harmonic polynomial $p_j : \mathbb{R}^n \to \mathbb{R}$ of degree d, $x_j \in \Sigma_{p_j}$, and $r_j > 0$ such that $\Theta^{\mathcal{H}_{n,k}}_{\Sigma_{p_j}}(x_j, r_j) < 1/j$, but $\hat{\zeta}_k(p_j, x_j, r_j) \ge j$. Replacing each p_j with $\tilde{p}_j$,

$$\tilde{p}_j(y) = H(p_j)^{-1} \cdot p(r_j(y + x_j)) \quad \text{for all } y \in \mathbb{R}^n;$$

that is, left translating by x_j, dilating by $1/r_j$, and scaling by $1/H(p_j)$, we may assume without loss of generality that $x_j = 0$, $r_j = 1$, and $H(p_j) = 1$ for all $j \ge 1$. Therefore, there exists a sequence $(p_j)_{j=1}^{\infty}$ of harmonic polynomials in $\mathbb{R}^n$ of degree d and height 1 with $p_j(0) = 0$ such that $\Theta^{\mathcal{H}_{n,k}}_{\Sigma_{p_j}}(0, 1) \le 1/j$, and $\hat{\zeta}_k(p_j, 0, 1) \ge j$. Passing to a subsequence, we may assume that $p_j \to p$ in coefficients to some harmonic polynomial $p : \mathbb{R}^n \to \mathbb{R}$ with height 1. By Lemma 4.3, $\Sigma_{p_j} \to \Sigma_p$, as well. On one hand,

$$\Theta^{\mathcal{H}_{n,k}}_{\Sigma_p}\big(0, \tfrac{1}{2}\big) \le 2 \liminf_{j\to\infty} \Theta^{\mathcal{H}_{n,k}}_{\Sigma_{p_j}}(0, 1) = 0. \tag{4-6}$$

(For a primer on the interaction of limits and approximation numbers, see Appendix A.) On the other hand, by Lemma 2.1 and the fact that $\hat{\zeta}_k(p_j, 0, 1) \ge j$, it must be that the height of the polynomial p_j is obtained from the coefficient of some term of p_j of degree at least $k + 1$, provided that j is sufficiently large. In particular, we conclude that p has degree at least $k + 1$. Hence $\hat{\zeta}_k(p, 0, 1)$ is well defined and

$\hat{\zeta}_k(p,0,1)=\lim_{j\to\infty}\hat{\zeta}_k(p_j,0,1)=\infty$ by Lemma 2.7. Thus, the low-order part of p at 0 (that is, the terms of degree at most k) vanishes and p has the form

$$p=p_d^{(0)}+p_{d-1}^{(0)}+\cdots+\cdots+p_i^{(0)},\quad p_i^{(0)}\neq 0 \text{ for some } i\geq k+1. \tag{4-7}$$

We shall now show that (4-6) and (4-7) are incompatible with Lemma 4.7:

By (4-6), there exists $\Sigma_q\in\overline{\mathcal{H}}_{n,k}=\mathcal{H}_{n,k}$ such that $\Sigma_p\cap B\big(0,\tfrac{1}{2}\big)=\Sigma_q\cap B\big(0,\tfrac{1}{2}\big)$, say

$$q=q_k^{(0)}+q_{k-1}^{(0)}+\cdots+q_l^{(0)},\quad q_l^{(0)}\neq 0 \text{ for some } 1\leq l\leq k. \tag{4-8}$$

Choose any sequence $r_m\downarrow 0$ as $m\to\infty$. By (4-7), $r_m^{-i}p(r_m\,\cdot\,)\to p_i^{(0)}$ in coefficients and by (4-8), $r_m^{-l}q(r_m\,\cdot\,)\to q_l^{(0)}$ in coefficients also. Hence $r_m^{-1}\Sigma_p=\Sigma_{r_m^{-i}p(r_m\,\cdot\,)}\to\Sigma_{p_i^{(0)}}\in\mathcal{F}_{n,i}$ and $r_m^{-1}\Sigma_q=\Sigma_{r_m^{-l}p(r_m\,\cdot\,)}\to\Sigma_{q_l^{(0)}}\in\mathcal{F}_{n,l}$ by Lemma 4.3. By the weak quasitriangle inequality (applied twice as in (4-5)),

$$\widetilde{\mathrm{D}}^{0,1}[\Sigma_{p_i^{(0)}},\Sigma_{q_i^{(0)}}]\leq 2\widetilde{\mathrm{D}}^{0,2}[\Sigma_{p_i^{(0)}},r_m^{-1}\Sigma_p]+4\widetilde{\mathrm{D}}^{0,4}[r_m^{-1}\Sigma_p,r_m^{-1}\Sigma_q]+4\widetilde{\mathrm{D}}^{0,4}[r_m^{-1}\Sigma_q,\Sigma_{q_l^{(0)}}].$$

As $m\to\infty$, the first and the last term vanish, because $r_m^{-1}\Sigma_p\to\Sigma_{p_i^{(0)}}$ and $r_m^{-1}\Sigma_q\to\Sigma_{q_l^{(0)}}$, respectively. Thus,

$$\widetilde{\mathrm{D}}^{0,1}[\Sigma_{p_i^{(0)}},\Sigma_{q_l^{(0)}}]\leq\liminf_{m\to\infty}4\widetilde{\mathrm{D}}^{0,4}[r_m^{-1}\Sigma_p,r_m^{-1}\Sigma_q]=\liminf_{m\to\infty}4\widetilde{\mathrm{D}}^{0,4r_m}[\Sigma_p,\Sigma_q]=0,$$

where the ultimate equality holds because $\Sigma_p\cap B\big(0,\tfrac{1}{2}\big)=\Sigma_q\cap B\big(0,\tfrac{1}{2}\big)$ and $4r_m\downarrow 0$. But by Lemma 4.7 $\widetilde{\mathrm{D}}^{0,1}[\Sigma_{p_i^{(0)}},\Sigma_{q_l^{(0)}}]>0$, because $\Sigma_{p_i^{(0)}}\in\mathcal{F}_{n,i}$, $\Sigma_{q_l^{(0)}}\in\mathcal{F}_{n,l}$, and $i>l$. We have reached a contradiction. Therefore, for all $n\geq 2$ and $1\leq k<d$, there exists $j\geq 1$ such that if $p:\mathbb{R}^n\to\mathbb{R}$ is a harmonic polynomial of degree d and $\Theta_{\Sigma_p}^{\mathcal{H}_{n,k}}(x,r)<1/j$ for some $x\in\Sigma_p$ and $r>0$, then $\hat{\zeta}_k(p,x,r)<j$. □

We now have all the ingredients required to prove Theorem 1.4.

Proof of Theorem 1.4. Given $n\geq 2$ and $1\leq k<d$, let $\delta_{n,d,k}>0$ denote the constant from Lemma 4.8. Let $p:\mathbb{R}^n\to\mathbb{R}$ be a harmonic polynomial of degree d and let $x\in\Sigma_p$. Write $\tilde{p}=p_k^{(x)}+\cdots+p_1^{(x)}$ for the part of p of terms of degree at most k, so that $\partial^\alpha p(x)\neq 0$ for some $|\alpha|\leq k$ if and only if $\tilde{p}\not\equiv 0$. On one hand, if $\tilde{p}\not\equiv 0$, then $\hat{\zeta}_k(p,x,1)<\infty$, whence

$$\Theta_{\Sigma_p}^{\mathcal{H}_{n,k}}(x,r)\lesssim_{n,d}\hat{\zeta}_k(p,x,r)^{1/k}\lesssim_{n,d}r^{1/k}\hat{\zeta}_k(p,x,1)^{1/k}\to 0\quad\text{as } r\to 0$$

by Lemmas 4.1 and 2.5. In particular, if $\tilde{p}\not\equiv 0$, then $\Theta_{\Sigma_p}^{\mathcal{H}_{n,k}}(x,r)<\delta_{n,d,k}$ for some $r>0$. On the other hand, if $\Theta_{\Sigma_p}^{\mathcal{H}_{n,k}}(x,r)<\delta_{n,d,k}$ for some $r>0$, then

$$\hat{\zeta}_k(p,x,r)<\delta_{n,d,k}^{-1}<\infty \tag{4-9}$$

by Lemma 4.8, whence $\tilde{p}\not\equiv 0$. Moreover, in this case,

$$\Theta_{\Sigma_p}^{\mathcal{H}_{n,k}}(x,sr)\lesssim_{n,d}\hat{\zeta}_k(p,x,sr)^{1/k}\lesssim_{n,d}s^{1/k}\hat{\zeta}_k(p,x,r)^{1/k}\lesssim_{n,d,k}s^{1/k}\quad\text{for all } s\in(0,1)$$

by Lemmas 4.1 and 2.5, and (4-9). □

Proof of Corollary 1.5. From (1-2) in Theorem 1.4, it immediately follows that $\mathcal{H}_{n,k}$ points are (ϕ,Φ) detectable in $\mathcal{H}_{n,d}$ for $\phi=\min\{\delta_{n,k+1,k},\ldots,\delta_{n,d,k}\}>0$ and some function Φ of the form $\Phi(s)=Cs^{1/k}$ for all $s\in(0,1)$ (see Definition A.12). □

5. Structure of sets locally bilaterally well approximated by $\mathcal{H}_{n,d}$

Now that we know $\mathcal{H}_{n,k}$ points are detectable in $\mathcal{H}_{n,d}$, we may obtain Theorem 1.1 from repeated use of Theorem A.14.

Proof of Theorem 1.1. Let $n \geq 2$ and $d \geq 2$ be given. By Remark 4.5 and Corollary 4.4, $\mathcal{H}_{n,k}$ and $\mathcal{F}_{n,k}$ are closed local approximation classes and $\mathcal{H}_{n,k}$ is also translation invariant for all $k \geq 1$. Thus, we may freely make use of the technology in the last three subsections of Appendix A. Using Definition A.13, Theorem 1.4 yields

$$\mathcal{H}_{n,k} \cap \mathcal{H}_{n,k-1}^{\perp} = \left\{ \Sigma_p \in \mathcal{H}_{n,k} : \liminf_{r \downarrow 0} \Theta_{\Sigma_p}^{\mathcal{H}_{n,k-1}}(0, r) > 0 \right\} = \mathcal{F}_{n,k} \quad \text{for all } k \geq 2.$$

Suppose that $A \subseteq \mathbb{R}^n$ is locally bilaterally well approximated by $\mathcal{H}_{n,d}$ and put $U_d = A$. Since $\mathcal{H}_{n,d-1}$ points are detectable in $\mathcal{H}_{n,d}$ (by Corollary 1.5) and U_d is locally bilaterally well approximated by $\mathcal{H}_{n,d}$, by Theorem A.14 we can write

$$U_d = (U_d)_{\mathcal{H}_{n,d-1}} \cup (U_d)_{\mathcal{H}_{n,d-1}^{\perp}} =: U_{d-1} \cup A_d,$$

where U_{d-1} and A_d are disjoint, U_{d-1} is relatively open in U_d, U_{d-1} is locally bilaterally well approximated by $\mathcal{H}_{n,d-1}$, and U_d is locally bilaterally well approximated *along* A_d by $\mathcal{H}_{n,d} \cap \mathcal{H}_{n,d-1}^{\perp} = \mathcal{F}_{n,d}$, that is, $\limsup_{r\downarrow 0} \sup_{x \in K} \Theta_{U_d}^{\mathcal{F}_{n,d}}(x, r) = 0$ for every compact set $K \subseteq A_d$. In particular, the latter property implies that every $x \in A_d$ is an $\mathcal{F}_{n,d}$ point of U_d by Theorem A.11. Next, since $\mathcal{H}_{n,d-2}$ points are detectable in $\mathcal{H}_{n,d-1}$, we may repeat the argument, mutatis mutandis, to write

$$U_{d-1} = (U_{d-1})_{\mathcal{H}_{n,d-2}} \cup (U_{d-1})_{\mathcal{H}_{n,d-2}^{\perp}} =: U_{d-2} \cup A_{d-1},$$

where U_{d-2} and A_{d-1} are disjoint, U_{d-2} is relatively open in U_{d-1}, U_{d-2} is locally bilaterally well approximated by $\mathcal{H}_{n,d-2}$, U_{d-1} is locally bilaterally well approximated along A_{d-1} by $\mathcal{F}_{n,d-1}$, and every $x \in A_{d-1}$ is an $\mathcal{F}_{n,d-1}$ point of U_{d-1}. In fact, since U_{d-1} is relatively open in U_d, we have U_{d-2} is relatively open in U_d, U_d is locally bilaterally well approximated along A_{d-1} by $\mathcal{F}_{n,d-1}$, and every $x \in A_{d-1}$ is an $\mathcal{F}_{n,d-1}$ point of U_d, as well. After a finite number of repetitions, this argument shows that

$$A = U_d = U_{d-1} \cup A_d = \cdots = U_1 \cup A_2 \cup \cdots \cup A_d,$$

where the sets $U_1, A_2, \ldots, A_d$ are pairwise disjoint, U_1 is relatively open in A, U_1 is locally bilaterally well approximated by $\mathcal{H}_{n,1}$, $U_k = U_1 \cup A_2 \cup \cdots \cup A_k$ is relatively open in A for all $2 \leq k \leq d$, U_k is locally bilaterally well approximated by $\mathcal{H}_{n,k}$ for all $2 \leq k \leq d$, A is locally bilaterally well approximated along A_k by $\mathcal{F}_{n,k}$ for all $2 \leq k \leq d$, and every $x \in A_k$ is an $\mathcal{F}_{n,k}$ point of A for all $2 \leq k \leq d$. Finally, assign $A_1 = U_1$. Since A_1 relatively open in A, A_1 is locally bilaterally well approximated by $\mathcal{H}_{n,1}$, and $\mathcal{H}_{n,1} = \mathcal{F}_{n,1}$, we conclude that every $x \in A_1$ is an $\mathcal{F}_{n,1}$ point of A by Theorem A.11. This verifies (i)–(iv) of Theorem 1.1 and (v) follows immediately from (ii) and (iii).

Next, we want to prove that A_1 is relatively dense in A. Suppose that $x \in A \setminus A_1$, say $x \in A_k$ for some $k \geq 2$. To find points in A_1 nearby x, we will rely on the following fact: by Remark A.15, since

$\mathcal{H}_{n,1}$ points are detectable in $\mathcal{H}_{n,d}$, there exist $\alpha, \beta > 0$ such that

$$\text{if } \Theta_A^{\mathcal{H}_{n,d}}(y, r') < \alpha \text{ for all } 0 < r' \le r \text{ and } \Theta_A^{\mathcal{H}_{n,1}}(y, r) < \beta \text{ for some } y \in A \text{ and } r > 0, \text{ then } y \in A_1. \quad (5\text{-}1)$$

To proceed, since x is an $\mathcal{F}_{n,k}$ point of A and $\mathcal{F}_{n,k}$ is closed, we can find a homogeneous harmonic polynomial $p : \mathbb{R}^n \to \mathbb{R}$ and sequence of scales $r_i \downarrow 0$ such that $r_i^{-1}(\bar{A} - x) \to \Sigma_p$ in the Attouch–Wets topology (Σ_p is a tangent set of $\bar{A}$ at x). Pick any $z \in \Sigma_p$ such that $|Dp|(z) \neq 0$. (That we can always find such a point is evident, because the singular set of a polynomial has dimension at most $n-2$, while $\dim \Sigma_p = n-1$.) Then $\lim_{s\downarrow 0} \Theta_{\Sigma_p}^{\mathcal{H}_{n,1}}(z, s) = 0$ by Theorem 1.4. In particular, there exists $s_1 > 0$ such that

$$\Theta_{\Sigma_p}^{\mathcal{H}_{n,1}}\big(z, \tfrac{3}{2}s_1\big) \le \tfrac{1}{18}\beta. \quad (5\text{-}2)$$

Since $r_i^{-1}(\bar{A} - x) \to \Sigma_p$, there exist $y_i \in \bar{A}$ such that $z_i := (y_i - x)/r_i \to z$. Replacing each y_i with $y_i' \in A$ such that $|y_i' - y_i| \le r_i/i$, say, we may assume without loss of generality that $y_i \in A$ for all i (because $\widetilde{\mathrm{D}}^{0,r}[r_i^{-1}(\bar{A} - y_i'), r_i^{-1}(\bar{A} - y_i)] \le 1/(ir) \to 0$ for all $r > 0$). Necessarily, $y_i \to x$, and thus, there exists $s_2 > 0$ such that

$$\sup_{i\ge 1} \Theta_A^{\mathcal{H}_{n,d}}(y_i, s) \le \tfrac{1}{2}\alpha < \alpha \quad \text{for all } s \le s_2, \quad (5\text{-}3)$$

because A is locally bilaterally well approximated by $\mathcal{H}_{n,d}$. Now, by quasimonotonicity of bilateral approximation numbers (see Lemma A.10) and (5-2),

$$\Theta_{\Sigma_p}^{\mathcal{H}_{n,1}}\big(z_i, \tfrac{1}{2}s_1\big) \le 2t + 2(1+t)\Theta_{\Sigma_p}^{\mathcal{H}_{n,1}}(z, (1+t)s_1) \le 2t + 3\Theta_{\Sigma_p}^{\mathcal{H}_{n,1}}\big(z, \tfrac{3}{2}s_1\big) \le 2t + \tfrac{1}{6}\beta$$

whenever $|z_i - z| \le ts_1 \le \frac{1}{2}s_1$. With $t = |z_i - z|/s_1$, this yields

$$\Theta_{\Sigma_p}^{\mathcal{H}_{n,1}}\big(z_i, \tfrac{1}{2}s_1\big) \le \frac{2|z_i - z|}{s_1} + \tfrac{1}{6}\beta$$

for all i sufficiently large that $|z_i - z| \le \frac{1}{2}s_1$. Hence, for all i sufficiently large that $|z_i - z| < \frac{1}{6}s_1$ $\big($guaranteeing $z \in \Sigma_p \cap B\big(z_i, \frac{1}{6}s_1\big) \neq \varnothing\big)$,

$$\Theta_{r_i^{-1}(\bar{A}-x)}^{\mathcal{H}_{n,1}}\big(z_i, \tfrac{1}{6}s_1\big) \le 3\widetilde{\mathrm{D}}^{z_i, s_1/2}\left[\frac{\bar{A} - x}{r_i}, \Sigma_p\right] + 3\Theta_{\Sigma_p}^{\mathcal{H}_{n,1}}\big(z_i, \tfrac{1}{2}s_1\big) \le 6\widetilde{\mathrm{D}}^{z, s_1}\left[\frac{\bar{A} - x}{r_i}, \Sigma_p\right] + \frac{6|z - z_i|}{s_1} + \tfrac{1}{2}\beta,$$

where we used the weak quasitriangle inequality in the first line and we used the quasimonotonicity of the relative Walkup–Wets distance in the second line (see Lemma A.1). Since $z_i \to z$ and $r_i^{-1}(\bar{A} - x) \to \Sigma_p$, we conclude that

$$\limsup_{i\to\infty} \Theta_A^{\mathcal{H}_{n,1}}\big(y_i, \tfrac{1}{6}r_i s_1\big) = \limsup_{i\to\infty} \Theta_{r_i^{-1}(\bar{A}-x)}^{\mathcal{H}_{n,1}}\big(z_i, \tfrac{1}{6}s_1\big) \le \tfrac{2}{3}\beta < \beta. \quad (5\text{-}4)$$

Note that $\frac{1}{6}r_i s_1 \le s_2$ for all $i \gg 1$, since $r_i \to 0$. Therefore, by (5-1), (5-3), and (5-4), we have $y_i \in A_1$ for all sufficiently large i. Recalling that $y_i \to x$, it follows that $x \in \bar{A}_1$. Since $x \in A \setminus A_1$ was fixed arbitrarily, this proves (vi).

We now aim to prove dimension bounds on A and $A \setminus A_1$ assuming that A is closed and nonempty. Since $\mathcal{H}_{n,d}$ is a closed, translation invariant approximation class and $\mathcal{H}_{n,1}$ points are detectable in $\mathcal{H}_{n,d}$, the set

$$\operatorname{sing}_{\mathcal{H}_{n,1}} \mathcal{H}_{n,d} = \left\{ (\Sigma_p)_{\mathcal{H}_{n,1}^{\perp}} : \Sigma_p \in \mathcal{H}_{n,d} \text{ and } 0 \in (\Sigma_p)_{\mathcal{H}_{n,1}^{\perp}} \right\}$$

is also a local approximation class and $A \setminus A_1$ is locally *unilaterally* well approximated by $\operatorname{sing}_{\mathcal{H}_{n,1}} \mathcal{H}_{n,d}$ by Theorem A.17. By Theorem 1.4, applied with $k = 1$, the class $\operatorname{sing}_{\mathcal{H}_{n,1}} \mathcal{H}_{n,d}$ is precisely the class $\mathcal{SH}_{n,d} = \{S_p = \Sigma_p \cap |Dp|^{-1}(0) : \Sigma_p \in \mathcal{H}_{n,d},\ 0 \in S_p\}$ of all singular sets of nonconstant harmonic polynomials of degree at most d that include the origin. Recall from the Introduction that

$$\operatorname{Vol}\left(\left\{x \in B\left(0, \tfrac{1}{2}\right) : \operatorname{dist}(x, \Sigma_p) \le r\right\}\right) \le (C(n)d)^d\, r \quad \text{for all } \Sigma_p \in \mathcal{H}_{n,d},$$
$$\operatorname{Vol}\left(\left\{x \in B\left(0, \tfrac{1}{2}\right) : \operatorname{dist}(x, S_p) \le r\right\}\right) \le C(n)^{d^2} r^2 \quad \text{for all } S_p \in \mathcal{SH}_{n,d}$$

by work of Naber and Valtorta [2014]. Using an elementary Vitali covering argument (e.g., see [Mattila 1995, (5.4) and (5.6)]), it follows that $\mathcal{H}_{n,d}$ has an $(n-1, C(n,d), 1)$ covering profile and $\mathcal{SH}_{n,d}$ has an $(n-2, C(n,d), 1)$ covering profile in the sense of Definition A.19.

Assume that A is a nonempty closed subset of $\mathbb{R}^n$. Since $A \setminus A_1$ is relatively closed in A by (v), $A \setminus A_1$ is closed in $\mathbb{R}^n$, as well. By Theorem A.20, A has upper Minkowski dimension at most $n-1$, since A is closed, A is locally unilaterally well approximated by $\mathcal{H}_{n,d}$, and $\mathcal{H}_{n,d}$ has an $(n-1, C(n,d), 1)$ covering profile. Also, by Theorem A.20, $A \setminus A_1$ has upper Minkowski dimension at most $n-2$, since $A \setminus A_1$ is closed, $A \setminus A_1$ is locally unilaterally well approximated by $\mathcal{SH}_{n,d}$, and $\mathcal{SH}_{n,d}$ has an $(n-2, C(n,d), 1)$ covering profile. This establishes (viii) and the upper bound in (vii). To wrap up, observe that A_1 is nonempty by (vi), A_1 is locally closed by (ii), and A_1 is locally Reifenberg vanishing by (iii). Therefore, by Reifenberg's topological disk theorem (see, e.g., [David and Toro 2012]), A_1 is a topological $(n-1)$-manifold (and more, see Remark 1.3). Therefore, A_1 has Hausdorff and upper Minkowski dimension at least $n-1$. This completes the proof of (vii). □

By examining the proof that A_1 is relatively dense in A in the proof of Theorem 1.1, one sees the only essential property about the cones $\mathcal{H}_{n,1}$ and $\mathcal{H}_{n,d}$, beyond detectability, is that for every $\Sigma_p \in \mathcal{F}_{n,k}$ there exist some $z \in \Sigma_p$ such that $\liminf_{s\downarrow 0} \Theta_{\Sigma_p}^{\mathcal{H}_{n,1}}(z, s) = 0$. Thus, abstracting the argument, one obtains the following result.

Theorem 5.1. *Let $\mathcal{T}$ and $\mathcal{S}$ be local approximation classes. Suppose $\mathcal{T}$ points are detectable in $\mathcal{S}$, and*

$$\textit{for all } S \in \bar{\mathcal{S}} \cap \mathcal{T}^{\perp} \textit{ there exists } x \in S \textit{ such that } \liminf_{r\downarrow 0} \Theta_S^{\mathcal{T}}(x, r) = 0. \tag{5-5}$$

If A is locally bilaterally well approximated by $\mathcal{S}$, then the set $A_{\mathcal{T}}$ described by Theorem A.14 is relatively dense in A, i.e., $\bar{A}_{\mathcal{T}} \cap A = A$.

6. Dimension bounds in the presence of good topology

We now focus our attention on sets A that separate $\mathbb{R}^n$ into two connected components. When $A = \Sigma_p$ and $p : \mathbb{R}^n \to \mathbb{R}$ is harmonic, this occurs precisely when the positive set $\Omega_p^+ = \{x \in \mathbb{R}^n : p(x) > 0\}$ of p and the negative set $\Omega_p^- = \{x \in \mathbb{R}^n : p(x) < 0\}$ of p are pathwise connected. To start, let us prove Lemma 1.7 from the Introduction, which implies that $\mathcal{F}_{n,k}$ contains zero sets Σ_p that separate $\mathbb{R}^n$ into two components for all dimensions $n \ge 4$ and for all degrees $k \ge 2$.

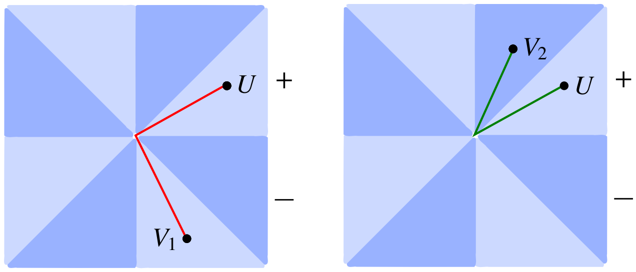

Figure 2. Let $q : \mathbb{R}^2 \to \mathbb{R}$ denote a nonconstant homogeneous harmonic polynomial (illustrated with degree 4). The light blue cells denote the positive set of q and the medium blue cells denote the negative set of q. Suppose that $q(U) > 0$, $q(V_1) > 0$, and $p(V_1, V_2) > 0$, where $p(W_1, W_2) \equiv q(W_1) + q(W_2)$. To move from (V_1, V_2) to (U, U) inside the positive set of p, first send V_2 to U along the green path and then move V_1 to U along the red path.

Proof of Lemma 1.7. We sketch the argument when $a = b = 1$, with the other cases following from an obvious modification. Let $q : \mathbb{R}^2 \to \mathbb{R}$ be a homogeneous harmonic polynomial of degree $k \geq 2$. Note that by elementary complex analysis, q can be written as the real part of a complex polynomial $\tilde{q} : \mathbb{C} \to \mathbb{C}$, $\tilde{q}(z) = cz^k$. Thus, Σ_q is the union of k equiangular lines through the origin and the chambers of $\mathbb{R}^2 \setminus \Sigma_q$ alternate between the positive and negative sets of q. Let $U = (x_1, y_1)$ be any point such that $q(U) > 0$. Then $p(U, U) > 0$, as well, where $p(W_1, W_2) \equiv q(W_1) + q(W_2)$. To show that the positive set of p is connected, it suffices to show that any point $(V_1, V_2) \in \mathbb{R}^2 \times \mathbb{R}^2$ such that $p(V_1, V_2) > 0$ can be connected to (U, U) by a piecewise linear path in the positive set. If $p(V_1, V_2) > 0$, then $q(V_1) > 0$ or $q(V_2) > 0$, say without loss of generality that $q(V_1) > 0$. Then the desired path from (V_1, V_2) to (U, U) is described in Figure 2. A similar argument verifies that the negative set of p is connected and we are done. $\square$

Our goal for the remainder of this section is to prove Theorem 1.8, which requires the following notion of nontangential accessibility.

Definition 6.1 [Jerison and Kenig 1982]. A domain (i.e., a connected open set) $\Omega \subset \mathbb{R}^n$ is called *NTA* or *nontangentially accessible* if there exist constants $M > 1$ and $R > 0$ for which the following are true:

(i) Ω satisfies the *corkscrew condition*: for all $Q \in \partial\Omega$ and $0 < r < R$, there exists $x \in \Omega \cap B(Q, r)$ such that $\operatorname{dist}(x, \partial\Omega) > M^{-1} r$.

(ii) $\mathbb{R}^n \setminus \Omega$ satisfies the corkscrew condition.

(iii) Ω satisfies the *Harnack chain condition*: if $x_1, x_2 \in \Omega \cap B\big(Q, \tfrac{1}{4} r\big)$ for some $Q \in \partial\Omega$ and $0 < r < R$, and $\operatorname{dist}(x_1, \partial\Omega) > \delta$, $\operatorname{dist}(x_2, \partial\Omega) > \delta$, and $|x_1 - x_2| < 2^l \delta$ for some $\delta > 0$ and $l \geq 1$, then there exists a chain of no more than Ml overlapping balls connecting x_1 to x_2 in Ω such that for each ball $B = B(x, s)$ in the chain

$$M^{-1} s < \operatorname{gap}(B, \partial\Omega) < Ms, \qquad \operatorname{gap}(B, \partial\Omega) = \inf_{x \in B} \inf_{y \in \partial\Omega} |x - y|,$$

$$\operatorname{diam} B > M^{-1} \min\{\operatorname{dist}(x_1, \partial\Omega), \operatorname{dist}(x_2, \partial\Omega)\}, \qquad \operatorname{diam} B = \sup_{x, y \in B} |x - y|.$$

We refer to M and R as *NTA constants* of the domain Ω. When $\partial\Omega$ is unbounded, $R=\infty$ is allowed. To distinguish between conditions (i) and (ii), the former may be called the *interior corkscrew condition* and the latter may be called the *exterior corkscrew condition*.

Remark 6.2. In the definition of NTA domains, the additional restriction $R=\infty$ when Ω is unbounded is sometimes imposed (e.g., see [Kenig and Toro 1999; 2006; Kenig et al. 2009]) in order to obtain globally uniform harmonic measure estimates on unbounded domains, but that restriction is not essential in the geometric context of Theorem 1.8, and thus, we omit it.

An essential feature of NTA domains that we need below is that the NTA properties persist under limits (with slightly different constants). When $\Gamma_i = r_i^{-1}(\partial\Omega - Q_i)$ is a sequence of pseudoblowups of the boundary $\partial\Omega$ of a 2-sided NTA domain $\Omega\subset\mathbb{R}^n$ for some $Q_i\in\partial\Omega$ and $r_i>0$ such that $Q_i\to Q\in\partial\Omega$ and $r_i\downarrow 0$, we have the following lemma, due to Kenig and Toro [2006, Theorem 4.1]; also see [Azzam and Mourgoglou 2015, Lemma 1.5] for a recent variant on uniform domains. For the proof of Lemma 6.3, see Appendix B below.

Lemma 6.3. *Suppose that $\Gamma_i\subset\mathbb{R}^n$ is a sequence of closed sets such that $\mathbb{R}^n\setminus\Gamma_i=\Omega_i^+\cup\Omega_i^-$ is the union of complimentary NTA domains Ω_i^+ and Ω_i^- with NTA constants M and R independent of i. If $\Gamma_i\to\Gamma\neq\varnothing$ in the Attouch–Wets topology, then $\mathbb{R}^n\setminus\Gamma=\Omega^+\cup\Omega^-$ is the union of complementary NTA domains Ω^+ and Ω^- with NTA constants $2M$ and R.*

In the remainder of this section, we work with subclasses of $\mathcal{H}_{n,d}$ and $\mathcal{F}_{n,k}$ whose zero sets Σ_p separate $\mathbb{R}^n$ into two distinct NTA components with uniform NTA constants.

Definition 6.4 (2-sided NTA restricted classes $\mathcal{H}^*_{n,d}$, $\mathcal{H}^{**}_{n,d}$, $\mathcal{F}^*_{n,k}$, $\mathcal{F}^{**}_{n,k}$). For all $n\geq 2$ and $d\geq 1$, let $\mathcal{H}^*_{n,d}$ denote the collection of all $\Sigma_p\in\mathcal{H}_{n,d}$ such that $\Omega_p^\pm=\{x\in\mathbb{R}^n:\pm p(x)>0\}$ are NTA domains with NTA constants $M^*=M$ and $R^*=\infty$ for some fixed $M>1$. (We deliberately suppress the choice of M^* from the notation.) Also, let $\mathcal{H}^{**}_{n,d}$ denote the collection of all $\Sigma_p\in\mathcal{H}_{n,d}$ such that $\Omega_p^\pm$ are NTA domains with NTA constants $M^{**}=2M^*$ and $R^{**}=\infty$. Finally, set $\mathcal{F}^*_{n,k}=\mathcal{H}^*_{n,k}\cap\mathcal{F}_{n,k}$ and $\mathcal{F}^{**}_{n,k}=\mathcal{H}^{**}_{n,k}\cap\mathcal{F}_{n,k}$ for all $k\geq 1$.

Remark 6.5. The classes $\mathcal{H}^*_{n,d}$ (hence $\mathcal{H}^{**}_{n,d}$) and $\mathcal{F}^*_{n,k}$ (hence $\mathcal{F}^{**}_{n,k}$) are local approximation classes (see Definition A.7), because $R^*=\infty$, and it is apparent that $\mathcal{H}^*_{n,d}$ is translation invariant in the sense that $\Sigma_p - x\in\mathcal{H}^*_{n,d}$ for all $\Sigma_p\in\mathcal{H}^*_{n,d}$ and $x\in\Sigma_p$. Hence $\overline{\mathcal{H}}^*_{n,d}$ is also translation invariant. By Corollary 4.4 and Lemma 6.3, $\overline{\mathcal{H}}^*_{n,d}\subseteq\mathcal{H}^{**}_{n,d}$ and $\overline{\mathcal{F}}^*_{n,k}\subseteq\mathcal{F}^{**}_{n,k}$. Since $\mathcal{H}_{n,k}$ points are detectable in $\mathcal{H}_{n,d}$ for all $1\leq k\leq d$ by Corollary 1.5 and $\mathcal{H}^*_{n,d}\subseteq\mathcal{H}_{n,d}$, we have $\mathcal{H}_{n,k}$ points are detectable in $\mathcal{H}^*_{n,d}$, as well. Finally, we reiterate that $\mathcal{F}^*_{n,k}$ is nonempty for some $M^*>1$ if and only if $k=1$ and $n\geq 2$; $k\geq 2$ is even and $n\geq 4$; or, $k\geq 3$ is odd and $n\geq 3$. See Remark 1.10. The assertion that the interiors of the two connected components of $\mathbb{R}^n\setminus\Sigma_p$ are NTA domains when $n=3$ and $p=p(x,y,z)$ is Szulkin's polynomial (or any of Lewy's odd-degree polynomials) and when $n=4$ and $p=p(x_1,y_1,x_2,y_2)$ is the zero set of one of the polynomials from Lemma 1.7 follows from the fact that in each case $\Sigma_p\cap\partial B(0,1)$ is a smooth hypersurface in the unit sphere and Σ_p is a cone.

