ANALYSIS & PDE

Volume 18 No. 1 2025

msp

Analysis & PDE

msp.org/apde

Cover image: Eric J. Heller: "Linear Ramp"

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

The subscription price for 2025 is US $475/year for the electronic version, and $735/year (+$70, 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.

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. 18 (2025), No. 1, pp. 1–92
DOI: 10.2140/apde.2025.18.1

MATHEMATICS OF INTERNAL WAVES IN A TWO-DIMENSIONAL AQUARIUM

SEMYON DYATLOV, JIAN WANG AND MACIEJ ZWORSKI

Following theoretical and experimental work of Maas et al. (*Nature* **288**:6642 (1997), 557–561) we consider a linearized model for internal waves in effectively two-dimensional aquaria. We provide a precise description of singular profiles appearing in long-time wave evolution and associate them to classical attractors. That is done by microlocal analysis of the spectral Poincaré problem, leading in particular to a limiting absorption principle. Some aspects of the paper (for instance Section 6) can be considered as a natural microlocal continuation of the work of John (*Amer. J. Math.* **63** (1941), 141–154) on the Dirichlet problem for hyperbolic equations in two dimensions.

1. Introduction

Internal waves are a central topic in oceanography and the theory of rotating fluids — see [Maas 2005; Sibgatullin and Ermanyuk 2019] for reviews and references. They can be described by linear perturbations of the initial state of rest of a stable-stratified fluid (dense fluid lies everywhere below less-dense fluid and the isodensity surfaces are all horizontal). Forcing can take place at linear level by pushing fluid away from this equilibrium state either mechanically, by wind, a piston, a moving boundary, or thermodynamically, by spatially differential heating or evaporation/rain.

The mechanism behind formation of internal waves comes from ray dynamics of the classical system which underlies wave equations — see Section 1.1 for the case of nonlinear ray dynamics relevant to the case we consider. When parameters of the system produce hyperbolic dynamics, attractors are observed in wave evolution — see Figure 1. This phenomenon is both physically and theoretically more accessible in dimension 2. The analysis in the physics literature, see [Maas 2005; Troitskaya 2017], has focused on constructions of standing and propagating waves and did not address the evolution problem analytically. (See, however, [Bajars et al. 2013] for an analysis of a numerical approach to the evolution problem.) In this paper we prove the emergence of singular profiles in the long-time evolution of linear waves for two-dimensional domains.

The model we consider is described as follows. Let $\Omega \subset \mathbb{R}^2 = \{x = (x_1, x_2) : x_j \in \mathbb{R}\}$ be a bounded simply connected open set with C^∞ boundary $\partial\Omega$. Following the fluid mechanics literature we consider the following evolution problem, sometimes referred to as the Poincaré problem:

$$(\partial_t^2 \Delta + \partial_{x_2}^2) u = f(x) \cos \lambda t, \quad u|_{t=0} = \partial_t u|_{t=0} = 0, \quad u|_{\partial\Omega} = 0, \tag{1-1}$$

MSC2020: primary 35G16, 76B55; secondary 35B40, 76M22.
Keywords: internal wave, attractor, stratified fluid, limiting absorption.

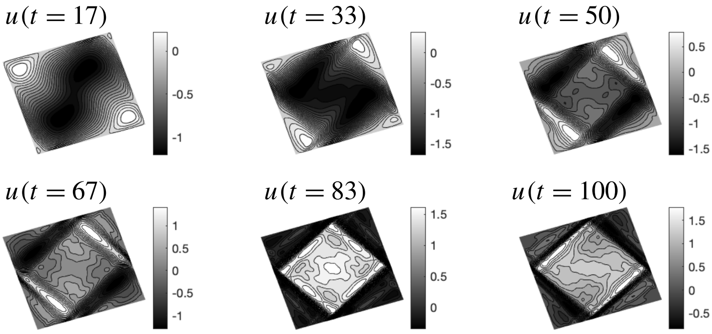

Figure 1. Contour plots of a numerical solution to (1-1) for Ω given by a unit square tilted by $\frac{\pi}{10}$ (see Section 2.5), with $f(x) = e^{-5\pi^2(x-x^0)^2}$, where x^0 is the center and $\lambda = 0.8$. In that case the rotation number of the billiard ball map is $\frac{1}{2}$ (see Figure 9) and the classical attractor is given by a parallelogram on which u develops a singularity — see Theorem 1.3.

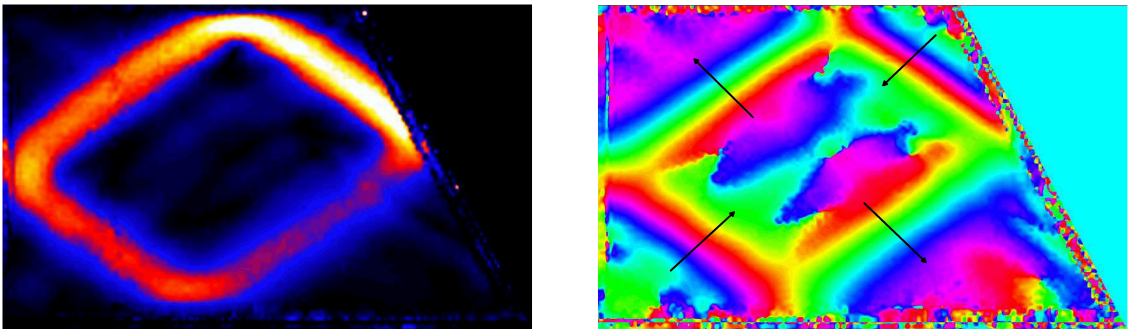

Figure 2. Experimental results of [Hazewinkel et al. 2010]: horizontal component of the observed perturbation buoyancy gradient projected onto a field that oscillates at the forcing frequency, thus reducing the time series to an amplitude field (left) and a phase field (right). In terms of our Theorem 1.3 this corresponds to amplitude and phase of u^+. The arrows indicate directions of phase propagation in agreement with our analysis, shown in Figure 4.

where $\lambda \in (0, 1)$ and $\Delta := \partial_{x_1}^2 + \partial_{x_2}^2$; see [Sobolev 1954, equation (48); Ralston 1973, p. 374; Maas et al. 1997; Brouzet 2016, §1.1.2–3; Dauxois et al. 2018; Colin de Verdière and Saint-Raymond 2020; Sibgatullin and Ermanyuk 2019]. It models internal waves in a stratified fluid in an effectively two-dimensional aquarium Ω with an oscillatory forcing term (here we follow [Colin de Verdière and Saint-Raymond 2020] rather than change the boundary condition). The geometry of Ω and the forcing frequency λ can produce concentration of the fluid velocity $\boldsymbol{v} = (\partial_{x_2} u, -\partial_{x_1} u)$ on attractors. This phenomenon was predicted by Maas and Lam [1995] and was then observed experimentally by Maas, Benielli, Sommeria and Lam [Maas et al. 1997]; see Figure 2 for experimental data from the more recent [Hazewinkel et al. 2010]. (See also the earlier work [Wunsch 1968], which studied the case of an internal wave converging to a corner, along a trajectory of the type pictured in Figure 7.) In this paper we provide a mathematical explanation: as mentioned above, the physics papers concentrated on the analysis of modes and classical dynamics rather than on the long-time behavior of solutions to (1-1).

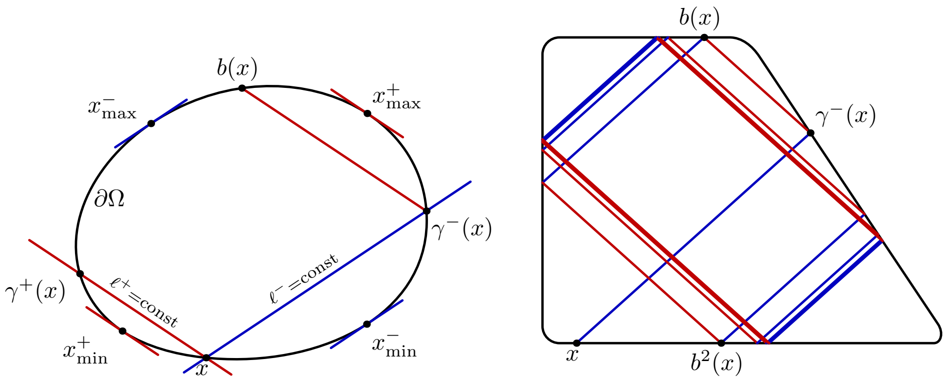

Figure 3. Left: the involutions $\gamma^\pm$ and the chess billiard map b. Right: a forward trajectory of the map b on a trapezium with rounded corners, converging to a periodic trajectory. We remark that the effect of smoothed corner on classical dynamics was investigated by Manders, Duistermaat and Maas [Manders et al. 2003], see also Section 2.4.

1.1. ***Assumptions on Ω and λ.*** The assumptions on Ω and λ which guarantee existence of singular profiles (internal waves) in long-time evolution of (1-1) are formulated using a "chess billiard" — see [Nogueira and Troubetzkoy 2022; Lenci et al. 2023] for recent studies and references. It was first considered in similar context by John [1941] (see also the later work of Aleksandrjan [1960]) and was the basis of the analysis in [Maas and Lam 1995]. It is defined as the reflected bicharacteristic flow for $(1-\lambda^2)\xi_2^2 - \lambda^2\xi_1^2$, which is the Hamiltonian for the $1+1$ wave equation with x_2 corresponding to time and the speed given by $c = \lambda/\sqrt{1-\lambda^2}$ — see Figure 3 and Section 2.1. This flow has a simple reduction to the boundary, which we describe using a factorization of the quadratic form dual to $(1-\lambda^2)\xi_2^2 - \lambda^2\xi_1^2$:

$$-\frac{x_1^2}{\lambda^2} + \frac{x_2^2}{1-\lambda^2} = \ell^+(x,\lambda)\ell^-(x,\lambda), \quad \ell^\pm(x,\lambda) := \pm\frac{x_1}{\lambda} + \frac{x_2}{\sqrt{1-\lambda^2}}. \tag{1-2}$$

We often suppress the dependence on λ, writing simply $\ell^\pm(x)$. Same applies to other λ-dependent objects introduced below.

Definition 1.1. Let $0 < \lambda < 1$. We say that Ω is λ-*simple* if each of the functions $\partial\Omega \ni x \mapsto \ell^\pm(x,\lambda)$ has only two critical points, which are both nondegenerate. We denote these minimum/maximum points by $x^\pm_{\min}(\lambda)$, $x^\pm_{\max}(\lambda)$.

Under the assumption of λ-simplicity we define the following two smooth orientation-reversing involutions on the boundary (see Section 2.1 for more details):

$$\gamma^\pm(\bullet,\lambda) : \partial\Omega \to \partial\Omega, \quad \ell^\pm(x) = \ell^\pm(\gamma^\pm(x)). \tag{1-3}$$

These maps correspond to interchanging intersections of the boundary with lines with slopes $\mp 1/c$, respectively — see Figure 3. The *chess billiard map* $b(\bullet,\lambda)$ is defined as the composition

$$b := \gamma^+ \circ \gamma^- \tag{1-4}$$

and is a C^∞ orientation-preserving diffeomorphism of $\partial\Omega$.

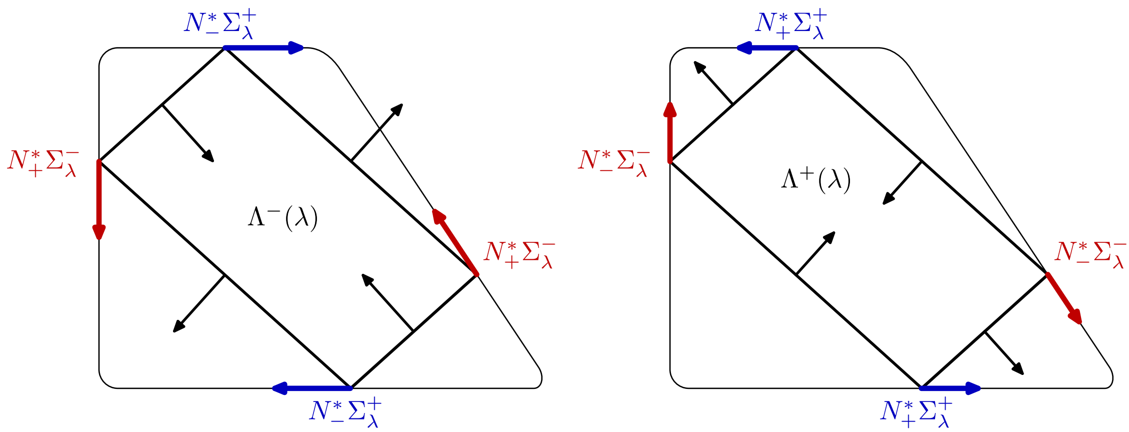

Figure 4. A visualization of the Lagrangian submanifolds (1-9) corresponding to attractive and repulsive cycles of b given in (1-4). The parallelogram represents the projection of the attractive ($+$) and repulsive ($-$) Lagrangians $\Lambda^\pm(\lambda)$ and the arrows perpendicular to the sides represent the conormal directions distinguishing the two Lagrangians. We also indicate the corresponding sets on the boundary: $\Sigma^\pm_\lambda$ are the attractive ($+$) and repulsive ($-$) periodic points of b given by (1-4) and the arrows indicate the sign of the conormal directions.

Denoting by b^n the n-th iterate of b, we consider the set of periodic points

$$\Sigma_\lambda := \{x \in \partial\Omega \mid b^n(x, \lambda) = x \text{ for some } n \geq 1\}. \tag{1-5}$$

If $\Sigma_\lambda \neq \varnothing$, then all the periodic points in Σ_λ have the same minimal period; see Section 2.1.

We are now ready to state the dynamical assumptions on the chess billiard:

Definition 1.2. Let $0 < \lambda < 1$. We say that λ satisfies the *Morse–Smale conditions* if:

(1) Ω is λ-simple.

(2) The map b has periodic points, that is, $\Sigma_\lambda \neq \varnothing$.

(3) The periodic points are hyperbolic, that is, $\partial_x b^n(x, \lambda) \neq 1$ for all $x \in \Sigma_\lambda$, where n is the minimal period.

Under the Morse–Smale conditions we have $\Sigma_\lambda = \Sigma^+_\lambda \sqcup \Sigma^-_\lambda$, where Σ^+_λ, Σ^-_λ are the sets of attractive, respectively repulsive, periodic points of b:

$$\Sigma^+_\lambda := \{x \in \Sigma_\lambda \mid \partial_x b^n(x, \lambda) < 1\}, \quad \Sigma^-_\lambda := \{x \in \Sigma_\lambda \mid \partial_x b^n(x, \lambda) > 1\}. \tag{1-6}$$

Moreover, each of the involutions $\gamma^\pm$ exchanges Σ^+_λ with Σ^-_λ; see (2-2).

For $y \in \partial\Omega$, let

$$\Gamma^\pm_\lambda(y) := \{x \in \Omega \mid \ell^\pm(x, \lambda) = \ell^\pm(y, \lambda)\} \tag{1-7}$$

be the open line segment connecting y with $\gamma^\pm(y, \lambda)$. Define $\Gamma_\lambda(y) := \Gamma^+_\lambda(y) \cup \Gamma^-_\lambda(y)$. Then $\Gamma_\lambda(\Sigma_\lambda)$ gives the closed trajectories of the chess billiard inside Ω.

For $y \in \partial\Omega$ which is not a critical point of ℓ^+, we split the conormal bundle $N^*\Gamma_\lambda^+(y)$ into the positive/negative directions:

$$\begin{aligned} N^*\Gamma_\lambda^+(y) \setminus 0 &= N_+^*\Gamma_\lambda^+(y) \sqcup N_-^*\Gamma_\lambda^+(y), \\ N_\pm^*\Gamma_\lambda^+(y) &:= \{(x, \tau d\ell^+(x)) \mid x \in \Gamma_\lambda^+(y),\ \pm(\partial_\theta \ell^+(y))\tau > 0\} \end{aligned} \tag{1-8}$$

and similarly for $N^*\Gamma_\lambda^-(y)$. Here ∂_θ is the derivative with respect to a positively oriented (that is, counterclockwise when Ω is convex) parametrization of the boundary $\partial\Omega$. Note that the orientation depends on the choice of y and not just on $\Gamma_\lambda^\pm(y)$: we have $N_+^*\Gamma_\lambda^\pm(\gamma^\pm(y)) = N_-^*\Gamma_\lambda^\pm(y)$.

We now define Lagrangian submanifolds $\Lambda^\pm(\lambda) \subset T^*\Omega \setminus 0$ by

$$\Lambda^\pm(\lambda) := N_+^*\Gamma_\lambda^-(\Sigma_\lambda^\pm) \sqcup N_-^*\Gamma_\lambda^+(\Sigma_\lambda^\mp); \tag{1-9}$$

see Figure 4. We note that $\pi(\Lambda^\pm(\lambda)) = \Gamma_\lambda(\Sigma_\lambda)$ and $N_-^*\Gamma_\lambda^\pm(\Sigma_\lambda^+) = N_+^*\Gamma_\lambda^\pm(\Sigma_\lambda^-)$.

1.2. ***Statement of results.*** The main result of this paper is formulated using the concept of *wave front set*; see [Hörmander 1990, §8.1; 1994, Theorem 18.1.27]. The wave front set of a distribution, $\mathrm{WF}(u)$, is a closed subspace of the cotangent bundle of $T^*\Omega \setminus 0$ and it provides phase space information about singularities. Its projection to the base, $\pi(\mathrm{WF}(u))$, is the singular support, $\operatorname{sing\,supp} u$.

Theorem 1.3. *Suppose that Ω and $\lambda \in (0, 1)$ satisfy the Morse–Smale conditions of Definition 1.2. Assume that $f \in C_c^\infty(\Omega; \mathbb{R})$. Then the solution to (1-1) is decomposed as*

$$\begin{aligned} &u(t) = \operatorname{Re}(e^{i\lambda t}u^+) + r(t) + e(t), \quad u^+ \in H^{(1/2)-}(\Omega),\ \mathrm{WF}(u^+) \subset \Lambda^+(\lambda), \\ &r(t) \in H^1(\Omega), \quad \|r(t)\|_{H^1(\Omega)} \le C, \quad \|e(t)\|_{H^{(1/2)-}(\Omega)} \to 0 \quad \text{as } t \to \infty, \end{aligned} \tag{1-10}$$

where $\Lambda^+(\lambda) \subset T^*\Omega \setminus 0$ is the attracting Lagrangian — see (1-9) and Figure 4. In particular, $\operatorname{sing\,supp} u^+$ *is contained in the union of closed orbits of the chess billiard flow. In addition, u^+ is a Lagrangian distribution, $u^+ \in I^{-1}(\bar{\Omega}, \Lambda^+(\lambda))$ (see Section 3.2) and $u^+|_{\partial\Omega} = 0$ (well-defined because of the wave front set condition).*

For a numerical illustration of (1-10), see Figure 1. We remark that numerically it is easier to consider polygonal domains — see Section 2.4 for a discussion of the stability of our assumptions for smoothed out polygonal domains.

Theorem 1.3 is proved using spectral properties of a self-adjoint operator associated to the evolution equation (1-1). To define it, let Δ_Ω be the (negative definite) Dirichlet Laplacian of Ω with the inverse denoted by $\Delta_\Omega^{-1} : H^{-1}(\Omega) \to H_0^1(\Omega)$. Then

$$P := \partial_{x_2}^2 \Delta_\Omega^{-1} : H^{-1}(\Omega) \to H^{-1}(\Omega), \quad \langle u, w\rangle_{H^{-1}(\Omega)} := \langle \nabla\Delta_\Omega^{-1}u, \nabla\Delta_\Omega^{-1}w\rangle_{L^2(\Omega)}, \tag{1-11}$$

is a bounded nonnegative (hence self-adjoint) operator studied in [Aleksandrjan 1960; Ralston 1973] — see Section 7.1. Studying the spectrum of P is referred to as a *Poincaré problem*.

The evolution equation (1-1) is equivalent to

$$(\partial_t^2 + P)w = f\cos\lambda t, \quad w|_{t=0} = \partial_t w|_{t=0} = 0, \quad f \in C_c^\infty(\Omega; \mathbb{R}), \quad u = \Delta_\Omega^{-1}w. \tag{1-12}$$

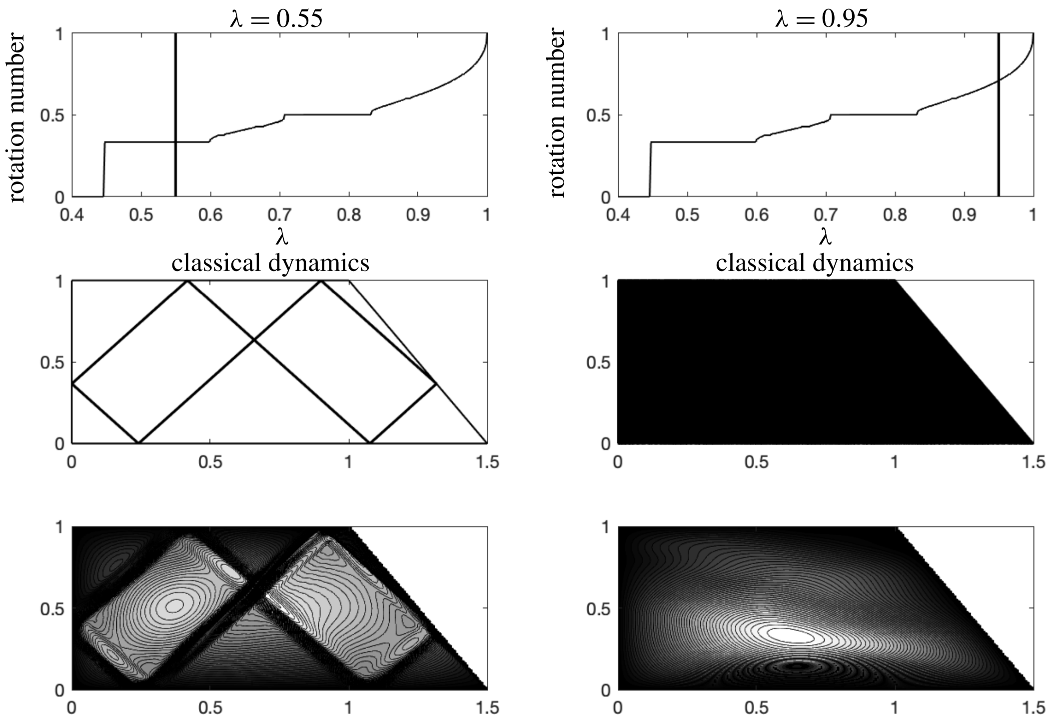

Figure 5. Numerical illustration of Theorem 1.4: contour plots of $|u(x)|$ for $u = (\partial_{x_2}^2 - (\lambda^2 + i\varepsilon)\Delta_\Omega)^{-1} f$, where $\varepsilon = 0.005$ and $f(x) = e^{-10(x-(1/2,1/2))^2}$ and $\Omega = \mathcal{T}_{0.5}$ (see Section 2.5). On the left, the rotation number is given by $\frac{1}{3}$ and we see concentration on an attractor; on the right, the rotation number is (nearly) irrational and, as $\varepsilon \to 0+$, u is expected to be uniformly distributed [Maas 2005]. Morse–Smale assumptions do not hold, at least not on scales relevant to numerical calculations and trajectories are uniformly distributed in the trapezium. In the contour plots of $|u(x)|$ black corresponds to 0.

This equation is easily solved using the functional calculus of P:

$$w(t) = \operatorname{Re}(e^{i\lambda t}\boldsymbol{W}_{t,\lambda}(P)f), \quad \text{where } \boldsymbol{W}_{t,\lambda}(z) = \int_0^t \frac{\sin(s\sqrt{z})}{\sqrt{z}} e^{-i\lambda s}\, ds = \sum_{\pm} \frac{1-e^{-it(\lambda\pm\sqrt{z})}}{2\sqrt{z}(\sqrt{z}\pm\lambda)}. \tag{1-13}$$

Using the Fourier transform of the Heaviside function (see (3-26)), we see that for any $\varphi \in C_c^\infty(\mathbb{R})$ we have

$$\int_{\mathbb{R}} \frac{1-e^{-it\zeta}}{\zeta}\varphi(\zeta)\, d\zeta = i\int_0^t \hat{\varphi}(\eta)\, d\eta \xrightarrow{t\to\infty} i\int_0^\infty \hat{\varphi}(\eta)\, d\eta = \int_{\mathbb{R}} (\zeta - i0)^{-1}\varphi(\zeta)\, d\zeta$$

and thus for any $\lambda \in (0, 1)$ we have the distributional limit

$$\boldsymbol{W}_{t,\lambda}(z) \to (z - \lambda^2 + i0)^{-1} \quad \text{as } t \to \infty, \quad \text{in } \mathcal{D}'_z((0, \infty)). \tag{1-14}$$

This suggests that, as long as we only look at the spectrum of P near λ^2 (the rest of the spectrum contributing the term $r(t)$ in Theorem 1.3), if the spectral measure of P applied to f is smooth in the spectral parameter z, then $\boldsymbol{W}_{t,\lambda}(P)f \to (P-\lambda^2+i0)^{-1}f$ as $t\to\infty$. By Stone's formula, it suffices to establish the *limiting absorption principle* for the operator P near λ^2 and that is the content of the following:

Theorem 1.4. *Suppose that $\mathcal{J}\subset(0,1)$ is an open interval such that each $\lambda\in\mathcal{J}$ satisfies the Morse–Smale conditions of Definition 1.2. Then for each $f\in C_c^\infty(\Omega)$ and $\lambda\in\mathcal{J}$ the limits*

$$(P-\lambda^2\pm i0)^{-1}f=\lim_{\varepsilon\to 0+}(P-(\lambda\mp i\varepsilon)^2)^{-1}f \quad \text{in } \mathcal{D}'(\Omega) \tag{1-15}$$

exist and the spectrum of P is purely absolutely continuous in $\mathcal{J}^2:=\{\lambda^2\mid\lambda\in\mathcal{J}\}$:

$$\sigma(P)\cap\mathcal{J}^2=\sigma_{\mathrm{ac}}(P)\cap\mathcal{J}^2. \tag{1-16}$$

Moreover,

$$(P-\lambda^2\pm i0)^{-1}f\in I^1(\bar{\Omega},\Lambda^\pm(\lambda))\subset H^{-(3/2)-}(\Omega), \tag{1-17}$$

where $\Lambda^\pm(\lambda)$ are given in (1-9) and the definition of the conormal spaces $I^1(\bar{\Omega},\Lambda^\pm(\lambda))$ is reviewed in Section 3.2.

Remarks. (1) The proof provides a more precise statement based on a reduction to the boundary — see Section 7. We also have smooth dependence on λ which plays a crucial role in proving Theorem 1.3 as in [Dyatlov and Zworski 2019b, §5] — see Section 8. This precise information is important in obtaining the $H^{(1/2)-}$ remainder in (1-10). The singular profile in Theorem 1.3 satisfies

$$u^+=\Delta_\Omega^{-1}(P-\lambda^2+i0)^{-1}f,$$

which agrees with the heuristic argument following (1-14).

(2) As noted in [Ralston 1973], $\sigma(P)=[0,1]$ but as emphasized there and in numerous physics papers the structure of the spectrum of P is far from clear. Here we only characterize the spectrum (1-16) under the Morse–Smale assumptions of Definition 1.2.

Rather than working with P, we consider the closely related stationary *Poincaré problem*

$$(\partial_{x_2}^2-\omega^2\Delta)u_\omega=f\in C_c^\infty,\quad u_\omega|_{\partial\Omega}=0,\quad \operatorname{Re}\omega\in(0,1),\quad \operatorname{Im}\omega>0.$$

Then $u_{\lambda+i\varepsilon}\in C^\infty(\bar{\Omega})$ has a limit in $\mathcal{D}'(\Omega)$ which satisfies $u_{\lambda+i0}\in I^{-1}(\bar{\Omega},\Lambda^-(\lambda))$, and we have $(P-\lambda^2-i0)^{-1}f=\Delta u_{\lambda+i0}$.

1.3. ***Related mathematical work.*** Motivated by the study of internal waves, results similar to Theorems 1.3 and 1.4 were obtained for self-adjoint 0th order pseudodifferential operators on two-dimensional tori with dynamical conditions in Definitions 1.1 and 1.2 replaced by demanding that a naturally defined flow is Morse–Smale. That was done first in [Colin de Verdière and Saint-Raymond 2020; Colin de Verdière 2020], with different proofs provided in [Dyatlov and Zworski 2019b]. The question of modes of viscosity limits in such models (addressing physics questions formulated for domains with boundary — see [Rieutord and Valdettaro 2018]) were investigated by Galkowski and Zworski [2022] and Wang [2022]. Finer questions related to spectral theory were also answered in [Wang 2023]. Unlike in the situation considered in this paper, embedded eigenvalues are possible in the case of 0th order pseudodifferential operators [Tao 2024].

The dynamical system (1-4) was recently studied in [Nogueira and Troubetzkoy 2022; Lenci et al. 2023]. We refer to those papers for additional references and dynamical results.

After this paper was accepted for publication we learned of significant contributions to the spectral theory of the operator P (see (1-11)) in the Russian literature. The study of its spectrum was known there as the *Sobolev problem*. We are grateful to Sergey Denisov for pointing this out to us. The most relevant (and translated to English) papers are those of [Fokin 1993; Troitskaya 2017]. The main result announced in Fokin's paper (proved in the longer Russian-language papers cited there) was the existence of singular continuous spectrum for some Ω: any Ω_0 with smooth boundary for which P has H_0^1 eigenfunctions (such as the disk) can be perturbed to obtain Ω with a smooth boundary and nonempty singular continuous spectrum for P. Troitskaya showed that the spectrum for any triangle is continuous. Both papers presented interesting results about long-time behavior of solutions of the Cauchy problem but did not seem to address the questions studied in this paper. These two papers and the papers cited by them contain however a wealth of ideas which may well have applications to our problem.

We should also mention that recently Li [2024] succeeded in providing analogues of Theorems 1.3 and 1.4 in the case of domains with corners. The statements are similar but more complicated as additional, weaker, singularities emanate from the corners.

1.4. ***Organization of the paper.*** In Section 2 we provide a self-contained analysis of the dynamical system given by the diffeomorphism (1-4). We emphasize properties needed in the analysis of the operator (1-11): properties of pushforwards by $\ell^\pm$ and existence of suitable escape/Lyapunov functions. Section 3 is devoted to a review of microlocal analysis used in this paper and in particular to definitions and properties of conormal/Lagrangian spaces used in the formulations of Theorems 1.3 and 1.4. In Section 4 we describe reduction to the boundary using $1+1$ Feynman propagators which arise naturally in the limiting absorption principles. Despite the presence of characteristic points, the restricted operator enjoys good microlocal properties — see Proposition 4.15. Microlocal analysis of that operator is given in Section 5 with the key estimate (5-19) motivated by Lasota–Yorke inequalities and radial estimates. The self-contained Section 6 analyses wave front set properties of distributions invariant under the diffeomorphisms (1-4). These results are combined in Section 7 to give the proof of the limiting absorption principle of Theorem 1.4. Finally, in Section 8 we follow the strategy of [Dyatlov and Zworski 2019b] to describe long-time properties of solutions to (1-1) — see Theorem 1.3.

2. Geometry and dynamics

In this section we assume that $\Omega \subset \mathbb{R}^2$ is an open bounded simply connected set with C^∞ boundary $\partial\Omega$ and review the basic properties of the involutions $\gamma^\pm$ and the chess billiard b defined in (1-3), (1-4). We orient $\partial\Omega$ in the positive direction as the boundary of Ω (that is, counterclockwise if Ω is convex).

2.1. ***Basic properties.*** Fix $\lambda \in (0, 1)$ such that Ω is λ-simple in the sense of Definition 1.1. We first show that the involutions $\gamma^\pm$ defined in (1-3) are smooth. Away from the critical set $\{x^\pm_{\min}, x^\pm_{\max}\}$ this is immediate. Next, we write

$$\begin{aligned}\ell^\pm(x) &= \ell^\pm(x^\pm_{\min}) + \theta^\pm_{\min}(x)^2 \quad \text{for } x \text{ near } x^\pm_{\min},\\ \ell^\pm(x) &= \ell^\pm(x^\pm_{\max}) - \theta^\pm_{\max}(x)^2 \quad \text{for } x \text{ near } x^\pm_{\max},\end{aligned} \tag{2-1}$$

where $\theta^\pm_{\min}, \theta^\pm_{\max}$ are local coordinate functions on $\partial\Omega$ which map $x^\pm_{\min}, x^\pm_{\max}$ to 0. Then for x near $x^\pm_{\min}$ the point $\gamma^\pm(x)$ satisfies the equation

$$\theta^\pm_{\min}(\gamma^\pm(x)) = -\theta^\pm_{\min}(x)$$

and similarly near $x^\pm_{\max}$. This shows the smoothness of $\partial\Omega \ni x \mapsto \gamma^\pm(x)$ near the critical points.

Next, note that since $\gamma^\pm$ are involutions, b is conjugate to its inverse:

$$b^{-1} = \gamma^\pm \circ b \circ \gamma^\pm. \tag{2-2}$$

Therefore $\Sigma^+_\lambda = \gamma^\pm(\Sigma^-_\lambda)$, where $\Sigma^\pm_\lambda$ are defined in (1-6). Since $x^\pm_{\min}, x^\pm_{\max}$ are fixed points of $\gamma^\pm$, the Morse–Smale conditions (see Definition 1.2) imply that there are no characteristic periodic points:

$$\Sigma_\lambda \cap \mathscr{C}_\lambda = \varnothing, \quad \text{where } \mathscr{C}_\lambda := \mathscr{C}^+_\lambda \sqcup \mathscr{C}^-_\lambda, \ \ \mathscr{C}^\pm_\lambda := \{x^\pm_{\min}(\lambda), x^\pm_{\max}(\lambda)\}. \tag{2-3}$$

2.1.1. *Useful identities.* For $x \in \partial\Omega$ and $\lambda \in (0, 1)$ we define the signs

$$\nu^\pm(x, \lambda) := \operatorname{sgn} \partial_\theta \ell^\pm(x, \lambda), \tag{2-4}$$

where ∂_θ is the derivative along $\partial\Omega$ with respect to a positively oriented parametrization.

Lemma 2.1. *Assume that Ω is λ-simple. Then for all $x \in \partial\Omega$*

$$\operatorname{sgn} \ell^\mp(\gamma^\pm(x) - x) = \pm\nu^\pm(x), \tag{2-5}$$

$$\nu^\pm(\gamma^\pm(x)) = -\nu^\pm(x), \tag{2-6}$$

$$\partial_\lambda \ell^\pm(x, \lambda) = \frac{2\lambda^2 - 1}{2\lambda(1-\lambda^2)}\ell^\pm(x, \lambda) + \frac{1}{2\lambda(1-\lambda^2)}\ell^\mp(x, \lambda). \tag{2-7}$$

Proof. To see (2-5), we first notice that it holds when $x \in \{x^\pm_{\min}, x^\pm_{\max}\}$, as then both sides are equal to 0. Now, assume that $\gamma^\pm(x) \neq x$ (that is, $x \notin \{x^\pm_{\min}, x^\pm_{\max}\}$). Denote by $v(x) \in \mathbb{R}^2$ the velocity vector of the parametrization at the point $x \in \partial\Omega$. The vector $\gamma^\pm(x) - x \in \mathbb{R}^2$ is pointing into Ω at the point $x \in \partial\Omega$. Since we use a positively oriented parametrization, the vectors $v(x), \gamma^\pm(x) - x$ form a positively oriented basis. We now note that ℓ^+, ℓ^- form a positively oriented basis of the dual space to $\mathbb{R}^2$, and hence

$$\det\begin{pmatrix}\ell^+(v(x)) & \ell^+(\gamma^\pm(x) - x)\\ \ell^-(v(x)) & \ell^-(\gamma^\pm(x) - x)\end{pmatrix} > 0.$$

Since $\partial_\theta \ell^\pm(x) = \ell^\pm(v(x))$, this gives (2-5). The identity (2-6) follows from (2-5), and (2-7) is verified by a direct computation. □

The next statement is used in the proof of Lemma 4.9.

Lemma 2.2. *Assume that Ω is λ-simple. Then for all $y \in \partial\Omega$ and $x \in \Omega$*

$$\nu^+(y)\ell^-(x - y) > 0 \quad \textit{or} \quad \nu^-(y)\ell^+(x - y) < 0 \quad (\textit{or both}). \tag{2-8}$$

Proof. Let $\Gamma_\lambda^\pm(y)$ be the sets defined in (1-7) and recall that they are open line segments with endpoints $y, \gamma^\pm(y)$. Then by (2-5),

$$\Omega \cap R^\pm(y) = \varnothing, \quad \text{where } R^\pm(y) := \{x \in \mathbb{R}^2 \mid \ell^\pm(x-y) = 0,\ \pm\nu^\pm(y)\ell^\mp(x-y) \le 0\}.$$

The sets $R^\pm(y)$ are closed rays starting at y when $\nu^\pm(y) \neq 0$ and lines passing through y when $\nu^\pm(y) = 0$. Any continuous curve starting at the set of $x \in \mathbb{R}^2$ satisfying (2-8) and ending in the complement of this set has to intersect $R^+(y) \cup R^-(y)$, as can be seen (in the case $\nu^\pm(y) \neq 0$) by applying the intermediate value theorem to the pullback to that curve of the function $x \mapsto \max(\nu^+(y)\ell^-(x-y), -\nu^-(y)\ell^+(x-y))$. Thus, since Ω is connected and contains at least one point x satisfying (2-8) (for instance, take any point in $\Gamma_\lambda^\pm(y)$), all points $x \in \Omega$ satisfy (2-8). □

2.1.2. *Properties of pushforwards.* We next show basic properties of pushforwards of smooth functions by the maps $\partial\Omega \ni x \mapsto \ell^\pm(x, \lambda)$, which are used in the proof of Lemma 4.8. Fix $\lambda \in (0, 1)$ such that Ω is λ-simple and define

$$\ell_{\min}^\pm := \ell^\pm(x_{\min}^\pm), \quad \ell_{\max}^\pm := \ell^\pm(x_{\max}^\pm), \tag{2-9}$$

so that $\ell^\pm$ maps $\partial\Omega$ onto the interval $[\ell_{\min}^\pm, \ell_{\max}^\pm]$. We again fix a positively oriented coordinate θ on $\partial\Omega$.

Lemma 2.3. (1) *Assume that* $f \in C^\infty(\partial\Omega)$ *and define* $\Pi_\lambda^\pm f \in \mathcal{E}'(\mathbb{R})$ *by the formula*

$$\int_{\mathbb{R}} \Pi_\lambda^\pm f(s)\varphi(s)\, ds = \int_{\partial\Omega} f(x)\varphi(\ell^\pm(x))\, d\theta(x) \quad \text{for all } \varphi \in C^\infty(\mathbb{R}). \tag{2-10}$$

Then $\operatorname{supp} \Pi_\lambda^\pm f \subset [\ell_{\min}^\pm, \ell_{\max}^\pm]$ *and*

$$\sqrt{(s-\ell_{\min}^\pm)(\ell_{\max}^\pm - s)}\, \Pi_\lambda^\pm f(s) \in C^\infty([\ell_{\min}^\pm, \ell_{\max}^\pm]). \tag{2-11}$$

(2) *Assume that* $f \in C^\infty(\partial\Omega)$ *and define the functions* $\Upsilon_\lambda^\pm f$ *on* $(\ell_{\min}^\pm, \ell_{\max}^\pm)$ *by*

$$\Upsilon_\lambda^\pm f(s) := \sum_{x \in \partial\Omega,\ \ell^\pm(x) = s} f(x), \quad s \in (\ell_{\min}^\pm, \ell_{\max}^\pm).$$

Then $\Upsilon_\lambda^\pm f \in C^\infty([\ell_{\min}^\pm, \ell_{\max}^\pm])$.

Proof. (1) The support property follows immediately from the definition: if $\operatorname{supp}\varphi \cap [\ell_{\min}^\pm, \ell_{\max}^\pm] = \varnothing$, then $\varphi \circ \ell^\pm = 0$ on $\partial\Omega$ and thus $\int (\Pi_\lambda^\pm f)\varphi = 0$.

To show (2-11), we compute

$$\Pi_\lambda^\pm f(s) = \sum_{x \in \partial\Omega,\ \ell^\pm(x) = s} \frac{f(x)}{|\partial_\theta \ell^\pm(x)|} \quad \text{for all } s \in (\ell_{\min}^\pm, \ell_{\max}^\pm). \tag{2-12}$$

It follows that $\Pi_\lambda^\pm f$ is smooth on the open interval $(\ell_{\min}^\pm, \ell_{\max}^\pm)$. Next, note that (2-11) does not depend on the choice of the parametrization θ since changing the parametrization amounts to multiplying f by a smooth positive function. Thus we can use the local coordinate $\theta = \theta_{\min}^\pm$ near $x_{\min}^\pm$ introduced in (2-1).

With this choice we have $\ell^\pm(x) = \ell^\pm_{\min} + \theta^2$ and the formula (2-12) gives for s near $\ell^\pm_{\min}$

$$\Pi^\pm_\lambda f(s) = \frac{f(\sqrt{s-\ell^\pm_{\min}}) + f(-\sqrt{s-\ell^\pm_{\min}})}{2\sqrt{s-\ell^\pm_{\min}}},$$

where we view f as a function of θ. It follows that $\sqrt{s-\ell^\pm_{\min}}\,\Pi^\pm_\lambda f(s)$ is smooth at the left endpoint of the interval $(\ell^\pm_{\min}, \ell^\pm_{\max})$. Similar analysis shows that $\sqrt{\ell^\pm_{\max}-s}\,\Pi^\pm_\lambda f(s)$ is smooth at the right endpoint of this interval.

(2) This is proved similarly to part (1), where we no longer have $|\partial_\theta \ell^\pm(x)|$ in the denominator in (2-12). □

2.1.3. *Dynamics of the chess billiard.* We now give a description of the dynamics of the orientation-preserving diffeomorphism $b = \gamma^+ \circ \gamma^-$ in the presence of periodic points.

Lemma 2.4. *Assume that* $\Sigma_\lambda \neq \varnothing$ (*see* (1-5)). *Then*:

(1) *All periodic points of b have the same minimal period.*

(2) *For each $x \in \partial\Omega$, the trajectory $b^k(x)$ converges to Σ_λ as $k \to \pm\infty$.*

(3) *If $\partial_x b^n \neq 1$ on Σ_λ, where n denotes the minimal period, then the set Σ_λ is finite.*

Proof. See for example [de Melo and van Strien 1993, §1.1] or [Walsh 1999] for the proof of the first two claims. The last claim follows from the fact that Σ_λ is the set of solutions to $b^n(x) = x$ and thus $\partial_x b^n(x) \neq 0$ implies that it consists of isolated points. □

We finally discuss the rotation number of b. Fix a positively oriented parametrization on $\partial\Omega$ which identifies it with the circle $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ and denote by $\pi : \mathbb{R} \to \partial\Omega$ the covering map. Consider a lift of $b(\bullet, \lambda)$ to $\mathbb{R}$, that is, an orientation-preserving diffeomorphism $\boldsymbol{b}(\bullet, \lambda) : \mathbb{R} \to \mathbb{R}$ such that

$$\pi(\boldsymbol{b}(\theta, \lambda)) = b(\pi(\theta), \lambda) \quad \text{for all } \theta \in \mathbb{R}.$$

Denote by $\boldsymbol{b}^k(\bullet, \lambda)$ the k-th iterate of $\boldsymbol{b}(\bullet, \lambda)$. Define the *rotation number* of $b(\bullet, \lambda)$ as

$$\boldsymbol{r}(\lambda) := \lim_{k\to\infty} \frac{\boldsymbol{b}^k(\theta, \lambda) - \theta}{k} \mod \mathbb{Z} \in \mathbb{R}/\mathbb{Z}. \tag{2-13}$$

The limit exists and is independent of the choice of $\theta \in \mathbb{R}$ and of the lift $\boldsymbol{b}$. We refer to [Walsh 1999] for a proof of this fact as well that of the following.

Lemma 2.5. *The rotation number $\boldsymbol{r}(\lambda)$ is rational if and only if $\Sigma_\lambda \neq \varnothing$. In this case $\boldsymbol{r}(\lambda) = q/n \mod \mathbb{Z}$, where $n > 0$ is the minimal period of the periodic points and $q \in \mathbb{Z}$ is coprime with n.*

We remark that $b(\bullet, \lambda)$ cannot have fixed points: indeed, if $x \in \partial\Omega$ and $b(x) = x$, then $\gamma^+(x) = \gamma^-(x)$, which is impossible. We then fix the lift $\boldsymbol{b}$ for which

$$0 < \boldsymbol{b}(0, \lambda) < 1. \tag{2-14}$$

With this choice we have $0 < \boldsymbol{b}^k(0, \lambda) < k$ for all $k \geq 0$ and thus (2-13) defines the rotation number $\boldsymbol{r}(\lambda)$ which satisfies $0 < \boldsymbol{r}(\lambda) < 1$.

2.2. ***Dependence on*** $\boldsymbol{\lambda}$. We now discuss the dependence of the dynamics of the chess billiard map $b(\bullet, \lambda)$ on λ. We first give a stability result:

Lemma 2.6. *The set of* $\lambda \in (0,1)$ *satisfying the Morse–Smale conditions (see Definition 1.2) is open. Moreover, the maps* $\gamma^\pm(x,\lambda)$ *and* $b(x,\lambda)$, *as well as the sets* Σ_λ, *depend smoothly on* λ *as long as* λ *satisfies the Morse–Smale conditions.*

Proof. Assume that λ_0 satisfies the Morse–Smale conditions. We need to show that all λ close enough to λ_0 satisfy this condition as well. From (1-2) we see that the functions $\ell^\pm(x,\lambda)$ depend smoothly on $x \in \partial\Omega$, $\lambda \in (0,1)$. Therefore, Ω is λ-simple for λ close to λ_0. Moreover, $\gamma^\pm(x,\lambda)$ and $b(x,\lambda)$ depend smoothly on λ as long as Ω is λ-simple.

Next, let $m > 0$ be the number of points in Σ_{λ_0} and let n be their minimal period under $b(\bullet,\lambda_0)$. Since $\partial_x b^n(x,\lambda_0) \neq 1$ on Σ_{λ_0}, by the implicit function theorem for λ close to λ_0 the equation $b^n(x,\lambda) = x$ has exactly m solutions, which depend smoothly on λ. It follows that λ satisfies the Morse–Smale conditions. □

Lemmas 2.5 and 2.6 imply in particular that when λ_0 satisfies the Morse–Smale conditions, the rotation number $\boldsymbol{r}$ is constant in a neighborhood of λ_0: indeed, the rotation number is determined by the combinatorial structure of the map b on each closed orbit (if the rotation number is equal to q/n with q, n coprime, then each closed orbit has period n and the action of b on this orbit is the shift by q points), which varies continuously with λ. A partial converse to this fact is given by the second part of the following.

Lemma 2.7. *Assume that* $\mathcal{J} \subset (0,1)$ *is an open interval such that* Ω *is* λ*-simple for each* $\lambda \in \mathcal{J}$. *Then*:

(1) $\boldsymbol{r}(\lambda)$ *is a continuous increasing function of* $\lambda \in \mathcal{J}$.

(2) *If* $\boldsymbol{r}$ *is constant on* $\mathcal{J}$, *then this constant is rational and the Morse–Smale conditions hold for Lebesgue almost every* $\lambda \in \mathcal{J}$.

Proof. (1) Fix a positively oriented coordinate θ on $\partial\Omega$. Using (1-3), (2-7) we compute

$$\partial_\lambda \gamma^\pm(x,\lambda) = \frac{\partial_\lambda \ell^\pm(x - \gamma^\pm(x,\lambda),\lambda)}{\partial_\theta \ell^\pm(\gamma^\pm(x,\lambda),\lambda)} = \frac{\ell^\mp(x - \gamma^\pm(x,\lambda),\lambda)}{2\lambda(1-\lambda^2)\partial_\theta \ell^\pm(\gamma^\pm(x,\lambda),\lambda)}.$$

By (2-5) and (2-6) we have

$$\partial_\lambda \gamma^+ > 0, \quad \partial_\lambda \gamma^- < 0.$$

We then compute

$$\partial_\lambda b(x,\lambda) = \partial_\lambda \gamma^+(\gamma^-(x,\lambda),\lambda) + \partial_\theta \gamma^+(\gamma^-(x,\lambda),\lambda)\partial_\lambda \gamma^-(x,\lambda).$$

Since $\partial\Omega \ni x \mapsto \gamma^+(x,\lambda) \in \partial\Omega$ is orientation-reversing, this gives

$$\partial_\lambda b(x,\lambda) > 0 \quad \text{for all } x \in \partial\Omega,\ \lambda \in \mathcal{J}. \tag{2-15}$$

Fix the lift $\boldsymbol{b}(\theta,\lambda)$ satisfying (2-14). Then (2-15) gives $\partial_\lambda \boldsymbol{b}(\theta,\lambda) > 0$. This implies that for each two-points $\lambda_1 < \lambda_2$ in $\mathcal{J}$ and every $k \geq 1$

$$\boldsymbol{b}^k(\theta,\lambda_1) < \boldsymbol{b}^k(\theta,\lambda_2).$$

Recalling the definition (2-13) of $\boldsymbol{r}(\lambda)$, we see that $\boldsymbol{r}(\lambda_1) \leq \boldsymbol{r}(\lambda_2)$; that is, $\boldsymbol{r}(\lambda)$ is an increasing function of $\lambda \in \mathcal{J}$.

(2) We now show that $\boldsymbol{r}(\lambda)$ is a continuous function of $\lambda \in \mathcal{J}$. Fix arbitrary $\lambda_0 \in \mathcal{J}$ and $\varepsilon > 0$; since $\boldsymbol{r}$ is an increasing function it suffices to show that there exists $\delta > 0$ such that

$$\boldsymbol{r}(\lambda_0 + \delta) < \boldsymbol{r}(\lambda_0) + \varepsilon, \quad \boldsymbol{r}(\lambda_0 - \delta) > \boldsymbol{r}(\lambda_0) - \varepsilon.$$

We show the first statement, with the second one proved similarly. Choose a rational number $q/n \in (\boldsymbol{r}(\lambda_0), \boldsymbol{r}(\lambda_0) + \varepsilon)$, where $n \in \mathbb{N}$ and $q \in \mathbb{Z}$ are coprime. Since $\boldsymbol{r}(\lambda_0) < q/n$, the definition (2-13) implies that there exists $k_0 > 0$ such that

$$\frac{\boldsymbol{b}^{k_0 n}(0, \lambda_0)}{n k_0} < \frac{q}{n},$$

that is, $\boldsymbol{b}^{k_0 n}(0, \lambda_0) < k_0 q$. Since $\boldsymbol{b}^{k_0 n}(0, \lambda)$ is continuous in λ, we can choose $\delta > 0$ small enough so that

$$\boldsymbol{b}^{k_0 n}(0, \lambda_0 + \delta) < k_0 q. \tag{2-16}$$

By induction on j we see that

$$\boldsymbol{b}^{j k_0 n}(0, \lambda_0 + \delta) < j k_0 q \quad \text{for all } j \geq 1. \tag{2-17}$$

Here the inductive step is proved as follows: using $\boldsymbol{b}^p(r) = \boldsymbol{b}^p(0) + r$, $r \in \mathbb{Z}$ ($\boldsymbol{b}^p$ is a lift of the orientation-preserving diffeomorphism b^p; we dropped $\lambda_0 + \delta$ in the notation),

$$\boldsymbol{b}^{(j+1)k_0 n}(0) = \boldsymbol{b}^{k_0 n}(\boldsymbol{b}^{j k_0 n}(0)) < \boldsymbol{b}^{k_0 n}(j k_0 q) = \boldsymbol{b}^{k_0 n}(0) + j k_0 q < (j+1) k_0 q.$$

Now, the definition (2-13) and (2-17) imply that $\boldsymbol{r}(\lambda_0 + \delta) \leq q/n < \boldsymbol{r}(\lambda_0) + \varepsilon$ as needed.

(3) Assume now that $\boldsymbol{r}$ is constant on $\mathcal{J}$. We first show that this constant is a rational number. Assume the contrary and take an arbitrary $\lambda_0 \in \mathcal{J}$. By (2-15) (shrinking $\mathcal{J}$ slightly if necessary) we may assume that $\partial_\lambda b(x, \lambda) \geq c > 0$ for some $c > 0$ and all $x \in \partial\Omega$, $\lambda \in \mathcal{J}$. Then $\partial_\lambda b^n(x, \lambda) \geq c$ for all $n \geq 0$ as well. Fix $\varepsilon > 0$ such that $\lambda_1 := \lambda_0 + \varepsilon/c$ lies in $\mathcal{J}$. Then $\boldsymbol{b}^n(\theta, \lambda_1) \geq \boldsymbol{b}^n(\theta, \lambda_0) + \varepsilon$ for all $\theta \in \mathbb{R}$.

Fix arbitrary $x_0 = \pi(\theta_0) \in \partial\Omega$, $\theta_0 \in \mathbb{R}$. Since $\boldsymbol{r}(\lambda_0)$ is irrational and $b(\bullet, \lambda_0)$ is smooth, by Denjoy's theorem [de Melo and van Strien 1993, §I.2] every orbit of $b(\bullet, \lambda_0)$ is dense; in particular the orbit $\{b^n(x_0, \lambda_0)\}_{n \geq 1}$ intersects the ε-sized interval on $\partial\Omega$ whose right endpoint is x_0. That is, there exist $n \in \mathbb{N}$, $m \in \mathbb{Z}$ such that

$$\theta_0 + m - \varepsilon \leq \boldsymbol{b}^n(\theta_0, \lambda_0) \leq \theta_0 + m.$$

It follows that

$$\boldsymbol{b}^n(\theta_0, \lambda_0) \leq \theta_0 + m \leq \boldsymbol{b}^n(\theta_0, \lambda_0) + \varepsilon \leq \boldsymbol{b}^n(\theta_0, \lambda_1).$$

By the intermediate value theorem, there exists $\lambda \in [\lambda_0, \lambda_1] \subset \mathcal{J}$ such that $\boldsymbol{b}^n(\theta_0, \lambda) = \theta_0 + m$. Then $x_0 = \pi(\theta_0)$ is a periodic orbit of $b(\bullet, \lambda)$, which contradicts our assumption that $\boldsymbol{r}(\lambda)$ is irrational for all $r \in \mathcal{J}$.

(4) Under the assumption of step (3), we now have $\boldsymbol{r}(\lambda) = q/n \bmod \mathbb{Z}$ for some coprime $q \in \mathbb{Z}$, $n \in \mathbb{N}$ and all $\lambda \in \mathcal{J}$. By Lemma 2.5, for each $\lambda \in \mathcal{J}$ the set of periodic points Σ_λ is nonempty and each such point has minimal period n. Define

$$\Sigma_{\mathcal{J}} := \{(x, \lambda) \mid \lambda \in \mathcal{J},\ x \in \Sigma_\lambda\} = \{(x, \lambda) \in \partial\Omega \times \mathcal{J} \mid b^n(x, \lambda) = x\}.$$

From (2-15) we see that $\partial_\lambda b^n(x,\lambda) > 0$ for all $x \in \partial\Omega$, $\lambda \in \mathcal{J}$. Shrinking $\mathcal{J}$ if needed, we may assume that $\Sigma_{\mathcal{J}}$ is a one-dimensional submanifold of $\mathcal{J} \times \partial\Omega$ projecting diffeomorphically onto the x-variable, that is, $\Sigma_{\mathcal{J}} = \{(x, \psi(x)) \mid x \in U\}$ for some open set $U \subset \partial\Omega$ and smooth function $\psi : U \to \mathcal{J}$, $\partial_x \psi(x) = (1 - \partial_x b^n(x,\lambda))/\partial_\lambda b^n(x,\lambda)$, $\lambda = \psi(x)$. Then $\lambda \in \mathcal{J}$ satisfies the Morse–Smale conditions if and only if λ is a regular value of ψ, which by the Morse–Sard theorem happens for Lebesgue almost every $\lambda \in \mathcal{J}$. □

2.3. *Escape functions.* We now construct an adapted parametrization of $\partial\Omega$ and a family of escape functions, which are used in Section 5 below. Throughout this section we assume that $\lambda \in (0,1)$ satisfies the Morse–Smale conditions of Definition 1.2. Recall the sets $\Sigma_\lambda^\pm$ of attractive/repulsive periodic points of the map $b(\bullet, \lambda)$ defined in (1-6). Let $n \in \mathbb{N}$ be the minimal period of the corresponding trajectories of b.

We first construct a parametrization of $\partial\Omega$ with a bound on $\partial_x b|_{\Sigma_\lambda^\pm}$ rather than on the derivative of the n-th iterate $\partial_x b^n|_{\Sigma_\lambda^\pm}$:

Lemma 2.8. *Let $\Sigma_\lambda^\pm$ be given by (1-6). There exists a positively oriented coordinate $\theta : \partial\Omega \to \mathbb{S}^1$ such that, taking derivatives on $\partial\Omega$ with respect to θ,*

$$\begin{aligned} \partial_x b(x,\lambda) &< 1 \quad \text{for all } x \in \Sigma_\lambda^+, \\ \partial_x b(x,\lambda) &> 1 \quad \text{for all } x \in \Sigma_\lambda^-. \end{aligned} \tag{2-18}$$

Proof. Fix any Riemannian metric g_0 on $\partial\Omega$ and consider the metric g on $\partial\Omega$ given by

$$|v|_{g(x)} := \sum_{j=0}^{n-1} |\partial_x b^j(x) v|_{g_0(b^j(x))} \quad \text{for all } (x,v) \in T(\partial\Omega).$$

We have for all $(x,v) \in T(\partial\Omega)$

$$|\partial_x b(x) v|_{g(b(x))} - |v|_{g(x)} = |\partial_x b^n(x) v|_{g_0(b^n(x))} - |v|_{g_0(x)}.$$

Thus by (1-6) we have for $v \neq 0$

$$\begin{aligned} |\partial_x b(x) v|_{g(b(x))} &< |v|_{g(x)}, \quad \text{when } x \in \Sigma_\lambda^+, \\ |\partial_x b(x) v|_{g(b(x))} &> |v|_{g(x)}, \quad \text{when } x \in \Sigma_\lambda^-. \end{aligned}$$

It remains to choose the coordinate θ so that $|\partial_\theta|_g$ is constant. □

We next use the global dynamics of $b(\bullet, \lambda)$ described in Lemma 2.4 to construct an *escape function* in Lemma 2.9 below. Fix a parametrization on $\partial\Omega$ which satisfies (2-18) and denote by

$$\Sigma_\lambda^\pm(\delta) \subset \partial\Omega$$

the open δ-neighborhoods of the sets $\Sigma_\lambda^\pm$ with respect to this parametrization. Here $\delta > 0$ is a constant small enough so that the closures $\overline{\Sigma_\lambda^+(\delta)}$ and $\overline{\Sigma_\lambda^-(\delta)}$ do not intersect each other. We also choose δ small enough so that

$$b(\overline{\Sigma_\lambda^+(\delta)}) \subset \Sigma_\lambda^+(\delta), \quad b^{-1}(\overline{\Sigma_\lambda^-(\delta)}) \subset \Sigma_\lambda^-(\delta); \tag{2-19}$$

this is possible by (2-18) and since $\Sigma_\lambda^\pm$ are b-invariant.

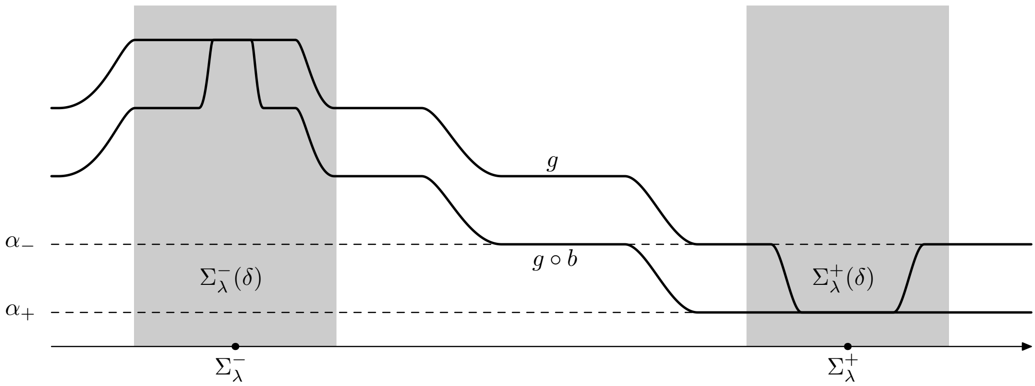

Figure 6. The escape function g constructed in Lemma 2.9 and the function $g \circ b$, where for simplicity we replace $\Sigma_\lambda^\pm$ by fixed points of the map b. The shaded regions correspond to $\Sigma_\lambda^\pm(\delta)$ and the dashed lines correspond to $\alpha_\pm$.

Lemma 2.9. *Let $\alpha_+ < \alpha_-$ be two real numbers. Then there exists a function $g \in C^\infty(\partial\Omega; \mathbb{R})$ such that*:

(1) $g(b(x)) \le g(x)$ *for all* $x \in \partial\Omega$.

(2) $g(b(x)) < g(x)$ *for all* $x \in \partial\Omega \setminus (\Sigma_\lambda^+(\delta) \cup \Sigma_\lambda^-(\delta))$.

(3) $g(x) \ge \alpha_+$ *for all* $x \in \partial\Omega$.

(4) $g(x) \ge \alpha_-$ *for all* $x \in \partial\Omega \setminus \Sigma_\lambda^+(\delta)$.

(5) $g = \alpha_+$ *on some neighborhood of* Σ_λ^+.

(6) *For* $M \gg 1$, *we have* $M(g(b(x)) - g(x)) + g(x) \le \alpha_+$ *for all* $x \in \partial\Omega \setminus \Sigma_\lambda^-(\delta)$.

See Figure 6.

Remark. We note that the same construction works for b^{-1} with the roles of $\Sigma_\lambda^\pm$ reversed. Hence for any real numbers $\alpha_- < \alpha_+$ we can find $g \in C^\infty(\partial\Omega; \mathbb{R})$ such that:

(1) $g(x) \le g(b(x))$ for all $x \in \partial\Omega$.

(2) $g(x) < g(b(x))$ for all $x \in \partial\Omega \setminus (\Sigma_\lambda^+(\delta) \cup \Sigma_\lambda^-(\delta))$.

(3) $g(x) \ge \alpha_-$ for all $x \in \partial\Omega$.

(4) $g(x) \ge \alpha_+$ for all $x \in \partial\Omega \setminus \Sigma_\lambda^-(\delta)$.

(5) $g = \alpha_-$ on some neighborhood of Σ_λ^-.

(6) For $M \gg 1$, we have $M(g(x) - g(b(x))) + g(b(x)) \le \alpha_-$ for all $x \in \partial\Omega \setminus \Sigma_\lambda^+(\delta)$.

Proof. In view of (2-19) there exists $0 < \delta_1 < \delta$ such that

$$b(\Sigma_\lambda^+(\delta)) \subset \Sigma_\lambda^+(\delta_1). \tag{2-20}$$

(1) We first show that there exists $N \ge 0$ such that

$$b^N(\partial\Omega \setminus \Sigma_\lambda^-(\delta)) \subset \Sigma_\lambda^+(\delta_1). \tag{2-21}$$

We argue by contradiction. Assume that (2-21) does not hold for any N. Then there exist sequences

$$x_j \in \partial\Omega \setminus \Sigma_\lambda^-(\delta), \quad m_j \to \infty, \quad b^{m_j}(x_j) \notin \Sigma_\lambda^+(\delta_1). \tag{2-22}$$

Passing to subsequences, we may assume that $x_j \to x_\infty$ for some $x_\infty \in \partial\Omega$. Since $x_j \notin \Sigma_\lambda^-(\delta)$, we have $x_\infty \notin \Sigma_\lambda^-(\delta)$ as well. Then by (2-19) the trajectory $b^k(x_\infty)$, $k \geq 0$, does not intersect $\Sigma_\lambda^-(\delta)$. On the other hand, by Lemma 2.4 this trajectory converges to $\Sigma_\lambda = \Sigma_\lambda^- \sqcup \Sigma_\lambda^+$ as $k \to \infty$. Thus this trajectory converges to Σ_λ^+; in particular

$$\text{there exists } k \geq 0 \text{ such that } b^k(x_\infty) \in \Sigma_\lambda^+(\delta_1).$$

Since $x_j \to x_\infty$, we have $b^k(x_j) \to b^k(x_\infty)$ as $j \to \infty$. Since $\Sigma_\lambda^+(\delta_1)$ is an open set, there exists $j \geq 0$ such that $m_j \geq k$ and $b^k(x_j) \in \Sigma_\lambda^+(\delta_1)$. But then by (2-20) we have $b^{m_j}(x_j) \in \Sigma_\lambda^+(\delta_1)$ which contradicts (2-22).

(2) Choose N such that (2-21) holds and fix a cutoff function

$$\chi_+ \in C_c^\infty(\Sigma_\lambda^+(\delta); [0, 1]), \qquad \chi_+ = 1 \quad \text{on } \Sigma_\lambda^+(\delta_1).$$

Define the function $\tilde{g} \in C^\infty(\partial\Omega; \mathbb{R})$ as an ergodic average of χ_+:

$$\tilde{g}(x) := \frac{1}{N} \sum_{j=0}^{N-1} \chi_+(b^j(x)) \quad \text{for all } x \in \partial\Omega.$$

It follows from the definition and (2-20) that

$$\begin{aligned} 0 \leq \tilde{g}(x) \leq 1 &\quad \text{for all } x \in \partial\Omega, \\ \tilde{g}(x) = 1 &\quad \text{for all } x \in \Sigma_\lambda^+(\delta_1), \\ \tilde{g}(x) \leq 1 - \frac{1}{N} &\quad \text{for all } x \in \partial\Omega \setminus \Sigma_\lambda^+(\delta). \end{aligned} \tag{2-23}$$

Next, we compute

$$\tilde{g}(b(x)) - \tilde{g}(x) = \frac{1}{N}(\chi_+(b^N(x)) - \chi_+(x)).$$

It follows that

$$\begin{aligned} \tilde{g}(b(x)) \geq \tilde{g}(x) &\quad \text{for all } x \in \partial\Omega, \\ \tilde{g}(b(x)) = \tilde{g}(x) + \frac{1}{N} &\quad \text{for all } x \in \partial\Omega \setminus (\Sigma_\lambda^+(\delta) \cup \Sigma_\lambda^-(\delta)). \end{aligned} \tag{2-24}$$

Indeed, take arbitrary $x \in \partial\Omega$. We have $\chi_+(x) = 0$ unless $x \in \Sigma_\lambda^+(\delta)$. By (2-21), we have $\chi_+(b^N(x)) = 1$ unless $x \in \Sigma_\lambda^-(\delta)$. Recalling that $0 \leq \chi_+ \leq 1$ and $\Sigma_\lambda^+(\delta) \cap \Sigma_\lambda^-(\delta) = \varnothing$, we get (2-24).