The following technical proposition, alluded to in the Introduction after the statement of Theorem 1.8, is a consequence of Lemma 6.3.

Lemma 6.6. *Suppose that $A \subseteq \mathbb{R}^n$ is closed and $\mathbb{R}^n \setminus A = \Omega^+ \cup \Omega^-$ is a union of complementary NTA domains. If A is locally bilaterally well approximated by $\mathcal{H}_{n,d}$ for some $n \geq 2$ and $d \geq 1$, then A is locally bilaterally well approximated by $\mathcal{H}^*_{n,d}$ for some $M^* > 1$ depending only on the NTA constants of Ω^+ and Ω^-.*

Proof. Suppose that A is closed, A is locally bilaterally well approximated by $\mathcal{H}_{n,d}$, and $\mathbb{R}^n \setminus A = \Omega^+ \cup \Omega^-$ is a union of complementary NTA domains with uniform NTA constants M and R. On one hand, by Theorem A.11 and Corollary 4.4, $\Psi\text{-}\mathrm{Tan}(A, x) \subseteq \overline{\mathcal{H}}_{n,d} = \mathcal{H}_{n,d}$ for all $x \in A$, where $\Psi\text{-}\mathrm{Tan}(A, x)$ is the collection of all pseudotangent sets of A at x. On the other hand, for every $x \in A$ and $r > 0$, the set $(A - x)/r = \Omega^+_{x,r} \cup \Omega^-_{x,r}$ is a union of complementary NTA domains $\Omega^+_{x,r}$ and $\Omega^-_{x,r}$ with NTA constants $M_{x,r} = M$ and $R_{x,r} = R/r$. Thus, every pseudotangent set $T = \lim_{i\to 0}(A - x_i)/r_i \in \Psi\text{-}\mathrm{Tan}(A, x)$ separates $\mathbb{R}^n$ into two NTA domains with NTA constants $M_T = 2M$ and $R_T = \infty$ by Lemma 6.3, since $R_{x_i,r_i} = R/r_i \to \infty$ as $r_i \to 0$. Therefore, $\Psi\text{-}\mathrm{Tan}(A, x) \subseteq \mathcal{H}^*_{n,d}$ for every $x \in A$ with $M^* = 2M$. By Theorem A.11, it follows that A is locally bilaterally well approximated by $\mathcal{H}^*_{n,d}$, as desired. □

In view of Lemma 6.6, Theorem 1.8 is a special case of the following theorem.

Theorem 6.7. *Let $n \geq 2$, $d \geq 2$, and $M^* > 1$. If $A \subseteq \mathbb{R}^n$ is closed and locally bilaterally well approximated by $\mathcal{H}^*_{n,d}$, then*

(i) *$A \setminus A_1 = A_2 \cup \cdots \cup A_d$ has upper Minkowski dimension at most $n - 3$; and,*

(ii) *the even singular set $A_2 \cup A_4 \cup A_6 \cup \cdots$ has Hausdorff dimension at most $n - 4$.*

To prove Theorem 6.7 using the technology of [Badger and Lewis 2015], we need to show the existence of "covering profiles" (see Definition A.19) for the classes $\mathrm{sing}_{\mathcal{H}_{n,1}} \overline{\mathcal{H}}^*_{n,d}$ and $\mathrm{sing}_{\mathcal{H}_{n,d-1}} \overline{\mathcal{H}}^*_{n,d}$ (see Definition A.16), which are well defined because $\overline{\mathcal{H}}^*_{n,d}$ is translation invariant and $\mathcal{H}_{n,k}$ points are detectable in $\mathcal{H}^*_{n,d}$ by Remark 6.5. The following lemma proves the existence of good covering profiles for $\mathrm{sing}_{\mathcal{H}_{n,k-1}} \overline{\mathcal{H}}^*_{n,k}$ for all degrees $k \geq 2$.

Lemma 6.8. *Let $k \geq 2$ and assume that $n + (k \bmod 2) \geq 4$. For every k-homogeneous harmonic polynomial $p : \mathbb{R}^n \to \mathbb{R}$ such that $\mathbb{R}^n \setminus \Sigma_p$ has two connected components,*

$$(\Sigma_p)_{\mathcal{H}^{\perp}_{n,k-1}} = \left\{x \in \Sigma_p : \liminf_{r\to 0} \Theta^{\mathcal{H}_{n,k-1}}_{\Sigma_p}(x, r) > 0\right\}$$

is a linear subspace V of $\mathbb{R}^n$ with $\dim V \leq n - 4 + (k \bmod 2)$. In particular,

$$\mathrm{sing}_{\mathcal{H}_{n,k-1}} \overline{\mathcal{H}}^*_{n,k} = \left\{(\Sigma_p)_{\mathcal{H}^{\perp}_{n,k-1}} : \Sigma_p \in \overline{\mathcal{H}}^*_{n,k},\ 0 \in (\Sigma_p)_{\mathcal{H}^{\perp}_{n,k-1}}\right\}$$

admits an $(n - 4 + (k \bmod 2), C(n), 1)$ covering profile.

Proof. Suppose that k and n satisfy the hypothesis of the lemma and let $p : \mathbb{R}^n \to \mathbb{R}$ be a k-homogeneous harmonic polynomial. We will show that $(\Sigma_p)^{\perp}_{\mathcal{H}_{n,k-1}}$ coincides with

$$V = \{x_0 \in \mathbb{R}^n : p(x + x_0) = p(x) \text{ for all } x \in \mathbb{R}^n\},$$

which is a linear subspace of $\mathbb{R}^n$ because p is k-homogeneous. To start, note that

$$\begin{aligned} x_0 \in (\Sigma_p)_{\mathcal{H}^\perp_{n,k-1}} &\iff \partial^\alpha p(x_0) = 0 \quad \text{for all } |\alpha| \le k-1 \\ &\iff p(x+x_0) \equiv q(x) \quad \text{for some } q, \text{ where } q : \mathbb{R}^n \to \mathbb{R} \text{ is } k\text{-homogeneous,} \end{aligned}$$

where the first equivalence holds by Theorem 1.4 and the second equivalence holds by Taylor's theorem. Hence $V \subseteq (\Sigma_p)_{\mathcal{H}^\perp_{n,k-1}}$, since p is k-homogeneous. Conversely, using the homogeneity of p and q, at any $x_0 \in (\Sigma_p)_{\mathcal{H}^\perp_{n,k-1}}$ we obtain

$$p(x+x_0) = q(x) = \lambda^k q(x/\lambda) = \lambda^k p(x/\lambda + x_0) = p(x + \lambda x_0) \quad \text{for all } \lambda \in \mathbb{R} \setminus \{0\}.$$

Letting $\lambda \to 0$, we conclude that $p(x+x_0) = p(x)$ for all $x \in \mathbb{R}^n$ whenever $x \in (\Sigma_p)_{\mathcal{H}^\perp_{n,k-1}}$. Thus, $(\Sigma_p)_{\mathcal{H}^\perp_{n,k-1}} \subseteq V$, as well.

To continue, suppose that Σ_p separates $\mathbb{R}^n$ into two components. Let $\tilde{p} : V^\perp \to \mathbb{R}$ be the image of p under the quotient map $\mathbb{R}^n \to \mathbb{R}^n/V \cong V^\perp$. Because V is the space of invariant directions for p, the map $\tilde{p}$ is still a degree-k homogeneous harmonic polynomial (in orthonormal coordinates for $V^\perp$) and

$$\Sigma_p = \Sigma_{\tilde{p}} \oplus V = \{x + v : x \in \Sigma_{\tilde{p}} \subseteq V^\perp, v \in V\}.$$

Hence $\Sigma_{\tilde{p}}$ separates $V^\perp$ into two components, since Σ_p separates $\mathbb{R}^n$ into two components. It follows that $\dim V^\perp \ge 4$ if $k \ge 2$ is even, and $\dim V^\perp \ge 3$ if $k \ge 3$ is odd; e.g., see the paragraph immediately preceding the statement of Lemma 1.7. Therefore, $\dim V \le n-4$ if $k \ge 2$ is even, and $\dim V \le n-3$ if $k \ge 3$ is odd.

Finally, by Theorem 1.4, Remark 6.5, and the first part of the lemma,

$$\operatorname{sing}_{\mathcal{H}_{n,k-1}} \overline{\mathcal{H}}^*_{n,k} = \left\{(\Sigma_p)_{\mathcal{H}^\perp_{n,k-1}} : \Sigma_p \in \overline{\mathcal{F}}^*_{n,k}\right\} \subseteq \left\{(\Sigma_p)_{\mathcal{H}^\perp_{n,k-1}} : \Sigma_p \in \mathcal{F}^{**}_{n,k}\right\} \subseteq \bigcup_{i=0}^{j} G(n,i),$$

where $j = n-4$ if $k \ge 2$ is even, and $j = n-3$ if $k \ge 3$ is odd. Here each $G(n,i)$ denotes the Grassmannian of dimension-i linear subspaces of $\mathbb{R}^n$, which possesses an $(i, C(i), 1)$ covering profile; that is, $V \cap B(0,r)$ can be covered by $C(i)s^{-i}$ balls $B(v_i, sr)$ centered in $V \cap B(0,r)$ for all planes $V \in G(n,i)$, $r > 0$, and $0 < s \le 1$. (For example, this follows from the fact that the Lebesgue measure of any ball of radius r in $\mathbb{R}^i$ is proportional to r^i.) It follows that the class $\operatorname{sing}_{\mathcal{H}_{n,k-1}} \overline{\mathcal{H}}^*_{n,k}$ has an $(n-4, C(n), 1)$ covering profile when $k \ge 2$ is even, and $\operatorname{sing}_{\mathcal{H}_{n,k-1}} \overline{\mathcal{H}}^*_{n,k}$ has an $(n-3, C(n), 1)$ covering profile when $k \ge 3$ is odd. □

The covering profiles for $\operatorname{sing}_{\mathcal{H}_{n,k-1}} \overline{\mathcal{H}}^*_{n,k}$ from Lemma 6.8 will enable us to prove (ii) in Theorem 6.7 and also to prove that $A \setminus A_1$ has Hausdorff dimension at most $n-3$. However, to show that $A \setminus A_1$ has upper Minkowski dimension at most $n-3$, we need to find covering profiles for $\operatorname{sing}_{\mathcal{H}_{n,1}} \overline{\mathcal{H}}^*_{n,d}$, whose existence does not automatically follow from the covering profiles in Lemma 6.8. To proceed, we use the quantitative stratification and volume estimates for singular sets of harmonic functions developed by Cheeger, Naber, and Valtorta [Cheeger et al. 2015]. The following description of the stratification combines several definitions from §1 of their paper; see Definitions 1.4, 1.7 and 1.9 and Remark 1.8 of the same work.

Definition 6.9 ([Cheeger et al. 2015]; quantitative stratification by symmetry). A smooth function $u : \mathbb{R}^n \to \mathbb{R}$ is called 0-*symmetric* if u is a homogeneous polynomial and u is called k-*symmetric* if u is 0-symmetric and there exists a k-dimensional subspace V such that

$$u(x+y) = u(x) \quad \text{for all } x \in \mathbb{R}^n \text{ and } y \in V.$$

For all smooth $u : B(0,1) \to \mathbb{R}$, and for all $x \in B(0, 1-r)$, define

$$T_{x,r}u(y) = \frac{u(x+ry) - u(x)}{\left(\fint_{\partial B(0,1)} |u(x+rz) - u(x)|^2 \, d\sigma(z) \right)^{1/2}} \quad \text{for all } y \in B(0,1).$$

(If the denominator vanishes, set $T_{x,r} = \infty$.) A harmonic function $u : B(0,1) \to \mathbb{R}$ is called (k, ε, r, x)-*symmetric* if there exists a harmonic k-symmetric function p with $\fint_{\partial B(0,1)} |p|^2 \, d\sigma = 1$ such that

$$\fint_{B(0,1)} |T_{x,r}u - p|^2 < \varepsilon.$$

For all harmonic $u : B(0,1) \to \mathbb{R}$, define the (k, η, r)-*effective singular stratum* by

$$\mathcal{S}^k_{\eta,r}(u) = \{x \in B(0,1) : u \text{ is not } (k{+}1, \eta, s, x)\text{-symmetric for all } s \geq r\}.$$

For harmonic functions, [Cheeger et al. 2015, Theorem 1.10] gives the following Minkowski-type estimates for effective singular strata. In the statement, $N(1, 0, u)$ denotes Almgren's frequency function with $r=1$, $x_0=0$, and $f=u$ (recall Definition 3.3 above).

Theorem 6.10 [Cheeger et al. 2015]. *If $u : B(0,1) \to \mathbb{R}$ is a harmonic function with $u(0) = 0$ and $N(1,0,u) \leq \Lambda < \infty$, then for every $\eta > 0$ and $k \leq n-2$,*

$$\operatorname{Vol}\left(\left\{x \in B\left(0, \tfrac{1}{2}\right) : \operatorname{dist}(x, \mathcal{S}^k_{\eta,r}(u)) < r\right\}\right) \leq C(n, \Lambda, \eta) r^{n-k-\eta}. \tag{6-1}$$

We now show that if η is small enough depending on n, d, and M^*, then the singular set of $\Sigma_p \in \mathcal{H}^*_{n,d}$ is contained in $\mathcal{S}^{n-3}_{\eta,r}(p)$.

Lemma 6.11. *For all $n \geq 2$, $d \geq 2$, and $M^* > 1$, there exists $\bar{\eta} > 0$ with the following property. If $\Sigma_p \in \mathcal{H}^*_{n,d}$, $x_0 \in \Sigma_p$, and p is $(n-2, \eta, r, x_0)$-symmetric for some $\eta \in (0, \bar{\eta})$ and $r > 0$, then x_0 is an $\mathcal{F}_{n,1}$ point of Σ_p. Consequently, the set of all singular points of Σ_p (that is, $\mathcal{F}_{n,2} \cup \cdots \cup \mathcal{F}_{n,d}$ points of Σ_p) belongs to $\mathcal{S}^{n-3}_{\eta,r}(p)$ for all $\eta \in (0, \bar{\eta})$ and $r > 0$.*

Proof. Let $n \geq 2$, $d \geq 2$, and $M^* > 1$ be given. Assume in order to obtain a contradiction that for all $i \geq 1$, there exist $\Sigma_{p_i} \in \mathcal{H}^*_{n,d}$, $\eta_i < 1/i$, $x_i \in \Sigma_{p_i}$, and $r_i > 0$ such that p_i is $(n-2, \eta_i, r_i, x_i)$-symmetric and x_i is not an $\mathcal{F}_{n,1}$ point of Σ_{p_i}. Equivalently, by Theorem 1.4, $Dp_i(x_i) = 0$. That is, the Taylor expansion for p_i at x_i has no nonzero linear terms. By definition of almost symmetry, there exist $(n{-}2)$-symmetric homogeneous harmonic polynomials h_i such that $\fint_{\partial B(0,1)} |h_i|^2 \, d\sigma = 1$ and

$$\fint_{B(0,1)} |T_{x_i,r_i} p_i - h_i|^2 < \frac{1}{i}. \tag{6-2}$$

As everything is translation, dilation, and rotation invariant, we may assume without loss of generality that for all $i \geq 1$, we have $x_i = 0$, $r_i = 1$, and $h_i(y_1, y_2, \ldots, y_n) = h_i(y_1, y_2, 0, \ldots, 0)$ for all $y \in \mathbb{R}^n$. To ease notation, let us abbreviate $q_i \equiv T_{0,1} p_i$. We note that

$$\|q_i\|_{L^2(B(0,1))} \sim_{n,d} \|q_i\|_{L^2(\partial B(0,1))} \sim_{n,d} 1 \quad \text{for all } i \geq 1, \tag{6-3}$$

where the first comparison holds by Lemma 3.8 and the second holds by the definition of $T_{0,1} p_i$.

We now claim that $\deg h_i \leq d$ for all i sufficiently large. To see this, suppose to the contrary that $l := \deg h_i > d$ for some $i \geq 1$. Recalling both that spherical harmonics of different degrees are orthogonal on spheres centered at the origin and that h_i is l-homogeneous with $l > \deg q_i$, we have

$$1 \sim_{n,d} \|q_i\|^2_{L^2(B(0,1))} \lesssim_{n,d} \fint_{B(0,1)} (q_i^2 + h_i^2) = \fint_{B(0,1)} |q_i - h_i|^2 < \frac{1}{i}$$

by (6-2) and (6-3). This is impossible if i is sufficient large depending only on n and d. Thus, $\deg h_i \leq d$ for all i sufficient large, as claimed. In particular,

$$\|h_i\|_{L^2(B(0,1))} \sim_{n,d} \|h_i\|_{L^2(\partial B(0,1))} \sim_{n,d} 1 \quad \text{for all } i \gtrsim_{n,d} 1. \tag{6-4}$$

By (6-3), (6-4), Lemma 2.1, and Corollary 3.5, we conclude that $H(q_i) \sim_{n,d} 1$ and $H(q_i) \sim_{n,d} 1$ for all sufficiently large i. Therefore, by passing to a subsequence of the pair $(q_i, h_i)_{i=1}^{\infty}$ (which we relabel), we may assume that $q_i \to q$ in coefficients and $h_i \to h$ in coefficients for some nonconstant harmonic polynomials q and h of degree at most d. On one hand, we have $\Sigma_q \in \overline{\mathcal{H}}^*_{n,d} \subseteq \mathcal{H}^{**}_{n,d}$ by Lemma 4.3 and $Dq(0) = 0$, since $Dq_i(0) = 0$ for all i. Hence q has degree at least 2. On the other hand, we have h is homogeneous and $h(y_1, y_2, \ldots, y_n) = h(y_1, y_2, 0, \ldots, 0)$ for all $y \in \mathbb{R}^n$, because the same are true of the polynomial h_i for all $i \gtrsim_{n,d} 1$.

We are now ready to obtain a contradiction. Since $q_i \to q$ and $h_i \to h$ uniformly on compact sets, we have $q \equiv h$ by (6-2). Thus, $\Sigma_q \in \mathcal{F}^{**}_{n,k}$ for some $2 \leq k \leq d$ — in particular, Σ_q is the zero set of a homogeneous harmonic polynomial of degree at least 2 that separates $\mathbb{R}^n$ into two components — and q depends on at most two variables. No such polynomial q exists (e.g., see Remark 6.5)! Therefore, for all $n \geq 2$, $d \geq 2$, and $M^* > 1$, there exists $\bar{\eta} > 0$ such that if $\Sigma_p \in \mathcal{H}^*_{n,d}$, $x_0 \in \Sigma_p$, and p is $(n-2, \eta, r, x_0)$-symmetric for some $\eta \in (0, \bar{\eta})$ and $r > 0$, then x_0 is an $\mathcal{F}_{n,1}$ point of Σ_p. Consequently, if $\Sigma_p \in \mathcal{H}^*_{n,d}$ and $x_0 \in \Sigma_p$ belongs to the singular set of p, then p is not $(n-2, \eta, r, x_0)$ symmetric for all $\eta \in (0, \bar{\eta})$ and $r > 0$. By definition of the singular strata, we conclude that for all $\Sigma_p \in \mathcal{H}^*_{n,d}$ the set of all singular points of Σ_p belongs to $\mathcal{S}^{n-3}_{\eta,r}(p)$ for all $\eta \in (0, \bar{\eta})$ and $r > 0$. □

At last, we are ready to prove Theorems 6.7 and 1.8.

Proof of Theorems 6.7 and 1.8. As noted earlier, Theorem 6.7 implies Theorem 1.8 by Lemma 6.6. Thus, it suffices to establish the former. Assume $A \subseteq \mathbb{R}^n$ is closed and locally bilaterally well approximated by $\mathcal{H}^*_{n,d}$ for some $M^* > 1$. Then A can be written as $A = A_1 \cup A_2 \cup \cdots \cup A_d$ according to Theorem 1.1. In particular, $U_k = A_1 \cup \cdots \cup A_k$ is relatively open in A and locally bilaterally well approximated by $\mathcal{H}_{n,k}$ for all $1 \leq k \leq d$. Hence U_k is also locally bilaterally well approximated by $\mathcal{H}^{**}_{n,k}$ for all $1 \leq k \leq d$, because $\Psi\text{-}\mathrm{Tan}(A, x) \subseteq \overline{\mathcal{H}}^*_{n,d} \cap \mathcal{H}_{n,k} \subseteq \mathcal{H}^{**}_{n,k}$ for all $x \in U_k$ by Theorem A.11 and Remark 6.5. Also,

$A \setminus A_1$ is closed in $\mathbb{R}^n$, because A_1 is relatively open in A and A is closed in $\mathbb{R}^n$, and A_k is σ-compact for each $k \geq 1$, because A_k is relatively closed in U_k, U_k is relatively open in A, and A is closed in $\mathbb{R}^n$. Our goal is to prove that (i) $\overline{\dim}_M\, A \setminus A_1 \leq n-3$ and (ii) $\dim_H A_k \leq n-4$ for all even $k \geq 2$.

We begin with a proof of (i). By Remark 6.5, $\overline{\mathcal{H}}^{**}_{n,d}$ is translation invariant and $\mathcal{H}_{n,1}$ points are detectable in $\mathcal{H}^{**}_{n,d}$. Thus, $A \setminus A_1$ is locally unilaterally well approximated by $\operatorname{sing}_{\mathcal{H}_{n,1}} \overline{\mathcal{H}}^{**}_{n,d}$ by Theorem A.17. By Lemma 6.11 and Theorem 6.10, the class $\operatorname{sing}_{\mathcal{H}_{n,1}} \overline{\mathcal{H}}^{**}_{n,d}$ admits an $(n-3+\eta, C(n, d, \eta, M^{**}), 1)$ covering profile for all $\eta > 0$. Thus, since $A \setminus A_1$ is closed, we have $\overline{\dim}_M\, A \setminus A_1 \leq n-3+\eta$ for all $\eta > 0$ by Theorem A.20. Letting $\eta \downarrow 0$, we conclude $\overline{\dim}_M\, A \setminus A_1 \leq n-3$, as desired.

We now prove (ii). Let $k \geq 2$ be even. By Remark 6.5, $\overline{\mathcal{H}}^{**}_{n,k-1}$ is translation invariant and $\mathcal{H}_{n,k-1}$ points are detectable in $\mathcal{H}^{**}_{n,k}$. Thus, $A_k = U_k \setminus U_{k-1}$ is locally unilaterally well approximated by $\operatorname{sing}_{\mathcal{H}_{n,k-1}} \overline{\mathcal{H}}^{**}_{n,k}$ by Theorem A.17. By Lemma 6.8, the class $\operatorname{sing}_{\mathcal{H}_{n,k-1}} \overline{\mathcal{H}}^{**}_{n,k}$ admits an $(n-4, C(n), 1)$ covering profile. Thus, since A_k is σ-compact, we have $\dim_H A_k \leq n-4$ by Theorem A.21, as desired. Because Hausdorff dimension is stable under countable unions, $\dim_H A_2 \cup A_4 \cup \cdots \leq n-4$, as well. $\square$

7. Boundary structure in terms of interior and exterior harmonic measures

Harmonic measure arises in classical analysis from the solution of the Dirichlet problem and in probability as the exit distribution of Brownian motion. For nice introductions to harmonic measure, see the books of Garnett and Marshall [2005] and Mörters and Peres [2010]. One of our motivations for this work is the desire to understand the extent to which the structure of the boundary of a domain in $\mathbb{R}^n$, $n \geq 2$, is determined by the relationship between harmonic measures in the interior and the exterior of the domain. This problem can be understood as a free boundary regularity problem for harmonic measure. For an in-depth introduction to free boundary problems for harmonic measure, see the book of Capogna, Kenig, and Lanzani [Capogna et al. 2005].

Given a simply connected domain $\Omega \subset \mathbb{R}^2$, bounded by a Jordan curve, let ω^+ and ω^- denote the harmonic measures associated to $\Omega^+ = \Omega$ and $\Omega^- = \mathbb{R}^2 \setminus \overline{\Omega}$, respectively, which are supported on their common boundary $\partial\Omega = \partial\Omega^+ = \partial\Omega^-$. Together, the theorems of McMillan, Makarov, and Pommerenke (see [Garnett and Marshall 2005, Chapter VI]) show that

$$\omega^+ \ll \omega^- \ll \omega^+ \quad \Longrightarrow \quad \omega^+ \ll \mathcal{H}^1|_G \ll \omega^+ \quad \text{and} \quad \omega^- \ll \mathcal{H}^1|_G \ll \omega^-$$

for some set $G \subseteq \partial\Omega$ with σ-finite 1-dimensional Hausdorff measure and $\omega^\pm(\partial\Omega \setminus G) = 0$; furthermore, in this case, $\partial\Omega$ possesses a unique tangent line at Q for $\omega^\pm$-a.e. $Q \in \partial\Omega$. Here $\mathcal{H}^s$ denotes the s-dimensional Hausdorff measure of sets in $\mathbb{R}^n$. Motivated by this result, Bishop [1992] asked whether if on a domain in $\mathbb{R}^n$, $n \geq 3$,

$$\omega^+ \ll \omega^- \ll \omega^+ \quad \Longrightarrow \quad \omega^+ \ll \mathcal{H}^{n-1}|_G \ll \omega^+ \quad \text{and} \quad \omega^- \ll \mathcal{H}^{n-1}|_G \ll \omega^- \tag{7-1}$$

for some $G \subseteq \partial\Omega$ with σ-finite $(n-1)$-dimensional Hausdorff measure and $\omega^\pm(\partial\Omega \setminus G) = 0$. In [Kenig et al. 2009], Kenig, Preiss, and Toro proved that when $\Omega^+ = \Omega \subset \mathbb{R}^n$ and $\Omega^- = \mathbb{R}^n \setminus \overline{\Omega}$ are NTA domains in $\mathbb{R}^n$, $n \geq 3$, the mutual absolute continuity of ω^+ and ω^- on a set $E \subseteq \partial\Omega$ implies that $\omega^\pm|_E$ has *upper Hausdorff dimension* $n-1$: there exists a set $E' \subseteq E$ of Hausdorff dimension $n-1$ such that

$\omega^\pm(E \setminus E') = 0$, and $\omega^\pm(E \setminus E'') > 0$ for every set $E'' \subset E$ with $\dim_H E'' < n-1$. Moreover, in this case $\omega^\pm|_E \ll \mathcal{H}^{n-1}|_E \ll \omega^\pm|_E$ provided that $\mathcal{H}^{n-1}|_{\partial\Omega}$ is locally finite (see [Badger 2012; 2013, Remark 6.19]). However, at present it is still unknown whether or not (7-1) holds on domains for which $\mathcal{H}^{n-1}|_{\partial\Omega}$ is not locally finite. For some related inquiries, see the work of Lewis, Verchota, and Vogel [Lewis et al. 2005], Azzam and Mourgoglou [2015], Bortz and Hofmann [2016].

Remark 7.1 (added in February 2017). Several months after the first version of this paper appeared on the arXiv in September 2015, a solution to Bishop's conjecture (7-1) was furnished by Azzam, Mourgoglou, and Tolsa [Azzam et al. 2017b] and by Azzam, Mourgoglou, Tolsa, and Volberg [Azzam et al. 2016]. An important tool in these works is a new "bounded Riesz transform" to "uniform rectifiability" criterion of Girela-Sarrión and Tolsa [2016].

Finer information about the structure and size of the boundary under more stringent assumptions on the relationship between ω^+ and ω^- has been obtained in [Kenig and Toro 2006; Badger 2011; 2013; Badger and Lewis 2015; Engelstein 2016]. We summarize these results in Theorem 7.3 after recalling the definition of the space $\mathrm{VMO}(d\omega)$ of functions of vanishing mean oscillation, which extends the space of uniformly continuous bounded functions on $\partial\Omega$.

Definition 7.2 [Kenig and Toro 2006, Definitions 4.2 and 4.3]. Let $\Omega \subset \mathbb{R}^n$ be an NTA domain (with the NTA constant $R = \infty$ when $\partial\Omega$ is unbounded) equipped with harmonic measure ω. We say that $f \in L^2_{\mathrm{loc}}(d\omega)$ belongs to $\mathrm{BMO}(d\omega)$ if and only if

$$\sup_{r>0}\sup_{Q\in\partial\Omega}\left(\fint_{B(Q,r)} |f - f_{Q,r}|^2\, d\omega\right)^{1/2} < \infty,$$

where $f_{Q,r} = \fint_{B(Q,r)} f\, d\omega$ denotes the average of f over the ball. We denote by $\mathrm{VMO}(d\omega)$ the closure in $\mathrm{BMO}(d\omega)$ of the set of uniformly continuous bounded functions on $\partial\Omega$.

Theorem 7.3. *Assume that $\Omega^+ = \Omega \subset \mathbb{R}^n$ and $\Omega^- = \mathbb{R}^n \setminus \overline{\Omega}$ are NTA domains* (*with the NTA constant $R = \infty$ when $\partial\Omega$ is unbounded*), *equipped with harmonic measures $\omega^\pm$ on $\Omega^\pm$. If $\omega^+ \ll \omega^- \ll \omega^+$ and the Radon–Nikodym derivative $f = d\omega^-/d\omega^+$ satisfies* $\log f \in \mathrm{VMO}(d\omega^+)$, *then the boundary $\partial\Omega$ satisfies the following properties.*

- *There exist $d \geq 1$ and $M^* > 1$ depending on at most n and the NTA constants of Ω^+ and Ω^- such that $\partial\Omega$ is locally bilaterally well approximated by $\mathcal{H}^*_{n,d}$* [Kenig and Toro 2006].
- *$\partial\Omega$ can be partitioned into disjoint sets Γ_k, $1 \leq k \leq d$, such that $x \in \Gamma_k$ if and only if x is an $\mathcal{F}_{n,k}$ point of $\partial\Omega$. Moreover, Γ_1 is dense in $\partial\Omega$ and $\omega^\pm(\partial\Omega \setminus \Gamma_1) = 0$* [Badger 2011].
- *Γ_1 is relatively open in $\partial\Omega$, Γ_1 is locally bilaterally well approximated by $\mathcal{H}_{n,1}$, and Γ_1 has Hausdorff dimension $n-1$* [Badger 2013].
- *$\partial\Omega$ has upper Minkowski dimension $n-1$ and $\partial\Omega \setminus \Gamma_1 = \Gamma_2 \cup \cdots \cup \Gamma_d$ has upper Minkowski dimension at most $n-2$* [Badger and Lewis 2015].
- *If* $\log f \in C^{l,\alpha}$ *for some $l \geq 0$ and $\alpha > 0$* (*resp.* $\log f \in C^\infty$, $\log f$ *real analytic*), *then Γ_1 is a $C^{l+1,\alpha}$* (*resp. C^∞, real analytic*) *$(n-1)$-dimensional manifold* [Engelstein 2016].

Remark 7.4. The statements from [Kenig and Toro 2006] and [Badger 2011] recorded in Theorem 7.3 were obtained by showing that the pseudotangent measures of the harmonic measures $\omega^{\pm}$ of $\Omega^{\pm}$ are "polynomial harmonic measures" in [Kenig and Toro 2006] and by studying the "separation at infinity" of cones of polynomial harmonic measures associated to polynomials of different degrees in [Badger 2011] (also see [Kenig et al. 2009]). The statements from [Badger 2013] and [Badger and Lewis 2015] are forerunners to and motivated the statement and proof of Theorem 1.1 in this paper. However, we wish to emphasize that the structure theorem [Badger 2013, Theorem 5.10] and dimension estimate on the singular set $\partial\Omega \setminus \Gamma_1$ in [Badger and Lewis 2015, Theorem 9.3] required existence of the decomposition from [Badger 2011] as part of their hypotheses. By contrast, in this paper, we are able to establish the decomposition $A = A_1 \cup \cdots \cup A_d$ and obtain dimension estimates on the singular set $A \setminus A_1$ in Theorem 1.1 *directly, without any reference to harmonic measure* or dependence on [Badger 2011].

Theorem 1.1 and 1.8 of the present paper yield several new pieces of information about the boundary of complimentary NTA domains with $\log f \in \mathrm{VMO}(d\omega^+)$, which we record in Theorem 7.5.

Theorem 7.5. *Under the hypothesis of Theorem 7.3, the boundary $\partial\Omega = \Gamma_1 \cup \cdots \cup \Gamma_d$ satisfies the following additional properties*:

(i) *For all $1 \le k \le d$, the set $U_k := \Gamma_1 \cup \cdots \cup \Gamma_k$ is relatively open in $\partial\Omega$ and $\Gamma_{k+1} \cup \cdots \cup \Gamma_d$ is closed.*

(ii) *For all $1 \le k \le d$, the set U_k is locally bilaterally well approximated by $\mathcal{H}^{**}_{n,k}$.*

(iii) *For all $1 \le k \le d$, the boundary $\partial\Omega$ is locally bilaterally well approximated along Γ_k by $\mathcal{F}^{**}_{n,k}$, i.e., $\limsup_{r\downarrow 0} \sup_{x\in K} \Theta^{\mathcal{F}^{**}_{n,k}}_{\partial\Omega}(x,r) = 0$ for every compact set $K \subseteq \Gamma_k$.*

(iv) *For all $1 \le l < k \le d$, the set U_l is relatively open in U_k and $\Gamma_{l+1} \cup \cdots \cup \Gamma_k$ is relatively closed in U_k.*

(v) *$\partial\Omega \setminus \Gamma_1 = \Gamma_2 \cup \cdots \cup \Gamma_d$ has upper Minkowski dimension at most $n-3$.*

(vi) *The even singular set $\Gamma_2 \cup \Gamma_4 \cup \cdots$ has Hausdorff dimension at most $n-4$.*

(vii) *When $n \ge 3$, the singular set $\partial\Omega \setminus \Gamma_1$ has Newtonian capacity zero.*

Proof. Parts (i) and (iv) of the theorem are a direct consequence of Theorem 1.1. Parts (ii) and (iii) follow from Theorem 1.1 in conjunction with Lemma 6.6, Theorem A.11, and Remark 6.5 (see the proof of Theorem 1.8). Parts (v) and (vi) are a direct consequence of Theorem 1.8. Newtonian capacity in $\mathbb{R}^n$, $n \ge 3$, is precisely the Riesz $(n-2)$-capacity. Thus, part (vii) follows from (v) and the fact that sets of finite s-dimensional Hausdorff measure have Riesz s-capacity zero (see, e.g., [Mörters and Peres 2010, Chapter 4] or [Mattila 1995, Chapter 8]). □

Remark 7.6. The dimension bounds (v) and (vi) in Theorem 7.5 are sharp by example. See Remark 1.10 and Remark 6.5.

Remark 7.7. The fact that $\partial\Omega \setminus \Gamma_1$ has Newtonian capacity zero implies $\omega^{\pm}(\partial\Omega \setminus \Gamma_1) = 0$; see [Mörters and Peres 2010, Chapter 8].

Appendix A: Local set approximation

A general framework for describing bilateral and unilateral approximations of a set $A \subseteq \mathbb{R}^n$ by a class $\mathcal{S}$ of closed "model" sets is developed in [Badger and Lewis 2015]. In this appendix, we give a brief, self-contained abstract of the main definitions and theorems from this framework as used above, but refer the reader to [Badger and Lewis 2015] for full details and further results. The principal results are two structure theorems for Reifenberg-type sets; see Theorems A.14 and A.17.

Distances between sets. If $A, B \subseteq \mathbb{R}^n$ are nonempty sets, the *excess of A over B* is the asymmetric quantity defined by $\mathrm{ex}(A, B) = \sup_{a\in A} \inf_{b\in B} |a-b| \in [0, \infty]$. By convention, one also defines $\mathrm{ex}(\varnothing, B) = 0$, but leaves $\mathrm{ex}(A, \varnothing)$ undefined. The excess is monotone,

$$\mathrm{ex}(A, B) \le \mathrm{ex}(A', B') \quad \text{whenever } A \subseteq A' \text{ and } B \supseteq B',$$

and satisfies the triangle inequality,

$$\mathrm{ex}(A, C) \le \mathrm{ex}(A, B) + \mathrm{ex}(B, C).$$

When $A = \{x\}$ for some $x \in \mathbb{R}^n$, the excess $\mathrm{ex}(\{x\}, B)$ is usually called the *distance of x to B* and is denoted by $\mathrm{dist}(x, B)$.

For all $x \in \mathbb{R}^n$ and $r > 0$, let $B(x,r)$ denote the open ball with center x and radius r. (In [Badger and Lewis 2015], $B(x,r)$ denotes the closed ball, but see Remark 2.4 in that paper.) For arbitrary sets $A, B \subseteq \mathbb{R}^n$ with B nonempty and for all $x \in \mathbb{R}^n$ and $r > 0$, define the *relative excess of A over B in $B(x,r)$* by

$$\tilde{\mathrm{d}}^{x,r}(A, B) = r^{-1}\,\mathrm{ex}\big(A \cap B(x,r), B\big) \in [0, \infty).$$

Also, for all sets $A, B \subseteq \mathbb{R}^n$ with A and B nonempty and for all $x \in \mathbb{R}^n$ and $r > 0$, define the *relative Walkup–Wets distance between A and B in $B(x,r)$* by

$$\widetilde{\mathrm{D}}^{x,r}[A, B] = \max\big\{\tilde{\mathrm{d}}^{x,r}(A, B), \tilde{\mathrm{d}}^{x,r}(B, A)\big\} \in [0, \infty).$$

Observe that $\widetilde{\mathrm{D}}^{x,r}[A, B] \le 2$ if both $A \cap B(x,r)$ and $B \cap B(x,r)$ are nonempty; and $\widetilde{\mathrm{D}}^{x,r}[A, B] \le 1$ if both $x \in A$ and $x \in B$.

Lemma A.1 [Badger and Lewis 2015, Lemma 2.2, Remark 2.4]. *Let $A, B, C \subseteq \mathbb{R}^n$ be nonempty sets, let $x, y \in \mathbb{R}^n$, and let $r, s > 0$. Then we have the following properties*:

- *closure*: $\widetilde{\mathrm{D}}^{x,r}[A, B] = \widetilde{\mathrm{D}}^{x,r}[A, \bar{B}] = \widetilde{\mathrm{D}}^{x,r}[\bar{A}, \bar{B}] = \widetilde{\mathrm{D}}^{x,r}[\bar{A}, B]$.
- *containment*: $\widetilde{\mathrm{D}}^{x,r}[A, B] = 0$ *if and only if* $\bar{A} \cap B(x,r) = \bar{B} \cap B(x,r)$.
- *quasimonotonicity*: *If* $B(x,r) \subseteq B(y,s)$, *then* $\widetilde{\mathrm{D}}^{x,r}[A, B] \le (s/r)\widetilde{\mathrm{D}}^{y,s}[A, B]$.
- *strong quasitriangle inequality*: *If* $\tilde{\mathrm{d}}^{x,r}(A, B) \le \varepsilon_1$ *and* $\tilde{\mathrm{d}}^{x,r}(C, B) \le \varepsilon_2$, *then*

$$\widetilde{\mathrm{D}}^{x,r}[A, C] \le (1+\varepsilon_2)\widetilde{\mathrm{D}}^{x,(1+\varepsilon_2)r}[A, B] + (1+\varepsilon_1)\widetilde{\mathrm{D}}^{x,(1+\varepsilon_1)r}[B, C].$$

- *weak quasitriangle inequalities*: *If $x \in B$, then*

$$\widetilde{\mathrm{D}}^{x,r}[A, C] \le 2\widetilde{\mathrm{D}}^{x,2r}[A, B] + 2\widetilde{\mathrm{D}}^{x,2r}[B, C].$$

If $B \cap B(x, r) \neq \varnothing$, then

$$\widetilde{\mathrm{D}}^{x,r}[A, B] \le 3\widetilde{\mathrm{D}}^{x,3r}[A, B] + 3\widetilde{\mathrm{D}}^{x,3r}[B, C].$$

- *scale invariance*: $\widetilde{\mathrm{D}}^{x,r}[A, B] = \widetilde{\mathrm{D}}^{\lambda x,\lambda r}[\lambda A, \lambda B]$ *for all* $\lambda > 0$.
- *translation invariance*: $\widetilde{\mathrm{D}}^{x,r}[A, B] = \widetilde{\mathrm{D}}^{x+z,r}[z + A, z + B]$ *for all* $z \in \mathbb{R}^n$.

Remark A.2. The *relative Hausdorff distance between A and B in $B(x, r)$*, defined by

$$\mathrm{D}^{x,r}[A, B] = r^{-1} \max \left\{ \operatorname{ex}\big(A \cap B(x, r), B \cap B(x, r)\big), \operatorname{ex}\big(B \cap B(x, r), A \cap B(x, r)\big) \right\}$$

whenever $A \cap B(x, r)$ and $B \cap B(x, r)$ are both nonempty, is a common, better-known variant of the relative Walkup–Wets distance. We note that $\widetilde{\mathrm{D}}^{x,r}[A, B] \le \mathrm{D}^{x,r}[A, B]$ whenever both quantities are defined. Although the relative Hausdorff distance satisfies the triangle inequality rather than just the weak and strong quasitriangle inequalities enjoyed by the relative Walkup–Wets distance, the relative Hausdorff distance fails to be quasimonotone (see [Badger and Lewis 2015, Remark 2.3]). This makes the relative Hausdorff distance unsuitable for use in the local set approximation framework below. The use of the relative Walkup–Wets distance is deliberate and ensures that one can obtain structure theorems for Reifenberg-type sets.

Attouch–Wets topology, tangent sets, and pseudotangent sets. Let $\mathfrak{C}(\mathbb{R}^n)$ denote the collection of all nonempty closed sets in $\mathbb{R}^n$. Let $\mathfrak{C}(0)$ denote the subcollection of all nonempty closed sets in $\mathbb{R}^n$ containing the origin. We endow $\mathfrak{C}(\mathbb{R}^n)$ and $\mathfrak{C}(0)$ with the *Attouch–Wets topology* (see [Beer 1993, Chapter 3] or [Rockafellar and Wets 1998, Chapter 4], i.e., the topology described by the following theorem.

Theorem A.3 [Badger and Lewis 2015, Theorem 2.5]. *There exists a metrizable topology on $\mathfrak{C}(\mathbb{R}^n)$ in which a sequence $(A_i)_{i=1}^\infty$ in $\mathfrak{C}(\mathbb{R}^n)$ converges to a set $A \in \mathfrak{C}(\mathbb{R}^n)$ if and only if*

$$\lim_{i\to\infty} \operatorname{ex}\big(A_i \cap B(0, r), A\big) = 0 \quad \textit{and} \quad \lim_{i\to\infty} \operatorname{ex}\big(A \cap B(0, r), A_i\big) = 0 \quad \textit{for all } r > 0.$$

Moreover, *in this topology*, $\mathfrak{C}(0)$ *is sequentially compact*; *i.e.*, *for any sequence $(A_i)_{i=1}^\infty$ in $\mathfrak{C}(0)$ there exists a subsequence $(A_{i_j})_{j=1}^\infty$ and $A \in \mathfrak{C}(0)$ such that $(A_{i_j})_{j=1}^\infty$ converges to A in the sense above.*

We write $A_i \to A$ or $A = \lim_{i\to\infty} A$ (in $\mathfrak{C}(\mathbb{R}^n)$) to denote that a sequence of $(A_i)_{i=1}^\infty$ in $\mathfrak{C}(\mathbb{R}^n)$ converges to a set $A \in \mathfrak{C}(\mathbb{R}^n)$ in the Attouch–Wets topology. If each set $A_i \in \mathfrak{C}(0)$, then we may write $A_i \to A$ in $\mathfrak{C}(0)$ to emphasize that $A \in \mathfrak{C}(0)$, as well.

Lemma A.4 [Badger and Lewis 2015, Lemma 2.6]. *Let $A, A_1, A_2, \ldots \in \mathfrak{C}(\mathbb{R}^n)$. The following statements are equivalent*:

(i) $A_i \to A$ *in* $\mathfrak{C}(\mathbb{R}^n)$.

(ii) $\lim_{i\to\infty} \widetilde{\mathrm{D}}^{x,r}[A_i, A] = 0$ *for all* $x \in \mathbb{R}^n$ *and for all* $r > 0$.

(iii) $\lim_{i\to\infty} \widetilde{\mathrm{D}}^{x_0,r_j}[A_i, A] = 0$ *for some* $x_0 \in \mathbb{R}^n$ *and for some sequence* $r_j \to \infty$.

The notions of tangent sets and pseudotangent sets of a closed set in the following definition are modeled on notions of tangent measures (introduced by Preiss [1987]) and pseudotangent measures (introduced by Kenig and Toro [1999]) of a Radon measure.

Definition A.5 [Badger and Lewis 2015, Definition 3.1]. Let $T \in \mathfrak{C}(0)$, let $A \in \mathfrak{C}(\mathbb{R}^n)$, and let $x \in A$. We say that T is a *pseudotangent set of A at x* if there exist sequences $x_i \in A$ and $r_i > 0$ such that $x_i \to x$, $r_i \to 0$, and

$$\frac{A - x_i}{r_i} \to T \quad \text{in } \mathfrak{C}(0).$$

If $x_i = x$ for all i, then we call T a *tangent set of A at x*. Let $\Psi\text{-}\operatorname{Tan}(A, x)$ and $\operatorname{Tan}(A, x)$ denote the collections of all pseudotangent sets of A at x and all tangent sets of A at x, respectively.

Lemma A.6 [Badger and Lewis 2015, Remark 3.3, Lemmas 3.4 and 3.5]. $\operatorname{Tan}(A, x)$ *and* $\Psi\text{-}\operatorname{Tan}(A, x)$ *are closed in* $\mathfrak{C}(0)$ *and are nonempty for all* $A \in \mathfrak{C}(\mathbb{R}^n)$ *and* $x \in A$. *Moreover*,

- *If* $T \in \operatorname{Tan}(A, x)$ *and* $\lambda > 0$, *then* $\lambda A \in \operatorname{Tan}(A, x)$.
- *If* $T \in \Psi\text{-}\operatorname{Tan}(A, x)$ *and* $\lambda > 0$, *then* $\lambda T \in \Psi\text{-}\operatorname{Tan}(A, x)$.
- *If* $T \in \Psi\text{-}\operatorname{Tan}(A, x)$ *and* $y \in T$, *then* $T - y \in \Psi\text{-}\operatorname{Tan}(A, x)$.

Reifenberg-type sets and Mattila–Vuorinen-type sets.

Definition A.7 [Badger and Lewis 2015, Definitions 4.1 and 4.7]. Let $A \subseteq \mathbb{R}^n$ be nonempty.

(i) A *local approximation class* $\mathcal{S}$ is a nonempty collection of closed sets in $\mathfrak{C}(0)$ such that $\mathcal{S}$ is a cone; that is, for all $S \in \mathcal{S}$ and $\lambda > 0$, $\lambda S \in \mathcal{S}$.

(ii) For every $x \in \mathbb{R}^n$ and $r > 0$, define the *bilateral approximability* $\Theta_A^{\mathcal{S}}(x, r)$ of A by $\mathcal{S}$ at location x and scale r by

$$\Theta_A^{\mathcal{S}} = \inf_{S \in \mathcal{S}} \widetilde{\mathrm{D}}^{x,r}[A, x + S] \in [0, \infty).$$

(iii) We say that $x \in A$ is an *$\mathcal{S}$ point* of A if $\lim_{r \downarrow 0} \Theta_A^{\mathcal{S}}(x, r) = 0$.

(iv) We say that A is *locally bilaterally ε-approximable by $\mathcal{S}$* if for every compact set $K \subseteq A$ there exists r_K such that $\Theta_A^{\mathcal{S}}(x, r) \le \varepsilon$ for all $x \in K$ and $0 < r \le r_K$.

(v) We say that A is *locally bilaterally well approximated by $\mathcal{S}$* if A is locally bilaterally ε-approximable by $\mathcal{S}$ for all $\varepsilon > 0$.

(vi) For every $x \in \mathbb{R}^n$ and $r > 0$, define the *unilateral approximability* $\beta_A^{\mathcal{S}}(x, r)$ of A by $\mathcal{S}$ at location x and scale r by

$$\beta_A^{\mathcal{S}}(x, r) = \inf_{S \in \mathcal{S}} \tilde{\mathrm{d}}^{x,r}(A, x + S) \in [0, 1].$$

(vii) We say that A is *locally unilaterally ε-approximable by $\mathcal{S}$* if for every compact set $K \subseteq A$ there exists r_K such that $\beta_A^{\mathcal{S}}(x, r) \le \varepsilon$ for all $x \in K$ and $0 < r \le r_K$.

(viii) We say that A is *locally unilaterally well approximated by $\mathcal{S}$* if A is locally unilaterally ε-approximable by $\mathcal{S}$ for all $\varepsilon > 0$.

Remark A.8. Sets that are bilaterally approximated by $\mathcal{S}$ are called *Reifenberg-type sets* and sets that are unilaterally approximated by $\mathcal{S}$ are called *Mattila–Vuorinen-type sets* with deference to pioneering work of Reifenberg [1960] and Mattila and Vuorinen [1990], which investigated, respectively, regularity of sets that admit locally uniform *bilateral* and *unilateral* approximations by $\mathcal{S} = G(n, m)$, the Grassmannian of m-dimensional subspaces of $\mathbb{R}^n$. The concept of (unilateral) approximation numbers first appeared in the work of Jones [1990] in connection with the analyst's traveling salesman theorem. For additional background, including examples of Reifenberg-type sets that have appeared in the literature, see the introduction of [Badger and Lewis 2015].

Remark A.9. For any nonempty closed set $A \subseteq \mathbb{R}^n$ and point $x \in A$, the set $\operatorname{Tan}(A, x)$ of tangent sets of A at x and the set $\Psi\text{-}\operatorname{Tan}(A, x)$ of pseudotangent sets of A at x are local approximation classes by Lemma A.6. We also note that from the definitions, it is immediate that any set $A \subseteq \mathbb{R}^n$ which is locally bilaterally well approximated by some local approximation class $\mathcal{S}$ is also locally unilaterally well approximated by $\mathcal{S}$.

The following essential properties of bilateral approximation numbers appear across a number of lemmas in [Badger and Lewis 2015, §4], which we consolidate into a single theorem statement; see Lemma 7.2 of that paper for the analogous properties of unilateral approximation numbers.

Lemma A.10 [Badger and Lewis 2015, §4, Remark 2.4]. *Let $\mathcal{S}$ be a local approximation class, let $A \subseteq \mathbb{R}^n$ be nonempty, let $x, y \in \mathbb{R}^n$, and let $r, s > 0$. Then we have the following properties*:

- *size*: $0 \le \Theta_A^{\mathcal{S}}(x, r) - \operatorname{dist}(x, A)/r \le 1$; *thus*, $0 \le \Theta_A^{\mathcal{S}}(x, r) \le 1$ *for all* $x \in A$.
- *scale invariance*: $\Theta_A^{\mathcal{S}}(x, r) = \Theta_{\lambda A}^{\mathcal{S}}(\lambda x, \lambda r)$ *for all* $\lambda > 0$.
- *translation invariance*: $\Theta_A^{\mathcal{S}}(x, r) = \Theta_{A+z}^{\mathcal{S}}(x + z, r)$ *for all* $z \in \mathbb{R}^n$.
- *closure*: $\Theta_A^{\mathcal{S}}(x, r) = \Theta_{\bar{A}}^{\mathcal{S}}(x, r)$.
- *quasimonotonicity*: *If* $B(x, r) \subseteq B(y, s)$ *and* $|x - y| \le ts$, *then*
$$\Theta_A^{\mathcal{S}}(x, r) \le \frac{s}{r}\big[t + (1 + t)\Theta_A^{\mathcal{S}}(y, (1 + t)s)\big].$$
In particular, *if* $r < s$, *then* $\Theta_A^{\mathcal{S}}(x, r) \le (s/r)\Theta_A^{\mathcal{S}}(x, s)$.
- *limits*: *If* $A, A_1, A_2, \ldots \in \mathfrak{C}(\mathbb{R}^n)$ *and* $A_i \to A$ *in* $\mathfrak{C}(\mathbb{R}^n)$, *then*
$$\frac{1}{1+\varepsilon} \limsup_{i\to\infty} \Theta_{A_i}^{\mathcal{S}}\left(x, \frac{r}{1+\varepsilon}\right) \le \Theta_A^{\mathcal{S}}(x, r) \le (1+\varepsilon) \liminf_{i\to\infty} \Theta_{A_i}^{\mathcal{S}}(x, r(1+\varepsilon)) \quad \textit{for all } \varepsilon > 0.$$

The notions of $\mathcal{S}$ points and locally bilaterally and unilaterally well-approximated sets admit the following characterizations in terms of tangent sets and pseudotangent sets. Here $\bar{\mathcal{S}}$ denotes the closure of $\mathcal{S}$ in $\mathfrak{C}(0)$ with respect to the Attouch–Wets topology.

Theorem A.11 [Badger and Lewis 2015, Corollaries 4.12 and 4.15, Lemma 7.7, Theorem 7.10]. *Let $\mathcal{S}$ be a local approximation class and let $A \subseteq \mathbb{R}^n$ be a nonempty set and let $x_0 \in A$. Then*

(i) *x_0 is an $\mathcal{S}$ point of A if and only if* $\operatorname{Tan}(\bar{A}, x_0) \subseteq \bar{\mathcal{S}}$;

(ii) *A is locally bilaterally well approximated by $\mathcal{S}$ if and only if*

$$\Psi\text{-}\operatorname{Tan}(\bar{A}, x) \subseteq \bar{\mathcal{S}} \quad \textit{for all } x \in A;$$

(iii) *A is locally unilaterally well approximated by $\mathcal{S}$ if and only if*

$$\Psi\text{-}\operatorname{Tan}(\bar{A}, x) \subseteq \{T \in \mathfrak{C}(0) : T \subseteq S \textit{ for some } S \in \bar{\mathcal{S}}\} \quad \textit{for all } x \in A.$$

Detectability and structure theorems for Reifenberg-type sets.

Definition A.12 [Badger and Lewis 2015, Definition 5.8]. Let $\mathcal{T}$ and $\mathcal{S}$ be local approximation classes. We say that *$\mathcal{T}$ points are detectable in $\mathcal{S}$* if there exist a constant $\phi > 0$ and a function $\Phi : (0, 1) \to (0, \infty)$ with $\liminf_{s\to 0+} \Phi(s) = 0$ such that if $S \in \mathcal{S}$ and $\Theta^{\mathcal{T}}_{S}(0, r) < \phi$, then $\Theta^{\mathcal{T}}_{S}(0, sr) < \Phi(s)$ for all $s \in (0, 1)$. To emphasize a choice of ϕ and Φ, we may say that *$\mathcal{T}$ points are (ϕ, Φ) detectable in $\mathcal{S}$*.

Definition A.13 [Badger and Lewis 2015, Definition 5.1]. Let $\mathcal{T}$ be a local approximation class. The *bilateral singular class of $\mathcal{T}$* is the local approximation class $\mathcal{T}^{\perp}$ given by

$$\mathcal{T}^{\perp} = \Big\{Z \in \mathfrak{C}(0) : \liminf_{r\downarrow 0} \Theta^{\mathcal{T}}_{Z}(0, r) > 0\Big\} = \{Z \in \mathfrak{C}(0) : \operatorname{Tan}(Z, 0) \cap \overline{\mathcal{T}} = \varnothing\}.$$

The following structure theorem decomposes a set $A \subseteq \mathbb{R}^n$ that is locally bilaterally well approximated by $\mathcal{S}$ into an open "regular part" $A_{\overline{\mathcal{T}}}$ and closed "singular part" $A_{\mathcal{T}^{\perp}}$, on the condition that "regular" $\mathcal{T}$ points are detectable in $\mathcal{S}$.

Theorem A.14 [Badger and Lewis 2015, Theorem 6.2, Corollaries 6.6 and 5.12]. *Let $\mathcal{T}$ and $\mathcal{S}$ be local approximation classes. Suppose $\mathcal{T}$ points are (ϕ, Φ) detectable in $\mathcal{S}$. If $A \subseteq \mathbb{R}^n$ is locally bilaterally well approximated by $\mathcal{S}$, then A can be written as a disjoint union*

$$A = A_{\overline{\mathcal{T}}} \cup A_{\mathcal{T}^{\perp}} \quad (A_{\overline{\mathcal{T}}} \cap A_{\mathcal{T}^{\perp}} = \varnothing),$$

where

(i) $\Psi\text{-}\operatorname{Tan}(\bar{A}, x) \subseteq \bar{\mathcal{S}} \cap \overline{\mathcal{T}}$ *for all $x \in A_{\overline{\mathcal{T}}}$, and*

(ii) $\operatorname{Tan}(\bar{A}, x) \subseteq \bar{\mathcal{S}} \cap \mathcal{T}^{\perp} = \{S \in \bar{\mathcal{S}} : \Theta^{\mathcal{T}}_{S}(0, r) \geq \phi \textit{ for all } r > 0\}$ *for all $x \in A_{\mathcal{T}^{\perp}}$.*

Moreover:

(iii) *$A_{\overline{\mathcal{T}}}$ is relatively open in A and $A_{\overline{\mathcal{T}}}$ is locally bilaterally well approximated by $\mathcal{T}$.*

(iv) *A is locally bilaterally well approximated **along** $A_{\mathcal{T}^{\perp}}$ by $\bar{\mathcal{S}} \cap \mathcal{T}^{\perp}$ in the sense that*

$$\limsup_{r\downarrow 0} \sup_{x\in K} \Theta^{\bar{\mathcal{S}}\cap\mathcal{T}^{\perp}}_{A}(x, r) = 0$$

for all compact sets $K \subseteq A_{\mathcal{T}^{\perp}}$.

Remark A.15. Suppose $\mathcal{T}$ points are (ϕ, Φ) detectable in $\mathcal{S}$ and A is locally bilaterally well approximated by $\mathcal{S}$. From the proof that $A_{\overline{\mathcal{T}}}$ is open in the proof of [Badger and Lewis 2015, Theorem 6.2], there exist constants $\alpha, \beta > 0$ depending only on ϕ and Φ such that if $\Theta^{\mathcal{S}}_{A}(x, r') < \alpha$ for all $0 < r' \leq r$ and $\Theta^{\mathcal{T}}_{A}(x, r) < \beta$ for some $x \in A$ and $r > 0$, then $x \in A_{\overline{\mathcal{T}}}$.

A local approximation class $\mathcal{S}$ is called *translation invariant* if for all $S \in \mathcal{S}$ and $x \in S$, $S - x \in \mathcal{S}$. It is an exercise to show that if $\mathcal{S}$ is translation invariant, then its closure $\bar{\mathcal{S}}$ is translation invariant, as well. If $\mathcal{T}$ and $\mathcal{S}$ are local approximation classes such that

$$\bar{\mathcal{S}} \text{ is translation invariant, and } \mathcal{T} \text{ points are } (\phi, \Phi) \text{ detectable in } \mathcal{S}, \tag{A-1}$$

then every set $X \in \bar{\mathcal{S}}$ is locally (in fact, globally) bilaterally well approximated by $\mathcal{S}$, whence $X = X_{\mathcal{T}} \cup X_{\mathcal{T}^\perp}$ and $X_{\mathcal{T}^\perp}$ is closed (since X is closed) by Theorem A.14.

Definition A.16 [Badger and Lewis 2015, Definition 7.12]. Let $\mathcal{T}$ and $\mathcal{S}$ be local approximation classes. Assume (A-1). We define the local approximation class of $\mathcal{T}$ *singular parts of sets in* $\bar{\mathcal{S}}$ by $\operatorname{sing}_{\mathcal{T}} \bar{\mathcal{S}} = \{X_{\mathcal{T}^\perp} : X \in \bar{\mathcal{S}} \text{ and } 0 \in X_{\mathcal{T}^\perp}\}$.

Theorem A.17 [Badger and Lewis 2015, Theorem 7.14]. *Let $\mathcal{T}$ and $\mathcal{S}$ be local approximation classes. Assume* (A-1). *If $A \subseteq \mathbb{R}^n$ is locally bilaterally well approximated by $\mathcal{S}$, then $A_{\mathcal{T}^\perp}$ is locally unilaterally well approximated by* $\operatorname{sing}_{\mathcal{T}} \bar{\mathcal{S}}$.

Covering profiles and dimension bounds for Mattila–Vuorinen-type sets. Finally, we record two upper bounds on the dimension of sets that are locally unilaterally well approximated by a local approximation class $\mathcal{S}$ with a uniform covering profile. Additional quantitative bounds for locally unilaterally ε-approximable sets may be found in [Badger and Lewis 2015, §8].

For reference, let us recall a definition of Minkowski dimension; e.g., see [Mattila 1995].

Definition A.18. Let $A \subseteq \mathbb{R}^n$, let $x \in \mathbb{R}^n$, and let $r, s > 0$. The (*intrinsic*) *s-covering number of A* is defined by

$$\mathrm{N}(A, s) := \min\left\{k \geq 0 : A \subseteq \bigcup_{i=1}^{k} B(a_i, s) \text{ for some } a_i \in A\right\}.$$

For bounded sets $A \subseteq \mathbb{R}^n$, the *upper Minkowski dimension of A* is given by

$$\overline{\dim}_M(A) = \limsup_{s \downarrow 0} \frac{\log(\mathrm{N}(A, s))}{\log(1/s)}.$$

For unbounded sets $A \subseteq \mathbb{R}^n$, the *upper Minkowski dimension of A* is given by

$$\overline{\dim}_M(A) = \lim_{t \uparrow \infty} (\overline{\dim}_M A \cap B(0, t)).$$

Letting $\dim_H(A)$ denote the usual *Hausdorff dimension* of a set $A \subseteq \mathbb{R}^n$,

$$0 \leq \dim_H(A) \leq \overline{\dim}_M(A) \leq n \quad \text{for all } A \subseteq \mathbb{R}^n,$$

with $\dim_H(A) < \overline{\dim}_M(A)$ for certain sets. For the definition of Hausdorff dimension, several equivalent definitions of Minkowski dimension, and related results, we refer the reader to [Mattila 1995].

Definition A.19 [Badger and Lewis 2015, Definitions 8.2 and 8.4]. Let $\mathcal{S}$ be a local approximation class. We say that *$\mathcal{S}$ has an (α, C, s_0) covering profile* for some $\alpha > 0$, $C > 0$, and $s_0 \in (0, 1]$ provided $\mathrm{N}(S \cap B(0, r), sr) \leq C s^{-\alpha}$ for all $S \in \mathcal{S}$, $r > 0$, and $s \in (0, s_0]$.

Theorem A.20 [Badger and Lewis 2015, Corollary 8.9]. *Let $\mathcal{S}$ be a local approximation class such that $\mathcal{S}$ has an (α, C, s_0) covering profile. If $A \subseteq \mathbb{R}^n$ is closed and A is locally unilaterally well approximated by $\mathcal{S}$, then* $\overline{\dim}_M(A) \leq \alpha$.

Theorem A.21 [Badger and Lewis 2015, Corollary 8.12]. *Let $\mathcal{S}$ be a local approximation class such that $\mathcal{S}$ has an (α, C, s_0) covering profile. If the subspace topology on $A \subseteq \mathbb{R}^n$ is σ-compact and A is locally unilaterally well approximated by $\mathcal{S}$, then* $\dim_H(A) \leq \alpha$.

Appendix B: Limits of complimentary NTA domains

For reference, let us recall that a connected open set $\Omega \subset \mathbb{R}^n$ is called an *NTA domain* (see Definition 6.1 and Remark 6.2) if there exist constants $M > 1$ and $R > 0$ for which the following are true:

(i) Ω satisfies the *corkscrew condition*: for all $Q \in \partial\Omega$ and $0 < r < R$, there exists $x \in \Omega \cap B(Q, r)$ such that $\operatorname{dist}(x, \partial\Omega) > M^{-1}r$.

(ii) $\mathbb{R}^n \setminus \Omega$ satisfies the corkscrew condition.

(iii) Ω satisfies the *Harnack chain condition*: if $x_1, x_2 \in \Omega \cap B\big(Q, \frac{1}{4}r\big)$ for some $Q \in \partial\Omega$ and $0 < r < R$, and $\operatorname{dist}(x_1, \partial\Omega) > \delta$, $\operatorname{dist}(x_2, \partial\Omega) > \delta$, and $|x_1 - x_2| < 2^l\delta$ for some $\delta > 0$ and $l \geq 1$, then there exists a chain of no more than Ml overlapping balls connecting x_1 to x_2 in Ω such that for each ball $B = B(x, s)$ in the chain

$$M^{-1}s < \operatorname{gap}(B, \partial\Omega) < Ms, \qquad \operatorname{gap}(B, \partial\Omega) = \inf_{x \in B} \inf_{y \in \partial\Omega} |x - y|,$$

$$\operatorname{diam} B > M^{-1} \min\{\operatorname{dist}(x_1, \partial\Omega), \operatorname{dist}(x_2, \partial\Omega)\}, \quad \operatorname{diam} B = \sup_{x, y \in B} |x - y|.$$

The constants M and R are called *NTA constants* of Ω, and the value $R = \infty$ is allowed when $\partial\Omega$ is unbounded. Lemma 6.3 asserts that if $\mathbb{R}^n \setminus \Gamma_i = \Omega_i^+ \cup \Omega_i^-$, where Ω_i^+ and Ω_i^- are complimentary NTA domains with NTA constants M and R independent of i, and $\Gamma_i \to \Gamma \neq \varnothing$ in the Attouch–Wets topology, then $\mathbb{R}^n \setminus \Gamma = \Omega^+ \cup \Omega^-$, where Ω^+ and Ω^- are complimentary NTA domains with constants $2M$ and R.

Proof of Lemma 6.3. Assume that we are given a sequence $(\Gamma_i, \Omega_i^+, \Omega_i^-)$, constants M and R, and a set Γ satisfying the hypothesis of the lemma. We note and will frequently use below that $\mathbb{R}^n \setminus \overline{\Omega}_i^{\pm} = \Omega_i^{\mp}$, $\Gamma_i = \partial\Omega_i^{\pm}$, and $\mathbb{R}^n = \Omega_i^+ \cup \Gamma_i \cup \Omega_i^-$ by the separation condition on Γ_i and the corkscrew conditions for $\Omega_i^{\pm}$.

Step 0 (*definition of Ω^+ and Ω^-*). Because the sequence $(\Gamma_i)_{i=1}^{\infty}$ does not escape to infinity (as $\Gamma_i \to \Gamma$), neither do $(\overline{\Omega}_i^{\pm})_{i=1}^{\infty}$. Thus, there is a subsequence of $(\Gamma_i, \Omega_i^+, \Omega_i^-)$ (which we relabel) and nonempty closed sets $F^+, F^- \subseteq \mathbb{R}^n$ such that $\overline{\Omega}_i^{\pm} \to F^{\pm}$. Here and below, convergence of a sequence of nonempty closed sets in $\mathbb{R}^n$ is always taken with respect to the Attouch–Wets topology; we refer the reader to the first two subsections of Appendix A for a brief introduction to this topology and to [Rockafellar and Wets 1998, Chapter 4] or [Beer 1993, Chapter 3] for the rest of the story. Consider the open sets Ω^+ and Ω^- defined by

$$\Omega^+ = \mathbb{R}^n \setminus F^- \quad \text{and} \quad \Omega^- = \mathbb{R}^n \setminus F^+.$$

We will show that $\mathbb{R}^n \setminus \Gamma = \Omega^+ \cup \Omega^-$ and Ω^+ and Ω^- are complementary NTA domains with NTA constants $2M$ and R.