(3) Now put

$$g(x) := N\alpha_- - (N-1)\alpha_+ - N(\alpha_- - \alpha_+)\tilde{g}(x). \tag{2-25}$$

Using (2-23) and (2-24), we see that the function g satisfies the first five properties, with the following quantitative versions of parts (2) and (5):

$$\begin{aligned} g(b(x)) - g(x) = \alpha_+ - \alpha_- < 0 &\quad \text{for all } x \in \partial\Omega \setminus (\Sigma_\lambda^+(\delta) \cup \Sigma_\lambda^-(\delta)), \\ g(x) = \alpha_+ &\quad \text{for all } x \in \Sigma_\lambda^+(\delta_1). \end{aligned} \tag{2-26}$$

To prove part (6) we first use (2-25) and (2-26) to see that, for all $M \geq N$ and $x \in \partial\Omega \setminus (\Sigma_\lambda^+(\delta) \cup \Sigma_\lambda^-(\delta))$,

$$M(g(b(x)) - g(x)) + g(x) \leq \alpha_+. \tag{2-27}$$

To establish (2-27) for $x \in \Sigma_\lambda^+(\delta)$ we use (2-20) and the fact that $g|_{\Sigma_\lambda^+(\delta_1)} = \alpha_+$ by (2-26). Then, for $M \geq 1$ and $x \in \Sigma_\lambda^+(\delta)$, property (1) gives

$$M(g(b(x)) - g(x)) + g(x) \leq g(b(x)) = \alpha_+,$$

which completes the proof of the lemma. □

Remark. We discuss here the dependence of the objects in this section on the parameter λ. The parametrization θ constructed in Lemma 2.8 depends smoothly on λ as follows immediately from its construction (recalling from the proof of Lemma 2.6 that the period n is locally constant in λ). Next, for each $\lambda_0 \in (0, 1)$ satisfying the Morse–Smale conditions there exists a neighborhood $U(\lambda_0)$ such that we can construct a function $g(x, \lambda)$ for each $\lambda \in U(\lambda_0)$ satisfying the conclusions of Lemma 2.9 in such a way that it is smooth in λ. Indeed, the sets $\Sigma_\lambda^\pm$ depend smoothly on λ by Lemma 2.6, so the cutoff function χ_+ can be chosen λ-independent. The function $g(x, \lambda)$ is constructed explicitly using this cutoff, the map $b(\bullet, \lambda)$, and the number N. The latter can be chosen λ-independent as well: if (2-21) holds for some λ, then it holds with the same N and all nearby λ.

2.4. *Domains with corners.* We now discuss the case when the boundary of $\partial\Omega$ has corners. This includes the situation when $\partial\Omega$ is a convex polygon, which is the setting of the experiments. Our results do not apply to such domains; however, they apply to appropriate "roundings" of these domains described below.

We first define domains with corners. Let $\Omega \subset \mathbb{R}^2$ be an open set of the form

$$\Omega = \{x \in \mathbb{R}^2 \mid F_1(x) > 0, \dots, F_k(x) > 0\},$$

where $F_1, \dots, F_k : \mathbb{R}^2 \to \mathbb{R}$ are C^∞ functions such that:

(1) The set $\overline{\Omega} := \{F_1 \geq 0, \dots, F_k \geq 0\}$ is compact and simply connected.

(2) For each $x \in \overline{\Omega}$, at most two of the functions $F_1, \dots, F_k$ vanish at x.

If only one of the functions $F_1, \dots, F_k$ vanishes at $x \in \overline{\Omega}$, then we call x a *regular point* of the boundary $\partial\Omega := \overline{\Omega} \setminus \Omega$. If two of the functions $F_1, \dots, F_k$ vanish at $x \in \overline{\Omega}$, then we call x a *corner* of Ω. We make the following natural nondegeneracy assumptions:

(3) If $x \in \partial\Omega$ is a regular point and $F_j(x) = 0$, then $dF_j(x) \neq 0$.

(4) If $x \in \partial\Omega$ is a corner and $F_j(x) = F_{j'}(x) = 0$, where $j \neq j'$, then $dF_j(x)$, $dF_{j'}(x)$ are linearly independent.

We call Ω a *domain with corners* if it satisfies the assumptions (1)–(4) above.

Since Ω is simply connected, the boundary $\partial\Omega$ is a Lipschitz continuous piecewise smooth curve. We parametrize $\partial\Omega$ in the positively oriented direction by a Lipschitz continuous map

$$\theta \in \mathbb{S}^1 := \mathbb{R}/\mathbb{Z} \;\mapsto\; \boldsymbol{x}(\theta) \in \partial\Omega \subset \mathbb{R}^2, \tag{2-28}$$

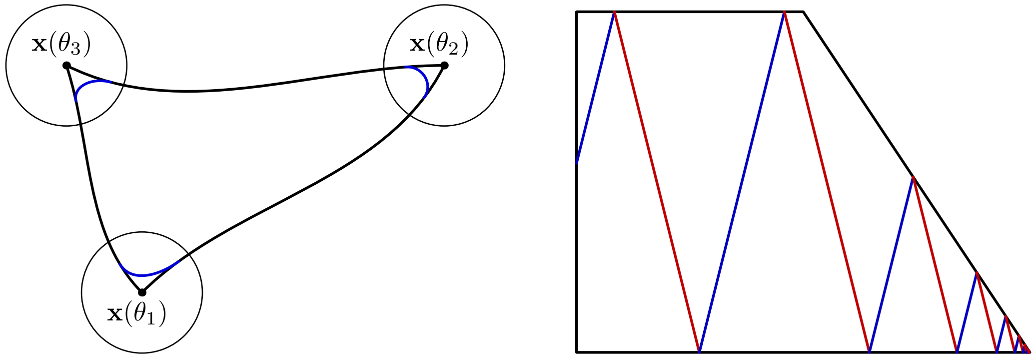

Figure 7. Left: a domain with corners and its ε-rounding (in blue). The circles have radius ε. Right: a trajectory on a trapezium which converges to a corner.

where the corners are given by $\boldsymbol{x}(\theta_j)$ for some $\theta_1 < \cdots < \theta_m$ and the map (2-28) is smooth on each interval $[\theta_j, \theta_{j+1}]$. See Figure 7.

We next extend the concept of λ-simplicity to domains with corners. Let $\ell \in C^\infty(\mathbb{R}^2; \mathbb{R})$ and $x = \boldsymbol{x}(\theta_j)$ be a corner of Ω. Consider the one-sided derivatives $\partial_\theta(\ell \circ \boldsymbol{x})(\theta_j \pm 0)$. There are three possible cases:

(1) Both derivatives are nonzero and have the same sign — then we call x *not a critical point* of ℓ.

(2) Both derivatives are nonzero and have opposite signs — then we call x a *nondegenerate critical point* of ℓ.

(3) At least one of the derivatives is zero — then we call x a *degenerate critical point* of ℓ.

If $x = \boldsymbol{x}(\theta)$ is instead a regular point of the boundary, then we use the standard definition of critical points: x is a critical point of ℓ if $\partial_\theta(\ell \circ \boldsymbol{x})(\theta) = 0$, and a critical point is nondegenerate if $\partial_\theta^2(\ell \circ \boldsymbol{x})(\theta) \neq 0$. With the above convention for critical points, we follow Definition 1.1: we say that a domain with corners Ω is λ-*simple* if each of the functions $\ell^\pm(\bullet, \lambda)$ defined in (1-2) has exactly two critical points on $\partial\Omega$, which are both nondegenerate.

If Ω is λ-simple, then the involutions $\gamma^\pm(\bullet, \lambda) : \partial\Omega \to \partial\Omega$ from (1-3) are well-defined and Lipschitz continuous. Thus $b = \gamma^+ \circ \gamma^-$ is an orientation-preserving bi-Lipschitz homeomorphism of $\partial\Omega$. We now revise the Morse–Smale conditions of Definition 1.2 as follows:

Definition 2.10. Let Ω be a domain with corners. We say that $\lambda \in (0, 1)$ satisfies the *Morse–Smale conditions* if:

(1) Ω is λ-simple.

(2) The set Σ_λ of periodic points of the map $b(\bullet, \lambda)$ is nonempty.

(3) The set Σ_λ does not contain any corners of Ω.

(4) For each $x \in \Sigma_\lambda$, we have $\partial_x b^n(x, \lambda) \neq 1$, where n is the minimal period.

The new condition (3) in Definition 2.10 ensures that b is smooth near the $\gamma^\pm$-invariant set Σ_λ, so condition (4) makes sense. Without this condition we could have trajectories of b converging to a corner; see Figure 7.

We finally show that if Ω is a domain with corners satisfying the Morse–Smale conditions of Definition 2.10 then an appropriate "rounding" of Ω satisfies the Morse–Smale conditions of Definition 1.2:

Proposition 2.11. *Let Ω be a domain with corners and $\lambda \in (0, 1)$ satisfy the Morse–Smale conditions for Ω. Then there exists $\varepsilon > 0$ such that for any open simply connected $\widehat{\Omega} \subset \mathbb{R}^2$ with C^∞ boundary and such that*

- *$\widehat{\Omega}$ is an **ε-rounding** of Ω in the sense that, for each $x \in \mathbb{R}^2$ which lies distance $\geq \varepsilon$ from all the corners of Ω, we have $x \in \Omega$ if and only if $x \in \widehat{\Omega}$,*
- *the domain $\widehat{\Omega}$ is λ-simple in the sense of Definition 1.1, and*

the Morse–Smale conditions are satisfied for λ and $\widehat{\Omega}$.

Proof. Fix a parametrization $\boldsymbol{x}(\theta)$ of $\partial\Omega$ as in (2-28). Take a parametrization

$$\theta \in \mathbb{S}^1 \quad \mapsto \quad \hat{\boldsymbol{x}}(\theta) \in \partial\widehat{\Omega},$$

which coincides with $\boldsymbol{x}(\theta)$ except ε-close to the corners:

$$\hat{\boldsymbol{x}}(\theta) = \boldsymbol{x}(\theta) \quad \text{for all } \theta \notin \bigcup_{j=1}^{m} I_j(\varepsilon), \qquad I_j(\varepsilon) := [\theta_j - C\varepsilon, \theta_j + C\varepsilon]. \tag{2-29}$$

Here C denotes a constant depending on Ω and the parametrization $\boldsymbol{x}(\theta)$, but not on $\widehat{\Omega}$ or ε, whose precise value might change from place to place in the proof.

Denote by $\gamma^\pm$, $\hat{\gamma}^\pm$ the involutions (1-3) corresponding to Ω, $\widehat{\Omega}$, and consider them as homeomorphisms of $\mathbb{S}^1$ using the parametrizations $x, \hat{x}$. Then by (2-29)

$$\gamma^\pm(\theta) = \hat{\gamma}^\pm(\theta) \quad \text{if } \theta, \gamma^\pm(\theta) \notin \bigcup_{j=1}^{m} I_j(\varepsilon). \tag{2-30}$$

Let $b = \gamma^+ \circ \gamma^-$, $\hat{b} = \hat{\gamma}^+ \circ \hat{\gamma}^-$ be the chess billiard maps of Ω, $\widehat{\Omega}$ and let Σ_λ, $\widehat{\Sigma}_\lambda$ be the corresponding sets of periodic trajectories. Choose $\varepsilon > 0$ such that the intervals $I_j(\varepsilon)$ do not intersect Σ_λ; this is possible since Σ_λ does not contain any corners of Ω. Since Σ_λ is invariant under $\gamma^\pm$, we see from (2-30) that $b = \hat{b}$ in a neighborhood of Σ_λ and thus $\Sigma_\lambda \subset \widehat{\Sigma}_\lambda$. That is, the periodic points for the original domain Ω are also periodic points for the rounded domain $\widehat{\Omega}$, with the same period n. It also follows that $\partial_x \hat{b}^n(x, \lambda) = \partial_x b^n(x, \lambda) \neq 1$ for all $x \in \Sigma_\lambda$.

It remains to show that $\widehat{\Sigma}_\lambda \subset \Sigma_\lambda$, that is the rounding does not create any new periodic points for $\hat{b}$. Note that all periodic points have the same period n, and it is enough to show that

$$\hat{b}^n(\theta) \neq \theta \quad \text{for all } \theta \in \bigcup_{j=1}^{m} I_j(\varepsilon). \tag{2-31}$$

From (2-30), the monotonicity of $\gamma^\pm$, $\hat{\gamma}^\pm$, and the Lipschitz continuity of $\gamma^\pm$ we have

$$|\gamma^\pm(\theta) - \hat{\gamma}^\pm(\theta)| \leq C\varepsilon \quad \text{for all } \theta \in \mathbb{S}^1.$$

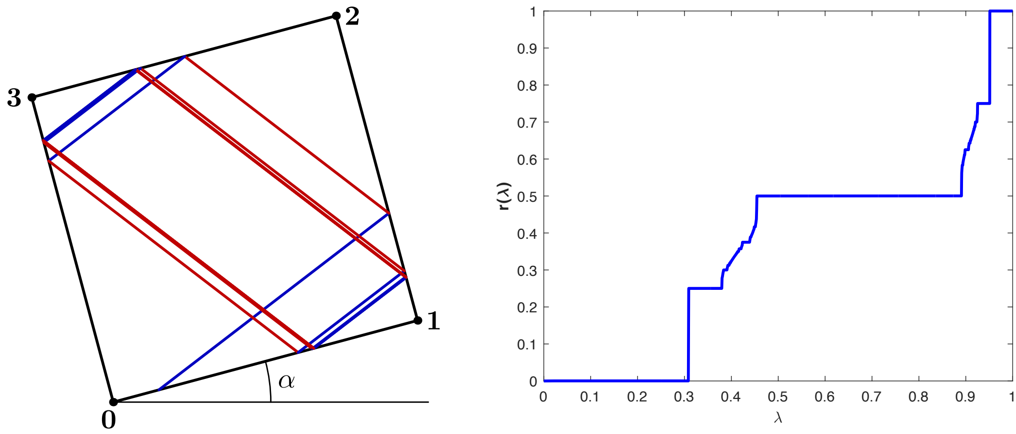

Figure 8. Left: the chess billiard for Ω_α in Example 2.12. The numbers in bold mark the values of the coordinate θ at the vertices. Right: rotation numbers as functions of λ for $\Omega_{\pi/10}$.

Iterating this and using the Lipschitz continuity of $\gamma^\pm$ again, we get

$$|b^n(\theta) - \hat{b}^n(\theta)| \le C\varepsilon \quad \text{for all } \theta \in \mathbb{S}^1.$$

Since $b^n(\theta_j) \neq \theta_j$ for all $j = 1, \dots, m$, taking ε small enough we get (2-31), finishing the proof. □

2.5. ***Examples of Morse–Smale chess billiards.*** Here we present two examples of Morse–Smale chess billiards.

Example 2.12. For $\alpha \in \left(0, \frac{\pi}{2}\right)$, let $\Omega_\alpha \subset \mathbb{R}^2$ be the open square with vertices $(0, 0)$, $(\cos\alpha, \sin\alpha)$, $\sqrt{2}\left(\cos\left(\alpha + \frac{\pi}{4}\right), \sin\left(\alpha + \frac{\pi}{4}\right)\right)$, $\left(\cos\left(\alpha + \frac{\pi}{2}\right), \sin\left(\alpha + \frac{\pi}{2}\right)\right)$. (See Figure 8.) We parametrize $\partial\Omega_\alpha$ by $\theta \in \mathbb{R}/4\mathbb{Z}$ so that the parametrization $\boldsymbol{x}(\theta)$ is affine on each side of the square and the vertices listed above correspond to $\theta = 0, 1, 2, 3$ respectively. For $\lambda \in (0, 1)$, we define

$$\beta \in \left(0, \tfrac{\pi}{2}\right), \quad \tan\beta = \sqrt{1-\lambda^2}/\lambda, \quad t_1 := \tan(\beta - \alpha), \quad t_2 := \tan(\beta + \alpha).$$

We will show that if

$$0 < \alpha < \tfrac{\pi}{8}, \quad \tfrac{\pi}{4} - \alpha < \beta < \tfrac{\pi}{4} + \alpha,$$

or equivalently

$$0 < \alpha < \tfrac{\pi}{8}, \quad \cos\left(\tfrac{\pi}{4} + \alpha\right) < \lambda < \cos\left(\tfrac{\pi}{4} - \alpha\right), \tag{2-32}$$

then λ and Ω_α satisfy the Morse–Smale conditions (Definition 2.10). Moreover, for α, λ satisfying (2-32), we have (identifying θ with $\boldsymbol{x}(\theta)$)

$$\Sigma_\lambda = \left\{ \frac{1-t_1}{t_2-t_1},\ 1 + \frac{t_1(t_2-1)}{t_2-t_1},\ 2 + \frac{1-t_1}{t_2-t_1},\ 3 + \frac{t_1(t_2-1)}{t_2-t_1} \right\}, \tag{2-33}$$

and the rotation number is $\boldsymbol{r}(\lambda) = \frac{1}{2}$.

In fact, assume α, λ satisfy (2-32); then $\ell^+(\bullet, \lambda)$ has exactly two nondegenerate critical points $\boldsymbol{x}(0)$, $\boldsymbol{x}(2)$ on $\partial\Omega_\alpha$, and $\ell^-(\bullet, \lambda)$ also has two nondegenerate critical points $\boldsymbol{x}(1)$, $\boldsymbol{x}(3)$ on $\partial\Omega$. This shows that Ω_α is λ-simple.

We have the following partial computation of the reflection maps $\gamma^\pm$ (note that $0 < t_1 < 1 < t_2 < \infty$ by (2-32)):

$$\gamma^+(\theta) = \begin{cases} t_2^{-1}(2-\theta)+2, & 1 \le \theta \le 2, \\ t_2^{-1}(4-\theta), & 3 \le \theta \le 4, \end{cases} \qquad \gamma^-(\theta) = \begin{cases} t_1(1-\theta)+1, & 0 \le \theta \le 1, \\ t_1(3-\theta)+3, & 2 \le \theta \le 3. \end{cases} \tag{2-34}$$

This in particular implies that we have the mapping properties

$$[0,1] \xrightarrow{\gamma^-} [1,2] \xrightarrow{\gamma^+} [2,3] \xrightarrow{\gamma^-} [3,4] \xrightarrow{\gamma^+} [0,1]. \tag{2-35}$$

Recall that $b = \gamma^+ \circ \gamma^-$. We compute

$$b^2(\theta) = \left(\frac{t_1}{t_2}\right)^2 \theta + \frac{(t_1+t_2)(1-t_1)}{t_2^2}, \qquad \theta \in [0,1]. \tag{2-36}$$

By solving $b^2(\theta_0) = \theta_0$, $\theta_0 \in [0,1]$, we find $\theta_0 = (1-t_1)/(t_2-t_1)$ and

$$\{\theta_0,\ \gamma^-(\theta_0),\ b(\theta_0),\ \gamma^+(\theta_0)\} \subset \Sigma_\lambda.$$

This shows that the right-hand side of (2-33) lies in Σ_λ and that the rotation number is $\boldsymbol{r}(\lambda) = \frac{1}{2}$. On the other hand, suppose $\theta_1 \in \mathbb{R}/4\mathbb{Z}$ and $\theta_1 \in \Sigma_\lambda$. If $\theta_1 \in [0,1]$, then $\theta_1 = \theta_0$ by (2-36). If $\theta_1 \in [2,3]$, then $b(\theta_1) \in \Sigma_\lambda \cap [0,1]$ and thus $\theta_1 = b(\theta_0)$. If $\theta_1 \in [1,2]$, then $\gamma^+(\theta_1) \in \Sigma_\lambda \cap [2,3]$ and thus $\theta_1 = \gamma^-(\theta_0)$. Finally, if $\theta_1 \in [3,4]$, then $\gamma^+(\theta_1) \in \Sigma_\lambda \cap [0,1]$ and thus $\theta_1 = \gamma^+(\theta_0)$. This shows (2-33).

Using (2-36) and the fact that b^2 commutes with b and is conjugated by $\gamma^\pm$ to b^{-2} we compute

$$\partial_\theta b^2(\theta) = \begin{cases} t_1^2/t_2^2 < 1, & \theta \in \{\theta_0, b(\theta_0)\}, \\ t_2^2/t_1^2 > 1, & \theta \in \{\gamma^-(\theta_0), \gamma^+(\theta_0)\}. \end{cases}$$

We have now checked that under the condition (2-32), Ω_α and λ satisfy all conditions in Definition 2.10.

Example 2.13. Let $\mathcal{T}_d \subset \mathbb{R}^2$ be the open trapezium with vertices $(0,0)$, $(1+d,0)$, $(1,1)$, $(0,1)$, $d>0$. (See Figure 9.) We parametrize $\partial\mathcal{T}_d$ by $\theta \in \mathbb{R}/4\mathbb{Z}$ so that the parametrization $\boldsymbol{x}(\theta)$ is affine on each side of the trapezium and the vertices listed above correspond to $\theta = 0, 1, 2, 3$ respectively.

For $\lambda \in (0,1)$, we put $c = \lambda/\sqrt{1-\lambda^2}$. We assume that

$$\max(1,d) < c < d+1. \tag{2-37}$$

Under the condition (2-37) we know $\ell^+(\bullet, \lambda)$ has exactly two nondegenerate critial points $\boldsymbol{x}(0)$, $\boldsymbol{x}(2)$; $\ell^-(\bullet, \lambda)$ also has two nondegenerate critical points $\boldsymbol{x}(1)$, $\boldsymbol{x}(3)$. Hence $\mathcal{T}_d$ is λ-simple.

We have the following partial computation of the reflection maps $\gamma^\pm$:

$$\gamma^+(\theta) = \begin{cases} 2+(c-d)(2-\theta), & 1 \le \theta \le 2, \\ \dfrac{c}{1+d}(4-\theta), & 3 \le \theta \le 4, \end{cases} \qquad \gamma^-(\theta) = \begin{cases} \dfrac{1+d}{c+d}(1-\theta)+1, & 0 \le \theta \le 1, \\ \dfrac{1}{c}(3-\theta)+3, & 2 \le \theta \le 3. \end{cases} \tag{2-38}$$

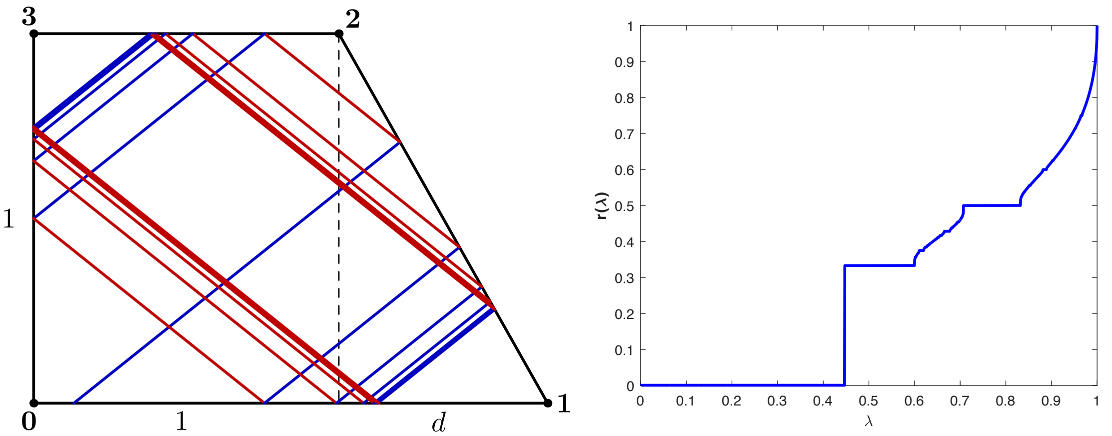

Figure 9. Left: the chess billiard for $\mathcal{T}_d$ in Example 2.13. Right: rotation numbers as functions of λ for $\mathcal{T}_{1/2}$.

This in particular implies that we again have the mapping properties (2-35). From here we compute

$$b^2(\theta) = \frac{c-d}{c+d}\theta + \frac{2c(c-1)}{(1+d)(c+d)}, \qquad \theta \in [0, 1].$$

The fixed point of this map is

$$\theta_0 = \frac{c(c-1)}{d(1+d)}.$$

Arguing as in Example 1, we see that $\mathcal{T}_d$, λ satisfy the conditions of Definition 2.10, with

$$\Sigma_\lambda = \left\{ \frac{c(c-1)}{d(1+d)},\ 2 - \frac{c-1}{d},\ 3 - c + \frac{c(c-1)}{d},\ 4 - \frac{c-1}{d} \right\}.$$

3. Microlocal preliminaries

In this section we present some general results needed in the proof. Most of the microlocal analysis in this paper takes place on the one-dimensional boundary $\partial\Omega$; we review the basic notions in Section 3.1. In Section 3.2 we review definitions and basic properties of conormal distributions (needed in dimensions 1 and 2). These are used to prove and formulate Theorem 1.3: the singularities of $(P - \lambda^2 \mp i0)^{-1} f$ using conormal distributions. In our approach, this structure of $(P - \lambda^2 \mp i0)^{-1} f$ is essential for describing the long-time evolution profile in Theorem 1.4. Finally, Sections 3.3–3.4 contain technical results needed in Section 4.

3.1. ***Microlocal analysis on $\partial\Omega$.*** We first briefly discuss pseudodifferential operators on the circle $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$, referring to [Hörmander 1994, §18.1] for a detailed introduction to the theory of pseudodifferential operators. Pseudodifferential operators on $\mathbb{S}^1$ are given by quantizations of 1-periodic symbols. More precisely, if $0 \le \delta < \frac{1}{2}$ and $m \in \mathbb{R}$, then we say that $a \in C^\infty(\mathbb{R}^2)$ lies in $S^m_\delta(T^*\mathbb{S}^1)$ if (letting $\langle \xi \rangle := \sqrt{1 + |\xi|^2}$)

$$a(x+1, \xi) = a(x, \xi), \qquad |\partial_x^\alpha \partial_\xi^\beta a(x, \xi)| \le C_{\alpha\beta} \langle \xi \rangle^{m + \delta\alpha - (1-\delta)\beta}. \tag{3-1}$$

For brevity we just write $S^m_\delta := S^m_\delta(T^*\mathbb{S}^1)$. Each $a \in S^m_\delta$ is quantized by the operator $\operatorname{Op}(a) : C^\infty(\mathbb{S}^1) \to C^\infty(\mathbb{S}^1),\ \mathcal{D}'(\mathbb{S}^1) \to \mathcal{D}'(\mathbb{S}^1)$ defined by

$$\operatorname{Op}(a)u(x) = \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{i(x-y)\xi} a(x, \xi) u(y)\, dy\, d\xi, \tag{3-2}$$

where $u \in C^\infty(\mathbb{R})$ is 1-periodic and the integral is understood in the sense of oscillatory integrals [Hörmander 1990, §7.8]. We introduce the following spaces of pseudodifferential operators:

$$\Psi^m_\delta := \{\operatorname{Op}(a) : a \in S^m_\delta\}, \quad \Psi^{m+}_\delta = \bigcap_{m'>m} \Psi^{m'}_\delta, \quad \Psi^m_{\delta+} = \bigcap_{\delta'>\delta} \Psi^m_{\delta'},$$
$$S^{m+}_\delta = \bigcap_{m'>m} S^{m'}_\delta, \quad S^m_{\delta+} = \bigcap_{\delta'>\delta} S^m_{\delta'}.$$

We remark that $S^m_{\delta+} \subset S^{m+}_\delta$; moreover, $a \in S^{m+}_\delta$ lies in $S^m_{\delta+}$ if and only if $a(x,\xi) = \mathcal{O}(\langle\xi\rangle^m)$. We henceforth define $\Psi^m := \Psi^m_0$. The space $\Psi^{-\infty} := \bigcap_m \Psi^m$ consists of smoothing operators.

In terms of Fourier series on $\mathbb{S}^1$, we have

$$\begin{aligned} \operatorname{Op}(a)u(x) &= \sum_{k,n\in\mathbb{Z}} e^{2\pi i n x} a_{n-k}(k) u_k, \\ a_\ell(k) &:= \int_0^1 a(x, 2\pi k) e^{-2\pi i \ell x}\, dx, \quad u_k := \int_0^1 u(x) e^{-2\pi i k x}\, dx. \end{aligned} \tag{3-3}$$

This shows that $\operatorname{Op}(a)$ does not determine a uniquely. This representation also shows boundedness on Sobolev spaces $\operatorname{Op}(a) : H^s(\mathbb{S}^1) \to H^{s-m}(\mathbb{S}^1),\ s \in \mathbb{R},\ a \in S^m_\delta$. Indeed, smoothness of a in x shows that $a_\ell(k) = \mathcal{O}(\langle\ell\rangle^{-\infty}\langle k\rangle^m)$ and the bound on the norm follows from the Schur criterion [Dyatlov and Zworski 2019a, (A.5.3)]. Despite the fact that $A := \operatorname{Op}(a)$ does not determine a uniquely, it does determine its *essential support*, which is the right-hand side in the definition of the wave front set of a pseudodifferential operator:

$$\operatorname{WF}(A) := \complement\left\{(x,\xi) : \xi \neq 0,\ \text{there exists } \rho > 0 \text{ such that } a(y,\eta) = \mathcal{O}(\langle\eta\rangle^{-\infty}) \text{ when } |x-y| < \rho,\ \frac{\eta}{\xi} > 0\right\};$$

see [Dyatlov and Zworski 2019a, §E.2]. We refer to that section and [Hörmander 1994, §18.1] for a discussion of wave front sets. We also recall a definition of the wave front set of a distribution,

$$\operatorname{WF}(u) := \bigcap_{Au\in C^\infty,\, A\in\Psi^0} \operatorname{Char}(A),$$
$$\operatorname{Char}(A) := \complement\left\{(x,\xi) : \xi \neq 0, \text{there exists } \rho, c > 0 \text{ such that } |a(y,\eta)| > c,\ |x-y| < \rho,\ \frac{\eta}{\xi} > 0,\ |\eta| > \frac{1}{\rho}\right\}$$

The symbol calculus on $\mathbb{S}^1$ translates directly from the symbol calculus of pseudodifferential operators on $\mathbb{R}$. We record in particular the composition formula [Hörmander 1994, Theorem 18.1.8]: for $b_1 \in S^{m_1}_\delta$, $b_2 \in S^{m_2}_\delta$,

$$\begin{aligned} &\operatorname{Op}(b_1)\operatorname{Op}(b_2) = \operatorname{Op}(b), \quad b \in S^{m_1+m_2}_\delta, \\ &b(x,\xi) = \exp(-i\partial_y\partial_\eta)[b_1(x,\eta)b_2(y,\xi)]|_{(y,\eta)=(x,\xi)}, \\ &b(x,\xi) = \sum_{0\le k<N} \frac{(-i)^k}{k!}\, \partial^k_\xi b_1(x,\xi)\, \partial^k_x b_2(x,\xi) + b_N(x,\xi), \quad b_N \in S^{m-N(1-2\delta)}_\delta, \end{aligned} \tag{3-4}$$

where expanding the exponential gives an asymptotic expansion of b.

We record here a norm bound for pseudodifferential operators at high frequency:

Lemma 3.1. *Assume that $a \in S^0_\delta$, $r \in S^{-1+}$, and $\sup |a| \le R$. Then, for all $N, \nu > 0$, and $u \in L^2(\mathbb{S}^1)$ we have*

$$\| \operatorname{Op}(a+r)u\|_{L^2} \le (R+\nu)\|u\|_{L^2} + C\|u\|_{H^{-N}}, \tag{3-5}$$

where the constant C depends on R, ν, N, and some seminorms of a and r but not on u.

Proof. By [Grigis and Sjöstrand 1994, Lemma 4.6] we can write

$$(R+\nu)^2 I = \operatorname{Op}(a+r)^* \operatorname{Op}(a+r) + \operatorname{Op}(b)^* \operatorname{Op}(b) + \operatorname{Op}(q)$$

for some $b \in S^0_\delta$ and $q \in S^{-\infty}$. The bound (3-5) follows. ☐

Although a in (3-3) is not unique, the principal *symbol* of $\operatorname{Op}(a)$ defined as

$$\sigma(\operatorname{Op}(a)) = [a] \in S^m_\delta / S^{m-1+2\delta}_\delta \tag{3-6}$$

is, and we have a short exact sequence $0 \to \Psi^{m-1+2\delta}_\delta \to \Psi^m_\delta \xrightarrow{\sigma} S^m_\delta / S^{m-1+2\delta}_\delta \to 0$. Somewhat informally, we write $\sigma(\operatorname{Op}(a)) = b$ for any b satisfying $a - b \in S^{m-1+2\delta}_\delta$.

In our analysis, we also consider families $\varepsilon \mapsto a_\varepsilon$, $\varepsilon \ge 0$, such that $a_\varepsilon \in S^{-\infty}$ for $\varepsilon > 0$ and $a_0 \in S^m_\delta$. In that case, for $A_\varepsilon = \operatorname{Op}(a_\varepsilon)$,

$$\sigma(A_\varepsilon) = [b_\varepsilon], \quad b_\varepsilon - a_\varepsilon \in S^{m-1+2\delta}_\delta \text{ uniformly for } \varepsilon \ge 0. \tag{3-7}$$

Again, we drop $[\bullet]$ when writing $\sigma(A)$ for a specific operator.

We will crucially use mild exponential weights which result in pseudodifferential operators of varying order — see [Unterberger 1971], and in a related context [Faure et al. 2008].

Lemma 3.2. *Suppose that (in the sense of (3-1)) $m_j \in S^0$, m_0 is real-valued, and*

$$G(x,\xi) := m_0(x,\xi)\log\langle\xi\rangle + m_1(x,\xi), \quad m_0(x,t\xi) = m_0(x,\xi), \quad t, |\xi| \ge 1. \tag{3-8}$$

Then

$$e^G \in S^M_{0+}, \quad e^{-G} \in S^{-m}_{0+}, \quad M := \max_{|\xi|=1} m_0(x,\xi), \quad m := \min_{|\xi|=1} m_0(x,\xi), \tag{3-9}$$

and there exists $r_G \in S^{-1+}$ such that

$$\operatorname{Op}(e^G)\operatorname{Op}(e^{-G}(1+r_G)) - I,\ \operatorname{Op}(e^{-G}(1+r_G))\operatorname{Op}(e^G) - I \in \Psi^{-\infty}. \tag{3-10}$$

Also, if $G_j(x,\xi)$ are given by (3-8) with m_0 and m_1 replaced by m_{0j}, m_{1j}, respectively, then, for $a_j \in S^0$, $r_j \in S^{-1+}$, $j = 1, 2$, there exists $r_3 \in S^{-1+}$ such that

$$\operatorname{Op}(e^{G_1}(a_1+r_1))\operatorname{Op}(e^{G_2}(a_2+r_2)) = \operatorname{Op}(e^{G_1+G_2}(a_1a_2+r_3)). \tag{3-11}$$

Proof. Since $\log\langle\xi\rangle = \mathcal{O}_\varepsilon(\langle\xi\rangle^\varepsilon)$ for all $\varepsilon > 0$, (3-9) follows from (3-1). In fact, we have the stronger bound

$$|\partial_x^\alpha \partial_\xi^\beta (e^{\pm G(x,\xi)})| \le C_{\alpha\beta\varepsilon} e^{\pm G(x,\xi)} \langle\xi\rangle^{\varepsilon-|\beta|}, \quad \varepsilon > 0. \tag{3-12}$$

This gives (3-11). Indeed, the remainder in the expansion (3-4) is in $S^{M_1+M_2-N+}$ and the k-th term is in $e^{G_1+G_2}S^{-k+}$ by (3-12); it suffices to take $N \geq M + M_2 + 1$.

To obtain (3-10) we note that (3-11) gives $\operatorname{Op}(e^{\pm G})\operatorname{Op}(e^{\mp G}) = I - \operatorname{Op}(r_\pm)$, $r_\pm \in S^{-1+}$. We then have parametrices for the operators $I - \operatorname{Op}(r_\pm)$ [Hörmander 1994, Theorem 18.1.9], $I + \operatorname{Op}(b_\pm)$, which give left and right approximate inverses (in the sense of (3-10)) $(I + \operatorname{Op}(b_-))\operatorname{Op}(e^{-G})$, $\operatorname{Op}(e^{-G})(I + \operatorname{Op}(b_+))$. Those have the required form by (3-11) (where one of G_1, G_2 is equal to $-G$ and the other one is equal to 0). □

We also record a change of variables formula. Suppose $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is a diffeomorphism with a lift $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}$, $\boldsymbol{f}(x + 1) = \boldsymbol{f}(x) \pm 1$ (with "+" for orientation-preserving f and "−" otherwise). For symbols 1-periodic in x we can use the standard formula given in [Hörmander 1994, Theorem 18.1.17] and an argument similar to (3-11). That gives, for G given by (3-8), and $r \in S^{-1+}$,

$$\begin{aligned} f^* \circ \operatorname{Op}(e^G(1 + r)) &= \operatorname{Op}(e^{G_f}(1 + r_f)) \circ f^*, \\ G_f(x, \xi) &:= G(f(x), f'(x)^{-1}\xi), \quad r_f \in S^{-1+}. \end{aligned} \tag{3-13}$$

In Section 4.6 below we will use pseudodifferential operators acting on 1-forms on $\mathbb{S}^1$. Using the canonical 1-form dx, $x \in \mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$, we identify 1-forms with functions, and this gives an identification of the class $\Psi^m_\delta(\mathbb{S}^1; T^*\mathbb{S}^1)$ (operators acting on 1-forms) with $\Psi^m_\delta(\mathbb{S}^1)$ (operators acting on functions). This defines the principal symbol map, which we still denote by σ.

Fixing a positively oriented coordinate $\theta : \partial\Omega \to \mathbb{S}^1$, we can identify functions/distributions on $\partial\Omega$ with functions/distributions on $\mathbb{S}^1$. The change of variables formula used for (3-13) also shows the invariance of $\sigma(A)$ under changes of variables and allows pseudodifferential operators acting on section of bundles — see [Hörmander 1994, Definition 18.1.32]. In particular, we can define the class of pseudodifferential operators $\Psi^m_\delta(\partial\Omega; T^*\partial\Omega)$ acting on 1-forms on $\partial\Omega$ and the symbol map

$$\sigma : \Psi^m_\delta(\partial\Omega; T^*\partial\Omega) \to S^m_\delta(T^*\partial\Omega)/S^{m-1+2\delta}_\delta(T^*\partial\Omega), \tag{3-14}$$

with the class Ψ^m_δ and the map σ independent of the choice of coordinate on $\partial\Omega$.

3.2. ***Conormal distributions.*** We now review conormal distributions associated to hypersurfaces, referring the reader to [Hörmander 1994, §18.2] for details. Although we consider the case of manifolds with boundaries, the hypersurfaces are assumed to be transversal to the boundaries and conormal distributions are defined as restrictions of conormal distributions in the no-boundary case.

Let M be a compact m-dimensional manifold with boundary and $\Sigma \subset M$ be a compact hypersurface transversal to the boundary (that is, Σ is a compact codimension-1 submanifold of M with boundary $\partial\Sigma = \Sigma \cap \partial M$ and $T_x\Sigma \neq T_x\partial M$ for all $x \in \partial\Sigma$). We should emphasize that in our case, the hypersurfaces Σ take a particularly simple form: we either have $M = \overline{\Omega}$ and Σ given by straight lines transversal to $\partial\Omega$ (see Theorems 1.3 and 1.4) or $M = \partial\Omega \simeq \mathbb{S}^1$ and Σ is given by points (see Propositions 7.3 and 7.4).

The conormal bundle to Σ is given by $N^*\Sigma := \{(x, \xi) \in T^*M : x \in \Sigma,\ \xi|_{T_x\Sigma} = 0\}$, which is a Lagrangian submanifold of T^*M and a one-dimensional vector bundle over Σ. For $k \in \mathbb{R}$, define the symbol class $S^k(N^*\Sigma)$ consisting of functions $a \in C^\infty(N^*\Sigma)$ satisfying the derivative bounds

$$|\partial_x^\alpha \partial_\theta^\beta a(x, \theta)| \leq C_{\alpha\beta}\langle\theta\rangle^{k-|\beta|}, \tag{3-15}$$

where we use local coordinates $(x,\theta)\in\mathbb{R}^{m-1}\times\mathbb{R}\simeq N^*\Sigma$. Here x is a coordinate on Σ and θ is a linear coordinate on the fibers of $N^*\Sigma$; $\langle\theta\rangle:=\sqrt{1+|\theta|^2}$. The estimates (3-15) are supposed to be valid uniformly up to the boundary of Σ. In other words we can consider a as a restriction of a symbol defined on an extension of Σ.

Denote by $I^s(M,N^*\Sigma)\subset\overline{\mathcal{D}'}(M^\circ)$ the space of extendible distributions on the interior M° (see [Hörmander 1994, §B.2]) which are conormal to Σ of order $s\in\mathbb{R}$ smoothly up to the boundary of M. To describe the class I^s we first consider two model cases:

- If $M=\mathbb{R}^m$, we write points in $\mathbb{R}^m$ as $(x_1,x')\in\mathbb{R}\times\mathbb{R}^{m-1}$, and $\Sigma=\{x_1=0\}$, then a compactly supported distribution $u\in\mathcal{E}'(\mathbb{R}^m)$ lies in $I^s(M,N^*\Sigma)$ if and only if its Fourier transform in the x_1-variable, $\check{u}(\xi_1,x')$, lies in $S^{m/4-1/2+s}(N^*\Sigma)$, where $N^*\Sigma=\{(0,x',\xi_1,0)\mid x'\in\mathbb{R}^{m-1},\ \xi_1\in\mathbb{R}\}$.
- If $M=\mathbb{R}_{x_1}\times[0,\infty)_{x_2}\times\mathbb{R}^{m-2}_{x''}$ and $\Sigma=\{x_1=0,\ x_2\ge 0\}$, then a distribution $u\in\mathcal{D}'(M^\circ)$ with bounded support lies in $I^s(M,N^*\Sigma)$ if and only if $u=\tilde{u}|_{M^\circ}$ for some $\tilde{u}\in\mathcal{E}'(\mathbb{R}^m)$ which lies in $I^s(\mathbb{R}^m,N^*\widetilde{\Sigma})$, with $\widetilde{\Sigma}:=\{x_1=0\}\subset\mathbb{R}^m$. Alternatively, $\check{u}(\xi_1,x')$ lies in $S^{m/4-1/2+s}(N^*\Sigma)$, where the derivative bounds are uniform up to the boundary.

In those model cases, elements of $I^s(M,N^*\Sigma)$ are given by the oscillatory integrals (where we use the prefactor from [Hörmander 1994, Theorem 18.2.9])

$$u(x)=(2\pi)^{-m/4-1/2}\int_{\mathbb{R}}e^{ix_1\xi_1}a(x',\xi_1)\,d\xi_1,\quad a\in S^{m/4-1/2+s}(N^*\{x_1=0\}). \tag{3-16}$$

We note that in both of the above cases the distribution u is in $C^\infty(M)$ (up to the boundary in the second case) outside of any neighborhood of Σ.

For the case of general compact manifold M and hypersurface Σ transversal to the boundary of M, we say that $u\in I^s(M,N^*\Sigma)$ if (see [Hörmander 1994, Theorem 18.2.8])

(1) u is in $C^\infty(M)$ (up to the boundary) outside of any neighborhood of Σ, and

(2) the localizations of u to the model cases using coordinates lie in I^s as defined above.

Note that the wavefront set of any $u\in I^s(M,N^*\Sigma)$, considered as a distribution on the interior M°, is contained in $N^*\Sigma$.

In addition we define the space

$$I^{s+}(M,N^*\Sigma):=\bigcap_{s'>s}I^{s'}(M,N^*\Sigma).$$

Such spaces are characterized in terms of the Sobolev spaces (where for simplicity assume that M has no boundary, since this is the only case used in this paper), $H^{s-}(M):=\bigcap_{s'<s}H^{s'}(M)$, as follows:

$$u\in I^{s+}(M,N^*\Sigma)\iff$$
$$\text{for any vector fields } X_1,\dots,X_\ell \text{ on } M \text{ tangent to } \Sigma, \text{ we have } X_1\cdots X_\ell\, u\in H^{-m/4-s-}(M); \tag{3-17}$$

see [Hörmander 1994, Definition 18.2.6 and Theorem 18.2.8].

Assume now that the conormal bundle $N^*\Sigma$ is oriented; for $(x,\xi)\in N^*\Sigma\setminus 0$ we say that $\xi>0$ if ξ is positively oriented and $\xi<0$ if ξ is negatively oriented. This gives the splitting

$$N^*\Sigma\setminus 0=N^*_+\Sigma\sqcup N^*_-\Sigma,\qquad N^*_\pm\Sigma:=\{(x,\xi)\in N^*\Sigma\mid \pm\xi>0\}. \tag{3-18}$$

Denote by $I^s(M,N^*_\pm\Sigma)$ the space of distributions $u\in I^s(M,N^*\Sigma)$ such that $\mathrm{WF}(u)\subset N^*_\pm\Sigma$, up to the boundary. Since Σ is transversal to the boundary this means that an extension of u satisfies this condition. In the model case (and effectively in the cases considered in this paper) $M=\mathbb{R}^m$, $\Sigma=\{x_1=0\}$ they can be characterized as follows: $\check{u}(\xi_1,x')$ lies in $S^{m/4-1/2+s}(N^*\Sigma)$ and $\check{u}(\xi_1,x')=\mathcal{O}(\langle\xi_1\rangle^{-\infty})$ as $\xi_1\to\mp\infty$.

In the present paper we will often study the case when $M=\partial\Omega$, identified with $\mathbb{S}^1$ by a coordinate θ, and we are given two finite sets $\Sigma^+,\Sigma^-\subset\partial\Omega$, with $\Sigma^+\cap\Sigma^-=\varnothing$. We define

$$I^s(\partial\Omega,N^*_+\Sigma^-\sqcup N^*_-\Sigma^+):=I^s(\partial\Omega,N^*_+\Sigma^-)+I^s(\partial\Omega,N^*_-\Sigma^+). \tag{3-19}$$

Put $\Sigma:=\Sigma^+\sqcup\Sigma^-$. Then $I^s(\partial\Omega,N^*_+\Sigma^-\sqcup N^*_-\Sigma^+)$ consists of the elements of $I^s(\partial\Omega;N^*\Sigma)$ with wavefront set contained in $N^*_+\Sigma^-\sqcup N^*_-\Sigma^+$.

Assume that Σ has an even number of points (which will be the case in our application) and fix a defining function $\rho\in C^\infty(\partial\Omega;\mathbb{R})$ of Σ: that is, $\Sigma=\rho^{-1}(0)$ and $d\rho\neq 0$ on Σ. Fix also a pseudodifferential operator $A_\Sigma\in\Psi^0(\partial\Omega)$ such that $\mathrm{WF}(A_\Sigma)\cap(N^*_+\Sigma^-\sqcup N^*_-\Sigma^+)=\varnothing$ and A_Σ is elliptic on $N^*_-\Sigma^-\sqcup N^*_+\Sigma^+$. Using (3-17), we see that $u\in\mathcal{D}'(\partial\Omega)$ lies in $I^{s+}(\partial\Omega,N^*_+\Sigma^-\sqcup N^*_-\Sigma^+)$ if and only if the following seminorms are finite:

$$\|(\rho\partial_\theta)^N u\|_{H^{-1/4-s-\beta}},\quad \|A_\Sigma u\|_{H^N}\quad\text{for all } N\in\mathbb{N}_0,\ \beta>0. \tag{3-20}$$

Choosing different ρ and A_Σ leads to an equivalent family of seminorms (3-20). In particular, if ρ, A_Σ are as above and $\tilde{A}_\Sigma\in\Psi^0(\partial\Omega)$ satisfies $\mathrm{WF}(\tilde{A}_\Sigma)\cap(N^*_+\Sigma^-\sqcup N^*_-\Sigma^+)=\varnothing$ then $\mathrm{WF}(\tilde{A}_\Sigma)$ lies in the union of $\{\rho\neq 0\}$ and the elliptic set of A_Σ; thus by the elliptic estimate we have for $N_0\geq N+\frac14+s+\beta$

$$\|\tilde{A}_\Sigma u\|_{H^N}\leq C(\|A_\Sigma u\|_{H^N}+\|(\rho\partial_\theta)^{N_0}u\|_{H^{-1/4-s-\beta}}+\|u\|_{H^{-1/4-s-\beta}}). \tag{3-21}$$

Moreover, the operator $\rho\partial_\theta$ is bounded with respect to the seminorms (3-20), as are pseudodifferential operators in $\Psi^0(\partial\Omega)$ [Hörmander 1994, Theorem 18.2.7].

We will also need the notion of conormal distributions depending smoothly on a parameter — see [Dyatlov and Zworski 2019b, Lemma 4.4] for a more general Lagrangian version. Here we restrict ourselves to the specific conormal distributions appearing in this paper and define relevant smooth families of conormal distributions in Proposition 7.4 and Lemma 8.3.

We will not discuss principal symbols of general conormal distributions to avoid introducing half-densities; however, we give here a special case of the way the principal symbol changes under pseudodifferential operators and under pullbacks:

Lemma 3.3. *Assume that $u\in\mathcal{E}'(\mathbb{R})$ lies in $I^s(\mathbb{R},\{0\})$, that is, $\hat{u}\in S^{s-1/4}(\mathbb{R})$. Then*:

(1) *If $a(x,\xi)\in S^0(T^*\mathbb{R})$ is compactly supported in the x-variable and* $\mathrm{Op}(a)$ *is defined by* (3-2), *then*

$$\widehat{\mathrm{Op}(a)u}(\xi)=a(0,\xi)\hat{u}(\xi)+S^{s-5/4}(\mathbb{R}). \tag{3-22}$$

(2) *If f is a diffeomorphism of open subsets of $\mathbb{R}$ such that $f(0)=0$ and the range of f contains* $\operatorname{supp} u$, *then*

$$\widehat{f^*u}(\xi) = \frac{1}{|f'(0)|}\hat{u}\left(\frac{\xi}{f'(0)}\right) + S^{s-5/4}(\mathbb{R}). \tag{3-23}$$

Proof. Since these statements are standard, we only sketch the proofs, referring to [Hörmander 1994, Theorems 18.2.9 and 18.2.12] for details. To see (3-22) we use the formula

$$\widehat{\operatorname{Op}(a)u}(\xi) = \frac{1}{2\pi}\int_{\mathbb{R}^2} e^{ix(\eta-\xi)}a(x,\eta)\hat{u}(\eta)\,d\eta\,dx.$$

Assume that $|\xi| \geq 1$. Using a smooth partition of unity, we split the integral above into two pieces: one where $|\eta-\xi| \geq \frac{1}{4}|\xi|$ and another one where $|\eta-\xi| \leq \frac{1}{2}|\xi|$. The first piece is rapidly decaying in ξ by integration by parts in x. The second piece is equal to $a(0,\xi)\hat{u}(\xi) + S^{s-5/4}(\mathbb{R})$ by the method of stationary phase.

To see (3-23) we fix a cutoff $\chi \in C_c^\infty(\mathbb{R})$ such that $\operatorname{supp}\chi$ lies in the range of f and $\chi = 1$ near $\operatorname{supp} u$. Using the Fourier inversion formula, we write

$$\widehat{f^*u}(\xi) = \frac{1}{2\pi}\int_{\mathbb{R}^2} e^{i(f(x)\eta - x\xi)}\chi(f(x))\hat{u}(\eta)\,d\eta\,dx.$$

Now (3-23) is proved similarly to (3-22). Here in the application of the method of stationary phase, the critical point is given by $x=0$, $\eta = \xi/f'(0)$ and the Hessian of the phase at the critical point has signature 0 and determinant $-f'(0)^2$. □

3.3. ***Convolution with logarithm.*** In Section 4 we need information about mapping properties between spaces of conormal distributions on the boundary and conormal distributions in the interior. In preparation for Lemma 4.8 below we now prove the following:

Lemma 3.4. *Let $f \in C_c^\infty(\mathbb{R})$ and define*

$$g(x) := \int_0^\infty \log|x-y|\frac{f(y)}{\sqrt{y}}\,dy, \quad x > 0. \tag{3-24}$$

Then $g \in C^\infty([0,\infty))$.

Remark. In general g is *not* smooth on $(-\infty, 0]$. In fact, changing variables $y = s^2|x|$, we obtain (see (3-25) below)

$$g'(x) = -2|x|^{-1/2}\int_0^\infty \frac{f(s^2|x|)}{1+s^2}\,ds, \quad x < 0,$$

which blows up as $x \to 0-$ if $f(0) \neq 0$. This, and the conclusion of Lemma 3.4, can also be seen from analysis on the Fourier transform side.

Proof. Let $x_+^{-1/2} := H(x)x^{-1/2}$, where $H(x)$ is the Heaviside function: $H(x)=1$ for $x>0$ and $H(x)=0$ for $x<0$. Since g is the convolution of $\log|x|$ and $x_+^{-1/2}f(x)$, which are both smooth except at $x=0$, the function g is smooth on $(0,\infty)$. Thus it suffices to prove that g is smooth on $[0,1]$ up to the boundary.

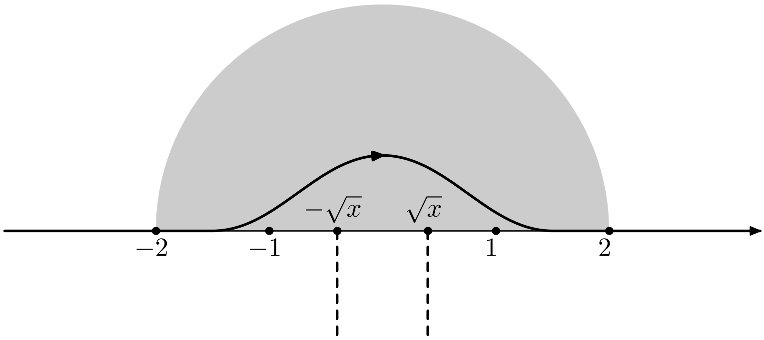

Figure 10. The contour Γ used in the proof of Lemma 3.4. The dashed lines are the cuts needed to define the function $\psi_x(z)$.

(1) Assume first that f is real-valued and extends holomorphically to the disk $\{|z| < 4\}$. Making the change of variables $y := t^2$, we write

$$g(x) = 2\int_0^\infty \log|t^2 - x| f(t^2)\, dt = \int_{\mathbb{R}} \log|t^2 - x| f(t^2)\, dt. \tag{3-25}$$

Assume that $x \in (0, 1]$ and consider the holomorphic function

$$\psi_x(z) := \log(z - \sqrt{x}) + \log(z + \sqrt{x}), \quad z \in \mathbb{C} \setminus \big((\sqrt{x} - i[0, \infty)) \cup (-\sqrt{x} - i[0, \infty))\big),$$

where we use the branch of the logarithm on $\mathbb{C} \setminus -i[0, \infty)$ which takes real values on $(0, \infty)$. Then $\operatorname{Re}\psi_x(t) = \log|t^2 - x|$ for all $t \in \mathbb{R} \setminus \{\sqrt{x}, -\sqrt{x}\}$.

Fix an x-independent contour $\Gamma = \{t + iw(t) \mid t \in \mathbb{R}\} \subset \mathbb{C}$ such that $w(t) \geq 0$ everywhere, $w(t) = 0$ for $|t| \geq \frac{3}{2}$, $|t + iw(t)| < 2$ for $|t| \leq \frac{3}{2}$, and $w(t) > 0$ for $|t| \leq 1$. (See Figure 10.) Deforming the contour in (3-25), we get

$$g(x) = \operatorname{Re}\int_\Gamma \psi_x(z) f(z^2)\, dz \quad \text{for all } x \in (0, 1].$$

Since $\partial_x \psi_x(z) = (x - z^2)^{-1}$, the function $\psi_x(z)$ and all its x-derivatives are bounded uniformly in $x \in (0, 1]$ and locally uniformly in $z \in \Gamma$. It follows that g is smooth on the interval $[0, 1]$.

(2) For the general case, fix a cutoff $\chi \in C_c^\infty(\mathbb{R})$ such that $\chi = 1$ near $[-4, 4]$. Take arbitrary $N \in \mathbb{N}$. Using the Taylor expansion of f at 0, we write

$$f(x) = f_1(x) + f_2(x), \quad f_1(x) = p(x)\chi(x),$$

where p is a polynomial of degree at most N and $f_2 \in C_c^\infty(\mathbb{R})$ satisfies $f_2(x) = \mathcal{O}(|x|^{N+1})$ as $x \to 0$. We write $g = g_1 + g_2$, where g_j are constructed from f_j using (3-24).

By step (1) of the present proof, we see that g_1 is smooth on $[0, 1]$. On the other hand, $x_+^{-1/2} f_2(x) \in C^N(\mathbb{R})$; since $\log|x|$ is locally integrable we get $g_2 \in C^N([0, 1])$. Since N can be chosen arbitrary, this gives $g \in C^\infty([0, 1])$ and finishes the proof. $\square$

We also give the following general mapping property of convolution with logarithm on conormal spaces used in Lemma 4.9 below:

Lemma 3.5. *Let* $\Sigma \subset \mathbb{R}$ *be a finite set and put* $\log_\pm x := \log(x \pm i0)$. *Then*

$$f \in I^s(\mathbb{R}, N^*_\pm \Sigma) \cap \mathcal{E}'(\mathbb{R}) \quad \Longrightarrow \quad \begin{cases} \log_\pm * f \in I^{s-1}(\mathbb{R}, N^*_\pm \Sigma), \\ \log_\mp * f \in C^\infty(\mathbb{R}). \end{cases}$$

Proof. Assume that $f \in I^s(\mathbb{R}, N^*_+ \Sigma) \cap \mathcal{E}'(\mathbb{R})$ (the case of $N^*_- \Sigma$ is handled similarly). We may reduce to the case $\Sigma = \{0\}$. Since ∂_x is an elliptic operator, the local definition (3-16) (with no x'-variable) shows that it is enough to prove

$$\partial_x \log_+ * f \in I^s(\mathbb{R}, N^*_+\{0\}), \quad \partial_x \log_- * f \in C^\infty(\mathbb{R}).$$

It remains to use that $\partial_x \log_\pm = (x \pm i0)^{-1}$ and we have the Fourier transform formula (see [Hörmander 1990, Example 7.1.17]; here H is the Heaviside function)

$$\boldsymbol{u}_{\pm 0}(x) := (x \pm i0)^{-1} \quad \Longrightarrow \quad \hat{\boldsymbol{u}}_{\pm 0}(\xi) = \mp 2\pi i H(\pm\xi), \tag{3-26}$$

completing the proof. □

3.4. *Microlocal structure of* $(x \pm i\varepsilon\psi(x))^{-1}$. In this section we study the behavior as $\varepsilon \to 0+$ of functions of the form

$$\chi(x, \varepsilon)(x \pm i\varepsilon\psi(x, \varepsilon))^{-1} \in C^\infty(J), \quad 0 < \varepsilon < \varepsilon_0, \tag{3-27}$$

where $J \subset \mathbb{R}$ is an open interval containing 0 and

$$\chi, \psi \in C^\infty(J \times [0, \varepsilon_0); \mathbb{C}), \qquad \operatorname{Re} \psi > 0 \quad \text{on } J \times [0, \varepsilon_0). \tag{3-28}$$

We first decompose (3-27) into the sum of $r(x \pm i\varepsilon z)^{-1}$, where $r, z \in \mathbb{C}$ depend on ε but not on x, and a function which is smooth uniformly in ε:

Lemma 3.6. *Under the conditions* (3-28) *we have for all* $\varepsilon \in (0, \varepsilon_0)$

$$\chi(x, \varepsilon)(x \pm i\varepsilon\psi(x, \varepsilon))^{-1} = r^\pm(\varepsilon)(x \pm i\varepsilon z^\pm(\varepsilon))^{-1} + q^\pm(x, \varepsilon), \tag{3-29}$$

where $r^\pm, z^\pm \in C^\infty([0, \varepsilon_0))$ *and* $q^\pm \in C^\infty(J \times [0, \varepsilon_0))$ *are complex-valued*, $\operatorname{Re} z^\pm > 0$ *on* $[0, \varepsilon_0)$, $z^\pm(0) = \psi(0, 0)$, *and* $\chi(x, 0) = xq^\pm(x, 0) + r^\pm(0)$.

Proof. Since $(x \pm i\varepsilon\psi(x, \varepsilon))^{-1}$ is a smooth function of $(x, \varepsilon) \in J \times [0, \varepsilon_0)$ outside of $(0, 0)$, it is enough to show that (3-29) holds for $|x|, \varepsilon$ small enough.

The complex-valued function $F^\pm(x, \varepsilon) := x \pm i\varepsilon\psi(x, \varepsilon)$ is smooth in $(x, \varepsilon) \in J \times [0, \varepsilon_0)$ and satisfies $F^\pm(0, 0) = 0$ and $\partial_x F^\pm(0, 0) = 1$. Thus by the Malgrange preparation theorem [Hörmander 1990, Theorem 7.5.6], we have for (x, ε) in some neighborhood of $(0, 0)$ in $J \times [0, \varepsilon_0)$

$$x = q_1^\pm(x, \varepsilon)(x \pm i\varepsilon\psi(x, \varepsilon)) + r_1^\pm(\varepsilon),$$

where $q_1^\pm, r_1^\pm$ are smooth. Taking $\varepsilon = 0$ we get $r_1^\pm(0) = 0$ and $q_1^\pm(x, 0) = 1$; differentiating in ε and then putting $x = \varepsilon = 0$ we get $\partial_\varepsilon r_1^\pm(0) = \mp i\psi(0, 0)$. We put $z^\pm(\varepsilon) := \pm i\varepsilon^{-1} r_1^\pm(\varepsilon)$, so that when $\varepsilon > 0$

$$(x \pm i\varepsilon\psi(x, \varepsilon))^{-1} = q_1^\pm(x, \varepsilon)(x \pm i\varepsilon z^\pm(\varepsilon))^{-1}. \tag{3-30}$$

Note that $z^\pm(0) = \psi(0, 0)$ and thus $\operatorname{Re} z^\pm(\varepsilon) > 0$ for small ε.

Now, we use the Malgrange preparation theorem again, this time for the function $F^{\pm}(x,\varepsilon) := x \pm i\varepsilon z^{\pm}(\varepsilon)$, to get for (x,ε) in some neighborhood of $(0,0)$ in $J\times[0,\varepsilon_0)$

$$\chi(x,\varepsilon)q_1^{\pm}(x,\varepsilon) = q^{\pm}(x,\varepsilon)(x\pm i\varepsilon z^{\pm}(\varepsilon)) + r^{\pm}(\varepsilon),$$

where $q^{\pm}$, $r^{\pm}$ are again smooth. Taking $\varepsilon = 0$ we get $\chi(x,0) = xq^{\pm}(x,0) + r^{\pm}(0)$. Together with (3-30) this gives the decomposition (3-29). □

As an application of Lemma 3.6, we give:

Lemma 3.7. *Assume that ψ satisfies (3-28). Then we have for all $s < -\frac{1}{2}$*

$$(x \pm i\psi(x,\varepsilon))^{-1} \to (x\pm i0)^{-1} \quad \text{as } \varepsilon \to 0+ \quad \text{in } H^s_{\text{loc}}(J). \tag{3-31}$$

Proof. Put $\chi \equiv 1$ and let $z^{\pm}(\varepsilon)$, $r^{\pm}(\varepsilon)$, $q^{\pm}(x,\varepsilon)$ be given by Lemma 3.6; note that $1 = xq^{\pm}(x,0) + r^{\pm}(0)$; thus $r^{\pm}(0) = 1$ and $q^{\pm}(x,0) = 0$. We have

$$(x\pm i\varepsilon z^{\pm}(\varepsilon))^{-1} \to (x\pm i0)^{-1} \quad \text{in } H^s(\mathbb{R}).$$

Indeed, the Fourier transform of the right-hand side is equal to $\mp 2\pi i H(\pm\xi)$ by (3-26) and the Fourier transform of the left-hand side is equal to $\mp 2\pi i H(\pm\xi)e^{-\varepsilon z^{\pm}(\varepsilon)|\xi|}$ by

$$\boldsymbol{u}_{\pm z}(x) := (x\pm iz)^{-1}, \quad \operatorname{Re} z > 0 \quad \Longrightarrow \quad \hat{\boldsymbol{u}}_{\pm z}(\xi) = \mp 2\pi i H(\pm\xi)e^{-z|\xi|}. \tag{3-32}$$

We have convergence of these Fourier transforms in $L^2(\mathbb{R};\langle\xi\rangle^{2s}\,d\xi)$ by the dominated convergence theorem.

By (3-29) this implies that the left-hand side of (3-31) converges in $H^s_{\text{loc}}(J)$ to

$$r^{\pm}(0)(x\pm i0)^{-1} + q^{\pm}(x,0) = (x\pm i0)^{-1},$$

which finishes the proof. □

The functions $r^{\pm}$, $z^{\pm}$, $q^{\pm}$ in Lemma 3.6 are not uniquely determined by χ, ψ; however, they are unique up to $\mathcal{O}(\varepsilon^{\infty})$:

Lemma 3.8. *Assume that $r_j^{\pm}, z_j^{\pm} \in C^{\infty}([0,\varepsilon_0))$, $j = 1,2$, are complex-valued functions such that $\operatorname{Re} z_j^{\pm} > 0$ on $[0,\varepsilon_0)$, $r_j^{\pm}(0) \neq 0$, and*

$$\tilde{q}^{\pm}(x,\varepsilon) := r_1^{\pm}(\varepsilon)(x\pm i\varepsilon z_1^{\pm}(\varepsilon))^{-1} - r_2^{\pm}(\varepsilon)(x\pm i\varepsilon z_2^{\pm}(\varepsilon))^{-1} \in C^{\infty}(J\times[0,\varepsilon_0)).$$

Then $r_1^{\pm}(\varepsilon) - r_2^{\pm}(\varepsilon)$, $z_1^{\pm}(\varepsilon) - z_2^{\pm}(\varepsilon)$, and $\tilde{q}^{\pm}(x,\varepsilon)$ are $\mathcal{O}(\varepsilon^{\infty})$, that is all their derivatives in ε vanish at $\varepsilon = 0$.

Proof. Differentiating $k-1$ times in x and then putting $x = 0$ we see that for all $k \geq 1$

$$\frac{r_1^{\pm}(\varepsilon)}{\varepsilon^k z_1^{\pm}(\varepsilon)^k} - \frac{r_2^{\pm}(\varepsilon)}{\varepsilon^k z_2^{\pm}(\varepsilon)^k} \in C^{\infty}([0,\varepsilon_0)).$$

Therefore

$$\partial_\varepsilon^\ell|_{\varepsilon=0}\left(\frac{r_1^{\pm}(\varepsilon)}{z_1^{\pm}(\varepsilon)^k} - \frac{r_2^{\pm}(\varepsilon)}{z_2^{\pm}(\varepsilon)^k}\right) = 0 \quad \text{for all } 0 \leq \ell < k. \tag{3-33}$$

Taking $\ell = 0$ and $k = 1, 2$, we see that $r_1^{\pm}(0) = r_2^{\pm}(0)$ and $z_1^{\pm}(0) = z_2^{\pm}(0)$. Arguing by induction on ℓ and using $k = \ell + 1, \ell + 2$ in (3-33) we see that $\partial_\varepsilon^\ell r_1^{\pm}(0) = \partial_\varepsilon^\ell r_2^{\pm}(0)$ and $\partial_\varepsilon^\ell z_1^{\pm}(0) = \partial_\varepsilon^\ell z_2^{\pm}(0)$. Thus $r_1^{\pm}(\varepsilon) - r_2^{\pm}(\varepsilon)$ and $z_1^{\pm}(\varepsilon) - z_2^{\pm}(\varepsilon)$ are $\mathcal{O}(\varepsilon^\infty)$, which implies that $\tilde{q}^{\pm}(x, \varepsilon)$ is $\mathcal{O}(\varepsilon^\infty)$ as well. □

Remark. If χ and ψ depend smoothly on some additional parameter y, then the proof of Lemma 3.6 shows that $r^{\pm}$, $z^{\pm}$, $q^{\pm}$ can be chosen to depend smoothly on y as well. Lemma 3.7 also holds, with convergence locally uniform in y, as does Lemma 3.8. In Section 4.6 below we use this to study expressions of the form

$$\chi(\theta, \theta', \varepsilon)(\theta - \theta' \pm i\varepsilon\psi(\theta, \theta', \varepsilon))^{-1}, \tag{3-34}$$

where (θ, θ') is in some neighborhood of 0 and we put $x := \theta - \theta'$, $y := \theta$.