Step $\frac{1}{2}$ (Ω^+, Γ, *and* Ω^- *are disjoint*). First, because $\Gamma_i \subseteq \overline{\Omega}_i^{\pm}$ for all $i \geq 1$, $\Gamma_i \to \Gamma$, and $\overline{\Omega}_i^{\pm} \to F^{\pm}$, we have $\Gamma \subseteq F^{\pm}$, as well. Hence, by the definition of $\Omega^{\pm}$,

$$\Gamma \cap \Omega^{\pm} \subseteq F^{\mp} \cap \Omega^{\pm} = F^{\mp} \setminus F^{\mp} = \varnothing.$$

Next, suppose that $x \in \Omega^{\pm}$. Then $x \notin F^{\mp}$, whence $\operatorname{dist}(x, F^{\mp}) = \delta$ for some $\delta > 0$. Since $\overline{\Omega}_i^{\mp} \to F^{\mp}$, it follows that $\operatorname{dist}(x, \overline{\Omega}_i^{\mp}) \geq \frac{1}{2}\delta$ for all $i \gg 1$. In particular, $x \in \Omega_i^{\pm} \subseteq \overline{\Omega}_i^{\pm}$ for all $i \gg 1$, because $\mathbb{R}^n \setminus \overline{\Omega}_i^{\mp} = \Omega_i^{\pm}$. Since $\overline{\Omega}_i^{\pm} \to F^{\pm}$, we obtain $x \in F^{\pm}$. Thus, $x \notin \Omega^{\mp}$ whenever $x \in \Omega^{\pm}$. We conclude that $\Omega^+ \cap \Omega^- = \varnothing$.

Step 1 ($\mathbb{R}^n = \Omega^+ \cup \Gamma \cup \Omega^-$). Let $x \in \mathbb{R}^n$. Because $\mathbb{R}^n = \overline{\Omega}_i^+ \cup \overline{\Omega}_i^-$, at least one of the following alternatives occur: $x \in \overline{\Omega}_i^+$ for infinitely many i or $x \in \overline{\Omega}_i^-$ for infinitely many i. Hence $x \in F^+$ or $x \in F^-$, since $\overline{\Omega}_i^+ \to F^+$ and $\overline{\Omega}_i^- \to F^-$. As x was arbitrary, we have

$$\mathbb{R}^n = F^+ \cup F^- = (F^+ \setminus F^-) \cup (F^+ \cap F^-) \cup (F^- \setminus F^+) = \Omega^+ \cup (F^+ \cap F^-) \cap \Omega^-.$$

Therefore, as soon as we show that $F^+ \cap F^- = \Gamma$, we will have $\mathbb{R}^n = \Omega^+ \cup \Gamma \cup \Omega^+$.

To prove that $F^+ \cap F^- \subseteq \Gamma$, suppose that $y \in F^+ \cap F^-$. Since $\overline{\Omega}_i^{\pm} \to F^{\pm}$, we can locate points $y_i^{\pm} \in \overline{\Omega}_i^{\pm}$ such that $y_i^{\pm} \to y$. The line segment between y^+ and y^- must intersect $\Gamma_i = \overline{\Omega}_i^+ \cap \overline{\Omega}_i^-$, say $Q_i \in [y_i^+, y_i^-] \cap \Gamma_i$. Then $Q_i \to y$, and because $\Gamma_i \to \Gamma$, we obtain $y \in \Gamma$. Thus, $F^+ \cap F^- \subseteq \Gamma$.

To prove that $\Gamma \subseteq F^+ \cap F^-$, suppose that $z \in \Gamma$. Since $\Gamma_i \to \Gamma$, there exists $z_i \in \Gamma_i$ such that $z_i \to \Gamma$. Because $\Gamma_i = \partial\Omega^+ = \partial\Omega^-$, we can locate points $z_i^{\pm} \in \Omega_i^{\pm} \cap B(z_i, 1/i)$. Then $z_i^{\pm} \to z$, and because $\overline{\Omega}_i^{\pm} \to F^{\pm}$, we obtain $z \in F^+ \cap F^-$. Thus, $\Gamma \subseteq F^+ \cap F^-$.

Step $\frac{3}{2}$ ($\partial\Omega^{\pm} \subseteq \Gamma$). Since Ω^+ and Ω^- are open and disjoint by Steps 0 and $\frac{1}{2}$, $\Omega^{\pm}$ coincides with the interior of $\Omega^{\pm}$ and $\Omega^{\mp}$ is contained in the exterior of $\Omega\pm$. Therefore, the boundary of $\Omega^{\pm}$ must be contained in $\mathbb{R}^n \setminus (\Omega^{\pm} \cup \Omega^{\mp}) = \Gamma$ by Step 1.

Step 2 (*Corkscrew condition for* $\Omega^{\pm}$). Suppose that $Q \in \partial\Omega^{\pm}$ and $0 < r < R$. By Step $\frac{3}{2}$, $Q \in \Gamma$. Since $\Gamma_i \to \Gamma$, there exists $Q_i \in \Gamma_i = \partial\Omega_i^{\pm}$ such that $Q_i \to Q$. By the corkscrew condition for $\Omega_i^{\pm}$, there exists a point $y_i^{\pm} \in \Omega_i^{\pm} \cap B\big(Q_i, \frac{3}{4}r\big)$ such that

$$\operatorname{dist}(y_i^{\pm}, \overline{\Omega}_i^{\mp}) = \operatorname{dist}(y_i^{\pm}, \partial\Omega_i^{\pm}) > \frac{3r}{4M}.$$

Assume $i \geq 1$ is sufficiently large that

$$y_i^{\pm} \in B\big(Q_i, \tfrac{3}{4}r\big) \subset B\big(Q, \tfrac{4}{5}r\big) \quad \text{and} \quad \operatorname{dist}(y_i^{\pm}, F^{\mp}) \leq |y_i^{\pm} - Q| < \tfrac{4}{5}r.$$

Then $\operatorname{dist}(y_i^{\pm}, F^{\mp}) = \operatorname{dist}\big(y_i^{\pm}, F^{\mp} \cap B\big(Q, \frac{4}{5}r\big)\big)$. Hence, by the triangle inequality for excess,

$$\begin{aligned} \operatorname{dist}(y_i^{\pm}, \overline{\Omega}_i^{\mp}) &\leq \operatorname{dist}\big(y_i^{\pm}, F^{\mp} \cap B\big(Q, \tfrac{4}{5}r\big)\big) + \operatorname{ex}\big(F^{\mp} \cap B\big(Q, \tfrac{4}{5}r\big), \overline{\Omega}_i^{\mp}\big) \\ &= \operatorname{dist}(y_i^{\pm}, F^{\mp}) + \operatorname{ex}\big(F^{\mp} \cap B\big(Q, \tfrac{4}{5}r\big), \overline{\Omega}_i^{\mp}\big). \end{aligned}$$

The last term vanishes as $i \to \infty$, since $\overline{\Omega}_i^{\mp} \to F^{\mp}$ in the Attouch–Wets topology. Thus,

$$\operatorname{dist}(y_i^{\pm}, F^{\mp}) \geq \operatorname{dist}(y_i^{\pm}, \overline{\Omega}_i^{\mp}) - \operatorname{ex}\big(F^{\mp} \cap B\big(Q, \tfrac{4}{5}r\big), \overline{\Omega}_i^{\mp}\big) > \frac{2r}{3M} \quad \text{for all } i \gg 1. \tag{B-1}$$

By compactness, we can choose subsequences $(y_{ij}^{\pm})_{j=1}^{\infty}$ of $(y_i^{\pm})_{i=1}^{\infty}$ such that $y_{ij}^{\pm} \to y^{\pm}$ for some $y^{\pm} \in \overline{B\big(Q, \frac{4}{5}r\big)} \subset B(Q, r)$. By (B-1), it follows that $\operatorname{dist}(y^{\pm}, F^{\mp}) \geq 2r/(3M) > r/(2M)$. Thus, $y^{\pm} \in \Omega^{\pm} \cap B(Q, r)$ and

$$\operatorname{dist}(y^{\pm}, \partial\Omega^{\pm}) = \operatorname{dist}(y^{\pm}, F^{\mp}) > \frac{r}{2M}.$$

Therefore, $\Omega^{\pm}$ satisfies the corkscrew condition with constants $2M$ and R. We note that by an obvious modification of the argument, one can show that $\Omega^{\pm}$ satisfies the corkscrew condition with constants M' and R for all $M' > M$.

Step $\frac{5}{2}$ ($\partial\Omega^{\pm} = \Gamma$). By Step $\frac{3}{2}$, $\partial\Omega^{\pm} \subseteq \Gamma$. To see that $\Gamma \subseteq \partial\Omega^{\pm}$, suppose that $Q \in \Gamma$. By the proof of Step 2, the ball $B(Q, r)$ contains points in $\Omega^{\pm}$ for all $0 < r < R$. Because $\Omega^{\mp}$ is disjoint from $\Omega^{\pm}$, it follows that $Q \in \partial\Omega^{\pm}$. We conclude that $\partial\Omega^{\pm} = \Gamma$.

Step 3 (*Harnack chain condition for* $\Omega^{\pm}$). Assume that $x_1, x_2 \in \Omega^{\pm} \cap B\big(Q, \frac{1}{4}r\big)$ for some $Q \in \Gamma = \partial\Omega^{\pm}$ and $0 < r < R$. Furthermore, assume that $\delta_1 := \operatorname{dist}(x_1, \partial\Omega) > \delta$, $\delta_2 := \operatorname{dist}(x_2, \partial\Omega) > \delta$, and $|x_1 - x_2| < 2^l\delta$ for some $\delta > 0$ and $l \geq 1$. We must show that x_1 can be connected to x_2 in $\Omega^{\pm}$ by a "short" chain of balls in $\Omega^{\pm}$ remaining "far away" from the boundary $\partial\Omega^{\pm}$, or equivalently, remaining "far away" from $F^{\mp}$. Since $\Gamma_i \to \Gamma$, there exists $Q_i \in \Omega_i^{\pm}$ such that $Q_i \to Q$. Because $\overline{\Omega}_i^{\mp} \to F^{\mp}$ in the Attouch–Wets topology, for all $i \geq 1$ sufficiently large, $r(1 + |Q - Q_i|) < R$, $x_1, x_2 \in \Omega_i^{\pm} \cap B\big(Q_i, \frac{1}{4}r(1 + |Q - Q_i|)\big)$, and

$$\operatorname{dist}(x_1, \partial\Omega_i^{\pm}) > \tfrac{1}{2}\delta_1 > \tfrac{1}{2}\delta, \quad \operatorname{dist}(x_2, \partial\Omega_i^{\pm}) > \tfrac{1}{2}\delta_2 > \tfrac{1}{2}\delta, \quad |x_1 - x_2| < 2^l\delta = 2^{l+1}\tfrac{1}{2}\delta.$$

(The details are similar to those written in the proof of the corkscrew condition in Step 2.) By the Harnack chain condition for $\Omega_i^{\pm}$, we can find a chain of no more than $M(l+1) \leq 2Ml$ balls connecting x_1 to x_2 in $\Omega_i^{\pm}$ such that for each ball $B = B(x, s)$ in the chain,

$$M^{-1}s < \operatorname{gap}(B, \partial\Omega_i^{\pm}) < Ms$$

and

$$\operatorname{diam} B > M^{-1} \min\{\operatorname{dist}(x_1, \partial\Omega_i^{\pm}), \operatorname{dist}(x_2, \partial\Omega_i^{\pm})\}.$$

Since $\overline{\Omega}_i^{\mp} \to F^{\mp}$ in the Attouch–Wets topology, it follows that for all sufficiently large i,

$$(2M)^{-1}s < \operatorname{gap}(B, \partial\Omega^{\pm}) < 2Ms$$

and

$$\operatorname{diam} B > (2M)^{-1} \min\{\operatorname{dist}(x_1, \partial\Omega^{\pm}), \operatorname{dist}(x_2, \partial\Omega^{\pm})\}.$$

(Again, the details are similar to those in Step 2.) By the gap condition, we also know each ball in the chain belongs to $\Omega^{\pm}$. Therefore, $\Omega^{\pm}$ satisfies the Harnack chain condition with constants $2M$ and R. We remark that given the discrete nature of the constant in the Harnack chain condition (counting balls), we cannot expect to be able to replace $2M$ by λM for arbitrary $\lambda > 1$.

Step 4 (Ω^+ *and* Ω^- *are connected*). It is well known that every NTA domain is a uniform domain with constants that depend only on the interior corkscrew condition and Harnack chain condition; e.g., see [Azzam et al. 2017a, Theorem 2.15]. Explicitly, this means that for every $M > 1$ and $R > 0$, there exist $C > 1$ and $c \in (0, 1)$ such that for every NTA domain $\Omega \subseteq \mathbb{R}^n$ with NTA constants M and R, and for every $x_0, x_1 \in \Omega$, there exists a continuous path $\gamma : [0, 1] \to \Omega$ such that $\gamma(0) = x_0$, $\gamma(1) = x_1$, $\operatorname{length}(\gamma) \le C|x_0 - x_1|$, and $\operatorname{dist}(\gamma(t), \partial\Omega) \ge c \min\{\operatorname{dist}(x_0, \partial\Omega), \operatorname{dist}(x_1, \partial\Omega)\}$ for all $t \in [0, 1]$.

Let x_0 and x_1 be arbitrary distinct points in $\Omega^\pm$, and set

$$\delta = \min\bigl\{\operatorname{dist}(x_0, \partial\Omega^\pm), \operatorname{dist}(x_1, \partial\Omega^\pm)\bigr\} = \min\bigl\{\operatorname{dist}(x_0, F^\mp), \operatorname{dist}(x_1, F^\mp)\bigr\}.$$

Assign $B = B(x_0, 3C|x_0 - x_1| + 3\delta)$, where C is the constant from the previous paragraph. Note that B contains x_0, x_1, and every path passing through x_0 of length no greater than $C|x_0 - x_1|$, and the closest point in $F^\mp$ for each item listed above, with room to spare. Since $\overline{\Omega}_i^\mp \to F^\mp$ in the Attouch–Wets topology,

$$\operatorname{ex}(\overline{\Omega}_i^\mp \cap B, F^\mp) < \tfrac{1}{3}c\delta \quad \text{and} \quad \operatorname{ex}(F^\mp \cap B, \overline{\Omega}_i^\mp) < \tfrac{1}{3}c\delta \quad \text{for all } i \gg 1, \tag{B-2}$$

where c is the constant from the previous paragraph. Pick any i such that (B-2) holds. Then $\operatorname{dist}(x_0, \overline{\Omega}_i^\mp) \ge \bigl(1 - \frac{1}{3}c\bigr)\delta > \frac{2}{3}\delta$ and $\operatorname{dist}(x_1, \overline{\Omega}_i^\mp) \ge \bigl(1 - \frac{1}{3}c\bigr)\delta > \frac{2}{3}\delta$. In particular, $x_0, x_1 \in \Omega_i^\pm$ and

$$\min\bigl\{\operatorname{dist}(x_0, \partial\Omega_i^\pm), \operatorname{dist}(x_1, \partial\Omega_i^\pm)\bigr\} > \tfrac{2}{3}\delta.$$

Since $\Omega_i^\pm$ is an NTA domain with NTA constants M and R, by the previous paragraph we can find a continuous path $\gamma : [0, 1] \to \Omega_i^\pm$ such that $\gamma(0) = x_0$, $\gamma(1) = x_1$, $\operatorname{length}(\gamma) \le C|x_0 - x_1|$, and $\operatorname{dist}(\gamma(t), \overline{\Omega}_i^\mp) = \operatorname{dist}(\gamma(t), \partial\Omega_i^\pm) > \frac{2}{3}c\delta$ for all $t \in [0, 1]$. Using (B-2) once again, we obtain $\operatorname{dist}(\gamma(t), F^\mp) > \frac{1}{3}c\delta$ for all $t \in [0, 1]$. In particular, $\gamma(t) \in \Omega^\pm$ for all $t \in [0, 1]$. Thus, γ is a continuous path joining x_0 to x_1 inside the set $\Omega^\pm$. Since x_0 and x_1 were fixed arbitrarily, we conclude that $\Omega^\pm$ is connected.

Conclusion. We have shown that $\mathbb{R}^n \setminus \Gamma = \Omega^+ \cup \Omega^-$ (Step 1), where Ω^+ and Ω^- are open (Step 0), connected (Step 4), and satisfy corkscrew (Step 2) and Harnack chain conditions (Step 3) with constants $2M$ and R. Therefore, $\mathbb{R}^n \setminus \Gamma = \Omega^+ \cup \Omega^-$ is the union of complimentary NTA domains Ω^+ and Ω^- with NTA constants $2M$ and R, as desired. □

Acknowledgements

A portion of this research was completed while Engelstein was visiting the University of Washington during the spring of 2015. He thanks the Mathematics Department at UW for their hospitality. Badger acknowledges and thanks Stephen Lewis for many insightful conversations about local set approximation, which have duly influenced the present manuscript. The authors would like to thank an anonymous referee for his or her critical feedback, which has led to an improved exposition of these results.

References

[Almgren 1979] F. J. Almgren, Jr., "Dirichlet's problem for multiple valued functions and the regularity of mass minimizing integral currents", pp. 1–6 in *Minimal submanifolds and geodesics: proceedings of the Japan-United States Seminar* (Tokyo, 1977), edited by M. Obata, North-Holland, New York, 1979. MR Zbl

[Axler et al. 2001] S. Axler, P. Bourdon, and W. Ramey, *Harmonic function theory*, 2nd ed., Graduate Texts in Mathematics **137**, Springer, 2001. MR Zbl

[Azzam and Mourgoglou 2015] J. Azzam and M. Mourgoglou, "Tangent measures and absolute continuity of harmonic measure", preprint, 2015. To appear in *Rev. Mat. Iberoam.* arXiv

[Azzam et al. 2016] J. Azzam, M. Mourgoglou, X. Tolsa, and A. Volberg, "On a two-phase problem for harmonic measure in general domains", preprint, 2016. arXiv

[Azzam et al. 2017a] J. Azzam, S. Hofmann, J. M. Martell, K. Nyström, and T. Toro, "A new characterization of chord-arc domains", *J. Eur. Math. Soc.* (*JEMS*) **19**:4 (2017), 967–981. MR Zbl

[Azzam et al. 2017b] J. Azzam, M. Mourgoglou, and X. Tolsa, "Mutual absolute continuity of interior and exterior harmonic measure implies rectifiability", *Comm. Pure Appl. Math.* (online publication February 2017).

[Badger 2011] M. Badger, "Harmonic polynomials and tangent measures of harmonic measure", *Rev. Mat. Iberoam.* **27**:3 (2011), 841–870. MR Zbl

[Badger 2012] M. Badger, "Null sets of harmonic measure on NTA domains: Lipschitz approximation revisited", *Math. Z.* **270**:1-2 (2012), 241–262. MR Zbl

[Badger 2013] M. Badger, "Flat points in zero sets of harmonic polynomials and harmonic measure from two sides", *J. Lond. Math. Soc.* (2) **87**:1 (2013), 111–137. MR Zbl

[Badger and Lewis 2015] M. Badger and S. Lewis, "Local set approximation: Mattila–Vuorinen type sets, Reifenberg type sets, and tangent sets", *Forum Math. Sigma* **3** (2015), art. id. e24. MR Zbl

[Beer 1993] G. Beer, *Topologies on closed and closed convex sets*, Mathematics and its Applications **268**, Kluwer Academic Publishers Group, Dordrecht, 1993. MR Zbl

[Bishop 1992] C. J. Bishop, "Some questions concerning harmonic measure", pp. 89–97 in *Partial differential equations with minimal smoothness and applications* (Chicago, IL, 1990), edited by B. Dahlberg et al., IMA Vol. Math. Appl. **42**, Springer, 1992. MR Zbl

[Bortz and Hofmann 2016] S. Bortz and S. Hofmann, "A singular integral approach to a two phase free boundary problem", *Proc. Amer. Math. Soc.* **144**:9 (2016), 3959–3973. MR Zbl

[Capogna et al. 2005] L. Capogna, C. E. Kenig, and L. Lanzani, *Harmonic measure: geometric and analytic points of view*, University Lecture Series **35**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[Cheeger et al. 2015] J. Cheeger, A. Naber, and D. Valtorta, "Critical sets of elliptic equations", *Comm. Pure Appl. Math.* **68**:2 (2015), 173–209. MR Zbl

[David and Toro 1999] G. David and T. Toro, "Reifenberg flat metric spaces, snowballs, and embeddings", *Math. Ann.* **315**:4 (1999), 641–710. MR Zbl

[David and Toro 2012] G. David and T. Toro, *Reifenberg parameterizations for sets with holes*, Mem. Amer. Math. Soc. **1012**, Amer. Math. Soc., Providence, RI, 2012. MR Zbl

[David et al. 2001] G. David, C. Kenig, and T. Toro, "Asymptotically optimally doubling measures and Reifenberg flat sets with vanishing constant", *Comm. Pure Appl. Math.* **54**:4 (2001), 385–449. MR Zbl

[Engelstein 2016] M. Engelstein, "A two-phase free boundary problem for harmonic measure", *Ann. Sci. Éc. Norm. Supér.* (4) **49**:4 (2016), 859–905. MR Zbl

[Garnett and Marshall 2005] J. B. Garnett and D. E. Marshall, *Harmonic measure*, New Mathematical Monographs **2**, Cambridge University Press, 2005. MR Zbl

[Girela-Sarrión and Tolsa 2016] D. Girela-Sarrión and X. Tolsa, "The Riesz transform and quantitative rectifiability for general Radon measures", preprint, 2016. arXiv

[Han 2007] Q. Han, "Nodal sets of harmonic functions", *Pure Appl. Math. Q.* **3**:3 (2007), 647–688. MR Zbl

[Jerison and Kenig 1982] D. S. Jerison and C. E. Kenig, "Boundary behavior of harmonic functions in nontangentially accessible domains", *Adv. in Math.* **46**:1 (1982), 80–147. MR Zbl

[Jones 1990] P. W. Jones, "Rectifiable sets and the traveling salesman problem", *Invent. Math.* **102**:1 (1990), 1–15. MR Zbl

[Kenig and Toro 1999] C. E. Kenig and T. Toro, "Free boundary regularity for harmonic measures and Poisson kernels", *Ann. of Math.* (2) **150**:2 (1999), 369–454. MR Zbl

[Kenig and Toro 2006] C. Kenig and T. Toro, "Free boundary regularity below the continuous threshold: 2-phase problems", *J. Reine Angew. Math.* **596** (2006), 1–44. MR Zbl

[Kenig et al. 2009] C. Kenig, D. Preiss, and T. Toro, "Boundary structure and size in terms of interior and exterior harmonic measures in higher dimensions", *J. Amer. Math. Soc.* **22**:3 (2009), 771–796. MR Zbl

[Kollár 1999] J. Kollár, "An effective Łojasiewicz inequality for real polynomials", *Period. Math. Hungar.* **38**:3 (1999), 213–221. MR Zbl

[Lewis et al. 2005] J. L. Lewis, G. C. Verchota, and A. L. Vogel, "Wolff snowflakes", *Pacific J. Math.* **218**:1 (2005), 139–166. MR Zbl

[Lewy 1977] H. Lewy, "On the minimum number of domains in which the nodal lines of spherical harmonics divide the sphere", *Comm. Partial Differential Equations* **2**:12 (1977), 1233–1244. MR Zbl

[Logunov and Malinnikova 2015] A. Logunov and E. Malinnikova, "On ratios of harmonic functions", *Adv. Math.* **274** (2015), 241–262. MR Zbl

[Łojasiewicz 1959] S. Łojasiewicz, "Sur le problème de la division", *Studia Math.* **18** (1959), 87–136. MR Zbl

[Mattila 1995] P. Mattila, *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability*, Cambridge Studies in Advanced Mathematics **44**, Cambridge University Press, 1995. MR Zbl

[Mattila and Vuorinen 1990] P. Mattila and M. Vuorinen, "Linear approximation property, Minkowski dimension, and quasiconformal spheres", *J. London Math. Soc.* (2) **42**:2 (1990), 249–266. MR Zbl

[Mörters and Peres 2010] P. Mörters and Y. Peres, *Brownian motion*, Cambridge Series in Statistical and Probabilistic Mathematics **30**, Cambridge University Press, 2010. MR Zbl

[Naber and Valtorta 2014] A. Naber and D. Valtorta, "Volume estimates on the critical sets of solutions to elliptic PDEs", preprint, 2014. arXiv

[Phạm 2012] T. S. Phạm, "An explicit bound for the Łojasiewicz exponent of real polynomials", *Kodai Math. J.* **35**:2 (2012), 311–319. MR Zbl

[Preiss 1987] D. Preiss, "Geometry of measures in $\mathbb{R}^n$: distribution, rectifiability, and densities", *Ann. of Math.* (2) **125**:3 (1987), 537–643. MR Zbl

[Reifenberg 1960] E. R. Reifenberg, "Solution of the Plateau Problem for m-dimensional surfaces of varying topological type", *Acta Math.* **104**:1-2 (1960), 1–92. MR Zbl

[Rockafellar and Wets 1998] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*, Grundlehren der Mathematischen Wissenschaften **317**, Springer, 1998. MR Zbl

[Szulkin 1978] A. Szulkin, "An example concerning the topological character of the zero-set of a harmonic function", *Math. Scand.* **43**:1 (1978), 60–62. MR Zbl

Received 1 Feb 2017. Accepted 24 Apr 2017.

MATTHEW BADGER: matthew.badger@uconn.edu
Department of Mathematics, University of Connecticut, Storrs, CT 06269-1009, United States

MAX ENGELSTEIN: maxe@mit.edu
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, United States

TATIANA TORO: toro@uw.edu
Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195-4350, United States

ANALYSIS AND PDE
Vol. 10, No. 6, 2017
dx.doi.org/10.2140/apde.2017.10.1497

FUGLEDE'S SPECTRAL SET CONJECTURE FOR CONVEX POLYTOPES

RACHEL GREENFELD AND NIR LEV

Let Ω be a convex polytope in $\mathbb{R}^d$. We say that Ω is spectral if the space $L^2(\Omega)$ admits an orthogonal basis consisting of exponential functions. There is a conjecture, which goes back to Fuglede (1974), that Ω is spectral if and only if it can tile the space by translations. It is known that if Ω tiles then it is spectral, but the converse was proved only in dimension $d = 2$, by Iosevich, Katz and Tao.

By a result due to Kolountzakis, if a convex polytope $\Omega \subset \mathbb{R}^d$ is spectral, then it must be centrally symmetric. We prove that also all the facets of Ω are centrally symmetric. These conditions are necessary for Ω to tile by translations.

We also develop an approach which allows us to prove that in dimension $d = 3$, any spectral convex polytope Ω indeed tiles by translations. Thus we obtain that Fuglede's conjecture is true for convex polytopes in $\mathbb{R}^3$.

1. Introduction

1A. Let $\Omega \subset \mathbb{R}^d$ be a bounded, measurable set of positive Lebesgue measure. A countable set $\Lambda \subset \mathbb{R}^d$ is called a *spectrum* for Ω if the system of exponential functions

$$E(\Lambda) = \{e_\lambda\}_{\lambda \in \Lambda}, \quad e_\lambda(x) = e^{2\pi i \langle \lambda, x \rangle}, \tag{1-1}$$

constitutes an orthogonal basis in $L^2(\Omega)$, that is, the system is orthogonal and complete in the space. A set Ω which admits a spectrum Λ is called a *spectral set*.

The classical example of such a situation is when Ω is the unit cube in $\mathbb{R}^d$, and Λ is the integer lattice $\mathbb{Z}^d$. Which other sets Ω are spectral? The study of this problem was initiated by Fuglede [1974]. For example, in that paper it was shown that a triangle and a disk in the plane are not spectral sets.

The set Ω is said to *tile* the space by translations along a countable set $\Lambda \subset \mathbb{R}^d$ if the family of sets $\Omega + \lambda$, $\lambda \in \Lambda$, constitutes a partition of $\mathbb{R}^d$ up to measure zero. In this case we will say that $\Omega + \Lambda$ is a *tiling*. Fuglede [1974] observed the following connection between the concepts of spectrality and tiling:

Let Λ be a lattice. If $\Omega + \Lambda$ is a tiling, then the dual lattice Λ^* is a spectrum for Ω, and also the converse is true.

Here, by a lattice we mean the image of $\mathbb{Z}^d$ under some invertible linear transformation, and the dual lattice is the set of all vectors λ^* such that $\langle \lambda, \lambda^* \rangle \in \mathbb{Z}$, $\lambda \in \Lambda$.

Fuglede conjectured that the spectral sets could be characterized in geometric terms using the concept of tiling in the following way: *the set Ω is spectral if and only if it can tile the space by translations*. This

Research supported by ISF Grant no. 225/13 and ERC Starting Grant no. 713927.
MSC2010: 42B10, 52C22.
Keywords: Fuglede's conjecture, spectral set, tiling, convex polytope.

conjecture inspired extensive research over the years, and a number of interesting results supporting the conjecture were obtained. See, for example, the survey given in [Kolountzakis 2004, Section 3].

On the other hand, it turned out that there also exist counterexamples to Fuglede's conjecture. Tao [2004] constructed in dimensions 5 and higher an example of a set Ω which is spectral, but cannot tile by translations. Subsequently, also examples of nonspectral sets which can tile by translations were found, and eventually the dimension in these examples was reduced up to $d \geqslant 3$; see [Kolountzakis and Matolcsi 2010, Section 4]. In all these examples the set Ω is the union of a finite number of unit cubes centered at points of the integer lattice $\mathbb{Z}^d$.

1B. It is nevertheless believed that Fuglede's conjecture should be true if the set Ω is assumed to be *convex*. There is a well-known characterization due to Venkov [1954], which was rediscovered by McMullen [1980; 1981], of the convex bodies (compact convex sets with nonempty interior) that can tile the space by translations:

Let Ω be a convex body in $\mathbb{R}^d$. Then Ω tiles by translations if and only if the following four conditions are satisfied:

(i) *Ω is a polytope.*

(ii) *Ω is centrally symmetric.*

(iii) *All the facets of Ω are centrally symmetric.*

(iv) *Each "belt" of Ω consists of exactly* 4 *or* 6 *facets.*

By a *belt* of a convex polytope $\Omega \subset \mathbb{R}^d$ with centrally symmetric facets one means the collection of its facets which contain a translate of a given subfacet, that is, a $(d-2)$-dimensional face, of Ω.

It was also proved in [Venkov 1954; McMullen 1980] that if a convex polytope Ω can tile by translations, then it admits a face-to-face tiling by translates along a certain *lattice*. Hence, combined with Fuglede's theorem above this yields the following result:

Let $\Omega \subset \mathbb{R}^d$ be a convex body. If Ω tiles by translations, then Ω is spectral.

The converse to this result, however, is known only in dimension $d = 2$. It is due to Iosevich, Katz and Tao [Iosevich et al. 2003], who showed that a spectral convex body in $\mathbb{R}^2$ must be either a parallelogram or a centrally symmetric hexagon, and hence it tiles by translations.

The situation in dimensions $d \geqslant 3$ is much less understood. It is known that the ball is not a spectral set [Iosevich et al. 1999; Fuglede 2001], nor any convex body with a smooth boundary [Iosevich et al. 2001]. We established in [Greenfeld and Lev 2016] that if Ω is a cylindric convex body whose base has a smooth boundary, then it can neither be spectral.

Kolountzakis [2000] proved the following result:

Let Ω be a convex body in $\mathbb{R}^d$. If Ω is spectral, then it must be centrally symmetric.

1C. In this paper we will focus on the case when Ω is a *convex polytope*. Our first result shows that in this case, not only the central symmetry of Ω, but also the central symmetry of all the facets of Ω, is a necessary condition for spectrality:

Theorem 1.1. *Let Ω be a convex polytope in $\mathbb{R}^d$. If Ω is a spectral set, then all the facets of Ω must be centrally symmetric.*

Our proof of this result is inspired by the paper [Kolountzakis and Papadimitrakis 2002].

Together with the result from [Kolountzakis 2000] we thus obtain that a spectral convex polytope $\Omega \subset \mathbb{R}^d$ must satisfy the conditions (ii) and (iii) in the Venkov–McMullen theorem above. So this supports the conjecture that any such Ω can tile by translations.

Our next theorem, which is the main result of this paper, confirms that this is indeed the case in dimension $d = 3$:

Theorem 1.2. *Let Ω be a convex polytope in $\mathbb{R}^3$. If Ω is a spectral set, then it can tile by translations.*

Combined with the above-mentioned results, we thus obtain that Fuglede's conjecture is true for convex polytopes $\Omega \subset \mathbb{R}^3$.

1D. In two dimensions, the convex polygons which can tile by translations are precisely the parallelograms and the centrally symmetric hexagons. The three-dimensional convex polytopes which can tile by translations were classified by Fedorov [1885] into five distinct combinatorial types: the parallelepiped, the hexagonal prism, the rhombic dodecahedron, the elongated dodecahedron and the truncated octahedron (see, for example, [Gruber 2007, Figure 32.4] for a graphical illustration of these types). Thus, for a convex polytope $\Omega \subset \mathbb{R}^3$ to tile by translations, it is necessary and sufficient that it belongs to one of these five types, and that Ω, as well as all its facets, are centrally symmetric. A detailed exposition of this result can be found in [Alexandrov 2005, Section 8.1].

Theorem 1.2 therefore yields that these conditions are also necessary and sufficient for a convex polytope $\Omega \subset \mathbb{R}^3$ to be spectral.

(The requirement that Ω is centrally symmetric is in fact redundant in this characterization: it is known [Alexandrov 1933] that if all the facets of a convex polytope $\Omega \subset \mathbb{R}^d$, $d \geqslant 3$, are centrally symmetric, then Ω itself must also be centrally symmetric.)

1E. As mentioned above, the Venkov–McMullen and Fuglede results imply not only that a convex polytope $\Omega \subset \mathbb{R}^d$ which can tile by translations is necessarily spectral, but also that Ω admits a *lattice* spectrum. Our approach allows us to establish that for certain convex polytopes, this spectrum is the *unique* one, up to translation.

First we have the following result in two dimensions:

Theorem 1.3. *Let Ω be a centrally symmetric hexagon in $\mathbb{R}^2$. Then Ω has a unique spectrum up to translation.*

This result is essentially contained in [Iosevich et al. 2003], although it was not stated explicitly in that paper.

The three-dimensional version of the result is the following:

Theorem 1.4. *Let Ω be a convex polytope in $\mathbb{R}^3$ which is spectral (and hence it can tile by translations), but which is neither a parallelepiped nor a hexagonal prism. Then Ω has a unique spectrum up to translation.*

Remark that it is necessary in these results to exclude the parallelograms in $\mathbb{R}^2$, and the parallelepipeds and the centrally symmetric hexagonal prisms in $\mathbb{R}^3$. Indeed, these convex polytopes admit infinitely many non translation-equivalent spectra (see [Jorgensen and Pedersen 1999, Section 2]).

1F. The paper is organized as follows.

In Section 2 we present some preliminary background.

In Section 3 we give a proof of the fact that a spectral convex polytope $\Omega \subset \mathbb{R}^d$ must be centrally symmetric. The proof given is based on the argument from [Kolountzakis and Papadimitrakis 2002].

In Section 4 we prove that also all the facets of such an Ω are centrally symmetric (Theorem 1.1).

In Sections 5–7 we develop an approach to show that a spectral convex polytope $\Omega \subset \mathbb{R}^d$ can tile by translations. In Section 8 we give a proof, based on this approach, of the result that a spectral convex polygon $\Omega \subset \mathbb{R}^2$ can tile by translations.

The proof of the three-dimensional Theorem 1.2 is given through Sections 9–15.

In Section 16 the results concerning the uniqueness of the spectrum up to translation are deduced (Theorems 1.3 and 1.4).

In Section 17 we give additional remarks and discuss some open problems.

2. Preliminaries

2A. ***Notation.*** We fix some notation that will be used throughout the paper.

We shall denote by $\vec{e}_1, \dots, \vec{e}_d$ the standard basis vectors in $\mathbb{R}^d$.

As usual, $\langle \cdot, \cdot \rangle$ and $|\cdot|$ are the Euclidean scalar product and norm in $\mathbb{R}^d$.

For a set $A \subset \mathbb{R}^d$ and a vector $x \in \mathbb{R}^d$, we use $\langle A, x \rangle$ to denote the set $\{\langle a, x \rangle : a \in A\}$.

We denote by $|\Omega|$ the Lebesgue measure of a measurable set $\Omega \subset \mathbb{R}^d$.

The Fourier transform in $\mathbb{R}^d$ will be normalized as

$$\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)\, e^{-2\pi i \langle \xi, x \rangle}\, dx.$$

2B. ***Properties of spectra.*** We recall some basic properties of spectra that will be used in the paper.

Let $\Omega \subset \mathbb{R}^d$ be a bounded, measurable set of positive measure. A countable set $\Lambda \subset \mathbb{R}^d$ is a spectrum for Ω if the system of exponential functions $E(\Lambda)$ defined by (1-1) is an orthogonal basis in the space $L^2(\Omega)$. Since we have

$$\langle e_\lambda, e_{\lambda'} \rangle_{L^2(\Omega)} = \int_\Omega e^{-2\pi i \langle \lambda' - \lambda, x \rangle}\, dx = \hat{\mathbb{1}}_\Omega(\lambda' - \lambda),$$

it follows that the orthogonality of $E(\Lambda)$ in $L^2(\Omega)$ is equivalent to the condition

$$\Lambda - \Lambda \subset \{\hat{\mathbb{1}}_\Omega = 0\} \cup \{0\}. \tag{2-1}$$

A set $\Lambda \subset \mathbb{R}^d$ is said to be *uniformly discrete* if there is $\delta > 0$ such that $|\lambda' - \lambda| \geqslant \delta$ for any two distinct points λ, λ' in Λ. The maximal constant δ with this property is called the *separation constant* of Λ, and will be denoted by $\delta(\Lambda)$.

The condition (2-1) implies that if Λ is a spectrum for Ω then it is a uniformly discrete set, with separation constant $\delta(\Lambda)$ not smaller than

$$\chi(\Omega) := \min\big\{|\xi| : \xi \in \mathbb{R}^d,\ \hat{\mathbb{1}}_\Omega(\xi) = 0\big\} > 0. \tag{2-2}$$

It is easy to verify that the property of Λ being a spectrum for Ω is invariant under translations of both Ω and Λ. It is also easy to check that if Λ is a spectrum for Ω, and if A is an invertible $d \times d$ matrix, then the set $(A^{-1})^\top(\Lambda)$ is a spectrum for $A(\Omega)$.

2C. ***Limits of spectra.*** Let Λ_n be a sequence of uniformly discrete sets in $\mathbb{R}^d$ such that $\delta(\Lambda_n) \geqslant \delta > 0$. The sequence Λ_n is said to *converge weakly* to a set Λ if for every $\varepsilon > 0$ and every R there is N such that

$$\Lambda_n \cap B_R \subset \Lambda + B_\varepsilon \quad \text{and} \quad \Lambda \cap B_R \subset \Lambda_n + B_\varepsilon$$

for all $n \geqslant N$, where by B_r we denote the open ball of radius r centered at the origin. In this case, the weak limit Λ is also uniformly discrete, and moreover, $\delta(\Lambda) \geqslant \delta$.

By a standard diagonalization argument one can show that given any sequence Λ_n satisfying $\delta(\Lambda_n) \geqslant \delta > 0$, there is a subsequence Λ_{n_j} which converges weakly to some (possibly empty) set Λ.

It is known that if for each n the set Λ_n is a spectrum for Ω, and if Λ_n converges weakly to a limit Λ, then also Λ is a spectrum for Ω. See, for example, [Greenfeld and Lev 2016, Section 3], where a simple proof of this fact can be found.

The latter fact easily implies that any spectrum Λ of Ω must be a *relatively dense* set in $\mathbb{R}^d$; namely, there is $R > 0$ such that every ball of radius R intersects Λ. Moreover, the constant $R = R(\Omega)$ does not depend on the spectrum Λ. Indeed, if this was not true then there would exist a sequence Λ_n of spectra for Ω which converges weakly to the empty set, which contradicts the fact that the weak limit must also be a spectrum for Ω.

2D. ***Fourier expansion with respect to a spectrum.*** If Λ is a spectrum for Ω, then each $f \in L^2(\Omega)$ admits a Fourier expansion with respect to the orthogonal basis $E(\Lambda)$. If we extend such a function f to the whole $\mathbb{R}^d$ by defining it to be zero outside of Ω, then we have $\langle f, e_\lambda\rangle_{L^2(\Omega)} = \hat{f}(\lambda)$; hence the Fourier expansion of f has the form

$$f = \frac{1}{|\Omega|} \sum_{\lambda \in \Lambda} \hat{f}(\lambda) e_\lambda, \tag{2-3}$$

and the series converges in $L^2(\Omega)$. Furthermore, Parseval's equality holds; namely,

$$\|f\|^2_{L^2(\Omega)} = \frac{1}{|\Omega|} \sum_{\lambda \in \Lambda} |\hat{f}(\lambda)|^2.$$

The following fact will be useful for us:

Lemma 2.1. *For each function $f \in L^2(\Omega)$ (extended to be zero outside of Ω) the series* (2-3) *converges unconditionally in L^2 on any bounded set to some measurable function $\tilde{f}$ defined a.e. on the whole $\mathbb{R}^d$, and f coincides with $\tilde{f}$ a.e. on Ω.*

This is a simple consequence of the following:

Lemma 2.2. *Let $\Lambda \subset \mathbb{R}^d$ be a uniformly discrete set and $\{c(\lambda)\}$ be a sequence in $\ell^2(\Lambda)$. Then the series*

$$\sum_{\lambda \in \Lambda} c(\lambda) e_\lambda \tag{2-4}$$

converges unconditionally in $L^2(S)$ for every bounded set $S \subset \mathbb{R}^d$.

The latter fact is well known; see, for instance, [Young 2001, Section 4.3, Theorem 4], where it is proved in dimension one. For the reader's convenience we provide a self-contained proof in arbitrary dimension d.

Proof of Lemma 2.2. First we show that if S is a bounded set then there is a constant $C = C(\Lambda, S)$ such that for every sequence $\{c(\lambda)\}$ with only finitely many nonzero terms we have

$$\Big\| \sum_{\lambda \in \Lambda} c(\lambda) e_\lambda \Big\|^2_{L^2(S)} \leqslant C \sum_{\lambda \in \Lambda} |c(\lambda)|^2. \tag{2-5}$$

Indeed, let $\delta > 0$ denote the separation constant of Λ, and choose a smooth function φ supported on a ball of radius $\delta/2$ around the origin such that $\int |\varphi(t)|^2 \, dt = 1$ and

$$\eta := \inf_{x \in S} |\hat{\varphi}(x)| > 0.$$

Then the left-hand side of (2-5) is not greater than $1/\eta^2$ times

$$\int_{\mathbb{R}^d} \Big| \hat{\varphi}(x) \sum_{\lambda \in \Lambda} c(\lambda) e_\lambda(x) \Big|^2 dx = \int_{\mathbb{R}^d} \Big| \sum_{\lambda \in \Lambda} c(\lambda) \varphi(t + \lambda) \Big|^2 dt = \sum_{\lambda \in \Lambda} |c(\lambda)|^2;$$

hence (2-5) holds with $C = 1/\eta^2$.

Now it follows from (2-5) that given an arbitrary sequence $\{c(\lambda)\}$ in $\ell^2(\Lambda)$, the partial sums of the series (2-4) constitute a Cauchy sequence in $L^2(S)$ for every arrangement of the terms of the series, and the limit in $L^2(S)$ of these partial sums is the same for every such arrangement. This confirms the assertion of the lemma. □

2E. *Convex polytopes.* By a *convex polytope* Ω in $\mathbb{R}^d$ we mean a compact set which is the convex hull of a finite number of points. By a *facet* of Ω we refer to a $(d-1)$-dimensional face of Ω, while a *subfacet* is a $(d-2)$-dimensional face.

If G is a k-dimensional face of Ω $(0 \leqslant k \leqslant d)$, then $|G|$ denotes the k-dimensional volume of G. For a facet F of Ω we denote by σ_F the surface measure on F.

The interior of Ω will be denoted by $\mathrm{int}(\Omega)$.

We say that Ω is *centrally symmetric* if there is a point $x \in \mathbb{R}^d$ (the center) such that $\Omega - x = -\Omega + x$. The following theorem, due to Minkowski, gives a criterion for the central symmetry of a convex polytope Ω in terms of the areas of its facets:

Theorem 2.3 (Minkowski). *A convex polytope Ω is centrally symmetric if and only if for each facet F of Ω there is a parallel facet F' such that $|F| = |F'|$.*

This is a consequence of the classical Minkowski's uniqueness theorem; see, for example, [Gruber 2007, Section 18.2].

We shall need some well-known facts about Fourier transforms related to convex polytopes in $\mathbb{R}^d$ (actually, in some of these results the convexity is not necessary). Since the proofs are not difficult, they are included for completeness.

Lemma 2.4. *Let Ω be a convex polytope in $\mathbb{R}^d$ $(d \geqslant 1)$. For each facet F of Ω, let n_F denote the outward unit normal to Ω on F. Then*

$$-2\pi i\xi\, \hat{\mathbb{1}}_\Omega(\xi) = \sum n_F\, \hat{\sigma}_F(\xi), \quad \xi \in \mathbb{R}^d, \tag{2-6}$$

where the sum is over all the facets F of Ω.

Proof. Fix two vectors ξ and u in $\mathbb{R}^d$, and let

$$\Phi(x) := ue^{-2\pi i\langle \xi, x\rangle}, \quad x \in \mathbb{R}^d.$$

Then we have

$$\operatorname{div} \Phi(x) = -2\pi i\langle \xi, u\rangle e^{-2\pi i\langle \xi, x\rangle}.$$

By the divergence theorem,

$$\int_\Omega \operatorname{div} \Phi(x)\, dx = \int_{\partial\Omega} \langle \Phi(x), n(x)\rangle\, d\sigma(x),$$

where σ denotes the surface measure on the boundary $\partial\Omega$, and $n(x) := n_F$ if x belongs to the relative interior of a facet F of Ω. This means that

$$-2\pi i\langle \xi, u\rangle \hat{\mathbb{1}}_\Omega(\xi) = \sum \langle n_F, u\rangle \hat{\sigma}_F(\xi),$$

where the sum is over all the facets F of Ω. But since ξ and u were arbitrary vectors in $\mathbb{R}^d$, this implies (2-6). □

Corollary 2.5. *If Ω is a convex polytope in $\mathbb{R}^d$ $(d \geqslant 1)$, then*

$$|\hat{\mathbb{1}}_\Omega(\xi)| \leqslant \frac{|\partial\Omega|}{2\pi} \cdot |\xi|^{-1},$$

where $|\partial\Omega|$ denotes the total surface area of Ω.

This follows from Lemma 2.4 using that the right-hand side of (2-6) is bounded in norm by $|\partial\Omega|$.

Lemma 2.6. *Let Ω be a convex polytope in $\mathbb{R}^d$ $(d \geqslant 2)$, andF be a facet of Ω. Let $\theta(\xi, F)$ denote the angle between a nonzero vector $\xi \in \mathbb{R}^d$ and the outward normal vector to Ω on F. Then*

$$|\hat{\sigma}_F(\xi)| \leqslant \frac{|\partial F|}{2\pi} \cdot \frac{|\xi|^{-1}}{|\sin\theta(\xi, F)|},$$

where $|\partial F|$ is the $(d-2)$-dimensional volume of the relative boundary of F.

Proof. By applying a rotation and a translation, we may assume F is contained in the hyperplane $\{x_1 = 0\}$, and that the outward unit normal to Ω on F is $\vec{e}_1$. Hence

$$\hat{\sigma}_F(\xi) = \varphi_F(\xi_2, \xi_3, \ldots, \xi_d),$$

where φ_F denotes the Fourier transform of the indicator function of the polytope in $\mathbb{R}^{d-1}$ obtained by projecting the facet F on the $(x_2, x_3, \ldots, x_d)$-coordinates. Using Corollary 2.5, this implies

$$|\hat{\sigma}_F(\xi)| \leqslant \frac{|\partial F|}{2\pi} \left(\sum_{j=2}^{d} \xi_j^2 \right)^{-1/2}.$$

However, since we have

$$\xi_1 = \langle \xi, \vec{e}_1 \rangle = |\xi| \cos\theta(\xi, F),$$

it follows that

$$\sum_{j=2}^{d} \xi_j^2 = |\xi|^2 - \xi_1^2 = |\xi|^2 (1 - \cos^2\theta(\xi, F)) = |\xi|^2 \sin^2\theta(\xi, F),$$

so this proves the claim. □

The previous lemmas imply the following result, which will be used in the next sections:

Lemma 2.7. *Let Ω be a convex polytope in $\mathbb{R}^d$ $(d \geqslant 2)$. Assume that A and B are two parallel facets of Ω, and that the outward unit normals to Ω on A and B are respectively the vectors $\vec{e}_1$ and $-\vec{e}_1$ (we also allow A to be a facet which does not have a parallel facet, in which case we understand B to be the empty set). Then there is $\alpha = \alpha(\Omega) > 0$ such that*

$$-2\pi i \xi_1 \hat{\mathbb{1}}_\Omega(\xi) = \hat{\sigma}_A(\xi) - \hat{\sigma}_B(\xi) + O(|\xi_1|^{-1}), \quad |\xi_1| \to \infty, \tag{2-7}$$

in the cone

$$K(\alpha) := \left\{ \xi \in \mathbb{R}^d : |\xi_j| \leqslant \alpha |\xi_1| \ (2 \leqslant j \leqslant d) \right\}. \tag{2-8}$$

Proof. By Lemma 2.4 we have

$$-2\pi i \xi_1 \, \hat{\mathbb{1}}_\Omega(\xi) = \hat{\sigma}_A(\xi) - \hat{\sigma}_B(\xi) + \sum \langle n_F, \vec{e}_1 \rangle \, \hat{\sigma}_F(\xi), \tag{2-9}$$

where the sum is over all the facets F of Ω other than A and B. If α is sufficiently small, then the angle between any vector in $K(\alpha)$ and the outward normal to Ω on any facet F other than A and B is bounded away from 0 and π. Hence by Lemma 2.6, the sum on the right-hand side of (2-9) is $O(|\xi|^{-1})$ as $|\xi| \to \infty$ in the cone $K(\alpha)$. But since the ratio $|\xi_1|/|\xi|$ is bounded from below in $K(\alpha)$, this implies (2-7). □

3. Spectral convex polytopes are symmetric

3A. In this section we give a proof of the following result:

Theorem 3.1 [Kolountzakis 2000]. *Let Ω be a convex polytope in $\mathbb{R}^d$ $(d \geqslant 2)$. If Ω is spectral then Ω is centrally symmetric.*

In fact, it was proved in [Kolountzakis 2000] that any convex body (not assumed to be a polytope) which is spectral must be centrally symmetric. This supports the conjecture that a spectral convex body Ω can tile by translations, as the central symmetry is a necessary condition for Ω to tile, by the Venkov–McMullen theorem.

There is another approach to prove Theorem 3.1, which was introduced in [Kolountzakis and Papadimitrakis 2002]. This approach is specific for polytopes, but on the other hand it does not require Ω to be convex. The main result in that paper gives a certain condition on a polytope $\Omega \subset \mathbb{R}^d$ that is necessary for its spectrality. If the polytope Ω is convex, then this condition coincides with the requirement that Ω is centrally symmetric.

For the completeness of our exposition, below we give a proof of Theorem 3.1 based on the argument in [Kolountzakis and Papadimitrakis 2002]. See also [Kolountzakis 2004, pp. 184–185]. The proof may also serve as a preparation for the next section, where the argument will be further developed.

3B. ***Proof of Theorem 3.1.*** By Minkowski's theorem, Theorem 2.3, it would be enough to show that for each facet A of Ω there is a parallel facet B such that $|A| = |B|$. If this is not true, then there is a facet A of Ω whose parallel facet B satisfies $|A| > |B|$, where we understand B to be the empty set if A is a facet of Ω with no parallel facet.

By applying an affine transformation, we may assume that A is contained in the hyperplane $\{x_1 = 0\}$, that B is contained in the hyperplane $\{x_1 = -1\}$, and that the outward unit normals to Ω on A and B are respectively the vectors $\vec{e}_1$ and $-\vec{e}_1$. It follows that

$$\hat{\sigma}_A(\xi) = \varphi_A(\xi_2, \xi_3, \dots, \xi_d), \tag{3-1}$$

$$\hat{\sigma}_B(\xi) = e^{2\pi i \xi_1} \varphi_B(\xi_2, \xi_3, \dots, \xi_d), \tag{3-2}$$

where φ_A, φ_B are respectively the Fourier transforms of the indicator functions of the polytopes in $\mathbb{R}^{d-1}$ obtained by projecting the facets A, B on the $(x_2, x_3, \dots, x_d)$-coordinates. In particular, φ_A and φ_B are continuous functions, and

$$\varphi_A(0) = |A|, \quad \varphi_B(0) = |B|. \tag{3-3}$$

For any $r > 0$ we denote by $S(r)$ the cylinder of radius r along the x_1-axis; namely

$$S(r) := \{t\vec{e}_1 + w : t \in \mathbb{R},\ w \in \mathbb{R}^d,\ |w| < r\}.$$

Notice that

$$S(r) - S(r) = S(2r). \tag{3-4}$$

By assumption, we have $|A| > |B|$. Choose a number η such that

$$0 < \eta < |A| - |B|.$$

It follows from (3-1), (3-2) and (3-3) that there is $\varepsilon > 0$ such that

$$|\hat{\sigma}_A(\xi) - \hat{\sigma}_B(\xi)| \geqslant |\hat{\sigma}_A(\xi)| - |\hat{\sigma}_B(\xi)| > \eta, \quad \xi \in S(2\varepsilon).$$

By Lemma 2.7 we have

$$-2\pi i \xi_1 \hat{\mathbb{1}}_\Omega(\xi) = \hat{\sigma}_A(\xi) - \hat{\sigma}_B(\xi) + O(|\xi_1|^{-1}), \quad |\xi_1| \to \infty,$$

in the cylinder $S(2\varepsilon)$. It follows that there is $R > 0$ such that

$$\hat{\mathbb{1}}_\Omega(\xi) \neq 0, \quad \xi \in S(2\varepsilon) \setminus B_R, \tag{3-5}$$

where B_R denotes the ball of radius R centered at the origin.

Now let Λ be a spectrum for Ω. We claim that for any $\tau \in \mathbb{R}^d$, if λ, λ' are two points in $\Lambda \cap (S(\varepsilon) + \tau)$, then $|\lambda' - \lambda| \leqslant R$. Indeed, if not then using (3-4) we get

$$\lambda' - \lambda \in S(2\varepsilon) \setminus B_R,$$

but due to (3-5) this implies $\hat{\mathbb{1}}_\Omega(\lambda' - \lambda) \neq 0$, a contradiction.

Since Λ is a uniformly discrete set, it follows that $\Lambda \cap (S(\varepsilon) + \tau)$ is a finite set for every $\tau \in \mathbb{R}^d$. Since Λ is a relatively dense set, there is $M > 0$ such that every ball of radius M intersects Λ. The cylinder $S(M)$ may be covered by a finite number of cylinders $S(\varepsilon) + \tau_j$ $(1 \leqslant j \leqslant N)$; hence $\Lambda \cap S(M)$ is also a finite set. But this implies that $S(M)$ must contain a ball of radius M free from points of Λ, a contradiction. This completes the proof of Theorem 3.1. □

4. Spectral convex polytopes have symmetric facets

4A. The result in Section 3 shows that the central symmetry is a necessary condition for a convex polytope $\Omega \subset \mathbb{R}^d$ to be spectral. In the present section we prove that also the central symmetry of all the facets of Ω is necessary for spectrality:

Theorem 4.1. *Let Ω be a convex, centrally symmetric polytope in $\mathbb{R}^d$ $(d \geqslant 3)$. If Ω is spectral then all the facets of Ω are also centrally symmetric.*

Recall that by the Venkov–McMullen theorem, the central symmetry of the facets is also a necessary condition for Ω to tile by translations. Hence this result further supports the conjecture that any spectral convex polytope Ω can tile by translations.

Notice that the conclusion cannot be further improved by showing that also all the k-dimensional faces of Ω, for some $2 \leqslant k \leqslant d-2$, are centrally symmetric. Indeed, this would imply [McMullen 1970] that all the faces of Ω of every dimension are centrally symmetric. However, the 24-cell in $\mathbb{R}^4$ is a well-known example of a convex polytope which tiles by translations, and hence is spectral, but which does not satisfy this property.

The rest of this section is devoted to the proof of Theorem 4.1. The proof is based on a development of the argument in [Kolountzakis and Papadimitrakis 2002].

4B. Let F be one of the facets of Ω. As before, to prove that F is centrally symmetric it would be enough, by Minkowski's theorem, Theorem 2.3, to show that for each subfacet A of F there is a parallel subfacet B of F such that $|A| = |B|$. So, again, suppose to the contrary that A, B are two parallel subfacets of F such that $|A| > |B|$, with the agreement that B is empty if A has no parallel subfacet of F.

By applying an affine transformation, we may assume

$$\Omega = -\Omega; \tag{4-1}$$

namely, Ω is symmetric about the origin,

$$F \subset \left\{x_1 = \tfrac{1}{2}\right\} \tag{4-2}$$

and the outward unit normal to Ω on F is $\vec{e}_1$,

$$A \subset \left\{x_1 = \tfrac{1}{2},\ x_2 = 0\right\}, \tag{4-3}$$

$$B \subset \left\{x_1 = \tfrac{1}{2},\ x_2 = -1\right\}, \tag{4-4}$$

and the outward unit normals to F on A and B are respectively $\vec{e}_2$ and $-\vec{e}_2$.

4C. Let φ_F (respectively, φ_A and φ_B) denote the Fourier transform of the indicator function of the polytope in $\mathbb{R}^{d-1}$ (respectively, $\mathbb{R}^{d-2}$) obtained by projecting the facet F on the $(x_2, x_3, \dots, x_d)$-coordinates (respectively, the subfacets A and B on the $(x_3, \dots, x_d)$-coordinates). Define

$$\psi(\xi) := \mathrm{Re}\big[e^{-\pi i \xi_1}\big(\varphi_A(\xi_3, \dots, \xi_d) - e^{2\pi i \xi_2}\varphi_B(\xi_3, \dots, \xi_d)\big)\big], \quad \xi \in \mathbb{R}^d. \tag{4-5}$$

Also, for any three positive real numbers L, δ and α, we let

$$K(L, \delta, \alpha) := \big\{\xi \in \mathbb{R}^d : L \leqslant |\xi_2| \leqslant \delta|\xi_1|,\ |\xi_j| \leqslant \alpha|\xi_2|\ (3 \leqslant j \leqslant d)\big\}.$$

Lemma 4.2. *There is $\alpha > 0$ such that given any $\eta > 0$ one can find $\delta > 0$ and L such that*

$$\big|2\pi^2\xi_1\xi_2\,\hat{\mathbb{1}}_\Omega(\xi) + \psi(\xi)\big| < \eta, \quad \xi \in K(L, \delta, \alpha). \tag{4-6}$$

Proof. Due to (4-1), the facet of Ω parallel to F is $-F$. If $0 < \delta \leqslant \alpha < 1$, then the set $K(L, \delta, \alpha)$ is contained in the cone

$$\big\{|\xi_j| \leqslant \alpha|\xi_1|,\ 2 \leqslant j \leqslant d\big\}. \tag{4-7}$$

Hence by Lemma 2.7, if α is sufficiently small then

$$-2\pi i\xi_1\,\hat{\mathbb{1}}_\Omega(\xi) = \hat{\sigma}_F(\xi) - \hat{\sigma}_{-F}(\xi) + O(|\xi_1|^{-1}), \quad |\xi_1| \to \infty, \tag{4-8}$$

in the cone (4-7). Observe that by (4-2) we have

$$\hat{\sigma}_F(\xi) - \hat{\sigma}_{-F}(\xi) = 2i\,\mathrm{Im}[\hat{\sigma}_F(\xi)] = 2i\,\mathrm{Im}\big[e^{-\pi i \xi_1}\varphi_F(\xi_2, \xi_3, \dots, \xi_d)\big]. \tag{4-9}$$

Now if $\xi \in K(L, \delta, \alpha)$ then the vector $(\xi_2, \xi_3, \dots, \xi_d)$ belongs to the cone

$$\big\{|\xi_j| \leqslant \alpha|\xi_2|,\ 3 \leqslant j \leqslant d\big\} \subset \mathbb{R}^{d-1}, \tag{4-10}$$

so again by Lemma 2.7 and by (4-3), (4-4) it follows that if α is sufficiently small then

$$-2\pi i\xi_2\varphi_F(\xi_2, \xi_3, \dots, \xi_d) = \varphi_A(\xi_3, \dots, \xi_d) - e^{2\pi i \xi_2}\varphi_B(\xi_3, \dots, \xi_d) + O(|\xi_2|^{-1}), \quad |\xi_2| \to \infty, \tag{4-11}$$

in the cone (4-10). Combining (4-8), (4-9) and (4-11) shows that there is $\alpha > 0$ and a positive constant C such that for any $0 < \delta \leqslant \alpha$ and any $L > 0$ we have

$$\left|2\pi^2 \xi_1 \xi_2 \hat{\mathbb{1}}_\Omega(\xi) + \psi(\xi)\right| \leqslant C\big(|\xi_2/\xi_1| + |1/\xi_2|\big), \quad \xi \in K(L, \delta, \alpha).$$

But for $\xi \in K(L, \delta, \alpha)$ we have $|\xi_2/\xi_1| \leqslant \delta$ and $|1/\xi_2| \leqslant L^{-1}$. Hence given any $\eta > 0$, by choosing δ sufficiently small and L sufficiently large, we obtain (4-6). □

4D. Recall that, by assumption, we have $|A| > |B|$. Choose a number η such that

$$0 < 2\eta < |A| - |B|.$$

Use Lemma 4.2 to find L, δ and α such that (4-6) holds. Define the vector

$$v_\delta := 2\vec{e}_1 + \delta\vec{e}_2 = (2, \delta, 0, 0, \dots, 0). \tag{4-12}$$

For any $r > 0$ we denote by $E(r, \delta)$ the union of balls of radius r centered at the integral multiples of the vector v_δ; that is,

$$E(r, \delta) := \{k v_\delta + w : k \in \mathbb{Z},\ w \in \mathbb{R}^d,\ |w| < r\}. \tag{4-13}$$

Notice that

$$E(r, \delta) - E(r, \delta) = E(2r, \delta). \tag{4-14}$$

Since φ_A, φ_B are continuous functions satisfying

$$\varphi_A(0) = |A|, \quad \varphi_B(0) = |B|,$$

it follows from (4-5), (4-12) and (4-13) that there is $\varepsilon > 0$ such that

$$\left|\psi(\xi) - \mathrm{Re}\big[|A| - e^{2\pi i \xi_2}|B|\big]\right| < \eta, \quad \xi \in E(2\varepsilon, \delta).$$

In particular, this implies

$$|\psi(\xi)| \geqslant |A| - |B| - \eta > \eta, \quad \xi \in E(2\varepsilon, \delta). \tag{4-15}$$

4E. Lemma 4.3. *There is $R > 0$ such that*

$$E(2\varepsilon, \delta) \setminus B_R \subset K(L, \delta, \alpha), \tag{4-16}$$

where B_R denotes the ball of radius R centered at the origin.

This can be verified easily, so we skip the proof.

4F. Now suppose that Λ is a spectrum for Ω. Use Lemma 4.3 to choose R such that (4-16) holds. We claim that for any $\tau \in \mathbb{R}^d$, if λ, λ' are two points in $\Lambda \cap (E(\varepsilon, \delta) + \tau)$, then $|\lambda' - \lambda| \leqslant R$. Indeed, if not then using (4-14) we get

$$\lambda' - \lambda \in E(2\varepsilon, \delta) \setminus B_R \subset K(L, \delta, \alpha).$$

It thus follows from (4-6) and (4-15) that $\hat{\mathbb{1}}_\Omega(\lambda' - \lambda) \neq 0$, a contradiction.

Since Λ is a uniformly discrete set, it follows that $\Lambda \cap (E(\varepsilon, \delta) + \tau)$ is a finite set for every $\tau \in \mathbb{R}^d$. Since Λ is a relatively dense set, there is $M > 0$ such that every ball of radius M intersects Λ. Let $S(M, \delta)$ denote the cylinder of radius M along the vector v_δ,

$$S(M, \delta) := \{t v_\delta + w : t \in \mathbb{R},\ w \in \mathbb{R}^d,\ |w| < M\}.$$

Then $S(M, \delta)$ may be covered by a finite number of sets $E(\varepsilon, \delta) + \tau_j$ $(1 \leqslant j \leqslant N)$; hence $\Lambda \cap S(M, \delta)$ is also a finite set. It follows that $S(M, \delta)$ contains a ball of radius M free from points of Λ, a contradiction. This completes the proof of Theorem 4.1. □

5. Covering and packing

It was shown in Sections 3 and 4 that if a convex polytope $\Omega \subset \mathbb{R}^d$ is spectral, then it must be centrally symmetric and have centrally symmetric facets. In order to prove that Ω tiles by translations, a conceivable strategy may therefore be to try and show that every belt of Ω must consist of either 4 or 6 facets. Indeed, this would imply that Ω tiles, by the Venkov–McMullen theorem.

Our approach, however, will not be based on such a strategy. Instead, we will use another condition, given in terms of the spectrum Λ, which implies that Ω tiles by translations. In this section, we prove the sufficiency of this condition (Corollary 5.3).

5A. Let $\Omega \subset \mathbb{R}^d$ be a convex polytope, which is centrally symmetric and has centrally symmetric facets. If F is any facet of Ω, then by the central symmetry, the opposite facet F' is a translate of F. We shall denote by τ_F the translation vector in $\mathbb{R}^d$ which carries F' onto F.

Following [Venkov 1954; McMullen 1980], we consider the set

$$T = T(\Omega) = \left\{ \sum_F k_F\, \tau_F \ :\ k_F \in \mathbb{Z} \right\}; \tag{5-1}$$

that is, T is the set of all linear combinations with integer coefficients of the vectors τ_F, where F goes through all the facets of Ω. The set T is a countable subgroup of $\mathbb{R}^d$.

Theorem 5.1 [Venkov 1954; McMullen 1980]. *$\Omega + T$ is a covering; that is, each point in $\mathbb{R}^d$ belongs to at least one of the sets $\Omega + \tau$, $\tau \in T$.*

This is a part of the Venkov–McMullen theorem, which characterizes the convex bodies that tile by translations by the four conditions (i)–(iv) mentioned in Section 1B. In the sufficiency part of the theorem it is shown that these four conditions imply that $\Omega + T$ is a tiling. However the last condition, namely the requirement (iv) that each belt consists of exactly 4 or 6 facets, is not used in that part of the proof where it is shown that $\Omega + T$ is a covering; see [McMullen 1980, pp. 115–116], where the latter fact is also mentioned explicitly. Hence the proof yields that the first three conditions (i)–(iii) are enough to conclude that $\Omega + T$ is a covering, as stated in Theorem 5.1.

Observe that Theorem 5.1 implies that T is a relatively dense set in $\mathbb{R}^d$.

It also follows from this theorem that, in order to prove that Ω tiles by translations, it would be enough to show that $\Omega + T$ is a *packing*, which means that the sets $\Omega + \tau$, $\tau \in T$, are disjoint up to measure zero.

Indeed, in such a case $\Omega + T$ is simultaneously a covering and a packing; hence Ω tiles by translations along the set T.

Notice that if $\Omega + T$ is a packing (and hence a tiling), then T must be a uniformly discrete set in $\mathbb{R}^d$. So in this case T is a subgroup of $\mathbb{R}^d$ which is both uniformly discrete and relatively dense, and it follows that T must be a *lattice*. As mentioned in [McMullen 1980], the tiling by translations of Ω along the lattice T constitutes a *face-to-face* tiling.

5B. The next lemma gives a sufficient condition for $\Omega + T$ to be a packing:

Lemma 5.2. *Suppose that $\Lambda \subset \mathbb{R}^d$ is a set satisfying the condition*

$$\langle \Lambda - \Lambda, \tau_F \rangle \subset \mathbb{Z} \tag{5-2}$$

for every facet F of Ω. If the system of exponentials $E(\Lambda)$ is complete in $L^2(\Omega)$, then $\Omega + T$ is a packing.