For the use in Section 4 we record the fact that operators with Schwartz kernels of the form (3-34) are pseudodifferential:

Lemma 3.9. *Assume that $c_\varepsilon^{\pm}(\theta')$ and $z_\varepsilon^{\pm}(\theta')$ are complex-valued functions smooth in $\theta' \in \mathbb{S}^1 := \mathbb{R}/\mathbb{Z}$ and $\varepsilon \in [0, \varepsilon_0)$ and such that $\operatorname{Re} z_\varepsilon^{\pm} > 0$. Let $\chi \in C^\infty(\mathbb{S}^1 \times \mathbb{S}^1)$ be supported in a neighborhood of the diagonal and equal to 1 on a smaller neighborhood of the diagonal. Consider the operator $A_\varepsilon^{\pm}$ on $C^\infty(\mathbb{S}^1)$ given by*

$$\begin{aligned} A_\varepsilon^{\pm} f(\theta) &= \int_{\mathbb{S}^1} K_\varepsilon^{\pm}(\theta, \theta') f(\theta')\, d\theta', \\ K_\varepsilon^{\pm}(\theta, \theta') &= \begin{cases} \chi(\theta, \theta') c_\varepsilon^{\pm}(\theta')(\theta - \theta' \pm i\varepsilon z_\varepsilon^{\pm}(\theta'))^{-1}, & \varepsilon > 0, \\ \chi(\theta, \theta') c_\varepsilon^{\pm}(\theta')(\theta - \theta' \pm i0)^{-1}, & \varepsilon = 0. \end{cases} \end{aligned} \tag{3-35}$$

Then $A_\varepsilon^{\pm} \in \Psi^0(\mathbb{S}^1)$ uniformly in ε and we have, uniformly in ε,

$$\mathrm{WF}(A_\varepsilon^{\pm}) \subset \{\pm\xi > 0\}, \quad \sigma(A_\varepsilon^{\pm})(\theta, \xi) = \mp 2\pi i c_\varepsilon^{\pm}(\theta) e^{-\varepsilon z_\varepsilon^{\pm}(\theta)|\xi|} H(\pm\xi), \tag{3-36}$$

where for $\varepsilon > 0$ the principal symbol is understood as in (3-7) and H is the Heaviside function (with the symbol considered for $|\xi| > 1$).

Remark. We note that the definition (3-7) of the symbol of a family of operators and Lemma 3.8 show that the principal symbol is independent of the (not unique) $c_\varepsilon^{\pm}$, $z_\varepsilon^{\pm}$.

Proof. Using the formulas (3-26) and (3-32), we write the kernel $K_\varepsilon^{\pm}(\theta, \theta')$ as an oscillatory integral (where $a_\varepsilon(\theta, \theta', \xi)$ is supported near $\{\theta = \theta'\}$ where $\theta - \theta' \in \mathbb{R}$ is well-defined)

$$\begin{aligned} K_\varepsilon^{\pm}(\theta, \theta') &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{i(\theta - \theta')\xi} a_\varepsilon(\theta, \theta', \xi)\, d\xi, \\ a_\varepsilon(\theta, \theta', \xi) &= \mp 2\pi i \chi(\theta, \theta') c_\varepsilon^{\pm}(\theta') e^{-\varepsilon z_\varepsilon^{\pm}(\theta')|\xi|} H(\pm\xi). \end{aligned}$$

Fix a cutoff function $\tilde{\chi} \in C_c^\infty(\mathbb{R})$ such that $\tilde{\chi} = 1$ near 0 and split $a_\varepsilon = \tilde{\chi}(\xi) a_\varepsilon + (1 - \tilde{\chi}(\xi)) a_\varepsilon$. The integral corresponding to $\tilde{\chi}(\xi) a_\varepsilon$ gives a kernel which is in C^∞ in θ, θ', and $\varepsilon \in [0, \varepsilon_0)$. Next, $(1 - \tilde{\chi}(\xi)) a_\varepsilon$ is a symbol of class S^0: for each α, β there exists $C_{\alpha\beta}$ such that for all θ, θ', ξ and $\varepsilon \in [0, \varepsilon_0)$, we have

$$\left| \partial_{(\theta, \theta')}^\alpha \partial_\xi^\beta \big((1 - \tilde{\chi}(\xi)) a_\varepsilon(\theta, \theta', \xi) \big) \right| \le C_{\alpha\beta} \langle \xi \rangle^{-\beta}.$$

Therefore (see [Hörmander 1994, Lemma 18.2.1] or [Grigis and Sjöstrand 1994, Theorem 3.4]) we see that $A_\varepsilon^\pm \in \Psi^0$ uniformly in ε, its wavefront set is contained in $\{\pm\xi > 0\}$, and its principal symbol is the equivalence class of $a_\varepsilon(\theta, \theta, \xi)$. □

Remark. The fine analysis in Section 3.4 is strictly speaking not necessary for our application in Section 4.6: indeed, in the proof of Proposition 4.15 one could instead use a version of Lemma 3.9 which allows $c_\varepsilon^\pm$ and $z_\varepsilon^\pm$ to depend on both θ and θ'. Moreover, ultimately one just uses that the exponential in (3-36) is bounded in absolute value by 1; see the proof of Lemma 5.2. However, we feel that using the results of Section 3.4 leads to nicer expressions for the kernels of the restricted single layer potentials in Sections 4.6.5–4.6.6 which could be useful elsewhere.

4. Boundary layer potentials

In this section we describe microlocal properties of boundary layer potentials for the operator $P - \omega^2 = \partial_{x_2}^2 \Delta_\Omega^{-1} - \omega^2$, or rather for the related partial differential operator $P(\omega)$ defined in (4-1). The key issue is the transition from elliptic to hyperbolic behavior as $\operatorname{Im}\omega \to 0$. To motivate the results we explain the analogy with the standard boundary layer potentials in Section 4.2. In Section 4.3 we compute fundamental solutions for $P(\omega)$ on $\mathbb{R}^2$ and in Section 4.4 we use these to study the Dirichlet problem for $P(\omega)$ on Ω. This will lead us to single layer potentials: in Section 4.5 we study their mapping properties (in particular relating Lagrangian distributions on the boundary to Lagrangian distributions in the interior) and in Section 4.6 we give a microlocal description of their restriction to $\partial\Omega$ uniformly as $\operatorname{Im}\omega \to 0$, which is crucially used in Section 5.

In Sections 4–7 we generally use the letter λ to denote the spectral parameter when it is real and the letter ω for complex values of the spectral parameter, often taking the limit $\omega \to \lambda \pm i0$.

4.1. *Basic properties.*

Consider the second-order constant coefficient differential operator on $\mathbb{R}^2_{x_1,x_2}$

$$P(\omega) := (1 - \omega^2)\partial_{x_2}^2 - \omega^2 \partial_{x_1}^2, \quad \text{where } \omega \in \mathbb{C},\ 0 < \operatorname{Re}\omega < 1. \tag{4-1}$$

Formally,

$$P(\omega) = (P - \omega^2)\Delta_\Omega, \quad P(\omega)^{-1} = \Delta_\Omega^{-1}(P - \omega^2)^{-1}. \tag{4-2}$$

We note that $P(\omega)$ is hyperbolic when $\omega \in (0, 1)$ and elliptic otherwise. We factorize $P(\omega)$ as

$$P(\omega) = 4L_\omega^+ L_\omega^-, \quad L_\omega^\pm := \tfrac{1}{2}(\pm\omega\, \partial_{x_1} + \sqrt{1 - \omega^2}\, \partial_{x_2}). \tag{4-3}$$

Here $\sqrt{1 - \omega^2}$ is defined by taking the branch of the square root on $\mathbb{C} \setminus (-\infty, 0]$ which takes positive values on $(0, \infty)$. We note that for $\lambda \in (0, 1)$ the operators $L_\lambda^\pm$ are two linearly independent constant vector fields on $\mathbb{R}^2$. For $\operatorname{Im}\omega \neq 0$, $L_\omega^\pm$ are Cauchy–Riemann-type operators.

The definition (1-2) of the functions $\ell^\pm(x, \lambda)$ extends to complex values of λ:

$$\ell^\pm(x, \omega) := \pm\frac{x_1}{\omega} + \frac{x_2}{\sqrt{1 - \omega^2}}, \quad x = (x_1, x_2) \in \mathbb{R}^2,\ \omega \in \mathbb{C},\ 0 < \operatorname{Re}\omega < 1.$$

Then the linear functions $\ell^\pm(\bullet, \omega)$ are dual to the operators $L^\pm_\omega$:

$$L^\pm_\omega \ell^\pm(x,\omega) = 1, \quad L^\mp_\omega \ell^\pm(x,\omega) = 0. \tag{4-4}$$

We record here the following statement:

Lemma 4.1. *Assume that* $\operatorname{Im}\omega > 0$. *Then the map* $x \in \mathbb{R}^2 \mapsto \ell^\pm(x,\omega) \in \mathbb{C}$ *is orientation-preserving in the case of* ℓ^+ *and orientation reversing in the case of* ℓ^-. *If* $\operatorname{Im}\omega < 0$ *then a similar statement holds with the roles of* $\ell^\pm$ *switched.*

Proof. This follows immediately from the sign identity

$$\operatorname{sgn}\operatorname{Im}\frac{\omega}{\sqrt{1-\omega^2}} = \operatorname{sgn}\operatorname{Im}\omega, \quad \omega \in \mathbb{C},\ 0 < \operatorname{Re}\omega < 1, \tag{4-5}$$

which can be verified by noting that $\operatorname{Re}\sqrt{1-\omega^2} > 0$ and $\operatorname{sgn}\operatorname{Im}\sqrt{1-\omega^2} = -\operatorname{sgn}\operatorname{Im}\omega$. □

4.2. ***Motivational discussion.*** When $\operatorname{Im}\omega > 0$ the decomposition (4-3) is similar to the factorization of the Laplacian,

$$\Delta = 4\partial_z\partial_{\bar z}, \quad \partial_z = \tfrac{1}{2}(\partial_x - i\partial_y), \quad \partial_{\bar z} = \tfrac{1}{2}(\partial_x + i\partial_y).$$

The functions $z = x + iy$ and $\bar z$ play the role of $\ell^\pm(x,\omega)$ for $\pm$ respectively (which matches the orientation in Lemma 4.1) and $\partial_z, \partial_{\bar z}$ play the role of L^+_ω, L^-_ω. Hence to explain the structure of the fundamental solution of $P(\omega)$ and to motivate the restricted boundary layer potential in Section 4.6 we review the basic case when $\Omega = \{y > 0\}$ and $P(\omega)$ is replaced by Δ. The fundamental solution is given by (see, e.g., [Hörmander 1990, Theorem 3.3.2])

$$\Delta E = \delta_0, \quad E := c\log(z\bar z), \quad \partial_z E = \frac{c}{z}, \quad \partial_{\bar z} E = \frac{c}{\bar z}, \quad c = \frac{1}{4\pi}.$$

We consider the *single layer* potential $S : C^\infty_c(\mathbb{R}) \to \mathcal{D}'(\mathbb{R}^2) \cap C^\infty(\mathbb{R}^2 \setminus \{y = 0\})$,

$$Sv(x,y) := \int_{\mathbb{R}} E(x - x', y)v(x')\,dx', \quad v \in C^\infty_c(\mathbb{R}).$$

We then have limits as $y \to 0\pm$,

$$\mathcal{C}_\pm v(x) = \mathcal{C}v(x) := \frac{1}{2\pi}\int_{\mathbb{R}} \log|x - x'|v(x')\,dx',$$

and we consider

$$\partial_x \mathcal{C}v(x) = \lim_{y\to 0\pm} \partial_x Sv(x,y) = \lim_{y\to 0\pm}(\partial_z + \partial_{\bar z})Sv(x,y).$$

Then $\big($where we recall $c = \frac{1}{4\pi}\big)$

$$\lim_{y\to 0\pm} \partial_z Sv(x,y) = \lim_{y\to 0\pm}\int_{\mathbb{R}} \frac{c}{x - x' + iy}v(x')\,dx' = \int_{\mathbb{R}} \frac{c}{x - x' \pm i0}v(x')\,dx,$$

and similarly,

$$\lim_{y\to 0\pm} \bar\partial_z Sv(x,y) = \lim_{y\to 0\pm}\int_{\mathbb{R}} \frac{c}{x - x' - iy}v(x')\,dx' = \int_{\mathbb{R}} \frac{c}{x - x' \mp i0}v(x')\,dx,$$

where the right-hand sides are understood as distributional pairings. This gives

$$\partial_x \mathcal{C} v(x) = \frac{1}{4\pi} \int_{\mathbb{R}} \sum_{\pm} (x - x' \pm i0)^{-1} v(x')\, dx', \tag{4-6}$$

which is $\frac{1}{2}$ times the Hilbert transform, that is, the Fourier multiplier with symbol $-i\,\mathrm{sgn}(\xi)$. (We note that $\sum_{\pm}(x - x' \pm i0)^{-1} = 2\partial_x \log|x|$ which is the principal value of $2/x$.)

In Section 4.6 we describe the analogue of $\partial_x \mathcal{C}$ in our case. It is similar to (4-6) when $\operatorname{Im}\omega > 0$ but when $\operatorname{Im}\omega \to 0+$, it has additional singularities described using the chess billiard map $b(x, \lambda)$, $\lambda = \operatorname{Re}\omega$, or rather its building components $\gamma^{\pm}$. The operator becomes an elliptic operator of order 0 (just as is the case in (4-6) if we restrict our attention to compact sets) plus a Fourier integral operator — see Proposition 4.15.

4.3. *Fundamental solutions.* We now construct a *fundamental solution* of the operator $P(\omega)$ defined in (4-1), that is, a distribution $E_\omega \in \mathcal{D}'(\mathbb{R}^2)$ such that

$$P(\omega) E_\omega = \delta_0. \tag{4-7}$$

For that we use the complex-valued quadratic form

$$A(x, \omega) := \ell^+(x, \omega)\ell^-(x, \omega) = -\frac{x_1^2}{\omega^2} + \frac{x_2^2}{1 - \omega^2}.$$

Since $0 < \operatorname{Re}\omega < 1$, we have $\operatorname{sgn}\operatorname{Im}(-\omega^{-2}) = \operatorname{sgn}\operatorname{Im}((1 - \omega^2)^{-1}) = \operatorname{sgn}\operatorname{Im}\omega$; thus

$$\operatorname{sgn}\operatorname{Im} A(x, \omega) = \operatorname{sgn}\operatorname{Im}\omega \quad \text{for all } x \in \mathbb{R}^2 \setminus \{0\}. \tag{4-8}$$

4.3.1. *The nonreal case.* We first consider the case $\operatorname{Im}\omega \neq 0$. In this case our fundamental solution is the locally integrable function

$$E_\omega(x) := c_\omega \log A(x, \omega), \quad x \in \mathbb{R}^2 \setminus \{0\}, \quad c_\omega := \frac{i\,\operatorname{sgn}\operatorname{Im}\omega}{4\pi\omega\sqrt{1 - \omega^2}}. \tag{4-9}$$

Here we use the branch of logarithm on $\mathbb{C} \setminus (-\infty, 0]$ which takes real values on $(0, \infty)$. Note that the function E_ω is smooth on $\mathbb{R}^2 \setminus \{0\}$.

Lemma 4.2. *The function E_ω defined in (4-9) solves (4-7).*

Proof. We first check that $P(\omega)E_\omega = 0$ on $\mathbb{R}^2 \setminus \{0\}$: this follows from (4-3), (4-4), and the identities

$$L_\omega^{\pm} \log A(x, \omega) = \frac{1}{\ell^{\pm}(x, \omega)} \quad \text{for all } x \in \mathbb{R}^2 \setminus \{0\}. \tag{4-10}$$

Next, denote by B_ε the ball of radius $\varepsilon > 0$ centered at 0 and orient ∂B_ε in the counterclockwise direction. Using the divergence theorem twice, we compute for each $\varphi \in C_c^\infty(\mathbb{R}^2)$

$$\begin{aligned}\int_{\mathbb{R}^2} E_\omega(x)(P(\omega)\varphi(x))\, dx &= 4 \lim_{\varepsilon\to 0+} \int_{\mathbb{R}^2\setminus B_\varepsilon} E_\omega(x)(L_\omega^+ L_\omega^- \varphi(x))\, dx = -4c_\omega \lim_{\varepsilon\to 0+} \int_{\mathbb{R}^2\setminus B_\varepsilon} \frac{L_\omega^- \varphi(x)}{\ell^+(x, \omega)}\, dx \\ &= -2c_\omega \lim_{\varepsilon\to 0+} \int_{\partial B_\varepsilon} \frac{\varphi(x)(\sqrt{1 - \omega^2}\, dx_1 + \omega\, dx_2)}{\ell^+(x, \omega)} \\ &= -2c_\omega \omega\sqrt{1 - \omega^2}\, \varphi(0) \lim_{\varepsilon\to 0+} \int_{\partial B_\varepsilon} \frac{d\ell^+(x, \omega)}{\ell^+(x, \omega)} = \varphi(0)\end{aligned}$$

which gives (4-7). Here in the last equality we make the change of variables $z = \ell^+(x, \omega)$ and use Lemma 4.1. ☐

4.3.2. *The real case.* We now discuss the case $\lambda \in (0, 1)$. Define the fundamental solutions $E_{\lambda \pm i0} \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$ as

$$E_{\lambda\pm i0}(x) := \pm c_\lambda \log(A(x,\lambda) \pm i0), \quad c_\lambda := \frac{i}{4\pi\lambda\sqrt{1-\lambda^2}},$$
$$\log(A(x,\lambda)\pm i0) = \begin{cases} \log A(x,\lambda), & A(x,\lambda) > 0, \\ \log(-A(x,\lambda)) \pm i\pi, & A(x,\lambda) < 0. \end{cases} \tag{4-11}$$

The next lemma shows that $E_{\lambda\pm i\varepsilon} \to E_{\lambda\pm i0}$ in $\mathcal{D}'(\mathbb{R}^2)$ as $\varepsilon \to 0+$. In fact it gives a stronger convergence statement with derivatives in λ. To make this statement we introduce the following notation: if $\mathcal{J} \subset (0, 1)$ is an open interval then

$$\mathscr{O}(\mathcal{J} + i[0,\infty)) \subset C^\infty(\mathcal{J} + i[0,\infty)) \tag{4-12}$$

consists of C^∞ functions on $\mathcal{J} + i[0, \infty)$ which are holomorphic in the interior $\mathcal{J} + i(0, \infty)$. Similarly one can define $\mathscr{O}(\mathcal{J} - i[0, \infty))$.

Lemma 4.3. *The maps*

$$\omega \in (0,1) \pm i[0,\infty) \mapsto \begin{cases} E_\omega, & \operatorname{Im}\omega \neq 0, \\ E_{\lambda\pm i0}, & \omega = \lambda \in (0,1), \end{cases} \tag{4-13}$$

lie in $\mathscr{O}((0, 1) \pm i[0, \infty); \mathcal{D}'(\mathbb{R}^2))$ in the following sense: the distributional pairing of (4-13) with any $\varphi \in C^\infty_c(\mathbb{R}^2)$ lies in $\mathscr{O}((0, 1) \pm i[0, \infty))$.

Proof. We consider the case of $\operatorname{Im}\omega \geq 0$, with the case $\operatorname{Im}\omega \leq 0$ handled similarly. Fix $\varphi \in C^\infty_c(\mathbb{R}^2)$.

(1) We will prove the following limiting statement: for each $\lambda \in (0, 1)$

$$\int_{\mathbb{R}^2} E_{\omega_j}(x)\varphi(x)\,dx \to \int_{\mathbb{R}^2} E_{\lambda+i0}(x)\varphi(x)\,dx \quad \text{for all } \omega_j \to \lambda,\ \operatorname{Im}\omega_j > 0. \tag{4-14}$$

We write $\omega_j = \lambda_j + i\varepsilon_j$, where $\lambda_j \to \lambda$ and $\varepsilon_j \to 0+$. We first show a bound on $E_{\omega_j}(x) = c_{\omega_j} \log A(x, \omega_j)$ which is uniform in j. Taking the Taylor expansion of $\ell^+(x, \lambda + i\varepsilon)$ in ε, we get

$$\ell^+(x,\omega_j) = \ell^+(x,\lambda_j) + i\varepsilon_j \partial_\lambda \ell^+(x,\lambda_j) + \mathcal{O}(\varepsilon_j^2|x|),$$

where the constant in $\mathcal{O}(\bullet)$ is independent of j. Since $\partial_\lambda \ell^+(x, \lambda_j)$ is real, we bound

$$|\ell^+(x,\omega_j)| \geq \tfrac{1}{2}(|\ell^+(x,\lambda_j)| + \varepsilon_j|\partial_\lambda\ell^+(x,\lambda_j)|) - C\varepsilon_j^2|x|. \tag{4-15}$$

As $\ell^+(x, \lambda_j)$, $\partial_\lambda \ell^+(x, \lambda_j)$ are linearly independent linear forms in x (see (2-7)), we have

$$|x| \leq C(|\ell^+(x,\lambda_j)| + |\partial_\lambda\ell^+(x,\lambda_j)|). \tag{4-16}$$

Together (4-15) and (4-16) show that for j large enough

$$|\ell^+(x,\omega_j)| \geq \tfrac{1}{3}|\ell^+(x,\lambda_j)|. \tag{4-17}$$

The same bound holds for ℓ^-. Since $A(x,\omega_j)=\ell^+(x,\omega_j)\ell^-(x,\omega_j)$, we then have

$$C^{-1}|A(x,\lambda_j)|\le|A(x,\omega_j)|\le C|x|^2,$$

which implies the following bound for j large enough, some j-independent constant C, and all $x\in\mathbb{R}^2$:

$$|E_{\omega_j}(x)|\le C(|E_{\lambda_j+i0}(x)|+\log(2+|x|)). \tag{4-18}$$

(2) To pin the zero set of $A(x,\lambda_j)$, which depends on λ_j, we introduce the linear isomorphism $\Phi_\lambda:\mathbb{R}^2_y\to\mathbb{R}^2_x$ such that $\Phi_\lambda^{-1}(x)=(\ell^+(x,\lambda),\ell^-(x,\lambda))$. Then $A(\Phi_\lambda(y),\lambda)=y_1y_2$, so the pullback of E_{λ_j+i0} by Φ_{λ_j} is given by

$$\Phi_{\lambda_j}^*E_{\lambda_j+i0}(y)=c_{\lambda_j}\log(y_1y_2+i0), \tag{4-19}$$

which is a locally integrable function on $\mathbb{R}^2$.

We can now show (4-14). For each $y\in\mathbb{R}^2$, we have $A(\Phi_{\lambda_j}(y),\omega_j)\to A(\Phi_\lambda(y),\lambda)=y_1y_2$ and $\varphi(\Phi_{\lambda_j}(y))\to\varphi(\Phi_\lambda(y))$. Using (4-8) we then get the pointwise limit

$$\Phi_{\lambda_j}^*(E_{\omega_j}\varphi)(y)\to\Phi_\lambda^*(E_{\lambda+i0}\varphi)(y)\quad\text{for all } y\in\mathbb{R}^2,\ y_1y_2\neq 0.$$

Now (4-14) follows from the dominated convergence theorem applied to the sequence of functions $\Phi_{\lambda_j}^*(E_{\omega_j}\varphi)$, where the dominant is given by the locally integrable function $C(1+|\log(y_1y_2+i0)|)$ as follows from the bound (4-18) and the identity (4-19).

(3) Denote by $F_\varphi(\omega)$ the pairing of (4-13) with φ. Since $A(x,\omega)$ is a quadratic form depending holomorphically on $\omega\in(0,1)+i(0,\infty)$, which has a positive definite imaginary part by (4-8), we see that F_φ is holomorphic on $(0,1)+i(0,\infty)$. Moreover, the restriction of F_φ to $(0,1)$ is smooth, as can be seen by writing

$$F_\varphi(\lambda)=\int_{\mathbb{R}^2}E_{\lambda+i0}(x)\varphi(x)\,dx=c_\lambda|\det\Phi_\lambda|\int_{\mathbb{R}^2}\log(y_1y_2+i0)\varphi(\Phi_\lambda(y))\,dy,\qquad\lambda\in(0,1),$$

and using that the function $(y,\lambda)\mapsto\varphi(\Phi_\lambda(y))$ is smooth in (y,λ). By (4-14) F_φ is continuous at the boundary interval $(0,1)$. Since F_φ is holomorphic, it is harmonic, so by boundary regularity for the Dirichlet problem for the Laplacian (see the references in the proof of Lemma 4.4 below) we see that $F_\varphi\in C^\infty((0,1)+i[0,\infty))$. □

Passing to the limit in (4-7) we see that

$$P(\lambda)E_{\lambda\pm i0}=\delta_0\quad\text{for all }\lambda\in(0,1). \tag{4-20}$$

Note that $E_{\lambda\pm i0}(x)$ is smooth except on the union of the two lines $\{\ell^+(x,\lambda)=0\}$ and $\{\ell^-(x,\lambda)=0\}$. We remark that $E_{\lambda\pm i0}$ are the Feynman propagators in dimension 1; see [Hörmander 1990, (6.2.1) and p. 141] for the formula in all dimensions.

4.4. ***Reduction to the boundary.*** We now let $\Omega\subset\mathbb{R}^2$ be a bounded open set with C^∞ boundary and consider the elliptic boundary value problem

$$P(\omega)u=f,\quad u|_{\partial\Omega}=0,\quad \operatorname{Re}\omega\in(0,1),\quad \operatorname{Im}\omega\neq 0. \tag{4-21}$$

Lemma 4.4. *For each $f \in C_c^\infty(\Omega)$, the problem (4-21) has a unique solution $u \in C^\infty(\bar{\Omega})$.*

Remark. The proof shows that if f is fixed, then $u \in C^\infty(\bar{\Omega})$ depends holomorphically on $\omega \in (0, 1) \pm i(0, \infty)$.

Proof. (1) We first show that for each $\mu \in \mathbb{C} \setminus [1, \infty)$ and $s \geq 2$, the map

$$\bar{H}^s(\Omega) \ni u \mapsto ((\Delta - \mu\partial_{x_2}^2)u, u|_{\partial\Omega}) \in \bar{H}^{s-2}(\Omega) \oplus H^{s-1/2}(\partial\Omega) \tag{4-22}$$

is a Fredholm operator. (Here $\bar{H}^s(\Omega)$ denotes the space of distributions on Ω which extend to H^s distributions on $\mathbb{R}^2$.) We apply [Hörmander 1994, Theorem 20.1.2]. The operator $\Delta - \mu\partial_{x_2}^2$ is elliptic, so it remains to verify that the Shapiro–Lopatinski condition [Hörmander 1994, Definition 20.1.1(ii)] holds for any domain Ω. (An example of an operator for which this condition fails is $(\partial_{x_1} + i\partial_{x_2})^2$.) In our specific case the Shapiro–Lopatinski condition can be reformulated as follows: for each basis (ξ, η) of $\mathbb{R}^2$, if we denote by $\mathcal{M}$ the space of all bounded solutions on $[0, \infty)$ to the ODE

$$p(\xi - i\eta\partial_t)u(t) = 0, \quad p(\xi) := \xi_1^2 + (1-\mu)\xi_2^2$$

then the map $u \in \mathcal{M} \mapsto u(0)$ is an isomorphism. This is equivalent to the requirement that the quadratic equation $p(\xi + z\eta) = 0$ have two roots, one with $\operatorname{Im} z > 0$ and one with $\operatorname{Im} z < 0$. To see that the latter condition holds, we argue by continuity: since $\Delta - \mu\partial_{x_2}^2$ is elliptic, the equation $p(\xi + z\eta) = 0$ cannot have any real roots z, so the condition either holds for all μ, ξ, η or fails for all μ, ξ, η. However, it is straightforward to check that the condition holds when $\mu = 0$, $\xi = (1, 0)$, $\eta = (0, 1)$, as the roots are $\pm i$.

(2) We next claim that the Fredholm operator (4-22) is invertible. We first show that it has index 0, arguing by continuity: since the operator (4-22) is continuous in μ in the operator norm topology, its index should be independent of μ. However, for $\mu = 0$ we get the Dirichlet problem for the Laplacian, where (4-22) is invertible.

To show that (4-22) is invertible it remains to prove injectivity, namely

$$u \in H^2(\Omega), \ \ (\Delta - \mu\partial_{x_2}^2)u = 0, \ \ u|_{\partial\Omega} = 0 \quad \Longrightarrow \quad u = 0. \tag{4-23}$$

Multiplying the equation $(\Delta - \mu\partial_{x_2}^2)u = 0$ by $\bar{u}$ and integrating by parts over Ω, we get $\|\nabla u\|_{L^2(\Omega)}^2 = \mu\|\partial_{x_2}u\|_{L^2(\Omega)}^2$. Since $0 \leq \|\partial_{x_2}u\|_{L^2(\Omega)}^2 \leq \|\nabla u\|_{L^2(\Omega)}^2$ and $\mu \notin [1, \infty)$, we see that $\|\nabla u\|_{L^2(\Omega)} = 0$, which implies that $u = 0$, giving (4-23).

(3) Writing

$$P(\omega) = \partial_{x_2}^2 - \omega^2\Delta = -\omega^2(\Delta - \mu\partial_{x_2}^2), \quad \mu := \omega^{-2} \in \mathbb{C} \setminus [1, \infty),$$

and using the invertibility of (4-22), we see that, for each $s \geq 2$ and $f \in \bar{H}^{s-2}(\Omega)$, the problem (4-21) has a unique solution $u \in \bar{H}^s(\Omega)$. When $f \in C_c^\infty(\Omega)$, we may take an arbitrary s which gives that $u \in C^\infty(\bar{\Omega})$. $\square$

We will next express the solution to (4-21) in terms of boundary data and single layer potentials. Let us first define the operators used below. Let $T^*\partial\Omega$ be the cotangent bundle of the boundary $\partial\Omega$. Sections of this bundle are differential 1-forms on $\partial\Omega$ (where we use the positive orientation on $\partial\Omega$); they can be identified with functions on $\partial\Omega$ by fixing a coordinate θ. Define the operator $\mathcal{I}:\mathcal{D}'(\partial\Omega;T^*\partial\Omega)\to\mathcal{E}'(\mathbb{R}^2)$ as follows: for $v\in\mathcal{D}'(\partial\Omega;T^*\partial\Omega)$ and $\varphi\in C^\infty(\mathbb{R}^2)$,

$$\int_{\mathbb{R}^2}\mathcal{I}v(x)\varphi(x)\,dx:=\int_{\partial\Omega}\varphi v. \tag{4-24}$$

Note that $\operatorname{supp}(\mathcal{I}v)\subset\partial\Omega$ and we can think of $\mathcal{I}v$ as multiplying v by the delta function on $\partial\Omega$. Next, let E_ω be the fundamental solution constructed in (4-9) and define the convolution operator

$$R_\omega:\mathcal{E}'(\mathbb{R}^2)\to\mathcal{D}'(\mathbb{R}^2),\qquad R_\omega g:=E_\omega * g. \tag{4-25}$$

Using the limiting fundamental solutions $E_{\lambda\pm i0}$ constructed in (4-11), we similarly define the operators $R_{\lambda\pm i0}$ for $\lambda\in(0,1)$ which will be used later. Finally, for $\omega\in(0,1)+i\mathbb{R}$, define the "Neumann data" operator

$$\mathcal{N}_\omega:C^\infty(\bar\Omega)\to C^\infty(\partial\Omega;T^*\partial\Omega),\quad \mathcal{N}_\omega u:=-2\omega\sqrt{1-\omega^2}\,\boldsymbol{j}^*(L^+_\omega u\,d\ell^+(\bullet,\omega)), \tag{4-26}$$

where $\boldsymbol{j}:\partial\Omega\to\bar\Omega$ is the embedding map and $\boldsymbol{j}^*$ is the pullback on 1-forms. We can now reduce the problem (4-21) to the boundary:

Lemma 4.5. *Assume that $u\in C^\infty(\bar\Omega)$ is the solution to (4-21) for some $f\in C^\infty_c(\Omega)$. Put $U:=\mathbb{1}_\Omega u\in\mathcal{E}'(\mathbb{R}^2)$ and $v:=\mathcal{N}_\omega u$. Then*

$$P(\omega)U=f-\mathcal{I}v, \tag{4-27}$$
$$U=R_\omega f-R_\omega\mathcal{I}v. \tag{4-28}$$

Remark. Note that we also have

$$v=2\omega\sqrt{1-\omega^2}\,\boldsymbol{j}^*(L^-_\omega u\,d\ell^-(\bullet,\omega)).$$

Indeed, $0=\boldsymbol{j}^*du=\boldsymbol{j}^*(L^+_\omega u\,d\ell^++L^-_\omega u\,d\ell^-)$ since $u|_{\partial\Omega}=0$ and by (4-4).

Proof. Let $\varphi\in C^\infty_c(\mathbb{R}^2)$. Then by (4-3)

$$\begin{aligned}\int_{\mathbb{R}^2}(P(\omega)U)\varphi\,dx&=4\int_\Omega uL^+_\omega L^-_\omega\varphi\,dx=-4\int_\Omega(L^+_\omega u)(L^-_\omega\varphi)\,dx\\&=\int_\Omega f\varphi\,dx-4\int_\Omega L^-_\omega(\varphi L^+_\omega u)\,dx\\&=\int_\Omega f\varphi\,dx+2\omega\sqrt{1-\omega^2}\int_{\partial\Omega}\varphi L^+_\omega u\,d\ell^+=\int_\Omega(f-\mathcal{I}v)\varphi\,dx,\end{aligned}$$

which gives (4-27). The identity (4-28) follows from (4-27), the fundamental solution equation (4-7), and the fact that U is a compactly supported distribution: $E_\omega * P(\omega)U=(P(\omega)E_\omega)*U=U$. □

In the notation of Lemma 4.5, define $S_\omega v:=(R_\omega\mathcal{I}v)|_\Omega\in\mathcal{D}'(\Omega)$. Then (4-28) implies that

$$u=(R_\omega f)|_\Omega-S_\omega v. \tag{4-29}$$

Since $R_\omega f \in C^\infty(\mathbb{R}^2)$, we have $S_\omega v \in C^\infty(\overline{\Omega})$. Denote by $\mathcal{C}_\omega v := (S_\omega v)|_{\partial\Omega}$ its boundary trace; then the boundary condition $u|_{\partial\Omega} = 0$ gives the following equation on v:

$$\mathcal{C}_\omega v = (R_\omega f)|_{\partial\Omega}. \tag{4-30}$$

This motivates the study of the operator S_ω in Section 4.5 and of the operator $\mathcal{C}_\omega$ in Section 4.6.

4.5. ***Single layer potentials.*** We now introduce single layer potentials. For $\omega \in \mathbb{C}$ with $0 < \operatorname{Re}\omega < 1$ and $\operatorname{Im}\omega \neq 0$ the single layer potential is the operator $S_\omega : \mathcal{D}'(\partial\Omega; T^*\partial\Omega) \to \mathcal{D}'(\Omega)$ given by

$$S_\omega v := (E_\omega * \mathcal{I}v)|_\Omega, \quad v \in \mathcal{D}'(\partial\Omega; T^*\partial\Omega). \tag{4-31}$$

Here $E_\omega \in \mathcal{D}'(\mathbb{R}^2)$ is the fundamental solution defined in (4-9) and the operator $\mathcal{I} : \mathcal{D}'(\partial\Omega; T^*\partial\Omega) \to \mathcal{E}'(\mathbb{R}^2)$ is defined in (4-24). Similarly, if $\lambda \in (0, 1)$ and Ω is λ-simple (see Definition 1.1) then we can define operators

$$S_{\lambda \pm i0} : \mathcal{D}'(\partial\Omega; T^*\partial\Omega) \to \mathcal{D}'(\Omega) \tag{4-32}$$

by the formula (4-31), using the limiting distributions $E_{\lambda\pm i0}$ defined in (4-11).

If we fix a positively oriented coordinate θ on $\partial\Omega$ and use it to identify $\mathcal{D}'(\partial\Omega; T^*\partial\Omega)$ with $\mathcal{D}'(\partial\Omega)$, then the action of S_ω on smooth functions is given by

$$S_\omega(f\, d\theta)(x) = \int_{\partial\Omega} E_\omega(x-y) f(y)\, d\theta(y), \quad f \in C^\infty(\partial\Omega),\ x \in \Omega, \tag{4-33}$$

and similarly for $S_{\lambda\pm i0}$.

We now discuss the mapping properties of S_ω, in particular showing that $S_\omega v, S_{\lambda\pm i0} v \in C^\infty(\overline{\Omega})$ when $v \in C^\infty(\partial\Omega; T^*\partial\Omega)$. We break the latter into two cases:

4.5.1. *The nonreal case.* We first consider the case $\operatorname{Im}\omega \neq 0$. We use the following standard result, which is a version of the Sochocki–Plemelj theorem:

Lemma 4.6. *Assume that $\Omega_0 \subset \mathbb{C}$ is a bounded open set with C^∞ boundary (oriented in the positive direction). For $f \in C^\infty(\partial\Omega_0)$, define $u \in C^\infty(\Omega_0)$ by*

$$u(z) = \int_{\partial\Omega_0} \frac{f(w)\, dw}{w - z}, \quad z \in \Omega_0.$$

Then u extends smoothly to the boundary and the operator $f \mapsto u$ is continuous $C^\infty(\partial\Omega_0) \to C^\infty(\overline{\Omega}_0)$.

Remark. In the (unbounded) model case $\Omega_0 = \{\operatorname{Im} z > 0\}$, we have for each $f \in C^\infty_c(\mathbb{R})$

$$u(x+iy) = \int_{\mathbb{R}} \frac{f(t)\, dt}{t - x - iy}, \quad y > 0.$$

We see in particular that the function $x \mapsto \lim_{y\to 0+} \partial_y^k u(x+iy)$ is given by the convolution of f with $(-1)^{k+1} i^k k!\, (x+i0)^{-k-1}$.

Proof. Let $\tilde f \in C_c^\infty(\mathbb{C})$ be an almost analytic extension of f; that is, $\tilde f|_{\partial\Omega_0} = f$ and $\partial_{\bar z}\tilde f$ vanishes to infinite order on $\partial\Omega_0$. (See for example [Dyatlov and Zworski 2019a, Lemma 4.30] for the existence of such an extension.) Denote by dm the Lebesgue measure on $\mathbb{C}$. By the Cauchy–Green formula (see for instance [Hörmander 1990, (3.1.11)]), we have

$$u(z) = 2\pi i \tilde f(z) + 2i \int_{\Omega_0} \frac{\partial_{\bar w}\tilde f(w)}{w-z}\, dm(w), \qquad z \in \Omega_0,$$

and this extends smoothly to $z \in \mathbb{C}$: indeed, the second term on the right-hand side is the convolution of the distribution $-2iz^{-1} \in L^1_{\mathrm{loc}}(\mathbb{C})$, with $\mathbb{1}_{\Omega_0}\partial_{\bar z}\tilde f \in C_c^\infty(\mathbb{C})$. □

We now come back to the mapping properties of single layer potentials:

Lemma 4.7. *Assume* $0 < \operatorname{Re}\omega < 1$ *and* $\operatorname{Im}\omega \neq 0$. *Then* S_ω *is a continuous operator from* $C^\infty(\partial\Omega; T^*\partial\Omega)$ *to* $C^\infty(\bar\Omega)$.

Remark. With more work, it is possible to show that S_ω is actually continuous $C^\infty(\partial\Omega; T^*\partial\Omega) \to C^\infty(\bar\Omega)$ *uniformly* as $\operatorname{Im}\omega \to 0$, with limits being the operators $S_{\lambda\pm i0}$, $\lambda = \operatorname{Re}\omega$. However, our proof of Lemma 4.7 only shows the mapping property for any fixed nonreal ω. This is enough for our purposes since we have weak convergence of $S_{\lambda\pm i\varepsilon}$ to $S_{\lambda\pm i0}$ (Lemma 4.3; see also Lemmas 4.10 and 4.16 below) and in Section 4.6 we analyze the behavior of the *restricted* single layer potentials uniformly as $\operatorname{Im}\omega \to 0$.

Proof. Let $v \in C^\infty(\partial\Omega; T^*\partial\Omega)$. Since E_ω is smooth on $\mathbb{R}^2 \setminus \{0\}$ and $\mathcal{I}v$ is supported on $\partial\Omega$, we have $S_\omega v \in C^\infty(\Omega)$. It remains to show that $S_\omega v$ is smooth up to the boundary, and for this it is enough to verify the smoothness of the derivatives $L_\omega^\pm S_\omega v$, where $L_\omega^\pm$ are defined in (4-3). By (4-10) we have (suppressing the dependence of $\ell^\pm$ on ω in the notation)

$$L_\omega^\pm S_\omega v(x) = c_\omega \int_{\partial\Omega} \frac{v(y)}{\ell^\pm(x-y)}, \qquad x \in \Omega.$$

Since $\operatorname{Im}\omega \neq 0$, the maps $x \mapsto \ell^\pm(x)$ are linear isomorphisms from $\mathbb{R}^2$ onto $\mathbb{C}$ (considered as a real vector space). Using this we write

$$L_\omega^\pm S_\omega v(x) = \pm\operatorname{sgn}(\operatorname{Im}\omega) c_\omega \int_{\partial\Omega_\pm} \frac{f_\pm(w)\,dw}{z-w}, \qquad z := \ell^\pm(x) \in \Omega_\pm, \tag{4-34}$$

where we put $\Omega_\pm := \ell^\pm(\Omega) \subset \mathbb{C}$ and define the functions $f_\pm \in C^\infty(\partial\Omega_\pm)$ by the equality of differential forms $v(y) = f_\pm(\ell^\pm(y))\, d\ell^\pm(y)$ on $\partial\Omega$. Here $\partial\Omega_\pm$ are positively oriented and the sign factor $\pm\operatorname{sgn}(\operatorname{Im}\omega)$ accounts for the orientation of the map $\ell^\pm$; see Lemma 4.1.

Now $L_\omega^\pm S_\omega v$ extends smoothly to the boundary by Lemma 4.6. □

4.5.2. *The real case.* We now consider the case $\lambda \in (0, 1)$:

Lemma 4.8. *Assume that* $\lambda \in (0, 1)$ *and* Ω *is* λ*-simple* (*see Definition 1.1*). *Then* $S_{\lambda\pm i0}$ *are continuous operators from* $C^\infty(\partial\Omega; T^*\partial\Omega)$ *to* $C^\infty(\bar\Omega)$.

Proof. (1) We focus on the operator $S_{\lambda+i0}$, noting that $S_{\lambda-i0}$ is related to it by the identity

$$S_{\lambda-i0}\bar v = \overline{S_{\lambda+i0} v} \quad \text{for all } v \in \mathcal{D}'(\partial\Omega; T^*\partial\Omega).$$

We again suppress the dependence on λ in the notation, writing simply $\ell^\pm(x)$ and $A(x)$. Denoting by $H(x)=\mathbb{1}_{(0,\infty)}(x)$ the Heaviside function, we can rewrite (4-11) as

$$\log(A(x)+i0)=\log|\ell^+(x)|+\log|\ell^-(x)|+i\pi H(-A(x)).$$

We then take the decomposition

$$S_{\lambda+i0}=c_\lambda(S_\lambda^+ + S_\lambda^- + i\pi S_\lambda^0), \tag{4-35}$$

where for all $x\in\Omega$ and $v\in C^\infty(\partial\Omega;T^*\partial\Omega)$

$$S_\lambda^\pm v(x)=\int_{\partial\Omega}\log|\ell^\pm(x-y)|v(y),$$
$$S_\lambda^0 v(x)=\int_{\partial\Omega}H(-A(x-y))v(y).$$

(2) Let $v\in C^\infty(\partial\Omega;T^*\partial\Omega)$. Fix a positively oriented coordinate θ on $\partial\Omega$ and write $v=f\,d\theta$ for some $f\in C^\infty(\partial\Omega)$. We first analyze $S_\lambda^\pm v$, writing it as

$$S_\lambda^\pm v(x)=g_\pm(\ell^\pm(x)),\quad g_\pm(t):=\int_{\mathbb{R}}(\Pi_\lambda^\pm f)(s)\log|t-s|\,ds,$$

where $\Pi_\lambda^\pm f\in\mathcal{E}'(\mathbb{R})$ are the pushforwards of f by the maps $\ell^\pm$ defined in (2-10). Let $\ell^\pm_{\min}<\ell^\pm_{\max}$ be defined in (2-9). By part (1) of Lemma 2.3, $\Pi_\lambda^\pm f$ is supported in $[\ell^\pm_{\min},\ell^\pm_{\max}]$ and

$$\sqrt{(s-\ell^\pm_{\min})(\ell^\pm_{\max}-s)}\,\Pi_\lambda^\pm f(s)\in C^\infty([\ell^\pm_{\min},\ell^\pm_{\max}]).$$

Using Lemma 3.4, we then get

$$g_\pm\in C^\infty([\ell^\pm_{\min},\ell^\pm_{\max}]),$$

which implies that $S_\lambda^\pm v\in C^\infty(\bar\Omega)$.

(3) It remains to show that $S_\lambda^0 v\in C^\infty(\bar\Omega)$. We may assume that $v=dF$ for some $F\in C^\infty(\partial\Omega)$, that is, $\int_{\partial\Omega}v=0$. Indeed, if we are studying $S_\lambda^0 v$ near some point $x_0\in\bar\Omega$ then we may take $y_0\in\partial\Omega$ such that $A(x_0-y_0)>0$ and change v in a small neighborhood of y_0 so that $S_\lambda^0 v(x)$ does not change for x near x_0 and v integrates to 0.

For $s\in(\ell^\pm_{\min},\ell^\pm_{\max})$, define $x^\pm_{(1)}(s),x^\pm_{(2)}(s)\in\partial\Omega$ by

$$\ell^\pm(x^\pm_{(1)}(s))=\ell^\pm(x^\pm_{(2)}(s))=s,\quad \ell^\mp(x^\pm_{(1)}(s))<\ell^\mp(x^\pm_{(2)}(s)).$$

Then for any $x\in\Omega$, the set of $y\in\partial\Omega$ such that $A(x-y)<0$ consists of two intervals of the circle $\partial\Omega$, from $x^+_{(1)}(\ell^+(x))$ to $x^-_{(2)}(\ell^-(x))$ (with respect to the positive orientation on $\partial\Omega$) and from $x^+_{(2)}(\ell^+(x))$ to $x^-_{(1)}(\ell^-(x))$ — see Figure 11. Since $v=dF$, we compute for $x\in\Omega$

$$S_\lambda^0 v(x)=F_-(\ell^-(x))-F_+(\ell^+(x)),\quad F_\pm(s):=F(x^\pm_{(1)}(s))+F(x^\pm_{(2)}(s)).$$

By part (2) of Lemma 2.3, we have $F_\pm=\Upsilon_\lambda^\pm F\in C^\infty([\ell^\pm_{\min},\ell^\pm_{\max}])$. Thus $S_\lambda^0 v\in C^\infty(\bar\Omega)$ as needed. $\square$

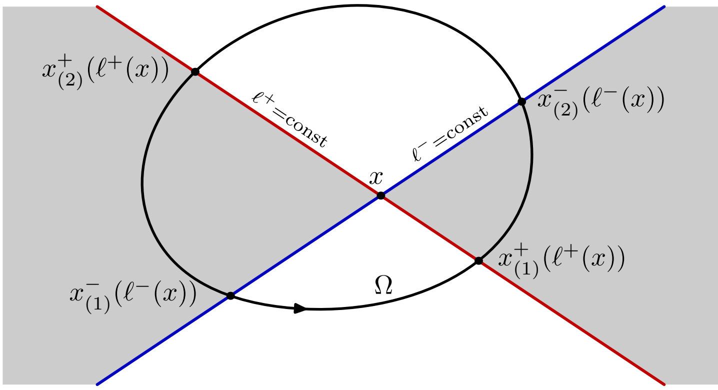

Figure 11. A point $x \in \Omega$ and the corresponding projections $x^\pm_{(j)}(\ell^\pm(x)) \in \partial\Omega$. The shaded region is the set of $y \in \mathbb{R}^2$ such that $A(x-y) < 0$.

4.5.3. *Conormal singularities.* We now study the action of $S_{\lambda+i0}$ on conormal distributions (see Section 3.2):

Lemma 4.9. *Assume that $\lambda \in (0,1)$ and Ω is λ-simple. Fix $y_0 \in \partial\Omega \setminus \mathscr{C}_\lambda$, where the characteristic set $\mathscr{C}_\lambda$ was defined in (2-3). Then for each $v \in \mathcal{D}'(\partial\Omega; T^*\partial\Omega)$ we have*

$$v \in I^s(\partial\Omega, N^*_\pm\{y_0\}) \quad \Longrightarrow \quad S_{\lambda+i0}v \in I^{s-5/4}(\bar{\Omega}, N^*_\pm\Gamma^\pm_\lambda(y_0)).$$

Here the positive/negative halves of the conormal bundle $N^*_\pm\{y_0\} \subset T^*\partial\Omega$ are defined using the positive orientation on $\partial\Omega$, the line segments $\Gamma^\pm_\lambda(y_0) \subset \Omega$ are defined in (1-7) and transverse to the boundary $\partial\Omega$, and $N^*_\pm\Gamma^\pm_\lambda(y_0)$ are defined in (1-8).

Proof. (1) By Lemma 4.8 and since v is smooth away from y_0, we may assume that

$$\operatorname{supp} v \subset U := \{y \in \partial\Omega \mid \nu^+(y) = \nu^+(y_0),\ \ \nu^-(y) = \nu^-(y_0)\}, \tag{4-36}$$

where $\nu^\pm(y) = \operatorname{sgn} \partial_\theta \ell^\pm(y)$; see (2-4). We define $\nu^+ := \nu^+(y_0)$, $\nu^- := \nu^-(y_0)$.

We claim that for all $y \in U$ and $x \in \Omega \setminus \Gamma_\lambda(y)$, where $\Gamma_\lambda(y) := \Gamma^+_\lambda(y) \cup \Gamma^-_\lambda(y)$,

$$\begin{gathered}\log(A(x-y)+i0) = \log(\ell^+(x-y)+i\nu^+0) + \log(\ell^-(x-y)-i\nu^-0) + c_0,\\ c_0 = \begin{cases} 2\pi i, & \text{if } \nu^+ = -1 \text{ and } \nu^- = 1,\\ 0, & \text{otherwise.}\end{cases}\end{gathered} \tag{4-37}$$

(Here as always we use the branch of log real on the positive real axis.) Indeed, fix x and y. By Lemma 2.2, we have

$$\nu^+\ell^-(x-y) > 0 \quad \text{or} \quad \nu^-\ell^+(x-y) < 0 \quad \text{(or both)}.$$

Then there exist $\alpha^+, \alpha^- > 0$ such that $\alpha^+\nu^+\ell^-(x-y) - \alpha^-\nu^-\ell^+(x-y) = 1$. This implies that for all $\varepsilon > 0$

$$\begin{aligned} A(x-y)+i\varepsilon &= \ell^+(x-y)\ell^-(x-y)+i\varepsilon\\ &= (\ell^+(x-y)+i\alpha^+\nu^+\varepsilon)(\ell^-(x-y)-i\alpha^-\nu^-\varepsilon) + \mathcal{O}(\varepsilon^2).\end{aligned}$$

Letting $\varepsilon \to 0+$, we obtain (4-37).

(2) Fix a coordinate θ on $\partial\Omega$ and write $v = f(\theta)\,d\theta$. Similarly to step (2) in the proof of Lemma 4.8 we get from (4-37)

$$S_{\lambda+i0}v(x) = c_\lambda\Big(g_+(\ell^+(x)) + g_-(\ell^-(x)) + c_0\int_{\partial\Omega} f(\theta)\,d\theta\Big),$$

where, letting $\log_+ x := \log(x+i0)$, $\log_- x := \log(x-i0)$ and using (2-10),

$$g_+ := (\Pi_\lambda^+ f) * \log_{\nu^+}, \qquad g_- := (\Pi_\lambda^- f) * \log_{-\nu^-}. \tag{4-38}$$

(Here, $\pm\nu^\bullet$ is meant as $\pm$ if $\nu^\bullet = 1$ and $\mp$ when $\nu^\bullet = -1$.)

By (4-36), we have $\operatorname{supp} v \subset U$, where $\ell^\pm : U \to \mathbb{R}$ are diffeomorphisms onto their ranges. Since $v \in I^s(\partial\Omega, N_\pm^*\{y_0\})$ and recalling (2-12), we then have

$$\Pi_\lambda^+ f \in I^s(\mathbb{R}, N_{\pm\nu^+}^*\{\ell^+(y_0)\}), \qquad \Pi_\lambda^- f \in I^s(\mathbb{R}, N_{\pm\nu^-}^*\{\ell^-(y_0)\}).$$

By Lemma 3.5, we see that

$$g_\pm \in I^{s-1}(\mathbb{R}, N_{\pm\nu^\pm}^*\{\ell^\pm(y_0)\}), \qquad g_\mp \in C^\infty.$$

Using the Fourier characterization of conormal distributions reviewed in Section 3.2, we see that $S_{\lambda+i0}f \in I^{s-5/4}(\overline{\Omega}, N_\pm^*\Gamma_\lambda^\pm(y_0))$ as needed. ☐

Remark. In Section 7 we will apply this result to elements of $I^s(\partial\Omega, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+)$, defined in (3-19), where $\Sigma_\lambda^\pm$ are defined in (1-6). Lemma 4.9 gives

$$S_{\lambda+i0} : I^s(\partial\Omega, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+) \to I^{s-5/4}(\overline{\Omega}, \Lambda^-(\lambda)), \tag{4-39}$$

where $\Lambda^-(\lambda) = N_-^*\Gamma_\lambda^+(\Sigma_\lambda^+) \sqcup N_+^*\Gamma_\lambda^-(\Sigma_\lambda^-) = N_+^*\Gamma_\lambda^+(\Sigma_\lambda^-) \sqcup N_-^*\Gamma_\lambda^-(\Sigma_\lambda^+)$ is defined in (1-9). Here we define the conormal spaces on the right-hand side similarly to (3-19):

$$I^s(\overline{\Omega}, \Lambda^-(\lambda)) := I^s(\overline{\Omega}, N_+^*\Gamma_\lambda^+(\Sigma_\lambda^-)) + I^s(\overline{\Omega}, N_-^*\Gamma_\lambda^-(\Sigma_\lambda^+)).$$

4.6. *The restricted single layer potentials.* We now study the restricted operators

$$\mathcal{C}_\omega : C^\infty(\partial\Omega; T^*\partial\Omega) \to C^\infty(\partial\Omega), \qquad \mathcal{C}_\omega v := (S_\omega v)|_{\partial\Omega}, \tag{4-40}$$

given by the boundary trace of $S_\omega v \in C^\infty(\overline{\Omega})$; see Lemma 4.7. When λ is real and Ω is λ-simple (see Definition 1.1), we have two operators $\mathcal{C}_{\lambda\pm i0}$ obtained by restricting $S_{\lambda\pm i0}$; see Lemma 4.8. From (4-33) we have for $v \in C^\infty(\partial\Omega; T^*\partial\Omega)$

$$\mathcal{C}_\omega v(x) = \int_{\partial\Omega} E_\omega(x-y)\,v(y), \qquad x \in \partial\Omega, \tag{4-41}$$

with the integration in y, and same is true for ω replaced with $\lambda \pm i0$. Later in (4-77) we show that $\mathcal{C}_\omega$ and $\mathcal{C}_{\lambda\pm i0}$ extend to continuous operators $\mathcal{D}'(\partial\Omega; T^*\partial\Omega) \to \mathcal{D}'(\partial\Omega)$.

Composing $\mathcal{C}_\omega$ with the differential $d : C^\infty(\partial\Omega) \to C^\infty(\partial\Omega; T^*\partial\Omega)$ we get the operator

$$d\mathcal{C}_\omega : C^\infty(\partial\Omega; T^*\partial\Omega) \to C^\infty(\partial\Omega; T^*\partial\Omega).$$

In this section we assume that

$$\omega = \lambda + i\varepsilon, \quad \varepsilon > 0, \tag{4-42}$$

where $\lambda \in (0, 1)$ is chosen so that Ω is λ-simple. Our main result here is a microlocal description of $d\mathcal{C}_\omega$ uniformly as $\varepsilon \to 0+$; see Proposition 4.15 below. (This description is also locally uniform in λ; see Remark (1) after Proposition 4.15.)

For convenience, we fix a positively oriented coordinate $\theta \in \mathbb{S}^1$ on $\partial\Omega$ and identify 1-forms on $\partial\Omega$ with functions on $\mathbb{S}^1$ by writing $v = f(\theta)\, d\theta$. Let $\boldsymbol{x} : \mathbb{S}^1 \to \partial\Omega$ be the corresponding parametrization map. Let

$$\gamma_\lambda^\pm : \mathbb{S}^1 \to \mathbb{S}^1, \quad \gamma^\pm(\boldsymbol{x}(\theta), \lambda) = \boldsymbol{x}(\gamma_\lambda^\pm(\theta)), \tag{4-43}$$

be the orientation reversing involutions on $\mathbb{S}^1$ induced by the maps $\gamma^\pm(\bullet, \lambda)$ defined in (1-3).

4.6.1. *A weak convergence statement.* Before starting the microlocal analysis of $d\mathcal{C}_\omega$, we show that $\mathcal{C}_{\lambda+i\varepsilon} \to \mathcal{C}_{\lambda+i0}$ as $\varepsilon \to 0+$ in a weak sense in x, y but uniformly with all derivatives in λ. A stronger convergence will be shown later in Lemma 4.16. We use the letter $\mathscr{O}(\mathcal{J} + i[0, \infty))$ for spaces of holomorphic functions that are smooth up to the boundary interval $\mathcal{J}$, introduced in (4-12) and in the statement of Lemma 4.3.

Lemma 4.10. *Let $\mathcal{J} \subset (0, 1)$ be an open interval such that Ω is λ-simple for all $\lambda \in \mathcal{J}$. Then the Schwartz kernel of the operator*

$$\omega \in \mathcal{J} + i[0, \infty) \mapsto \begin{cases} \mathcal{C}_\omega, & \operatorname{Im}\omega > 0, \\ \mathcal{C}_{\lambda+i0}, & \omega = \lambda \in \mathcal{J}, \end{cases} \tag{4-44}$$

lies in $\mathscr{O}(\mathcal{J} + i[0, \infty); \mathcal{D}'(\partial\Omega \times \partial\Omega))$.

Proof. (1) The holomorphy of (4-44) when $\operatorname{Im}\omega > 0$ follows by differentiating (4-41) (one can cut away from the singularity at $x = y$ and represent the pairing of (4-44) with any element of $C^\infty(\partial\Omega \times \partial\Omega)$ as the locally uniform limit of a sequence of holomorphic functions). The smoothness of the restriction of (4-44) to $\mathcal{J}$ can be shown using the decomposition (4-35) and the λ-dependent local coordinates introduced in step (2) of the present proof. Arguing similarly to step (3) in the proof of Lemma 4.3 and recalling (4-41), we then see that it suffices to show the following convergence statement for all $\varphi \in C^\infty(\partial\Omega \times \partial\Omega)$:

$$\int_{\partial\Omega\times\partial\Omega} E_{\omega_j}(x-y)\varphi(x, y)\, d\theta(x)\, d\theta(y) \\ \to \int_{\partial\Omega\times\partial\Omega} E_{\lambda+i0}(x-y)\varphi(x, y)\, d\theta(x)\, d\theta(y) \quad \text{for all } \omega_j \to \lambda \in \mathcal{J},\ \operatorname{Im}\omega_j > 0.$$

Similarly to (4-35) we take the decomposition

$$E_\omega = E_\omega^+ + E_\omega^- + E_\omega^0, \quad E_\omega^\pm(x) := c_\omega \log|\ell^\pm(x, \omega)|, \quad E_\omega^0 := ic_\omega \operatorname{Im}\log A(x, \omega),$$

and similarly for $E_{\lambda+i0}$. It suffices to show that for $\bullet = +, -, 0$ we have

$$\int_{\partial\Omega\times\partial\Omega} E_{\omega_j}^\bullet(x-y)\varphi(x, y)\, d\theta(x)\, d\theta(y) \to \int_{\partial\Omega\times\partial\Omega} E_{\lambda+i0}^\bullet(x-y)\varphi(x, y)\, d\theta(x)\, d\theta(y). \tag{4-45}$$

(2) We have $E^\bullet_{\omega_j}(x-y)\to E^\bullet_{\lambda+i0}(x-y)$ for almost every $(x,y)\in\partial\Omega\times\partial\Omega$, more specifically for all (x,y) such that $y\notin\{x,\gamma^+(x,\lambda),\gamma^-(x,\lambda)\}$. This gives (4-45) for $\bullet=0$ by the dominated convergence theorem since $|\operatorname{Im}\log A(x,\omega_j)|\le\pi$.

To see (4-45) for $\bullet=+$ (a similar argument works for $\bullet=-$), we follow step (2) of the proof of Lemma 4.3. Instead of the family of linear isomorphisms Φ_λ used there we choose a specific local coordinate θ_j on $\partial\Omega$ which depends on $\lambda_j=\operatorname{Re}\omega_j$. More precisely, using a partition of unity we see that it suffices to show that each $(x_0,y_0)\in\partial\Omega\times\partial\Omega$ has a neighborhood U such that (4-45) holds for all $\varphi\in C^\infty_{\mathrm{c}}(U)$. Now we consider four cases (corresponding to Sections 4.6.3–4.6.6 below):

- $\ell^+(x_0,\lambda)\neq\ell^+(y_0,\lambda)$: we can use the dominated convergence theorem since $E^+_{\omega_j}(x-y)$ is bounded uniformly in j and in $(x,y)\in U$ by (4-17).
- $y_0=x_0\neq\gamma^+(x_0,\lambda)$: we choose the coordinate $\theta_j=\ell^+(x,\lambda_j)$ near x_0. Then the argument in the proof of Lemma 4.3 goes through, using that $\log|\theta-\theta'|$ is a locally integrable function of $(\theta,\theta')\in\mathbb{R}^2$.
- $y_0=\gamma^+(x_0,\lambda)\neq x_0$: we again choose the coordinate $\theta_j=\ell^+(x,\lambda_j)$ near x_0 and near y_0, and the argument goes through as in the previous case.
- $x_0=y_0=\gamma^+(x_0,\lambda)$: assume that $x_0=x^+_{\min}(\lambda)$ is the minimum point of $\ell^+(\bullet,\lambda)$ on $\partial\Omega$ (the case when x_0 is the maximum point is handled similarly). We choose the coordinate θ_j near x_0 given by (2-1):

$$\ell^+(x,\lambda_j)=\ell^+(x^+_{\min}(\lambda_j),\lambda_j)+\theta_j(x)^2.$$

Then the argument in the proof of Lemma 4.3 goes through, using that $\log|\theta^2-(\theta')^2|$ is a locally integrable function of $(\theta,\theta')\in\mathbb{R}^2$. □

4.6.2. *Decomposition into $T^\pm_\omega$.* Since the linear functions $\ell^\pm(x,\omega)$ are dual to the vector fields $L^\pm_\omega$ (see (4-4)), we have

$$d\mathcal{C}_\omega=T^+_\omega+T^-_\omega,\tag{4-46}$$

where the operators $T^\pm_\omega:C^\infty(\partial\Omega;T^*\partial\Omega)\to C^\infty(\partial\Omega;T^*\partial\Omega)$ are given by (with j the embedding map)

$$T^\pm_\omega v=j^*((L^\pm_\omega S_\omega v)\,d\ell^\pm),\qquad j:\partial\Omega\to\overline{\Omega}.\tag{4-47}$$

Let $K^\pm_\omega(\theta,\theta')\in\mathcal{D}'(\mathbb{S}^1\times\mathbb{S}^1)$ be the Schwartz kernel of $T^\pm_\omega$, that is,

$$T^\pm_\omega v(\theta)=(\partial_\theta\ell^\pm(\boldsymbol{x}(\theta),\omega))L^\pm_\omega S_\omega v(\boldsymbol{x}(\theta))\,d\theta=\left(\int_{\mathbb{S}^1}K^\pm_\omega(\theta,\theta')f(\theta')\,d\theta'\right)d\theta,\tag{4-48}$$

where we put $v=f(\theta)\,d\theta$. Recalling the integral definition (4-33) of S_ω, the formula (4-9) for E_ω (which in particular shows that E_ω is smooth on $\mathbb{R}^2\setminus\{0\}$), and the identity (4-10), we see that $K^\pm_\omega$ is smooth on $(\mathbb{S}^1\times\mathbb{S}^1)\setminus\{\theta\neq\theta'\}$ and

$$K^\pm_\omega(\theta,\theta')=c_\omega\frac{\partial_\theta\ell^\pm(\boldsymbol{x}(\theta),\omega)}{\ell^\pm(\boldsymbol{x}(\theta)-\boldsymbol{x}(\theta'),\omega)},\qquad\theta\neq\theta'.\tag{4-49}$$

4.6.3. *Away from the singularities.* Define the sets

$$\begin{aligned}\mathrm{Diag} &:= \{(\theta,\theta) \mid \theta \in \mathbb{S}^1\},\\ \mathrm{Ref}_\lambda^\pm &:= \{(\theta, \gamma_\lambda^\pm(\theta)) \mid \theta \in \mathbb{S}^1\}.\end{aligned} \tag{4-50}$$

Note that the intersection

$$\mathrm{Diag} \cap \mathrm{Ref}_\lambda^\pm = \{(\theta,\theta) \mid \theta \in \mathbb{S}^1,\ \partial_\theta \ell^\pm(\boldsymbol{x}(\theta),\lambda) = 0\} \tag{4-51}$$

corresponds to the critical points $x_{\min}^\pm(\lambda)$, $x_{\max}^\pm(\lambda)$ of $\ell^\pm(\bullet,\lambda)$ on $\partial\Omega$ (see Definition 1.1). At these points the operator $P(\lambda)$ is characteristic with respect to $\partial\Omega$.

We start the analysis of the uniform behavior of $K_\omega^\pm$ as $\varepsilon = \operatorname{Im}\omega \to 0$ by showing that the singularities are contained in $\mathrm{Diag} \cup \mathrm{Ref}_\lambda^\pm$:

Lemma 4.11. *We have*

$$K_\omega^\pm|_{(\mathbb{S}^1\times\mathbb{S}^1)\setminus(\mathrm{Diag}\cup\mathrm{Ref}_\lambda^\pm)} \in C^\infty((\mathbb{S}^1\times\mathbb{S}^1)\setminus(\mathrm{Diag}\cup\mathrm{Ref}_\lambda^\pm))$$

smoothly in ε up to $\varepsilon = 0$.

Proof. This follows immediately from (4-49). Indeed, for $(\theta,\theta') \notin \mathrm{Diag}\cup\mathrm{Ref}_\lambda^\pm$, we have $\ell^\pm(\boldsymbol{x}(\theta),\lambda) \neq \ell^\pm(\boldsymbol{x}(\theta'),\lambda)$ and thus the denominator in (4-49) is nonvanishing when $\varepsilon = 0$. □

4.6.4. *Noncharacteristic diagonal.* We next consider the singularities of $K_\omega^\pm(\theta,\theta')$ near the diagonal but away from the characteristic set $\mathrm{Diag}\cap\mathrm{Ref}_\lambda^\pm$. In that case the structure of the kernel is similar to the model case (4-6):

Lemma 4.12. *Take $\theta_0 \in \mathbb{S}^1$ such that $\gamma_\lambda^\pm(\theta_0) \neq \theta_0$. Then, for θ,θ' in some neighborhood U of θ_0 and $\varepsilon = \operatorname{Im}\omega > 0$ small enough, we have*

$$K_\omega^\pm(\theta,\theta') = c_\omega(\theta-\theta' \pm i0)^{-1} + \mathscr{K}_\omega^\pm(\theta,\theta'), \tag{4-52}$$

where $\mathscr{K}_\omega^\pm \in C^\infty(U\times U)$ is smooth in $\theta,\theta',\varepsilon$ up to $\varepsilon = 0$.

Proof. (1) Fix some smooth vector field $\boldsymbol{v}(\theta)$ on $\partial\Omega$ which points inwards. We have for all $v = f(\theta)\,d\theta \in C^\infty(\partial\Omega; T^*\partial\Omega)$,

$$\begin{aligned}\int_{\mathbb{S}^1} K_\omega^\pm(\theta,\theta') f(\theta')\,d\theta' &= (\partial_\theta \ell^\pm(\boldsymbol{x}(\theta),\omega)) \lim_{\delta\to 0+} L_\omega^\pm S_\omega v(\boldsymbol{x}(\theta)+\delta\boldsymbol{v}(\theta))\\ &= c_\omega(\partial_\theta \ell^\pm(\boldsymbol{x}(\theta),\omega)) \lim_{\delta\to0+}\int_{\mathbb{S}^1} \frac{f(\theta')}{\ell^\pm(\boldsymbol{x}(\theta)-\boldsymbol{x}(\theta')+\delta\boldsymbol{v}(\theta),\omega)}\,d\theta',\end{aligned}$$

where the limit is in $C^\infty(\mathbb{S}^1)$. Here in the first equality we use the definition (4-48) of $K_\omega^\pm$ (recalling that $S_\omega v \in C^\infty(\overline{\Omega})$ by Lemma 4.7). In the second equality we use the definition (4-33) of S_ω, the formula (4-9) for E_ω, and the identity (4-10).

Since $\partial_\theta \ell^\pm(\boldsymbol{x}(\theta),\lambda) \neq 0$ at $\theta = \theta_0$, we factorize for θ,θ' in some neighborhood U of θ_0 and $\varepsilon = \operatorname{Im}\omega$ small enough

$$\ell^\pm(\boldsymbol{x}(\theta)-\boldsymbol{x}(\theta'),\omega) = G_\omega^\pm(\theta,\theta')(\theta-\theta'),$$

where $G^{\pm}_{\omega}(\theta,\theta')$ is a nonvanishing smooth function of $\theta, \theta', \varepsilon$ up to $\varepsilon = 0$ and

$$G^{\pm}_{\omega}(\theta,\theta) = \partial_\theta \ell^{\pm}(\boldsymbol{x}(\theta),\omega), \quad \theta \in U. \tag{4-53}$$

Therefore, for $(\theta,\theta') \in U \times U$ we have

$$K^{\pm}_{\omega}(\theta,\theta') = \frac{c_\omega \partial_\theta \ell^{\pm}(\boldsymbol{x}(\theta),\omega)}{G^{\pm}_{\omega}(\theta,\theta')} \lim_{\delta\to 0+} \left(\theta - \theta' + \delta \frac{\ell^{\pm}(\boldsymbol{v}(\theta),\omega)}{G^{\pm}_{\omega}(\theta,\theta')}\right)^{-1}, \tag{4-54}$$

with the limit in $\mathcal{D}'(U \times U)$.

(2) We next claim that if U is a small enough neighborhood of θ_0, then for all $(\theta,\theta') \in U \times U$ and $\operatorname{Im}\omega = \varepsilon > 0$ small enough

$$\pm \operatorname{Im} \frac{\ell^{\pm}(\boldsymbol{v}(\theta),\omega)}{G^{\pm}_{\omega}(\theta,\theta')} > 0. \tag{4-55}$$

When $\omega = \lambda$ is real, the expression (4-55) is equal to 0. Thus it suffices to check that for all $(\theta,\theta') \in U \times U$

$$\pm \partial_\varepsilon|_{\varepsilon=0} \operatorname{Im} \frac{\ell^{\pm}(\boldsymbol{v}(\theta),\lambda + i\varepsilon)}{G^{\pm}_{\lambda+i\varepsilon}(\theta,\theta')} > 0. \tag{4-56}$$

It is enough to consider the case $\theta = \theta' = \theta_0$, in which case the left-hand side of (4-56) equals

$$\pm \partial_\varepsilon|_{\varepsilon=0} \operatorname{Im} \frac{\ell^{\pm}(\boldsymbol{v}(\theta_0),\lambda + i\varepsilon)}{\ell^{\pm}(\partial_\theta \boldsymbol{x}(\theta_0),\lambda + i\varepsilon)}.$$

By (2-7) and since $\ell^{\pm}$ is holomorphic in ω it then suffices to check that

$$\pm\big(\ell^{\mp}(\boldsymbol{v}(\theta_0),\lambda)\ell^{\pm}(\partial_\theta \boldsymbol{x}(\theta_0),\lambda) - \ell^{\pm}(\boldsymbol{v}(\theta_0),\lambda)\ell^{\mp}(\partial_\theta \boldsymbol{x}(\theta_0),\lambda)\big) > 0. \tag{4-57}$$

The inequality (4-57) follows from the fact that $x \mapsto (\ell^{+}(x,\lambda), \ell^{-}(x,\lambda))$ is an orientation-preserving linear map on $\mathbb{R}^2$ and $\partial_\theta \boldsymbol{x}(\theta_0)$, $\boldsymbol{v}(\theta_0)$ form a positively oriented basis of $\mathbb{R}^2$ since the parametrization $\boldsymbol{x}(\theta)$ is positively oriented and $\boldsymbol{v}(\theta)$ points inside Ω. This finishes the proof of (4-55).

(3) By Lemma 3.7 (see also (3-34)), with δ taking the role of ε, the distributional limit on the right-hand side of (4-54) is equal to $(\theta - \theta' \pm i0)^{-1}$. Therefore

$$K^{\pm}_{\omega}(\theta,\theta') = \frac{c_\omega \partial_\theta \ell^{\pm}(\boldsymbol{x}(\theta),\omega)}{G^{\pm}_{\omega}(\theta,\theta')} (\theta - \theta' \pm i0)^{-1}. \tag{4-58}$$

By (4-53) we can write for some $\mathscr{K}^{\pm}_{\omega}(\theta,\theta')$ which is smooth in $\theta, \theta', \varepsilon$ up to $\varepsilon = 0$,

$$\frac{c_\omega \partial_\theta \ell^{\pm}(\boldsymbol{x}(\theta),\omega)}{G^{\pm}_{\omega}(\theta,\theta')} = c_\omega + \mathscr{K}^{\pm}_{\omega}(\theta,\theta')(\theta - \theta'),$$

which gives (4-52) since $(\theta - \theta')(\theta - \theta' \pm i0)^{-1} = 1$. □

4.6.5. *Noncharacteristic reflection.* We now move to the singularities on the reflection sets $\operatorname{Ref}^{\pm}_{\lambda}$, again staying away from the characteristic set $\operatorname{Diag} \cap \operatorname{Ref}^{\pm}_{\lambda}$:

Lemma 4.13. *Take $\theta_0 \in \mathbb{S}^1$ such that $\gamma_\lambda^\pm(\theta_0) \neq \theta_0$. Then there exists neighborhoods $U, U' = \gamma_\lambda^\pm(U)$ of $\gamma_\lambda^\pm(\theta_0), \theta_0$ such that, for $(\theta, \theta') \in U \times U'$ and $\varepsilon = \operatorname{Im}\omega > 0$ small enough, we have*

$$K_\omega^\pm(\theta, \theta') = \tilde{c}_\omega^\pm(\theta')\big(\gamma_\lambda^\pm(\theta) - \theta' \pm i\varepsilon z_\omega^\pm(\theta')\big)^{-1} + \mathscr{K}_\omega^\pm(\theta, \theta'), \tag{4-59}$$

where $\mathscr{K}_\omega^\pm \in C^\infty(U \times U')$ is smooth in $\theta, \theta', \varepsilon$ up to $\varepsilon = 0$, the functions $\tilde{c}_\omega^\pm(\theta')$ and $z_\omega^\pm(\theta')$ are smooth in θ', ε up to $\varepsilon = 0$, and

$$\tilde{c}_\omega^\pm(\theta') = \frac{c_\omega}{\partial_{\theta'}\gamma_\lambda^\pm(\theta')} + \mathcal{O}(\varepsilon), \quad \operatorname{Re} z_\omega^\pm(\theta') \geq c > 0, \tag{4-60}$$

where c is independent of ε, θ'.

Proof. (1) Recall that $\omega = \lambda + i\varepsilon$. We take Taylor expansions of $\ell^\pm(x, \omega)$ at $\varepsilon = 0$, using its holomorphy in ω:

$$\ell^\pm(x, \omega) = \ell^\pm(x, \lambda) + i\varepsilon\ell_1^\pm(x, \lambda) + \varepsilon^2\ell_2^\pm(x, \lambda, \varepsilon), \quad \ell_1^\pm(x, \lambda) := \partial_\lambda\ell^\pm(x, \lambda), \tag{4-61}$$

where the coefficients of the linear maps $x \mapsto \ell_2^\pm(x, \lambda, \varepsilon)$ are smooth in ε up to $\varepsilon = 0$. Since we have $\partial_\theta\ell^\pm(\boldsymbol{x}(\theta), \lambda) \neq 0$ at $\theta = \theta_0$, we factorize for θ, θ' in some neighborhoods $U, U' = \gamma_\lambda^\pm(U)$ of $\gamma_\lambda^\pm(\theta_0), \theta_0$

$$\ell^\pm(\boldsymbol{x}(\theta) - \boldsymbol{x}(\theta'), \lambda) = \ell^\pm(\boldsymbol{x}(\gamma_\lambda^\pm(\theta)) - \boldsymbol{x}(\theta'), \lambda) = G_\lambda^\pm(\theta, \theta')(\gamma_\lambda^\pm(\theta) - \theta'),$$

where $G_\lambda^\pm \in C^\infty(U \times U'; \mathbb{R})$ is nonvanishing and

$$G_\lambda^\pm(\gamma_\lambda^\pm(\theta'), \theta') = \partial_{\theta'}\ell^\pm(\boldsymbol{x}(\theta'), \lambda). \tag{4-62}$$

Hence for $(\theta, \theta') \in U \times U'$

$$\ell^\pm(\boldsymbol{x}(\theta) - \boldsymbol{x}(\theta'), \omega) = G_\lambda^\pm(\theta, \theta')(\gamma_\lambda^\pm(\theta) - \theta' \pm i\varepsilon\psi_\omega^\pm(\theta, \theta')),$$

$$\psi_\omega^\pm(\theta, \theta') := \pm\frac{\ell_1^\pm(\boldsymbol{x}(\theta) - \boldsymbol{x}(\theta'), \lambda) - i\varepsilon\ell_2^\pm(\boldsymbol{x}(\theta) - \boldsymbol{x}(\theta'), \lambda, \varepsilon)}{G_\lambda^\pm(\theta, \theta')}.$$

By (4-49) we have for $(\theta, \theta') \in U \times U'$

$$K_\omega^\pm(\theta, \theta') = F_\omega^\pm(\theta, \theta')\big(\gamma_\lambda^\pm(\theta) - \theta' \pm i\varepsilon\psi_\omega^\pm(\theta, \theta')\big)^{-1}, \quad F_\omega^\pm(\theta, \theta') := c_\omega\frac{\partial_\theta\ell^\pm(\boldsymbol{x}(\theta), \omega)}{G_\lambda^\pm(\theta, \theta')}.$$

Note that $\psi_\omega^\pm(\theta, \theta')$ and $F_\omega^\pm(\theta, \theta')$ are smooth in $\theta, \theta', \varepsilon$ up to $\varepsilon = 0$.

(2) We next claim that $\operatorname{Re}\psi_\omega^\pm(\theta, \theta') \geq c > 0$ for ε small enough and $(\theta, \theta') \in U \times U'$, if U, U' are sufficiently small neighborhoods of $\gamma_\lambda^\pm(\theta_0), \theta_0$. For that it suffices to show that

$$\pm\frac{\ell_1^\pm(\boldsymbol{x}(\gamma_\lambda^\pm(\theta_0)) - \boldsymbol{x}(\theta_0), \lambda)}{G_\lambda^\pm(\gamma_\lambda^\pm(\theta_0), \theta_0)} > 0. \tag{4-63}$$

By (2-7) and (4-62), and since $\ell^\pm(\boldsymbol{x}(\gamma_\lambda^\pm(\theta_0)) - \boldsymbol{x}(\theta_0), \lambda) = 0$, the left-hand side of (4-63) has the same sign as

$$\pm\frac{\ell^\mp(\boldsymbol{x}(\gamma_\lambda^\pm(\theta_0)) - \boldsymbol{x}(\theta_0), \lambda)}{\partial_\theta\ell^\pm(\boldsymbol{x}(\theta), \lambda)|_{\theta=\theta_0}},$$

which is positive by (2-5) with $x := \boldsymbol{x}(\theta_0)$.

(3) Now (4-59) and the second part of (4-60) follow from Lemma 3.6, see also the remark following Lemma 3.8 where we replace θ with $\gamma_\lambda^\pm(\theta)$. Finally, by (4-62) and differentiating the identity $\ell^\pm(\boldsymbol{x}(\gamma_\lambda^\pm(\theta')),\lambda)=\ell^\pm(\boldsymbol{x}(\theta'),\lambda)$ in θ' we compute

$$F_\omega^\pm(\gamma_\lambda^\pm(\theta'),\theta')=c_\omega\frac{\partial_\theta\ell^\pm(\boldsymbol{x}(\theta),\lambda)|_{\theta=\gamma_\lambda^\pm(\theta')}}{\partial_{\theta'}\ell^\pm(\boldsymbol{x}(\theta'),\lambda)}+\mathcal{O}(\varepsilon)=\frac{c_\omega}{\partial_{\theta'}\gamma_\lambda^\pm(\theta')}+\mathcal{O}(\varepsilon),$$

which gives the first part of (4-60). □

4.6.6. *Characteristic points.* We finally study the singularities of $K_\omega^\pm$ near the characteristic set $\mathrm{Diag}\cap\mathrm{Ref}_\lambda^\pm$. Recalling (4-51), we see that this set consists of two points $(\theta_{\min}^\pm,\theta_{\min}^\pm)$ and $(\theta_{\max}^\pm,\theta_{\max}^\pm)$, where $\boldsymbol{x}(\theta_{\min}^\pm)=x_{\min}^\pm(\lambda)$, $\boldsymbol{x}(\theta_{\max}^\pm)=x_{\max}^\pm(\lambda)$ are the critical points of $\ell^\pm(\bullet,\lambda)$ (see Definition 1.1).

Lemma 4.14. *Assume that $\theta_0\in\{\theta_{\min}^\pm,\theta_{\max}^\pm\}$. Then there exists a neighborhood $U=\gamma_\lambda^\pm(U)$ of θ_0 such that, for $(\theta,\theta')\in U\times U$ and $\varepsilon=\operatorname{Im}\omega>0$ small enough, we have*

$$K_\omega^\pm(\theta,\theta')=c_\omega(\theta-\theta'\pm i0)^{-1}+\tilde c_\omega^\pm(\theta')\big(\gamma_\lambda^\pm(\theta)-\theta'\pm i\varepsilon z_\omega^\pm(\theta')\big)^{-1}+\mathscr{K}_\omega^\pm(\theta,\theta'),\tag{4-64}$$

where $\mathscr{K}_\omega^\pm\in C^\infty(U\times U)$ is smooth in $\theta,\theta',\varepsilon$ up to $\varepsilon=0$, $\tilde c_\omega^\pm(\theta')$ and $z_\omega^\pm(\theta')$ are smooth in θ',ε up to $\varepsilon=0$, and (4-60) holds.

Remarks. (1) Note that Lemma 4.14 implies Lemmas 4.12 and 4.13 in a neighborhood of the characteristic set, since the first term on the right-hand side of (4-64) is smooth away from the diagonal Diag and the second term is smooth (uniformly in ε) away from the reflection set $\mathrm{Ref}_\lambda^\pm$.

(2) Since keeping track of the signs is frustrating we present a model situation: $\ell^+(x)=x_1+i\varepsilon x_2$, $\ell^-(x)=x_2+i\varepsilon x_1$ (which is compatible with Lemma 4.1) and $\partial\Omega$ which near $(0,0)$ is given by

$$x_1=q(x_2),\quad q(0)=q'(0)=0,\quad q''(0)<0.$$

This corresponds to the point $\theta_{\max}^+$, since when $\varepsilon=0$ the function $\ell^+(x)=x_1$ has a nondegenerate maximum on $\partial\Omega$.

We can use $\theta=x_2$ as a positively oriented parametrization of $\partial\Omega$ near $(0,0)$. In that case the involution $\gamma^+(\theta)$ is given by

$$q(\gamma^+(\theta))=q(\theta),\quad \gamma^+(\theta)=-\theta+\mathcal{O}(\theta^2).$$

This gives

$$q(\theta)-q(\theta')=Q(\theta,\theta')(\theta-\theta')(\gamma^+(\theta)-\theta'),\quad Q(0,0)=-\frac{q''(0)}{2}>0.$$

The Schwartz kernel of the model restricted single layer potential $\mathcal{C}$ is given by (with $Q=Q(\theta,\theta')$ and neglecting the overall constant c_ω in (4-9))

$$\begin{aligned}K(\theta,\theta')&=\log\big(\ell^+(\boldsymbol{x}(\theta)-\boldsymbol{x}(\theta'))\ell^-(\boldsymbol{x}(\theta)-\boldsymbol{x}(\theta'))\big)\\&=\log\big((q(\theta)-q(\theta')+i\varepsilon(\theta-\theta'))(\theta-\theta'+i\varepsilon(q(\theta)-q(\theta')))\big)\\&=\log\big((\theta-\theta')^2(Q(\gamma^+(\theta)-\theta')+i\varepsilon)(1+i\varepsilon Q(\gamma^+(\theta)-\theta'))\big)\\&=2\log|\theta-\theta'|+\log(\gamma^+(\theta)-\theta'+i\varepsilon Q^{-1})+\log(1+i\varepsilon Q(\gamma^+(\theta)-\theta'))+\log Q.\end{aligned}$$

Hence (see Section 4.2) the Schwartz kernel of $\partial_\theta \mathcal{C}$ is

$$\partial_\theta K(\theta,\theta') = \sum_{\pm}(\theta-\theta'\pm i0)^{-1} + \frac{\partial_\theta\gamma^+(\theta)+i\varepsilon\partial_\theta Q^{-1}(\theta,\theta')}{\gamma^+(\theta)-\theta'+i\varepsilon Q^{-1}(\theta,\theta')} + \mathscr{K}(\theta,\theta'),$$

where $Q(0,0)>0$ and $\mathscr{K}\in C^\infty$ uniformly in ε. This is consistent with (4-64) and (4-60), where we use Lemma 3.6 and recall that by (4-46) we have $\partial_\theta K = K^+ + K^-$.

Proof of Lemma 4.14. (1) Recall that $\omega=\lambda+i\varepsilon$, $\varepsilon>0$ and consider the expansion (4-61):

$$\ell^\pm(x,\omega)=\ell^\pm(x,\lambda)+i\varepsilon\ell_1^\pm(x,\lambda)+\varepsilon^2\ell_2^\pm(x,\lambda,\varepsilon).$$

We have for θ,θ' in a sufficiently small neighborhood U of $\theta_0\in\{\theta^\pm_{\min},\theta^\pm_{\max}\}$

$$\begin{aligned}\ell^\pm(\boldsymbol{x}(\theta)-\boldsymbol{x}(\theta'),\lambda)&=G_0(\theta,\theta')(\gamma_\lambda^\pm(\theta)-\theta')(\theta-\theta'),\\ \ell_1^\pm(\boldsymbol{x}(\theta)-\boldsymbol{x}(\theta'),\lambda)&=G_1(\theta,\theta')(\theta-\theta'),\\ \ell_2^\pm(\boldsymbol{x}(\theta)-\boldsymbol{x}(\theta'),\lambda,\varepsilon)&=G_2(\theta,\theta',\varepsilon)(\theta-\theta'),\end{aligned}\tag{4-65}$$

where G_0, G_1, G_2 are smooth in $\theta,\theta',\varepsilon$ up to $\varepsilon=0$, and G_0, G_1 are real-valued and nonvanishing. Indeed, the first decomposition follows from (2-1) and the second one, from (2-7) and the fact that $\partial_\theta\ell^\mp(\boldsymbol{x}(\theta),\lambda)\neq 0$ at $\theta=\theta_0$. We have now (with $G_j=G_j(\theta,\theta')$)

$$\ell^\pm(\boldsymbol{x}(\theta)-\boldsymbol{x}(\theta'),\omega)=(\theta-\theta')(G_0(\gamma_\lambda^\pm(\theta)-\theta')+i\varepsilon G_1+\varepsilon^2G_2).\tag{4-66}$$

(2) The argument in the proof of Lemma 4.12 (see (4-58)) shows that for any fixed small $\varepsilon>0$

$$K_\omega^\pm(\theta,\theta')=\frac{c_\omega\partial_\theta\ell^\pm(\boldsymbol{x}(\theta),\omega)}{G_0(\gamma_\lambda^\pm(\theta)-\theta')+i\varepsilon G_1+\varepsilon^2G_2}(\theta-\theta'\pm i0)^{-1}.\tag{4-67}$$

To apply this argument we need to check the condition (4-55), which we rewrite as

$$\pm\operatorname{Im}\frac{G_0(\gamma_\lambda^\pm(\theta)-\theta')+i\varepsilon G_1+\varepsilon^2G_2}{\ell^\pm(\boldsymbol{v}(\theta),\omega)}<0\tag{4-68}$$

for θ,θ' near θ_0, $\varepsilon=\operatorname{Im}\omega>0$ small enough, and $\boldsymbol{v}(\theta)$ an inward-pointing vector field on $\partial\Omega$. Here the denominator is separated away from zero since $\ell^\pm(\boldsymbol{v}(\theta_0),\lambda)\neq 0$.

For $\varepsilon=0$, the expression (4-68) is equal to 0. Thus it suffices to check the sign of its derivative in ε at $\varepsilon=0$ and $\theta=\theta'=\theta_0$, that is, show that (where we use (2-7))

$$\pm\ell^\pm(\boldsymbol{v}(\theta_0),\lambda)\ell^\mp(\partial_\theta\boldsymbol{x}(\theta_0),\lambda)<0.\tag{4-69}$$

The latter follows from the fact that $\ell^\pm(\partial_\theta\boldsymbol{x}(\theta_0),\lambda)=0$, $x\mapsto(\ell^+(x,\lambda),\ell^-(x,\lambda))$ is an orientation-preserving linear map on $\mathbb{R}^2$, and $\partial_\theta\boldsymbol{x}(\theta_0)$, $\boldsymbol{v}(\theta_0)$ form a positively oriented basis of $\mathbb{R}^2$.

(3) Differentiating (4-66) in θ to get a formula for $\partial_\theta\ell^\pm(\boldsymbol{x}(\theta),\omega)$ and substituting into (4-67) we get the following identity for $\theta,\theta'\in U$:

$$K_\omega^\pm(\theta,\theta')=c_\omega(\theta-\theta'\pm i0)^{-1}+\frac{c_\omega\partial_\theta(G_0(\gamma_\lambda^\pm(\theta)-\theta')+i\varepsilon G_1+\varepsilon^2G_2)}{G_0(\gamma_\lambda^\pm(\theta)-\theta')+i\varepsilon G_1+\varepsilon^2G_2},\tag{4-70}$$

where, as before, $G_j = G_j(\theta,\theta')$. Dividing the numerator and denominator of the last term on the right-hand side by G_0, we see that the second term on the right-hand side of (4-70) is equal to $F^\pm_\omega(\theta,\theta')(\gamma^\pm_\lambda(\theta)-\theta' \pm i\varepsilon\psi^\pm_\omega(\theta,\theta'))^{-1}$, where the functions

$$\psi^\pm_\omega(\theta,\theta') := \pm\frac{G_1(\theta,\theta') - i\varepsilon G_2(\theta,\theta',\varepsilon)}{G_0(\theta,\theta')},$$

$$F^\pm_\omega(\theta,\theta') := \frac{c_\omega\partial_\theta\big(G_0(\theta,\theta')(\gamma^\pm_\lambda(\theta)-\theta') + i\varepsilon G_1(\theta,\theta') + \varepsilon^2 G_2(\theta,\theta',\varepsilon)\big)}{G_0(\theta,\theta')}$$

are smooth in $\theta,\theta',\varepsilon$ up to $\varepsilon = 0$ and $\psi^\pm_\omega$ is real and nonzero when $\varepsilon = 0$.

To get (4-64) we can now use Lemma 3.6 (and the remark following Lemma 3.8) similarly to step (3) in the proof of Lemma 4.13. Here the sign condition $\operatorname{Re}\psi^\pm_\omega \ge c > 0$ and (4-60) can be verified by a direct computation using (2-7), definitions (4-65) and (4-69); note that for the sign condition it suffices to check the sign of G_1/G_0 at $\theta = \theta' = \theta_0$. □

4.6.7. *Summary.* We summarize the findings of this section in microlocal terms. Consider the pullback operator by $\gamma^\pm_\lambda$ on 1-forms on $\mathbb{S}^1$,

$$(\gamma^\pm_\lambda)^* : C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1) \to C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1).$$

In terms of the identification of functions with 1-forms, $f \mapsto f\,d\theta$, we have

$$(\gamma^\pm_\lambda)^*(f\,d\theta) = ((f\circ\gamma^\pm_\lambda)\partial_\theta\gamma^\pm_\lambda)\,d\theta. \tag{4-71}$$

Proposition 4.15. *Assume that $\omega = \lambda + i\varepsilon$, where $\lambda \in (0,1)$, $\varepsilon \ge 0$, and Ω is λ-simple in the sense of Definition 1.1. Let $\mathcal{C}_\omega$ be the operator defined in (4-40), where for $\varepsilon = 0$ we understand it as the operator $\mathcal{C}_{\lambda+i0}$. Using the coordinate θ, we treat $d\mathcal{C}_\omega$ as an operator on $C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1)$. Then for all ε small enough, we can write*

$$\mathcal{E}_\omega d\mathcal{C}_\omega = I + (\gamma^+_\lambda)^* A^+_\omega + (\gamma^-_\lambda)^* A^-_\omega, \tag{4-72}$$

where $\mathcal{E}_\omega$, $A^\pm_\omega$ are pseudodifferential operators in $\Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1)$ bounded uniformly in ε and such that, uniformly in ε (see (3-7)),

$$\sigma(\mathcal{E}_\omega)(\theta,\xi) = \frac{i\operatorname{sgn}\xi}{2\pi c_\omega}, \quad \operatorname{WF}(A^\pm_\omega) \subset \{\pm\xi > 0\}, \quad \sigma(A^\pm_\omega)(\theta,\xi) = a^\pm_\omega(\theta)H(\pm\xi)e^{-\varepsilon z^\pm_\omega(\theta)|\xi|},$$

where $H(\xi)$ denotes the Heaviside function, $a^\pm_\omega$ and $z^\pm_\omega$ are smooth in θ,ε up to $\varepsilon = 0$, $\operatorname{Re} z^\pm_\omega(\theta) \ge c > 0$, and $a^\pm_\omega(\theta) = -1 + \mathcal{O}(\varepsilon)$.

Remarks. (1) Proposition 4.15 is stated for a fixed value of $\lambda = \operatorname{Re}\omega$. However, its proof still works when λ varies in some open interval $\mathcal{J} \subset (0,1)$ such that Ω is λ-simple for all $\lambda \in \mathcal{J}$. The conclusions of Proposition 4.15 hold locally uniformly in $\lambda \in \mathcal{J}$ and the functions $a^\pm_\omega(\theta)$, $z^\pm_\omega(\theta)$ can be chosen depending smoothly on $\theta \in \mathbb{S}^1$, $\lambda \in \mathcal{J}$, and $\varepsilon = \operatorname{Im}\omega \ge 0$. Moreover, the operators $A^\pm_\omega$ and $\mathcal{E}_\omega$ depend smoothly on λ and all their λ-derivatives are in Ψ^0 uniformly in ε; the same is true for the pseudodifferential operators featured in the decomposition (4-75) below.

(2) One can formulate a version of (4-72) directly on $\partial\Omega$ which does not depend on the choice of the (positively oriented) coordinate θ, using the fact that the principal symbol (3-14) is invariantly defined.

Proof. (1) Recall from (4-46) that $d\mathcal{C}_\omega = T_\omega^+ + T_\omega^-$, where $T_\omega^\pm$ are defined in (4-47). As with $d\mathcal{C}_\omega$, we use the coordinate θ to think of $T_\omega^\pm$ as operators on $C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1)$. We will write $T_\omega^\pm$ as a sum of a pseudodifferential operator and a composition of a pseudodifferential operator with $(\gamma_\lambda^\pm)^*$; see (4-75) below. The singular supports of the Schwartz kernels of these two operators will lie in the sets Diag and $\mathrm{Ref}_\lambda^\pm$ defined in (4-50).

Fix a cutoff $\chi_{\mathrm{Diag}} \in C^\infty(\mathbb{S}^1 \times \mathbb{S}^1)$ supported in a small neighborhood of the diagonal Diag and equal to 1 on a smaller neighborhood of Diag. Define the (ω-dependent) operator

$$T_{\mathrm{Diag}}^\pm : C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1) \to C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1),$$

with the Schwartz kernel $c_\omega \chi_{\mathrm{Diag}}(\theta, \theta')(\theta - \theta' \pm i0)^{-1}$. Here Schwartz kernels are defined in (4-48). By Lemma 3.9 we have

$$T_{\mathrm{Diag}}^\pm \in \Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1), \quad \sigma(T_{\mathrm{Diag}}^\pm)(\theta, \xi) = \mp 2\pi i c_\omega H(\pm\xi). \tag{4-73}$$

(2) Next, define the reflected operators

$$T_{\mathrm{Ref}}^\pm := T_\omega^\pm - T_{\mathrm{Diag}}^\pm, \quad \widehat{T}_{\mathrm{Ref}}^\pm := (\gamma_\lambda^\pm)^* T_{\mathrm{Ref}}^\pm.$$

Denote by $K_{\mathrm{Ref}}^\pm$, $\widehat{K}_{\mathrm{Ref}}^\pm$ the corresponding Schwartz kernels. Combining Lemmas 4.11, 4.12, 4.13, and 4.14 we see that, putting $\chi_{\mathrm{Ref}}^\pm(\theta, \theta') := \chi_{\mathrm{Diag}}(\gamma_\lambda^\pm(\theta), \theta')$,

$$K_{\mathrm{Ref}}^\pm(\theta, \theta') = \chi_{\mathrm{Ref}}^\pm(\theta, \theta')\tilde{c}_\omega^\pm(\theta')(\gamma_\lambda^\pm(\theta) - \theta' \pm i\varepsilon z_\omega^\pm(\theta'))^{-1} + \mathscr{K}_\omega^\pm(\theta, \theta'), \quad 0 < \varepsilon < \varepsilon_0,$$

where $\mathscr{K}_\omega^\pm$ is smooth in $\theta, \theta', \varepsilon$ up to $\varepsilon = 0$, $\tilde{c}_\omega^\pm(\theta')$ and $z_\omega^\pm(\theta')$ are smooth in θ', ε up to $\varepsilon = 0$, $\operatorname{Re} z_\omega^\pm(\theta') \geq c > 0$ for some constant c, and $\tilde{c}_\omega^\pm(\theta') = c_\omega / \partial_{\theta'}\gamma_\lambda^\pm(\theta') + \mathcal{O}(\varepsilon)$. Here we use a partition of unity and Lemma 3.8 to patch together different local representations from Lemmas 4.13 and 4.14 and get globally defined $\tilde{c}_\omega^\pm$, $z_\omega^\pm$. Recalling (4-71), we have

$$\widehat{K}_{\mathrm{Ref}}^\pm(\theta, \theta') = (\partial_\theta \gamma_\lambda^\pm(\theta)) K_{\mathrm{Ref}}^\pm(\gamma_\lambda^\pm(\theta), \theta').$$

Thus by Lemma 3.9 the operator $\widehat{T}_{\mathrm{Ref}}^\pm$ is pseudodifferential: we have uniformly in $\varepsilon > 0$

$$\begin{gathered} \widehat{T}_{\mathrm{Ref}}^\pm \in \Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1), \quad \mathrm{WF}(\widehat{T}_{\mathrm{Ref}}^\pm) \subset \{\pm\xi > 0\}, \\ \sigma(\widehat{T}_{\mathrm{Ref}}^\pm)(\theta, \xi) = \mp 2\pi i \tilde{c}_\omega^\pm(\theta)(\partial_\theta \gamma_\lambda^\pm(\theta)) e^{-\varepsilon z_\omega^\pm(\theta)|\xi|} H(\pm\xi). \end{gathered} \tag{4-74}$$

(3) We now have the decomposition for $\varepsilon > 0$

$$d\mathcal{C}_\omega = T_{\mathrm{Diag}}^+ + T_{\mathrm{Diag}}^- + (\gamma_\lambda^+)^* \widehat{T}_{\mathrm{Ref}}^+ + (\gamma_\lambda^-)^* \widehat{T}_{\mathrm{Ref}}^-. \tag{4-75}$$

Taking the limit as $\varepsilon \to 0+$ and using Lemma 3.7 (see also (3-34)) and Lemma 4.10 we see that the same decomposition holds for $\varepsilon = 0$, where we have

$$K_{\mathrm{Ref}}^\pm(\theta, \theta') = \chi_{\mathrm{Ref}}^\pm(\theta, \theta')\tilde{c}_\omega^\pm(\theta')(\gamma_\lambda^\pm(\theta) - \theta' \pm i0)^{-1} + \mathscr{K}_\omega^\pm(\theta, \theta'), \quad \text{when } \varepsilon = 0,$$

and by Lemma 3.9 the properties (4-74) hold for $\varepsilon = 0$.

The operator $T_{\mathrm{Diag}} := T^+_{\mathrm{Diag}} + T^-_{\mathrm{Diag}}$ lies in $\Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1)$ and has principal symbol $-2\pi i c_\omega \operatorname{sgn}\xi$ (away from $\xi = 0$), which is elliptic. Let $\mathcal{E}_\omega$ be the elliptic parametrix of T_{Diag}, so that $\mathcal{E}_\omega T_{\mathrm{Diag}} = I + \Psi^{-\infty}$ (see [Hörmander 1994, Theorem 18.1.9]). We have $\sigma(\mathcal{E}_\omega) = 1/\sigma(T_{\mathrm{Diag}}) = i \operatorname{sgn}\xi/(2\pi c_\omega)$. Multiplying (4-75) on the left by $\mathcal{E}_\omega$ we get (4-72) where the operators $A^\pm_\omega$ have the form

$$A^\pm_\omega = (\gamma^\pm_\lambda)^* \mathcal{E}_\omega (\gamma^\pm_\lambda)^* \widehat{T}^\pm_{\mathrm{Ref}}.$$

By [Hörmander 1994, Theorem 18.1.17], $(\gamma^\pm_\lambda)^* \mathcal{E}_\omega (\gamma^\pm_\lambda)^* \in \Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1)$ has the principal symbol $-i \operatorname{sgn}\xi/(2\pi c_\omega)$ (as $\gamma^\pm_\lambda$ is orientation-reversing), so from (4-74) we get the needed properties of $A^\pm_\omega$, with

$$a^\pm_\omega(\theta) = -\frac{\tilde{c}^\pm_\omega(\theta)}{c_\omega} \partial_\theta \gamma^\pm_\lambda(\theta) = -1 + \mathcal{O}(\varepsilon). \qquad \square$$

4.6.8. *A strong convergence statement.* A corollary of Lemma 4.10 and Proposition 4.15 is the following limiting statement:

Lemma 4.16. *Assume that* $\lambda \in (0, 1)$, Ω *is* λ*-simple*, $k \in \mathbb{N}_0$, *and* $s + 1 > t$. *Then*

$$\|\partial^k_\omega \mathcal{C}_{\omega_j} - \partial^k_\lambda \mathcal{C}_{\lambda+i0}\|_{H^{s+k}(\partial\Omega; T^*\partial\Omega) \to H^t(\partial\Omega)} \to 0 \quad \textit{for all } \omega_j \to \lambda,\ \operatorname{Im}\omega_j > 0. \tag{4-76}$$

Proof. (1) Fix k. We first show the following uniform bound: for each s there exists C_s such that, for all large j,

$$\|\partial^k_\omega \mathcal{C}_{\omega_j}\|_{H^{s+k}(\partial\Omega; T^*\partial\Omega) \to H^{s+1}(\partial\Omega)} \le C_s. \tag{4-77}$$

Indeed, Proposition 4.15 (more precisely, (4-75)) and Remark (1) after it imply that

$$\|d\partial^k_\omega \mathcal{C}_{\omega_j}\|_{H^{s+k}(\partial\Omega; T^*\partial\Omega) \to H^s(\partial\Omega; T^*\partial\Omega)} \le C_s \quad \text{for all } s, \tag{4-78}$$

where the loss of k derivatives comes from differentiating the pullback operators $\gamma^\pm_\lambda$ in $\lambda = \operatorname{Re}\omega$. On the other hand Lemma 4.10 shows that, for each $\varphi \in C^\infty(\partial\Omega \times \partial\Omega)$, and denoting by $\partial^k_\omega \mathcal{C}_\omega(x, y)$ the Schwartz kernel of the operator $\partial^k_\omega \mathcal{C}_\omega$, the sequence

$$\int_{\partial\Omega\times\partial\Omega} \partial^k_\omega \mathcal{C}_{\omega_j}(x, y)\varphi(x, y)\, d\theta(x)\, d\theta(y)$$

converges (to the same integral for $\partial^k_\lambda \mathcal{C}_{\lambda+i0}$) and thus in particular is bounded. By the Banach–Steinhaus theorem in the Fréchet space $C^\infty(\partial\Omega \times \partial\Omega)$, we see that there exists N_k such that

$$\|\partial^k_\omega \mathcal{C}_{\omega_j}\|_{H^s(\partial\Omega; T^*\partial\Omega) \to H^t(\partial\Omega)} \le C_{s,t} \quad \text{for all } s \ge N_k,\ t \le -N_k. \tag{4-79}$$

(Another way to show (4-79), avoiding Banach–Steinhaus, would be to carefully examine the proof of Lemma 4.10.)

Together (4-78), (4-79), and the elliptic estimate for ∂_θ imply that (4-77) holds for all $s \ge N_k - k$. The operator $\partial^k_\omega \mathcal{C}_{\omega_j}$ is its own transpose under the natural bilinear pairing on $C^\infty(\partial\Omega; T^*\partial\Omega) \times C^\infty(\partial\Omega)$. Since H^{-s} is dual to H^s under this pairing, (4-77) holds for all $s \le -N_k - 1$. Then (4-79) holds for all s, t such that $t \le \min(s + 1 - k, -N_k)$. Together with (4-78) and the elliptic estimate for ∂_θ, this implies that (4-77) holds in general. Same bound holds for the operator $\partial^k_\lambda \mathcal{C}_{\lambda+i0}$.

(2) We now show that

$$\partial_\omega^k \mathcal{C}_{\omega_j} v \to \partial_\lambda^k \mathcal{C}_{\lambda+i0} v \quad \text{in } C^\infty(\partial\Omega), \text{ for all } v \in C^\infty(\partial\Omega; T^*\partial\Omega). \tag{4-80}$$

Indeed, by (4-77) the sequence $\partial_\omega^k \mathcal{C}_{\omega_j} v$ is precompact in H^s for every s, and any convergent subsequence has to converge to $\partial_\lambda^k \mathcal{C}_{\lambda+i0} v$ since $\partial_\omega^k \mathcal{C}_{\omega_j} v \to \partial_\lambda^k \mathcal{C}_{\lambda+i0} v$ in $\mathcal{D}'$ by Lemma 4.10.

Since C^∞ is dense in H^{s+k}, we get from (4-77), (4-80), and a standard argument in functional analysis the strong-operator convergence

$$\partial_\omega^k \mathcal{C}_{\omega_j} v \to \partial_\lambda^k \mathcal{C}_{\lambda+i0} v \quad \text{in } H^{s+1}(\partial\Omega), \text{ for all } v \in H^{s+k}(\partial\Omega; T^*\partial\Omega). \tag{4-81}$$

We are now ready to prove (4-76). Let $s+1 > t$. Assume that (4-76) fails; then by passing to a subsequence we may assume that there exists some $c > 0$ and a sequence

$$v_j \in H^{s+k}(\partial\Omega; T^*\partial\Omega), \quad \|v_j\|_{H^{s+k}} = 1, \quad \|(\partial_\omega^k \mathcal{C}_{\omega_j} - \partial_\lambda^k \mathcal{C}_{\lambda+i0}) v_j\|_{H^t} \ge c.$$

Since H^{s+k} embeds compactly into H^{t-1+k}, passing to a subsequence we may assume that $v_j \to v_0$ in H^{t-1+k}. But then

$$\|(\partial_\omega^k \mathcal{C}_{\omega_j} - \partial_\lambda^k \mathcal{C}_{\lambda+i0}) v_j\|_{H^t} \le \|(\partial_\omega^k \mathcal{C}_{\omega_j} - \partial_\lambda^k \mathcal{C}_{\lambda+i0})(v_j - v_0)\|_{H^t} + \|(\partial_\omega^k \mathcal{C}_{\omega_j} - \partial_\lambda^k \mathcal{C}_{\lambda+i0}) v_0\|_{H^t}.$$

Now the first term on the right-hand side goes to 0 as $j \to \infty$ by (4-77), and the second term goes to 0 by (4-81), giving a contradiction. □

4.6.9. *Action on conormal distributions.* We finish this section by showing that $\mathcal{C}_\omega$ is bounded uniformly as $\operatorname{Im}\omega \to 0+$ on conormal spaces $I^s(\partial\Omega; N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+)$ defined in (3-19), where $\Sigma_\lambda^\pm$ are the attractive/repulsive sets of the chess billiard $b(\bullet, \lambda)$ defined in (1-6) and $\lambda = \operatorname{Re}\omega$ — see Lemma 4.17 below. Moreover, we get similar estimates on all the derivatives $\partial_\omega^k \mathcal{C}_\omega$. This is used in the proof of Proposition 7.4 below.

Since the conormal spaces above depend on λ, we introduce a λ-dependent system of coordinates which maps $\Sigma_\lambda^\pm$ to λ-independent sets. Assume that $\mathcal{J} \subset (0, 1)$ is an open interval such that the Morse–Smale conditions hold for each $\lambda \in \mathcal{J}$ (see Definition 1.2). Recall from Lemma 2.6 that the points in the sets $\Sigma_\lambda^\pm$ depend smoothly on $\lambda \in \mathcal{J}$. Fix any finite set $\widetilde{\Sigma} \subset \mathbb{S}^1$ with the same number of points as $\Sigma_\lambda = \Sigma_\lambda^+ \sqcup \Sigma_\lambda^-$ and a family of orientation-preserving diffeomorphisms depending smoothly on λ

$$\Theta_\lambda : \mathbb{S}^1 \to \partial\Omega, \quad \lambda \in \mathcal{J}, \quad \Theta_\lambda(\widetilde{\Sigma}) = \Sigma_\lambda.$$

We may take the decomposition $\widetilde{\Sigma} = \widetilde{\Sigma}^+ \sqcup \widetilde{\Sigma}^-$, where $\widetilde{\Sigma}^\pm$ are λ-independent sets and

$$\Theta_\lambda(\widetilde{\Sigma}^\pm) = \Sigma_\lambda^\pm \quad \text{for all } \lambda \in \mathcal{J}. \tag{4-82}$$

Note that for any fixed $\lambda \in \mathcal{J}$ the pullback Θ_λ^* gives an isomorphism

$$\Theta_\lambda^* : I^s(\partial\Omega, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+) \to I^s(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+)$$

and the space on the right-hand side is independent of λ.

For $\omega \in \mathcal{J} + i(0, \infty)$ define the conjugated operator (here Θ_λ^{-*} is the pullback by Θ_λ^{-1})

$$\widetilde{\mathcal{C}}_\omega := \Theta_\lambda^* \mathcal{C}_\omega \Theta_\lambda^{-*} : C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1) \to C^\infty(\mathbb{S}^1), \quad \text{where } \lambda := \operatorname{Re} \omega. \tag{4-83}$$

We write $\omega = \lambda + i\varepsilon$ and define $\partial_\lambda^k \widetilde{\mathcal{C}}_\omega$ by differentiating in λ with ε fixed. (Note that $\mathcal{C}_\omega$ is holomorphic in ω by Lemma 4.10 but $\widetilde{\mathcal{C}}_\omega$ is not holomorphic.)

We say that a sequence of operators

$$T_j : I^{s+}(\mathbb{S}^1, N_+^* \widetilde{\Sigma}^- \sqcup N_-^* \widetilde{\Sigma}^+) \to I^{t+}(\mathbb{S}^1, N_+^* \widetilde{\Sigma}^- \sqcup N_-^* \widetilde{\Sigma}^+)$$

is bounded uniformly in j if for each sequence $\tilde{v}_j \in I^{s+}(\mathbb{S}^1, N_+^* \widetilde{\Sigma}^- \sqcup N_-^* \widetilde{\Sigma}^+)$ with every seminorm (3-20) bounded uniformly in j, the sequence $T_j \tilde{v}_j$ also has all the seminorms (3-20) bounded uniformly in j. Similarly we consider operators acting on differential forms on $\mathbb{S}^1$, which are identified with functions using the canonical coordinate θ.

Lemma 4.17. *Assume that $\lambda \in \mathcal{J}$ and $\omega_j \to \lambda$, $\operatorname{Im} \omega_j > 0$. Then for each k and s, the sequence of operators*

$$\partial_\lambda^k \widetilde{\mathcal{C}}_{\omega_j} : I^{s+}(\mathbb{S}^1, N_+^* \widetilde{\Sigma}^- \sqcup N_-^* \widetilde{\Sigma}^+) \to I^{s-1+}(\mathbb{S}^1, N_+^* \widetilde{\Sigma}^- \sqcup N_-^* \widetilde{\Sigma}^+)$$

is bounded uniformly in j.

Proof. (1) From (4-77) we see that for each r, $\partial_\lambda^k \widetilde{\mathcal{C}}_{\omega_j}$ is bounded $H^r \to H^{r-k+1}$ uniformly in j. By elliptic regularity, it then suffices to show that the sequence of operators

$$d\partial_\lambda^k \widetilde{\mathcal{C}}_{\omega_j} : I^{s+}(\mathbb{S}^1, N_+^* \widetilde{\Sigma}^- \sqcup N_-^* \widetilde{\Sigma}^+) \to I^{s+}(\mathbb{S}^1, N_+^* \widetilde{\Sigma}^- \sqcup N_-^* \widetilde{\Sigma}^+)$$

is bounded uniformly in j. Using the decomposition (4-75), we write

$$d\widetilde{\mathcal{C}}_\omega = \widetilde{T}_{\mathrm{Diag},\omega} + (\tilde{\gamma}_\lambda^+)^* \widetilde{T}_{\mathrm{Ref},\omega}^+ + (\tilde{\gamma}_\lambda^-)^* \widetilde{T}_{\mathrm{Ref},\omega}^-, \tag{4-84}$$

where $\omega = \lambda + i\varepsilon$,

$$\widetilde{T}_{\mathrm{Diag},\omega} := \Theta_\lambda^* (T_{\mathrm{Diag}}^+ + T_{\mathrm{Diag}}^-) \Theta_\lambda^{-*}, \quad \widetilde{T}_{\mathrm{Ref},\omega}^\pm := \Theta_\lambda^* \widehat{T}_{\mathrm{Ref}}^\pm \Theta_\lambda^{-*}$$

are families of pseudodifferential operators in $\Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1)$ smooth in $\lambda \in \mathcal{J}$ uniformly in ε (see Remark (1) following Proposition 4.15), and

$$\tilde{\gamma}_\lambda^\pm := \Theta_\lambda^{-1} \circ \gamma^\pm(\bullet, \lambda) \circ \Theta_\lambda$$

is a family of orientation-reversing involutive diffeomorphisms of $\mathbb{S}^1$ depending smoothly on $\lambda \in \mathcal{J}$ and such that by (2-2) and (4-82)

$$\tilde{\gamma}_\lambda^\pm(\widetilde{\Sigma}^+) = \widetilde{\Sigma}^-, \quad \tilde{\gamma}_\lambda^\pm(\widetilde{\Sigma}^-) = \widetilde{\Sigma}^+. \tag{4-85}$$

(2) Differentiating (4-84) in $\lambda = \operatorname{Re} \omega$, we see that it suffices to show that for each k and s the sequences of operators

$$\partial_\lambda^k \widetilde{T}_{\mathrm{Diag},\omega_j}, \ \partial_\lambda^k \widetilde{T}_{\mathrm{Ref},\omega_j}^\pm, \ \partial_\lambda^k (\tilde{\gamma}_{\lambda_j}^\pm)^* : I^{s+}(\mathbb{S}^1, N_+^* \widetilde{\Sigma}^- \sqcup N_-^* \widetilde{\Sigma}^+) \to I^{s+}(\mathbb{S}^1, N_+^* \widetilde{\Sigma}^- \sqcup N_-^* \widetilde{\Sigma}^+)$$

are bounded uniformly in j. The operators $\partial_\lambda^k \widetilde{T}_{\mathrm{Diag},\omega_j}$ and $\partial_\lambda^k \widetilde{T}_{\mathrm{Ref},\omega_j}^\pm$ are bounded in Ψ^0 uniformly in j and thus bounded on any space of conormal distributions [Hörmander 1994, Theorem 18.2.7], so it remains to show the boundedness of $\partial_\lambda^k (\tilde{\gamma}_{\lambda_j}^\pm)^*$.

Instead of pullback on 1-forms we study pullback on functions, since the two differ by a multiplication operator which can be put into $\widetilde{T}^{\pm}_{\mathrm{Ref},\omega}$. We then have for all $\lambda \in \mathcal{J}$

$$\partial_\lambda (\tilde{\gamma}^{\pm}_{\lambda})^* = X^{\pm}_{\lambda} (\tilde{\gamma}^{\pm}_{\lambda})^*,$$

where $X^{\pm}_{\lambda}$ is the vector field on $\mathbb{S}^1$ given by

$$X^{\pm}_{\lambda}(\theta) = \frac{\partial_\lambda \tilde{\gamma}^{\pm}_{\lambda}(\theta)}{\partial_\theta \tilde{\gamma}^{\pm}_{\lambda}(\theta)}\, \partial_\theta.$$

We note that $X^{\pm}_{\lambda}$ vanishes on $\widetilde{\Sigma}$ by (4-85).

It follows that $\partial^k_\lambda (\tilde{\gamma}^{\pm}_{\lambda})^*$ is a linear combination with constant coefficients of operators of the form

$$(\partial^{k_1}_\lambda X^{\pm}_{\lambda}) \cdots (\partial^{k_\ell}_\lambda X^{\pm}_{\lambda})(\tilde{\gamma}^{\pm}_{\lambda})^*, \quad k_1 + \cdots + k_\ell + \ell = k.$$

Thus it remains to show that for all k the operators

$$\partial^k_\lambda X^{\pm}_{\lambda_j},\ (\tilde{\gamma}^{\pm}_{\lambda_j})^* : I^{s+}(\mathbb{S}^1, N^*_+\widetilde{\Sigma}^- \sqcup N^*_-\widetilde{\Sigma}^+) \to I^{s+}(\mathbb{S}^1, N^*_+\widetilde{\Sigma}^- \sqcup N^*_-\widetilde{\Sigma}^+) \tag{4-86}$$

are bounded uniformly in j.

Each $\partial^k_\lambda X^{\pm}_{\lambda}$ is a vector field which vanishes on $\widetilde{\Sigma}$ and thus can be written in the form $a\rho\partial_\theta$ for some $a \in C^\infty(\mathbb{S}^1)$ depending smoothly on λ and ρ which is a defining function of $\widetilde{\Sigma}$ (see the discussion preceding (3-20)). Thus $\partial^k_\lambda X^{\pm}_{\lambda_j}$ is bounded on the spaces (4-86) uniformly in j. Finally, $(\tilde{\gamma}^{\pm}_{\lambda_j})^*$ is bounded on these spaces uniformly in j by the mapping property (4-85) and since $\tilde{\gamma}^{\pm}_{\lambda_j}$ is orientation reversing, its symplectic lift maps $N^*_+\widetilde{\Sigma}^-$ and $N^*_-\widetilde{\Sigma}^+$ to each other. □

5. High frequency analysis on the boundary

In this section, we take

$$\omega = \lambda + i\varepsilon, \quad 0 < \varepsilon \ll 1,$$

where $\lambda \in (0, 1)$ satisfies the Morse–Smale conditions on Ω (see Definition 1.2), and consider the elliptic boundary value problem (4-21):

$$P(\omega)u_\omega = f, \quad u_\omega|_{\partial\Omega} = 0.$$

Here $f \in C^\infty_{\mathrm{c}}(\Omega)$ is fixed and the solution u_ω lies in $C^\infty(\overline{\Omega})$ (see Lemma 4.4). Our goal is to prove high-frequency estimates on u_ω which are uniform in the limit $\varepsilon \to 0+$, when the operator $P(\omega)$ becomes hyperbolic. To do this we combine the detailed analysis of Section 4.6 with the dynamical properties following from the Morse–Smale conditions.

5.1. ***Splitting into positive and negative frequencies.*** Fix a positively oriented coordinate $\theta : \partial\Omega \to \mathbb{S}^1$ to identify $\partial\Omega$ with $\mathbb{S}^1$. Recall from (4-30) that

$$\mathcal{C}_\omega v_\omega = G_\omega := (R_\omega f)|_{\partial\Omega}. \tag{5-1}$$

Here the 1-form $v_\omega := \mathcal{N}_\omega u_\omega \in C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1)$ is the "Neumann data" of u_ω defined using (4-26); however, we do not have uniform bounds on v_ω in C^∞ as $\varepsilon = \operatorname{Im}\omega \to 0+$. The function G_ω lies in $C^\infty(\mathbb{S}^1)$

uniformly in ε since $f \in C_c^\infty$ and R_ω is the convolution operator with the fundamental solution E_ω, which has a distributional limit as $\varepsilon \to 0+$ by Lemma 4.3.

Let $\gamma_\lambda^\pm$ be defined in (4-43). By Proposition 4.15 we have

$$\mathcal{E}_\omega d G_\omega = \mathcal{E}_\omega d \mathcal{C}_\omega v_\omega = v_\omega + (\gamma_\lambda^+)^* A_\omega^+ v_\omega + (\gamma_\lambda^-)^* A_\omega^- v_\omega.$$

We rewrite this equation as

$$v_\omega = -\mathcal{A}_\omega v_\omega + \mathcal{E}_\omega d G_\omega, \quad \mathcal{A}_\omega := (\gamma_\lambda^+)^* A_\omega^+ + (\gamma_\lambda^-)^* A_\omega^-. \tag{5-2}$$

The operator $\mathcal{A}_\omega$ exchanges positive and negative frequencies, since $A_\omega^\pm$ are pseudodifferential and the maps $\gamma_\lambda^\pm$ are orientation reversing. We thus study the square of $\mathcal{A}_\omega$, which maps positive and negative frequencies to themselves. It is expressed in terms of the pullback of the chess billiard map $b = \gamma^+ \circ \gamma^-$ to $\mathbb{S}^1$:

$$b_\lambda := \gamma_\lambda^+ \circ \gamma_\lambda^-, \quad b_\lambda^{-1} = \gamma_\lambda^- \circ \gamma_\lambda^+, \tag{5-3}$$

which is an orientation-preserving diffeomorphism of $\mathbb{S}^1$. Denote the pullback operators by b_λ and b_λ^{-1} on 1-forms by

$$b_\lambda^*, b_\lambda^{-*} : C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1) \to C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1).$$

Lemma 5.1. *We have*

$$\mathcal{A}_\omega^2 = B_\omega^+ b_\lambda^* + B_\omega^- b_\lambda^{-*}, \tag{5-4}$$

where $B_\omega^\pm$ are pseudodifferential; more precisely we have uniformly in $\varepsilon \geq 0$ (see (3-7))

$$\begin{gathered} B_\omega^\pm \in \Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1), \quad \mathrm{WF}(B_\omega^\pm) \subset \{\pm\xi > 0\}, \\ \sigma(B_\omega^\pm)(\theta, \xi) = \tilde{a}_\omega^\pm(\theta) H(\pm\xi) e^{-\varepsilon \tilde{z}_\omega^\pm(\theta)|\xi|}, \end{gathered} \tag{5-5}$$

where H denotes the Heaviside function, the functions $\tilde{a}_\omega^\pm(\theta)$, $\tilde{z}_\omega^\pm(\theta)$ are smooth in $\theta \in \mathbb{S}^1$ and $\varepsilon \geq 0$, $\operatorname{Re} \tilde{z}_\omega^\pm \geq c > 0$, and $\tilde{a}_\omega^\pm(\theta) = 1 + \mathcal{O}(\varepsilon)$.

Remark. From Remark (1) after Proposition 4.15 we see that $B_\omega^\pm$ are smooth in λ (where $\omega = \lambda + i\varepsilon$), with λ-derivatives of all orders lying in Ψ^0 uniformly in ε.

Proof. From Proposition 4.15 and the change of variables formula for pseudodifferential operators [Hörmander 1994, Theorem 18.1.17] we see that $(\gamma_\lambda^\pm)^* A_\omega^\pm (\gamma_\lambda^\pm)^*$ lies in $\Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1)$ and has wavefront set inside $\{\mp\xi > 0\}$ uniformly in ε. Since products of pseudodifferential operators with nonintersecting wavefront sets are smoothing, we see that

$$((\gamma_\lambda^\pm)^* A_\omega^\pm)^2 \in \Psi^{-\infty}(\mathbb{S}^1; T^*\mathbb{S}^1) \quad \text{uniformly in } \varepsilon \geq 0.$$

Recalling (5-2) we see that (with $\Psi^{-\infty}$ denoting smoothing operators uniformly in ε)

$$\mathcal{A}_\omega^2 = (\gamma_\lambda^+)^* A_\omega^+ (\gamma_\lambda^-)^* A_\omega^- + (\gamma_\lambda^-)^* A_\omega^- (\gamma_\lambda^+)^* A_\omega^+ + \Psi^{-\infty}.$$

This gives the decomposition (5-4) with

$$B_\omega^\pm = ((\gamma_\lambda^\mp)^* A_\omega^\mp (\gamma_\lambda^\mp)^*)((b_\lambda^{\pm 1})^* A_\omega^\pm (b_\lambda^{\mp 1})^*) + \Psi^{-\infty}.$$

Using the properties of $A_\omega^\pm$ in Proposition 4.15 together with the product formula and the change of variables formula for pseudodifferential operators, we see that $B_\omega^\pm \in \Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1)$ and $\mathrm{WF}(B_\omega^\pm) \subset \{\pm\xi > 0\}$ uniformly in ε. This also gives

$$\sigma(B_\omega^\pm)(\theta,\xi) = \sigma(A_\omega^\mp)\Big(\gamma_\lambda^\mp(\theta), \frac{\xi}{\partial_\theta \gamma_\lambda^\mp(\theta)}\Big)\sigma(A_\omega^\pm)\Big(b_\lambda^{\pm 1}(\theta), \frac{\xi}{\partial_\theta b_\lambda^{\pm 1}(\theta)}\Big)$$

in the sense of (3-7), which implies the formula for the principal symbol in (5-5), with

$$\tilde a_\omega^\pm(\theta) = a_\omega^\mp(\gamma_\lambda^\mp(\theta)) a_\omega^\pm(b_\lambda^{\pm 1}(\theta)), \quad \tilde z_\omega^\pm(\theta) = \frac{z_\omega^\mp(\gamma_\lambda^\mp(\theta))}{|\partial_\theta \gamma_\lambda^\mp(\theta)|} + \frac{z_\omega^\pm(b_\lambda^{\pm 1}(\theta))}{\partial_\theta b_\lambda^{\pm 1}(\theta)},$$

where $a_\omega^\pm$, $z_\omega^\pm$ are given in Proposition 4.15. □

Applying (5-2) twice, we get the equation

$$v_\omega = B_\omega^+ b_\lambda^* v_\omega + B_\omega^- b_\lambda^{-*} v_\omega + g_\omega, \tag{5-6}$$

where

$$g_\omega := (I - \mathcal{A}_\omega)\mathcal{E}_\omega d G_\omega$$

is in $C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1)$ uniformly in $\varepsilon > 0$.

We now split v_ω into positive and negative frequencies. Consider a pseudodifferential partition of unity

$$\begin{gathered} I = \Pi^+ + \Pi^-, \quad \Pi^\pm \in \Psi^0(\mathbb{S}^1, T^*\mathbb{S}^1), \\ \mathrm{WF}(\Pi^\pm) \subset \{\pm\xi > 0\}, \quad \sigma(\Pi^\pm)(\theta,\xi) = H(\pm\xi). \end{gathered} \tag{5-7}$$

Put

$$v_\omega^\pm := \Pi^\pm v_\omega, \quad g_\omega^\pm := \Pi^\pm g_\omega, \tag{5-8}$$

with $g_\omega^\pm$ in $C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1)$ uniformly in ε, and apply $\Pi^\pm$ to (5-6) to get

$$v_\omega^\pm = B_\omega^\pm (b_\lambda^{\pm 1})^* v_\omega^\pm + \mathscr{R}_\omega^\pm v_\omega + g_\omega^\pm, \tag{5-9}$$

where the operator

$$\mathscr{R}_\omega^\pm := \big([\Pi^\pm, B_\omega^\pm] + B_\omega^\pm(\Pi^\pm - (b_\lambda^{\pm 1})^* \Pi^\pm (b_\lambda^{\mp 1})^*)\big)(b_\lambda^{\pm 1})^* + \Pi^\pm B_\omega^\mp (b_\lambda^{\mp 1})^*$$

is in $\Psi^{-\infty}(\mathbb{S}^1; T^*\mathbb{S}^1)$ uniformly in ε, as follows from (5-7) and the fact that $\mathrm{WF}(B_\omega^\pm) \subset \{\pm\xi > 0\}$.

5.2. ***Microlocal Lasota–Yorke inequalities.*** We now show that $B_\omega^\pm (b_\lambda^{\pm 1})^*$ featured in (5-9) are contractions at high frequencies on appropriately chosen inhomogeneous Sobolev spaces, and use this to prove a high-frequency estimate on v_ω; see Proposition 5.3 below. This is reminiscent of Lasota–Yorke inequalities (see [Baladi 2018]) and could be considered a simple version of radial estimates (see [Dyatlov and Zworski 2019a, §E.4.3]) for Fourier integral operators. It is also related to microlocal weights used by Faure, Roy and Sjöstrand [Faure et al. 2008].

Unlike applications to volume-preserving Anosov maps in [Baladi 2018; Faure et al. 2008], where critical regularity is given by L^2, for us the critical regularity space is $H^{-1/2}$. This can be informally explained as follows: if we have $v_\omega^\pm = df^\pm$ for some functions $f^\pm$ then the *flux* $\operatorname{Im}\int_{\mathbb{S}^1} \overline{f^\pm}\, df^\pm$ is invariant under replacing $f^\pm$ with the pullback $(b_\lambda^{\pm 1})^* f^\pm$ and is well-defined for $f^\pm \in H^{1/2}$. When $\mathrm{WF}(f^\pm) \subset \{\pm\xi > 0\}$, the flux is related to $\|f^\pm\|^2_{H^{1/2}} \sim \|v_\omega^\pm\|^2_{H^{-1/2}}$.

To simplify notation, we only study in detail the case of the "+" sign. The case of the "−" sign is handled similarly by replacing b_λ with b_λ^{-1}, switching Σ_λ^+ with Σ_λ^-, and using the escape function in Lemma 2.9 (rather than in the remark following it).

We identify $\partial\Omega$ with $\mathbb{S}^1$ using the adapted coordinate θ constructed in Lemma 2.8, which satisfies for $\delta > 0$ small enough

$$\mp \log \partial_\theta b_\lambda > 0 \quad \text{on } \overline{\Sigma_\lambda^\pm(\delta)}, \tag{5-10}$$

where $\Sigma_\lambda^\pm \subset \mathbb{S}^1$ are the attractive (+) and repulsive (−) periodic points of b_λ defined in (1-6), and $\Sigma_\lambda^\pm(\delta)$ are their open δ-neighborhoods.