Proof. By translating Λ we may assume that it contains the origin; hence $\langle \Lambda, \tau_F \rangle \subset \mathbb{Z}$ for every facet F. It follows that the exponential functions e_λ $(\lambda \in \Lambda)$ are periodic with respect to T; namely

$$e_\lambda(x + \tau) = e_\lambda(x)$$

for every $\tau \in T$. If $\Omega + T$ is not a packing then there exist distinct vectors $\tau', \tau'' \in T$ such that the set $(\Omega + \tau') \cap (\Omega + \tau'')$ has positive measure. Thus the set E defined by

$$E := \Omega \cap (\Omega - \tau), \quad \tau := \tau'' - \tau',$$

is a set of positive measure, and E, $E + \tau$ are both contained in Ω. Hence the function $f := \mathbb{1}_E - \mathbb{1}_{E+\tau}$ is supported by Ω, and since $\tau \neq 0$, the function f does not vanish identically a.e. On the other hand, for every $\lambda \in \Lambda$ we have

$$\langle e_\lambda, f \rangle_{L^2(\Omega)} = \int_E e_\lambda(x)\,dx - \int_{E+\tau} e_\lambda(x)\,dx = 0,$$

due to the periodicity of e_λ. Hence f is orthogonal in $L^2(\Omega)$ to all the exponentials $\{e_\lambda\}$, $\lambda \in \Lambda$, which contradicts the completeness of the system $E(\Lambda)$ in the space $L^2(\Omega)$. $\square$

5C. Combining Theorem 5.1 and Lemma 5.2 we obtain the following:

Corollary 5.3. *Let $\Omega \subset \mathbb{R}^d$ be a convex polytope which is centrally symmetric and has centrally symmetric facets. Suppose that Ω admits a spectrum Λ satisfying* (5-2) *for every facet F of Ω. Then $\Omega + T$ is a tiling, and so Ω can tile by translations.*

Moreover, in this case the set T defined by (5-1) is a lattice in $\mathbb{R}^d$, and Ω tiles face-to-face by translations along the lattice T.

Remark. The formulation of Corollary 5.3 is inspired by [Iosevich et al. 2003, p. 568], where the assertion was proved in dimension $d = 2$ by directly showing that Ω must be either a parallelogram or a centrally symmetric hexagon. The proof in arbitrary dimension that we have given above is based on different considerations than the one in that paper.

6. Structure of spectrum, I

We obtained in Section 5 a sufficient condition for a spectral convex polytope Ω in $\mathbb{R}^d$ to tile by translations. This condition (Corollary 5.3) requires the existence of a spectrum Λ admitting a certain structure. In the present section we start to develop an approach to analyze the structure of a given spectrum Λ.

6A. Let $\Omega \subset \mathbb{R}^d$ be a convex polytope which is centrally symmetric and has centrally symmetric facets. We will assume $\Omega = -\Omega$; that is, Ω is symmetric about the origin. Let F be one of the facets of Ω, and assume that $F \subset \{x_1 = \frac{1}{2}\}$, and that the center of F is the point $(\frac{1}{2}, 0, 0, \dots, 0)$.

These assumptions are made merely for convenience. Later on, we will reduce the general situation to this more specific one by applying an affine transformation.

The assumptions imply that

$$F = \{\tfrac{1}{2}\} \times \Sigma,$$

where Σ is a convex polytope in $\mathbb{R}^{d-1}$ such that

$$\Sigma = -\Sigma.$$

The facet opposite to F is therefore

$$-F = \{-\tfrac{1}{2}\} \times \Sigma.$$

6B. For $\alpha > 0$ we consider the cone

$$K(\alpha) := \{\xi \in \mathbb{R}^d : |\xi_j| \leqslant \alpha|\xi_1| \ (2 \leqslant j \leqslant d)\}. \tag{6-1}$$

Lemma 6.1. *There is $\alpha = \alpha(\Omega) > 0$ such that*

$$\pi\xi_1 \hat{\mathbb{1}}_\Omega(\xi) = \sin \pi\xi_1 \cdot \hat{\mathbb{1}}_\Sigma(\xi_2, \xi_3, \dots, \xi_d) + O(|\xi_1|^{-1}), \quad |\xi_1| \to \infty, \tag{6-2}$$

in the cone $K(\alpha)$.

Proof. By Lemma 2.7, if α is sufficiently small then

$$-2\pi i \xi_1 \hat{\mathbb{1}}_\Omega(\xi) = \hat{\sigma}_F(\xi) - \hat{\sigma}_{-F}(\xi) + O(|\xi_1|^{-1}), \quad |\xi_1| \to \infty,$$

in $K(\alpha)$. But we have

$$\hat{\sigma}_F(\xi) = e^{-\pi i \xi_1} \hat{\mathbb{1}}_\Sigma(\xi_2, \xi_3, \dots, \xi_d) \quad \text{and} \quad \hat{\sigma}_{-F}(\xi) = e^{\pi i \xi_1} \hat{\mathbb{1}}_\Sigma(\xi_2, \xi_3, \dots, \xi_d),$$

which yields the conclusion of the lemma. □

6C. Assume now that we are given a set $\Lambda \subset \mathbb{R}^d$ which is a spectrum for Ω. To this spectrum Λ we associate a set $\Pi \subset \mathbb{R}^{d-1}$ defined as follows: Π is the set of all points $s \in \mathbb{R}^{d-1}$ such that for every open ball B containing s, the cylinder $\mathbb{R} \times B$ contains infinitely many points of Λ.

If we denote a point in $\mathbb{R}^d$ as $(t, s) \in \mathbb{R} \times \mathbb{R}^{d-1}$, then one can check that a point $s \in \mathbb{R}^{d-1}$ belongs to Π if and only if there is a sequence $(t_n, s_n) \in \Lambda$ such that

$$|t_n| \to \infty, \quad s_n \to s \quad (n \to \infty). \tag{6-3}$$

It is also not difficult to verify that Π is a *closed* subset of $\mathbb{R}^{d-1}$.

The motivation for introducing the set Π is the following observation:

Lemma 6.2. *For each $s \in \Pi$ there is a (unique) number $0 \leqslant \theta(s) < 1$ such that*

$$\Lambda \cap (\mathbb{R} \times (s + U_\Sigma)) \subset (\mathbb{Z} + \theta(s)) \times \mathbb{R}^{d-1},$$

where

$$U_\Sigma := \{\hat{\mathbb{1}}_\Sigma \neq 0\}.$$

In other words, if $(t', s') \in \Lambda$ and if $\hat{\mathbb{1}}_\Sigma(s' - s) \neq 0$, then $t' \in \mathbb{Z} + \theta(s)$.

Proof. It would be enough to show that if (t', s') and (t'', s'') are two points in $\Lambda \cap (\mathbb{R} \times (s + U_\Sigma))$, then $t'' - t' \in \mathbb{Z}$. Since $s \in \Pi$, there is a sequence $(t_n, s_n) \in \Lambda$ such that $|t_n| \to \infty$, $s_n \to s$. The vectors $(t' - t_n, s' - s_n)$ and $(t'' - t_n, s'' - s_n)$ belong to the set $(\Lambda - \Lambda) \setminus \{0\}$ for all large enough n; hence they lie in the zero set of $\hat{\mathbb{1}}_\Omega$. Using Lemma 6.1 it follows that

$$\sin \pi(t' - t_n) \cdot \hat{\mathbb{1}}_\Sigma(s' - s_n) \to 0, \quad \sin \pi(t'' - t_n) \cdot \hat{\mathbb{1}}_\Sigma(s'' - s_n) \to 0$$

as $n \to \infty$. Recall that $s' - s$ and $s'' - s$ are not in the zero set of $\hat{\mathbb{1}}_\Sigma$. Hence $|\hat{\mathbb{1}}_\Sigma(s' - s_n)|$ and $|\hat{\mathbb{1}}_\Sigma(s'' - s_n)|$ remain bounded away from zero as $n \to \infty$. We conclude that

$$\sin \pi(t' - t_n), \quad \sin \pi(t'' - t_n)$$

both tend to zero as $n \to \infty$, or equivalently,

$$\operatorname{dist}(t' - t_n, \mathbb{Z}), \quad \operatorname{dist}(t'' - t_n, \mathbb{Z})$$

both tend to zero. But

$$\operatorname{dist}(t'' - t', \mathbb{Z}) \leqslant \operatorname{dist}(t' - t_n, \mathbb{Z}) + \operatorname{dist}(t'' - t_n, \mathbb{Z}),$$

which implies $t'' - t' \in \mathbb{Z}$. □

Corollary 6.3. *Let $s', s'' \in \Pi$. If $\theta(s') \neq \theta(s'')$, then $\hat{\mathbb{1}}_\Sigma(s'' - s') = 0$.*

Proof. Let $(t_n, s_n) \in \Lambda$ be a sequence such that $|t_n| \to \infty$, $s_n \to s''$. If $\hat{\mathbb{1}}_\Sigma(s'' - s') \neq 0$, then for large enough n we would have $\hat{\mathbb{1}}_\Sigma(s_n - s') \neq 0$. By Lemma 6.2 it follows that $t_n \in \mathbb{Z} + \theta(s')$. On the other hand, for all large enough n we also have $\hat{\mathbb{1}}_\Sigma(s_n - s'') \neq 0$, since

$$\hat{\mathbb{1}}_\Sigma(0) = |\Sigma| > 0.$$

Hence, again by Lemma 6.2, we have $t_n \in \mathbb{Z} + \theta(s'')$. So we must have $\theta(s') = \theta(s'')$. □

6D. Lemma 6.2 allows us to define an equivalence relation on Π by saying that $s' \sim s''$ if $\theta(s') = \theta(s'')$. It follows from Corollary 6.3 that if Π' and Π'' are two distinct equivalence classes, then

$$\Pi'' - \Pi' \subset \{\hat{\mathbb{1}}_\Sigma = 0\}.$$

The set $\{\hat{\mathbb{1}}_\Sigma = 0\}$ is disjoint from the open ball of radius $\chi(\Sigma) > 0$ centered at the origin; see (2-2). It follows that each equivalence class is a closed set, and that there can be at most countably many equivalence classes. So we may enumerate them as $\Pi_0, \Pi_1, \Pi_2, \dots$ (finitely or infinitely many), and we denote by $\theta_0, \theta_1, \theta_2, \dots$ respectively the values of the function $\theta(s)$ on these equivalence classes.

6E. To illustrate the construction above, let us consider two representative examples.

Example 6.4. Assume that Ω tiles face-to-face along a lattice T of translation vectors, which in this case is given by (5-1). Since the facet F has the form $F = \left\{\frac{1}{2}\right\} \times \Sigma$, we have $\tau_F = (1, 0, 0, \dots, 0) \in T$. Let Λ be a spectrum of Ω given by the dual lattice; that is, $\Lambda = T^*$. Then $\langle \Lambda, \tau \rangle \subset \mathbb{Z}$ for any $\tau \in T$. In particular this is true for $\tau = \tau_F$; hence

$$\Lambda \subset \mathbb{Z} \times \mathbb{R}^{d-1}.$$

It follows that $\theta(s) = 0$ for all $s \in \Pi$. Thus in this case the set Π consists of a single equivalence class, namely $\Pi = \Pi_0$, and we have $\theta_0 = 0$.

Example 6.5. Assume that $\Omega = I \times \Sigma$, where I denotes the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$. Then Ω is a prism with base Σ. Suppose that Σ is a spectral set, and let $\Gamma \subset \mathbb{R}^{d-1}$ be a spectrum for Σ. For each $\gamma \in \Gamma$, let $\theta(\gamma)$ be an arbitrary real number, $0 \leqslant \theta(\gamma) < 1$, and define

$$\Lambda := \bigcup_{\gamma \in \Gamma} (\mathbb{Z} + \theta(\gamma)) \times \{\gamma\}.$$

It is known, see [Jorgensen and Pedersen 1999, Theorem 4], that Λ is a spectrum for Ω. In this case we clearly have $\Pi = \Gamma$, and the numbers $\theta(\gamma)$ coincide with the ones given by Lemma 6.2. The equivalence classes Π_j depend on the specific choice of the numbers $\theta(\gamma)$, but in the case when all the $\theta(\gamma)$ are distinct, the sets Π_j are singletons. Observe that we have

$$\Pi_j - \Pi_k \subset \{\hat{\mathbb{1}}_\Sigma = 0\} \quad (k \neq j)$$

since Γ is a spectrum for Σ. This is in accordance with Corollary 6.3.

7. Structure of spectrum, II

In this section we continue to work under the same assumptions as in Section 6. Namely, we assume $\Omega \subset \mathbb{R}^d$ is a convex polytope which is centrally symmetric, $\Omega = -\Omega$ and has centrally symmetric facets, F is one of the facets of Ω, and $F = \left\{\frac{1}{2}\right\} \times \Sigma$, where Σ is a convex polytope in $\mathbb{R}^{d-1}$ such that $\Sigma = -\Sigma$.

We also assume Λ is a spectrum for Ω, and to this spectrum Λ we associate the set $\Pi \subset \mathbb{R}^{d-1}$ that was defined in Section 6.

7A. From the given spectrum Λ one can construct a new spectrum Λ' for Ω in the following way. Consider the sequence of translates of Λ given by

$$\Lambda - k \cdot (1, 0, 0, \dots, 0), \quad k = 1, 2, 3, \dots .$$

Each one of these sets is a spectrum for Ω, and they are uniformly discrete with the same separation constant. Hence one may extract from this sequence a subsequence

$$\Lambda - k_n \cdot (1, 0, 0, \dots, 0), \quad k_n \to \infty,$$

which converges weakly to some set Λ', which is also a spectrum for Ω (see Section 2C). Notice that we do not make any claim concerning the uniqueness of the weak limit Λ', which in general may depend on the particular subsequence that was selected.

Lemma 7.1. *We have*

$$\Lambda' \subset \bigcup_{j \geqslant 0} (\mathbb{Z} + \theta_j) \times \Pi_j. \tag{7-1}$$

We remind that by θ_j $(j \geqslant 0)$ we denote the distinct values attained by the function $\theta(s)$ defined on Π, given in Lemma 6.2, and

$$\Pi_j = \{s \in \Pi : \theta(s) = \theta_j\}. \tag{7-2}$$

Recall also that according to Corollary 6.3 we have

$$\Pi_k - \Pi_j \subset \{\hat{\mathbb{1}}_\Sigma = 0\} \quad (j \neq k); \tag{7-3}$$

hence Lemma 7.1 reveals a certain structure satisfied by the new spectrum Λ'.

Proof of Lemma 7.1. The claim is equivalent to the statement that for every $(t', s') \in \Lambda'$ we have $s' \in \Pi$ and $t' \in \mathbb{Z} + \theta(s')$. Let therefore $(t', s') \in \Lambda'$. Since Λ' is the weak limit of the sequence $\Lambda - k_n \cdot (1, 0, 0, \dots, 0)$, there exist $(t_n, s_n) \in \Lambda$ such that

$$(t_n - k_n, s_n) \to (t', s'), \quad n \to \infty.$$

Hence $s_n \to s'$, and $t_n \to \infty$ since $k_n \to \infty$. This implies $s' \in \Pi$. For all sufficiently large n we have

$$\hat{\mathbb{1}}_\Sigma(s_n - s') \neq 0;$$

thus by Lemma 6.2 we have $t_n \in \mathbb{Z} + \theta(s')$. Since $t_n - k_n \to t'$ and the k_n are integers, this implies that also $t' \in \mathbb{Z} + \theta(s')$. $\square$

7B. Given a point (t_0, s_0) in $\mathbb{R} \times \mathbb{R}^{d-1}$, we associate with it a function f defined by

$$f(x, y) := \mathbb{1}_I(x) e^{2\pi i t_0 x} \mathbb{1}_\Sigma(y) e^{2\pi i \langle s_0, y \rangle}, \quad (x, y) \in \mathbb{R} \times \mathbb{R}^{d-1}, \tag{7-4}$$

where I denotes again the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$. Notice that the function f is supported by the prism $I \times \Sigma$. This prism is contained in Ω since $\left\{\frac{1}{2}\right\} \times \Sigma$ and $\left\{-\frac{1}{2}\right\} \times \Sigma$ are facets of Ω and Ω is convex. Hence f is also supported by Ω.

It follows from the definition (7-4) of f that its Fourier transform is given by

$$\hat{f}(t,s)=\hat{\mathbb{1}}_I(t-t_0)\,\hat{\mathbb{1}}_\Sigma(s-s_0),\quad (t,s)\in\mathbb{R}\times\mathbb{R}^{d-1}. \tag{7-5}$$

Using the function f thus defined, we can prove a result similar to Lemma 7.1 but which is concerned with the originally given spectrum Λ. However, the conclusion is somewhat weaker, as the right-hand side of (7-1) is replaced by a larger set:

Lemma 7.2. *We have*

$$\Lambda\subset\bigcup_{j\geqslant 0}(\mathbb{Z}+\theta_j)\times(\Pi_j+U_\Sigma),$$

where, *as before*, *we let*

$$U_\Sigma=\{\hat{\mathbb{1}}_\Sigma\neq 0\}.$$

Proof. By Lemma 6.2 we have

$$\Lambda\cap(\mathbb{R}\times(\Pi_j+U_\Sigma))\subset(\mathbb{Z}+\theta_j)\times(\Pi_j+U_\Sigma)$$

for every j. Hence, to prove the claim it would be enough to show that the sets Π_j+U_Σ cover the whole $\mathbb{R}^{d-1}$. Suppose to the contrary that there is a point $s_0\in\mathbb{R}^{d-1}$ which lies outside all the sets Π_j+U_Σ. Since $U_\Sigma=-U_\Sigma$, this means that

$$\hat{\mathbb{1}}_\Sigma(s-s_0)=0,\quad s\in\Pi.$$

Let t_0 be an arbitrary real number, and consider the function f defined by (7-4). Then f is supported by Ω, and by (7-5) its Fourier transform $\hat{f}$ vanishes on $\mathbb{R}\times\Pi$. In particular we have $\hat{f}(\lambda)=0$ for all $\lambda\in\Lambda'$, due to Lemma 7.1. That is,

$$\langle f,e_\lambda\rangle_{L^2(\Omega)}=\hat{f}(\lambda)=0,\quad \lambda\in\Lambda'.$$

Hence f is orthogonal in $L^2(\Omega)$ to all the exponentials $\{e_\lambda\}$, $\lambda\in\Lambda'$, which contradicts the completeness of the system $E(\Lambda')$ in the space $L^2(\Omega)$. ☐

Corollary 7.3. *Assume that the function $\theta(s)$ is constant on Π. Then*

$$\Lambda-\Lambda\subset\mathbb{Z}\times\mathbb{R}^{d-1}. \tag{7-6}$$

Proof. It is assumed that $\Pi=\Pi_0$ and $\theta(s)=\theta_0$ for all $s\in\Pi$. Hence by Lemma 7.2, the set Λ is contained in $(\mathbb{Z}+\theta_0)\times(\Pi_0+U_\Sigma)$, which implies (7-6). ☐

7C. Corollary 7.3 is an important point in our approach to the proof that Ω can tile by translations. Let us clarify its role. Recall that a sufficient condition for Ω to tile was given by Corollary 5.3; namely, it is enough to know that the spectrum Λ satisfies condition (5-2) for every facet F of Ω. For the facet $F=\left\{\frac{1}{2}\right\}\times\Sigma$ we have $\tau_F=(1,0,0,\dots,0)$; hence for this facet the condition (5-2) is the same as (7-6). It thus follows from Corollary 7.3 that in order to establish (5-2) for the facet $F=\left\{\frac{1}{2}\right\}\times\Sigma$, it would be sufficient to prove that the function $\theta(s)$ is constant on Π.

8. Spectral convex polygons tile the plane

8A. In this section we will demonstrate how the tools developed so far can be useful in our problem by showing that at this point they already enable us to give an alternative proof of the following result in dimension $d = 2$:

Theorem 8.1 [Iosevich et al. 2003]. *Let Ω be a convex polygon in $\mathbb{R}^2$. If Ω is spectral, then Ω tiles by translations.*

We remark that the paper [Iosevich et al. 2003] actually contains a proof of a more general result, which yields the same conclusion for any convex body $\Omega \subset \mathbb{R}^2$ (not assumed a priori to be a polygon).

8B. In order to prove Theorem 8.1, we now restrict ourselves to dimension $d = 2$. Let Ω be a convex polygon in $\mathbb{R}^2$. Assume that Ω is spectral, and let Λ be a spectrum for Ω. We must prove that Ω tiles by translations. This is obvious if Ω is a parallelogram, so in what follows we will assume Ω is not a parallelogram.

By Theorem 3.1 the polygon Ω is centrally symmetric, and since the facets of Ω are line segments, then automatically also all the facets of Ω are centrally symmetric.

Lemma 8.2. *Let Ω be a convex, centrally symmetric polygon in $\mathbb{R}^2$, and assume Ω is not a parallelogram. If Λ is a spectrum of Ω, then*

$$\langle \Lambda - \Lambda, \tau_F \rangle \subset \mathbb{Z} \tag{8-1}$$

for every facet F of Ω.

Theorem 8.1 follows immediately from a combination of Lemma 8.2 and Corollary 5.3. Hence, it only remains to prove the lemma.

Lemma 8.2 was proved in [Iosevich et al. 2003, Proposition 3.1], and was also used there to deduce that Ω tiles by translations. However, both our proof of Lemma 8.2, and the argument we use to deduce Theorem 8.1 from Lemma 8.2, are different from theirs.

8C. Now we give our proof of Lemma 8.2.

Proof of Lemma 8.2. Let F be a facet of Ω. We must show that if Λ is a spectrum of Ω, then it satisfies condition (8-1). By applying an affine transformation we may assume Ω is symmetric about the origin, $\Omega = -\Omega$, and that $F = \left\{\frac{1}{2}\right\} \times I$, where I is the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$. Hence we have $\Sigma = I,\ \tau_F = (1, 0)$, and condition (8-1) becomes

$$\Lambda - \Lambda \subset \mathbb{Z} \times \mathbb{R}. \tag{8-2}$$

Let $\Pi \subset \mathbb{R}$ be the set associated to the spectrum Λ defined as in Section 6, and $\theta(s)$ be the function on Π given by Lemma 6.2. By Corollary 7.3, to establish (8-2) it would be enough to show that $\theta(s)$ is constant on Π.

Let us first consider the case when

$$\Pi - \Pi \subset \mathbb{Z}. \tag{8-3}$$

We will show that in this case we must have $\Omega = I \times I$, that is, Ω is the unit cube, which is not possible as we have assumed that Ω is not a parallelogram. Indeed, suppose that (8-3) holds, and let Λ' be the spectrum constructed from Λ in Section 7. Fix a point $\lambda_0 = (t_0, s_0) \in \Lambda'$. It follows from Lemma 7.1 and (8-3) that if $\lambda' = (t', s')$ is any point in Λ' other than λ_0, then at least one of the numbers $t' - t_0$ and $s' - s_0$ must be in $\mathbb{Z} \setminus \{0\}$. Now consider the function f defined by (7-4). This function is supported by Ω, and by (7-5) its Fourier transform $\hat{f}$ vanishes on all the points of Λ' except for λ_0, since $\hat{\mathbb{1}}_I$ vanishes on $\mathbb{Z} \setminus \{0\}$. Hence f is orthogonal in $L^2(\Omega)$ to all the exponentials $\{e_\lambda\}$, $\lambda \in \Lambda' \setminus \{\lambda_0\}$. Since the system $E(\Lambda')$ is orthogonal and complete in $L^2(\Omega)$, this implies that f must coincide a.e. on Ω with a constant (nonzero) multiple of e_{λ_0}. In particular, f cannot vanish on any subset of Ω of positive measure. On the other hand, by the definition of f it does vanish on $\Omega \setminus (I \times I)$. This is possible only if $\Omega = I \times I$.

We thus conclude that (8-3) is not possible, so we must have

$$\Pi - \Pi \not\subset \mathbb{Z}. \tag{8-4}$$

Let us then show that $\theta(s)$ is a constant function on Π. Indeed, due to (8-4) there exist $s', s'' \in \Pi$ such that $s'' - s' \notin \mathbb{Z}$. Since $\{\hat{\mathbb{1}}_I = 0\} = \mathbb{Z} \setminus \{0\}$, Corollary 6.3 implies $\theta(s') = \theta(s'')$. Observe that for any $s \in \Pi$ we must have $s - s' \notin \mathbb{Z}$ or $s - s'' \notin \mathbb{Z}$, and in either case we obtain, again by Corollary 6.3, that $\theta(s) = \theta(s') = \theta(s'')$. This shows that $\theta(s)$ must be a constant function on Π. □

9. Prisms and cylindric sets

9A. The proof presented in Section 8 that a spectral convex polygon in the plane $\mathbb{R}^2$ can tile by translations eventually relied on showing that the function $\theta(s)$ is constant on the set Π. In order to show this we had to exclude the case when Ω is a parallelogram, but since a parallelogram automatically tiles by translations, this loss of generality was innocuous in the proof.

In dimension $d = 3$, however, the situation is more complicated. Even if we exclude the case when Ω is a parallelepiped, one still cannot expect to be able to prove that $\theta(s)$ is a constant function on Π. Indeed, we have seen in Example 6.5 above that if Ω is a prism whose base is a spectral set, then the function $\theta(s)$ may attain countably many arbitrary distinct values. Hence, the role of the parallelogram in dimension $d = 2$ will be played not by the parallelepiped, but by the prism, in dimension $d = 3$.

We remind the reader that by a *prism* in $\mathbb{R}^d$ one means a polytope Ω which can be expressed as the Minkowski sum of a $(d-1)$-dimensional polytope and a line segment.

Notice, however, that while a parallelogram in $\mathbb{R}^2$ automatically tiles by translations, this is not so for a prism in $\mathbb{R}^3$. Hence it is yet required to prove — necessarily by a different method — that a spectral convex prism in $\mathbb{R}^3$ can tile by translations.

Let us formulate this result explicitly:

Theorem 9.1. *Let Ω be a convex prism in $\mathbb{R}^3$. If Ω is spectral, then it tiles by translations.*

9B. A bounded, measurable set $\Omega \subset \mathbb{R}^d$ $(d \geqslant 2)$ will be called a *cylindric set* if it has the form $\Omega = I \times \Sigma$, where I is an interval in $\mathbb{R}$, and Σ is a measurable set in $\mathbb{R}^{d-1}$. In this case, the set Σ will be called the *base* of the cylindric set Ω.

If the base Σ is a convex polytope in $\mathbb{R}^{d-1}$, then the set $\Omega = I \times \Sigma$ is a convex prism. Conversely, any convex prism in $\mathbb{R}^d$ is the affine image of some set of the form $I \times \Sigma$, where I is an interval and Σ is a convex polytope in $\mathbb{R}^{d-1}$.

We will deduce Theorem 9.1 from the following result, proved in our paper [Greenfeld and Lev 2016]. The result is valid in all dimensions $d \geqslant 2$ (not just $d = 3$).

Theorem 9.2 [Greenfeld and Lev 2016]. *A cylindric set $\Omega = I \times \Sigma$ is spectral* (*as a set in $\mathbb{R}^d$*) *if and only if its base Σ is a spectral set* (*as a set in $\mathbb{R}^{d-1}$*).

This result thus provides a characterization of the cylindric spectral sets Ω in terms of the spectrality of their base Σ.

The "if" part of Theorem 9.2 is obvious. Suppose for simplicity that $I = \left[-\frac{1}{2}, \frac{1}{2}\right]$. If $\Gamma \subset \mathbb{R}^{d-1}$ is a spectrum for Σ, then it is easy to check that $\mathbb{Z} \times \Gamma$ is a spectrum for Ω; hence Ω is spectral.

On the other hand, the converse, "only if" part of the theorem (which is what we shall need for our present goal), is nontrivial. Roughly speaking, the difficulty lies in that knowing Ω to have a spectrum Λ in no way implies that Λ has a product structure as $\mathbb{Z} \times \Gamma$. In particular, we do not have any obvious candidate for a set $\Gamma \subset \mathbb{R}^{d-1}$ that might serve as a spectrum for Σ.

Remark. In [Greenfeld and Lev 2016] we also gave a similar characterization of the cylindric sets Ω in $\mathbb{R}^d$ which can tile the space by translations. Namely, it was proved there that a cylindric set $\Omega = I \times \Sigma$ tiles if and only if its base Σ tiles.

9C. Theorem 9.1 can now be obtained by a combination of Theorem 9.2 and the result from [Iosevich et al. 2003] that a spectral convex polygon in $\mathbb{R}^2$ can tile by translations, namely, Theorem 8.1 (for which we have provided an independent proof in Section 8).

Proof of Theorem 9.1. By applying an affine transformation we can assume $\Omega = I \times \Sigma$, where I is the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$ and Σ is a convex polygon in $\mathbb{R}^2$. Since Ω is spectral, it follows by Theorem 9.2 that also Σ is spectral. Hence by Theorem 8.1, Σ tiles by translations, so there is a set $\Gamma \subset \mathbb{R}^2$ such that $\Sigma + \Gamma$ is a tiling of $\mathbb{R}^2$. It is then clear that Ω tiles $\mathbb{R}^3$ with the translation set $\mathbb{Z} \times \Gamma$. $\square$

10. Prisms and zonotopes

In Section 9 we explained why the case when the convex polytope $\Omega \subset \mathbb{R}^3$ is a prism requires a special treatment in our approach. In this case we obtained a complete solution to our problem; namely, it was proved that if a convex prism in $\mathbb{R}^3$ is a spectral set, then it tiles by translations (Theorem 9.1). Hence, in what follows we will be mainly interested in the case when Ω is not a prism. The goal of the present section is to point out some geometric properties of such an Ω that will be useful in the analysis of the spectrum later on.

10A. Let $\Omega \subset \mathbb{R}^3$ be a convex polytope, centrally symmetric and with centrally symmetric facets. Let F be a facet of Ω, and F' be the opposite facet. Recall that by the central symmetry, F' is a translate of F, and that we have denoted by τ_F the translation vector in $\mathbb{R}^3$ which carries F' onto F, that is, $F = F' + \tau_F$.

Suppose now that A is a subfacet of F. Then A is the image under the translation by τ_F of a subfacet A' of F', that is, $A = A' + \tau_F$. We denote by $H_{F,A}$ the hyperplane which contains the subfacets A and A'.

Lemma 10.1. *If Ω is not a prism, then for any facet F of Ω there is a subfacet A such that* $\operatorname{int}(\Omega)$ *intersects each one of the two open half-spaces bounded by $H_{F,A}$.*

Proof. Let F be a facet of Ω. By applying an affine transformation we may assume

$$\Omega = -\Omega, \quad F = \left\{\tfrac{1}{2}\right\} \times \Sigma, \quad F' = \left\{-\tfrac{1}{2}\right\} \times \Sigma,$$

where Σ is a convex polygon in $\mathbb{R}^2$ such that $\Sigma = -\Sigma$. Suppose to the contrary that for any subfacet A of F, $\operatorname{int}(\Omega)$ entirely lies within one of the open half-spaces bounded by $H_{F,A}$. The intersection of the closures of all these half-spaces with the set $I \times \mathbb{R}^2$, where $I = \left[-\tfrac{1}{2}, \tfrac{1}{2}\right]$, is equal to $I \times \Sigma$. Hence Ω is contained in $I \times \Sigma$. But Ω also contains $I \times \Sigma$, since $I \times \Sigma$ is the convex hull of the facets F and F'. We conclude that $\Omega = I \times \Sigma$, which is not possible unless Ω is a prism. This contradiction ends the proof. $\square$

10B. By a *zonotope* in $\mathbb{R}^d$ one means a polytope which can be represented as the Minkowski sum of a finite number of line segments. A zonotope is a convex, centrally symmetric polytope, and all its facets are also zonotopes. In particular, all the facets of a zonotope are also centrally symmetric.

It is known, see, e.g., [Schneider 1993, Theorem 3.5.1], that in dimension $d = 3$, a convex polytope which has centrally symmetric facets must be a zonotope.

Remark, by the way, that this is not true in dimensions $d \geqslant 4$. A well-known example is the 24-cell in $\mathbb{R}^4$, a convex polytope which tiles by translations, and hence is centrally symmetric and has centrally symmetric facets, but which is not a zonotope.

10C. Let again $\Omega \subset \mathbb{R}^3$ be a convex polytope, centrally symmetric and with centrally symmetric facets (and hence a zonotope). Let F be a facet of Ω, and A, B be two parallel subfacets of F. Let F' and A', B' be the facet and its two subfacets which are carried onto F and A, B respectively by the translation vector τ_F. We denote by $S_{F,A,B}$ the closed slab which lies between the two parallel hyperplanes $H_{F,A}$ and $H_{F,B}$.

Lemma 10.2. *Assume that the intersection of Ω and $S_{F,A,B}$ coincides with the convex hull of the facets F and F'. Then Ω is a prism.*

Proof. By applying an affine transformation we may assume

$$\Omega = -\Omega, \quad F \subset \left\{x_1 = \tfrac{1}{2}\right\},$$

F is symmetric about the point $\left(\tfrac{1}{2}, 0, 0\right)$, and

$$A = \left\{\tfrac{1}{2}\right\} \times \left\{\tfrac{1}{2}\right\} \times I, \quad B = \left\{\tfrac{1}{2}\right\} \times \left\{-\tfrac{1}{2}\right\} \times I,$$

where I denotes as usual the interval $\left[-\tfrac{1}{2}, \tfrac{1}{2}\right]$. Hence $F = \left\{\tfrac{1}{2}\right\} \times \Sigma$, where Σ is a convex polygon in $\mathbb{R}^2$ such that $\Sigma = -\Sigma$, and such that $\left\{\tfrac{1}{2}\right\} \times I$, $\left\{-\tfrac{1}{2}\right\} \times I$ are facets of Σ.

The assumption in the lemma thus means that

$$\Omega \cap (\mathbb{R} \times I \times \mathbb{R}) = I \times \Sigma. \tag{10-1}$$

Since Ω is a zonotope, it can be represented as the Minkowski sum of several line segments $S_1, S_2, \dots, S_n$. Thus we have $\Omega = S_1 + S_2 + \cdots + S_n$. As Ω is symmetric about the origin, we can assume that the same is true for each line segment S_j; that is, $S_j = -S_j$. We can also assume that no two of the segments S_j are parallel.