Take arbitrary $\alpha_- < \alpha_+$ and small $\delta > 0$ (in particular, so that (5-10) holds). Let $\boldsymbol{g} \in C^\infty(\mathbb{S}^1; \mathbb{R})$ be the escape function defined in the remark following Lemma 2.9. We have

$$\alpha_- \le \boldsymbol{g}(\theta) \le N_0 \quad \text{for some } N_0. \tag{5-11}$$

Define the symbol

$$G(\theta, \xi) := \boldsymbol{g}(\theta)(1 - \chi_0(\xi)) \log|\xi|, \quad (\theta, \xi) \in T^*\mathbb{S}^1, \tag{5-12}$$

where $\chi_0 \in C_c^\infty((-1, 1))$ is equal to 1 near 0. We use Lemma 3.2 to construct

$$\begin{gathered} E_G := \mathrm{Op}(e^G) \in \Psi_{0+}^{N_0}(\mathbb{S}^1; T^*\mathbb{S}^1), \quad \widetilde{E}_{-G} := \mathrm{Op}(e^{-G}(1 + r_G)) \in \Psi_{0+}^{-\alpha_-}(\mathbb{S}^1; T^*\mathbb{S}^1), \\ r_G \in S^{-1+}, \quad \widetilde{E}_{-G} E_G - I,\, E_G \widetilde{E}_{-G} - I \in \Psi^{-\infty}. \end{gathered} \tag{5-13}$$

By property (4) in the remark following Lemma 2.9 we have $\boldsymbol{g} \ge \alpha_+$ on $\mathbb{S}^1 \setminus \Sigma_\lambda^-(\delta)$. Therefore by (3-11)

$$\chi \widetilde{E}_{-G} \in \Psi_{0+}^{-\alpha_+}(\mathbb{S}^1; T^*\mathbb{S}^1) \quad \text{for all } \chi \in C^\infty(\mathbb{S}^1), \quad \operatorname{supp} \chi \cap \overline{\Sigma_\lambda^-(\delta)} = \varnothing. \tag{5-14}$$

We now apply E_G to (5-9) (with the "+" sign) to get

$$\begin{gathered} v_G = T_G v_G + \mathscr{R}_G v_\omega + g_G, \quad \text{where } v_G := E_G v_\omega^+,\ g_G := E_G g_\omega^+, \\ T_G := E_G B_\omega^+ b_\lambda^* \widetilde{E}_{-G}, \quad \mathscr{R}_G := E_G B_\omega^+ b_\lambda^* (I - \widetilde{E}_{-G} E_G)\Pi^+ + E_G \mathscr{R}_\omega^+. \end{gathered} \tag{5-15}$$

Here $g_G \in C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1)$ and $\mathscr{R}_G \in \Psi^{-\infty}(\mathbb{S}^1; T^*\mathbb{S}^1)$, both uniformly in ε. The function v_G lies in $C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1)$ for $\varepsilon > 0$, but it is not bounded in this space uniformly in ε. We also have the following bounds for each N, which follow from (5-13) and (5-14) (writing $v_\omega^+ = \widetilde{E}_{-G} v_G + (I - \widetilde{E}_{-G} E_G) v_\omega^+$):

$$\|v_\omega^+\|_{H^{\alpha_-}} \le C\|v_G\|_{L^2} + C_N \|v_\omega\|_{H^{-N}}, \tag{5-16}$$

$$\|\chi v_\omega^+\|_{H^{\alpha_+}} \le C\|v_G\|_{L^2} + C_N \|v_\omega\|_{H^{-N}} \quad \text{if } \operatorname{supp}\chi \cap \overline{\Sigma_\lambda^-(\delta)} = \varnothing, \tag{5-17}$$

$$\|g_G\|_{L^2} \le C\|g_\omega\|_{H^{N_0}}. \tag{5-18}$$

The key result in this section is the following lemma. The point is that for $\alpha_- < -\frac{1}{2} < \alpha_+$, we can obtain a contraction property of the microlocally conjugated operator T_G:

Lemma 5.2. *Suppose that G is given by (5-12) (using a coordinate θ in which (5-10) holds) with $\boldsymbol{g}$ defined with parameters $\alpha_- < \alpha_+$, $\delta > 0$, and that T_G is defined in (5-15). Define the norm on $L^2(\mathbb{S}^1; T^*\mathbb{S}^1)$ using the coordinate θ. Then for any N and $\nu > 0$ there exists C_N such that, for all small $\varepsilon = \operatorname{Im}\omega > 0$ and all $w \in C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1)$,*

$$\|T_G w\|_{L^2} \le \Big(\max_{\pm} \sup_{\Sigma_\lambda^\pm(\delta)} (\partial_\theta b_\lambda)^{1/2+\alpha_\pm} + \nu\Big)\|w\|_{L^2} + C_N\|w\|_{H^{-N}}. \tag{5-19}$$

Proof. (1) Recalling the formula (4-71) for pullback operators on 1-forms, we see that the operator

$$(b_\lambda^{-*})(\partial_\theta b_\lambda)^{1/2} : L^2(\mathbb{S}^1; T^*\mathbb{S}^1) \to L^2(\mathbb{S}^1; T^*\mathbb{S}^1)$$

is unitary. Multiplying T_G by this operator on the right, we see that it suffices to show that

$$\|\widetilde{T}_G w\|_{L^2} \le \Big(\max_{\pm} \sup_{\Sigma_\lambda^\pm(\delta)} (\partial_\theta b_\lambda)^{1/2+\alpha_\pm} + \nu\Big)\|w\|_{L^2} + C_N\|w\|_{H^{-N}},$$
$$\text{where } \widetilde{T}_G := E_G B_\omega^+ b_\lambda^* \widetilde{E}_{-G} b_\lambda^{-*}(\partial_\theta b_\lambda)^{1/2}. \tag{5-20}$$

By (3-13) we have $b_\lambda^* \widetilde{E}_{-G} b_\lambda^{-*} = \operatorname{Op}(e^{-G_b}(1+r))$ for $G_b(\theta,\xi) := G(b_\lambda(\theta), \xi/\partial_\theta b_\lambda(\theta))$ and some $r \in S^{-1+}$. Recalling the definition (5-12) of G, we compute for $|\xi|$ large enough

$$G(\theta,\xi) - G_b(\theta,\xi) = (\boldsymbol{g}(\theta) - \boldsymbol{g}(b_\lambda(\theta)))\log|\xi| + \boldsymbol{g}(b_\lambda(\theta))\log\partial_\theta b_\lambda(\theta). \tag{5-21}$$

Since $\boldsymbol{g}(\theta) - \boldsymbol{g}(b_\lambda(\theta)) \le 0$ by property (1) in the remark following Lemma 2.9, we see that $G - G_b$ is bounded above by some constant. By (3-11) and Lemma 5.1 we then see that $\widetilde{T}_G \in \Psi^0_{0+}(\mathbb{S}^1; T^*\mathbb{S}^1)$ uniformly in ε and its principal symbol is (in the sense of (3-7))

$$\sigma(\widetilde{T}_G)(\theta,\xi) = \tilde{a}_\omega^+(\theta)H(\xi)e^{-\varepsilon\tilde{z}_\omega^+(\theta)\xi}(\partial_\theta b_\lambda(\theta))^{1/2}e^{G(\theta,\xi)-G_b(\theta,\xi)}, \quad |\xi| \ge 1.$$

Thus (5-20) follows from Lemma 3.1 once we show that there exists $C_1 > 0$ such that for all $\xi \ge C_1$

$$|\tilde{a}_\omega^+(\theta)|e^{-\varepsilon\operatorname{Re}\tilde{z}_\omega^+(\theta)\xi}(\partial_\theta b_\lambda(\theta))^{1/2}e^{G(\theta,\xi)-G_b(\theta,\xi)} \le \max_{\pm} \sup_{\Sigma_\lambda^\pm(\delta)} (\partial_\theta b_\lambda)^{1/2+\alpha_\pm}. \tag{5-22}$$

(2) Since $\tilde{a}_\omega^+(\theta) = 1 + \mathcal{O}(\varepsilon)$ and $\operatorname{Re}\tilde{z}_\omega^+(\theta) \ge c > 0$, for $\xi \ge C_1$ and C_1 large enough we have $|\tilde{a}_\omega^+(\theta)|e^{-\varepsilon\operatorname{Re}\tilde{z}_\omega^+(\theta)\xi} \le 1$. Thus (5-22) reduces to showing that for all $\xi \ge C_1$

$$\widetilde{G}(\theta,\xi) \le \max_{\pm} \sup_{\Sigma_\lambda^\pm(\delta)} \big(\tfrac{1}{2} + \alpha_\pm\big)\log\partial_\theta b_\lambda,$$
$$\text{where } \widetilde{G}(\theta,\xi) := (\boldsymbol{g}(\theta) - \boldsymbol{g}(b_\lambda(\theta)))\log\xi + \big(\tfrac{1}{2} + \boldsymbol{g}(b_\lambda(\theta))\big)\log\partial_\theta b_\lambda(\theta). \tag{5-23}$$

This in turn is proved if we show that there exists $c_0 > 0$ such that for ξ large enough

$$\widetilde{G}(\theta,\xi) \le \begin{cases} -c_0\log\xi, & \theta \in \mathbb{S}^1 \setminus (\Sigma_\lambda^-(\delta) \cup \Sigma_\lambda^+(\delta)), \\ \big(\tfrac{1}{2} + \alpha_+\big)\log\partial_\theta b_\lambda(\theta), & \theta \in \Sigma_\lambda^+(\delta), \\ \big(\tfrac{1}{2} + \alpha_-\big)\log\partial_\theta b_\lambda(\theta), & \theta \in \Sigma_\lambda^-(\delta). \end{cases} \tag{5-24}$$

We now prove (5-24) using properties (1)–(6) in Lemma 2.9 (or rather the remark which follows it). The first inequality follows from property (2), since $\boldsymbol{g}(\theta) - \boldsymbol{g}(b_\lambda(\theta)) \le -2c_0$ for some $c_0 > 0$. The second inequality follows from properties (1) and (4) together with (5-10). Finally, the third inequality follows from property (6) with $M := (\log \xi)/(\log \partial_\theta b_\lambda(\theta)) \gg 1$, where we again use (5-10). □

With Lemma 5.2 in place we give a basic high-frequency estimate on solutions to (5-6) which is uniform as $\operatorname{Im}\omega \to 0$. An upgraded version of this estimate (Proposition 5.4) is used in the proof of the limiting absorption principle in Section 7 below.

Proposition 5.3. *Fix $\beta > 0$, N, and some functions $\chi^\pm \in C^\infty(\mathbb{S}^1)$ such that* $\operatorname{supp}\chi^\pm \cap \Sigma_\lambda^\mp = \varnothing$. *Then there exist N_0 and C such that for all small $\varepsilon = \operatorname{Im}\omega > 0$ and each solution $v_\omega \in C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1)$ to (5-6) we have*

$$\|v_\omega\|_{H^{-1/2-\beta}} \le C(\|g_\omega\|_{H^{N_0}} + \|v_\omega\|_{H^{-N}}), \tag{5-25}$$

$$\|\chi^\pm \Pi^\pm v_\omega\|_{H^N} \le C(\|g_\omega\|_{H^{N_0}} + \|v_\omega\|_{H^{-N}}). \tag{5-26}$$

Remarks. (1) The a priori assumption that v_ω is smooth (without any uniformity as $\varepsilon \to 0+$) is important in the argument because it ensures that the norm $\|v_G\|_{L^2}$ is finite.

(2) Using the notation (3-18), we see that (5-26) implies that, assuming that the right-hand side of this inequality is bounded uniformly in ε for each N_0 and some N, we have $\operatorname{WF}(v_\omega) \subset N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+$ uniformly in ε.

Proof. (1) Fix $\alpha_\pm$ satisfying

$$-\tfrac{1}{2} - \beta \le \alpha_- < -\tfrac{1}{2} < \alpha_+, \quad \alpha_+ \ge N.$$

Next, fix $\delta > 0$ in the construction of the escape function $\boldsymbol{g}$ small enough so that (5-10) holds and $\operatorname{supp}\chi^+ \cap \overline{\Sigma_\lambda^-(\delta)} = \varnothing$. By (5-10) and since $\alpha_- < -\frac{1}{2} < \alpha_+$ we may choose τ such that

$$\max_\pm \sup_{\Sigma_\lambda^\pm(\delta)} (\partial_\theta b_\lambda)^{1/2+\alpha_\pm} < \tau < 1.$$

Take N_0 so that (5-11) holds. We use (5-15) and (5-18) to get

$$\|v_G\|_{L^2} \le \|T_G v_G\|_{L^2} + C(\|g_\omega\|_{H^{N_0}} + \|v_\omega\|_{H^{-N}}).$$

Applying Lemma 5.2 to $w := v_G$, we see that

$$\|T_G v_G\|_{L^2} \le \tau \|v_G\|_{L^2} + C\|v_\omega\|_{H^{-N}}.$$

Since $\tau < 1$, together these two inequalities give

$$\|v_G\|_{L^2} \le C(\|g_\omega\|_{H^{N_0}} + \|v_\omega\|_{H^{-N}}). \tag{5-27}$$

(2) From (5-27) and (5-16) we have

$$\|v_\omega^+\|_{H^{-1/2-\beta}} \le C(\|g_\omega\|_{H^{N_0}} + \|v_\omega\|_{H^{-N}}). \tag{5-28}$$

The bound (5-26) for the "+" sign follows from (5-27) and (5-17). Similar analysis (replacing b_λ with b_λ^{-1}, switching the roles of Σ_λ^+ and Σ_λ^-, and using Lemma 2.9 instead of the remark that follows it) shows that (5-28) holds for v_ω^- and (5-26) holds for the "−" sign. Since $v_\omega = v_\omega^+ + v_\omega^-$, we obtain (5-25). □

5.3. *Conormal regularity.* We now upgrade Proposition 5.3 to obtain iterated conormal regularity uniformly as $\operatorname{Im}\omega \to 0+$. We also relax the assumptions on the right-hand side g_ω: instead of being smooth uniformly in ε it only needs to be bounded in a certain conormal space uniformly in ε. This is the high-frequency estimate used in the proof of Lemma 7.1 below.

As before we identify $\partial\Omega$ with $\mathbb{S}^1$ and 1-forms on $\mathbb{S}^1$ with functions using the coordinate θ constructed in Lemma 2.8, which in particular makes it possible to define the operator ∂_θ on 1-forms. Fix some defining function ρ of $\Sigma_\lambda = \Sigma_\lambda^+ \sqcup \Sigma_\lambda^-$ and an operator $A_{\Sigma_\lambda} \in \Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1)$ such that $\operatorname{WF}(A_{\Sigma_\lambda}) \cap (N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+) = \varnothing$ and A_{Σ_λ} is elliptic on $N_-^*\Sigma_\lambda^- \sqcup N_+^*\Sigma_\lambda^+$. The estimate (5-29) below features the seminorms (3-20) for the space $I^{(1/4)+}(\mathbb{S}^1, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+)$ defined in (3-19). The proposition below applies to any $v_\omega, g_\omega \in C^\infty$ solving (5-6), not just to v_ω discussed in Section 5.1.

Proposition 5.4. *Fix $\beta > 0$, $k \in \mathbb{N}_0$, and N. Then there exist $N_0 = N_0(\beta, k)$ and $C = C(\beta, k, N)$ such that for all small $\varepsilon = \operatorname{Im}\omega > 0$ and any solution $v_\omega \in C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1)$ to (5-6) we have*

$$\|(\rho\partial_\theta)^k v_\omega\|_{H^{-1/2-\beta}} + \|A_{\Sigma_\lambda} v_\omega\|_{H^k} \le C\Big(\max_{0\le\ell\le N_0} \|(\rho\partial_\theta)^\ell g_\omega\|_{H^{-1/2-\beta}} + \|A_{\Sigma_\lambda} g_\omega\|_{H^{N_0}} + \|v_\omega\|_{H^{-N}}\Big). \quad (5\text{-}29)$$

Remark. From the Remark at the end of Section 2.3 we see that the statement of Proposition 5.4 holds locally uniformly in λ. More precisely, assume that $\mathcal{J} \subset (0, 1)$ is an open interval such that each $\lambda \in \mathcal{J}$ satisfies the Morse–Smale conditions of Definition 1.2. We may choose the coordinate θ, the defining function ρ of Σ_λ, and the operator A_{Σ_λ} depending smoothly on λ. Then for each compact set $\mathcal{K} \subset \mathcal{J}$ we may choose constants N_0 and C so that (5-29) holds for all $\lambda = \operatorname{Re}\omega \in \mathcal{K}$. This can be seen from the remark following Lemma 5.1 and the fact that the escape function $\boldsymbol{g}$ can be chosen to depend smoothly on λ.

Proof. (1) By the discussion following (3-20) it suffices to show (5-29) for one specific choice of ρ. We choose ρ such that

$$\rho^{-1}(0) = \Sigma_\lambda, \quad |\partial_\theta\rho| = 1 \quad \text{on } \Sigma_\lambda. \quad (5\text{-}30)$$

Recalling the formula (4-71) for pullback on 1-forms we have the commutation identity of operators on $C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1)$

$$\rho\partial_\theta b_\lambda^* = \varphi b_\lambda^* \rho\partial_\theta + \psi b_\lambda^*, \quad \varphi(\theta) = \frac{\rho(\theta)\partial_\theta b_\lambda(\theta)}{\rho(b_\lambda(\theta))}, \quad \psi(\theta) = \frac{\rho(\theta)\partial_\theta^2 b_\lambda(\theta)}{\partial_\theta b_\lambda(\theta)}. \quad (5\text{-}31)$$

By (5-30) and since $b_\lambda(\Sigma_\lambda) = \Sigma_\lambda$ we have $|\varphi| = 1$ on Σ_λ.

As in (5-8), let $v_\omega^\pm := \Pi^\pm v_\omega$ and $g_\omega^\pm := \Pi^\pm g_\omega$. Since $\operatorname{WF}(A_{\Sigma_\lambda}) \cap N_\pm^*\Sigma_\lambda^\mp = \varnothing$, we may fix $\chi^\pm \in C^\infty(\mathbb{S}^1)$ such that

$$\operatorname{supp}\chi^\pm \cap \Sigma_\lambda^\mp = \varnothing, \quad \chi^\pm = 1 \text{ near } \{\theta \in \mathbb{S}^1 \mid (\theta, \pm 1) \in \operatorname{WF}(A_{\Sigma_\lambda})\}.$$

We will show that there exist $N_0 = N_0(\beta, k)$ and $\tilde\chi^\pm \in C^\infty(\mathbb{S}^1)$ such that $\operatorname{supp}\tilde\chi^\pm \cap \Sigma_\lambda^\mp = \varnothing$ and

$$\|(\rho\partial_\theta)^k v_\omega^\pm\|_{H^{-1/2-\beta}} + \|\chi^\pm v_\omega^\pm\|_{H^k} \le C\Big(\max_{0\le\ell\le k} \|(\rho\partial_\theta)^\ell g_\omega^\pm\|_{H^{-1/2-\beta}} + \|\tilde\chi^\pm g_\omega^\pm\|_{H^{N_0}} + \|v_\omega\|_{H^{-N}}\Big). \quad (5\text{-}32)$$

Adding these together and using that $v_\omega = v_\omega^+ + v_\omega^-$, we get

$$\|(\rho\partial_\theta)^k v_\omega\|_{H^{-1/2-\beta}} + \|\chi^+ v_\omega^+ + \chi^- v_\omega^-\|_{H^k} \le C\Big(\max_{0\le\ell\le k} \|(\rho\partial_\theta)^\ell g_\omega\|_{H^{-1/2-\beta}} + \|\tilde\chi^+ g_\omega^+\|_{H^{N_0}} + \|\tilde\chi^- g_\omega^-\|_{H^{N_0}} + \|v_\omega\|_{H^{-N}}\Big).$$

Since $\chi^+\Pi^+ + \chi^-\Pi^-$ is elliptic on $\mathrm{WF}(A_{\Sigma_\lambda})$, we may estimate $\|A_{\Sigma_\lambda} v_\omega\|_{H^k}$ in terms of $\|\chi^+ v_\omega^+ + \chi^- v_\omega^-\|_{H^k}$. Since $\mathrm{WF}(\tilde\chi^\pm\Pi^\pm) \cap (N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+) = \varnothing$, we may estimate $\|\tilde\chi^\pm g_\omega^\pm\|_{H^{N_0}}$ by (3-21). Thus (5-32) implies (5-29) (possibly with a larger value of N_0).

(2) It remains to show (5-32). We show an estimate on v_ω^+, with the case of v_ω^- handled similarly. We start with the case $k=0$. Let E_G, $\widetilde{E}_{-G}$ be constructed in (5-13), where the escape function $\boldsymbol{g}$ is constructed using parameters $\alpha_- < \alpha_+$, $\delta > 0$ such that

$$\alpha_- = -\tfrac{1}{2} - \beta, \quad \operatorname{supp}\chi^+ \cap \overline{\Sigma_\lambda^-(\delta)} = \varnothing, \tag{5-33}$$

$$\alpha_+ \geq 0, \quad \max_{\pm} \sup_{\Sigma_\lambda^\pm(\delta)} (\partial_\theta b_\lambda)^{1/2+\alpha_\pm} < 1. \tag{5-34}$$

Using (5-15) and Lemma 5.2 similarly to the proof of Proposition 5.3, we get the inequality

$$\|v_G\|_{L^2} \leq C(\|g_G\|_{L^2} + \|v_\omega\|_{H^{-N}}), \tag{5-35}$$

where $v_G := E_G v_\omega^+$, $g_G := E_G g_\omega^+$. By (5-16) and (5-17) we have

$$\|v_\omega^+\|_{H^{-1/2-\beta}} + \|\chi^+ v_\omega^+\|_{L^2} \leq C(\|v_G\|_{L^2} + \|v_\omega\|_{H^{-N}}). \tag{5-36}$$

By property (5) in the remark following Lemma 2.9 we have $\boldsymbol{g} = \alpha_-$ on some neighborhood of Σ_λ^-. Thus we can choose $\tilde\chi^+ \in C^\infty(\mathbb{S}^1)$ such that $\operatorname{supp}\tilde\chi^+ \cap \Sigma_\lambda^- = \varnothing$ and

$$\boldsymbol{g} = \alpha_- \quad \text{near } \operatorname{supp}(1-\tilde\chi^+).$$

Then $E_G(1-\tilde\chi^+) \in \Psi_{0+}^{\alpha_-}$ by (3-11). Fix N_0 such that (5-11) holds, so that $E_G \in \Psi_{0+}^{N_0}$. Writing $g_G = E_G(1-\tilde\chi^+)g_\omega^+ + E_G\tilde\chi^+ g_\omega^+$, we get

$$\|g_G\|_{L^2} \leq C(\|g_\omega^+\|_{H^{-1/2-\beta}} + \|\tilde\chi^+ g_\omega^+\|_{H^{N_0}}). \tag{5-37}$$

Putting together (5-35)–(5-37), we get (5-32) for $k=0$.

(3) We next show (5-32) for $k=1$. Put for $j \in \mathbb{N}_0$

$$v^j := (\rho\partial_\theta)^j v_\omega^+ \in C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1), \quad v_G^j := E_G v^j, \quad g_G^j := E_G(\rho\partial_\theta)^j g_\omega^+.$$

We apply $\rho\partial_\theta$ to (5-9) and use (5-31) to get a similar equation on $v^1 = \rho\partial_\theta v_\omega^+$ which also involves $v^0 = v_\omega^+$:

$$\begin{aligned} v^1 &= B_\omega^+\varphi b_\lambda^* v^1 + Q_\omega^+ b_\lambda^* v^0 + \rho\partial_\theta(\mathscr{R}_\omega^+ v_\omega + g_\omega^+), \\ Q_\omega^+ &= [\rho\partial_\theta, B_\omega^+] + B_\omega^+\psi \in \Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1) \quad \text{uniformly in } \varepsilon. \end{aligned} \tag{5-38}$$

Applying E_G to (5-38), we get similarly to (5-15)

$$v_G^1 = T_G^1 v_G^1 + Q_G v_G^0 + \mathscr{R}_G^1 v_\omega + g_G^1, \tag{5-39}$$

where $\mathscr{R}_G^1 \in \Psi^{-\infty}$ uniformly in ε and

$$T_G^1 := E_G B_\omega^+\varphi b_\lambda^* \widetilde{E}_{-G}, \quad Q_G := E_G Q_\omega^+ b_\lambda^* \widetilde{E}_{-G}.$$

We fix the parameters $\alpha_\pm$, δ in the construction of the escape function $\boldsymbol{g}$ such that we have (5-33) and the following strengthening of (5-34):

$$\alpha_+ \ge 1, \quad \max_{\pm} \sup_{\Sigma_\lambda^\pm(\delta)} \max(1, |\varphi|)(\partial_\theta b_\lambda)^{1/2+\alpha_\pm} < 1. \tag{5-40}$$

This is possible by (5-10) and since $|\varphi| = 1$ on $\Sigma_\lambda^\pm$.

Arguing similarly to the proof of Lemma 5.2, we get the bounds for some $\tau < 1$

$$\|T_G^1 v_G^1\|_{L^2} \le \tau \|v_G^1\|_{L^2} + C\|v_\omega\|_{H^{-N}}, \quad \|Q_G v_G^0\|_{L^2} \le C\|v_G^0\|_{L^2}.$$

Combining these with (5-39) and recalling (5-35) we get

$$\begin{aligned} \|v_G^1\|_{L^2} &\le C(\|v_G^0\|_{L^2} + \|g_G^1\|_{L^2} + \|v_\omega\|_{H^{-N}}), \\ \|v_G^0\|_{L^2} &\le C(\|g_G^0\|_{L^2} + \|v_\omega\|_{H^{-N}}). \end{aligned} \tag{5-41}$$

Similarly to (5-36)–(5-37) we have for $j = 0, 1$

$$\begin{aligned} \|v^j\|_{H^{-1/2-\beta}} + \|\chi^+ v^j\|_{H^1} &\le C(\|v_G^j\|_{L^2} + \|v_\omega\|_{H^{-N}}), \\ \|g_G^j\|_{L^2} &\le C(\|(\rho\partial_\theta)^j g_\omega^+\|_{H^{-1/2-\beta}} + \|\tilde\chi^+ g_\omega^+\|_{H^{N_0}}). \end{aligned} \tag{5-42}$$

Together (5-41)–(5-42) give (5-32) for $k = 1$.

(4) The case of general k is handled similarly to $k = 1$. We write similarly to (5-38)

$$v^k = B_\omega^+ \varphi^k b_\lambda^* v^k + \sum_{j=0}^{k-1} Q_{\omega,k,j}^+ b_\lambda^* v^j + (\rho\partial_\theta)^k (\mathscr{R}_\omega^+ v_\omega + g_\omega^+).$$

Here $Q_{\omega,k,j}^+ \in \Psi^0$ uniformly in ε is defined inductively as

$$Q_{\omega,k,j}^+ := \big([\rho\partial_\theta, B_\omega^+ \varphi^{k-1}] + B_\omega^+ \varphi^{k-1}\psi\big)\delta_{j,k-1} + Q_{\omega,k-1,j-1}^+ \varphi + [\rho\partial_\theta, Q_{\omega,k-1,j}^+] + Q_{\omega,k-1,j}^+ \psi$$

and we use the notation $\delta_{a,b} = 1$ if $a = b$ and 0 otherwise, and $Q_{\omega,k-1,j}^+ = 0$ when $j \in \{-1, k-1\}$. The argument in step (3) of this proof now goes through, replacing (5-40) with

$$\alpha_+ \ge k, \quad \max_{\pm} \sup_{\Sigma_\lambda^\pm(\delta)} \max(1, |\varphi|^k)(\partial_\theta b_\lambda)^{1/2+\alpha_\pm} < 1 \tag{5-43}$$

and gives (5-32) for any value of k. □

We will also need a refinement concerning Lagrangian regularity. Let $B_{\lambda+i0}^\pm$ be the operators $B_\omega^\pm$ from Lemma 5.1 with $\varepsilon := 0$.

Lemma 5.5. *Suppose that $v \in \mathcal{D}'(\mathbb{S}^1; T^*\mathbb{S}^1)$ satisfies (5-6) with $\varepsilon = 0$:*

$$v = B_{\lambda+i0}^+ b_\lambda^* v + B_{\lambda+i0}^- b_\lambda^{-*} v + g \quad \textit{for some } g \in C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1). \tag{5-44}$$

Similarly to (5-8) define $v^\pm := \Pi^\pm v$. Then

$$v^\pm \in I^{(1/4)+}(\mathbb{S}^1, N^*\Sigma_\lambda^\mp) \quad \Longrightarrow \quad v^\pm \in I^{1/4}(\mathbb{S}^1, N^*\Sigma_\lambda^\mp). \tag{5-45}$$

Proof. (1) Let us consider v^+, with v^- handled similarly. Similarly to (5-9) we have from (5-44)

$$v^+ = B^+_{\lambda+i0} b^*_\lambda v^+ + g_1, \quad \text{where } g_1 \in C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1).$$

Iterating this n times, where n is the period of the closed trajectories of b_λ, we see that

$$v^+ = Bf^* v^+ + g_2, \quad \text{where } f := b^n_\lambda : \mathbb{S}^1 \to \mathbb{S}^1,\ g_2 \in C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1), \tag{5-46}$$

and the pseudodifferential operator

$$B := B^+_{\lambda+i0}(b^*_\lambda B^+_{\lambda+i0} b^{-*}_\lambda) \cdots ((b^{n-1}_\lambda)^* B^+_{\lambda+i0} (b^{n-1}_\lambda)^{-*}) \in \Psi^0(\mathbb{S}^1; T^*\mathbb{S}^1)$$

satisfies $\sigma(B) = H(\xi)$ by Lemma 5.1.

Take arbitrary $x_0 \in \Sigma^-_\lambda$ and assume that the coordinate θ is chosen so that $\theta(x_0) = 0$. Note that $f(0) = 0$. Fix $\chi \in C^\infty(\mathbb{R})$ supported on a small neighborhood of 0 which does not contain any other point in Σ^-_λ, and such that $\chi = 1$ near 0. We write

$$\chi v^+ = u(\theta)\, d\theta \quad \text{for some } u \in \mathcal{E}'(\mathbb{R}).$$

Then $u \in I^{\frac{1}{4}+}(\mathbb{R}, N^*\{0\})$ and by (5-46) we have

$$u = \widetilde{B} f^* u + g_3, \quad \text{where } g_3 \in C^\infty_c(\mathbb{R}) \tag{5-47}$$

and $\widetilde{B}$ is a compactly supported operator in $\Psi^0(\mathbb{R})$ such that $\sigma(\widetilde{B})(0,\xi) = f'(0)H(\xi)$. Here in (5-46) the operator f^* is the pullback on 1-forms, and in (5-47) the same symbol denotes the pullback on functions, with the two related by the formula (4-71). We can take arbitrary $\widetilde{B}$ which is equal to Bf' near $\theta = 0$, since u is smooth away from 0.

It suffices to show that $u \in I^{1/4}(\mathbb{R}, N^*\{0\})$, which (recalling (3-16)) is equivalent to $\hat{u} \in S^0(\mathbb{R})$. Note that $\hat{u}(\xi)$ is rapidly decaying as $\xi \to -\infty$ since $\mathrm{WF}(v^+) \subset \{\xi > 0\}$, so it suffices to study what happens for $\xi > 1$.

(2) We now use the invariance of the principal symbol of u coming from (5-47). More precisely, by Lemma 3.3, and since the Fourier transform $\hat{g}_3$ is rapidly decaying, (5-47) implies for $\xi > 1$

$$\hat{u}(\xi) = \hat{u}(\xi/R) + q(\xi), \quad \text{where } R := f'(0) > 1,\ q \in S^{-1+}(\mathbb{R}).$$

Iterating this, we see that for any $k \in \mathbb{N}_0$ and $\eta \geq 1$

$$\hat{u}(R^k\eta) = \hat{u}(\eta) + \sum_{\ell=1}^{k} q(R^\ell \eta). \tag{5-48}$$

We now estimate (using for simplicity that $q \in S^{-1/2}$ rather than $q \in S^{-1+}$)

$$\sup_{\xi\geq 1} |\hat{u}(\xi)| = \sup_{k\in\mathbb{N}_0} \sup_{1\leq\eta\leq R} |\hat{u}(R^k\eta)| \leq \sup_{1\leq\eta\leq R} |\hat{u}(\eta)| + C\sum_{\ell=1}^{\infty} R^{-\ell/2} < \infty.$$

Differentiating (5-48) m times in η, we similarly see that $\sup_{\xi\geq 1} \xi^m |\partial^m_\xi \hat{u}(\xi)| < \infty$. This gives $\hat{u} \in S^0(\mathbb{R})$ and finishes the proof. ☐

6. Microlocal properties of Morse–Smale maps

Here we prove properties of distributions invariant under Morse–Smale maps (see Definition 1.2). We start with a stand alone local result about distributions invariant under contracting maps. The quantum flux defined below (6-3) is reminiscent of similar quantities appearing in scattering theory — see [Dyatlov and Zworski 2019a, (3.6.17)]. The wave front condition (6-16) is an analogue of the outgoing condition in scattering theory — see [Dyatlov and Zworski 2019a, Theorem 3.37]. Although technically very different, Lemma 6.1 and Proposition 6.3 are analogous to [Dyatlov and Zworski 2017, Lemma 2.3] and play the role of that lemma in showing the absence of embedded eigenvalues — see [Dyatlov and Zworski 2019b, §3.2].

6.1. ***Local analysis.*** In this section we assume that $f : [-1, 1] \to (-1, 1)$ is a C^∞ map such that

$$f(0) = 0, \quad 0 < f'(x) < 1. \tag{6-1}$$

We also assume that

$$u \in \mathcal{D}'((-1, 1)), \quad \operatorname{singsupp} u \subset \{0\}, \quad f^*u = u \text{ on } (-1, 1). \tag{6-2}$$

For $\chi \in C_c^\infty(f(-1, 1))$, $\chi = 1$ near 0 we then define the *flux* of u (understood as an integral of a differential 1-form):

$$\boldsymbol{F}(u) := i \int_{(-1,1)} (f^*\chi - \chi)\bar{u}\, du. \tag{6-3}$$

The integral is well-defined since u is smooth on $\operatorname{supp}(f^*\chi - \chi) \subset (-1, 1) \setminus \{0\}$.

We note that $\boldsymbol{F}(u)$ is independent of χ. In fact, if $\chi_j \in C_c^\infty(f(-1, 1))$, $\chi_1 = \chi_2$ near 0, then the difference of the fluxes defined using χ_j's in place of χ, is given by (6-3) with $\tilde{\chi} = \chi_1 - \chi_2 \in C_c^\infty(f(-1, 1) \setminus \{0\})$ in place of χ. Since $\tilde{\chi}$ is supported away from 0 we can split the integral:

$$\int_{(-1,1)} (f^*\tilde{\chi} - \tilde{\chi})\bar{u}\, du = \int (f^*\tilde{\chi})\bar{u}\, du - \int f^*(\tilde{\chi}\bar{u}\, du) = \int (f^*\tilde{\chi})(\bar{u}\, du - f^*(\bar{u}\, du)) = 0.$$

Here in the first equality we made a change of variables by $f : (-1, 1) \to f(-1, 1)$ and in the last equality we used (6-2). In fact, this argument shows that we could take χ in (6-3) to be the indicator function of some interval $f(a_-, a_+)$ with $-1 < a_- < 0 < a_+ < 1$, obtaining

$$\boldsymbol{F}(u) = i \int_{[a_-, f(a_-)] \sqcup [f(a_+), a_+]} \bar{u}\, du. \tag{6-4}$$

Similarly we see that $\boldsymbol{F}(u)$ is real. For that we take χ real-valued so that

$$\begin{aligned} 2 \operatorname{Im} \boldsymbol{F}(u) &= 2 \int_{(-1,1)} (f^*\chi - \chi) \operatorname{Re}(\bar{u}\, du) = \int (f^*\chi - \chi)\, d(|u|^2) \\ &= \int |u|^2\, d(\chi - f^*\chi) = \int |u|^2\, d\chi - \int |u|^2 f^*\, d\chi = 0, \end{aligned}$$

where in the last line we used (6-2) and the fact that $\chi' = 0$ near 0.

The key local result is given in:

Lemma 6.1. *Suppose that (6-1) and (6-2) hold. Then*

$$\operatorname{WF}(u) \subset \{0\} \times \mathbb{R}_+, \quad \boldsymbol{F}(u) \geq 0 \quad \Longrightarrow \quad u = \text{const}. \tag{6-5}$$

Remark. The wavefront set restriction to positive frequencies is crucial: for example, if u is the Heaviside function, then (6-2) holds and $\boldsymbol{F}(u)=0$. A nontrivial example when (6-1), (6-2), and the wavefront set condition in (6-5) hold is $f(x)=e^{-2\pi}x$, $u(x)=(x+i0)^{ik}$, $k\in\mathbb{Z}\setminus\{0\}$, where $\boldsymbol{F}(u)=2\pi k(e^{-2\pi k}-1)<0$.

To prove Lemma 6.1 we use a standard one-dimensional linearization result [Sternberg 1957]. For the reader's convenience we present a variant of the proof from [Yoccoz 1995, Appendice 4].

Lemma 6.2. *Assume that f satisfies (6-1). Then there exists a unique C^∞ diffeomorphism $h:[-1,1]\to h([-1,1])\subset\mathbb{R}$ such that for all $x\in[-1,1]$*

$$h(f(x))=f'(0)h(x),\quad h(0)=0,\quad h'(0)=1. \tag{6-6}$$

Proof. (1) We first note that any C^1 diffeomorphism satisfying (6-6) is unique. In fact, suppose that h_j, $j=1,2$, are two such diffeomorphisms. With $a=f'(0)\in(0,1)$, $ah_j=h_j\circ f$ we have $ah_1\circ h_2^{-1}(x)=h_1\circ fh_2^{-1}(x)=h_1\circ h_2^{-1}(ax)$ for all $x\in h_2([-1,1])$, so that

$$h_1\circ h_2^{-1}(x)=a^{-n}h_1\circ h_2^{-1}(a^nx)=\lim_{n\to\infty}a^{-n}h_1\circ h_2^{-1}(a^nx)=(h_1\circ h_2^{-1})'(0)x=x.$$

Hence it is enough to show that for every n there exists a C^n diffeomorphism satisfying (6-6).

Using the fact that $a=f'(0)\in(0,1)$ we can construct a formal power series such that (6-6) holds for the Taylor series of f as an asymptotic expansion. Using Borel's lemma [Hörmander 1990, Theorem 1.2.6] we can then construct a diffeomorphism h_0 of $[-1,1]$ onto itself with that formal series as Taylor series at 0. Then $h_0\circ f\circ h_0^{-1}=ax(1+g(x))$, where $g\in C^\infty$ vanishes to infinite order at 0. Hence we can assume that

$$f(x)=ax(1+g(x)).$$

We might no longer have $f'<1$ but f is still eventually contracting: there exists $m>0$ such that the m-th iterate f^m satisfies

$$\partial_x(f^m(x))<1\quad\text{for all } x\in[-1,1]. \tag{6-7}$$

(2) We are now looking for $h(x)=x(1+\varphi(x))$, $\varphi(0)=0$ such that $h(ax(1+g(x)))=ah(x)$, that is, $ax(1+g(x))(1+\varphi(f(x)))=ax(1+\varphi(x))$, or

$$(1+g(x))(1+\varphi(f(x)))=1+\varphi(x).$$

A formal solution is then given by $1+\varphi(x)=\prod_{\ell=0}^{\infty}(1+g(f^\ell(x)))$. Rather than analyze this expression, we follow [Yoccoz 1995, Appendice 4] and use the contraction mapping principle for Banach spaces, B_n, of C^n functions on $[-\delta,\delta]$ vanishing to order $n\ge 2$ at 0: we look for $\varphi\in B_n$ such that

$$g(x)+(1+g(x))\varphi(f(x))=\varphi(x),\quad x\in[-\delta,\delta]. \tag{6-8}$$

We claim that for $\delta>0$ small enough,

$$\varphi(x)\mapsto(T\varphi)(x):=(1+g(x))\varphi(f(x))$$

is a contraction on B_n. The norm on B_n is given by

$$\|\varphi\|_{B_n}:=\sup_{|x|\le\delta}|\partial^n\varphi(x)|,\qquad \sup_{|x|\le\delta}|\partial^j\varphi(x)|\le C_n\delta^{n-j}\|\varphi\|_{B_n},\quad \varphi\in B_n,\ j\le n, \tag{6-9}$$

where the last inequality follows from Taylor's formula. Since $f(x) = ax(1+g(x))$, we have $f'(x) = a+\mathcal{O}(x^\infty)$ and $f^{(j)}(x) = \mathcal{O}(x^\infty)$ for $j > 1$. Hence, we obtain for $|x| \le \delta$, using (6-9) and with homogenous polynomials Q_j,

$$\begin{aligned}\partial^n[\varphi(f(x))] &= \partial^n\varphi(f(x))(\partial f(x))^n + \sum_{j=1}^{n-1}\partial^j\varphi(f(x))Q_j(\partial f(x), \dots, \partial^{n-j+1}f(x))\\ &= \partial^n\varphi(f(x))a^n(1+\mathcal{O}_n(\delta)) + \sum_{j=1}^{n-1}\mathcal{O}_n(\delta^{n-j})\|\varphi\|_{B_n}.\end{aligned}$$

It follows that $\|T\varphi\|_{B_n} \le (a^n + \mathcal{O}_n(\delta))\|\varphi\|_{B_n}$, which for δ small enough (depending on n) shows that T is a contraction on B_n. That gives a solution φ to (6-8). Consequently, we have shown that for every n there exist $\delta > 0$ and $\varphi \in C^n([-\delta, \delta])$ such that, for $h(x) = x(1+\varphi(x))$,

$$h(f(x)) = ah(x), \quad |x| \le \delta,\ h \in C^n([-\delta, \delta]).$$

By (6-7), there exists $N > 0$ such that $f^N([-1, 1]) \subset [-\delta, \delta]$. We extend h to $[-1, 1]$ by putting $h(x) := a^{-N}h(f^N(x))$ to obtain a C^n diffeomorphism $h : [-1, 1] \to h([-1, 1])$ satisfying (6-6). □

Proof of Lemma 6.1. (1) We first note that if $u \in C^\infty((-1, 1))$ then u is constant as follows from (6-2): for each $x \in (-1, 1)$ we have $u(x) = u(f^N(x)) \to u(0)$ as $N \to \infty$. Since we assumed that $\operatorname{singsupp} u \subset \{0\}$ it suffices to show that u is smooth in a neighborhood of 0.

Making the change of variable given by Lemma 6.2, we may assume that $f(x) = ax$ for small x, where $a := f'(0) \in (0, 1)$. Restricting to a neighborhood of 0, rescaling, and using (6-4) we reduce to the following statement: if

$$u \in \mathcal{D}'((-a^{-1}, a^{-1})), \quad \mathrm{WF}(u) \subset \{0\} \times \mathbb{R}_+, \quad u(ax) = u(x),\ |x| < a^{-1}, \tag{6-10}$$

$$\boldsymbol{F}(u) := i\int_{[-1,-a]\sqcup[a,1]} \bar{u}\, du \ge 0 \tag{6-11}$$

then $u \in C^\infty((-1, 1))$.

(2) We next extend u to a distribution on the entire $\mathbb{R}$. Fix

$$\psi \in C_c^\infty((-a^{-1}, a^{-1}) \setminus [-a, a]), \quad \sum_{k\in\mathbb{Z}}\psi(a^{-k}x) = 1, \quad x \ne 0.$$

Then $\psi u \in C_c^\infty(\mathbb{R} \setminus \{0\})$. Define

$$v(x) := \sum_{k\in\mathbb{Z}}(\psi u)(a^{-k}x) \in C^\infty(\mathbb{R} \setminus \{0\}). \tag{6-12}$$

Since $u(ax) = u(x)$ for $|x| < a^{-1}$, we have $u = v$ on $(-a^{-1}, a^{-1}) \setminus \{0\}$. Thus we may extend v to an element of $\mathcal{D}'(\mathbb{R})$ so that $u = v|_{(-a^{-1},a^{-1})}$. We note that

$$\begin{gathered}v \in \mathscr{S}'(\mathbb{R}), \quad v(ax) = v(x),\ x \in \mathbb{R},\\ \mathrm{WF}(v) \subset \{0\} \times \mathbb{R}_+, \quad \boldsymbol{F}(v) = \boldsymbol{F}(u) \ge 0.\end{gathered} \tag{6-13}$$

It remains to show that $v \in C^\infty$; in fact, we will show that v is constant.

(3) Fix $\chi \in C_c^\infty(\mathbb{R})$ such that $\chi = 1$ near 0 and write

$$v = v_1 + v_2, \quad v_1 := \chi v, \quad v_2 := (1-\chi)v.$$

From (6-12) we obtain uniformly in $x \neq 0$,

$$\partial_x^\ell v(x) = x^{-\ell} \sum_{k\in\mathbb{Z}} ((\bullet)^\ell (\psi u)^{(\ell)})(a^{-k}x) = \mathcal{O}(x^{-\ell}),$$

since $(\bullet)^\ell(\psi u)^{(\ell)} : x \mapsto x^\ell (\psi u)^{(\ell)}(x) \in C_c^\infty(\mathbb{R}\setminus\{0\})$ and the sum is locally finite with a uniformly bounded number of terms. Hence $\partial_x^\ell v_2(x) = \mathcal{O}(\langle x\rangle^{-\ell})$, which implies that $\hat{v}_2(\xi)$ (and thus $\hat{v}(\xi)$) is smooth when $\xi \in \mathbb{R}\setminus\{0\}$ and

$$\hat{v}_2(\xi) = \mathcal{O}(\langle\xi\rangle^{-\infty}), \quad |\xi| \to \infty.$$

On the other hand the assumption on $\mathrm{WF}(v)$ and [Hörmander 1990, Proposition 8.1.3] shows that $\hat{v}_1(\xi) = \mathcal{O}(\langle\xi\rangle^{-\infty})$, as $\xi \to -\infty$. From (6-13) we obtain for $\xi < 0$ and $k \in \mathbb{N}$,

$$\hat{v}(\xi) = a^{-1}\hat{v}(a^{-1}\xi) = a^{-k}\hat{v}(a^{-k}\xi) = \mathcal{O}_\xi(a^k) \quad \Longrightarrow \quad \hat{v}|_{\mathbb{R}_-} \equiv 0. \tag{6-14}$$

(4) It follows from (6-14) that the distributional pairing

$$V(z) := \hat{v}(e^{iz\bullet})/(2\pi), \quad \operatorname{Im} z > 0, \tag{6-15}$$

is well-defined and holomorphic in $\{\operatorname{Im} z > 0\}$ and $|V(z)| \leq C\langle z\rangle^N/(\operatorname{Im} z)^M$, $\operatorname{Im} z > 0$ (with more precise bounds possible). We also have $v(x) = V(x+i0)$ for $x \in \mathbb{R}\setminus\{0\}$, and $V(az) = V(z)$ when $\operatorname{Im} z > 0$, which follows from (6-15). We will now use V to calculate $\boldsymbol{F}(v)$. We have

$$\boldsymbol{F}(v) = i\int_{\gamma_0} \overline{V(z)}\partial_z V(z)\,dz, \quad \gamma_0 := [-1,-a]\cup[a,1],$$

where the curve γ_0 is positively oriented. Let γ_α, $\alpha > 0$, be the half circle $|z| = \alpha$, $\operatorname{Im} z > 0$ oriented counterclockwise. Since $V(az) = V(z)$,

$$\int_{\gamma_1} \overline{V(z)}\partial_z V(z)\,dz = \int_{\gamma_1} \overline{V(az)}(\partial_z V)(az)\,d(az) = \int_{\gamma_a} \overline{V(z)}\partial_z V(z)\,dz.$$

If Γ is the semiannulus bounded by $\partial\Gamma := \gamma_0+\gamma_1-\gamma_a$ (see Figure 12) it follows from the Cauchy–Pompeiu formula [Hörmander 1990, (3.1.9)] that (with $z = x+iy$)

$$\boldsymbol{F}(v) = i\oint_{\partial\Gamma} \overline{V(z)}\partial_z V(z)\,dz = -2\int_\Gamma \partial_{\bar z}(\overline{V(z)}\partial_z V(z))\,dx\,dy = -2\int_\Gamma |\partial_z V(z)|^2\,dx\,dy.$$

Since we assumed $\boldsymbol{F}(v) \geq 0$ it follows that V is constant on Γ and thus on the entire upper half-plane, which implies that v is constant on $\mathbb{R}\setminus\{0\}$. Since functions supported at 0 are linear combinations of derivatives of the delta function and cannot solve the equation $v(ax) = v(x)$, we see that v is constant on $\mathbb{R}$, which finishes the proof. □

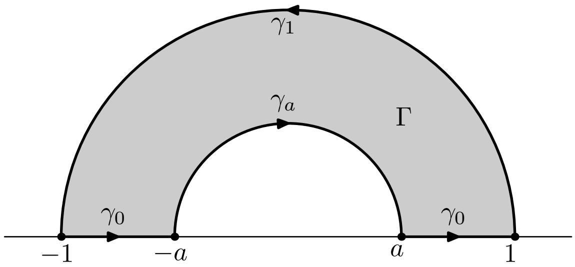

Figure 12. The domain Γ used in the proof of Lemma 6.1.

6.2. ***A global result.*** We now use the local result in Lemma 6.1 to obtain a global result for Morse–Smale diffeomorphisms of the circle.

Proposition 6.3. *Let $b : \partial\Omega \to \partial\Omega$ be a Morse–Smale diffeomorphism* (see Definition 1.2). *Denote by Σ^+, $\Sigma^- \subset \partial\Omega$ the sets of attractive, respectively repulsive, periodic points of b, and define $N^*_\pm\Sigma^\pm \subset T^*\partial\Omega$ by (3-18). Suppose that $u \in \mathcal{D}'(\partial\Omega)$ satisfies*

$$b^*u = u, \quad \mathrm{WF}(u) \subset N^*_+\Sigma^+ \sqcup N^*_-\Sigma^-. \tag{6-16}$$

Then u is constant.

Remark. The same conclusion holds when the wavefront set condition in (6-16) is replaced by $\mathrm{WF}(u) \subset N^*_+\Sigma^- \sqcup N^*_-\Sigma^+$, as can be seen by applying Proposition 6.3 to the complex conjugate $\bar{u}$.

Proof. We introduce fluxes associated to $g := b^n$, where n is the minimal period of periodic points of b. For that we take two arbitrary cutoff functions

$$\chi_\pm \in C^\infty(\partial\Omega), \quad \mathrm{supp}(1-\chi_\pm) \cap \Sigma^\pm = \varnothing, \quad \mathrm{supp}\,\chi_\pm \cap \Sigma^\mp = \varnothing.$$

Assume that u satisfies (6-16) and define the fluxes (where we again use positive orientation on $\partial\Omega$ to define the integrals of 1-forms):

$$\boldsymbol{F}_+(u) := i\int_{\partial\Omega} (g^*\chi_+ - \chi_+)\,\bar{u}\,du,$$
$$\boldsymbol{F}_-(u) := i\int_{\partial\Omega} ((g^{-1})^*\chi_- - \chi_-)\,\bar{u}\,du.$$

The integrals above are well-defined since $g^*\chi_+ - \chi_+$ and $(g^{-1})^*\chi_- - \chi_-$ are supported in $\partial\Omega\setminus(\Sigma^+\sqcup\Sigma^-)$, where u is smooth. Moreover as in the case of $\boldsymbol{F}(u)$ defined in (6-3), $\boldsymbol{F}_\pm(u)$ are real and do not depend on the choice of $\chi_\pm$. We also note that (by taking $\chi_\pm$ real-valued)

$$\boldsymbol{F}_\pm(\bar{u}) = -\overline{\boldsymbol{F}_\pm(u)} = -\boldsymbol{F}_\pm(u). \tag{6-17}$$

Since $\boldsymbol{F}_\pm(u)$ are independent of $\chi_\pm$, we may choose $\chi_+ := 1 - (g^{-1})^*\chi_-$ to get the identity

$$\boldsymbol{F}_+(u) = \boldsymbol{F}_-(u). \tag{6-18}$$

Let $\Sigma^+ = \{x_1^+, \ldots, x_m^+\}$. By taking $\chi_+ = \chi_+^1 + \cdots + \chi_+^m$, where each χ_+^j is supported near x_j^+, we can write $\boldsymbol{F}_+(u) = \boldsymbol{F}_+^1(u) + \cdots + \boldsymbol{F}_+^m(u)$. We may apply Lemma 6.1 with f defined by g in local coordinates near $x_j^+ \simeq 0$ to see that $\boldsymbol{F}_+^j(u) \le 0$ with equality only if u is constant near x_j^+. Adding these together, we see that $\boldsymbol{F}_+(u) \le 0$ with equality only if u is locally constant near Σ^+.

Arguing similarly near Σ^-, using $f := g^{-1}$ and replacing u by $\bar{u}$, with $\mathrm{WF}(\bar{u}) = \{(x, -\xi) \mid (x, \xi) \in \mathrm{WF}(u)\}$, we see that $\boldsymbol{F}_-(u) \ge 0$ with equality only if u is locally constant near Σ^-. By (6-18) we then see that u is locally constant near $\Sigma^+ \sqcup \Sigma^-$ and hence $u \in C^\infty(\partial\Omega)$. Since for $x \in \partial\Omega \setminus \Sigma^-$, $g^n(x) \to x_0$ for some x_0 in Σ^+, we conclude that $u \in C^\infty$ takes finitely many values and hence is constant. □

7. Limiting absorption principle

In this section we consider operators

$$\begin{aligned} P &:= \partial_{x_2}^2 \Delta_\Omega^{-1} : H^{-1}(\Omega) \to H^{-1}(\Omega), \\ P(\omega) &:= \partial_{x_2}^2 - \omega^2 \Delta = (P - \omega^2)\Delta_\Omega : H_0^1(\Omega) \to H^{-1}(\Omega). \end{aligned} \tag{7-1}$$

We prove the limiting absorption principle for P in the form presented in Theorem 1.4. To do this we follow Section 4.4 to reduce the equation $P(\omega)u_\omega = f$ to the boundary $\partial\Omega$. We next analyze the resulting "Neumann data" $v_\omega = \mathcal{N}_\omega u_\omega$ (see (4-26)) uniformly as $\varepsilon = \operatorname{Im}\omega \to 0+$, using the high-frequency estimates of Section 5 and the absence of embedded spectrum following from the results of Section 6. This is slightly nonstandard since the boundary has characteristic points and the problem changes from elliptic to hyperbolic as $\operatorname{Im}\omega \to 0+$.

7.1. ***Poincaré spectral problem.*** We recall (see for instance [Davies 1995, Chapter 6]) that $\Delta = \partial_{x_1}^2 + \partial_{x_2}^2$ with the domain $H^2(\Omega) \cap H_0^1(\Omega)$ ($H_0^1(\Omega)$ is the closure of $C_c^\infty(\Omega)$ with respect to the norm $\|\bullet\|_{H_0^1(\Omega)}$ below) is a negative definite unbounded self-adjoint operator on $L^2(\Omega)$. Its inverse is an isometry,

$$\Delta_\Omega^{-1} : H^{-1}(\Omega) \to H_0^1(\Omega),$$

with inner products on these Hilbert spaces given by

$$\langle u, w \rangle_{H_0^1(\Omega)} := \int_\Omega \nabla u \cdot \overline{\nabla w}\, dx, \quad \langle U, W \rangle_{H^{-1}(\Omega)} := \langle \Delta_\Omega^{-1} U, \Delta_\Omega^{-1} W \rangle_{H_0^1(\Omega)}.$$

Since $\partial_{x_2}^2 : H_0^1(\Omega) \to H^{-1}(\Omega)$ the operator P in (7-1) is indeed bounded on $H^{-1}(\Omega)$.

Let $\{e_\alpha\}_{\alpha \in A}$ be an $L^2(\Omega)$-orthonormal basis of eigenfunctions of $-\Delta_\Omega$:

$$-\Delta_\Omega e_\alpha = \mu_\alpha^2 e_\alpha, \quad e_\alpha|_{\partial\Omega} = 0, \quad \langle e_\alpha, e_\beta \rangle_{L^2(\Omega)} = \delta_{\alpha,\beta}.$$

Then $\{\mu_\alpha e_\alpha\}_{\alpha \in A}$ is an orthonormal basis of the Hilbert space $H^{-1}(\Omega)$. The matrix elements of P in this basis are given by

$$\begin{aligned} \langle P \mu_\alpha e_\alpha, \mu_\beta e_\beta \rangle_{H^{-1}} = \langle \Delta_\Omega^{-1} \partial_{x_2}^2 \mu_\alpha^{-1} e_\alpha, \mu_\beta^{-1} e_\beta \rangle_{H_0^1} &= -\mu_\alpha^{-1} \mu_\beta^{-1} \langle \partial_{x_2}^2 e_\alpha, e_\beta \rangle_{L^2} \\ &= \mu_\alpha^{-1} \mu_\beta^{-1} \langle \partial_{x_2} e_\alpha, \partial_{x_2} e_\beta \rangle_{L^2(\Omega)}, \end{aligned}$$

where the last integration by parts is justified as $e_\beta|_{\partial\Omega} = 0$. This shows that P is a bounded self-adjoint operator on $H^{-1}(\Omega)$. This representation is particularly useful in numerical calculations needed to produce Figure 1. Testing P against $\Delta^2(\psi(x)e^{i\langle n,x\rangle})$, $\psi \in C_c^\infty(\Omega)$, $n \in \mathbb{Z}^2$, shows that

$$\operatorname{Spec}(P) = [0, 1];$$

see [Ralston 1973, Theorem 2]. In particular, for $\omega^2 \in \mathbb{C} \setminus [0, 1]$,

$$\|P(\omega)^{-1}\|_{H^{-1}(\Omega)\to H_0^1(\Omega)} = \|(P-\omega^2)^{-1}\|_{H^{-1}(\Omega)\to H^{-1}(\Omega)} = \frac{1}{d(\omega^2, [0, 1])}. \tag{7-2}$$

Limiting absorption principle in its most basic form means showing we have limiting operators acting on smaller spaces with values in larger spaces: for $\lambda \in (0, 1)$ satisfying the Morse–Smale conditions

$$(P-\lambda^2-i0)^{-1} : C_c^\infty(\Omega) \to H^{-(3/2)-}(\Omega), \qquad P(\lambda+i0)^{-1} : C_c^\infty(\Omega) \to H^{(1/2)-}(\Omega). \tag{7-3}$$

7.2. *Regularity of limits as $\varepsilon \to 0+$.* In this section we use the results of Section 5 to get a conormal regularity statement for weak limits of boundary data. In Section 7.4 below we apply this to the Neumann data $v_\omega = \mathcal{N}_\omega u_\omega$, $P(\omega)u_\omega = f$.

Since the conormal spaces used below depend on $\lambda = \operatorname{Re}\omega$, we need to define what it means for a sequence of distributions to be bounded in these spaces uniformly in λ. Assume that $\mathcal{J} \subset (0, 1)$ is an open interval such that each $\lambda \in \mathcal{J}$ satisfies the Morse–Smale conditions of Definition 1.2. Let $\Sigma_\lambda^\pm$ be defined in (1-6) and $\Sigma_\lambda = \Sigma_\lambda^+ \sqcup \Sigma_\lambda^-$. Fix a defining function $\rho_\lambda \in C^\infty(\partial\Omega; \mathbb{R})$ of Σ_λ and a pseudodifferential operator $A_{\Sigma_\lambda} \in \Psi^0(\partial\Omega)$ such that $\operatorname{WF}(A_{\Sigma_\lambda}) \cap (N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+) = \varnothing$ and A_{Σ_λ} is elliptic on $N_-^*\Sigma_\lambda^- \sqcup N_+^*\Sigma_\lambda^+$. We choose both ρ_λ and A_{Σ_λ} depending smoothly on $\lambda \in \mathcal{J}$.

Given two sequences $\lambda_j \to \lambda \in \mathcal{J}$ and $v_j \in C^\infty(\partial\Omega)$, we say that

$$v_j \text{ is bounded in } I^{s+}(\partial\Omega, N_+^*\Sigma_{\lambda_j}^- \sqcup N_-^*\Sigma_{\lambda_j}^+) \text{ uniformly in } j$$

if each of the seminorms (3-20) is bounded uniformly in j. We can similarly talk about uniform boundedness of sequences of 1-forms $v_j \in C^\infty(\partial\Omega; T^*\partial\Omega)$, identifying these with scalar distributions using a coordinate θ.

Lemma 7.1. *Assume that $\omega_j \to \lambda \in \mathcal{J}$, $\operatorname{Im}\omega_j > 0$, and the sequence $v_j \in C^\infty(\partial\Omega; T^*\partial\Omega)$ has the following properties*:

$$v_j \to v_0 \textit{ in } H^{-N} \textit{ for some } N, \tag{7-4}$$

$$\mathcal{C}_{\omega_j} v_j \textit{ is bounded in } I^{-(3/4)+}(\partial\Omega, N_+^*\Sigma_{\lambda_j}^- \sqcup N_-^*\Sigma_{\lambda_j}^+) \textit{ uniformly in } j, \tag{7-5}$$

where $\lambda_j = \operatorname{Re}\omega_j$ and $\mathcal{C}_{\omega_j}$ was defined in Section 4.6. Then we have

$$v_j \to v_0 \textit{ in } H^{-(1/2)-\beta} \textit{ for all } \beta > 0, \tag{7-6}$$

$$v_0 \in I^{(1/4)+}(\partial\Omega, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+). \tag{7-7}$$

Remark. In fact we have $v_j \to v_0$ in $I^{(1/4)+}(\partial\Omega, N_+^*\Sigma_{\lambda_j}^- \sqcup N_-^*\Sigma_{\lambda_j}^+)$ where convergence is defined using the seminorms (3-20) — see the last paragraph of the proof below.

Proof. The function v_j satisfies (5-6):

$$v_j = B_{\omega_j}^+ b_{\lambda_j}^* v_j + B_{\omega_j}^- b_{\lambda_j}^{-*} v_j + g_j, \qquad g_j = (I - \mathcal{A}_{\omega_j})\mathcal{E}_{\omega_j} d\mathcal{C}_{\omega_j} v_j. \tag{7-8}$$

Here the operator $\mathcal{A}_{\omega_j} = (\gamma_{\lambda_j}^+)^* A_{\omega_j}^+ + (\gamma_{\lambda_j}^-)^* A_{\omega_j}^-$ is defined in (5-2).

Applying Proposition 5.4 to $v_\omega := v_j$ we see that for each $k \in \mathbb{N}_0$ and $\beta > 0$ there exist N_0, C independent of j such that

$$\|(\rho_{\lambda_j}\partial_\theta)^k v_j\|_{H^{-1/2-\beta}} + \|A_{\Sigma_{\lambda_j}} v_j\|_{H^k} \le C\Big(\max_{0\le \ell\le N_0} \|(\rho_{\lambda_j}\partial_\theta)^\ell g_j\|_{H^{-1/2-\beta}} + \|A_{\Sigma_{\lambda_j}} g_j\|_{H^{N_0}} + \|v_j\|_{H^{-N}}\Big). \quad (7\text{-}9)$$

The pseudodifferential operators $A^\pm_{\omega_j}, \mathcal{E}_{\omega_j}$ are bounded on $I^{(1/4)+}(\partial\Omega, N^*_+\Sigma^-_{\lambda_j} \sqcup N^*_-\Sigma^+_{\lambda_j})$ uniformly in j, as are the pullback operators $(\gamma^\pm_{\lambda_j})^*$ (see Remark (1) after Proposition 4.15 and the end of the proof of Lemma 4.17). Thus by (7-5) we see that g_j is bounded uniformly in j in the space $I^{(1/4)+}(\partial\Omega, N^*_+\Sigma^-_{\lambda_j} \sqcup N^*_-\Sigma^+_{\lambda_j})$. Moreover, by (7-4) $\|v_j\|_{H^{-N}}$ is bounded uniformly in j as well. It follows that the right-hand side of (7-9), and thus its left-hand side as well, is bounded uniformly in j for any choice of $k \in \mathbb{N}_0$, $\beta > 0$.

Take arbitrary $0 < \beta' < \beta$. Then $\|v_j\|_{H^{-1/2-\beta'}}$ is bounded in j. Using compactness of the embedding $H^{-1/2-\beta'} \hookrightarrow H^{-1/2-\beta}$ we see that each subsequence of $\{v_j\}$ has a subsequence converging in $H^{-1/2-\beta}$; the limit of this further subsequence has to be equal to v_0 by (7-4). This implies (7-6).

A similar argument using again the boundedness of the left-hand side of (7-9) shows that $(\rho_{\lambda_j}\partial_\theta)^k v_j \to (\rho_\lambda\partial_\theta)^k v_0$ in $H^{-1/2-\beta}$ for all $k \in \mathbb{N}_0$, $\beta > 0$ and $A_{\Sigma_{\lambda_j}} v_j \to A_{\Sigma_\lambda} v_0$ in C^∞. In particular, this implies that $(\rho_\lambda\partial_\theta)^k v_0 \in H^{-(1/2)-}$ and $A_{\Sigma_\lambda} v_0 \in C^\infty$, which by (3-20) gives (7-7). □

7.3. ***Uniqueness for the limiting problem.*** We next use the analysis of Section 6 to show a uniqueness result for the restricted single layer potential operator $\mathcal{C}_{\lambda+i0}$ (see Section 4.6) in the space of distributions satisfying additional conditions. This will give us the lack of embedded spectrum for the operator P in the Morse–Smale case. To formulate this result, we recall the operators $R_{\lambda+i0} : g \mapsto E_{\lambda+i0} * g$ defined in (4-25) and $\mathcal{I} : \mathcal{D}'(\partial\Omega; T^*\partial\Omega) \to \mathcal{E}'(\mathbb{R}^2)$ defined in (4-24).

Lemma 7.2. *Let $\lambda \in (0,1)$ satisfy the Morse–Smale conditions of Definition 1.2. Assume that $v \in \mathcal{D}'(\partial\Omega; T^*\partial\Omega)$ lies in $I^s(\partial\Omega, N^*_+\Sigma^-_\lambda \sqcup N^*_-\Sigma^+_\lambda)$ for some s (see (3-19)), where $\Sigma^\pm_\lambda$ are defined in (1-6). Then*

$$\mathcal{C}_{\lambda+i0} v = 0, \quad \operatorname{supp}(R_{\lambda+i0}\mathcal{I}v) \subset \overline{\Omega} \quad \Longrightarrow \quad v = 0. \quad (7\text{-}10)$$

Proof. (1) Put $U := R_{\lambda+i0}\mathcal{I}v \in \mathcal{D}'(\mathbb{R}^2)$. Since $P(\lambda)E_{\lambda+i0} = \delta_0$ by (4-20), we have

$$P(\lambda)U = \mathcal{I}v. \quad (7\text{-}11)$$

We first show that

$$\operatorname{supp} U \subset \partial\Omega. \quad (7\text{-}12)$$

By the second assumption in (7-10) we have $\operatorname{supp} U \subset \overline{\Omega}$; thus it suffices to show that $u = 0$, where $u := U|_\Omega = S_{\lambda+i0}v$ and $S_{\lambda+i0}$ is the limiting single layer potential defined in (4-32).

Since $\operatorname{supp}\mathcal{I}v \subset \partial\Omega$, from (7-11) we have $P(\lambda)u = 0$. As $\lambda \in (0,1)$, $P(\lambda)$ is a constant coefficient hyperbolic operator. In view of (4-3) and (4-4) we then have, letting $\ell^\pm(x) := \ell^\pm(x,\lambda)$, $\ell^\pm_{\min} := \ell^\pm(x^\pm_{\min})$, $\ell^\pm_{\max} := \ell^\pm(x^\pm_{\max})$

$$u(x) := u_+(\ell^+(x)) - u_-(\ell^-(x)), \quad x \in \Omega, \ u_\pm \in \mathcal{D}'((\ell^\pm_{\min}, \ell^\pm_{\max})). \quad (7\text{-}13)$$

From (4-39) we see that $u \in I^{s-5/4}(\bar{\Omega}, \Lambda^-(\lambda))$, in particular by (2-3) u is smooth up to the boundary near the characteristic set $\mathscr{C}_\lambda$. It follows that $u_\pm$ are smooth near the boundary points $\ell^\pm_{\min}$, $\ell^\pm_{\max}$ up to the boundary. Define the pullbacks of $u_\pm$ to $\partial\Omega$ by the maps $\ell^\pm$,

$$w_\pm = u_\pm(\ell^\pm(x))|_{\partial\Omega} \in \mathcal{D}'(\partial\Omega), \quad (\gamma^\pm)^* w_\pm = w_\pm.$$

From the proof of conormal regularity of u in Lemma 4.9 we see that $\mathrm{WF}(w_\pm) \subset N^*_+\Sigma^-_\lambda \sqcup N^*_-\Sigma^+_\lambda$.

The restriction $u|_{\partial\Omega}$ is equal to both $w_+ - w_-$ and $\mathcal{C}_{\lambda+i0}v$. Thus by the first assumption in (7-10) we have $w_+ = w_-$. Defining $w := w_+ = w_-$, we have

$$(\gamma^\pm)^* w = w, \quad \mathrm{WF}(w) \subset N^*_+\Sigma^-_\lambda \sqcup N^*_-\Sigma^+_\lambda.$$

This implies that $b^* w = w$ and we can apply Proposition 6.3 to see that w is constant. But then $u_\pm$ are constant and $u = 0$, giving (7-12).