Now we consider two distinct cases separately. Let us first consider the case when Σ is not the cube $I \times I$. In this case there must exist at least one vertex v of Σ which belongs to $\operatorname{int}(I \times \mathbb{R})$. Hence $I \times \{v\}$ is a subfacet of $I \times \Sigma$. By (10-1) it follows that $I \times \{v\}$ is also a subfacet of Ω. Each subfacet of Ω is a translate of one of the S_j's (see, for example, [McMullen 1971]). Hence one of the line segments, say S_1, must be equal to $I \times \{0\} \times \{0\}$. It then follows that all the other line segments $S_2, \dots, S_n$ must lie in $\{0\} \times \mathbb{R} \times \mathbb{R}$. Indeed, if this is not true for some S_j, then $S_1 + S_j$ is not contained in $I \times \mathbb{R} \times \mathbb{R}$. But $S_1 + S_j$ is contained in Ω, and Ω is contained in $I \times \mathbb{R} \times \mathbb{R}$, so this is not possible. Hence all the segments $S_2, \dots, S_n$ lie in $\{0\} \times \mathbb{R} \times \mathbb{R}$. It follows that $S_2 + \cdots + S_n = \{0\} \times \Sigma$, and $\Omega = I \times \Sigma$. This shows that Ω must be a prism.

Now we consider the remaining case, namely, when $\Sigma = I \times I$. In this case, the assumption (10-1) becomes

$$\Omega \cap (\mathbb{R} \times I \times \mathbb{R}) = I \times I \times I. \tag{10-2}$$

Hence $\mathbb{R} \times \mathbb{R} \times \{\frac{1}{2}\}$ and $\mathbb{R} \times \mathbb{R} \times \{-\frac{1}{2}\}$ are supporting hyperplanes of Ω, and thus $\Omega \subset \mathbb{R} \times \mathbb{R} \times I$. Since $A = \{\frac{1}{2}\} \times \{\frac{1}{2}\} \times I$ is a subfacet of Ω, then as before, one of the line segments, say again S_1, must be equal to $\{0\} \times \{0\} \times I$. It then follows that all the other line segments $S_2, \dots, S_n$ must lie in $\mathbb{R} \times \mathbb{R} \times \{0\}$, since if not, then as before, this would contradict the fact that $\Omega \subset \mathbb{R} \times \mathbb{R} \times I$. Hence $S_2 + \cdots + S_n = P \times \{0\}$ for a certain convex polygon $P \subset \mathbb{R}^2$, and $\Omega = P \times I$. Again we obtain that Ω must be a prism, so this proves the lemma. $\square$

11. Structure of spectrum, III

In this section our goal is to relate the geometric observations made in Section 10 to the spectrality problem for convex polytopes in dimension $d = 3$. More specifically, we will see how one can use the assumption that Ω is not a prism in order to obtain new information on the structure of the spectrum Λ.

11A. Let $\Omega \subset \mathbb{R}^3$ be a convex polytope, centrally symmetric and with centrally symmetric facets. Assume, as before, that $\Omega = -\Omega$; that is, Ω is symmetric about the origin, F is a facet of Ω contained in $\{x_1 = \frac{1}{2}\}$, and the center of F is the point $(\frac{1}{2}, 0, 0)$. Hence $F = \{\frac{1}{2}\} \times \Sigma$, where Σ is a convex polygon in $\mathbb{R}^2$ such that $\Sigma = -\Sigma$.

Suppose also that Λ is a spectrum for Ω. Let $\Pi \subset \mathbb{R}^2$ be the set associated to the spectrum Λ defined as in Section 6 and $\theta(s)$ be the function on Π given by Lemma 6.2. We also let Λ' be the new spectrum constructed from Λ in Section 7.

Recall that to each point $(t_0, s_0) \in \mathbb{R} \times \mathbb{R}^2$ we have associated a function f, supported by Ω, defined by (7-4). As an element of $L^2(\Omega)$, this function f admits a Fourier expansion with respect to the spectrum Λ', given by

$$f = \frac{1}{|\Omega|} \sum_{\lambda \in \Lambda'} \hat{f}(\lambda) e_\lambda. \tag{11-1}$$

By Lemma 2.1 the series on the right-hand side of (11-1) converges in L^2 on any bounded set to a measurable function $\tilde{f}$ on $\mathbb{R}^3$, and f coincides with $\tilde{f}$ a.e. on Ω.

We now observe that for certain values of (t_0, s_0), the Fourier expansion of f with respect to the spectrum Λ' consists of exceptionally few terms:

Lemma 11.1. *Let (t_0, s_0) be a point belonging to $(\mathbb{Z}+\theta_j) \times \Pi_j$ for some j, and let f be the function defined by* (7-4). *Then the Fourier expansion* (11-1) *of f with respect to the spectrum Λ' consists only of terms corresponding to $\lambda \in \Lambda' \cap (\{t_0\} \times \Pi_j)$.*

In other words, all the coefficients $\hat{f}(\lambda)$ in the expansion (11-1) must vanish except for possibly those which correspond to $\lambda = (t, s) \in \Lambda'$ such that $t = t_0$ and $s \in \Pi_j$.

Proof of Lemma 11.1. If $(t, s) \in \Lambda'$, then by Lemma 7.1 there is k such that $t \in \mathbb{Z}+\theta_k$ and $s \in \Pi_k$. If $k \neq j$ then $\hat{\mathbb{1}}_\Sigma(s - s_0) = 0$ due to (7-3), and it follows from (7-5) that $\hat{f}(t, s) = 0$. If $k = j$ then both t_0 and t belong to $\mathbb{Z}+\theta_j$; hence $t - t_0$ is an integer. Since $\hat{\mathbb{1}}_I$ vanishes on $\mathbb{Z} \setminus \{0\}$, it follows again by (7-5) that $\hat{f}(t, s) = 0$ unless $t = t_0$. This shows that in the series (11-1) the nonzero coefficients can only correspond to $\lambda = (t, s)$ such that $t = t_0$ and $s \in \Pi_j$. □

Remark. It may be interesting to notice that Lemma 11.1 implies that Λ' must contain points from *each one* of the sets $\{t_0\} \times \Pi_j$, where t_0 goes through the elements of $\mathbb{Z}+\theta_j$.

11B. Now suppose that Ω is *not a prism*. Then by Lemma 10.1 there is a subfacet A of F such that $\operatorname{int}(\Omega)$ intersects each one of the two open half-spaces bounded by the hyperplane $H_{F,A}$. Let us assume, for simplicity, that this subfacet is $A = \{\frac{1}{2}\} \times \{\frac{1}{2}\} \times I$, where $I = [-\frac{1}{2}, \frac{1}{2}]$ (later on, the general situation will be reduced to this case by applying an affine transformation). Thus $\{\frac{1}{2}\} \times I$ is a facet of the convex polygon Σ.

We can now use Lemma 11.1 to obtain some additional information on the structure of the components Π_j of the set Π.

Lemma 11.2. *For each j we have*

$$\Pi_j - \Pi_j \not\subset \mathbb{Z} \times \mathbb{R}. \tag{11-2}$$

Proof. Suppose that (11-2) is not true for some j. By translating the spectrum Λ we can assume Π_j contains the origin, and hence

$$\Pi_j \subset \mathbb{Z} \times \mathbb{R}. \tag{11-3}$$

Choose a point $(t_0, s_0) \in (\mathbb{Z}+\theta_j) \times \Pi_j$, and let f be the function associated to this point defined by (7-4). By Lemma 11.1 and due to (11-3), the Fourier expansion of f with respect to Λ' consists only of exponentials e_λ such that $\lambda \in \Lambda' \cap (\mathbb{R} \times \mathbb{Z} \times \mathbb{R})$. It follows (Lemma 2.1) that the right-hand side of (11-1) is a function $\tilde{f}$ on $\mathbb{R}^3$ which is periodic with respect to the vector $(0, 1, 0)$, and f coincides with $\tilde{f}$ a.e. on Ω.

Recall that we have chosen the subfacet A of F (using Lemma 10.1) such that $\operatorname{int}(\Omega)$ intersects each one of the two open half-spaces bounded by the hyperplane $H_{F,A}$. Since it was assumed that $A = \{\frac{1}{2}\} \times \{\frac{1}{2}\} \times I$, this means that $H_{F,A} = \{x_2 = \frac{1}{2}\}$, and hence

$$\Omega \not\subset \{x_2 \leqslant \tfrac{1}{2}\}. \tag{11-4}$$

Recall also that $F = \{\frac{1}{2}\} \times \Sigma$, where Σ is a convex polygon in $\mathbb{R}^2$, $\Sigma = -\Sigma$, and $\{\frac{1}{2}\} \times I$ is a face of Σ. By convexity, Σ contains the unit square $I \times I$, and hence $I \times \Sigma$ contains the unit cube $I \times I \times I$.

Thus $|\tilde{f}| = |f| = 1$ a.e. on $I \times I \times I$. By the periodicity of $\tilde{f}$ this implies $|\tilde{f}| = 1$ a.e. on $I \times \mathbb{R} \times I$. In particular, $|\tilde{f}| = 1$ a.e. on the set

$$\Omega \cap (I \times (\mathbb{R} \setminus I) \times I). \tag{11-5}$$

On the other hand, the set (11-5) is disjoint from $I \times \Sigma$; hence $|f| = 0$ on this set. It follows that the set (11-5) cannot have positive measure, and therefore

$$\Omega \cap (I \times \mathbb{R} \times I) = I \times I \times I.$$

This implies that $\{x_2 = \frac{1}{2}\}$ is a supporting hyperplane of Ω, which contradicts (11-4). □

Lemma 11.3. *For each j we have*

$$\Pi_j - \Pi_j \not\subset \mathbb{R} \times \mathbb{Z}. \tag{11-6}$$

Proof. We argue in a way similar to the proof of the previous lemma. If (11-6) is violated for some j, then by translating Λ we can assume

$$\Pi_j \subset \mathbb{R} \times \mathbb{Z}. \tag{11-7}$$

Hence, choosing a point $(t_0, s_0) \in (\mathbb{Z} + \theta_j) \times \Pi_j$, the corresponding function f defined by (7-4) coincides a.e. on Ω with a function $\tilde{f}$ on $\mathbb{R}^3$, which by (11-7) and Lemma 11.1 is periodic with respect to the vector $(0, 0, 1)$.

Since we have $|\tilde{f}| = |f| = 1$ a.e. on $I \times I \times I$, the periodicity of $\tilde{f}$ implies $|\tilde{f}| = 1$ a.e. on $I \times I \times \mathbb{R}$. In particular, $|\tilde{f}| = 1$ a.e. on the set

$$\Omega \cap (I \times ((I \times \mathbb{R}) \setminus \Sigma)). \tag{11-8}$$

But since this set is disjoint from $I \times \Sigma$, we have $|f| = 0$ on the set (11-8). So the set (11-8) cannot have positive measure, and therefore

$$\Omega \cap (I \times I \times \mathbb{R}) = I \times \Sigma.$$

By Lemma 10.2 this is possible only if Ω is a prism, so this concludes the proof. □

Lemma 11.4. *Let X be a subset of an abelian group G, and let H_1 and H_2 be two subgroups of G. Assume that*

$$X - X \subset H_1 \cup H_2. \tag{11-9}$$

Then $X - X \subset H_1$ or $X - X \subset H_2$.

Proof. Suppose that $X - X \not\subset H_1$, so there exist $x, y \in X$ such that $x - y \notin H_1$. Then by (11-9) we have $x - y \in H_2$. The property $x - y \notin H_1$ implies that for each $z \in X$ we must have $z - x \notin H_1$ or $z - y \notin H_1$. But in either case, it follows from (11-9) that $z \in x + H_2 = y + H_2$, so we conclude that $X \subset x + H_2 = y + H_2$. Thus $X - X \subset H_2$. □

Corollary 11.5. *For each j we have*

$$\Pi_j - \Pi_j \not\subset (\mathbb{Z} \times \mathbb{R}) \cup (\mathbb{R} \times \mathbb{Z}). \tag{11-10}$$

This is an immediate consequence of Lemmas 11.2, 11.3 and 11.4.

12. Structure of spectrum, IV

In the present section, we continue to analyze the structure of the spectrum of a convex polytope Ω in dimension $d = 3$. Although we are mainly interested in the case when Ω is not a prism, we will not need to assume this in the present section.

12A. Let Ω be a convex polytope in $\mathbb{R}^3$, centrally symmetric and with centrally symmetric facets. Assume that Ω is in our "standard position"; namely, $\Omega = -\Omega$, F is a facet of Ω contained in $\{x_1 = \frac{1}{2}\}$, and F is symmetric about the point $(\frac{1}{2}, 0, 0)$. Assume also that $A = \{\frac{1}{2}\} \times \{\frac{1}{2}\} \times I$ is a subfacet of F, where $I = [-\frac{1}{2}, \frac{1}{2}]$. Hence $F = \{\frac{1}{2}\} \times \Sigma$, where Σ is a convex polygon in $\mathbb{R}^2$, $\Sigma = -\Sigma$, and $\{\frac{1}{2}\} \times I$ is a facet of Σ.

Suppose that Λ is a spectrum for Ω. Let $\Pi \subset \mathbb{R}^2$ be the set associated to the spectrum Λ defined in Section 6 and $\theta(s)$ be the function on Π given by Lemma 6.2. Recall that in Section 7 a new spectrum Λ' was constructed from the given spectrum Λ by taking the weak limit of a sequence of translates of Λ. The new spectrum Λ' was shown (Lemma 7.1) to enjoy a particular structure, namely

$$\Lambda' \subset \bigcup_{j \geqslant 0} (\mathbb{Z} + \theta_j) \times \Pi_j, \tag{12-1}$$

where Π_j are the components of the set Π, and θ_j are respectively the values of the function $\theta(s)$ on these components. The sets Π_j were shown (Corollary 6.3) to satisfy

$$\Pi_k - \Pi_j \subset \{\hat{\mathbb{1}}_\Sigma = 0\} \quad (j \neq k). \tag{12-2}$$

When we want to further analyze the structure of the spectrum in dimension $d = 3$, a new complication arises that was not present in the case $d = 2$. Namely, the zero set $\{\hat{\mathbb{1}}_\Sigma = 0\}$ is not known explicitly, except in the special case when Σ is the cube $I \times I$. In order to address this difficulty, a further limiting procedure will now be performed on the spectrum Λ', yielding a third spectrum Λ'' of Ω.

12B. The new spectrum Λ'' is constructed as follows. Consider the sequence of translates of the spectrum Λ' given by

$$\Lambda' - r \cdot (0, 1, 0), \quad r = 1, 2, 3, \ldots.$$

As in Section 7 we may extract from this sequence a subsequence

$$\Lambda' - r_n \cdot (0, 1, 0), \quad r_n \to \infty, \tag{12-3}$$

which converges weakly to some set Λ'', which is again a spectrum of Ω.

According to (12-1) we may form a partition of the spectrum Λ' into sets defined by

$$\Lambda'_j := \Lambda' \cap ((\mathbb{Z} + \theta_j) \times \Pi_j). \tag{12-4}$$

It would be convenient for us to know that for each j, the sequence of translates

$$\Lambda'_j - r_n \cdot (0, 1, 0) \tag{12-5}$$

of each component Λ'_j has a weak limit as $n \to \infty$. This does not follow automatically from the weak convergence of the sequence (12-3), though, since we have not excluded the possibility that there may be infinitely many θ_j and that they may have accumulation points. Nevertheless, we can assume that (12-5) has a weak limit as $n \to \infty$ for each j, simply by selecting a further subsequence if necessary.

We shall denote by Λ''_j the weak limit of (12-5). Observe that a point $(t, u, v) \in \mathbb{R}^3$ belongs to Λ''_j if and only if there is a sequence $(t_n, u_n, v_n) \in \Lambda'_j$ such that

$$(t_n, u_n - r_n, v_n) \to (t, u, v), \quad n \to \infty.$$

Remark that while by Lemma 11.1 none of the components Λ'_j may be empty, this is not true for the sets Λ''_j that we cannot exclude some of which to be empty.

It follows from (12-4) that

$$\Lambda''_j \subset \Lambda'' \cap ((\mathbb{Z} + \theta_j) \times \mathbb{R}^2); \tag{12-6}$$

hence the sets Λ''_j are disjoint subsets of Λ''. Remark, however, that these sets do not necessarily form a partition of Λ''; namely, their union need not be equal to the whole Λ''. Again, this may happen only if there are infinitely many θ_j. An example of such a situation can be obtained if Ω is a prism whose base is a spectral set. Indeed, we have seen in Example 6.5 that in such a case the function $\theta(s)$ may attain countably many arbitrary distinct values, and that the components Π_j of the set Π may be singletons. This implies that every Λ''_j is empty, while Λ'' certainly cannot be empty being a spectrum for Ω.

This makes it necessary for us in general to consider also the subset of Λ'' defined by

$$\Lambda''_\infty := \Lambda'' \setminus \bigcup_{j \geqslant 0} \Lambda''_j.$$

Lemma 12.1. *Let $(t, u, v) \in \mathbb{R}^3$. Then (t, u, v) belongs to Λ''_∞ if and only if there is a sequence $k_n \to \infty$, and for each n there is a point $(t_n, u_n, v_n) \in \Lambda'_{k_n}$ such that*

$$(t_n, u_n - r_n, v_n) \to (t, u, v), \quad n \to \infty.$$

Proof. Suppose first that (t, u, v) is a point in Λ''_∞. Then $(t, u, v) \in \Lambda''$, and since Λ'' is the weak limit of (12-3), there exist $(t_n, u_n, v_n) \in \Lambda'$ such that $(t_n, u_n - r_n, v_n) \to (t, u, v)$. Due to (12-1) and (12-4), for each n there is $k_n \geqslant 0$ such that $(t_n, u_n, v_n) \in \Lambda'_{k_n}$. If $k_n \not\to \infty$, then k_n admits infinitely often a certain value, say $k_n = j$, for infinitely many n. But this implies that (t, u, v) must belong to the weak limit of (12-5), and hence $(t, u, v) \in \Lambda''_j$, so it cannot lie in Λ''_∞. Hence we must have $k_n \to \infty$.

Conversely, suppose that the point (t, u, v) satisfies the condition in the lemma. The condition implies that (t, u, v) belongs to the weak limit of (12-3); hence $(t, u, v) \in \Lambda''$. If (t, u, v) is not in Λ''_∞, then it belongs to one of the sets Λ''_j. But then we must have $k_n = j$ for all sufficiently large n, so $k_n \not\to \infty$, a contradiction. Hence $(t, u, v) \in \Lambda''_\infty$. □

We also point out that the inclusion (12-6) is not necessarily an equality, as the right-hand side of (12-6) may contain elements of Λ''_∞.

12C. Now we establish some properties satisfied by the new spectrum Λ'' and its components Λ''_k $(0 \leqslant k \leqslant \infty)$. The first property is derived from the condition (12-2).

Lemma 12.2. *For each* $0 \leqslant j, k \leqslant \infty,\ j \neq k$, *we have*

$$\Lambda''_k - \Lambda''_j \subset \mathbb{R} \times \{\hat{\mathbb{1}}_\Sigma = 0\}. \tag{12-7}$$

Proof. By symmetry we may assume $0 \leqslant j < k \leqslant \infty$. Let $(t, u, v) \in \Lambda''_j$ and $(t', u', v') \in \Lambda''_k$. Then there exist two sequences

$$(t_n, u_n, v_n) \in \Lambda'_j, \quad (t_n, u_n - r_n, v_n) \to (t, u, v),$$

and

$$(t'_n, u'_n, v'_n) \in \Lambda'_{k_n}, \quad (t'_n, u'_n - r_n, v'_n) \to (t', u', v'),$$

where $k_n = k$ in the case when k is finite and $k_n \to \infty$ if $k = \infty$ (Lemma 12.1). In any case we have $k_n \neq j$ for all sufficiently large n. Since by (12-4) we have

$$(u_n, v_n) \in \Pi_j, \quad (u'_n, v'_n) \in \Pi_{k_n},$$

it follows from (12-2) that

$$(t'_n, u'_n - r_n, v'_n) - (t_n, u_n - r_n, v_n) = (t'_n - t_n, u'_n - u_n, v'_n - v_n) \in \mathbb{R} \times \{\hat{\mathbb{1}}_\Sigma = 0\}.$$

Letting $n \to \infty$ we obtain

$$(t', u', v') - (t, u, v) \in \mathbb{R} \times \{\hat{\mathbb{1}}_\Sigma = 0\},$$

which confirms (12-7). □

Lemma 12.2 shows that the structure (12-2) is basically preserved in the new spectrum Λ'' and its components Λ''_k $(0 \leqslant k \leqslant \infty)$. However, our motivation for introducing this new spectrum is due to the following lemma:

Lemma 12.3. *Let* $0 \leqslant j < \infty,\ 0 \leqslant k \leqslant \infty,\ k \neq j$. *Then*

$$\Lambda''_k - \mathbb{R} \times \Pi_j \subset (\mathbb{R} \times \mathbb{Z} \times \mathbb{R}) \cup (\mathbb{R} \times \mathbb{R} \times (\mathbb{Z} \setminus \{0\})). \tag{12-8}$$

In other words, if $(u_0, v_0) \in \Pi_j$ *and if* $(t, u, v) \in \Lambda''_k$, *then* $u - u_0 \in \mathbb{Z}$ *or* $v - v_0 \in \mathbb{Z} \setminus \{0\}$.

This lemma is similar in spirit to Lemma 6.2. To see the resemblance between the two lemmas, recall that $\{\frac{1}{2}\} \times I$ is a facet of the polygon Σ, and $\{\hat{\mathbb{1}}_I = 0\} = \mathbb{Z} \setminus \{0\}$. The assertion of (12-8) is equivalent to the statement that if $(u_0, v_0) \in \Pi_j,\ (t, u, v) \in \Lambda''_k$, and if $\hat{\mathbb{1}}_I(v - v_0) \neq 0$, then $u \in \mathbb{Z} + u_0$. The proof is also similar to that of Lemma 6.2.

Proof of Lemma 12.3. Let $(u_0, v_0) \in \Pi_j$ and $(t, u, v) \in \Lambda''_k$. Regardless of whether k is finite or not, there is a sequence k_n and there are points $(t_n, u_n, v_n) \in \Lambda'_{k_n}$ such that

$$(t_n, u_n - r_n, v_n) \to (t, u, v), \quad n \to \infty. \tag{12-9}$$

Indeed, if $0 \leqslant k < \infty$ then $k_n = k$ for all n, while if $k = \infty$ then $k_n \to \infty$ (Lemma 12.1). In any case, we have $k_n \neq j$ for all sufficiently large n. Since $(t_n, u_n, v_n) \in \Lambda'_{k_n}$ we have $(u_n, v_n) \in \Pi_{k_n}$ by (12-4). Hence by (12-2) this implies

$$\hat{\mathbb{1}}_\Sigma(u_n - u_0, v_n - v_0) = 0 \tag{12-10}$$

for all sufficiently large n.

Observe that since $r_n \to \infty$, (12-9) implies that also $u_n \to \infty$. Hence using Lemma 6.1 for the polygon Σ and its facet $\left\{\frac{1}{2}\right\} \times I$, it follows from (12-10) that

$$\sin \pi(u_n - u_0) \cdot \hat{\mathbb{1}}_I(v_n - v_0) \to 0, \quad n \to \infty.$$

Indeed, the polygon Σ is centrally symmetric and it has centrally symmetric facets, as the facets of Σ are line segments; hence all the conditions of Lemma 6.1 are satisfied.

Now suppose that $v - v_0 \notin \mathbb{Z} \setminus \{0\}$. Then $v - v_0$ is not contained in the zero set of $\hat{\mathbb{1}}_I$, and hence $|\hat{\mathbb{1}}_I(v_n - v_0)|$ remains bounded away from zero as $n \to \infty$. So we must have $\sin \pi(u_n - u_0) \to 0$, or equivalently, $\operatorname{dist}(u_n - u_0, \mathbb{Z}) \to 0$. But since r_n is an integer, (12-9) implies that also $\operatorname{dist}(u_n - u, \mathbb{Z}) \to 0$. It follows that

$$\operatorname{dist}(u - u_0, \mathbb{Z}) \leqslant \operatorname{dist}(u_n - u_0, \mathbb{Z}) + \operatorname{dist}(u_n - u, \mathbb{Z}) \to 0.$$

We conclude that $u - u_0 \in \mathbb{Z}$ as required. □

From the previous lemma it is easy to deduce the next one:

Lemma 12.4. *For each* $0 \leqslant j, k \leqslant \infty$, $j \neq k$, *we have*

$$\Lambda''_k - \Lambda''_j \subset (\mathbb{R} \times \mathbb{Z} \times \mathbb{R}) \cup (\mathbb{R} \times \mathbb{R} \times (\mathbb{Z} \setminus \{0\})). \tag{12-11}$$

Actually we will not use Lemma 12.4 in what follows. We state it merely to demonstrate an essential advantage of the newly constructed spectrum Λ''. On one hand, according to (12-7) it basically inherits the structure of the previously constructed spectrum Λ', while on the other hand, condition (12-11) reveals an extra structure in Λ''.

Since the proof of Lemma 12.4 is quite short, we include it for completeness.

Proof of Lemma 12.4. By symmetry we may assume $0 \leqslant j < k \leqslant \infty$. Let $(t, u, v) \in \Lambda''_k$. Then by Lemma 12.3 the set $\mathbb{R} \times \Pi_j$ must be contained in

$$(\mathbb{R} \times (u + \mathbb{Z}) \times \mathbb{R}) \cup (\mathbb{R} \times \mathbb{R} \times (v + (\mathbb{Z} \setminus \{0\}))). \tag{12-12}$$

Due to (12-4) we have $\Lambda'_j \subset \mathbb{R} \times \Pi_j$; hence also the set Λ'_j is contained in (12-12). Since the set (12-12) is invariant under translations by vectors in $\{0\} \times \mathbb{Z} \times \{0\}$, it follows that all the sets (12-5) are also contained in (12-12), and hence the same is true for their weak limit Λ''_j. This implies that $\Lambda''_j - (t, u, v)$ is contained in the set on the right-hand side of (12-11). As (t, u, v) was an arbitrary element of Λ''_k, this establishes (12-11). □

13. Auxiliary lemmas

In this section we establish some specific facts about the spectrum of a convex polytope Ω that will be used later on. These facts are true in arbitrary dimension, so in the present section we do not restrict the discussion to three dimensions.

13A. Let $\Omega \subset \mathbb{R}^d$ be a convex polytope. Let F and F' be two parallel facets of Ω, and assume $F \subset \{x_1 = \frac{1}{2}\}$, $F' \subset \{x_1 = -\frac{1}{2}\}$, and that F is the image of F' under translation by the vector $\vec{e}_1$. These assumptions imply that

$$F = \{\tfrac{1}{2}\} \times \Sigma, \quad F' = \{-\tfrac{1}{2}\} \times \Sigma,$$

where Σ is a convex polytope in $\mathbb{R}^{d-1}$.

Assume also that Ω is spectral, and let Λ be a spectrum for Ω.

Lemma 13.1. *If Ω is not a prism, then Λ cannot contain any set of the form*

$$(\mathbb{Z}+\theta) \times \{s\}, \tag{13-1}$$

where $\theta \in \mathbb{R}$ and $s \in \mathbb{R}^{d-1}$.

Proof. Suppose to the contrary that Λ does contain a set of the form (13-1). This implies that the set $\Lambda - \Lambda$ contains $\mathbb{Z} \times \{0\}$. On the other hand, since Λ is a spectrum for Ω, the set $\Lambda - \Lambda$ must be contained in $\{\hat{\mathbb{1}}_\Omega = 0\} \cup \{0\}$. We conclude that

$$\hat{\mathbb{1}}_\Omega(k, 0) = 0, \quad k \in \mathbb{Z} \setminus \{0\}. \tag{13-2}$$

For each $x \in \mathbb{R}$ denote by Ω_x the $(d-1)$-dimensional polytope obtained by the intersection of Ω with the hyperplane $\{x\} \times \mathbb{R}^{d-1}$, and let $\varphi(x)$ be the $(d-1)$-dimensional volume of Ω_x. Then the function φ vanishes off the interval $I = [-\frac{1}{2}, \frac{1}{2}]$, it is continuous on I, and $\varphi(\frac{1}{2}) = \varphi(-\frac{1}{2}) = |\Sigma|$. Notice that, by convexity, Ω_x contains $\{x\} \times \Sigma$ for every $x \in I$. In particular this implies $\varphi(x) \geqslant |\Sigma|$, $x \in I$.

It follows from the definition of the function φ that its Fourier transform is given by

$$\hat{\varphi}(t) = \hat{\mathbb{1}}_\Omega(t, 0), \quad t \in \mathbb{R}.$$

Combining this with (13-2) we obtain that $\hat{\varphi}$ vanishes on $\mathbb{Z} \setminus \{0\}$. Since φ is supported on I, this implies that φ is orthogonal in $L^2(I)$ to all the exponentials $\{e_k\}$, $k \in \mathbb{Z} \setminus \{0\}$. But as the system $E(\mathbb{Z})$ is orthogonal and complete in $L^2(I)$, this is possible only if φ is constant on I. Hence $\varphi(x) = |\Sigma|$ for all $x \in I$. In turn, this implies $\Omega_x = \{x\} \times \Sigma$, $x \in I$. We conclude that $\Omega = I \times \Sigma$, and so Ω is a prism, a contradiction. $\square$

Remark. One can see from the proof that the only property of the set (13-1) that was actually used was that its difference set contains $\mathbb{Z} \times \{0\}$. Hence the lemma remains true if (13-1) is replaced by any other set for which the latter property is satisfied.

13B. Denote by $Q = I^{d-1}$ the unit cube in $\mathbb{R}^{d-1}$. As usual, I is the interval $[-\frac{1}{2}, \frac{1}{2}]$.

Lemma 13.2. *Assume that Σ contains Q. If Ω is not a prism, then Λ cannot be covered by the union of two translates of $\mathbb{Z}^d$.*

Proof. Suppose to the contrary that Λ is contained in the union of two translates of $\mathbb{Z}^d$. By translating Λ we may assume

$$\Lambda \subset \mathbb{Z}^d \cup (\mathbb{Z}^d + \tau) \tag{13-3}$$

for some $\tau \in \mathbb{R}^d$. According to Lemma 13.1, the spectrum Λ cannot contain the whole set $\mathbb{Z} \times \{0\}$. This implies that by further translating Λ by a certain vector in $\mathbb{Z} \times \{0\}$, we may additionally assume Λ does not contain the origin.

Since Σ is assumed to contain Q, and since by convexity Ω contains $I \times \Sigma$, it follows that Ω must contain $I \times Q$, the unit cube in $\mathbb{R}^d$. Hence the function $f = \mathbb{1}_{I \times Q}$ is supported by Ω. Consider the Fourier expansion (2-3) of this function f. Since $\hat{f}$ vanishes on all the points of $\mathbb{Z}^d$ except the origin, and since the origin does not belong to Λ, it follows from (13-3) that only exponentials e_λ such that $\lambda \in \Lambda \cap (\mathbb{Z}^d + \tau)$ may have a nonzero coefficient in the expansion (2-3). Hence by Lemma 2.1 the right-hand side of (2-3) represents a function $\tilde{f}$ of the form

$$\tilde{f}(x) = e^{2\pi i \langle \tau, x\rangle} g(x), \quad x \in \mathbb{R}^d,$$

where g is some $\mathbb{Z}^d$-periodic function, and f coincides with $\tilde{f}$ a.e. on Ω. Notice that $|g| = |\tilde{f}| = |f| = 1$ a.e. on $I \times Q$. By the periodicity of g this implies $|g| = 1$ a.e. on $\mathbb{R}^d$. Hence $|f| = |\tilde{f}| = |g| = 1$ a.e. on Ω. In particular, f cannot vanish on any subset of Ω of positive measure. On the other hand, by the definition of f it does vanish on $\Omega \setminus (I \times Q)$. This is possible only if $\Omega = I \times Q$; namely, Ω is the unit cube in $\mathbb{R}^d$. But this contradicts the assumption that Ω is not a prism, so the proof is complete. $\square$

14. Structure of spectrum, V

In this section we complete the analysis of the spectrum in dimension $d = 3$.

14A. Our assumptions will be the following.

Let $\Omega \subset \mathbb{R}^3$ be a convex polytope, centrally symmetric and with centrally symmetric facets. We assume Ω is not a prism. Suppose that Ω is in the "standard position"; namely, $\Omega = -\Omega$, F is a facet of Ω contained in $\{x_1 = \frac{1}{2}\}$, and F is symmetric about the point $(\frac{1}{2}, 0, 0)$. Hence $F = \{\frac{1}{2}\} \times \Sigma$, where Σ is a convex polygon in $\mathbb{R}^2$ such that $\Sigma = -\Sigma$. We assume $A = \{\frac{1}{2}\} \times \{\frac{1}{2}\} \times I$ is a subfacet of F, where $I = [-\frac{1}{2}, \frac{1}{2}]$, and therefore $\{\frac{1}{2}\} \times I$ is a facet of Σ. We also suppose that $\operatorname{int}(\Omega)$ intersects each one of the two open half-spaces $\{x_2 < \frac{1}{2}\}$ and $\{x_2 > \frac{1}{2}\}$.

Suppose now that Λ is a spectrum for Ω. Let Π be the set constructed from Λ in Section 6, and $\theta(s)$ be the function on Π given by Lemma 6.2. Let Λ' be the spectrum for Ω constructed from Λ in Section 7, and Λ'' be the spectrum constructed from Λ' in Section 12. We shall continue to use the notations Π_j, θ_j, Λ'_j, Λ''_j and Λ''_∞ with the same meaning as in the previous sections.

Our goal in the present section is to prove that, under the assumptions above, the function $\theta(s)$ is necessarily constant on Π.

14B. It will be convenient to introduce the following notation. Let

$$G := (\mathbb{Z} \times \mathbb{R}) \cup (\mathbb{R} \times \mathbb{Z}) \tag{14-1}$$

and

$$G_0 := (\mathbb{Z} \times \mathbb{R}) \cup (\mathbb{R} \times (\mathbb{Z} \setminus \{0\})). \tag{14-2}$$

Lemma 14.1. *Let* Π_j $(0 \leqslant j < \infty)$ *be one of the components of* Π, *and let* $0 \leqslant k \leqslant \infty$, $k \neq j$. *Then we have*

$$\Lambda_k'' \subset \mathbb{R} \times \bigcap_{s \in \Pi_j} (s + G_0). \tag{14-3}$$

Also, if the set Λ_k'' *is not empty, then we have*

$$\Pi_j \subset \bigcap_{(t,s) \in \Lambda_k''} (s + G_0). \tag{14-4}$$

In fact, each one of (14-3) and (14-4) is just a reformulation of condition (12-8). Hence Lemma 14.1 is a consequence of Lemma 12.3.