(2) We now show that $v = 0$ away from the characteristic set $\mathscr{C}_\lambda$ of $P(\lambda)$ on $\partial\Omega$ (see (2-3)). For each $x_0 \in \partial\Omega \setminus \mathscr{C}_\lambda$ we can find a neighborhood $V \subset \mathbb{R}^2$ of x_0 and coordinates (y_1, y_2) on V such that for some open interval $\mathscr{I} \subset \mathbb{R}$

$$\partial\Omega \cap V = \{y_1 = 0,\ y_2 \in \mathscr{I}\}, \quad P(\lambda)|_V = \sum_{|\alpha|\le 2} a_\alpha(y)\partial_y^\alpha, \quad a_{2,0} \neq 0.$$

(The noncharacteristic property means that the conormal bundle of $\{y_1 = 0\}$ is disjoint from the set of zeros of $\sum_{|\alpha|=2} a_\alpha \eta^\alpha$.) Now, by [Hörmander 1990, Theorem 2.3.5] we see that (7-12) implies $U|_V = \sum_{k\le K} u_k(y_2)\delta^{(k)}(y_1)$, $u_k \in \mathcal{D}'(\mathscr{I})$. Hence, for some $\tilde{u}_k \in \mathcal{D}'(\mathscr{I})$,

$$P(\lambda)U|_V = a_{2,0}(y)u_K(y_2)\delta^{(K+2)}(y_1) + \sum_{k\le K+1} \tilde{u}_k(y_2)\delta^{(k)}(y_1).$$

By (7-11) we have $P(\lambda)U|_V = \mathcal{I}v|_V = a(y_2)v(y_2)\delta(y_1)$, $a \neq 0$. Thus $u_K = 0$. (Here we use y_2 as a local coordinate on $\partial\Omega$ to identify $v|_{\partial\Omega\cap V}$ with a distribution on $\mathscr{I}$.) Iterating this argument shows that $U|_V = 0$, which means $v|_{V\cap\partial\Omega} = 0$.

(3) We have shown $\operatorname{supp} v$ is contained in the finite set $\mathscr{C}_\lambda$. On the other hand, $v \in I^s(\partial\Omega, N^*_+\Sigma^-_\lambda \sqcup N^*_-\Sigma^+_\lambda)$ is smooth away from Σ_λ. Since $\Sigma_\lambda \cap \mathscr{C}_\lambda = \varnothing$ by (2-3), we get $v = 0$. □

Remark. The proof would be simpler if we knew that the limiting single layer potential operators $S_{\lambda+i0}$ were injective acting on the conormal spaces (4-39) — this would imply the injectivity of $\mathcal{C}_{\lambda+i0}$ on $I^s(\partial\Omega, N^*_+\Sigma^-_\lambda \sqcup N^*_-\Sigma^+_\lambda)$ without the support condition in (7-10) as follows from step (1) of the proof above. However, that is not clear. Under the dynamical assumptions made here, the proof of Proposition 5.4 shows that $\ker S_{\lambda+i0} \subset \ker \mathcal{C}_{\lambda+i0}$ is finite-dimensional but injectivity seems to be a curious open problem.

7.4. ***Boundary data analysis.*** Fix $f \in C^\infty_c(\Omega)$ and let $\mathcal{J} \subset (0, 1)$ be an open interval such that each $\lambda \in \mathcal{J}$ satisfies the Morse–Smale conditions of Definition 1.2. Consider the solution to the boundary-value problem (4-21):

$$u_\omega \in C^\infty(\bar{\Omega}), \quad P(\omega)u_\omega = f, \quad u_\omega|_{\partial\Omega} = 0, \quad \omega \in \mathcal{J} + i(0, \infty).$$

In this section, we combine the results of Sections 7.2–7.3 to study the behavior as $\operatorname{Im}\omega \to 0+$ of the "Neumann data" defined using (4-26):

$$v_\omega := \mathcal{N}_\omega u_\omega \in C^\infty(\partial\Omega;\, T^*\partial\Omega).$$

We first show the following convergence statement:

Proposition 7.3. *Assume that $\omega_j \to \lambda \in \mathcal{J}$, $\operatorname{Im}\omega_j > 0$. Then for all $\beta > 0$*

$$v_{\omega_j} \to v_{\lambda+i0} \quad \text{in } H^{-1/2-\beta}(\partial\Omega;\, T^*\partial\Omega), \quad \text{as } j \to \infty, \tag{7-14}$$

where $v_{\lambda+i0} \in H^{-(1/2)-}(\partial\Omega;\, T^*\partial\Omega)$ is the unique distribution such that

$$\begin{gathered} v_{\lambda+i0} \in I^{(1/4)+}(\partial\Omega,\, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+), \\ \mathcal{C}_{\lambda+i0} v_{\lambda+i0} = (R_{\lambda+i0} f)|_{\partial\Omega}, \quad \operatorname{supp} R_{\lambda+i0}(f - \mathcal{I} v_{\lambda+i0}) \subset \overline{\Omega}. \end{gathered} \tag{7-15}$$

Moreover, $v_{\lambda+i0} \in I^{1/4}(\partial\Omega,\, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+)$.

Proof. (1) We start with a few general observations. Recall (5-1):

$$\mathcal{C}_\omega v_\omega = (R_\omega f)|_{\partial\Omega}, \quad R_\omega f = E_\omega * f \in C^\infty(\mathbb{R}^2). \tag{7-16}$$

Moreover, by (4-28) we have

$$\mathbb{1}_\Omega u_\omega = R_\omega(f - \mathcal{I} v_\omega). \tag{7-17}$$

By Lemma 4.3, $E_{\omega_j} \to E_{\lambda+i0}$ in $\mathcal{D}'(\mathbb{R}^2)$. Passing to the limit in (7-16) we see that

$$\mathcal{C}_{\omega_j} v_{\omega_j} \to (R_{\lambda+i0} f)|_{\partial\Omega} \quad \text{in } C^\infty(\partial\Omega). \tag{7-18}$$

(2) We now show a boundedness statement: for each $\beta > 0$ there exists a constant C (depending on f and β) such that for all j

$$\|v_{\omega_j}\|_{H^{-1/2-\beta}(\partial\Omega;T^*\partial\Omega)} \le C. \tag{7-19}$$

We proceed by contradiction. If (7-19) fails then we may pass to a subsequence to make $\|v_{\omega_j}\|_{H^{-1/2-\beta}} \to \infty$. We then put

$$v_j := v_{\omega_j}/\|v_{\omega_j}\|_{H^{-1/2-\beta}}, \quad u_j := u_{\omega_j}/\|v_{\omega_j}\|_{H^{-1/2-\beta}}.$$

By (7-18) we have

$$\mathcal{C}_{\omega_j} v_j \to 0 \quad \text{in } C^\infty(\partial\Omega). \tag{7-20}$$

By compactness of the embedding $H^{-1/2-\beta} \hookrightarrow H^{-N}$, where we fix $N > \frac{1}{2} + \beta$, we may pass to a subsequence to make

$$v_j \to v_0 \ \text{ in } \ H^{-N} \text{ for some } v_0 \in H^{-N}(\partial\Omega;\, T^*\partial\Omega).$$

Now Lemma 7.1 applies and gives

$$v_j \to v_0 \ \text{ in } \ H^{-1/2-\beta}(\partial\Omega;\, T^*\partial\Omega), \ v_0 \in I^{(1/4)+}(\partial\Omega,\, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+). \tag{7-21}$$

By Lemma 4.16 and passing to the limit in (7-17) using Lemma 4.3 we get

$$\mathcal{C}_{\omega_j} v_j \to \mathcal{C}_{\lambda+i0} v_0 \quad \text{in } \mathcal{D}'(\partial\Omega), \qquad \mathbb{1}_\Omega u_j \to -R_{\lambda+i0}\mathcal{I} v_0 \quad \text{in } \mathcal{D}'(\mathbb{R}^2).$$

Thus by (7-20) and since $\operatorname{supp}(\mathbb{1}_\Omega u_j) \subset \bar{\Omega}$ for all j we have

$$\mathcal{C}_{\lambda+i0} v_0 = 0, \quad \operatorname{supp}(R_{\lambda+i0}\mathcal{I} v_0) \subset \bar{\Omega}.$$

Now Lemma 7.2 gives $v_0 = 0$. On the other hand the first part of (7-21) and the fact that $\|v_j\|_{H^{-1/2-\beta}} = 1$ imply that $\|v_0\|_{H^{-1/2-\beta}} = 1$, which gives a contradiction.

(3) Fix $\beta > 0$ and take an arbitrary subsequence $v_{\omega_{j_\ell}}$ which converges to some v in $H^{-1/2-\beta}$. By Lemma 7.1 and (7-18) we have $v \in I^{(1/4)+}(\partial\Omega, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+)$. By Lemma 4.16 and (7-18) we have $\mathcal{C}_{\lambda+i0} v = (R_{\lambda+i0} f)|_{\partial\Omega}$. Finally, passing to the limit in (7-17) using Lemma 4.3 we have $\operatorname{supp} R_{\lambda+i0}(f - \mathcal{I} v) \subset \bar{\Omega}$. Thus v satisfies (7-15). By Lemma 7.2 there is at most one distribution which satisfies (7-15). This implies that all the limits of convergent subsequences of v_{ω_j} in $H^{-1/2-\beta}$ have to be the same.

On the other hand by (7-19) and compactness of the embedding $H^{-1/2-\beta'} \hookrightarrow H^{-1/2-\beta}$ when $0 < \beta' < \beta$ we see that the sequence v_{ω_j} is precompact in $H^{-1/2-\beta}$. Together with uniqueness of limit of subsequences this implies the convergence statement (7-14).

We finally show that $v_{\lambda+i0} \in I^{1/4}(\partial\Omega, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+)$. From (7-15) we get similarly to (5-6)

$$v_{\lambda+i0} = B_{\lambda+i0}^+ b_\lambda^* v_{\lambda+i0} + B_{\lambda+i0}^- b_\lambda^{-*} v_{\lambda+i0} + g_{\lambda+i0}, \quad \text{where} \quad g_{\lambda+i0} \in C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1).$$

It remains to apply Lemma 5.5. □

We now upgrade Proposition 7.3 to a convergence statement for all the derivatives $\partial_\omega^k v_\omega$. Here $v_\omega \in C^\infty(\partial\Omega; T^*\partial\Omega)$ is holomorphic in $\omega \in \mathcal{J} + i(0,\infty)$: indeed, $u_\omega \in C^\infty(\bar{\Omega})$ is holomorphic by the Remark following Lemma 4.4 and the operator $\mathcal{N}_\omega$ defined in (4-26) is holomorphic as well.

As in the proof of Proposition 7.3 we will use the spaces $I^s(\partial\Omega, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+)$ which depend on $\lambda = \operatorname{Re}\omega$. We recall from Section 4.6.9 the family of diffeomorphisms

$$\Theta_\lambda : \mathbb{S}^1 \to \partial\Omega, \quad \Theta_\lambda(\widetilde{\Sigma}^\pm) = \Sigma_\lambda^\pm, \quad \lambda \in \mathcal{J},$$

with the pullback operator Θ_λ^* mapping $I^s(\partial\Omega, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+)$ to the λ-independent space $I^s(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+)$. Define

$$\tilde{v}_\omega := \Theta_\lambda^* v_\omega \in C^\infty(\mathbb{S}^1; T^*\mathbb{S}^1), \quad \omega \in \mathcal{J} + i(0,\infty), \quad \lambda = \operatorname{Re}\omega.$$

If $v_{\lambda+i0}$ is defined in (7-15), then we also put

$$\tilde{v}_{\lambda+i0} := \Theta_\lambda^* v_{\lambda+i0} \in I^{(1/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+), \quad \lambda \in \mathcal{J}.$$

Writing $\omega = \lambda + i\varepsilon$, denote by $\partial_\lambda^\ell \tilde{v}_\omega$ the ℓ-th derivative of $\tilde{v}_\omega$ in λ with ε fixed. (Note that unlike v_ω, the function $\tilde{v}_\omega$ is not holomorphic in ω.)

We are now ready to give the main technical result of this section. The proof is similar to that of Proposition 7.3 (which already contains the key ideas), using additionally Lemma 4.17 which establishes regularity in λ of the operators $\mathcal{C}_\omega$ conjugated by Θ_λ.

Proposition 7.4. *We have*

$$\tilde{v}_{\lambda+i0} \in C^\infty(\mathcal{J}; I^{(1/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+)), \tag{7-22}$$

where the topology on $I^{(1/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+)$ *is defined using the seminorms* (3-20). *Moreover, for each* $\lambda \in \mathcal{J}$ *and* ℓ *we have as* $\varepsilon \to 0+$

$$\partial_\lambda^\ell \tilde{v}_{\lambda+i\varepsilon} \to \partial_\lambda^\ell \tilde{v}_{\lambda+i0} \quad \text{in } I^{(1/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+), \tag{7-23}$$

with convergence locally uniform in λ.

Remarks. (1) From (7-22) we get a regularity statement for $v_{\lambda+i0} = \Theta_\lambda^{-*}\tilde{v}_{\lambda+i0}$:

$$v_{\lambda+i0} \in C^\ell(\mathcal{J}; H^{-1/2-\ell-}(\partial\Omega; T^*\partial\Omega)) \quad \text{for all } \ell.$$

Here the loss of ℓ derivatives comes from differentiating Θ_λ^{-*} in λ.

(2) The property (7-22) can be reformulated as follows: the distribution $(\lambda, x) \mapsto v_{\lambda+i0}(x)$ lies in $I^{0+}(\mathcal{J} \times \partial\Omega, N_+^*\Sigma_{\mathcal{J}}^- \sqcup N_-^*\Sigma_{\mathcal{J}}^+)$, where $\Sigma_{\mathcal{J}}^\pm := \{(\lambda, x) \mid \lambda \in \mathcal{J},\ x \in \Sigma_\lambda^\pm\}$ and we orient the conormal bundles $N^*\Sigma_{\mathcal{J}}^\pm$ using the positive orientation on $\partial\Omega$.

Proof. (1) We start with a few identities on $\tilde{v}_\omega$, $\omega \in \mathcal{J} + i(0,\infty)$. Let $\widetilde{\mathcal{C}}_\omega = \Theta_\lambda^*\mathcal{C}_\omega\Theta_\lambda^{-*}$, $\lambda = \operatorname{Re}\omega$, be the conjugated restricted single layer potential defined in (4-83). Applying Θ_λ^* to (7-16) we get

$$\widetilde{\mathcal{C}}_\omega \tilde{v}_\omega = \widetilde{G}_\omega, \quad \text{where } \widetilde{G}_\omega := \Theta_\lambda^*((R_\omega f)|_{\partial\Omega}). \tag{7-24}$$

From (7-17) we have

$$\mathbb{1}_\Omega u_\omega = R_\omega(f - \mathcal{I}\Theta_\lambda^{-*}\tilde{v}_\omega). \tag{7-25}$$

Differentiating these identities ℓ times in $\lambda = \operatorname{Re}\omega$, we get

$$\widetilde{\mathcal{C}}_\omega \partial_\lambda^\ell \tilde{v}_\omega = \partial_\lambda^\ell \widetilde{G}_\omega - \sum_{r=0}^{\ell-1} \binom{\ell}{r} (\partial_\lambda^{\ell-r}\widetilde{\mathcal{C}}_\omega)(\partial_\lambda^r \tilde{v}_\omega), \tag{7-26}$$

$$R_\omega \mathcal{I}\Theta_\lambda^{-*}\partial_\lambda^\ell \tilde{v}_\omega = \partial_\omega^\ell R_\omega f - \mathbb{1}_\Omega\, \partial_\omega^\ell u_\omega - \sum_{r=0}^{\ell-1} \binom{\ell}{r} (\partial_\lambda^{\ell-r}(R_\omega \mathcal{I}\Theta_\lambda^{-*}))(\partial_\lambda^r \tilde{v}_\omega). \tag{7-27}$$

(2) Take an arbitrary sequence $\omega_j = \lambda_j + i\varepsilon_j \to \lambda \in \mathcal{J}$, $\operatorname{Im}\omega_j > 0$. We show that for each $\ell \in \mathbb{N}_0$

$$\partial_\lambda^\ell \tilde{v}_{\omega_j} \text{ is bounded uniformly in } j, \text{ in } I^{(1/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+). \tag{7-28}$$

We use induction on ℓ, showing (7-28) under the assumption

$$\partial_\lambda^r \tilde{v}_{\omega_j} \text{ is bounded uniformly in } j, \text{ in } I^{(1/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+), \text{ for all } r < \ell. \tag{7-29}$$

Fix arbitrary $\beta > 0$; the main task will be to show that

$$\|\partial_\lambda^\ell \tilde{v}_{\omega_j}\|_{H^{-1/2-\beta}} \quad \text{is bounded in } j. \tag{7-30}$$

We argue by contradiction: if (7-30) does not hold then we can pass to a subsequence to make $\|\partial_\lambda^\ell \tilde{v}_{\omega_j}\|_{H^{-1/2-\beta}} \to \infty$. Define

$$\tilde{v}_j := \partial_\lambda^\ell \tilde{v}_{\omega_j} / \|\partial_\lambda^\ell \tilde{v}_{\omega_j}\|_{H^{-1/2-\beta}}, \quad \|\tilde{v}_j\|_{H^{-1/2-\beta}} = 1. \tag{7-31}$$

Since $H^{-1/2-\beta}$ embeds compactly into H^{-N}, where we fix $N > \frac{1}{2} + \beta$, we may pass to a subsequence to get

$$\tilde{v}_j \to \tilde{v}_0 \quad \text{in } H^{-N} \text{ for some } \tilde{v}_0 \in H^{-N}(\mathbb{S}^1; T^*\mathbb{S}^1). \tag{7-32}$$

We now analyze the right-hand side of (7-26) for $\omega = \omega_j$. Since $R_\omega f = E_\omega * f$, $\partial_\omega^r E_{\omega_j} \to \partial_\lambda^r E_{\lambda+i0}$ in $\mathcal{D}'(\mathbb{R}^2)$ by Lemma 4.3, and $f \in C_c^\infty(\Omega)$ is independent of j, we see that

$$\partial_\lambda^\ell \widetilde{G}_{\omega_j} \text{ is bounded uniformly in } j, \text{ in } C^\infty(\mathbb{S}^1). \tag{7-33}$$

By Lemma 4.17 and (7-29) we next have for all $r < \ell$

$$(\partial_\lambda^{\ell-r} \widetilde{\mathcal{C}}_{\omega_j})(\partial_\lambda^r \tilde{v}_{\omega_j}) \quad \text{is bounded uniformly in } j, \text{ in } I^{-(3/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+). \tag{7-34}$$

Dividing (7-26) by $\|\partial_\lambda^\ell \tilde{v}_{\omega_j}\|_{H^{-1/2-\beta}}$, we then get

$$\widetilde{\mathcal{C}}_{\omega_j} \tilde{v}_j \to 0 \quad \text{in } I^{-(3/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+). \tag{7-35}$$

We now apply Lemma 7.1 to

$$v_j := \Theta_{\lambda_j}^{-*} \tilde{v}_j, \quad v_0 := \Theta_\lambda^{-*} \tilde{v}_0, \quad \mathcal{C}_{\omega_j} v_j = \Theta_{\lambda_j}^{-*} \widetilde{\mathcal{C}}_{\omega_j} \tilde{v}_j$$

and get

$$v_j \to v_0 \quad \text{in } H^{-1/2-\beta}, \quad v_0 \in I^{(1/4)+}(\partial\Omega, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+). \tag{7-36}$$

By Lemma 4.16 and (7-36) we have $\mathcal{C}_{\omega_j} v_j \to \mathcal{C}_{\lambda+i0} v_0$ in $\mathcal{D}'(\partial\Omega)$; thus by (7-35)

$$\mathcal{C}_{\lambda+i0} v_0 = 0. \tag{7-37}$$

We now obtain a support condition on $\mathcal{R}_{\lambda+i0}\mathcal{I} v_0$ by analyzing the right-hand side of (7-27) for $\omega = \omega_j$. Similarly to the proof of (7-33) we have

$$\partial_\omega^\ell R_{\omega_j} f \quad \text{is bounded uniformly in } j, \text{ in } C^\infty(\mathbb{R}^2).$$

By a similar argument using additionally (7-29) we get for all $r < \ell$

$$(\partial_\lambda^{\ell-r}(R_{\omega_j}\mathcal{I}\Theta_{\lambda_j}^{-*}))(\partial_\lambda^r \tilde{v}_{\omega_j}) \quad \text{is bounded uniformly in } j, \text{ in } \mathcal{D}'(\mathbb{R}^2),$$

where we define $\partial_\lambda^k(R_{\omega_j}\mathcal{I}\Theta_{\lambda_j}^{-*}) := \partial_\lambda^k(R_\omega \mathcal{I}\Theta_\lambda^{-*})|_{\omega=\omega_j}$.

By Lemma 4.3 and (7-36) we get

$$R_{\omega_j}\mathcal{I}\Theta_{\lambda_j}^{-*}\tilde{v}_j \to R_{\lambda+i0}\mathcal{I} v_0 \quad \text{in } \mathcal{D}'(\mathbb{R}^2).$$

Now, dividing (7-27) by $\|\partial_\lambda^\ell \tilde{v}_{\omega_j}\|_{H^{-1/2-\beta}}$ and using that $\operatorname{supp}(\mathbb{1}_\Omega\, \partial_\omega^\ell u_{\omega_j}) \subset \overline{\Omega}$ for all j we obtain

$$\operatorname{supp}(R_{\lambda+i0}\mathcal{I} v_0) \subset \overline{\Omega}. \tag{7-38}$$

Applying Lemma 7.2 and using (7-36)–(7-38) we now see that $v_0 = 0$. This gives a contradiction with (7-36), since $\|v_j\|_{H^{-1/2-\beta}}$ is bounded away from 0 by (7-31). This finishes the proof of (7-30).

The bound (7-30) implies the stronger boundedness statement (7-28). Indeed, the proof of Lemma 7.1 (more precisely, (7-9) and the paragraph following it) shows that any seminorm of $\Theta_{\lambda_j}^{-*}\partial_\lambda^\ell \tilde{v}_{\omega_j}$ in $I^{(1/4)+}(\partial\Omega, N_+^*\Sigma_{\lambda_j}^- \sqcup N_-^*\Sigma_{\lambda_j}^+)$ (see (3-20)) is bounded in terms of $\|\partial_\lambda^\ell \tilde{v}_{\omega_j}\|_{H^{-1/2-\beta}}$ (for any choice of β) and of some $I^{-(3/4)+}(\partial\Omega, N_+^*\Sigma_{\lambda_j}^- \sqcup N_-^*\Sigma_{\lambda_j}^+)$-seminorm of $\Theta_{\lambda_j}^{-*}\widetilde{\mathcal{C}}_{\omega_j}\partial_\lambda^\ell \tilde{v}_{\omega_j}$. The former is bounded in j by (7-30) and the latter is bounded in j by (7-26), (7-33), and (7-34).

(3) From (7-28) we see that (as before, using the seminorms (3-20)), the family of distributions $\tilde{v}_{\lambda+i\varepsilon}$ is bounded uniformly in $\varepsilon \in (0,1]$ in the space $C^\infty(\mathcal{J}; I^{(1/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+))$. By the Arzelà–Ascoli theorem [Munkres 2000, Theorem 47.1] and since any sequence which is bounded in $I^{(1/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+)$ is also precompact in this space (following from (3-20) and the compactness of embedding $H^s \subset H^t$ for $s > t$), it follows that $\tilde{v}_{\lambda+i\varepsilon}$ is also precompact in the space $C^\infty(\mathcal{J}; I^{(1/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+))$. Moreover, $\tilde{v}_{\lambda+i\varepsilon} \to \tilde{v}_{\lambda+i0}$ in the space $C^0(\mathcal{J}; H^{-\frac{1}{2}-}(\mathbb{S}^1; T^*\mathbb{S}^1))$ by Proposition 7.3. Together these two statements imply that as $\varepsilon \to 0+$

$$\tilde{v}_{\lambda+i\varepsilon} \to \tilde{v}_{\lambda+i0} \quad \text{in } C^\infty(\mathcal{J}; I^{(1/4)+}(\mathbb{S}^1, N_+^*\widetilde{\Sigma}^- \sqcup N_-^*\widetilde{\Sigma}^+)),$$

giving (7-22) and (7-23). □

7.5. *Proof of Theorem 1.4.* Fix $f \in C_c^\infty(\Omega)$, let $\omega = \lambda + i\varepsilon$, where $\lambda \in \mathcal{J}$ and $0 < \varepsilon \ll 1$. Without loss of generality we assume that f is real-valued. It suffices to show existence of the limit of $(P - (\lambda + i\varepsilon)^2)^{-1} f$, since $(P - (\lambda - i\varepsilon)^2)^{-1} f$ is given by its complex conjugate.

Let $u_\omega \in C^\infty(\bar{\Omega})$ be the solution to the boundary-value problem (4-21). Recalling (7-1) we see that

$$(P - \omega^2)^{-1} f = \Delta u_\omega \in C^\infty(\bar{\Omega}).$$

Next, by (4-29) we have

$$u_\omega = (R_\omega f)|_\Omega - S_\omega v_\omega, \tag{7-39}$$

where the "Neumann data" $v_\omega := \mathcal{N}_\omega u_\omega \in C^\infty(\partial\Omega; T^*\partial\Omega)$ is defined using (4-26).

By Proposition 7.3 we have

$$v_{\lambda+i\varepsilon} \to v_{\lambda+i0} \quad \text{in } H^{-(1/2)-}(\partial\Omega; T^*\partial\Omega), \text{ as } \varepsilon \to 0+,$$

with convergence locally uniform in $\lambda \in \mathcal{J}$. Using Lemma 4.3 and recalling that $R_\omega f = E_\omega * f$ and $S_\omega v_\omega = (R_\omega \mathcal{I} v_\omega)|_\Omega$, we pass to the limit in (7-39) to get

$$u_{\lambda+i\varepsilon} \to u_{\lambda+i0} := (R_{\lambda+i0} f)|_\Omega - S_{\lambda+i0} v_{\lambda+i0} \quad \text{in } \mathcal{D}'(\Omega), \text{ as } \varepsilon \to 0+,$$

with convergence again locally uniform in $\lambda \in \mathcal{J}$. This gives the convergence statement (1-15) with

$$(P - \lambda^2 - i0)^{-1} f = \Delta u_{\lambda+i0}.$$

Next, since $R_{\lambda+i0} f \in C^\infty(\mathbb{R}^2)$ and $v_{\lambda+i0} \in I^{1/4}(\partial\Omega, N_+^*\Sigma_\lambda^- \sqcup N_-^*\Sigma_\lambda^+)$, we apply the mapping property (4-39) to get $u_{\lambda+i0} \in I^{-1}(\bar{\Omega}, \Lambda^-(\lambda))$, which implies

$$(P - \lambda^2 - i0)^{-1} f \in I^1(\bar{\Omega}, \Lambda^-(\lambda)).$$

Since $C_c^\infty(\Omega)$ is dense in $H^{-1}(\Omega)$ (see for instance [Ralston 1973, Lemma 5]), it is then standard (see for instance [Cycon et al. 1987, Proposition 4.1]) that the spectrum of P in $\mathcal{J}^2$ is purely absolutely continuous. □

8. Large time asymptotic behavior

We will now adapt the analysis of [Dyatlov and Zworski 2019b, §5, 6] and use (1-13) to describe asymptotic behavior of solutions to (1-1), giving the proof of Theorem 1.3. Assume that $\lambda \in (0,1)$ satisfies the Morse–Smale conditions of Definition 1.2 and fix an open interval $\mathcal{J} \subset (0,1)$ containing λ such that each $\omega \in \mathcal{J}$ satisfies the Morse–Smale conditions as well (this is possible by Lemma 2.6). We emphasize that in this section, in contrast with Sections 4–7, we denote by λ the fixed real frequency featured in the forcing term in (1-1) and by ω an arbitrary real number (often lying in $\mathcal{J}$).

8.1. ***Reduction to the resolvent.*** Fix $f \in C_c^\infty(\Omega;\mathbb{R})$ and let u be the solution to (1-1). We first split (1-13) into two parts. Fix a cutoff function

$$\varphi \in C_c^\infty(\mathcal{J};[0,1]), \qquad \varphi = 1 \text{ on } [\lambda-\delta, \lambda+\delta] \text{ for some } \delta > 0. \tag{8-1}$$

By (1-13) we can write

$$u(t) = \Delta_\Omega^{-1}\operatorname{Re}(e^{i\lambda t}(w_1(t) + r_1(t))), \tag{8-2}$$

where, with $\boldsymbol{W}_{t,\lambda}$ defined in (1-13),

$$w_1(t) = \varphi(\sqrt{P})\boldsymbol{W}_{t,\lambda}(P)f, \qquad r_1(t) = (I - \varphi(\sqrt{P}))\boldsymbol{W}_{t,\lambda}(P)f. \tag{8-3}$$

The contribution of r_1 to u is bounded in $H^1(\Omega)$ uniformly as $t \to \infty$ as follows from:

Lemma 8.1. *We have*

$$\|\operatorname{Re}(e^{i\lambda t} r_1(t))\|_{H^{-1}(\Omega)} \le \frac{2}{\lambda\delta}\|f\|_{H^{-1}(\Omega)} \quad \textit{for all } t \ge 0. \tag{8-4}$$

Proof. We calculate $\operatorname{Re}(e^{i\lambda t} r_1(t)) = \boldsymbol{R}_{t,\lambda}(P)f$, where

$$\boldsymbol{R}_{t,\lambda}(z) = \operatorname{Re}\big(e^{i\lambda t}\boldsymbol{W}_{t,\lambda}(z)(1-\varphi(\sqrt{z}))\big) = \frac{(\cos(t\sqrt{z}) - \cos(t\lambda))(1-\varphi(\sqrt{z}))}{\lambda^2 - z}.$$

Since $\varphi = 1$ near λ, we see that $\sup_{[0,1]}|\boldsymbol{R}_{t,\lambda}| \le 2/(\lambda\delta)$. Now (8-4) follows from the functional calculus for the self-adjoint operator P on $H^{-1}(\Omega)$. □

Define for $\omega \in \mathcal{J}$ the limits in $\mathcal{D}'(\Omega)$ (which exist by Theorem 1.4; see Section 7.5)

$$u^\pm(\omega) := \Delta_\Omega^{-1}(P - \omega^2 \pm i0)^{-1} f. \tag{8-5}$$

Here $u^+(\omega)$ is the complex conjugate of $u^-(\omega)$ since f is real-valued.

By Stone's formula for the operator P (see for instance [Dyatlov and Zworski 2019a, Theorem B.10]) and a change of variables in the spectral parameter we have

$$\Delta_\Omega^{-1} w_1(t) = \frac{1}{\pi i}\int_{\mathbb{R}} \varphi(\omega)\boldsymbol{W}_{t,\lambda}(\omega^2)(u^-(\omega)-u^+(\omega))\omega\,d\omega$$
$$= \frac{1}{2\pi}\int_0^t\int_{\mathbb{R}} \varphi(\omega)(e^{i(\omega-\lambda)s}-e^{-i(\omega+\lambda)s})(u^+(\omega)-u^-(\omega))\,d\omega\,ds. \tag{8-6}$$

8.2. *Global geometry.* The proof of Theorem 1.4 in Section 7.5 shows that $u^\pm(\omega)$ are smooth families of conormal distributions associated to ω-dependent lines in $\mathbb{R}^2$, more precisely

$$u^\pm(\omega) \in I^{-1}(\overline{\Omega}, \Lambda^\pm(\omega)), \tag{8-7}$$

where $\Lambda^\pm(\omega)$ are defined in (1-9). To understand the behavior of $\Delta_\Omega^{-1}w_1(t)$ as $t\to\infty$ we present an explicit version of (8-7), relying on Proposition 7.4. The most confusing thing here are the signs defined in (1-8). Figures 4, 13, and 14 can be used for guidance here.

Let $\omega\in\mathcal{J}$ and $\Sigma_\omega^\pm \subset \partial\Omega$ be the attractive/repulsive sets of the chess billiard $b(\bullet,\omega)$ defined in (1-6). Recall that $b=\gamma^+\circ\gamma^-$ and the involutions $\gamma^\pm(\bullet,\omega)$ map Σ_ω^+ to Σ_ω^-. Let n be the minimal period of the periodic points of b. To simplify notation, we assume that each of the sets $\Sigma_\omega^\pm$ consists of exactly n points, that is, it is a single periodic orbit of b (as opposed to a union of several periodic orbits), but the analysis works in the same way in the general case. We write (with the cyclic convention that $x_{n+1}^\pm(\omega)=x_1^\pm(\omega)$, $x_0^\pm(\omega)=x_n^\pm(\omega)$)

$$\Sigma_\omega^\pm = \{x_k^\pm(\omega)\}_{k=1}^n, \quad \gamma^-(x_k^+)=x_k^-, \quad \gamma^+(x_k^-)=x_{k+1}^+, \tag{8-8}$$

and $b^{\pm1}(x_k^\pm(\omega),\omega)=x_{k+1}^\pm$. By Lemma 2.6, we can make $x_k^\pm(\omega)$ depend smoothly on $\omega\in\mathcal{J}$.

In the notation of (1-8) and (1-9),

$$\begin{gathered}\Lambda^-(\omega)=\bigsqcup_{k=1}^n N_+^*\Gamma_\omega^-(x_k^-(\omega))\sqcup\bigsqcup_{k=1}^n N_-^*\Gamma_\omega^+(x_k^+(\omega)),\\ N_+^*\Gamma_\omega^-(x_k^-(\omega))=\{(x,\tau d\ell_\omega^-):\ell_\omega^-(x-x_k^-(\omega))=0,\ \nu_k^-\tau>0\},\\ N_-^*\Gamma_\omega^+(x_k^+(\omega))=\{(x,\tau d\ell_\omega^+):\ell_\omega^+(x-x_k^+(\omega))=0,\ \nu_k^+\tau<0\},\\ \nu_k^\pm := \nu^\pm(x_k^\pm(\omega),\omega) := \operatorname{sgn}\partial_\theta\ell_\omega^\pm(x_k^\pm(\omega)),\end{gathered} \tag{8-9}$$

where $\ell_\omega^\pm(x):=\ell^\pm(x,\omega)$. We note that $\nu_k^\pm$ are independent of $\omega\in\mathcal{J}$. To obtain $\Lambda^+(\omega)$ we switch the sign of τ in (8-9) — see Figure 4.

We need the following geometric result (see Figure 14):

Lemma 8.2. *With the notation above we have for all* $\omega\in\mathcal{J}$

$$x\in\overline{\Omega},\quad \ell_\omega^\pm(x-x_k^\pm(\omega))=0 \quad\Longrightarrow\quad \operatorname{sgn}[\partial_\omega(\ell_\omega^\pm(x-x_k^\pm(\omega)))]=\mp\nu_k^\pm. \tag{8-10}$$

Proof. (1) We note that the definition (1-3) of $\gamma_\omega^\pm$ and (8-8) give

$$\ell_\omega^\pm(x-x_k^\pm(\omega))=\ell_\omega^\pm(x-x_\ell^\mp(\omega)),\quad \ell=\begin{cases}k-1, & +,\\ k, & -.\end{cases} \tag{8-11}$$

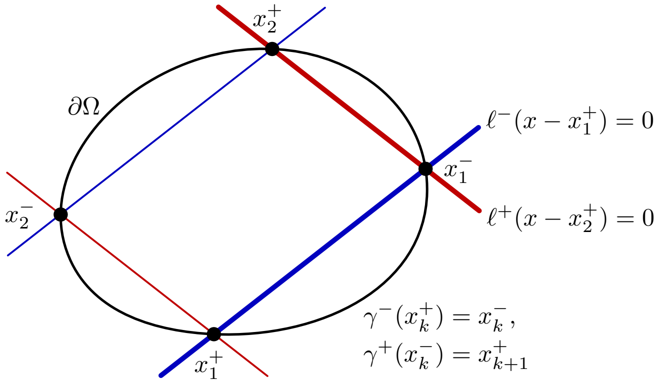

Figure 13. An illustration of (8-8) with $\Sigma_\lambda^\pm = \{x_1^\pm(\lambda), x_2^\pm(\lambda)\}$.

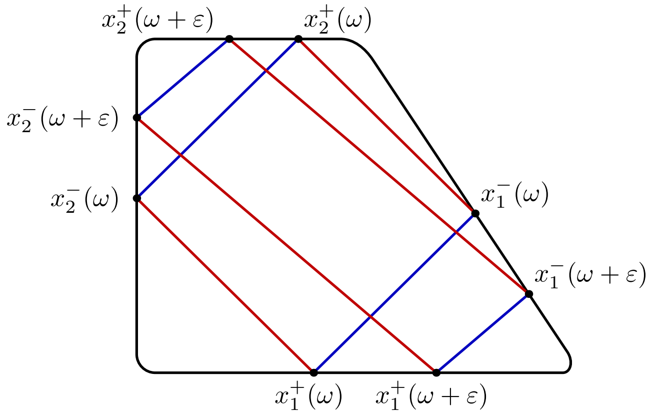

Figure 14. An illustration of Lemma 8.2, showing the periodic trajectory of the chess billiard for ω and for $\omega+\varepsilon$, $\varepsilon > 0$. In this example $\nu_1^+ = \nu_1^- = 1$, $\nu_2^+ = \nu_2^- = -1$. Lemma 8.2 shows in which direction the blue and red segments move as ω grows. For example, the entire red segment $\{x \in \Omega \mid \ell_\omega^+(x - x_1^+(\omega)) = 0\}$, which connects $x_1^+(\omega)$ to $x_2^-(\omega)$, lies inside the half-plane $\{x \mid \ell_{\omega+\varepsilon}^+(x - x_1^+(\omega+\varepsilon)) < 0\}$, which is consistent with (8-10).

We also note that (2-6) implies

$$\operatorname{sgn} \partial_\theta \ell_\omega^\pm(x_\ell^\mp(\omega)) =: \nu^\pm(x_\ell^\mp(\omega), \omega) = -\nu^\pm(x_k^\pm(\omega), \omega) = -\nu_k^\pm, \tag{8-12}$$

where $\theta \mapsto \boldsymbol{x}(\theta) \in \partial\Omega$ is a positive parametrization of $\partial\Omega$ by $\mathbb{S}_\theta^1$. It is sufficient to establish (8-10) with $x_k^\pm$ replaced by $x_\ell^\mp$, where ℓ is given in (8-11):

$$x \in \overline{\Omega}, \quad \ell_\omega^\pm(x - x_\ell^\mp(\omega)) = 0 \quad \Longrightarrow \quad \operatorname{sgn}[\partial_\omega(\ell_\omega^\pm(x - x_\ell^\mp(\omega)))] = \mp\nu_k^\pm. \tag{8-13}$$

(2) Using (2-7) and the condition on x in (8-13) we see that

$$\partial_\omega[\ell_\omega^\pm(x - x_\ell^\mp(\omega))] = \frac{\ell_\omega^\mp(x - x_\ell^\mp(\omega))}{2\omega(1-\omega^2)} - d_x\ell_\omega^\pm(\partial_\omega x_\ell^\mp(\omega)). \tag{8-14}$$

We start by considering the sign of the second term on the right-hand side:

$$-\operatorname{sgn} d_x \ell_\omega^\pm(\partial_\omega x_\ell^\mp(\omega)) = -\operatorname{sgn}[\partial_\theta \ell_\omega^\pm](x_\ell^\mp(\omega))\partial_\omega \theta(x_\ell^\mp(\omega)) = \nu_k^\pm \operatorname{sgn} \partial_\omega[\theta(x_\ell^\mp(\omega))], \tag{8-15}$$

where we used (8-12).

We now put $f := \theta \circ b^n \circ \theta^{-1}$, with n the primitive period. Then (2-15) and (1-6) give

$$f(\theta(x_\ell^\mp(\omega)), \omega) = \theta(x_\ell^\mp(\omega)), \quad \partial_\omega f(x,\omega) > 0, \quad \mp\big(1 - [\partial_\theta f](\theta(x_\ell^\mp(\omega)), \omega)\big) > 0.$$

Differentiating the first equality in ω gives

$$\partial_\omega[\theta(x_\ell^\mp(\omega))] = \partial_\omega f(\theta(x), \omega)|_{x = x_\ell^\mp(\omega)} / \big(1 - [\partial_\theta f](\theta(x_\ell^\mp(\omega)), \omega)\big),$$

and hence $\operatorname{sgn} \partial_\omega[\theta(x_\ell^\mp(\omega))] = \mp 1$. Returning to (8-15) we see that

$$-\operatorname{sgn} d_x \ell_\omega^\pm(\partial_\omega x_\ell^\mp(\omega)) = \mp \nu_k^\pm.$$

(3) We next claim that

$$x \in \overline{\Omega}, \quad \ell_\omega^\pm(x - x_\ell^\mp(\omega)) = 0 \quad \Longrightarrow \quad \operatorname{sgn} \ell_\omega^\mp(x - x_\ell^\mp(\omega)) \in \{\mp \nu_k^\mp, 0\}. \tag{8-16}$$

Combined with (8-14) and the conclusion of step (2), this will give (8-13) and hence (8-10). Since the set on the left-hand side of (8-16) is given by $x = (1-t)x_\ell^\mp(\omega) + t\gamma_\omega^\pm(x_\ell^\mp(\omega))$, $0 \le t \le 1$, it suffices to establish the conclusion in (8-16) for $x = \gamma_\omega^\pm(x_\ell^\mp(\omega))$. For that we use (2-5) and (8-12), which give

$$\operatorname{sgn} \ell_\omega^\mp(\gamma_\omega^\pm(x_\ell^\mp(\omega)) - x_\ell^\mp(\omega)) = \pm \nu^\pm(x_\ell^\mp(\omega), \omega) = \mp \nu_k^\pm,$$

completing the proof. □

In the notation of this section, Theorem 1.4 is reformulated as follows. Note that henceforth in this section, ε denotes a sign (either $+$ or $-$) in contrast with its use in the statement and proof of Theorem 1.4.

Lemma 8.3. *In the notation of (8-5), (8-9) and with $\varepsilon \in \{+,-\}$,*

$$\begin{gathered} u^\varepsilon(x,\omega) = \sum_{k=1}^n \sum_\pm g_{k,\pm}^\varepsilon(x,\omega) + u_0^\varepsilon(x,\omega), \quad u_0^\varepsilon \in C^\infty(\overline{\Omega} \times \mathcal{J}), \ g_{k,\pm}^\varepsilon \in \mathcal{D}'(\mathbb{R}^2), \\ g_{k,\pm}^\varepsilon(x,\omega) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{i\tau \ell_\omega^\pm(x - x_k^\pm(\omega))} a_{k,\pm}^\varepsilon(\tau,\omega)\, d\tau, \quad (x,\omega) \in \mathbb{R}^2 \times \mathcal{J}, \end{gathered} \tag{8-17}$$

where $a_{k,\pm}^\varepsilon \in S^{-1+}(\mathcal{J}_\omega \times \mathbb{R}_\tau)$ is supported in $\{\tau : \pm\varepsilon \nu_k^\pm \tau \ge 1\}$.

Proof. We consider the case of $\varepsilon = -$, with the case $\varepsilon = +$ following since $u^+(\omega) = \overline{u^-(\omega)}$. Recall from Section 7.5 that

$$u^-(\omega) = u_{\omega+i0} = (R_{\omega+i0} f)|_\Omega - S_{\omega+i0} v_{\omega+i0},$$

where $R_{\omega+i0} f \in C^\infty(\mathbb{R}^2 \times \mathcal{J})$ by Lemma 4.3 and

$$v_{\omega+i0} \in C^\infty(\mathcal{J}; I^{(1/4)+}(\partial\Omega, N_+^*\Sigma_\omega^- \sqcup N_-^*\Sigma_\omega^+)),$$

with smoothness in ω understood in the sense of Proposition 7.4. By the mapping property (4-39) we have

$$S_{\omega+i0}v_{\omega+i0} \in C^\infty(\mathcal{J}; I^{-1+}(\overline{\Omega}, \Lambda^-(\omega))).$$

Here smoothness in ω is obtained by following the proof of Lemma 4.9, which writes $S_{\omega+i0}v$ for $v \in C^\infty(\mathcal{J}; I^{(1/4)+}(\partial\Omega, N^*_\pm\{x^\mp_\ell(\omega)\}))$ as a sum of a function in $C^\infty(\overline{\Omega}\times\mathcal{J})$ and the pullback by $\ell^\pm_\omega$ of a conormal distribution to $\ell^\pm_\omega(x^\mp_\ell(\omega)) = \ell^\pm_\omega(x^\pm_k(\omega))$, with k, ℓ related by (8-11). This gives the representation (8-17). Here we can follow (1-9) and (8-9) to obtain an explicit parametrization of the conormal bundles $N^*_\mp\Gamma^\pm_\omega(x^\pm_k(\omega)) = N^*_\pm\Gamma^\pm_\omega(x^\mp_\ell(\omega))$ and check that $a^\varepsilon_{k,\pm}$ can be written as a sum of a symbol supported in $\{\tau : \pm\varepsilon\nu^\pm_k\tau \geq 1\}$ and a symbol which is rapidly decaying in τ, with the contribution of the latter lying in $C^\infty(\overline{\Omega}\times\mathcal{J})$. □

The next lemma disposes of the term u^ε_0:

Lemma 8.4. *Suppose that $u^\pm(x,\omega) \in C^\infty(\overline{\Omega}\times\mathcal{J})$. If w_1 is defined by (8-6) then, for any k, there exists C_k such that, for all $t \geq 0$, $\|\Delta_\Omega^{-1}w_1(t)\|_{C^k(\overline{\Omega})} \leq C_k$.*

Proof. Recalling (8-6), we see that it suffices to prove that for any $u \in C^\infty(\overline{\Omega}\times\mathcal{J})$

$$\sup_{t\geq 0}\|w(t)\|_{C^k(\overline{\Omega})} \leq C_k, \quad \text{where } w(t) := \int_0^t\int_{\mathbb{R}} \varphi(\omega)(e^{i(\omega-\lambda)s} - e^{-i(\omega+\lambda)s})u(\omega)\,d\omega\,ds.$$

Integrating by parts in ω, we get

$$\begin{aligned} w(x,t) &= \int_0^t\int_{\mathbb{R}} \varphi(\omega)u(x,\omega)[(1+s^2)^{-1}(1+D^2_\omega)](e^{i(\omega-\lambda)s} - e^{-i(\omega+\lambda)s})\,d\omega\,ds \\ &= \int_0^t\int_{\mathbb{R}} (1+D^2_\omega)[\varphi(\omega)u(x,\omega)](e^{i(\omega-\lambda)s} - e^{-i(\omega+\lambda)s})(1+s^2)^{-1}\,d\omega\,ds, \end{aligned}$$

which is bounded in $C^\infty(\overline{\Omega})$ uniformly in $t \geq 0$. □

Returning to (8-6) we see that we have to analyze the behavior of

$$w^{\varepsilon,\varepsilon'}_{k,\pm}(x,t) := \frac{1}{2\pi}\int_0^t\int_{\mathbb{R}} \varphi(\omega)g^\varepsilon_{k,\pm}(x,\omega)e^{is(\varepsilon'\omega-\lambda)}\,d\omega\,ds, \quad \varepsilon,\varepsilon' \in \{+,-\}, \tag{8-18}$$

as $t \to \infty$. More precisely, if the term u^ε_0 in the decomposition (8-17) were zero, then

$$\Delta_\Omega^{-1}w_1(x,t) = \sum_{k=1}^n \sum_\pm \sum_{\varepsilon,\varepsilon'\in\{+,-\}} \varepsilon\varepsilon' w^{\varepsilon,\varepsilon'}_{k,\pm}(x,t). \tag{8-19}$$

8.3. *Asymptotic behavior of $w^{\varepsilon,\varepsilon'}_{k,\pm}$.*

For $\tau \neq 0$, define

$$\begin{aligned} A^{\varepsilon,\varepsilon'}_{k,\pm}(x,t,\tau) &:= \frac{1}{2\pi}\int_0^t\int_{\mathbb{R}} e^{i\tau\ell^\pm_\omega(x-x^\pm_k(\omega))+i(\varepsilon'\omega-\lambda)s}\varphi(\omega)a^\varepsilon_{k,\pm}(\tau,\omega)\,d\omega\,ds \\ &= \frac{\tau}{2\pi}\int_0^{t/\tau}\int_{\mathbb{R}} e^{i\tau(\ell^\pm_\omega(x-x^\pm_k(\omega))+(\varepsilon'\omega-\lambda)r)}\varphi(\omega)a^\varepsilon_{k,\pm}(\tau,\omega)\,d\omega\,dr \end{aligned}$$

in the notation used for $g^{\varepsilon}_{k,\pm}$ in Lemma 8.3, where in the second line we made the change of variables $s=\tau r$. We then have

$$w^{\varepsilon,\varepsilon'}_{k,\pm}(x,t)=\frac{1}{2\pi}\int_{\mathbb{R}}A^{\varepsilon,\varepsilon'}_{k,\pm}(x,t,\tau)\,d\tau, \tag{8-20}$$

in the sense of oscillatory integrals (since $\partial_x\ell^{\pm}_{\omega}(x-x^{\pm}_k(\omega))=d\ell^{\pm}_{\omega}\neq 0$ the phase is nondegenerate — see [Hörmander 1990, §7.8]). From the support condition in Lemma 8.3 we get

$$A^{\varepsilon,\varepsilon'}_{k,\pm}(x,t,\tau)\neq 0 \quad\Longrightarrow\quad \pm\varepsilon\nu^{\pm}_k\tau\geq 1. \tag{8-21}$$

The lemma below shows that we only need to integrate over a compact interval in r:

Lemma 8.5. *There exist $\chi\in C^{\infty}_{\mathrm{c}}((0,\infty))$ and φ satisfying (8-1) such that for*

$$\tilde{A}^{\varepsilon,\varepsilon'}_{k,\pm}(x,t,\tau):=\frac{\tau}{2\pi}\int_0^{t/\tau}\int_{\mathbb{R}}e^{i\tau(\ell^{\pm}_{\omega}(x-x^{\pm}_k(\omega))+(\varepsilon'\omega-\lambda)r)}(1-\chi(\pm\varepsilon'\nu^{\pm}_k r))\varphi(\omega)a^{\varepsilon}_{k,\pm}(\tau,\omega)\,d\omega\,dr,$$

$$\tilde{w}^{\varepsilon,\varepsilon'}_{k,\pm}(x,t):=\frac{1}{2\pi}\int_{\mathbb{R}}\tilde{A}^{\varepsilon,\varepsilon'}_{k,\pm}(x,t,\tau)\,d\tau,$$

we have $\|\tilde{w}^{\varepsilon,\varepsilon'}_{k,\pm}(t)\|_{C^k(\overline{\Omega})}\leq C_k$ for every k and uniformly as $t\to\infty$.

Proof. (1) Put $F(x,\omega):=\ell^{\pm}_{\omega}(x-x^{\pm}_k(\omega))$. Lemma 8.2 shows that for all $\omega\in\mathcal{J}$

$$x\in\overline{\Omega},\ F(x,\omega)=0 \quad\Longrightarrow\quad \mp\nu^{\pm}_k\partial_{\omega}F(x,\omega)>0.$$

Fix a cutoff function $\psi\in C^{\infty}_{\mathrm{c}}(\overline{\Omega})$ such that $\psi=1$ in a neighborhood of $\{x\in\overline{\Omega}\mid F(x,\lambda)=0\}$ and $\mp\nu^{\pm}_k\partial_{\omega}F(x,\lambda)>0$ for all $x\in\operatorname{supp}\psi$. Choosing φ supported in a sufficiently small neighborhood of λ, we see that there exists $\chi\in C^{\infty}_{\mathrm{c}}((0,\infty))$ such that for all $\omega\in\operatorname{supp}\varphi$

$$x\in\overline{\Omega}\cap\operatorname{supp}(1-\psi) \quad\Longrightarrow\quad F(x,\omega)\neq 0, \tag{8-22}$$

$$x\in\overline{\Omega}\cap\operatorname{supp}\psi \quad\Longrightarrow\quad \mp\nu^{\pm}_k\partial_{\omega}F(x,\omega)\notin\operatorname{supp}(1-\chi). \tag{8-23}$$

(2) Using the singular support property of conormal distributions (see Section 3.2) and (8-22), we have

$$(1-\psi(x))\varphi(\omega)g^{\varepsilon}_{k,\pm}(x,\omega)\in C^{\infty}(\overline{\Omega}\times\mathbb{R}).$$

The proof of Lemma 8.4 shows that $\|(1-\psi(x))\tilde{w}^{\varepsilon,\varepsilon'}_{k,\pm}(t)\|_{C^k(\overline{\Omega})}\leq C_k$, uniformly as $t\to\infty$.

On the other hand, (8-23) implies that for some constant $c>0$

$$\omega\in\operatorname{supp}\varphi,\quad x\in\overline{\Omega}\cap\operatorname{supp}\psi,\quad r\in\operatorname{supp}(1-\chi(\pm\varepsilon'\nu^{\pm}_k\bullet)) \quad\Longrightarrow\quad |\partial_{\omega}F(x,\omega)+\varepsilon' r|\geq c\langle r\rangle.$$

Integration by parts in ω shows that

$$\partial^{\alpha}_x[\psi(x)\tilde{A}^{\varepsilon,\varepsilon'}_{k,\pm}(x,t,\tau)]=\mathcal{O}(\langle\tau\rangle^{-\infty}),$$

uniformly in t. But that gives uniform smoothness of $\psi(x)\tilde{w}^{\varepsilon,\varepsilon'}_{k,\pm}(x,t)$, finishing the proof. $\square$

The lemma shows that in the study of (8-18) we can replace A in (8-20) by

$$\begin{aligned} B^{\varepsilon,\varepsilon'}_{k,\pm}(x,t,\tau) &= A^{\varepsilon,\varepsilon'}_{k,\pm}(x,t,\tau) - \tilde{A}^{\varepsilon,\varepsilon'}_{k,\pm}(x,t,\tau) \\ &= \frac{\tau}{2\pi}\int_0^{t/\tau}\int_{\mathbb{R}} e^{i\tau(\ell^{\pm}_{\omega}(x-x^{\pm}_k(\omega))+(\varepsilon'\omega-\lambda)r)}\chi(\pm\varepsilon'\nu^{\pm}_k r)\varphi(\omega)a^{\varepsilon}_{k,\pm}(\tau,\omega)\,d\omega\,dr. \end{aligned}$$

Define the limit $B^{\varepsilon,\varepsilon'}_{k,\pm}(x,\infty,\tau)$ by replacing the integral $\int_0^{t/\tau} dr$ above by $\int_0^{(\operatorname{sgn}\tau)\infty} dr$, which is well-defined thanks to the cutoff $\chi(\pm\varepsilon'\nu^{\pm}_k r)$, where we recall that $\chi\in C^{\infty}_{\mathrm{c}}((0,\infty))$. The next lemma describes the behavior of this limit as $\tau\to\infty$:

Lemma 8.6. *Define $F(x,\omega):=\ell^{\pm}_{\omega}(x-x^{\pm}_k(\omega))$. Then $e^{-i\tau F(x,\lambda)}B^{\varepsilon,\varepsilon'}_{k,\pm}(x,\infty,\tau)$ lies in the symbol class $S^{-1+}(\bar{\Omega}_x\times\mathbb{R}_\tau)$ and*

$$e^{-i\tau F(x,\lambda)}B^{\varepsilon,\varepsilon'}_{k,\pm}(x,\infty,\tau)\in\begin{cases}\chi(\mp\nu^{\pm}_k\partial_\lambda F(x,\lambda))a^{\varepsilon}_{k,\pm}(\tau,\lambda)+S^{-2+}(\bar{\Omega}\times\mathbb{R}), & \varepsilon=\varepsilon'=+,\\ S^{-\infty}(\bar{\Omega}\times\mathbb{R}), & \textit{otherwise}.\end{cases}$$

Proof. We first note that if $\pm\varepsilon'\nu^{\pm}_k\tau<0$ then $B^{\varepsilon,\varepsilon'}_{k,\pm}(x,\infty,\tau)=0$. Hence we can assume that

$$\operatorname{sgn}\tau=\pm\varepsilon'\nu^{\pm}_k. \tag{8-24}$$

In that case we can replace limits of integration in r by $(-\infty,\infty)$, with τ replaced by $|\tau|$ in the prefactor $\tau/(2\pi)$. The method of stationary phase (see for instance [Hörmander 1990, Theorem 7.7.5]) can be applied to the double integral $\int_{\mathbb{R}^2} d\omega\,dr$ and the critical point in given by

$$\omega=\varepsilon'\lambda,\qquad r=-\varepsilon'\partial_\omega F(x,\omega).$$

Since $\omega=-\lambda$ lies outside of the support of φ, if $\varepsilon'=-$ then (by the method of nonstationary phase) we have $B^{\varepsilon,\varepsilon'}_{k,\pm}(x,\infty,\tau)\in S^{-\infty}(\bar{\Omega}\times\mathbb{R})$. We thus assume that $\varepsilon'=+$, which by (8-24) gives $\pm\nu^{\pm}_k\tau>0$. If $\varepsilon=-$ then the support property of $a^{\varepsilon}_{k,\pm}$ in Lemma 8.3 shows that $B^{\varepsilon,\varepsilon'}_{k,\pm}(x,\infty,\tau)=0$. Thus we may assume that $\varepsilon=\varepsilon'=+$. In the latter case the method of stationary phase gives the expansion for $B^{\varepsilon,\varepsilon'}_{k,\pm}(x,\infty,\tau)$. □

We now analyze the remaining term given by

$$v^{\varepsilon,\varepsilon'}_{k,\pm}(x,t):=\frac{1}{2\pi}\int_{\mathbb{R}} C^{\varepsilon,\varepsilon'}_{k,\pm}(x,t,\tau)\,d\tau, \tag{8-25}$$

where

$$C^{\varepsilon,\varepsilon'}_{k,\pm}(x,t,\tau)=\frac{\tau}{2\pi}\int_{t/\tau}^{(\operatorname{sgn}\tau)\infty}\int_{\mathbb{R}} e^{i\tau(\ell^{\pm}_{\omega}(x-x^{\pm}_k(\omega))+(\varepsilon'\omega-\lambda)r)}\chi^{\varepsilon'}_{k,\pm}(r)\varphi(\omega)a^{\varepsilon}_{k,\pm}(\tau,\omega)\,d\omega\,dr,$$

and $\chi^{\varepsilon'}_{k,\pm}(r):=\chi(\pm\varepsilon'\nu^{\pm}_k r)\in C^{\infty}_{\mathrm{c}}(\mathbb{R}\setminus\{0\})$. The last lemma deals with this term:

Lemma 8.7. *For $v^{\varepsilon,\varepsilon'}_{k,\pm}$ given by (8-25) we have, for every $\beta>0$,*

$$\|v^{\varepsilon,\varepsilon'}_{k,\pm}(t)\|_{H^{1/2-\beta}(\Omega)}\to 0\quad \textit{as } t\to\infty. \tag{8-26}$$

Proof. (1) We put $\varepsilon' = +$ as the other case is similar and simpler. To simplify notation we will often drop ε and k. Fix a cutoff function

$$\psi \in C_c^\infty(\mathbb{R}), \quad (\ell_\omega^\mp)^*\psi = 1 \text{ near } \overline{\Omega} \text{ for all } \omega \in \operatorname{supp}\varphi. \tag{8-27}$$

For $x \in \mathbb{R}^2$, $t > 0$, and $\omega \in \mathcal{J}$, define

$$U_\pm(x,t,\omega) = \psi(\ell_\omega^\mp(x)) V_\pm(\ell_\omega^\pm(x), t, \omega),$$

where (in the sense of oscillatory integrals)

$$\begin{aligned} V_\pm(y,t,\omega) &:= \int_{\mathbb{R}} \frac{\tau}{2\pi} \int_{t/\tau}^{(\operatorname{sgn}\tau)\infty} e^{i\tau(y-\ell_\omega^\pm(x_k^\pm(\omega)) + (\omega-\lambda)r)} \tilde{\chi}(r) b(\tau,\omega)\, dr\, d\tau, \\ b &:= \frac{a_{k,\pm}^\varepsilon}{2\pi} \in S^{-1+}(\mathcal{J}_\omega \times \mathbb{R}_\tau), \quad \tilde{\chi} := \chi_{k,\pm}^+ \in C_c^\infty(\mathbb{R} \setminus 0). \end{aligned} \tag{8-28}$$

Then we have for $x \in \Omega$

$$v_{k,\pm}^{\varepsilon,+}(x,t) = \int_{\mathbb{R}} \varphi(\omega) U_\pm(x,t,\omega)\, d\omega,$$

which together with the Fourier characterization of the Sobolev space $H^{1/2-\beta}(\mathbb{R}^2)$ implies the following bound, where $\widehat{U}_\pm$ denotes the Fourier transform of $U_\pm$ in the x-variable:

$$\|v_{k,\pm}^{\varepsilon,+}(t)\|_{H^{1/2-\beta}(\Omega)}^2 \le \int_{\mathbb{R}^2} \langle\xi\rangle^{1-2\beta} \left| \int_{\mathbb{R}} \varphi(\omega) \widehat{U}_\pm(\xi,t,\omega)\, d\omega \right|^2 d\xi. \tag{8-29}$$

(2) Thinking of $L_\omega^\pm$ (see (4-3)) as elements of the dual $\mathbb{R}^2$ of $(\mathbb{R}^2)^*$ we have by (4-4)

$$(\mathbb{R}^2)^* \ni \xi = L_\omega^+(\xi)\ell_\omega^+ + L_\omega^-(\xi)\ell_\omega^-, \quad \ell_\omega^\pm \in (\mathbb{R}^2)^*.$$

Hence, since $\det \partial(x_1,x_2)/\partial(\ell_\omega^+, \ell_\omega^-) = \frac{1}{2}\omega\sqrt{1-\omega^2}$,

$$\begin{aligned} \mathcal{F}_{x\to\xi}\big(f(\ell_\omega^+(x)) g(\ell_\omega^-(x))\big) &= \int_{\mathbb{R}^2} e^{-iL_\omega^+(\xi)\ell_\omega^+(x) - iL_\omega^-(\xi)\ell_\omega^-(x)} f(\ell_\omega^+(x)) g(\ell_\omega^-(x))\, dx \\ &= \tfrac{1}{2}\omega\sqrt{1-\omega^2}\, \hat{f}(L_\omega^+(\xi)) \hat{g}(L_\omega^-(\xi)). \end{aligned}$$

Consequently,

$$\widehat{U}_\pm(\xi,t,\omega) = L_\omega^\pm(\xi) D(L_\omega^\pm(\xi), t, \omega-\lambda) \hat{\psi}(L_\omega^\mp(\xi)) e^{-iL_\omega^\pm(\xi)\ell_\omega^\pm(x_k^\pm(\omega))} b(L_\omega^\pm(\xi), \omega), \tag{8-30}$$

where we absorbed the Jacobian into b and put

$$D(\tau,t,\rho) := \int_{t/\tau}^{(\operatorname{sgn}\tau)\infty} \tilde{\chi}(r) e^{i\tau\rho r}\, dr, \quad \tau \ne 0.$$

Since $\tilde{\chi} \in C_c^\infty(\mathbb{R})$, we have $D(t,\tau,\rho) \to 0$, for fixed $\tau \ne 0$ as $t \to \infty$, uniformly in ρ. In view of the support condition in Lemma 8.3 (which implies that $|L_\omega^\pm(\xi)| \ge 1$ on the support of $\widehat{U}_\pm$) we then get

$$\int_{\mathbb{R}} \varphi(\omega) \widehat{U}_\pm(\xi,t,\omega)\, d\omega \to 0 \quad \text{as } t \to \infty, \tag{8-31}$$

for all $\xi \in \mathbb{R}^2$. Using the dominated convergence theorem and (8-29) we see that to establish (8-26) it is enough to show that the integrand on the right-hand side of (8-29) is bounded by a t-independent integrable function of ξ.

(3) We have

$$|D(\tau, t, \rho)| \le C\langle \tau\rho\rangle^{-1}. \tag{8-32}$$

Indeed, since the support of $\tilde{\chi}$ is bounded, we have $D = \mathcal{O}(1)$. On the other hand, when $|\tau\rho| > 1$ we can integrate by parts using that $e^{i\tau\rho r} = (i\tau\rho)^{-1}\partial_r e^{i\tau\rho r}$, which gives the estimate.

Recalling (8-30), (8-32) and using that $\hat{\psi} \in \mathscr{S}(\mathbb{R})$ by (8-27) and $b(\tau,\omega) = \mathcal{O}(\langle\tau\rangle^{-1+\beta/2})$ by (8-28), we get

$$|\widehat{U}_\pm(\xi, t, \omega)| \le C\langle L^\mp_\omega(\xi)\rangle^{-10}\langle L^\pm_\omega(\xi)\rangle^{\beta/2}\langle(\omega-\lambda)L^\pm_\omega(\xi)\rangle^{-1}.$$

Thus is remains to show that

$$\left\| \int_{\mathbb{R}} \varphi(\omega) H(\xi,\omega)\, d\omega \right\|_{L^2(\mathbb{R}^2_\xi)} < \infty, \quad \text{where } H(\xi,\omega) := \langle\xi\rangle^{1/2-\beta}\langle L^\mp_\omega(\xi)\rangle^{-10}\langle L^\pm_\omega(\xi)\rangle^{\beta/2}\langle(\omega-\lambda)L^\pm_\omega(\xi)\rangle^{-1}.$$

Using the integral version of the triangle inequality for $L^2(\mathbb{R}^2_\xi)$, this reduces to

$$\int_{\mathbb{R}} \varphi(\omega)\|H(\xi,\omega)\|_{L^2(\mathbb{R}^2_\xi)}\, d\omega < \infty. \tag{8-33}$$

Fix $\omega \in \operatorname{supp}\varphi$ and make the linear change of variables $\xi \mapsto \eta = (\eta_+, \eta_-)$, $\eta_\pm = L^\pm_\omega(\xi)$. Then we see that

$$\|H(\xi,\omega)\|^2_{L^2(\mathbb{R}^2_\xi)} \le C\int_{\mathbb{R}^2} \langle\eta\rangle^{1-2\beta}\langle\eta_\mp\rangle^{-20}\langle\eta_\pm\rangle^{\beta}\langle(\omega-\lambda)\eta_\pm\rangle^{-2}\, d\eta.$$

Integrating out $\eta_\mp$ and making the change of variables $\zeta := (\omega-\lambda)\eta_\pm$, we get (for ω bounded and assuming $\beta < 1$)

$$\|H(\xi,\omega)\|^2_{L^2(\mathbb{R}^2_\xi)} \le C\int_{\mathbb{R}} \langle\eta_\pm\rangle^{1-\beta}\langle(\omega-\lambda)\eta_\pm\rangle^{-2}\, d\eta_\pm \le C|\omega-\lambda|^{\beta-2}.$$

Thus

$$\int_{\mathbb{R}} \varphi(\omega)\|H(\xi,\omega)\|_{L^2(\mathbb{R}^2_\xi)}\, d\omega \le C\int_0^1 |\omega-\lambda|^{\beta/2-1}\, d\omega < \infty,$$

giving (8-33) and finishing the proof. □

8.4. ***Proof of Theorem 1.3.*** We now review how the pieces presented in Sections 8.1–8.3 fit together to give the proof of Theorem 1.3.

In view of (8-2), Lemma 8.1, and (8-7) it suffices to show that

$$\begin{gathered} \Delta_\Omega^{-1} w_1(t) = u^+(\lambda) + r_2(t) + \tilde{e}(t), \\ \|r_2(t)\|_{H^1(\Omega)} = \mathcal{O}(1), \qquad \|\tilde{e}(t)\|_{H^{(1/2)-}(\Omega)} \to 0 \quad \text{as } t \to \infty. \end{gathered} \tag{8-34}$$

We use the formula (8-6) which expresses $\Delta_\Omega^{-1} w_1(t)$ as an integral featuring the distributions $u^\varepsilon(x, \omega)$, $\varepsilon \in \{+, -\}$. Lemma 8.3 gives a decomposition of u^ε into the conormal components $g_{k,\pm}^\varepsilon$ and the smooth component u_0^ε. Lemma 8.4 then shows that the contribution of u_0^ε to $\Delta_\Omega^{-1} w_1(t)$ can be absorbed into $r_2(t)$.

The contribution of conormal terms $g_{k,\pm}^\varepsilon$ to $\Delta_\Omega^{-1} w_1(t)$ is then given by (8-19). Restricting integration in r using the cut-off $1 - \chi$ in Lemma 8.5 produces other terms which can be absorbed into $r_2(t)$. The limit of the remaining terms as $t \to +\infty$ is described in Lemma 8.6: summing over k and $\pm$ gives the leading term as $u^+(\lambda)$ (as seen by returning to Lemma 8.3, where the cutoff χ does not matter by (8-22) and (8-23)) and terms which again can be absorbed in $r_2(t)$.

What is left is given by a sum of (8-25). Lemma 8.7 shows that those terms all go to 0 in $H^{(1/2)-}(\Omega)$ as $t \to \infty$ and their sum constitutes $\tilde{e}(t)$.

Acknowledgements

We would like to thank Peter Hintz and Hart Smith for helpful discussions concerning Section 6 and injectivity of $\mathcal{C}_\lambda$ (see the Remark in Section 7.3), respectively, and for their interest in this project. Special thanks are also due to Leo Maas for his comments on an early version of the paper, help with the Introduction and references and experimental data in Figure 2 (reproduced with permission from the American Institute of Physics). We are grateful to the anonymous referees for a careful reading of the paper and many comments to improve the presentation. We would also like to reiterate our gratitude to Sergey Denisov for alerting us to important results in the Russian literature. Dyatlov was partially supported by NSF CAREER grant DMS-1749858 and a Sloan Research Fellowship, while Wang and Zworski were partially supported by NSF grant DMS-1901462.

References

[Aleksandrjan 1960] R. A. Aleksandrjan, "Spectral properties of operators arising from systems of differential equations of Sobolev type", *Trudy Moskov. Mat. Obšč.* **9** (1960), 455–505. In Russian. MR Zbl

[Bajars et al. 2013] J. Bajars, J. Frank, and L. R. M. Maas, "On the appearance of internal wave attractors due to an initial or parametrically excited disturbance", *J. Fluid Mech.* **714** (2013), 283–311. MR Zbl

[Baladi 2018] V. Baladi, *Dynamical zeta functions and dynamical determinants for hyperbolic maps: a functional approach*, Ergebnisse der Mathematik (3) **68**, Springer, 2018. MR Zbl

[Brouzet 2016] C. Brouzet, *Internal wave attractors: from geometrical focusing to non-linear energy cascade and mixing*, Ph.D. thesis, Université de Lyon, 2016, available at https://theses.hal.science/tel-01361201/en.

[Cycon et al. 1987] H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon, *Schrödinger operators with application to quantum mechanics and global geometry*, Springer, 1987. MR Zbl

[Dauxois et al. 2018] T. Dauxois, S. Joubaud, P. Odier, and A. Venaille, "Instabilities of internal gravity wave beams", pp. 131–156 in *Annual review of fluid mechanics*, *Vol.* 50, edited by S. H. Davis, Annu. Rev. Fluid Mech. **50**, Annual Reviews, Palo Alto, CA, 2018. MR Zbl

[Davies 1995] E. B. Davies, *Spectral theory and differential operators*, Cambridge Studies in Advanced Mathematics **42**, Cambridge University Press, 1995. MR Zbl

[Dyatlov and Zworski 2017] S. Dyatlov and M. Zworski, "Ruelle zeta function at zero for surfaces", *Invent. Math.* **210**:1 (2017), 211–229. MR Zbl

[Dyatlov and Zworski 2019a] S. Dyatlov and M. Zworski, *Mathematical theory of scattering resonances*, Graduate Studies in Mathematics **200**, Amer. Math. Soc., Providence, RI, 2019. MR Zbl

[Dyatlov and Zworski 2019b] S. Dyatlov and M. Zworski, "Microlocal analysis of forced waves", *Pure Appl. Anal.* **1**:3 (2019), 359–384. MR Zbl

[Faure et al. 2008] F. Faure, N. Roy, and J. Sjöstrand, "Semi-classical approach for Anosov diffeomorphisms and Ruelle resonances", *Open Math. J.* **1** (2008), 35–81. MR Zbl

[Fokin 1993] M. V. Fokin, "The existence of the singular spectrum and the asymptotic behavior of solutions in a problem of S. L. Sobolev", *Dokl. Akad. Nauk* **333**:3 (1993), 304–307. In Russian. MR

[Galkowski and Zworski 2022] J. Galkowski and M. Zworski, "Viscosity limits for zeroth-order pseudodifferential operators", *Comm. Pure Appl. Math.* **75**:8 (2022), 1798–1869. MR Zbl

[Grigis and Sjöstrand 1994] A. Grigis and J. Sjöstrand, *Microlocal analysis for differential operators*: *an introduction*, London Mathematical Society Lecture Note Series **196**, Cambridge University Press, 1994. MR Zbl

[Hazewinkel et al. 2010] J. Hazewinkel, C. Tsimitri, L. R. M. Maas, and S. B. Dalziel, "Observations on the robustness of internal wave attractors to perturbations", *Physics of Fluids* **22**:11 (2010), art. id. 107102.

[Hörmander 1990] L. Hörmander, *The analysis of linear partial differential operators, I*: *Distribution theory and Fourier analysis*, 2nd ed., Grundl. Math. Wissen. **256**, Springer, 1990. MR Zbl

[Hörmander 1994] L. Hörmander, *The analysis of linear partial differential operators*, *III*: *Pseudo-differential operators*, Grundl. Math. Wissen. **274**, Springer, 1994. MR Zbl

[John 1941] F. John, "The Dirichlet problem for a hyperbolic equation", *Amer. J. Math.* **63** (1941), 141–154. MR Zbl

[Lenci et al. 2023] M. Lenci, C. Bonanno, and G. Cristadoro, "Internal-wave billiards in trapezoids and similar tables", *Nonlinearity* **36**:2 (2023), 1029–1052. MR Zbl

[Li 2024] Z. Li, "Internal waves in aquariums with characteristic corners", preprint, 2024. arXiv 2402.01896

[Maas 2005] L. R. M. Maas, "Wave attractors: linear yet nonlinear", *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **15**:9 (2005), 2557–2782. MR Zbl

[Maas and Lam 1995] L. R. M. Maas and F.-P. A. Lam, "Geometric focusing of internal waves", *J. Fluid Mech.* **300** (1995), 1–41. MR Zbl

[Maas et al. 1997] L. R. M. Maas, D. Benielli, J. Sommeria, and F.-P. A. Lam, "Observation of an internal wave attractor in a confined, stably stratified fluid", *Nature* **388**:6642 (1997), 557–561.

[Manders et al. 2003] A. M. M. Manders, J. J. Duistermaat, and L. R. M. Maas, "Wave attractors in a smooth convex enclosed geometry", *Phys. D* **186**:3-4 (2003), 109–132. MR Zbl

[de Melo and van Strien 1993] W. de Melo and S. van Strien, *One-dimensional dynamics*, Ergebnisse der Mathematik (3) **25**, Springer, 1993. MR Zbl

[Munkres 2000] J. R. Munkres, *Topology*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2000. MR Zbl

[Nogueira and Troubetzkoy 2022] A. Nogueira and S. Troubetzkoy, "Chess billiards", *Math. Intelligencer* **44**:4 (2022), 331–338. MR Zbl

[Ralston 1973] J. V. Ralston, "On stationary modes in inviscid rotating fluids", *J. Math. Anal. Appl.* **44** (1973), 366–383. MR Zbl

[Rieutord and Valdettaro 2018] M. Rieutord and L. Valdettaro, "Axisymmetric inertial modes in a spherical shell at low Ekman numbers", *J. Fluid Mech.* **844** (2018), 597–634. MR Zbl

[Sibgatullin and Ermanyuk 2019] I. N. Sibgatullin and E. V. Ermanyuk, "Internal and inertial wave attractors: a review", *Prikl. Mekh. Tekhn. Fiz.* **60**:2 (2019), 113–136. In Russian; translated in *J. Appl. Mech. Tech. Phys.* **60**:2 (2019), 284–302. MR Zbl

[Sobolev 1954] S. L. Sobolev, "On a new problem of mathematical physics", *Izv. Akad. Nauk SSSR Ser. Mat.* **18** (1954), 3–50. In Russian. MR Zbl

[Sternberg 1957] S. Sternberg, "Local C^n transformations of the real line", *Duke Math. J.* **24** (1957), 97–102. MR Zbl

[Tao 2024] Z. Tao, "On 0-th order pseudo-differential operators on the circle", *Proc. Amer. Math. Soc.* **152**:8 (2024), 3289–3297. MR Zbl

[Troitskaya 2017] S. Troitskaya, "Mathematical analysis of inertial waves in rectangular basins with one sloping boundary", *Stud. Appl. Math.* **139**:3 (2017), 434–456. MR Zbl

[Unterberger 1971] A. Unterberger, "Résolution d'équations aux dérivées partielles dans des espaces de distributions d'ordre de régularité variable", *Ann. Inst. Fourier* (*Grenoble*) **21**:2 (1971), 85–128. MR Zbl

[Colin de Verdière 2020] Y. Colin de Verdière, "Spectral theory of pseudodifferential operators of degree 0 and an application to forced linear waves", *Anal. PDE* **13**:5 (2020), 1521–1537. MR Zbl

[Colin de Verdière and Saint-Raymond 2020] Y. Colin de Verdière and L. Saint-Raymond, "Attractors for two-dimensional waves with homogeneous Hamiltonians of degree 0", *Comm. Pure Appl. Math.* **73**:2 (2020), 421–462. MR Zbl

[Walsh 1999] J. A. Walsh, "The dynamics of circle homeomorphisms: a hands-on introduction", *Math. Mag.* **72**:1 (1999), 3–13. MR Zbl

[Wang 2022] J. Wang, "Dynamics of resonances for 0th order pseudodifferential operators", *Comm. Math. Phys.* **391**:2 (2022), 643–668. MR Zbl

[Wang 2023] J. Wang, "The scattering matrix for 0th order pseudodifferential operators", *Ann. Inst. Fourier* (*Grenoble*) **73**:5 (2023), 2185–2237. MR Zbl

[Wunsch 1968] C. Wunsch, "On the propagation of internal waves up a slope", *Deep Sea Research Ocean. Abstracts* **15**:3 (1968), 251–258.

[Yoccoz 1995] J.-C. Yoccoz, "Centralisateurs et conjugaison différentiable des difféomorphismes du cercle", pp. 89–242 in *Petits diviseurs en dimension* 1, Astérisque **231**, Société Mathématique de France, Paris, 1995. MR Zbl

Received 21 Mar 2022. Revised 3 Aug 2023. Accepted 11 Sep 2023.

SEMYON DYATLOV: dyatlov@math.mit.edu
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, United States

JIAN WANG: wangjian@email.unc.edu
Department of Mathematics, University of North Carolina, Chapel Hill, NC, United States

MACIEJ ZWORSKI: zworski@math.berkeley.edu
Department of Mathematics, University of California, Berkeley, CA, United States

ANALYSIS AND PDE
Vol. 18 (2025), No. 1, pp. 93–108
DOI: 10.2140/apde.2025.18.93

LARGE SETS CONTAINING NO COPIES OF A GIVEN INFINITE SEQUENCE

MIHAIL N. KOLOUNTZAKIS AND EFFIE PAPAGEORGIOU

Suppose a_n is a real, nonnegative sequence that does not increase exponentially. For any $p < 1$, we construct a Lebesgue measurable set $E \subseteq \mathbb{R}$ which has measure at least p in any unit interval and which contains no affine copy $\{x + ta_n : n \in \mathbb{N}\}$ of the given sequence (for any $x \in \mathbb{R}$, $t > 0$). We generalize this to higher dimensions and also for some "nonlinear" copies of the sequence. Our method is probabilistic.

1. Introduction

In Euclidean Ramsey theory, one is interested in assuming some kind of largeness for sets E in Euclidean space $\mathbb{R}^d$, or sometimes in $\mathbb{Z}^d$, and concluding that E then contains a "copy" of a pattern. The most famous such example is perhaps Szemeredi's theorem [1975], which states that any subset of the integers with positive density contains arbitrarily long arithmetic progressions. Another well-known example is the theorem of Falconer and Marstrand [1986], Fürstenberg, Katznelson and Weiss [Fürstenberg et al. 1990] and Bourgain [1986] (see also [Kolountzakis 2004]): if the set $E \subseteq \mathbb{R}^d$ has positive Lebesgue density (this means there are arbitrarily large cubes where E takes up at least a constant fraction of the measure) then its points implement all sufficiently large distances (conjecture by Székely [1983]).

Another well-known problem, very much related to the contents of this paper, is the so-called *Erdős similarity problem*: a set $\mathbb{A} \subseteq \mathbb{R}$ is called *universal in measure* if, whenever $E \subseteq \mathbb{R}$ has positive Lebesgue measure, we can find an affine copy of A contained in E. In other words, $x + t\mathbb{A} \subseteq E$ for some $x \in \mathbb{R}$, $t > 0$. It is easy to see that every finite set $\mathbb{A}$ is universal (just look close enough to some point of density of E, shrink $\mathbb{A}$ enough and average the number of points of the copy of $\mathbb{A}$ that belong to E over translates of $\mathbb{A}$ nearby) but it has been conjectured [Erdős 2015] (see also [Croft et al. 1991, p. 183]) that no infinite set $\mathbb{A}$ can be universal in measure. This is known for many classes of infinite sets but not for all [Chlebik 2015; Falconer 1984; Gallagher et al. 2023; Humke and Laczkovich 1998; Komjáth 1983]. Clearly it would suffice to prove this for $\mathbb{A}$ a positive sequence a_n decreasing to 0, but if a_n decays fast to 0 (so it is in some sense sparse, hence not that hard to contain) this is still unknown. On the contrary, this is known when $\log 1/a_n = o(n)$. This is not known if $a_n = 2^{-n}$, for example.

In this paper we consider an analogue of the Erdős similarity problem "in the large". Let $\mathbb{A} \subseteq \mathbb{R}$ be a discrete, unbounded, infinite set in $\mathbb{R}$. Can we find a "large" measurable set $E \subseteq \mathbb{R}$ which does not contain any affine copy $x + t\mathbb{A}$ of $\mathbb{A}$ (for any $x \in \mathbb{R}$, $t > 0$)? Our attention was drawn to this problem by

Supported by the Hellenic Foundation for Research and Innovation, Project HFRI-FM17-1733 and by University of Crete Grant 4725.
MSC2020: 05D40, 28A80.
Keywords: Erdős similarity problem, Euclidean Ramsey theory, probabilistic method.

a recent paper by Bradford, Kohut and Mooroogen [Bradford et al. 2023] in which the authors prove that if $\mathbb{A}$ is an infinite arithmetic progression then this is indeed possible: for any $p \in [0, 1)$, they construct a Lebesgue measurable set E, with measure at least p in any interval of length 1, which does not contain any affine copy of $\mathbb{A}$. This is clearly equivalent to being able to obtain, for any $p \in [0, 1)$, a set E avoiding all infinite arithmetic progressions and having measure $\geq p$ in any interval of length 1 *whose endpoints are integers*. (Indeed, if the set E has measure at least p in every interval of the form $[n, n+1]$, $n \in \mathbb{Z}$, then, since for any x the interval $[x, x+1]$ is contained in the union of two such unit-length intervals with integer endpoints, we obtain that $[x, x+1] \setminus E$ has measure at most $2(1-p)$. Since p can be as close to 1 as we want, this implies that $[x, x+1] \setminus E$ has measure as close to 0 as we want.) From now on we follow this simplification, and we deal only with intervals with integer endpoints (in any dimension).

We generalize the result of [Bradford et al. 2023] to sequences of nonnegative numbers $\mathbb{A}$ which do not grow too fast. To state our result, we introduce the following class of sequences.

Definition 1.1. We say that a real sequence $\mathbb{A} = \{a_n : n \in \mathbb{N}\}$ is in the *class* (A) if

(1) $a_0 = 0$,

(2) $a_{n+1} - a_n \geq 1$ for every $n \in \mathbb{N}$,

(3) $\log a_n = o(n)$.

Remark 1.2. Since the problem we are studying is translation invariant, (1) in Definition 1.1 is unnecessary, but we keep it as it simplifies the proofs somewhat.

Writing

$$A(t) = |\mathbb{A} \cap [0, t]| \tag{1-1}$$

for the *counting function* of the set $\mathbb{A}$, notice that the growth (3) is equivalent to the limit, as $t \to +\infty$,

$$\frac{A(t)}{\log t} \to +\infty. \tag{1-2}$$

Our main result is the following.

Theorem 1.3. *Consider the sequence $\mathbb{A} = \{a_n : n \in \mathbb{N}\}$ which belongs to the class* (A). *Then, for each $0 \leq p < 1$, there exists a Lebesgue measurable set $E \subseteq \mathbb{R}$ such that*

$$|E \cap [m, m+1]| \geq p \quad \textit{for all } m \in \mathbb{Z},$$

but E does not contain any affine copy of $\mathbb{A}$.

As in the case of the Erdős similarity problem described above, the sparser the set $\mathbb{A}$, the easier it should be to contain it in large sets, so it is not surprising that we had to impose a growth condition (to belong to the class (A)). It remains an open question if a similar set E can be constructed when $\mathbb{A}$ grows exponentially or faster.

Question 1. *Is there a sequence $0 < a_n \to +\infty$ and a number $p \in [0, 1)$ such that one can find an affine copy of $\mathbb{A} = \{a_n : n \in \mathbb{N}\}$ in any set $E \subseteq \mathbb{R}$ which has measure more than p in any interval of length* 1?

Unlike the approach taken in [Bradford et al. 2023], our method of proof is probabilistic. We construct a family of random sets and we show that, with high probability, such a random set will have all the properties we want. This method turns out to be extremely flexible, and this allows us to generalize. Not only can we deal with essentially arbitrary and unstructured sequences $\mathbb{A}$, but we can also relax the sense in which we seek copies of $\mathbb{A}$ in the large set E. Instead of scaling the elements of $\mathbb{A}$ and translating them,

$$x + ta_n, \quad x \in \mathbb{R}, \quad t > 0,$$

we can allow for more general transformations

$$x + \phi(n,t) \cdot a_n, \quad x \in \mathbb{R}, \quad t > 0. \tag{1-3}$$

Theorem 1.4. *Consider the set* $\mathbb{A} = \{a_n : n \in \mathbb{N}\}$ *which belongs to the class* (A), *and let*

$$\phi(n,t) : \mathbb{N} \times (0, +\infty) \to (0, +\infty)$$

be such that, for each n, *the function* $\phi(n,t)$ *is increasing in* t, *and such that, for all* $n \in \mathbb{N}$, *we have*

$$C_1 t \le \phi(n+1,t)a_{n+1} - \phi(n,t)a_n \tag{1-4}$$

and

$$\phi(n,t) \le C_2 t \quad \text{for all } t > 0, \tag{1-5}$$

for some $C_1, C_2 > 0$. *Then, for each* $0 \le p < 1$, *there exists a Lebesgue measurable set* $E \subseteq \mathbb{R}$ *such that* E *intersects every interval of unit length in a set of measure at least* p, *but* E *does not contain the set*

$$\{x + \phi(n,t) \cdot a_n : n \in \mathbb{N}\}$$

for any choice of $x \in \mathbb{R}$, $t > 0$.

We adopt certain arguments from [Kolountzakis 1997, Section 3] where it is proved, on the Erdős similarity problem, that sequences with a finite limit, say 0, which are not decaying very fast (e.g., they decay polynomially or subexponentially but not, for instance, exponentially fast — compare to our growth condition (3)) cannot be universal in measure, by showing the existence of a randomly constructed set $E \subseteq [0, 1]$, avoiding all affine copies of the sequence.

The measure assumption makes this problem different than other "avoidance" problems, where the avoiding set is often taken to have zero Lebesgue measure but to have large Hausdorff dimension or Fourier dimension. For example, in [Keleti 2008], a compact subset of $\mathbb{R}$ is constructed that has full Hausdorff dimension but does not contain any 3-term arithmetic progression. See also [Cruz et al. 2022; Denson et al. 2021; Fraser and Pramanik 2018; Maga 2011; Máthé 2017; Shmerkin 2017; Yavicoli 2021].

We can also prove the following result in higher dimension (Theorem 1.5). We phrase it as avoiding linear images of a set in Euclidean space into another Euclidean space. In this manner we obtain easily some corollaries, Theorem 1.3 one of them, and its proof is rather simpler than that of Theorem 1.3 given in Section 3. But it does not extend easily to more complicated transformations such as those in Theorem 1.4, so we choose to stay with linear maps.

Theorem 1.5. *Let $d_1, d \geq 1,\ b,\ f > 0$ and $p \in [0, 1)$. Let also $\alpha(R)$ be a function satisfying*

$$\frac{\alpha(R)}{\log R} \to +\infty \quad \text{as } R \to +\infty.$$

Then, if $\mathbb{A} \subseteq \mathbb{R}^{d_1}$ is a discrete point set such that

$$|\mathbb{A} \cap B_R(0)| \leq C_2 R^b, \quad R > 0, \tag{1-6}$$

there is a set $E \subseteq \mathbb{R}^d$ such that:

(i) *$|E \cap (m + [0, 1]^d)| \geq p$ for all $m \in \mathbb{Z}^d$.*

(ii) *For any linear map $T : \mathbb{R}^{d_1} \to \mathbb{R}^d$, if, for arbitrarily large values of R,*

$$T(\mathbb{A}) \cap B_R(0) \tag{1-7}$$

contains at least $\alpha(R)$ points with separation R^{-f} then

$$T(\mathbb{A}) \text{ is not contained in } E. \tag{1-8}$$

Proof of Theorem 1.3 using Theorem 1.5. Apply Theorem 1.5 with $d_1 = 2,\ d = 1,\ b = 1,\ \alpha(x) = A(x^{1/2})$ (where $A(x)$ is the counting function of $\mathbb{A}$), $f = 1$ (there is great flexibility in choosing $\alpha(x),\ b,\ f$) and the set

$$P = \mathbb{A} \times \{1\} \subseteq \mathbb{R}^2$$

to obtain a set $E \subseteq \mathbb{R}$ satisfying $|E \cap [m, m+1]| \geq p$ for all $m \in \mathbb{Z}$. We see that (1-6) is satisfied. Let now $T : \mathbb{R}^2 \to \mathbb{R}$ be given by the 1×2-matrix $T = (t, x)$, so that

$$T(P) = x + t\mathbb{A}.$$

For any $x \in \mathbb{R},\ t > 0$, the set $(x + t\mathbb{A}) \cap [-R, R]$ contains at least $A(R/t)$ points of separation t, so, if R is large enough, it contains $\alpha(R) = A(R^{1/2})$ points with separation R^{-1}. It follows that $x + t\mathbb{A}$ is not contained in E. □

Corollary 1.6 (avoiding linear images of general sets in high dimension). *Let $p \in [0, 1),\ d \geq 1,\ a_n \in \mathbb{R}^d$ for $n \in \mathbb{N}$, with $\log|a_n| = o(n)$ and $|a_n - a_{n+1}| \geq 1$ for all $n \in \mathbb{N}$. Then there is a set $E \subseteq \mathbb{R}^d$ such that, for all $m \in \mathbb{Z}^d$, we have $|E \cap (m + [0, 1]^d)| \geq p$, and such that, for all $x \in \mathbb{R}^d$ and for all nonsingular linear $T : \mathbb{R}^d \to \mathbb{R}^d$, the set $\{x + Ta_n : n \in \mathbb{N}\}$ is not contained in E.*

Proof. Take $\mathbb{A} \subseteq \mathbb{R}^{2d}$ to be the set $A \times \{(\overbrace{1, 0, \ldots, 0}^{d})\}$, where $A = \{a_n : n \in \mathbb{N}\}$. Writing $A(s) = \#(A \cap B_s(0))$ for the counting function of A, we have $A(R)/\log R \to +\infty$. Use Theorem 1.5 with $d_1 = 2d,\ b = 1,\ \alpha(R) = A(R^{1/2}),\ f = 1$. Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be nonsingular, $x \in \mathbb{R}^d$, and define the linear map $S : \mathbb{R}^{2d} \to \mathbb{R}^d$ by

$$S(u, v) = S(u, v_1, v_2, \ldots, v_d) = Tu + v_1 x.$$

In other words the $d \times (2d)$-matrix of S is $(T \mid x \mid 0)$ in block form. It follows that

$$S(\mathbb{A}) = \{Ta_n + x : n \in \mathbb{N}\}.$$

Since T is nonsingular it follows that, if $R > 0$ is sufficiently large, the set $S(\mathbb{A}) \cap B_R(0)$ contains at least $\alpha(R)$ points with separation $\geq R^{-1}$, so the set $E \subseteq \mathbb{R}^d$ furnished by Theorem 1.5 does not contain $S(\mathbb{A})$, as we had to prove. □

Corollary 1.7 [Bradford et al. 2023, Corollary 6]. *If $p \in [0, 1)$ then there exists a set $E \subseteq \mathbb{R}^d$ such that $|E \cap (m + [0, 1]^d)| \geq p$ for all $m \in \mathbb{Z}^d$ and it does not contain any set of the form $x + \mathbb{N}\Delta$, with $x \in \mathbb{R}^d$ and $\Delta \in \mathbb{R}^d \setminus \{0\}$ (an arithmetic progression in $\mathbb{R}^d$).*

Proof. We use Corollary 1.6 with the sequence $a_n = (n, 0, \ldots, 0) \in \mathbb{R}^d$, $x \in \mathbb{R}^d$ and any nonsingular $d \times d$-matrix T that maps $(1, 0, \ldots, 0)$ to Δ. □

The outline of this note is as follows. In Section 3 we give the proof of Theorem 1.3 without using Theorem 1.5, and we indicate how the same proof also works for Theorem 1.4. In Section 4 we extend our technique to cover linear transformations of given sequences from one Euclidean space to another and prove Theorem 1.5 and some corollaries.

Added in revision: The results in [Burgin et al. 2023], which came after this paper was submitted, are very relevant to the results in this paper and contain some improvements.

2. Warm-up and some basic tools: no translational copies

In this section we introduce the basic probabilistic method by proving the more restricted Theorem 2.1: we can avoid all translations of a given infinite sequence $0 \leq a_n \to +\infty$ with a set which is arbitrarily large everywhere. This is considerably easier than avoiding all affine copies of the sequence, when scaling the sequence as well as translating it is allowed. For translations we have only one degree of freedom while for affine copies we have two. Still, some important ingredients of the method will be evident in the proof of Theorem 2.1 below. In Section 3 we will introduce the extra discretization in scaling space that will be required.

Theorem 2.1. *Let $\mathbb{A} = \{a_0 = 0 < a_1 < a_2 < \cdots\} \subseteq \mathbb{R}$ be a sequence with $a_n \to +\infty$, and let $p \in [0, 1)$. Then we can find a Lebesgue measurable set $E \subseteq \mathbb{R}$ such that no translate of $\mathbb{A}$,*

$$x + \mathbb{A}, \quad x \in \mathbb{R},$$

is contained in E, and such that, for each $m \in \mathbb{Z}$, we have

$$|E \cap [m, m+1]| \geq p.$$

Proof. Let $q < 1$ be defined by $1 - q = \frac{1}{2}(1 - p)$ $\big($or $q = \frac{1}{2}(1 + p)\big)$. Passing to a subsequence, we can assume that $a_{n+1} - a_n \geq 1$ for all n. We construct a random set E by breaking up each unit interval $[m, m+1]$, $m \in \mathbb{Z}$, into a number N_m of equal intervals and keeping each of these subintervals with probability q, independently, in our set E. See Figure 1 for an illustration of the set E. As $|m|$ increases, the number N_m will also have to increase, so let us take $N_m = \max\{K, |m|\}$ say, where the large positive integer K will be determined later.

Define now, for $x \in \mathbb{R}$, the random function

$$\phi(x) = \mathbb{1}(x + \mathbb{A} \subseteq E).$$

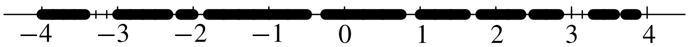

Figure 1. The random set E.

Since all points of $x+\mathbb{A}$ are in different random intervals, it follows, by independence, that $\mathbb{E}\phi(x) = \mathbb{P}[x+\mathbb{A} \subseteq E] = 0$. Let the set of "bad" x be

$$B = \{x \in \mathbb{R} : x + \mathbb{A} \subseteq E\}.$$

We have

$$\mathbb{E}|B| = \int \mathbb{E}\phi(x)\, dx = 0;$$

hence $|B|$ is almost surely 0.

It remains to make sure that $|E \cap [m, m+1]| \geq p$ for all $m \in \mathbb{Z}$. Fix m, and let $X_1, \ldots, X_{N_m}$ be 0/1 random variables such that X_i is 0 if we included the i-th subinterval of $[m, m+1]$ in the set E and is 1 otherwise. In other words, X_i denotes the absence of the i-th subinterval from the set E. Clearly $\mathbb{E}X_i = 1-q$, and the random variable

$$X = \sum_{i=1}^{N_m} X_i \text{ (the number of missing subintervals)}$$

is a sum of independent indicator random variables with $\mathbb{E}X = (1-q)N_m$ and we can use the very versatile large deviation *Chernoff inequality* (to be used repeatedly in Sections 3.1 and 4 below)

$$\mathbb{P}[|X - \mathbb{E}X| \geq \epsilon \mathbb{E}X] \leq 2e^{-c_\epsilon \mathbb{E}X} \tag{2-1}$$

(see [Alon and Spencer 1992; Chernoff 1952]) with $\epsilon = 1$ to obtain

$$\begin{aligned} \mathbb{P}[|E \cap [m, m+1]| < p] = \mathbb{P}[X > (1-p)N_m] &= \mathbb{P}[X - \mathbb{E}X > \mathbb{E}X] \\ &\leq 2\exp(-c_1(1-q)\max\{K, |m|\}). \end{aligned} \tag{2-2}$$

Define now the bad events $B_m = \{|E \cap [m, m+1]| < p\}$ which we do not want to hold, for all $m \in \mathbb{Z}$, and observe that the above inequality means that we can choose K large enough to achieve

$$\sum_{m \in \mathbb{Z}} \mathbb{P}[B_m] < \tfrac{1}{2}.$$

This means that, with probability at least $\frac{1}{2}$, none of the bad events B_m hold and, with the same probability, the set B has measure 0. We now amend our random set E by removing from it the set B (the set of first terms of those $x+\mathbb{A}$ which are contained in E). Thus arises a set E', which differs from E by a set of measure 0 and which contains no translate of $\mathbb{A}$. □

Remark 2.2. It is not necessary to assume that $a_n \to +\infty$ in Theorem 2.1. It suffices to assume that the set $\mathbb{A}$ is infinite. If $\mathbb{A}$ does not contain a sequence tending to infinity (for Theorem 2.1 to apply to it) then it will have a finite accumulation point, so a result of Komjáth [1983] guarantees the existence of a set $\widetilde{E} \subseteq [0, 1]$, of measure arbitrarily close to 1, which contains no translate of $\mathbb{A}$. Repeating $\widetilde{E}$ 1-periodically,

$$E = \bigcup_{n\in\mathbb{Z}} \widetilde{E} + n,$$

we obtain a set E with the required properties. For a probabilistic proof of this result in the spirit of the present paper, see [Kolountzakis 1997].

Remark 2.3. The Chernoff inequality (2-1) is extremely useful when one needs to control a random variable X (this means that one wants to ensure, with high probability, that X is near its mean $\mathbb{E}X$) which is a sum of indicator, independent random variables. The key is that the mean $\mathbb{E}X$ cannot be very small, as it appears in the exponent in the right-hand side of (2-1). Since one usually wants to do so *simultaneously* for a large number of random variables X, one key situation to keep in mind is the following: if the number of random variables to be controlled is polynomial in N (a parameter), it is enough that their mean is at least a large multiple of $\log N$.

With minor modifications of the proof we can get a progressively denser set E avoiding all translates. We throw in the whole negative half-line (as we could have done in Theorem 1.3 too).

Theorem 2.4. *Let $\mathbb{A} = \{a_0 = 0 < a_1 < a_2 < \cdots\} \subseteq \mathbb{R}$ be a sequence with $a_n \to +\infty$. Then we can find a Lebesgue measurable set $E \subseteq \mathbb{R}$ such that no translate of $\mathbb{A}$,*

$$x + \mathbb{A}, \quad x \in \mathbb{R},$$

is contained in E, and such that

$$(-\infty, 0] \subseteq E \quad \text{and} \quad |E \cap [m, m+1]| \to 1^- \text{ as } m \to +\infty.$$

Proof. We indicate the differences with the proof of Theorem 2.1 and omit some details.

Our random set E now will be of the same type as in the proof of Theorem 2.1 but with the probability of including the small subintervals tending slowly to 1 as we go out to $+\infty$ and with the negative half-line contained in E to begin with.

Let us view the probability of keeping an interval as a function $p(s)$ defined on the real line. In the proof of Theorem 2.1 this function was constant. Here it will be constant on all intervals of the form $[m, m+1]$, $m \in \mathbb{Z}$.

With $\phi(x) = \mathbb{1}(x + \mathbb{A} \subseteq E)$, we need again to ensure that $\mathbb{E}\phi(x) = 0$ for all $x \in \mathbb{R}$. After assuming, as in the previous proof, that the points of $\mathbb{A}$ differ by at least 1, we again have independence of all events $x + a \in E$ for $a \in \mathbb{A}$ so that $\mathbb{E}\phi(x) = 0$ becomes equivalent to

$$\prod_{a\in\mathbb{A}} p(x+a) = 0,$$

which, writing $q(s) = 1 - p(s)$, is equivalent to

$$\sum_{a\in\mathbb{A}} q(x+a) = +\infty. \tag{2-3}$$

Let $0 = k_1 < k_2 < \cdots$ be those positive integers for which

$$[k, k+1) \cap \mathbb{A} \neq \varnothing.$$

Define then $q(x)$ to be $1/i$ in the interval $[k_i, k_{i+1})$, $i = 1, 2, \ldots$. It follows easily that, for all $x \in \mathbb{R}$, we have (2-3): since the function $q(\cdot)$ is decreasing we have $q(x + a_n) \geq q(a_n)$ if $x \leq 0$, and if $x \geq 0$ we have $q(x + a_n) \geq q(a_{\lceil x \rceil + n})$ since $a_{k+1} - a_k \geq 1$ for all $k \in \mathbb{N}$. In both cases the series (2-3) contains a tail of the series $\sum_{a \in \mathbb{A}} q(a)$, which is divergent.

It remains to ensure that the random variables $|[m, m+1] \setminus E|$ tend to 0 with $m \to +\infty$. These random variables are $1/N_m$ times a sum of independent indicator random variables (one for each of the N_m subintervals into which we break up $[m, m+1]$) of mean $q(m)N_m$, so we can use the Chernoff bound (2-1) to obtain

$$\mathbb{P}[|[m, m+1] \setminus E| > 2q(m)] \leq 2\exp(-c_1 q(m) N_m).$$

To ensure that the sum, over all $m \in \mathbb{Z}$, of the left-hand side is < 1 we can of course pick the integers N_m to be very large, say $N_m = K|m|/q(m)$, with a sufficiently large constant $K > 0$. □

3. No affine copies for slowly increasing sequences

In this section we prove Theorem 1.3 and explain why the proof also gives the more general Theorem 1.4.

Lemma 3.1. *Let $\mathbb{A} \in$ (A). For all $0 < a < b$, $0 \leq p < 1$ and $\epsilon > 0$, there is $N_0 \in \mathbb{N}$ such that, for all $N \geq N_0$, there is a set $E \subseteq [-N, N]$ such that*

(i) *for all $m \in \{-N, -N+1, \ldots, N-1\}$, we have $|E \cap [m, m+1]| \geq p$ and*

(ii) *if the set B consists of all $x \in [-N, N]$ for which there is $t \in [a, b]$ such that*

  (a) *$(x + t\mathbb{A}) \cap [-N, N] \subseteq E$ and*

  (b) *$\#((x + t\mathbb{A}) \cap [-N, N]) \geq A(N/(10b))$,*

then $|B| < \epsilon$. Here, $A(\cdot)$ is the counting function (1-1) of the set $\mathbb{A}$ and

$$A\Big(\frac{N}{10b}\Big) = \Big|\mathbb{A} \cap \Big[0, \frac{N}{(10b)}\Big]\Big|.$$

Let us first show how one derives Theorem 1.3 from Lemma 3.1. We give the proof of Theorem 1.3 in two steps: the first verifies the result for a restricted scale, that is, for scales in a compact interval, and the second concludes for all positive scales, by writing the whole scaling interval $(0, +\infty)$ as a countable union of intervals of the above type.

Step 1. *For all $0 < a < b$ and for each $0 \leq p < 1$, there exists a set $E \subseteq \mathbb{R}$ such that $|E \cap [m, m+1]| \geq p$ for all $m \in \mathbb{Z}$, but E does not contain any affine copies of $\mathbb{A}$ with scale in $[a, b]$.*

Consider $0 \leq p < 1$ and a positive increasing sequence $\{p_n\}$, $n = 1, 2, \ldots$ such that $p_n \to 1^-$ and, moreover,

$$\sum_{n=0}^{\infty} (1 - p_n) < 1 - p. \tag{3-1}$$

Take also any positive sequence $\epsilon_n \to 0$. According to Lemma 3.1, for $0 < a < b$, we can choose an increasing sequence of natural numbers $N_n = N_n(p_n, \epsilon_n, a, b) \to \infty$ for which there exist sets $E_n \subseteq [-N_n, N_n]$ with the following properties:

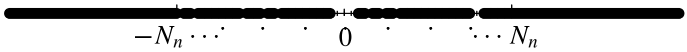

Figure 2. The set $\widetilde{E}_n$.

(i) for all $m = -N_n, \dots, N_n - 1$, we have $|E_n \cap [m, m+1]| \geq p_n$,

(ii) if

$$\mathbb{A}_n(x, t) = (x + t\mathbb{A}) \cap [-N_n, N_n]$$

and

$$B_n = \left\{ x \in [-N_n, N_n] : \exists t \in [a, b] \text{ s.t. } \mathbb{A}_n(x, t) \subseteq E_n \text{ and } \#\mathbb{A}_n(x, t) \geq A\left(\frac{N_n}{10b}\right) \right\},$$

then $|B_n| < \epsilon_n$.

Now take

$$\widetilde{E}_n = (-\infty, -N_n] \cup E_n \cup [N_n, +\infty)$$

and

$$E = \bigcap_{n=1}^{\infty} \widetilde{E}_n.$$

See Figure 2 for an illustration of the set $\widetilde{E}_n$.

Then, since $|\widetilde{E}_n \cap [m, m+1]| \geq p_n$ for all $m \in \mathbb{Z}$, we get from (3-1) that the set E has measure at least p at every unit interval with integer endpoints. Also, if there exist x and t such that $x + t\mathbb{A} \subseteq E$, then $x + t\mathbb{A}$ is also contained in each $\widetilde{E}_n$. Having fixed x and t we can then find n_0 large enough such that, for all $n \geq n_0$, we have $\#((x + t\mathbb{A}) \cap [-N_n, N_n]) \geq A(N_n/(10b))$. This implies that, for every $n \geq n_0$, we have $x \in B_n$. It follows that $|B_n| < \epsilon_n$ for every $n \geq n_0$. Since $\epsilon_n \to 0$, setting

$$B = \{x : \exists t \in [a, b] \text{ s.t. } x + t\mathbb{A} \subseteq E\},$$

we get $|B| = 0$. The null set of "bad" translates B is contained in E (since we assumed that $0 \in \mathbb{A}$), thus removing it from E results in a set E' which still has measure $|E' \cap [m, m+1]| \geq p$ for all $m \in \mathbb{Z}$ but contains no affine copy of $\mathbb{A}$ with scale in $[a, b]$.

Step 2. *Completion of the proof of Theorem 1.3.*

Take a positive sequence $p'_n \in [0, 1)$, $n \in \mathbb{Z}$, such that

$$\sum_{n \in \mathbb{Z}} (1 - p'_n) < 1 - p. \tag{3-2}$$

Consider the intervals $[a_n, b_n] = [2^{n-1}, 2^n]$, $n \in \mathbb{Z}$. Then, according to Step 1, for each p'_n, there exists a set E_n such that $|E_n \cap [m, m+1]| \geq p'_n$ for all $m \in \mathbb{Z}$, but, for all $x \in \mathbb{R}$ and for all $t \in [a_n, b_n]$, the set $x + t\mathbb{A}$ is not contained in E_n.

Take

$$E = \bigcap_{n \in \mathbb{Z}} E_n.$$

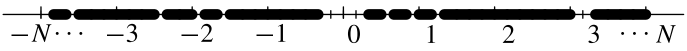

Figure 3. The new random set E.

Assume that, for some $x \in \mathbb{R}$ and some $t > 0$, we have $x + t\mathbb{A} \subseteq E$. Then, $x + t\mathbb{A} \subseteq E_n$ for all $n \in \mathbb{Z}$. However, since there is $n_0 \in \mathbb{Z}$ such that $t \in [2^{n_0-1}, 2^{n_0}]$, the inclusion $x + t\mathbb{A} \subseteq E_{n_0}$ cannot be true. Thus, E does not contain any affine copy of $\mathbb{A}$ with positive scale. Finally, due to (3-2), we have $|[m, m+1] \setminus E| < 1 - p$, or $|E \cap [m, m+1]| \geq p$.

3.1. ***Proof of Lemma 3.1.*** Fix the scale $t \in [a, b]$, and let $0 \leq p < 1$. Consider the positive sequence

$$p_N = 1 - \sqrt{\frac{\log\left(\frac{N}{10b}\right)}{A\left(\frac{N}{10b}\right)}}. \tag{3-3}$$

From (1-2) this implies $p_N \to 1^-$.

Partition $[-N, N]$ into unit intervals $[m, m+1]$, $m = -N, -N+1, \dots, N-1$. Divide each $[m, m+1]$ further into k_N equal subintervals

$$I_{i,m} = m + \left[\frac{i-1}{k_N}, \frac{i}{k_N}\right], \quad i = 1, \dots, k_N,$$

where

$$k_N = \left\lceil \frac{10}{a} \right\rceil \frac{N}{1 - p_N}. \tag{3-4}$$

Notice that $k_N/N \to +\infty$.

Construct a random set $E = E_N$ as follows: keep each $I_{i,m}$ in E independently of the other intervals and with probability p_N as in (3-3). Then, $\mathbb{P}(x \in E) = p_N$ for each $x \in [-N, N]$. See Figure 3 for an illustration of the new set E.

Let $M_N(x, t)$ be the number of elements of $(x + t\mathbb{A}) \cap [-N, N]$, and observe that

$$M_N(x, t) \leq A\left(\frac{2N}{a}\right) \quad \text{for } x \in [-N, N]. \tag{3-5}$$

For a given set $E \subseteq [-N, N]$, consider the set of "bad" translates

$$B = \left\{x \in [-N, N] : \exists t \in [a, b] \text{ s.t. } (x + t\mathbb{A}) \cap [-N, N] \subseteq E \text{ and } M_N(x, t) \geq A\left(\frac{N}{10b}\right)\right\}. \tag{3-6}$$

We first deal with the measure of B. We have

$$\begin{aligned} \mathbb{E}|B| &= \mathbb{E} \int_{-N}^{N} \mathbb{1}_B(x)\, dx \\ &= \int_{-N}^{N} \mathbb{P}\left[\exists t \in [a, b] : (x + t\mathbb{A}) \cap [-N, N] \subseteq E \text{ and } M_N(x, t) \geq A\left(\frac{N}{10b}\right)\right] dx. \end{aligned} \tag{3-7}$$

In what follows, we estimate from above the probability in (3-7), uniformly in $x \in [-N, N]$.

Fix $x \in [-N, N]$. To check whether there exists $t \in [a, b]$ such that $(x + t\mathbb{A}) \cap [-N, N] \subseteq E$, it is sufficient to check whether such a t exists in a finite set

$$S = S(x) = \{t_1, t_2, \dots, t_u\} \subseteq [a, b]. \tag{3-8}$$

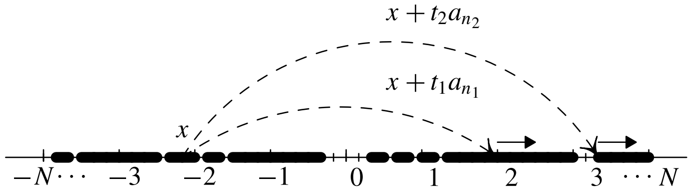

Figure 4. As x is held fixed and t grows, the points $x + ta_n$ cross over interval endpoints creating events that need to be checked.

Write $\alpha'_0 < \alpha'_1 < \cdots < \alpha'_{M_N(x,t)-1}$ for the elements of $(x + t\mathbb{A}) \cap [-N, N]$. Then, the set S consists exactly of those $t \in [a, b]$ for which some $\alpha'_j = x + ta_j$, $j = 0, \ldots, M_N(x, t) - 1$, is in the set

$$m + \left\{0, \frac{1}{k_N}, \frac{2}{k_N}, \ldots, \frac{k_N - 1}{k_N}, 1\right\}$$

for some $m \in \{-N, -N + 1, \ldots, N - 1\}$. Each of the points $\alpha'_j = x + ta_j$ traverses — as t moves from a to b and as long as the point α'_j remains in $[-N, N]$ — an interval of length at most $2N$, therefore it meets at most $2Nk_N$ interval endpoints of the intervals $I_{i,m}$. See Figure 4 for an illustration. Altogether, we have

$$u \le 2Nk_N \sup_{a \le t \le b} M_N(x, t) \le c(a)N^2(1 - p_N)^{-1} A\left(\frac{2N}{a}\right), \tag{3-9}$$

where, for the last inequality, we used (3-4) and (3-5).

Since $k_N \to +\infty$, we can take N large enough, say $N \ge N_0$, that $k_N > 1/a$ for every $N \ge N_0$. Then, the length of each $I_{i,m}$ is small enough, $\le a$, to ensure that, for each $t \in [a, b]$, the points α'_j, $j = 0, \ldots, M_N - 1$, all belong to different intervals $I_{i,m}$. Therefore, for any fixed x and t,

$$\begin{aligned}\mathbb{P}\left[(x + t\mathbb{A}) \cap [-N, N] \subseteq E \text{ and } M_N(x, t) \ge A\left(\frac{N}{10b}\right)\right] \\ \le \mathbb{P}\left[(x + t\mathbb{A}) \cap [-N, N] \subseteq E \,\middle|\, M_N(x, t) \ge A\left(\frac{N}{10b}\right)\right] \le p_N^{A(N/(10b))}.\end{aligned} \tag{3-10}$$

Thus, using the bound (3-9),

$$\mathbb{P}[\exists t \in S : (x + t\mathbb{A}) \cap [-N, N]) \subseteq E] \le c(a)N^2(1 - p_N)^{-1} A\left(\frac{2N}{a}\right) p_N^{A(N/(10b))}.$$

Thus, (3-7) yields

$$\mathbb{E}|B| \le 2c(a)N^3(1 - p_N)^{-1} A\left(\frac{2N}{a}\right) p_N^{A(N/(10b))}.$$

We want to have

$$N^3(1 - p_N)^{-1} A\left(\frac{2N}{a}\right) p_N^{A(N/(10b))} \to 0,$$

while $p_N \to 1^-$ as $N \to \infty$. Since $A(\cdot)$ grows at most linearly at infinity, it suffices to show that

$$A\left(\frac{N}{10b}\right) \log p_N \left(4\frac{\log N}{A\left(\frac{N}{10b}\right) \log p_N} - \frac{\log(1 - p_N)}{A\left(\frac{N}{10b}\right) \log p_N} + 1\right) \to -\infty. \tag{3-11}$$

To show (3-11), observe first that, since $\lim_{x\to+\infty} x\log(1-x^{-1/2}) = -\infty$, we have

$$\frac{A\left(\frac{N}{10b}\right)\log p_N}{\log N} \to -\infty \tag{3-12}$$

due to (3-3). Therefore, we also have $A(N/(10b))\log p_N \to -\infty$. Finally, by (3-3) and (3-12), we get

$$\frac{\log(1-p_N)}{A\left(\frac{N}{10b}\right)\log p_N} = -\frac{1}{2}\frac{\log A\left(\frac{N}{10b}\right)}{A\left(\frac{N}{10b}\right)\log p_N}\left\{1-\frac{\log\log\frac{N}{10b}}{\log A\left(\frac{N}{10b}\right)}\right\} \to 0.$$

In other words, we have shown that, for every $\epsilon > 0$, there is $N_1 \ge N_0$ such that, for all $N \ge N_1$, we have $\mathbb{E}|B| < \frac{1}{2}\epsilon$, which implies that

$$\mathbb{P}(|B| \ge \epsilon) < \tfrac{1}{2} \quad \text{for all } N \ge N_1. \tag{3-13}$$

We now turn to the measure of E in every unit interval with integer endpoints. Fix $m \in [-N, N]$. Let $X_1^m, X_2^m, \dots, X_{k_N}^m$ be independent indicator random variables, with $X_i^m = 1$ if and only if $I_{i,m} \subseteq E$. Let $Y_i^m = 1 - X_i^m$, and denote by $X^m = \sum_{i=1}^{k_N} X_i^m$ and $Y^m = \sum_{i=1}^{k_N} Y_i^m$ their sums. Then, $\mathbb{E}Y^m = (1-p_N)k_N$. Notice also that the total measure kept in $[m, m+1] \cap E$ is equal to X^m/k_N.

For any $\delta > 0$, we define the "bad" events

$$A_m = \{|Y^m - \mathbb{E}Y^m| > \delta\,\mathbb{E}Y^m\}, \quad m = -N, -N+1, \dots, N-1.$$

To control $\mathbb{P}[A_m]$, we use Chernoff's inequality [Alon and Spencer 1992; Chernoff 1952]: for all $\delta > 0$,

$$\mathbb{P}[A_m] \le 2e^{-c_\delta \mathbb{E}Y^m},$$

where $c_\delta = \min\{(1+\delta)\log(1+\delta) - \delta\log\delta,\ \frac{1}{2}\delta^2\}$. Take $\delta = \frac{1}{2}$. It follows that

$$\mathbb{P}\big[|Y^m - (1-p_N)k_N| > \tfrac{1}{2}(1-p_N)k_N\big] \le 2\exp\big(-\tfrac{1}{2}(1-p_N)k_N\big).$$

Thus, the probability that there is some $[m, m+1] \subseteq [-N, N]$ such that A_m holds is at most

$$4N\exp\big(-\tfrac{1}{2}(1-p_N)k_N\big),$$

and the right-hand side tends to zero as $N \to +\infty$ by our choice of k_N in (3-4). Thus, there is $N_2 \ge N_1$ such that

$$\mathbb{P}[\exists m \in \{-N, -N+1, \dots, N-1\} : A_m \text{ holds}] < \tfrac{1}{2} \tag{3-14}$$

for all $N \ge N_2$. Then, (3-13) and (3-14) imply the existence of a set $E \subseteq \mathbb{R}$ such that, on the one hand, it satisfies

$$|B| < \epsilon$$

and, on the other hand,

$$X^m - p_N k_N \ge -\tfrac{1}{2}(1-p_N)k_N$$

for all $m = -N, -N+1, \dots, N-1$, for all $N \ge N_2$. Thus the measure of E in each unit interval $[m, m+1]$ is at least $p_N - \frac{1}{2}(1-p_N) \to 1$ as $p_N \to 1^-$. In other words, for all $0 \le p < 1$, there is $N_3 \ge N_2$ such that, for all $N \ge N_3$, we have $|E \cap [m, m+1]| \ge p$. The proof of Lemma 3.1 is complete.

Remark 3.2. Let us indicate here why the proof of Theorem 1.3 just completed also applies to Theorem 1.4 without any essential changes. First of all, the implication from Lemma 3.1 to Theorem 1.5 (finite to infinite) remains true almost verbatim. So it suffices to ensure that Lemma 3.1 is true in this case. The main ingredients of the proof of Lemma 3.1 are the following. Having fixed x and varying t we have to make sure that the following conditions hold:

C.1 All points of the (x, t)-copy of the set remain well-separated, so that independence applies and we can multiply the probabilities that they belong to our random set. This is ensured by (1-4).

C.2 The number of points in the (x, t)-copy of the set in the interval $[-N, N]$ has to be large as this is the exponent in the upper bound (3-10). Condition (1-5) guarantees this.

C.3 The number of events that need to be checked so that we are certain that, for all t, no (x, t)-copy is contained in our random set is small. This is the number u in (3-8). What we are doing in the proof is to count how many times each of the points of our set (as x is held fixed and t increases from a to b) crosses over an interval boundary. Since the $\phi(n, t)$ are assumed increasing in t this remains as before.

It should be clear that the conditions imposed on the scaling functions $\phi(n, t)$ in Theorem 1.4 are far from optimal. They are rather indicative of what can be accomplished with the method, and it is clear that the method could work under different sorts of conditions.

4. The problem in higher dimension

We will derive Theorem 1.5 as a consequence of the more finitary theorem below.

Theorem 4.1. *Let $d_1, d \geq 1$, $\beta, \zeta > 0$, $p \in (0, 1)$. Let also $\alpha(N)$ be a function satisfying*

$$\frac{\alpha(N)}{\log N} \to +\infty.$$

Then, if N is sufficiently large and $P \subseteq \mathbb{R}^{d_1}$ is a point set with at most N^ζ points, there is a set $E_N \subseteq [-N, N]^d$ such that:

(1) *$|E_N \cap (m + [0, 1]^d)| \geq p$ for all $m = (m_1, \ldots, m_d) \in \mathbb{Z}^d$, with $-N \leq m_j < N$.*

(2) *For any linear map $T : \mathbb{R}^{d_1} \to \mathbb{R}^d$, if*

$$T(P) \cap [-N, N]^d \tag{4-1}$$

contains at least $\alpha(N)$ points with separation $\geq N^{-\beta}$, then

$$(T(P) \cap [-N, N]^d) \subsetneq E_N. \tag{4-2}$$

Proof. Let $\gamma > \beta$, and split the cube $[-N, N]^d$ with an $N^{-\gamma} \times \cdots \times N^{-\gamma}$-spaced grid of $O(dN^{1+\gamma})$ hyperplanes perpendicular to the d coordinate axes. Define the random set E to contain each of the $N^{-\gamma} \times \cdots \times N^{-\gamma}$-sized cubes independently with probability $p' \in (p, 1)$. We show that, with positive probability, one can take $E_N = E$.

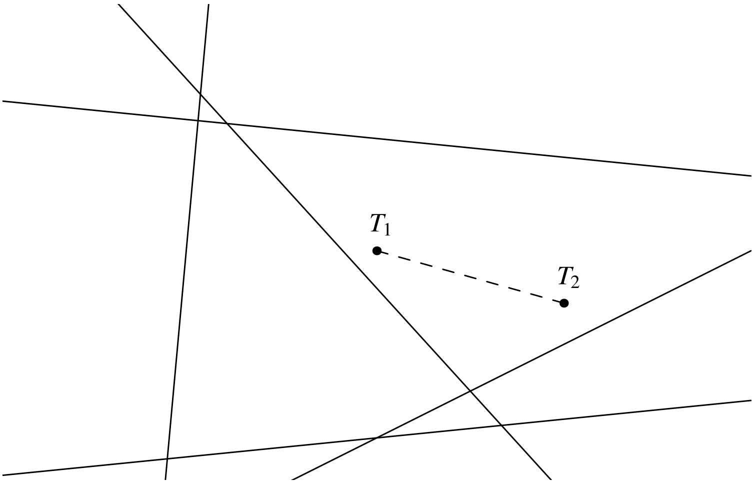

Figure 5. The regions defined in T-space by the equations $E(H, T(q))$ for all $H,\ q$. Only one of the transformations $T_1,\ T_2$ needs to be checked.

The first property of E is a simple consequence of Chernoff bounds and we can assume it holds with probability $> \frac{1}{2}$ working as in the proof of Theorem 1.3.

Let $T = (T_{i,j})$ be a linear map $\mathbb{R}^{d_1} \to \mathbb{R}^d$. This depends on $d \cdot d_1$ real variables $T_{i,j}$, so we view T as an element of $\mathbb{R}^{d \cdot d_1}$. Instead of checking condition (2) for all $T \in \mathbb{R}^{d \cdot d_1}$, we first show that there is a small number (polynomial in N) of transformations T that need to be checked.

Indeed, the set of $N^{-\gamma} \times \cdots \times N^{-\gamma}$-sized cubes that contain $T(P)$ does not change when T varies except when one or more of the points in $T(P)$ cross a dividing hyperplane of those that subdivide $[-N, N]^d$. Let H be one of those $O(dN^{1+\gamma})$ hyperplanes, and fix an arbitrary point $h \in H$. Let also u be a unit vector orthogonal to H. For a point $x \in \mathbb{R}^d$ to belong to H, it must satisfy the linear equation

$$E(H, x) : u \cdot x = u \cdot h.$$

Let $q \in P$. For the point $T(q)$ to belong to H, we must have

$$E(H, T(q)) : u \cdot T(q) = u \cdot h,$$

which is a linear equation in $T \in \mathbb{R}^{d \cdot d_1}$. Taking all such equations in T over all dividing hyperplanes H and all $q \in P$, we obtain a subdivision of $\mathbb{R}^{d \cdot d_1}$ by

$$n = O(d \cdot N^{1+\gamma} \cdot |P|)$$

hyperplanes. These n hyperplanes subdivide $\mathbb{R}^{d \cdot d_1}$ into $m = O(n^{d \cdot d_1})$ connected regions (this is easily proved by induction on the dimension or see [Buck 1943]). For any two points $T_1,\ T_2$ in the same region, condition (4-2) is either true for both or false for both since we can move continuously from T_1 to T_2 without leaving the region and, therefore, without any of the points $T(q)$ touching any of the dividing hyperplanes H. See Figure 5 for an illustration.

It suffices therefore to check condition (4-2) for one point per region. Let us call these points $T_1, \dots, T_m$. To guarantee that (4-2) holds for all T, it is enough for it to be true for all $T_j,\ j = 1, 2, \dots, m$. Define the bad events

$$B_j = \bigcap_{q \in P} \{T_j(q) \in E\}.$$

We need to ensure that none of the B_j holds, but we only need to check those B_j for which there is a T in the cell of T_j for which (4-1) holds. For such a j, the number of different $N^{-\gamma} \times \cdots \times N^{-\gamma}$-sized cubes touched by $T_j(P)$ is the same as the number touched by $T(P)$, which is at least $\alpha(N)$, so

$$\mathbb{P}[B_j] \le p'^{\alpha(N)},$$

and it is therefore enough to make sure that

$$n^{d \cdot d_1} p'^{\alpha(N)} = O(N^{\zeta \cdot d \cdot d_1} N^{(1+\gamma) d \cdot d_1} p'^{\alpha(N)})$$

can be made arbitrarily small by choosing N large. This is clearly possible since the term $p'^{\alpha(N)}$ decays faster than any power of N. □

Proof of Theorem 1.5. Let $p_n \in (0, 1)$ be such that

$$\sum_{n=1}^{\infty} (1 - p_n) < 1 - p. \tag{4-3}$$

Apply Theorem 4.1 successively for $N = n,\ p_n,\ \zeta = b,\ \alpha(N) = \alpha(R),\ \beta = f$ and the set $P = \mathbb{A} \cap [-n, n]^{d_1}$ to obtain sets $E_n \subseteq [-n, n]^d$. Define

$$E = \bigcap_{n=1}^{\infty} (E_n \cup (\mathbb{R}^d \setminus [-n, n]^d)).$$

It is easy to see because of (4-3) that, for any $m \in \mathbb{Z}^d$, we have $|E \cap m + [0, 1]^d| \ge p$. Let $T : \mathbb{R}^{d_1} \to \mathbb{R}^d$, and let R be such that $T(\mathbb{A}) \cap B_R(0)$ contains $\alpha(R)$ points which are R^{-f}-separated. Let $n = \lceil R \rceil$. It follows from Theorem 4.1 that $T(\mathbb{A}) \cap [-n, n]^d$ is not contained in $E_n \cup (\mathbb{R}^d \setminus [-n, n]^d)$ and therefore not contained in E, as we had to show. □

References

[Alon and Spencer 1992] N. Alon and J. H. Spencer, *The probabilistic method*, Wiley, New York, 1992. MR Zbl

[Bourgain 1986] J. Bourgain, "A Szemerédi type theorem for sets of positive density in $\mathbb{R}^k$", *Israel J. Math.* **54**:3 (1986), 307–316. MR Zbl

[Bradford et al. 2023] L. Bradford, H. Kohut, and Y. Mooroogen, "Large subsets of Euclidean space avoiding infinite arithmetic progressions", *Proc. Amer. Math. Soc.* **151**:8 (2023), 3535–3545. MR

[Buck 1943] R. C. Buck, "Partition of space", *Amer. Math. Monthly* **50**:9 (1943), 541–544. MR Zbl

[Burgin et al. 2023] A. Burgin, S. Goldberg, T. Keleti, C. MacMahon, and X. Wang, "Large sets avoiding infinite arithmetic/geometric progressions", *Real Anal. Exchange* **48**:2 (2023), 351–364.

[Chernoff 1952] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations", *Ann. Math. Statist.* **23** (1952), 493–507. MR Zbl

[Chlebik 2015] M. Chlebik, “On the Erdős similarity problem”, preprint, 2015. arXiv 1512.05607

[Croft et al. 1991] H. T. Croft, K. J. Falconer, and R. K. Guy, *Unsolved problems in geometry*, Unsolved Problems in Intuitive Math. **2**, Springer, 1991. MR Zbl

[Cruz et al. 2022] A. Cruz, C.-K. Lai, and M. Pramanik, “Large sets avoiding affine copies of infinite sequences”, preprint, 2022. arXiv 2204.12720

[Denson et al. 2021] J. Denson, M. Pramanik, and J. Zahl, “Large sets avoiding rough patterns”, pp. 59–75 in *Harmonic analysis and applications*, edited by M. T. Rassias, Springer Optim. Appl. **168**, Springer, 2021. MR Zbl

[Erdős 2015] P. Erdős, “My Scottish Book ‘problems”’, pp. 27–33 in *The Scottish Book*, 2nd ed., edited by R. D. Mauldin, Springer, 2015.

[Falconer 1984] K. J. Falconer, “On a problem of Erdős on sequences and measurable sets”, *Proc. Amer. Math. Soc.* **90**:1 (1984), 77–78. MR Zbl

[Falconer and Marstrand 1986] K. J. Falconer and J. M. Marstrand, “Plane sets with positive density at infinity contain all large distances”, *Bull. Lond. Math. Soc.* **18**:5 (1986), 471–474. MR Zbl

[Fraser and Pramanik 2018] R. Fraser and M. Pramanik, “Large sets avoiding patterns”, *Anal. PDE* **11**:5 (2018), 1083–1111. MR Zbl

[Fürstenberg et al. 1990] H. Fürstenberg, Y. Katznelson, and B. Weiss, “Ergodic theory and configurations in sets of positive density”, pp. 184–198 in *Mathematics of Ramsey theory*, edited by J. Nešetřil and V. Rödl, Algorithms Combin. **5**, Springer, 1990. MR Zbl

[Gallagher et al. 2023] J. Gallagher, C.-K. Lai, and E. Weber, “On a topological Erdős similarity problem”, *Bull. Lond. Math. Soc.* **55**:3 (2023), 1104–1119. MR Zbl

[Humke and Laczkovich 1998] P. D. Humke and M. Laczkovich, “A visit to the Erdős problem”, *Proc. Amer. Math. Soc.* **126**:3 (1998), 819–822. MR Zbl

[Keleti 2008] T. Keleti, “Construction of one-dimensional subsets of the reals not containing similar copies of given patterns”, *Anal. PDE* **1**:1 (2008), 29–33. MR Zbl

[Kolountzakis 1997] M. N. Kolountzakis, “Infinite patterns that can be avoided by measure”, *Bull. Lond. Math. Soc.* **29**:4 (1997), 415–424. MR Zbl

[Kolountzakis 2004] M. N. Kolountzakis, “Distance sets corresponding to convex bodies”, *Geom. Funct. Anal.* **14**:4 (2004), 734–744. MR Zbl

[Komjáth 1983] P. Komjáth, “Large sets not containing images of a given sequence”, *Canad. Math. Bull.* **26**:1 (1983), 41–43. MR Zbl

[Maga 2011] P. Maga, “Full dimensional sets without given patterns”, *Real Anal. Exchange* **36**:1 (2011), 79–90. MR Zbl

[Máthé 2017] A. Máthé, “Sets of large dimension not containing polynomial configurations”, *Adv. Math.* **316** (2017), 691–709. MR Zbl

[Shmerkin 2017] P. Shmerkin, “Salem sets with no arithmetic progressions”, *Int. Math. Res. Not.* **2017**:7 (2017), 1929–1941. MR Zbl

[Székely 1983] L. A. Székely, “Remarks on the chromatic number of geometric graphs”, pp. 312–315 in *Graphs and other combinatorial topics* (Prague, 1982), edited by M. Fiedler, Teubner-Texte Math. **59**, Teubner, Leipzig, 1983. MR Zbl

[Szemerédi 1975] E. Szemerédi, “On sets of integers containing no k elements in arithmetic progression”, *Acta Arith.* **27** (1975), 199–245. MR Zbl

[Yavicoli 2021] A. Yavicoli, “Large sets avoiding linear patterns”, *Proc. Amer. Math. Soc.* **149**:10 (2021), 4057–4066. MR Zbl

Received 4 Aug 2022. Revised 19 Aug 2023. Accepted 3 Oct 2023.

MIHAIL N. KOLOUNTZAKIS: kolount@gmail.com
Department of Mathematics and Applied Mathematics, University of Crete, Heraklion, Greece

EFFIE PAPAGEORGIOU: papageoeffie@gmail.com
Department of Mathematics and Applied Mathematics, University of Crete, Heraklion, Greece

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 18 (2025), No. 1, pp. 109–139
DOI: 10.2140/apde.2025.18.109

msp

THE 3D STRICT SEPARATION PROPERTY FOR THE NONLOCAL CAHN–HILLIARD EQUATION WITH SINGULAR POTENTIAL

ANDREA POIATTI

We consider the nonlocal Cahn–Hilliard equation with singular (logarithmic) potential and constant mobility in three-dimensional bounded domains and we establish the validity of the instantaneous strict separation property. This means that any weak solution, which is not a pure phase initially, stays uniformly away from the pure phases ± 1 from any positive time on. This work extends the result in dimension two for the same equation and gives a positive answer to the long-standing open problem of the validity of the strict separation property in dimensions higher than 2. In conclusion, we show how this property plays an essential role to achieve higher-order regularity for the solutions and to prove that any weak solution converges to a single equilibrium.