14C. Lemma 14.2. *If for some* $0 \leqslant k \leqslant \infty$, *the set* Λ_k'' *is not empty, then*

$$\Lambda_k'' - \Lambda_k'' \not\subset \mathbb{R} \times G. \tag{14-5}$$

Proof. The proof is very similar to that of Corollary 11.5, and therefore it will only be outlined. The proof involves several steps.

Step 1. Let (t_0, s_0) be a point in Λ_k'', and let f be the function defined by (7-4). Then the Fourier expansion

$$f = \frac{1}{|\Omega|} \sum_{\lambda \in \Lambda''} \hat{f}(\lambda) e_\lambda \tag{14-6}$$

of f with respect to the spectrum Λ'' consists only of terms corresponding to $\lambda \in \Lambda_k''$. This follows from Lemma 12.2 and the expression (7-5) for the Fourier transform of f.

Step 2. We have

$$\Lambda_k'' - \Lambda_k'' \not\subset \mathbb{R} \times \mathbb{Z} \times \mathbb{R}. \tag{14-7}$$

Indeed, if this is not true then by translating Λ we may assume $\Lambda_k'' \subset \mathbb{R} \times \mathbb{Z} \times \mathbb{R}$. Hence from the Fourier expansion (14-6) it follows (Lemma 2.1) that f coincides a.e. on Ω with a function $\tilde{f}$ on $\mathbb{R}^3$ which is periodic with respect to the vector $(0, 1, 0)$. As in the proof of Lemma 11.2, this leads to a contradiction to the assumption that $\operatorname{int}(\Omega)$ intersects both half-spaces $\{x_2 < \frac{1}{2}\}$ and $\{x_2 > \frac{1}{2}\}$.

Step 3. We have

$$\Lambda_k'' - \Lambda_k'' \not\subset \mathbb{R} \times \mathbb{R} \times \mathbb{Z}. \tag{14-8}$$

In the same way, if this does not hold then by translating Λ we can assume $\Lambda_k'' \subset \mathbb{R} \times \mathbb{R} \times \mathbb{Z}$. As in Step 2 this implies that f coincides a.e. on Ω with a function $\tilde{f}$ on $\mathbb{R}^3$ which is periodic with respect to the vector $(0, 0, 1)$. As in the proof of Lemma 11.3, this together with Lemma 10.2 implies that Ω must be a prism, a contradiction.

Step 4. We have

$$\Lambda_k'' - \Lambda_k'' \not\subset \mathbb{R} \times G.$$

This follows by combining (14-7), (14-8) and Lemma 11.4. □

14D. Lemma 14.3. *Let s, s', s'' be three points in $\mathbb{R}^2$, and*

$$X = (s+G) \cap (s'+G) \cap (s''+G). \tag{14-9}$$

If the points s, s', s'' are distinct modulo $\mathbb{Z}^2$, then $X - X \subset G$.

This is not difficult to verify, and we omit the details.

Lemma 14.4. *Suppose that there is a component Π_j of the set Π $(0 \leqslant j < \infty)$ such that for any $0 \leqslant k \leqslant \infty$, $k \neq j$, the set Λ_k'' is empty. Then $\Pi = \Pi_j$; namely Π_j is the unique component of Π, and so the function $\theta(s)$ is constant on Π.*

Proof. The assumption means that $\Lambda'' = \Lambda_j''$. By (12-6) we therefore have

$$\Lambda'' \subset (\mathbb{Z} + \theta_j) \times \mathbb{R}^2.$$

Consider the set of all points $s \in \mathbb{R}^2$ for which there is $t \in \mathbb{Z} + \theta_j$ such that $(t, s) \in \Lambda''$. We claim that this set must contain at least three points which are distinct modulo $\mathbb{Z}^2$. Indeed, if this is not true then the spectrum Λ'' is contained in a union of two sets of the form

$$(\mathbb{Z} + \theta_j) \times (\mathbb{Z}^2 + s), \quad s \in \mathbb{R}^2.$$

But this would imply that Λ'' can be covered by the union of two translates of $\mathbb{Z}^3$, which is not possible according to Lemma 13.2 since Ω is not a prism (notice that Σ contains the cube $I \times I$, so we may use Lemma 13.2). Hence there must exist three points (t, s), (t', s'), (t'', s'') in the spectrum Λ'' such that s, s', s'' are distinct modulo $\mathbb{Z}^2$.

Let Π_k, $0 \leqslant k < \infty$, be any one of the components of Π other than Π_j. Then by applying (14-4) (with j, k interchanged) we obtain

$$\Pi_k \subset (s+G) \cap (s'+G) \cap (s''+G).$$

Using Lemma 14.3 this implies $\Pi_k - \Pi_k \subset G$, which is impossible due to Corollary 11.5. It follows that Π_j must be the unique component of Π. This means that $\theta(s) = \theta_j$ for all $s \in \Pi$; thus $\theta(s)$ is constant on Π. □

14E. At this point it will be useful to introduce the following:

Definition 14.5. Let (s_0, s_0') be a pair of points in $\mathbb{R}^2$ such that $s_0' - s_0 \notin G$. If (s_1, s_1') is another pair of points in $\mathbb{R}^2$, then we say that (s_1, s_1') is *dual* to (s_0, s_0') if the following conditions are satisfied:

(i) $s_1 - s_0'$ and $s_1' - s_0$ are both in $\mathbb{Z} \times \mathbb{R}$.

(ii) $s_1 - s_0$ and $s_1' - s_0'$ are both in $\mathbb{R} \times \mathbb{Z}$.

For example, consider the pair (s_0, s_0') given by $s_0 = (0, 0)$, $s_0' = (\alpha, \beta)$, where α, β are two real numbers which are both not in $\mathbb{Z}$. Then the pair (s_1, s_1') given by $s_1 = (\alpha, 0)$, $s_1' = (0, \beta)$ is dual to (s_0, s_0').

It is not difficult to check that the duality relation just defined satisfies the following properties:

1. If (s_1, s_1') is dual to (s_0, s_0') then, since it was assumed that $s_0' - s_0 \notin G$, it follows that also $s_1' - s_1 \notin G$.
2. The duality relation is symmetric; that is, if (s_1, s_1') is dual to (s_0, s_0'), then also (s_0, s_0') is dual to (s_1, s_1').
3. Whether two given pairs are dual to each other or not depends only on the congruence classes of the points modulo $\mathbb{Z}^2$. In other words, if (s_1, s_1') and (s_2, s_2') are two pairs such that $s_2 - s_1$ and $s_2' - s_1'$ are both in $\mathbb{Z}^2$, and if (s_1, s_1') is dual to a certain pair (s_0, s_0'), then also (s_2, s_2') is dual to (s_0, s_0').
4. For every pair (s_0, s_0') such that $s_0' - s_0 \notin G$ there exists a dual pair (s_1, s_1'), and this dual pair is unique modulo $\mathbb{Z}^2$.

The reason for introducing the duality relation above is the following:

Lemma 14.6. *Let (s_0, s_0') be a pair of points in $\mathbb{R}^2$ such that $s_0' - s_0 \notin G$. Then*

$$(s_0 + G) \cap (s_0' + G) = \mathbb{Z}^2 + \{s_1, s_1'\}, \tag{14-10}$$

where (s_1, s_1') is any pair which is dual to (s_0, s_0').

This can be checked easily. It is also easy to see that Lemma 14.6 implies:

Lemma 14.7. *Let (s_0, s_0') and (s_1, s_1') be two pairs of points in $\mathbb{R}^2$ such that $s_0' - s_0$ and $s_1' - s_1$ are both not in G. If the pairs (s_0, s_0') and (s_1, s_1') are not dual to each other, then the set*

$$Y = (s_0 + G) \cap (s_0' + G) \cap (\mathbb{Z}^2 + \{s_1, s_1'\}) \tag{14-11}$$

is contained in a translate of $\mathbb{Z}^2$.

14F. Lemma 14.8. *Suppose that the set Π can be covered by the union of two translates of $\mathbb{Z}^2$. Then the function $\theta(s)$ is constant on Π.*

Proof. By the assumption of the lemma there exist two points $s_0, s_0' \in \mathbb{R}^2$ such that

$$\Pi \subset \mathbb{Z}^2 + \{s_0, s_0'\}. \tag{14-12}$$

Due to (12-1) we have $\Lambda' \subset \mathbb{R} \times \Pi$, and together with (14-12) this implies that Λ' is contained in the set

$$\mathbb{R} \times (\mathbb{Z}^2 + \{s_0, s_0'\}). \tag{14-13}$$

Hence all the sets in (12-3), as well as their weak limit Λ'', are also contained in (14-13).

The set Π has at least one component Π_0. Since by Corollary 11.5 we have $\Pi_0 - \Pi_0 \not\subset G$, we may assume that s_0, s_0' both belong to Π_0 and that $s_0' - s_0 \notin G$. Hence using (14-3) for $j = 0$ we conclude that

$$\Lambda_k'' \subset \mathbb{R} \times ((s_0 + G) \cap (s_0' + G))$$

for every $1 \leqslant k \leqslant \infty$. In turn, by Lemma 14.6 this implies that Λ_k'' is contained in a set of the form

$$\mathbb{R} \times (\mathbb{Z}^2 + \{s_1, s_1'\}), \tag{14-14}$$

where (s_1, s_1') is a pair which is dual to (s_0, s_0').

We conclude that for every $1 \leqslant k \leqslant \infty$, the set Λ_k'' is contained in both (14-13) and (14-14); hence Λ_k'' must be the empty set. Now Lemma 14.4 allows us to deduce that Π_0 is the unique component of Π, and that $\theta(s)$ is a constant function on Π. ☐

Lemma 14.9. *Suppose that one of the components Π_j of Π cannot be covered by the union of two translates of $\mathbb{Z}^2$. Then the function $\theta(s)$ is constant on Π.*

Proof. The assumption means that the component Π_j contains three points s, s', s'' which are distinct modulo $\mathbb{Z}^2$. Hence by Lemma 14.3 the set X defined by (14-9) satisfies $X - X \subset G$. By (14-3), for any $0 \leqslant k \leqslant \infty,\ k \neq j$, we have $\Lambda_k'' \subset \mathbb{R} \times X$, so it follows that

$$\Lambda_k'' - \Lambda_k'' \subset \mathbb{R} \times G.$$

But according to Lemma 14.2 this is possible only if Λ_k'' is empty. We conclude that all the sets Λ_k'' such that $0 \leqslant k \leqslant \infty,\ k \neq j$, are empty. By Lemma 14.4 this implies that Π_j is the unique component of Π, and $\theta(s)$ is constant on Π, as we had to show. ☐

14G. Lemma 14.10. *Suppose the function $\theta(s)$ is not constant on Π. Then there exist two components Π_{j_0} and Π_{j_1} $(j_0 \neq j_1)$ of the set Π, and there are points $s_0, s_0' \in \Pi_{j_0}$ and $s_1, s_1' \in \Pi_{j_1}$ such that*

(i) *Π_{j_0} is contained in the set*

$$X_0 := \mathbb{Z}^2 + \{s_0, s_0'\}, \tag{14-15}$$

while Π_{j_1} is contained in

$$X_1 := \mathbb{Z}^2 + \{s_1, s_1'\}; \tag{14-16}$$

(ii) *$\Lambda_{j_0}'' \subset (\mathbb{Z} + \theta_{j_0}) \times X_0$ and $\Lambda_{j_1}'' \subset (\mathbb{Z} + \theta_{j_1}) \times X_1$;*

(iii) *the two pairs (s_0, s_0') and (s_1, s_1') are dual to each other;*

(iv) *Λ_k'' is empty for every $0 \leqslant k \leqslant \infty,\ k \neq j_1,\ k \neq j_2$.*

Proof. Assume that the function $\theta(s)$ is not constant on Π. Let Π_{j_0} be one of the components of Π. By Corollary 11.5 we have $\Pi_{j_0} - \Pi_{j_0} \not\subset G$; hence there exist two points s_0, s_0' in Π_{j_0} such that $s_0' - s_0 \notin G$. Observe that by Lemma 14.9 the component Π_{j_0} must be contained in the union of two translates of $\mathbb{Z}^2$, which are necessarily given by $\mathbb{Z}^2 + s_0$ and $\mathbb{Z}^2 + s_0'$. That is,

$$\Pi_{j_0} \subset \mathbb{Z}^2 + \{s_0, s_0'\}. \tag{14-17}$$

By Lemma 14.8, the set Π cannot be covered by the union of two translates of $\mathbb{Z}^2$. Hence the set Π must contain some point s_1 which is distinct modulo $\mathbb{Z}^2$ from both s_0 and s_0'. According to (14-17), the new point s_1 cannot belong to Π_{j_0}; hence it belongs to some other component Π_{j_1}.

Using (14-3) it follows that for every $0 \leqslant k \leqslant \infty,\ k \neq j_0,\ k \neq j_1$, we have

$$\Lambda_k'' \subset \mathbb{R} \times \big((s_0 + G) \cap (s_0' + G) \cap (s_1 + G)\big).$$

But then Lemma 14.3 implies that $\Lambda_k'' - \Lambda_k'' \subset \mathbb{R} \times G$. According to Lemma 14.2 this is not possible unless Λ_k'' is empty. We conclude that all the sets Λ_k'', where $0 \leqslant k \leqslant \infty,\ k \neq j_0,\ k \neq j_1$, are empty.

Due to Corollary 11.5, the component Π_{j_1} cannot be contained in the set $\mathbb{Z}^2 + s_1$; hence there is another point s_1' in Π_{j_1} which is not congruent to s_1 modulo $\mathbb{Z}^2$. It then follows from Lemma 14.9 that

$$\Pi_{j_1} \subset \mathbb{Z}^2 + \{s_1, s_1'\}. \tag{14-18}$$

In turns, this implies that we must have $s_1' - s_1 \notin G$, again by Corollary 11.5.

Recalling the definition of the sets Λ''_{j_0} and Λ''_{j_1}, the conditions (14-17) and (14-18) now imply that the property (ii) in the lemma is satisfied.

It remains to show that the pairs (s_0, s_0') and (s_1, s_1') are dual to each other. If this is not the case, then by Lemma 14.7 the set Y defined by (14-11) is contained in a translate of $\mathbb{Z}^2$. But we have $\Lambda''_{j_1} \subset (\mathbb{Z}+\theta_{j_1}) \times Y$, due to (14-3) and property (ii). This implies $\Lambda''_{j_1} - \Lambda''_{j_1} \subset \mathbb{Z} \times \mathbb{Z}^2$, and consequently Λ''_{j_1} must be empty by Lemma 14.2. In a completely similar way we can also deduce that Λ''_{j_0} must be empty. But this yields that all the sets Λ''_k, for every $0 \leqslant k \leqslant \infty$, are empty, which is impossible since Λ'' cannot be empty being a spectrum for Ω. This contradiction confirms that (s_0, s_0') and (s_1, s_1') must be dual to each other, and concludes the proof. □

14H. Lemma 14.11. *The function $\theta(s)$ is necessarily constant on Π.*

Proof. Suppose to the contrary that this is not the case. Then by Lemma 14.10 there are two components Π_{j_0} and Π_{j_1} ($j_0 \neq j_1$) of the set Π, and there are points $s_0, s_0' \in \Pi_{j_0}$ and $s_1, s_1' \in \Pi_{j_1}$ satisfying properties (i)–(iv) of that lemma.

By translating the spectrum Λ by a vector in $\{0\} \times \mathbb{R}^2$ we may assume $s_0 = (0, 0)$. Since $s_0' - s_0 \notin G$, we have $s_0' = (\alpha, \beta)$ for certain real numbers α, β none of which is an integer. Since the pair (s_1, s_1') is dual to (s_0, s_0'), it follows that s_1 and s_1' are congruent modulo $\mathbb{Z}^2$ to the points $(\alpha, 0)$ and $(0, \beta)$ respectively. In other words, we have $s_1 \in \mathbb{Z}^2 + (\alpha, 0)$ and $s_1' \in \mathbb{Z}^2 + (0, \beta)$.

By further translating Λ by a vector in $\mathbb{R} \times \{(0, 0)\}$ we may also assume $\theta_{j_0} = 0$. It will be convenient to denote $\theta := \theta_{j_1}$ (notice that we then have $0 < \theta < 1$, since θ_{j_0} and θ_{j_1} are different numbers).

According to Lemma 13.1, the spectrum Λ'' cannot contain the whole set $\mathbb{Z} \times \{(0, 0)\}$. This implies that by translating Λ once more by some vector in $\mathbb{Z} \times \{(0, 0)\}$ we may additionally assume that Λ'' does not contain the origin $(0, 0, 0)$.

By property (ii) from Lemma 14.10 we have

$$\Lambda''_{j_0} \subset \mathbb{Z} \times \big(\mathbb{Z}^2 + \{(0, 0), (\alpha, \beta)\}\big). \tag{14-19}$$

Hence each point in Λ''_{j_0} belongs to one of two possible types:

1. Points of the form (k, n, m), where k, n, m are integers, not all of which are zero (that k, n, m cannot all be zero follows from the assumption that Λ'' does not contain the origin).
2. Points of the form $(k, n + \alpha, m + \beta)$, where k, n, m are integers.

By the same property (ii) from Lemma 14.10, we also have

$$\Lambda''_{j_1} \subset (\mathbb{Z} + \theta) \times \big(\mathbb{Z}^2 + \{(\alpha, 0), (0, \beta)\}\big). \tag{14-20}$$

Notice that so far, we have always used (14-3) and (14-4) with the set G_0 on the right-hand side actually replaced by G (which is valid since G_0 is a subset of G). However, at this point the fact that G_0, and not just G, appears on the right-hand side of (14-3) will be important. We apply (14-3) with $j = j_0$ and $k = j_1$, and use the assumption that $(0, 0) = s_0 \in \Pi_{j_0}$, to conclude that

$$\Lambda''_{j_1} \subset \mathbb{R} \times G_0. \tag{14-21}$$

It then follows from (14-20) and (14-21) that also each point in Λ''_{j_1} belongs to one of two possible types:

3. Points of the form $(k+\theta, n+\alpha, m)$, where k, n, m are integers, and m is nonzero (that m cannot be zero follows from (14-21) and the fact that α is not an integer).
4. Points of the form $(k+\theta, n, m+\beta)$, where k, n, m are integers.

By property (iv) of Lemma 14.10, the spectrum Λ'' is the union of the two disjoint sets Λ''_{j_0} and Λ''_{j_1}. We conclude that each point of Λ'' belongs to one of the four types 1, 2, 3 and 4 described above.

Now consider the function

$$f(x, y, z) := \mathbb{1}_I(x)\mathbb{1}_I(y)\mathbb{1}_I(z), \quad (x, y, z) \in \mathbb{R}^3,$$

where $I = \left[-\frac{1}{2}, \frac{1}{2}\right]$; namely, f is the indicator function of the unit cube in $\mathbb{R}^3$. Then f is supported by Ω. Consider the Fourier expansion

$$f = \frac{1}{|\Omega|} \sum_{\lambda \in \Lambda''} \hat{f}(\lambda) e_\lambda \tag{14-22}$$

of f with respect to the spectrum Λ''. Since we have

$$\hat{f}(t, u, v) = \hat{\mathbb{1}}_I(t)\hat{\mathbb{1}}_I(u)\hat{\mathbb{1}}_I(v), \quad (t, u, v) \in \mathbb{R}^3,$$

it follows that $\hat{f}(t, u, v) = 0$ whenever at least one of t, u, v is a nonzero integer. This implies that $\hat{f}$ vanishes on all the points of Λ'' which belong to types 1 and 3. Hence only exponentials e_λ such that λ is of type 2 or 4 may have a nonzero coefficient in the expansion (14-22).

It follows (Lemma 2.1) that the right-hand side of (14-22) is a function $\tilde{f}$ of the form

$$\tilde{f}(x, y, z) = e^{2\pi i(\alpha y + \beta z)} g(x, y, z) + e^{2\pi i(\theta x + \beta z)} h(x, y, z), \quad (x, y, z) \in \mathbb{R}^3, \tag{14-23}$$

where g and h are $\mathbb{Z}^3$-periodic functions, and f coincides with $\tilde{f}$ a.e. on Ω. Notice that it follows from (14-23) that the function $|\tilde{f}|$ is periodic with respect to the vector $(0, 0, 1)$. Since we have $|\tilde{f}| = |f| = 1$ a.e. on $I \times I \times I$, the periodicity of $|\tilde{f}|$ implies $|\tilde{f}| = 1$ a.e. on $I \times I \times \mathbb{R}$. Hence $|f| = |\tilde{f}| = 1$ a.e. on the set $\Omega \cap (I \times I \times \mathbb{R})$. On the other hand, by its definition f vanishes on the set

$$\Omega \cap (I \times I \times (\mathbb{R} \setminus I)),$$

so the latter set must have measure zero. We conclude that

$$\Omega \cap (I \times I \times \mathbb{R}) = I \times I \times I. \tag{14-24}$$

Since Ω contains the prism $I \times \Sigma$, and since $\left\{\frac{1}{2}\right\} \times I$ and $\left\{-\frac{1}{2}\right\} \times I$ are facets of Σ, it follows from (14-24) that $\Sigma = I \times I$. Moreover, we obtain that the intersection of Ω and the slab $\mathbb{R} \times I \times \mathbb{R}$ coincides with $I \times \Sigma$. However, by Lemma 10.2 this contradicts our assumption that Ω is not a prism. □

15. Spectral convex polytopes in $\mathbb{R}^3$ tile by translations

Based on the results obtained in the previous sections, we can now deduce:

Theorem 15.1. *Let Ω be a convex polytope in $\mathbb{R}^3$. If Ω is spectral, then it tiles by translations.*

15A. By Theorems 3.1 and 4.1, the polytope Ω must be centrally symmetric and have centrally symmetric facets. Since Theorem 15.1 was already proved in the case when Ω is a prism (Theorem 9.1), it remains to consider the case when Ω is not a prism.

Lemma 15.2. *Let Ω be a convex polytope in $\mathbb{R}^3$, centrally symmetric and with centrally symmetric facets, which is not a prism. If Λ is a spectrum of Ω, then*

$$\langle \Lambda - \Lambda, \tau_F \rangle \subset \mathbb{Z} \tag{15-1}$$

for every facet F of Ω.

This result is the three-dimensional analog of Lemma 8.2. By combining Lemma 15.2 with Corollary 5.3 we immediately obtain that Ω tiles by translations; hence it only remains to prove the lemma.

15B. Lemma 15.2 is a direct consequence of our previous results:

Proof of Lemma 15.2 . Let F be a facet of Ω. We must show that if Λ is a spectrum of Ω, then it satisfies condition (15-1). Since Ω is not a prism, we may use Lemma 10.1 to select a subfacet A of F such that $\mathrm{int}(\Omega)$ intersects each one of the two open half-spaces bounded by the hyperplane $H_{F,A}$.

By applying an affine transformation we may suppose that Ω is in our "standard position"; namely, $\Omega = -\Omega$, $F = \left\{\frac{1}{2}\right\} \times \Sigma$, where Σ is a convex polygon in $\mathbb{R}^2$, $\Sigma = -\Sigma$, and $A = \left\{\frac{1}{2}\right\} \times \left\{\frac{1}{2}\right\} \times I$, where $I = \left[-\frac{1}{2}, \frac{1}{2}\right]$. The hyperplane $H_{F,A}$ is therefore given by $\left\{x_2 = \frac{1}{2}\right\}$, and hence $\mathrm{int}(\Omega)$ intersects both half-spaces $\left\{x_2 < \frac{1}{2}\right\}$ and $\left\{x_2 > \frac{1}{2}\right\}$. We also have $\tau_F = (1, 0, 0)$, so that condition (15-1) becomes

$$\Lambda - \Lambda \subset \mathbb{Z} \times \mathbb{R}^2. \tag{15-2}$$

Let Π be the set constructed from Λ in Section 6, and $\theta(s)$ be the function on Π given by Lemma 6.2. Since all the assumptions of Section 14 are satisfied, we may apply Lemma 14.11, which yields that the function $\theta(s)$ is constant on Π. By Corollary 7.3 this implies that (15-2) holds, which concludes the proof. □

16. Uniqueness of the spectrum

The approach that was used above to prove that in dimensions $d = 2, 3$ any spectral convex polytope Ω can tile by translations also allows us to establish that, except in the case when Ω is a prism, the spectrum is unique up to translation.

16A. To prove this we use the following lemma, which is valid in any dimension d (not just $d = 2, 3$).

Lemma 16.1. *Let $\Omega \subset \mathbb{R}^d$ be a convex polytope, centrally symmetric and with centrally symmetric facets. Suppose that Ω has a spectrum Λ satisfying* (5-2) *for every facet F of Ω. Then Λ is a translate of the lattice T^*, the dual of the lattice T given by* (5-1).

Proof. By Corollary 5.3, the set T given by (5-1) is a lattice, and $\Omega + T$ is a tiling. Hence by Fuglede's theorem the dual lattice T^* is a spectrum for Ω. By translating Λ we may assume that it contains the origin. So (5-2) implies

$$\langle \Lambda, \tau \rangle \subset \mathbb{Z}, \quad \tau \in T.$$

This means that Λ is a subset of T^*. But since no proper subset of a spectrum can also be a spectrum, we must therefore have $\Lambda = T^*$. ☐

From this lemma we immediately obtain the following sufficient condition for a spectral convex polytope to admit a unique spectrum up to translation:

Corollary 16.2. *Let $\Omega \subset \mathbb{R}^d$ be a convex polytope, centrally symmetric and with centrally symmetric facets. Assume that Ω is spectral, and that condition* (5-2) *is satisfied for every spectrum Λ of Ω and every facet F of Ω. Then Ω has a unique spectrum up to translation. More specifically, every spectrum Λ of Ω is a translate of the lattice T^*.*

16B. The criterion just proved can now be applied to the following situations:

Theorem 16.3. *Let Ω be a spectral convex polygon in $\mathbb{R}^2$ which is not a parallelogram. Then Ω admits a unique spectrum up to translation.*

Theorem 16.4. *Let Ω be a spectral convex polytope in $\mathbb{R}^3$ which is not a prism. Then Ω admits a unique spectrum up to translation.*

Indeed, by Theorems 3.1 and 4.1, the polytope Ω must be centrally symmetric and have centrally symmetric facets. Hence Theorem 16.3 follows from Lemma 8.2 and Corollary 16.2, while Theorem 16.4 is a consequence of Lemma 15.2 and Corollary 16.2.

Remark that the assumptions that Ω is not a parallelogram in $\mathbb{R}^2$ and that it is not a prism in $\mathbb{R}^3$ are necessary in these results. Indeed, we have seen in Example 6.5 that if Ω is a prism, then it admits infinitely many non translation-equivalent spectra.

17. Remarks and open problems

17A. It would be interesting to extend Theorem 1.2 to dimensions $d \geqslant 4$.

Problem 17.1. Let Ω be a convex polytope in $\mathbb{R}^d$ $(d \geqslant 4)$. Prove that if Ω is spectral, then it can tile the space by translations.

We know (Theorems 3.1 and 4.1) that such an Ω must be centrally symmetric and have centrally symmetric facets.

Using our previous results, the assertion in Problem 17.1 can be verified for the class of *four-dimensional convex prisms* (the polytopes $\Omega \subset \mathbb{R}^4$ which can be expressed as the Minkowski sum of a three-dimensional convex polytope and a line segment):

Theorem 17.2. *Let Ω be a convex prism in $\mathbb{R}^4$. If Ω is spectral, then it can tile by translations.*

Indeed, this follows from a combination of Theorems 9.2 and 15.1 in the same way as we have deduced Theorem 9.1 from Theorems 8.1 and 9.2.

17B. It is conceivable that Problem 17.1 could be solved in the general case by an appropriate development of our approach. However, there are certain difficulties which should be addressed in extending our proof to higher dimensions.

One problem is to identify the class of polytopes that would play the role of the parallelograms in two dimensions, and of the prisms in three dimensions. The spectral polytopes in these classes do not have a unique spectrum up to translation, and it was therefore necessary to exclude them in Lemmas 8.2 and 15.2, and, for $d = 3$, to prove by a different method that they can tile by translations (Theorem 9.1).

Another problem in higher dimensions might be to obtain an analog of Lemma 12.3. In that lemma we have used the fact that in three dimensions, all the subfacets of Ω are line segments, and hence in particular they are also centrally symmetric. However, a spectral convex polytope Ω in $\mathbb{R}^d$ $(d \geqslant 4)$ need not have centrally symmetric k-dimensional faces for any $2 \leqslant k \leqslant d-2$ (see Section 4A).

The latter problem disappears, though, if we impose the extra assumption that the convex polytope Ω is a *zonotope*. Thus we propose the following restricted version of Problem 17.1.

Problem 17.3. Let Ω be a zonotope in $\mathbb{R}^d$ $(d \geqslant 4)$. Prove that if Ω is spectral, then it tiles by translations.

17C. It would also be interesting to know whether the conclusion of Theorem 1.2 is true for any convex body Ω (not assumed a priori to be a polytope). The paper [Iosevich et al. 2003] contains a proof that, in two dimensions, a spectral convex body Ω must be a polygon. As far as we know, no such a result has been proved in dimensions $d \geqslant 3$.

Problem 17.4. Let Ω be a convex body in $\mathbb{R}^d$. Prove that if Ω is a spectral set, then it must be a polytope.

It is known [Iosevich et al. 2001] that Ω cannot have a smooth boundary. Using the results in [Greenfeld and Lev 2016] it follows that the assertion is also true if Ω is a cylindric convex body whose base has a smooth boundary.

References

[Alexandrov 1933] A. D. Alexandrov, "A theorem on convex polyhedra", *Trudy Mat. Inst. Steklov* **4** (1933), 87. In Russian.

[Alexandrov 2005] A. D. Alexandrov, *Convex polyhedra*, Springer, 2005. MR Zbl

[Fedorov 1885] E. S. Fedorov, "Načala učeniya o figurah", *Zap. Mineral. Obšč.* **21**:2 (1885), 1–279.

[Fuglede 1974] B. Fuglede, "Commuting self-adjoint partial differential operators and a group theoretic problem", *J. Functional Analysis* **16** (1974), 101–121. MR Zbl

[Fuglede 2001] B. Fuglede, "Orthogonal exponentials on the ball", *Expo. Math.* **19**:3 (2001), 267–272. MR Zbl

[Greenfeld and Lev 2016] R. Greenfeld and N. Lev, "Spectrality and tiling by cylindric domains", *J. Funct. Anal.* **271**:10 (2016), 2808–2821. MR Zbl

[Gruber 2007] P. M. Gruber, *Convex and discrete geometry*, Grundlehren der Mathematischen Wissenschaften **336**, Springer, 2007. MR Zbl

[Iosevich et al. 1999] A. Iosevich, N. Katz, and S. Pedersen, "Fourier bases and a distance problem of Erdős", *Math. Res. Lett.* **6**:2 (1999), 251–255. MR Zbl

[Iosevich et al. 2001] A. Iosevich, N. H. Katz, and T. Tao, "Convex bodies with a point of curvature do not have Fourier bases", *Amer. J. Math.* **123**:1 (2001), 115–120. MR Zbl

[Iosevich et al. 2003] A. Iosevich, N. Katz, and T. Tao, "The Fuglede spectral conjecture holds for convex planar domains", *Math. Res. Lett.* **10**:5-6 (2003), 559–569. MR Zbl

[Jorgensen and Pedersen 1999] P. E. T. Jorgensen and S. Pedersen, "Spectral pairs in Cartesian coordinates", *J. Fourier Anal. Appl.* **5**:4 (1999), 285–302. MR Zbl

[Kolountzakis 2000] M. N. Kolountzakis, "Non-symmetric convex domains have no basis of exponentials", *Illinois J. Math.* **44**:3 (2000), 542–550. MR Zbl

[Kolountzakis 2004] M. N. Kolountzakis, "The study of translational tiling with Fourier analysis", pp. 131–187 in *Fourier analysis and convexity*, edited by L. Brandolini et al., Birkhäuser, Boston, 2004. MR Zbl

[Kolountzakis and Matolcsi 2010] M. N. Kolountzakis and M. Matolcsi, "Teselaciones por traslación", *Gac. R. Soc. Mat. Esp.* **13**:4 (2010), 725–746. English version available at https://arxiv.org/abs/1009.3799. MR

[Kolountzakis and Papadimitrakis 2002] M. N. Kolountzakis and M. Papadimitrakis, "A class of non-convex polytopes that admit no orthonormal basis of exponentials", *Illinois J. Math.* **46**:4 (2002), 1227–1232. MR Zbl

[McMullen 1970] P. McMullen, "Polytopes with centrally symmetric faces", *Israel J. Math.* **8** (1970), 194–196. MR Zbl

[McMullen 1971] P. McMullen, "On zonotopes", *Trans. Amer. Math. Soc.* **159** (1971), 91–109. MR Zbl

[McMullen 1980] P. McMullen, "Convex bodies which tile space by translation", *Mathematika* **27**:1 (1980), 113–121. MR Zbl

[McMullen 1981] P. McMullen, "Convex bodies which tile space by translation: acknowledgement of priority", *Mathematika* **28**:2 (1981), 191. MR Zbl

[Schneider 1993] R. Schneider, *Convex bodies: the Brunn–Minkowski theory*, Encyclopedia of Mathematics and its Applications **44**, Cambridge University Press, 1993. MR Zbl

[Tao 2004] T. Tao, "Fuglede's conjecture is false in 5 and higher dimensions", *Math. Res. Lett.* **11**:2-3 (2004), 251–258. MR Zbl

[Venkov 1954] B. A. Venkov, "On a class of Euclidean polyhedra", *Vestnik Leningrad. Univ. Ser. Mat. Fiz. Him.* **9**:2 (1954), 11–31. In Russian. MR

[Young 2001] R. M. Young, *An introduction to nonharmonic Fourier series*, 1st ed., Academic, San Diego, CA, 2001. MR Zbl

Received 5 Mar 2017. Accepted 29 May 2017.

RACHEL GREENFELD: rachelgrinf@gmail.com
Department of Mathematics, Bar-Ilan University, Ramat-Gan 52900, Israel

NIR LEV: levnir@math.biu.ac.il
Department of Mathematics, Bar-Ilan University, Ramat-Gan 52900, Israel

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.org/apde.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in APDE are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

ANALYSIS & PDE

Volume 10 No. 6 2017

2157-5045(2017)10:6;1-S