1. Introduction

The diffuse interface theory, also called the phase field method, is one of the oldest and most efficient approaches to multiphase problems. This approach is characterized by the notion of diffuse interface, meaning that the transition layer between the two phases or components has a narrow finite size. The interface is not explicitly tracked as in boundary integral and front-tracking methods. On the other hand, the phase state is incorporated into the macroscopic equations and the internal microstructures arise from the competition between the diffusion and aggregation mechanisms included in the free energy. The fundamental advantage of this theory is the natural representation of singular interfacial behaviors, such as topological change, self-intersection, merger and pinch-off.

Consider a mixture of two incompatible substances A and B, which is homogeneously distributed and isothermal. Under certain circumstances, namely if the temperature is above a critical threshold θ_c, this configuration is stable; however, if suddenly cooled down and kept at $\bar{\theta} < \theta_c$, the initially (macroscopically) homogeneous alloy evolves in a way such that A-rich and B-rich regions appear and grow. The Cahn–Hilliard equation was introduced in [Allen and Cahn 1979; Cahn and Hilliard 1958] to model this phenomenon in iron alloys, and it has now become a widespread model, since phase separation has become a paradigm also in cell biology (see, e.g., [Dolgin 2018]). Let Ω be a bounded domain in $\mathbb{R}^d$, $d = 2, 3$, filled with a binary solution consisting of A and B atoms, and let us fix a time horizon $T > 0$.

The present research has been supported by mur grant dipartimento di eccellenza 2023–2027.
MSC2020: 35B40, 35B65, 35Q82, 35R09.
Keywords: three-dimensional nonlocal Cahn–Hilliard equation, singular potential, strict separation property, regularization of weak solutions, convergence to equilibrium.

We define their relative mass fraction difference as ϕ, which is the phase-field variable, whose smooth but highly localized variation is associated with the (diffuse) interface. If the mixture is isothermal and the molar volume is uniform and independent on pressure, the system evolves in order to minimize the free energy functional

$$\mathcal{U}(\phi) := \int_\Omega \Big(\frac{\epsilon}{2}|\nabla\phi|^2 + \Psi(\phi)\Big)\,\mathrm{d}x, \tag{1-1}$$

where $\Psi(\phi)$ is the Helmholtz free energy density

$$\Psi(s) = \frac{\bar{\alpha}}{2}((1+s)\ln(1+s) + (1-s)\ln(1-s)) - \frac{\alpha_0}{2}s^2 = F(s) - \frac{\alpha_0}{2}s^2 \quad \text{for all } s \in [-1, 1], \tag{1-2}$$

with $\bar{\alpha}$ such that $0 < \bar{\alpha} < \alpha_0$, constants related to the temperature of the mixture. The term ϵ is called capillary coefficient, related to the thickness of interfaces. The potential defined in this way is called *singular*, whereas many authors (see, e.g., [Fife 2000]) considered a proper approximation, which avoids the fact that Ψ' is unbounded at the pure phases -1 and 1: namely, the significant potential is considered to be still a double-well, but with the two local minima coinciding with the pure phases. The most common choice is polynomial of even degree, like the case $\Psi(s) = \frac{1}{4}(s^2-1)^2$. However, in the case of polynomial potentials, it is worth recalling that it is not possible to guarantee the existence of physical solutions, that is, solutions for which $-1 \le \phi(x,t) \le 1$. Following, e.g., [Lowengrub and Truskinovsky 1998], we get a differential description of the phenomenon of the phase separation as

$$\partial_t\phi + \operatorname{div}\boldsymbol{J} = 0 \quad \text{in } \Omega \times (0,T), \tag{1-3}$$

where ϕ is the order parameter and $\boldsymbol{J}$ is the diffusional flux given by Fick's law,

$$\boldsymbol{J} = -M(\phi)\nabla\frac{\delta\mathcal{U}(\phi)}{\delta\phi} = -M(\phi)\nabla(-\epsilon\Delta\phi + \Psi'(\phi)),$$

where $\delta\mathcal{U}(\phi)/\delta\phi$ is the variational derivative of $\mathcal{U}(\phi)$. The function $M(\phi)$ is the mobility of the substances and in this work will be considered as a unitary constant (see, for instance, [Cherfils et al. 2011; Elliott and Garcke 1996] for an analysis of the case of nonconstant and degenerate mobility, i.e., vanishing at the pure phases). The Cahn–Hilliard equation with constant mobility then reads

$$\begin{cases} \partial_t\phi = \Delta\mu & \text{in } \Omega \times (0,T), \\ \mu = -\epsilon\Delta\phi + \Psi'(\phi) & \text{in } \Omega \times (0,T), \end{cases} \tag{1-4}$$

with the initial condition ϕ_0 and two boundary conditions which are generally the following:

$$\partial_{\boldsymbol{n}}\phi = 0, \qquad \partial_{\boldsymbol{n}}\mu = 0, \qquad \text{on } \partial\Omega \times (0,T), \tag{1-5}$$

with $\boldsymbol{n}$ as the outer normal vector. The former condition means that no mass flux occurs at the boundary, while the latter requires the interface to be orthogonal at the boundary.

It is worth noticing that the free energy $\mathcal{U}$ in (1-1) only focuses on short range interactions between particles. Indeed, the gradient square term accounts for the fact that the local interaction energy is spatially dependent and varies across the interfacial surface due to spatial inhomogeneities in the concentration. Going back to the general approach of statistical mechanics, the mutual short and long range interactions between particles is described through convolution integrals weighted by interactions kernels. Following this approach, Giacomin and Lebowitz [1996; 1997; 1998] observed that a physically more rigorous derivation leads to nonlocal dynamics, which is the nonlocal Cahn–Hilliard equation. In particular, this equation is rigorously justified as a macroscopic limit of microscopic phase segregation models with particles conserving dynamics. In this case, the gradient term is replaced by a nonlocal spatial interaction integral, namely, the energy is defined as

$$\mathcal{E}(\phi) := -\frac{1}{2}\int_\Omega\int_\Omega J(x-y)\phi(x)\phi(y)\,\mathrm{d}x\mathrm{d}y + \int_\Omega F(\phi(x))\,\mathrm{d}x, \tag{1-6}$$

where J is a sufficiently smooth symmetric interaction kernel. Note that this functional is characterized by a competition between the mixing entropy F and a nonlocal demixing term. As shown in [Giacomin and Lebowitz 1997] (see also [Gal et al. 2017; 2023a]), the energy $\mathcal{U}$ can be seen as an approximation of $\mathcal{E}$, as long as we suitably redefine F as $\widetilde{F}(x,s) = F(s) - \frac{1}{2}(J*1)(x)s^2$. In particular, we can rewrite $\mathcal{E}$ as

$$\begin{aligned}\mathcal{E}(\phi) &= \frac{1}{4}\int_\Omega\int_\Omega J(x-y)|\phi(y)-\phi(x)|^2\,\mathrm{d}x\mathrm{d}y + \int_\Omega\left(F(\phi(x)) - \frac{a(x)}{2}\phi^2(x)\right)\mathrm{d}x\\ &= \frac{1}{4}\int_\Omega\int_\Omega J(x-y)|\phi(y)-\phi(x)|^2\,\mathrm{d}x\mathrm{d}y + \int_\Omega \widetilde{F}(\phi(x))\,\mathrm{d}x,\end{aligned}$$

with $a(x) = (J*1)(x)$. If we formally interpret $\widetilde{F}$ as the potential Ψ of (1-1), we realize that the (formal) first approximation of the nonlocal interaction is $\frac{k}{2}|\nabla\phi|^2$, for some $k > 0$, as long as J is sufficiently peaked around 0. In the case $\Omega = \mathbb{T}^3$ (see, e.g., [Giacomin and Lebowitz 1998]), the term $J*1$ is a constant: thus $\mathcal{E}$ and $\mathcal{U}$ appear to be very similar. In particular, in this case, corresponding to set $a(x) = \alpha_0$, nonlocal-to-local asymptotics results have been obtained in [Davoli et al. 2021a; 2021b] (see also [Gal and Shomberg 2022]) for the nonlocal equation (1-7) below: namely, the solution to the nonlocal equation converges, under suitable conditions on the data of the problem, to the weak solution of (1-4)–(1-5).

The resulting nonlocal Cahn–Hilliard equation then reads (see [Gal et al. 2017; 2023a])

$$\begin{cases}\partial_t\phi - \Delta\mu = 0 & \text{in } \Omega\times(0,T),\\ \mu = F'(\phi) - J*\phi & \text{in } \Omega\times(0,T),\\ \partial_{\boldsymbol{n}}\mu = 0 & \text{on } \partial\Omega\times(0,T),\\ \phi(\cdot\,,0) = \phi_0 & \text{in } \Omega.\end{cases} \tag{1-7}$$

From now on we will refer to problem (1-4)–(1-5) as the local Cahn–Hilliard equation, in order to distinguish it from the nonlocal one in (1-7).

The well-posedness theory of Cahn–Hilliard equations with logarithmic (or singular) potential has been widely studied. The local Cahn–Hilliard equation (1-4)–(1-5) has been studied in [Abels and Wilke 2007; Debussche and Dettori 1995; Elliott and Luckhaus 1991; Giorgini et al. 2017; Londen and Petzeltová 2018; Miranville and Zelik 2004] (see also [Cherfils et al. 2011; Gal et al. 2023a] for a review and an insight analysis about this topic). Concerning the nonlocal Cahn–Hilliard equation, the physical relevance of nonlocal interactions was already pointed out in the pioneering paper [van der Waals 1982] (see also [Emmerich 2003, 4.2]) and studied for different kind of evolution equations, mainly Cahn–Hilliard and phase-field systems (see, e.g., [Bertozzi et al. 2007; Colli et al. 2007; Gajewski and Zacharias 2003; Gal and Grasselli 2014; Krejčí et al. 2007]). In particular, regarding the nonlocal system (1-7), the existence of weak solutions and their uniqueness, and the existence of the connected global attractor were proven in [Frigeri et al. 2016; Frigeri and Grasselli 2012a; 2012b]. Moreover, well-posedness and regularity of weak solutions are studied in [Gal et al. 2017], namely, in this work the authors establish the validity of the strict separation property in dimension two for the nonlocal Cahn–Hilliard equation (1-7) with constant mobility and singular potential. This means that if the initial state is not a pure phase (i.e., $\phi_0 \equiv 1$ or $\phi_0 \equiv -1$), then the corresponding solution stays away from the pure states in finite time, uniformly with respect to the initial datum. Exploiting this crucial property in dimension two, the authors derive straightforward consequences, such as further regularity results as well as the existence of regular finite-dimensional attractors and the convergence of a weak solution to a single equilibrium point. In [Gal et al. 2023a], the same authors propose an alternative argument to prove the strict separation property in dimension two, relying on a De Giorgi's iteration scheme (see [Gal et al. 2023a, Theorem 4.1]).

In the present work we extend the results of [Gal et al. 2023a] to the case of three-dimensional bounded domains, namely we prove the validity of the instantaneous strict separation property in dimension three for the system (1-7) with singular potential F. Our main result is the following: given a weak solution to (1-7),

$$\text{for all } \tau > 0 \text{ there exists } \delta > 0 \text{ such that } |\phi(x,t)| \le 1-\delta \quad \text{for a.e. } (x,t) \in \Omega \times (\tau, +\infty), \tag{1-8}$$

where δ depends on the parameters of the problem, the initial datum ϕ_0 and τ. Furthermore, we show that, if the initial datum ϕ_0 is more regular and already strictly separated from the pure phases, then (1-8) also holds with $\tau = 0$, i.e., the solution is uniformly strictly separated at almost any time $t \ge 0$. To assess the importance of property (1-8), similarly to [Gal et al. 2017], we infer some additional regularization results for any weak solution and we prove that each weak solution converges to a single stationary state.

As far as we are aware, this is the first time the instantaneous strict separation property is shown in three-dimensional bounded domains for the Cahn–Hilliard equation with constant mobility and singular (logarithmic) potential. Indeed, the only available result in dimension three regards the nonlocal Cahn–Hilliard equation with degenerate mobility and singular potential; see [Londen and Petzeltová 2011]. For the local Cahn–Hilliard equation the instantaneous separation property was first proven to hold in [Miranville and Zelik 2004], but only in dimension two. Concerning dimension three, only the asymptotic (i.e., from some positive time on, depending on the specific initial datum) separation property was proven in [Abels and Wilke 2007] for the local Cahn–Hilliard equation, but nothing is known about its instantaneous (i.e., from *any* positive time on) counterpart. The main issue which so far seemed to be hard to overcome in dimension three for both local and nonlocal cases is the use of the Trudinger–Moser inequality (see, e.g., [Nagai et al. 1997]), which, in dimension $d = 2, 3$, reads

$$\int_\Omega e^{|f(x)|}\,\mathrm{d}x \le C e^{C\|f\|^d_{W^{1,d}(\Omega)}} \quad \text{for all } f \in W^{1,d}(\Omega), \tag{1-9}$$

for some positive constant C independent of f, but depending on the dimension d and on the Lebesgue d-dimensional measure of Ω. In dimension two this inequality is easy to be handled, since it concerns only the $H^1(\Omega)$ norm of f. Indeed, if one assumes that

$$F''(s) \le Ce^{C|F'(s)|} \quad \text{for all } s \in (-1, 1), \tag{1-10}$$

for some constant $C > 0$ (see, e.g., [Gal et al. 2023a, (E2)] or [Gal et al. 2017]), which is satisfied by the logarithmic potential

$$F(s) = \frac{\bar{\alpha}}{2}\big((1+s)\ln(1+s) + (1-s)\ln(1-s)\big) \quad \text{for all } s \in [-1, 1], \tag{1-11}$$

then, exploiting (1-9) as done in [Gal et al. 2017] or adopting an argument as in [Gal et al. 2023a, Theorem 3.1], one can control the quantity $\|F''(\phi(t))\|_{L^p(\Omega)}$, for any $p \ge 2$, uniformly in time and this is the key tool to prove the validity of the separation property in two dimensions for example of the nonlocal Cahn–Hilliard equation with constant mobility and singular potential. In the case of three-dimensional bounded domains, (1-9) leads to the necessity of a control of the $W^{1,3}(\Omega)$ norm of f and this does not seem to be feasible in this context. Thus the proof proposed in [Gal et al. 2017] does not hold in dimension three. Moreover, also the alternative proof in [Gal et al. 2023a] to allow the control of $\|F''(\phi(t))\|_{L^p(\Omega)}$ is not viable in dimension three, due to the fact that the embedding $H^1(\Omega) \hookrightarrow L^q(\Omega)$ holds only for $q \in [2, 6]$, so that a result like [Gal et al. 2023a, (3.3)–(3.6)] cannot be obtained.

Here we are able to establish the (strict) separation property in three dimensions by avoiding the control of the quantity $F''(\phi(t))$ in any $L^p(\Omega)$ space. We do not assume condition (1-10) on F any more (see assumptions (H_2)–(H_3) and Remark 4.2 below), but we only rely on some natural growth conditions of F' and F'' near the endpoints ± 1. The idea is to perform a De Giorgi's iteration scheme on each interval of the form $(T - \tilde{\tau}, T)$, with $T > 0$ arbitrary and $\tilde{\tau}$ suitably chosen, similarly to the proof of [Gal et al. 2023a, Theorem 4.1], but modifying the argument in order to fully exploit the property that $F''(1-2\delta)^{-4} = O(\delta^4)$, for $\delta > 0$ sufficiently small (see (4-32)). This is possible in the estimates by treating in a suitable way all the terms leading to the presence of a quantity of the kind $F''(1-2\delta)^{-\gamma}$, with $0 \le \gamma < 4$ (see, e.g., the term Z_2 in the proof of [Gal et al. 2023a, Theorem 4.1]). To this aim, we first show the validity of a novel Poincaré-type inequality (Lemma 3.1), which is applied to a particular family of truncated functions obtained from the weak solution ϕ (namely, a family $\phi_\rho = (\phi - \rho)^+$, for some suitable $\rho \in (0, 1)$). This can be obtained heavily relying on the conservation of total mass (i.e.,

$$\int_\Omega \phi_0(x)\,\mathrm{d}x = \int_\Omega \phi(x,t)\,\mathrm{d}x$$

for any $t \ge 0$), that is one of the most important properties of the solution. By means of this Poincaré-type inequality, in the De Giorgi's scheme we get, at the end of the estimates, a term of the kind $F''(1-2\delta)^{-4}\delta^{-5} = O(\delta^{-1})$ and this, together with the use of the growth condition of F' near 1, permits to obtain the strict separation property by choosing a suitably small $\tilde{\tau}$ depending on δ. Since the size of δ and the related quantity $\tilde{\tau}$ do not depend on T, we repeat the same argument on each time interval $(T - \tilde{\tau}, T)$ for arbitrary $T > 0$, extending the result of the separation property on the entire interval $(\tau, +\infty)$, for $\tau > 0$ arbitrarily fixed at the beginning, completing in this way the proof of the validity of (1-8).

As future work, it is worth noticing that the strict separation property could pave the way for the study of other related problems with logarithmic potential in dimension three. For example, one could study the nonlocal Cahn–Hilliard–Oono equation (see, e.g., [Della Porta and Grasselli 2015]), the nonlocal Cahn–Hilliard–Hele–Shaw system (see, e.g., [Della Porta et al. 2018]) as well as other hydrodynamic phase-field models for binary fluid mixtures of incompressible viscous fluids (see also Remark 4.7).

The paper is organized as follows. In Section 2 we introduce the functional setting. Section 3 is devoted to the presentation some preliminaries, which are essential in the proofs, in particular the new Poincaré-type inequality. In the same section we also recall some already-known results concerning well-posedness of the nonlocal Cahn–Hilliard equation and we present a Lemma on geometric convergence of numerical sequences, which is a key tool for De Giorgi's type arguments. Section 4 contains the main result concerning the strict separation property in dimension three for the system (1-7), together with its proof. In conclusion, in Section 5 we present some consequences of the validity of the strict separation property, namely we show some regularization results and we prove that any weak solution to (1-7) converges to a single equilibrium.

2. Mathematical setting

Let Ω be a smooth bounded domain in $\mathbb{R}^3$. The Sobolev spaces are denoted as usual by $W^{k,p}(\Omega)$, where $k \in \mathbb{N}$ and $1 \le p \le \infty$, with norm $\|\cdot\|_{W^{k,p}(\Omega)}$. The Hilbert space $W^{k,2}(\Omega)$ is denoted by $H^k(\Omega)$ with norm $\|\cdot\|_{H^k(\Omega)}$. In particular, we will adopt the notation

$$H = L^2(\Omega), \quad V = H^1(\Omega), \quad V_2 = \{v \in H^2(\Omega) : \partial_{\boldsymbol{n}} v = 0 \text{ on } \partial\Omega\}.$$

Moreover, given a space X, we denote by $\boldsymbol{X}$ the space of vectors of three components, each one belonging to X. We then denote by $(\cdot\,,\cdot)$ the inner product in H and by $\|\cdot\|$ the induced norm. We indicate by $(\cdot\,,\cdot)_V$ and $\|\cdot\|_V$ the canonical inner product and its induced norm in V, respectively. We also define the integral mean of a function f as

$$\bar{f} := \frac{\int_\Omega f(x)\,\mathrm{d}x}{|\Omega|},$$

where $|\Omega|$ stands for the three-dimensional Lebesgue measure of the set Ω. We then introduce

$$H_0 = \{v \in H : \bar{f} = 0\}, \quad V_0 = \{v \in V : \bar{f} = 0\}, \quad V_0' = \left\{v \in V' : \frac{\langle f, 1\rangle}{|\Omega|} = 0\right\},$$

endowed with the norms of H, V and V'. Thanks to the Poincaré-Wirtinger inequality, it follows that $(\|\nabla u\|^2_{L^2(\Omega)} + |\bar{u}|^2)^{1/2}$ is a norm on V equivalent to $\|u\|_V$. The Laplace operator $A_0 : V_0 \to V_0'$ defined by $\langle A_0 u, v\rangle = (\nabla u, \nabla v)$ is an isomorphism. We denote by $\mathcal{N}$ its inverse map and we set $\|f\|_* := \|\nabla \mathcal{N} f\|$, which is a norm on V_0' equivalent to the canonical one. Moreover, we recall that

$$\|f - \bar{f}\|_*^2 + |\bar{f}|^2 \tag{2-1}$$

is a norm V' which is equivalent to the standard one. Next, we recall the following Gagliardo–Nirenberg inequality (see, e.g., [Brezis 2011, Chapter 9]):

$$\|u\|_{L^p(\Omega)} \le C(p)\|u\|^{\frac{6-p}{2p}} \|u\|_V^{\frac{3(p-2)}{2p}} \quad \text{for all } u \in V \text{ and } p \in [2,6], \tag{2-2}$$

where the constant $C(p)$ depends on Ω and p. From this inequality, in the case $p = \frac{10}{3}$ we get

$$\|u\|_{L^{10/3}(\Omega)} \le \widehat{C}\|u\|^{\frac{2}{5}} \|u\|_V^{\frac{3}{5}} \quad \text{for all } u \in V, \tag{2-3}$$

with $\widehat{C} > 0$ depending on Ω.

3. Preliminaries

Here we present some preliminary results, which are essential for the proof of our main theorem.

3.1. ***A Poincaré-type inequality.*** First we state the following generalized version of the well known Poincaré's inequality:

Lemma 3.1. *Let I be either a compact interval or an interval of the kind $[\tau, +\infty)$, with $\tau > 0$. Let $\mathcal{K} \subset \mathbb{R}$ be a set of indices and $\{f_\rho\}_{\rho\in\mathcal{K}} \subset L^\infty(I; V) \cap C(I; H)$. Assume also that, for any $\rho \in \mathcal{K}$ and for any $t \in I$, $f_\rho(t) \equiv 0$ on the set $E(t) := \{x \in \Omega : g(t,x) \le 1 - 2\delta\} \subset \Omega$, with $g \in C(I; L^q(\Omega))$, $q \ge 1$, and $\delta \in \left(0, \frac{1}{2}\right)$. Moreover, for a fixed $\varepsilon > 0$ sufficiently small, assume that for any $t \in I$ the set $\{x \in \Omega : g(t,x) \le 1 - 2\delta - \varepsilon\} \subset E(t)$ has strictly positive Lebesgue measure. In the case the interval I is $[\tau, +\infty)$, assume additionally that for any sequence $\{t_l\}_l$, such that $t_l \to \infty$ as $l \to \infty$, there exists a (nonrelabeled) subsequence $\{t_l\}_l$, a function $g^\star \in L^r(\Omega)$, $r \ge 1$, and $\tilde{\varepsilon} > 0$, such that $g(t_l) \to g^\star$ strongly in $L^r(\Omega)$ as $l \to \infty$ and the set $\{x \in \Omega : g^\star(x) \le 1 - 2\delta - \tilde{\varepsilon}\}$ has strictly positive Lebesgue measure.*

Then there exists a uniform (in ρ and t) constant $C_P > 0$ such that

$$\|f_\rho(t)\| \le C_P \|\nabla f_\rho(t)\| \quad \textit{for all } t \in I \textit{ and } \rho \in \mathcal{K}. \tag{3-1}$$

Remark 3.2. Since $\{f_\rho\}_\rho \subset C(I; H) \cap L^\infty(I; V) \hookrightarrow C_w(I; V)$, where $C_w(I; V)$ denotes the V-valued weakly continuous functions (see, e.g., [Boyer and Fabrie 2013, Lemma II.5.9]), it makes sense to ask for conditions at *any* time $t \in I$.

Proof. Due to $\{f_\rho\}_\rho \subset C_w(I; V)$, $f_\rho(t) \in V$ for any $\rho \in \mathcal{K}$ and any $t \in I$. Assume by contradiction that (3-1) is false. Then there exist a sequence $\{\rho_n\}_{n\in\mathbb{N}} \subset \mathcal{K}$ and a sequence $\{t_n\}_{n\in\mathbb{N}} \subset I$ such that

$$\|f_{\rho_n}(t_n)\| > n\|\nabla f_{\rho_n}(t_n)\| \quad \text{for all } n \in \mathbb{N}.$$

We then set

$$w_n := \frac{f_{\rho_n}(t_n)}{\|f_{\rho_n}(t_n)\|}, \qquad \text{with } \|w_n\| = 1.$$

We need to consider two cases:

(1) Either the interval I is compact or there exists a nonrelabeled subsequence of $\{t_n\}_n$ which is entirely contained in the set $[\tau, M] \subset I$, for some $M < +\infty$. In this case there exists another nonrelabeled subsequence of times and $t^\star \in I$, with $t^\star < +\infty$, such that $t_n \to t^\star$.

Now notice that, since $g \in C(I; L^q(\Omega))$, $q \geq 1$, we get $g(t_n) \to g(t^\star)$ in $L^q(\Omega)$. Therefore, there exists a subsequence $\{g(t_{n_j})\}_j$ such that, as $j \to \infty$,

$$g(t_{n_j}) \to g(t^\star) \quad \text{a.e. in } \Omega.$$

Let us now set $D := \{x \in \Omega : g(t^\star, x) \leq 1 - 2\delta - \varepsilon\}$, and

$$\alpha = |D| > 0,$$

which is possible by assumption. Then by the Severini–Egorov theorem (notice that Ω has finite measure, so this theorem can be applied), there exists a measurable subset $B \subset \Omega$ such that $|B| < \frac{\alpha}{2}$ and such that, as $j \to \infty$,

$$g(t_{n_j}) \to g(t^\star) \quad \text{uniformly on } \Omega \setminus B.$$

Therefore, we also deduce that $|D \setminus B| > \frac{\alpha}{2} > 0$ and that also

$$g(t_{n_j}) \to g(t^\star) \quad \text{uniformly on } D \setminus B.$$

This means that there exists a $\bar{J} \in \mathbb{N}$ such that, for any $x \in D \setminus B$,

$$|g(t_{n_j}, x) - g(t^\star, x)| < \varepsilon \quad \text{for all } j \geq \bar{J},$$

implying that, for any $x \in D \setminus B$, by definition of the set D,

$$g(t_{n_j}, x) = g(t_{n_j}, x) - g(t^\star, x) + g(t^\star, x) \leq \varepsilon + 1 - 2\delta - \varepsilon = 1 - 2\delta \quad \text{for all } j \geq \bar{J}.$$

This means, by the assumptions, that

$$D \setminus B \subset E(t_{n_j}) \subset \{x \in \Omega : w_{n_j}(x) = 0\} \quad \text{for all } j \geq \bar{J},$$

implying

$$D \setminus B \subset \bigcap_{j \geq \bar{J}} \{x \in \Omega : w_{n_j}(x) = 0\}, \quad |D \setminus B| > \frac{\alpha}{2}.$$

(2) The interval I is of the form $[\tau, +\infty)$ and there are no bounded subsequences of $\{t_n\}_n$, i.e., $t_n \to +\infty$ as $n \to \infty$. In this case we have by assumption that, up to a nonrelabeled subsequence, there exists $g^\star \in L^r(\Omega)$, $r \geq 1$, such that $g(t_n) \to g^\star$ strongly in $L^r(\Omega)$. Thus there exists a subsequence $\{g(t_{n_j})\}_j$ such that

$$g(t_{n_j}) \to g^\star \quad \text{a.e. in } \Omega.$$

As in case (1), we set $D := \{x \in \Omega : g^\star(x) \leq 1 - 2\delta - \tilde{\varepsilon}\}$, and

$$\alpha = |D| > 0,$$

which is again possible by assumption. Then we can repeat exactly the same arguments as in case (1) to obtain again that

$$D \setminus B \subset E(t_{n_j}) \subset \{x \in \Omega : \; w_{n_j}(x) = 0\} \quad \text{for all } j \geq \bar{J},$$

implying

$$D \setminus B \subset \bigcap_{j \geq \bar{J}} \{x \in \Omega : \; w_{n_j}(x) = 0\}, \quad |D \setminus B| > \frac{\alpha}{2}.$$

Clearly notice that in this case the set B will be such that there exists a $\bar{J} \in \mathbb{N}$ such that, for any $x \in D \setminus B$,

$$|g(t_{n_j}, x) - g^{\star}(x)| < \tilde{\varepsilon} \quad \text{for all } j \geq \bar{J}.$$

In both cases (1) and (2), since w_{n_j} is uniformly bounded in V, there exists $w \in V$ such that, by the Rellich–Kondrachov theorem, as $j \to \infty$,

$$w_{n_j} \rightharpoonup w \quad \text{in } V, \qquad w_{n_j} \to w \quad \text{in } H, \qquad \nabla w_{n_j} \rightharpoonup \nabla w \quad \text{in } H,$$

up to a nonrelabeled subsequence. Moreover, since $\|\nabla w_{n_j}\| < 1/n_j$, we deduce, by weak lower sequential semicontinuity of the L^2-norm, that $\nabla w \equiv 0$ almost everywhere in Ω and thus, being Ω connected, $w \equiv \kappa$ almost everywhere in Ω, with κ constant. Therefore, since also, up to another subsequence, $w_{n_j} \to w$ almost everywhere in Ω, we have $w \equiv 0$ on $D \setminus B$ (of positive Lebesgue measure) up to a zero measure set. But this clearly implies that $\kappa = 0$, which is a contradiction, since $\|w\| = 1$ (because $\|w_{n_j}\| = 1$ and $w_{n_j} \to w$ in H as $j \to \infty$). This concludes the proof. □

3.2. ***The state of the art for the three-dimensional nonlocal Cahn–Hilliard equation.*** For the sake of completeness we state here the already-known results concerning the nonlocal Cahn–Hilliard equation with constant mobility and singular potential in three-dimensional bounded domains. We first consider the following assumptions:

(H_1) $J \in W^{1,1}_{\text{loc}}(\mathbb{R}^3)$, with $J(x) = J(-x)$.

(H_2) $F \in C([-1, 1]) \cap C^2(-1, 1)$ fulfills

$$\lim_{s \to -1} F'(s) = -\infty, \quad \lim_{s \to 1} F'(s) = +\infty, \quad F''(s) \geq \alpha > 0 \quad \text{for all } s \in (-1, 1).$$

We extend $F(s) = +\infty$ for any $s \notin [-1, 1]$. Without loss of generality, $F(0) = 0$ and $F'(0) = 0$. In particular, this entails that $F(s) \geq 0$ for any $s \in [-1, 1]$. Also, we assume that there exists $\gamma \in (0, 1)$ such that F'' is nondecreasing in $[1 - \gamma, 1)$ and nonincreasing in $(-1, -1 + \gamma]$.

Theorem 3.3. *Assume that (H_1)–(H_2) hold and also that $\phi_0 \in L^\infty(\Omega)$ such that $\|\phi_0\|_{L^\infty} \leq 1$ and $|\bar{\phi}_0| = m < 1$. Then there exists a unique weak solution to (1-7) such that, for any $T > 0$,*

$$\phi \in L^\infty(\Omega \times (0, T)) : \quad \textit{for all } t > 0, \; |\phi(t)| < 1, \quad \textit{a.e. in } \Omega,$$
$$\phi \in L^2(0, T; V) \cap H^1(0, T; H),$$
$$\mu \in L^2(0, T; V), \quad F'(\phi) \in L^2(0, T; V),$$

such that

$$\langle \partial_t \phi, v\rangle + (\nabla\mu, \nabla v) = 0 \quad \text{for all } v \in V, \quad \text{a.e. in } (0, T), \tag{3-2}$$

$$\mu = F'(\phi) - J * \phi \quad \text{a.e. in } \Omega \times (0, T), \tag{3-3}$$

and $\phi(\cdot\,, 0) = \phi_0(\cdot)$ *in* Ω. *The weak solution also satisfies the energy identity* ($\mathcal{E}$ *is defined in* (1-6))

$$\mathcal{E}(\phi(t)) + \int_s^t \|\nabla\mu(\tau)\|^2 \,\mathrm{d}\tau = \mathcal{E}(\phi(s)) \quad \text{for all } 0 \le s \le t < \infty. \tag{3-4}$$

Moreover, *for any* $\tau > 0$,

$$\sup_{t\ge\tau} \|\partial_t\phi(t)\|_{V'} + \sup_{t\ge\tau} \|\partial_t\phi\|_{L^2(t,t+1,H)} \le \frac{K_0}{\sqrt{\tau}}, \tag{3-5}$$

$$\sup_{t\ge\tau} \|\mu(t)\|_V + \sup_{t\ge\tau} \|\phi(t)\|_V \le \frac{K_0}{\sqrt{\tau}}, \tag{3-6}$$

$$\|F'(\phi)\|_{L^\infty(\tau,t;V)} + \|\mu\|_{L^2(t,t+1,V_2)} \le K_1 \quad \text{for all } t \ge \tau, \tag{3-7}$$

$$\|\nabla\mu\|_{L^q(t,t+1;L^p(\Omega))} + \|\nabla\phi\|_{L^q(t,t+1;L^p(\Omega))} \le K_2 \quad \text{if } \frac{3p-6}{2p} = \frac{2}{q} \quad \text{for all } p \in [2,6] \text{ and } t \ge \tau, \tag{3-8}$$

where the positive constant K_0 *depends only on the initial datum energy* $\mathcal{E}(\phi_0)$, $\bar{\phi}_0$, Ω *and the parameters of the system*, *whereas* $K_1 = K_1(\tau)$ *and* $K_2 = K_2(\tau)$ *also depend on* τ. *Furthermore* K_2 *depends on also* q, p. *In conclusion*, *there holds the following continuous dependence estimate*: *for every two weak solutions* ϕ_1 *and* ϕ_2 *to* (1-7) *on* $[0, T]$, *with initial data* ϕ_{01} *and* ϕ_{02}, *respectively*, *we have*, *for all* $t \in [0, T]$,

$$\|\phi_1(t) - \phi_2(t)\|_{V'}^2 \le \|\phi_{01} - \phi_{02}\|_{V'}^2 + K|\bar{\phi}_{01} - \bar{\phi}_{02}|e^{CT},$$

where C *is a positive constant and*

$$K = C(\|F'(\phi_1)\|_{L^1(0,T;L^1(\Omega))} + \|F'(\phi_2)\|_{L^1(0,T;L^1(\Omega))}).$$

Remark 3.4. The proof of the above theorem can be found in [Gal et al. 2017, Theorems 3.4, 4.1, Proposition 4.2] and [Della Porta et al. 2018, Proposition 3.1]; see also [Gal et al. 2023b, Theorem 4.1] and [Poiatti and Signori 2024, Theorem 2.2] for a comprehensive result in the more general case of an advective nonlocal Cahn–Hilliard equation in two and three dimensions, respectively. In particular, we refer to [Gal et al. 2023b, Theorem 4.1, (4.4)] and [Della Porta et al. 2018, Proposition 3.1, (3.53)], which still hold in the nonadvective case $\boldsymbol{u} = \mathbf{0}$, for the validity of the energy identity (3-4), whereas (3-5) is shown in [Gal et al. 2017, Theorem 4.1, (4.2)]. Estimates (3-6)–(3-8) can be found in Theorem 4.1, (4.3), Proposition 4.2, (4.7), and Proposition 4.2, (4.9) of [Gal et al. 2017], respectively.

Remark 3.5. If we assume additionally that $\nabla F'(\phi_0) \in \boldsymbol{H}$ we can actually extend (3-5)–(3-8) to $\tau = 0$, since the initial datum is more regular and one can argue as in [Della Porta et al. 2018, Section 4] to obtain the desired regularity departing from the initial time. This means that the solution ϕ with initial datum ϕ_0 is indeed a strong solution to problem (1-7).

Remark 3.6. Notice that from condition (3-7) we can also deduce by Sobolev embeddings that

$$\|F'(\phi)\|_{L^\infty(\tau,\infty;L^p(\Omega))} \leq K_3(\tau, p) \quad \text{for all } p \in [1, 6], \tag{3-9}$$

where $K_3(\tau, p)$ depends on K_1, Ω and p.

Remark 3.7. We highlight that the previous theorem and our following main result concerning the strict separation property in dimension three heavily rely on the assumption $\bar{\phi}_0 \in (-1, 1)$ (see also [Kenmochi et al. 1995] for the local Cahn–Hilliard equation). This is physically reasonable since $\bar{\phi}_0 = 1$ (or $\bar{\phi}_0 = -1$) means that the initial condition is a pure phase, so that no phase separation takes place in Ω, unless we assume the existence of a source or reaction term (see, for instance [Grasselli et al. 2023]).

3.3. ***A lemma on geometric convergence of sequences.*** We present here one of the key tools for the application of De Giorgi's iteration argument. This lemma can be found, e.g., in [DiBenedetto 1993, Chapter I, Lemma 4.1], [Ladyženskaja et al. 1968, Chapter 2, Lemma 5.6], and it has also been proposed in [Gal et al. 2023a, Lemma 4.3].

Lemma 3.8. *Let $\{y_n\}_{n\in\mathbb{N}\cup\{0\}} \subset \mathbb{R}^+$ satisfy the recursive inequalities*

$$y_{n+1} \leq C b^n y_n^{1+\varepsilon} \quad \textit{for all } n \geq 0, \tag{3-10}$$

for some $C > 0$, $b > 1$ and $\varepsilon > 0$. If

$$y_0 \leq \theta := C^{-\frac{1}{\varepsilon}} b^{-\frac{1}{\varepsilon^2}}, \tag{3-11}$$

then

$$y_n \leq \theta b^{-\frac{n}{\varepsilon}} \quad \textit{for all } n \geq 0, \tag{3-12}$$

and consequently $y_n \to 0$ for $n \to \infty$.

Proof. The proof can be easily carried out directly by induction. Indeed, the case $n = 0$ is trivial. Then assume that (3-12) holds for n. We prove that it also holds for $n + 1$. In particular we have by (3-10) and recalling (3-11),

$$y_{n+1} \leq C b^n y_n^{1+\varepsilon} \leq C b^n \theta^{1+\varepsilon} b^{-\frac{n}{\varepsilon}(1+\varepsilon)} = C\theta^{1+\varepsilon} b^{-\frac{n}{\varepsilon}} = \theta b^{-\frac{n+1}{\varepsilon}} C\theta^\varepsilon b^{\frac{1}{\varepsilon}} \leq \theta b^{-\frac{n+1}{\varepsilon}},$$

where we exploited the definition of θ in (3-11). This means that (3-12) also holds for $n + 1$, concluding the proof by induction. □

We now present our main results, concerning the instantaneous strict separation property in three-dimensional bounded domains.

4. Main results

Let us assume, additionally to (H_2), the following hypotheses on the singular potential F:

(H_3) As $\delta \to 0^+$ we assume

$$\frac{1}{F'(1-2\delta)} = O\left(\frac{1}{|\ln(\delta)|}\right), \quad \frac{1}{F''(1-2\delta)} = O(\delta), \tag{4-1}$$

and analogously

$$\frac{1}{|F'(-1+2\delta)|}=O\left(\frac{1}{|\ln(\delta)|}\right),\quad \frac{1}{F''(-1+2\delta)}=O(\delta). \tag{4-2}$$

Remark 4.1. Notice that these conditions are verified by the logarithmic potential (1-11). Indeed,

$$F'(s)=\frac{\bar{\alpha}}{2}\ln\left(\frac{1+s}{1-s}\right),\quad F''(s)=\frac{\bar{\alpha}}{1-s^2};$$

thus

$$F'(1-2\delta)=\frac{\bar{\alpha}}{2}\ln\left(\frac{1-\delta}{\delta}\right),\qquad F''(1-2\delta)=\frac{\bar{\alpha}}{4\delta(1-\delta)},$$

$$F'(-1+2\delta)=\frac{\bar{\alpha}}{2}\ln\left(\frac{\delta}{1-\delta}\right),\quad F''(-1+2\delta)=\frac{\bar{\alpha}}{4\delta(1-\delta)},$$

clearly implying assumption (H_3).

Remark 4.2. As already pointed out in the Introduction, assumption (H_3) does not make any explicit reference to the typical extra condition (1-10). Indeed, as far as we know, this is the first proof of the instantaneous separation property concerning nonlocal Cahn–Hilliard equation with constant mobility and singular potential (problem (1-7)) in which it is not exploited any constraint on $\|F''(\phi(t))\|_{L^q(\Omega)}$, for some $q\geq 2$ and for almost any $t\geq\tau$, with $\tau>0$. Indeed, in our proof we simply rely on some natural growth conditions of F' and F'' near the endpoints ± 1. Note that assumptions (H_2)–(H_3) on the potential F are somehow minimal, in the sense that the proof of the separation property in dimension three works only in this case (or for more singular potentials than the logarithmic one). This seems to suggest that the use of the logarithmic potential when modeling phase separation phenomena with the help of the nonlocal Cahn–Hilliard equation with constant mobility could be a good choice, since it preserves all the basic physical properties expected from the solution.

We can now state our main theorem.

Theorem 4.3. *Let $\Omega\subset\mathbb{R}^3$ be a smooth bounded domain and let assumptions (H_1)–(H_3) hold. Assume that $\phi_0\in L^\infty(\Omega)$ such that $\|\phi_0\|_{L^\infty}\leq 1$ and $|\overline{\phi}_0|=m<1$. Then for any $\tau>0$ there exists $\delta\in(0,1)$, depending on τ, m and the initial datum, such that the unique weak solution to problem (1-7) given in Theorem 3.3 satisfies*

$$|\phi(x,t)|\leq 1-\delta\quad \textit{for a.e. } (x,t)\in\Omega\times(\tau,+\infty),$$

i.e., the instantaneous strict separation property from the pure phases ± 1 holds.

Remark 4.4. Observe that the quantity δ given in the theorem strongly depends on the specific entire trajectory, therefore, by the uniqueness of the solution, on the initial datum ϕ_0. This means that we cannot have an explicit dependence of δ, e.g., on the initial datum energy.

As a byproduct of the main theorem, we also prove that, if the initial datum ϕ_0 is more regular and already separated from the pure phases, i.e., there exists $\delta_0 \in (0, 1]$ such that

$$\|\phi_0\|_{L^\infty(\Omega)} \le 1 - \delta_0,$$

then the unique solution ϕ departing from ϕ_0, which is now strong from the time $t = 0$ (see Remark 3.5), is strictly separated on $[0, +\infty)$, i.e., it remains separated from the pure phases *uniformly* for almost any time $t \ge 0$.

Corollary 4.5. *Under the same hypotheses of Theorem 4.3, if we assume additionally that $\nabla F'(\phi_0) \in \boldsymbol{H}$, and that ϕ_0 is strictly separated, i.e., there exists $\delta_0 \in (0, 1]$ such that*

$$\|\phi_0\|_{L^\infty(\Omega)} \le 1 - \delta_0,$$

then there exists $\delta \in (0, 1)$, depending on τ, m, δ_0 and the initial datum, such that the unique strong solution to problem (1-7) given in Remark 3.5 satisfies

$$|\phi(x, t)| \le 1 - \delta, \quad \text{for a.e. } (x, t) \in \Omega \times [0, +\infty),$$

i.e., the instantaneous strict separation property from the pure phases ± 1 holds for almost any time $t \ge 0$.

Remark 4.6. Observe that, since by Theorem 4.3 the solution ϕ in Corollary 4.5 is strictly separated on time sets of the kind $(\tau, +\infty)$, for any $\tau > 0$, it is enough to show that there exists an interval $[0, T_1]$ ($T_1 > 0$) on which the solution is separated to obtain the strict separation over $[0, +\infty)$, choosing $\tau = T_1$. As it will be clear from the proof of Corollary 4.5, T_1 can be explicitly computed as a function of the parameters of the problem and the initial datum.

4.1. ***Proof of Theorem 4.3.*** We divide the proof into two steps. In the first we show that we can apply Lemma 3.1 to a specific family of functions, which will be of essential importance in the second step, when we adopt a De Giorgi's iteration scheme (as in [Gal et al. 2023a, Theorem 4.1]) to obtain the desired result.

Step 1. Application of Lemma 3.1 to a family of truncated functions. Let us consider the unique solution ϕ departing from ϕ_0, whose existence and regularity is stated in Theorem 3.3. We make the following observations: first fix any $\tau > 0$.

- Since $|\overline{\phi}_0| \le m < 1$, there exists $\hat{\delta} > 0$ and an $\varepsilon > 0$ such that

$$m \le 1 - 2\hat{\delta} - \varepsilon. \tag{4-3}$$

In particular we may choose $\varepsilon := (1 - m)/2 > 0$ and $\hat{\delta} := (1 - m)/4 > 0$. Thanks to the conservation of total mass, we have that for any $\rho \in \mathbb{R}^+$, $\rho \ge 1 - 2\hat{\delta}$, and for any $t \in [0, +\infty)$, the function

$$\phi_\rho(x, t) := (\phi(x, t) - \rho)^+ \tag{4-4}$$

vanishes on the set (independent of ρ)

$$E(t) := \{x \in \Omega : \phi(x, t) \le 1 - 2\hat{\delta}\}, \tag{4-5}$$

which is such that

$$|\{x \in \Omega : \phi(x, t) \le 1 - 2\hat{\delta} - \varepsilon\}| > 0 \quad \text{for all } t \ge 0. \tag{4-6}$$

Proof. To prove this observation, let us assume by contradiction that, for some $\tilde{t} \geq 0$,

$$|\{x \in \Omega : \phi(x, \tilde{t}) \leq 1 - 2\hat{\delta} - \varepsilon\}| = 0.$$

By the conservation of total mass we get, for any $t \geq 0$,

$$(1 - 2\hat{\delta} - \varepsilon)|\Omega| \geq m|\Omega| \geq \int_\Omega \phi_0(x)\,\mathrm{d}x = \int_\Omega \phi(x, t)\,\mathrm{d}x,$$

but then we get a contradiction, since $|\Omega| = |\{x \in \Omega : \phi(x, \tilde{t}) > 1 - 2\hat{\delta} - \varepsilon\}|$ and

$$(1 - 2\hat{\delta} - \varepsilon)|\Omega| \geq \int_\Omega \phi(x, \tilde{t})\,\mathrm{d}x > (1 - 2\hat{\delta} - \varepsilon)|\{x \in \Omega : \phi(x, \tilde{t}) > 1 - 2\hat{\delta} - \varepsilon\}|. \qquad \square$$

- We aim to apply Lemma 3.1 with $\mathcal{K} = [1 - 2\hat{\delta}, 1]$, $\{f_\rho\}_{\rho \in \mathcal{K}} = \{\phi_\rho\}_{\rho \in \mathcal{K}}$, $I = [\tau, +\infty)$, $g = \phi$, $\delta = \hat{\delta}$, $\tilde{\varepsilon} = \varepsilon$. Indeed we verify all the assumptions:

- We have $\{\phi_\rho\}_\rho \subset L^\infty(I; V) \cap C(I; H)$, $\phi \in C(I; H)$, and (4-5) and (4-6) hold for any $t \in I$.

- Let $\{t_l\}_l$ be any sequence such that $t_l \to \infty$. By (3-6), there exists a constant $C(\tau) > 0$ such that

$$\sup_{t \geq \tau} \|\phi\|_V \leq C(\tau).$$

Therefore, since V is reflexive, there exist a (nonrelabeled) subsequence $\{t_l\}_l$ and a function $g^\star \in V$ (which could depend on the subsequence) such that, as $l \to \infty$,

$$\phi(t_l) \rightharpoonup g^\star \quad \text{in } V,$$

implying by compactness that

$$\phi(t_l) \to g^\star \quad \text{in } H. \tag{4-7}$$

Now notice that this strong convergence also implies, by the conservation of total mass, that

$$\int_\Omega \phi_0(x)\,\mathrm{d}x = \int_\Omega \phi(x, t_l)\,\mathrm{d}x \to \int_\Omega g^\star(x)\,\mathrm{d}x,$$

and thus also $g^\star$ enjoys the same total mass as the initial datum ϕ_0:

$$\int_\Omega g^\star(x)\,\mathrm{d}x = \int_\Omega \phi_0(x)\,\mathrm{d}x.$$

This means that we can repeat exactly the same argument as the one adopted to get (4-6) to infer

$$|\{x \in \Omega : g^\star(x) \leq 1 - 2\hat{\delta} - \varepsilon\}| > 0, \tag{4-8}$$

so that, having chosen $\tilde{\varepsilon} = \varepsilon$ and $g = \phi$, thanks to (4-7)–(4-8), we have completed the verification of the assumptions of Lemma 3.1.

In the end we can conclude that there exists a uniform (in ρ and t) constant $C_{P,+} > 0$ such that

$$\|\phi_\rho(t)\| \leq C_{P,+}\|\nabla\phi_\rho(t)\| \tag{4-9}$$

for any $t \in [\tau, +\infty)$ and any $\rho \in [1 - 2\hat{\delta}, 1]$.

- Since in the last part of the proof we need to reproduce all the arguments on the functions

$$\tilde{\phi}_\rho(x,t) := (\phi(x,t)+\rho)^- = (-\phi(x,t)-\rho)^+, \tag{4-10}$$

with $\rho \geq 1-2\hat{\delta}$, we observe that (4-5) and (4-6) still hold substituting ϕ with $-\phi$, simply because, again by the conservation of mass, $m|\Omega| \geq \int_\Omega -\phi(x,t)\,dx$ for any $t \geq \tau$. Therefore again the assumptions of Lemma 3.1 are satisfied (with $g=-\phi$), and thus that there exists a uniform (in ρ and t) constant $C_{P,-} > 0$ (which is possibly different from $C_{P,+}$) such that

$$\|\tilde{\phi}_\rho(t)\| \leq C_{P,-}\|\nabla\tilde{\phi}_\rho(t)\| \tag{4-11}$$

for any $t \in [\tau,+\infty)$ and for any $\rho \in [1-2\hat{\delta},1]$. Thus we introduce the constant $C_P := \max\{C_{P,+}, C_{P,-}\}$ so that both (4-9) and (4-11) hold with the same constant C_P, i.e.,

$$\|\phi_\rho(t)\| \leq C_P\|\nabla\phi_\rho(t)\|, \quad \|\tilde{\phi}_\rho(t)\| \leq C_P\|\nabla\tilde{\phi}_\rho(t)\|, \tag{4-12}$$

for any $t \geq \tau$ and any $\rho \in [1-2\hat{\delta},1]$. Note that the constant C_P depends on the specific solution ϕ we used, thus, since ϕ is uniquely determined by ϕ_0, we have that C_P depends in a nontrivial way on the initial datum.

Step 2. De Giorgi's iteration scheme. We perform a De Giorgi's iteration scheme following the one presented in [Gal et al. 2023a, Lemma 4.1]. Let us fix δ sufficiently small such that $\delta \leq \hat{\delta}$, so that (4-12) holds for any $\rho \in [1-2\delta,1]$. Set then $\tilde{\tau} > 0$ such that

$$\tilde{\tau} = \frac{2^{-20}\delta^5(F''(1-2\delta))^4F'(1-2\delta)}{3C(\tau)\|\nabla J\|^5_{L^1(B_r)}\widehat{C}^{\frac{3}{2}}(1+C_P^2)^{\frac{3}{2}}}, \tag{4-13}$$

where C_P is given in (4-12), $\widehat{C}$ is defined in (2-3) and B_r is a ball centered at $\mathbf{0}$ of radius $r>0$ sufficiently large such that $x-\Omega \subset B_r$ for any $x \in \Omega$ (see also [Giorgini 2024] for this observation on B_r). Now observe that, since, by (4-1), there exists a positive constant $C_F > 0$ such that, for δ sufficiently small,

$$0 < \frac{1}{F''(1-2\delta)} \leq C_F\delta \quad \text{and} \quad 0 < \frac{1}{F'(1-2\delta)} \leq \frac{C_F}{|\ln(\delta)|},$$

we have

$$\begin{aligned}\frac{\frac{8\delta^2}{\tilde{\tau}}}{\frac{\|\nabla J\|^2_{L^1(B_r)}}{2F''(1-2\delta)}} &= \frac{16\delta^2F''(1-2\delta)}{\|\nabla J\|^2_{L^1(B_r)}}\frac{3C(\tau)\|\nabla J\|^5_{L^1(B_r)}\widehat{C}^{\frac{3}{2}}(1+C_P^2)^{\frac{3}{2}}}{2^{-20}\delta^5(F''(1-2\delta))^4F'(1-2\delta)}\\ &= \frac{3C(\tau)\|\nabla J\|^3_{L^1(B_r)}\widehat{C}^{\frac{3}{2}}(1+C_P^2)^{\frac{3}{2}}}{2^{-24}\delta^3(F''(1-2\delta))^3F'(1-2\delta)} \leq \frac{\widetilde{C}}{|\ln(\delta)|} \to 0 \quad \text{as } \delta \to 0^+,\end{aligned}$$

where

$$\widetilde{C} := \frac{3C(\tau)\|\nabla J\|^3_{L^1(B_r)}\widehat{C}^{\frac{3}{2}}(1+C_P^2)^{\frac{3}{2}}C_F^4}{2^{-24}} > 0,$$

so that

$$\frac{\frac{8\delta^2}{\tilde{\tau}}}{\frac{\|\nabla J\|^2_{L^1(B_r)}}{2F''(1-2\delta)}} = O\left(\frac{1}{|\ln(\delta)|}\right).$$

This means that we can find a sufficiently small $\delta > 0$ so that

$$\max\left\{\frac{\|\nabla J\|^2_{L^1(B_r)}}{2F''(1-2\delta)}, \frac{8\delta^2}{\tilde{\tau}}\right\} = \frac{\|\nabla J\|^2_{L^1(B_r)}}{2F''(1-2\delta)}. \tag{4-14}$$

Choose now $T > 0$ such that $T - 3\tilde{\tau} \geq \frac{\tau}{2}$ (for example, one can start with $T = 3\tilde{\tau} + \frac{\tau}{2}$). Up to reducing the size of δ, and thus of $\tilde{\tau}$, we can find $\tilde{\tau}$ such that

$$2\tilde{\tau} + \frac{\tau}{2} \leq \tau. \tag{4-15}$$

Let us then fix $\delta > 0$ (and thus $\tilde{\tau} > 0$) so that also (4-14) and (4-15) hold. Notice that the choice of δ and $\tilde{\tau}$ *does not* depend on the specific T, but clearly depends on τ.

We now define the sequence

$$k_n = 1 - \delta - \frac{\delta}{2^n} \quad \text{for all } n \geq 0, \tag{4-16}$$

where

$$1 - 2\delta < k_n < k_{n+1} < 1 - \delta \quad \text{for all } n \geq 1, \quad k_n \to 1 - \delta \quad \text{as } n \to \infty, \tag{4-17}$$

and the sequence of times

$$t_n = \begin{cases} T - 3\tilde{\tau} & \text{if } n = -1, \\ t_{n-1} + \dfrac{\tilde{\tau}}{2^n} & \text{if } n \geq 0, \end{cases} \tag{4-18}$$

which satisfies

$$t_{-1} < t_n < t_{n+1} < T - \tilde{\tau} \quad \text{for all } n \geq 0.$$

We now introduce a cutoff function $\eta_n \in C^1(\mathbb{R})$ by setting

$$\eta_n(t) := \begin{cases} 0 & \text{if } t \leq t_{n-1}, \\ 1 & \text{if } t \geq t_n \end{cases} \quad \text{and} \quad |\eta_n'(t)| \leq \frac{2^{n+1}}{\tilde{\tau}}, \tag{4-19}$$

on account of the above definition of the sequence $\{t_n\}_n$. Recalling (4-4), we then set $\rho = k_n$,

$$\phi_n(x,t) := (\phi - k_n)^+, \tag{4-20}$$

and, for any $n \geq 0$, we introduce the interval $I_n = [t_{n-1}, T]$ and the set

$$A_n(t) := \{x \in \Omega : \phi(x,t) - k_n \geq 0\} \quad \text{for all } t \in I_n.$$

Clearly, we have

$$\begin{aligned} I_{n+1} &\subseteq I_n && \text{for all } n \geq 0, \\ A_{n+1}(t) &\subseteq A_n(t) && \text{for all } n \geq 0 \text{ and } t \in I_{n+1}. \end{aligned}$$

In conclusion, we set

$$y_n = \int_{I_n}\int_{A_n(s)} 1 \,\mathrm{d}x\,\mathrm{d}s \quad \text{for all } n \geq 0.$$

Now, for any $n \geq 0$, we consider the test function $v = \phi_n \eta_n^2$, and integrate over $[t_{n-1}, t]$, $t_n \leq t \leq T$. Then

$$\int_{t_{n-1}}^{t} \langle \partial_t \phi, \phi_n \eta_n^2 \rangle \mathrm{d}s + \int_{t_{n-1}}^{t}\int_{A_n(s)} F''(\phi)\nabla\phi \cdot \nabla\phi_n \eta_n^2 \,\mathrm{d}x\,\mathrm{d}s = \int_{t_{n-1}}^{t}\int_{A_n(s)} \eta_n^2 (\nabla J * \phi)\cdot\nabla\phi_n \,\mathrm{d}x\,\mathrm{d}s, \tag{4-21}$$

since $\nabla F'(\phi(t)) = F''(\phi)\nabla\phi(t)$, for almost every $x \in \Omega$ and for any $t \geq \frac{\tau}{2}$, which can be proven, e.g., by a truncation argument as in [He and Wu 2021, Lemma 3.2], applied for any $t \geq \frac{\tau}{2}$. Indeed, as in [He and Wu 2021, (3.5)], we obtain $\nabla F'(\phi(t)) = F''(\phi)\nabla\phi(t)$ in the sense of distribution and thus, since $\nabla F'(\phi) \in L^\infty\big(\frac{\tau}{2}, \infty; \boldsymbol{H}\big)$, we immediately infer that the equality holds also almost everywhere in Ω, for any $t \geq \frac{\tau}{2}$. Now, as in [Gal et al. 2023a], for δ sufficiently small we obtain

$$\int_{t_{n-1}}^{t} \eta_n^2 \int_{A_n(s)} F''(\phi)\nabla\phi\cdot\nabla\phi_n \,\mathrm{d}x\,\mathrm{d}s \geq F''(1-2\delta)\int_{t_{n-1}}^{t} \eta_n^2 \|\nabla\phi_n\|^2 \,\mathrm{d}s, \tag{4-22}$$

and, for the right-hand side of (4-21), recalling that $|\phi| < 1$ a.e. in $\Omega \times (0, +\infty)$, we find

$$\begin{aligned}
&\int_{t_{n-1}}^{t}\int_{A_n(s)} (\nabla J * \phi)\cdot\nabla\phi_n \eta_n^2 \,\mathrm{d}x\,\mathrm{d}s \\
&\quad\leq \frac{1}{2}F''(1-2\delta)\int_{t_{n-1}}^{t} \eta_n^2\|\nabla\phi_n\|^2\,\mathrm{d}s + \frac{1}{2F''(1-2\delta)}\int_{t_{n-1}}^{t}\int_{A_n(s)} \eta_n^2 |\nabla J * \phi|^2 \,\mathrm{d}x\,\mathrm{d}s \\
&\quad\leq \frac{1}{2}F''(1-2\delta)\int_{t_{n-1}}^{t} \eta_n^2\|\nabla\phi_n\|^2\,\mathrm{d}s + \frac{1}{2F''(1-2\delta)}\int_{t_{n-1}}^{t} \|\nabla J * \phi\|^2_{L^\infty(\Omega)}\int_{A_n(s)} \mathrm{d}x\,\mathrm{d}s \\
&\quad\leq \frac{1}{2}F''(1-2\delta)\int_{t_{n-1}}^{t} \eta_n^2\|\nabla\phi_n\|^2\,\mathrm{d}s + \frac{\|\nabla J\|^2_{L^1(B_r)}}{2F''(1-2\delta)}\int_{t_{n-1}}^{t}\int_{A_n(s)} \mathrm{d}x\,\mathrm{d}s \\
&\quad\leq \frac{1}{2}F''(1-2\delta)\int_{t_{n-1}}^{t} \eta_n^2\|\nabla\phi_n\|^2\,\mathrm{d}s + \frac{\|\nabla J\|^2_{L^1(B_r)}}{2F''(1-2\delta)} y_n,
\end{aligned} \tag{4-23}$$

where we have applied (see, e.g., [Brezis 2011, Theorem 4.33])

$$\|\nabla J * \phi\|_{L^\infty(\Omega)} \leq \|\nabla J\|_{L^1(B_r)}\|\phi\|_{L^\infty(\Omega)} \leq \|\nabla J\|_{L^1(B_r)}. \tag{4-24}$$

Moreover, we have

$$\int_{t_{n-1}}^{t} \langle \partial_t\phi, \phi_n\eta_n^2\rangle \,\mathrm{d}s = \frac{1}{2}\|\phi_n(t)\|^2 - \int_{t_{n-1}}^{t} \|\phi_n(s)\|^2 \eta_n \partial_t \eta_n \,\mathrm{d}s. \tag{4-25}$$

Note that, since $|\phi| < 1$ a.e. in Ω, for any $t \geq \frac{\tau}{2}$,

$$0 \leq \phi_n \leq 2\delta \quad \text{a.e. in } \Omega \ \text{ for all } t \geq \frac{\tau}{2}. \tag{4-26}$$

Then, by the above inequality,

$$\int_{t_{n-1}}^{t} \|\phi_n(s)\|^2 \eta_n \partial_t \eta_n \,\mathrm{d}s = \int_{t_{n-1}}^{t} \int_{\Omega} \phi_n^2(s) \eta_n \partial_t \eta_n \,\mathrm{d}x\,\mathrm{d}s = \int_{t_{n-1}}^{t} \int_{A_n(s)} \phi_n^2(s) \eta_n \partial_t \eta_n \,\mathrm{d}x\,\mathrm{d}s \\ \leq \int_{t_{n-1}}^{t} \int_{A_n(s)} (2\delta)^2 \frac{2^{n+1}}{\tilde{\tau}} \,\mathrm{d}x\,\mathrm{d}s \leq \frac{2^{n+3}\delta^2}{\tilde{\tau}} y_n. \tag{4-27}$$

Plugging (4-22), (4-23), (4-25) and (4-27) into (4-21), we find

$$\frac{1}{2}\|\phi_n(t)\|^2 + \frac{1}{2} F''(1-2\delta) \int_{t_{n-1}}^{t} \eta_n^2 \|\nabla \phi_n(s)\|^2 \,\mathrm{d}s \leq 2^{n+1} \max\left\{ \frac{\|\nabla J\|_{L^1(B_r)}^2}{2F''(1-2\delta)}, \frac{8\delta^2}{\tilde{\tau}} \right\} y_n$$

for any $t \in [t_n, T]$. Thanks to the choice of δ and $\tilde{\tau}$, we recall (4-14), implying

$$\max_{t\in I_{n+1}} \|\phi_n(t)\|^2 \leq X_n, \quad F''(1-2\delta) \int_{I_{n+1}} \|\nabla \phi_n\|^2 \,\mathrm{d}s \leq X_n, \tag{4-28}$$

where

$$X_n := 2^{n+1} \frac{\|\nabla J\|_{L^1(B_r)}^2}{F''(1-2\delta)} y_n.$$

On the other hand, for any $t \in I_{n+1}$ and for almost any $x \in A_{n+1}(t)$, we get

$$\phi_n(x,t) = \phi(x,t) - \left[1-\delta-\frac{\delta}{2^n}\right] \\ = \underbrace{\phi(x,t) - \left[1-\delta-\frac{\delta}{2^{n+1}}\right]}_{\phi_{n+1}(x,t)\geq 0} + \delta\left[\frac{1}{2^n} - \frac{1}{2^{n+1}}\right] \geq \frac{\delta}{2^{n+1}},$$

which implies

$$\int_{I_{n+1}} \int_{\Omega} |\phi_n|^3 \,\mathrm{d}x\,\mathrm{d}s \geq \int_{I_{n+1}} \int_{A_{n+1}(s)} |\phi_n|^3 \,\mathrm{d}x\,\mathrm{d}s \geq \left(\frac{\delta}{2^{n+1}}\right)^3 \int_{I_{n+1}} \int_{A_{n+1}(s)} \mathrm{d}x\,\mathrm{d}s = \left(\frac{\delta}{2^{n+1}}\right)^3 y_{n+1}.$$

Then we have

$$\left(\frac{\delta}{2^{n+1}}\right)^3 y_{n+1} \leq \int_{I_{n+1}} \int_{\Omega} |\phi_n|^3 \,\mathrm{d}x\,\mathrm{d}s = \int_{I_{n+1}} \int_{A_n(s)} |\phi_n|^3 \,\mathrm{d}x\,\mathrm{d}s \\ \leq \left(\int_{I_{n+1}} \int_{\Omega} |\phi_n|^{\frac{10}{3}} \,\mathrm{d}x\,\mathrm{d}s\right)^{\frac{9}{10}} \left(\int_{I_{n+1}} \int_{A_n(s)} \mathrm{d}x\,\mathrm{d}s\right)^{\frac{1}{10}}. \tag{4-29}$$

Notice that, thanks to (2-3) and (4-12) (which holds thanks to (4-17)), we get

$$\int_{I_{n+1}} \int_{\Omega} |\phi_n|^{\frac{10}{3}} \,\mathrm{d}x\,\mathrm{d}s \leq \widehat{C} \int_{I_{n+1}} \|\phi_n\|_V^2 \|\phi_n\|^{\frac{4}{3}} \,\mathrm{d}s \leq \widehat{C} \int_{I_{n+1}} \left(\|\phi_n\|^2 + \|\nabla\phi_n\|^2\right) \|\phi_n\|^{\frac{4}{3}} \,\mathrm{d}s \\ \leq \widehat{C}(1+C_P^2) \int_{I_{n+1}} \|\nabla \phi_n\|^2 \|\phi_n\|^{\frac{4}{3}} \,\mathrm{d}s,$$

where we have chosen an equivalent norm on V. Observe now that, by (4-28),

$$\begin{aligned}\int_{I_{n+1}}\int_\Omega |\phi_n|^{\frac{10}{3}}\,\mathrm{d}x\,\mathrm{d}s &\le \widehat{C}(1+C_P^2)\int_{I_{n+1}}\|\nabla\phi_n\|^2\|\phi_n\|^{\frac{4}{3}}\,\mathrm{d}s\\ &\le \widehat{C}(1+C_P^2)\max_{t\in I_{n+1}}\|\phi_n(t)\|^{\frac{4}{3}}\int_{I_{n+1}}\|\nabla\phi_n\|^2\,\mathrm{d}s\\ &\le \frac{\widehat{C}(1+C_P^2)}{F''(1-2\delta)}X_n^{\frac{2}{3}}F''(1-2\delta)\int_{I_{n+1}}\|\nabla\phi_n\|^2\,\mathrm{d}s\\ &\le \frac{\widehat{C}(1+C_P^2)}{F''(1-2\delta)}X_n^{\frac{5}{3}} \le \frac{2^{\frac{5n}{3}+\frac{5}{3}}\|\nabla J\|_{L^1(B_r)}^{\frac{10}{3}}\widehat{C}(1+C_P^2)}{(F''(1-2\delta))^{\frac{8}{3}}}y_n^{\frac{5}{3}}.\end{aligned}$$

Coming back to (4-29), we immediately infer

$$\begin{aligned}\left(\frac{\delta}{2^{n+1}}\right)^3 y_{n+1} &\le \left(\int_{I_{n+1}}\int_\Omega|\phi_n|^{\frac{10}{3}}\,\mathrm{d}x\,\mathrm{d}s\right)^{\frac{9}{10}}\left(\int_{I_{n+1}}\int_{A_n(s)}\mathrm{d}x\,\mathrm{d}s\right)^{\frac{1}{10}}\\ &\le \frac{2^{\frac{3}{2}n+\frac{3}{2}}\|\nabla J\|_{L^1(B_r)}^3\widehat{C}^{\frac{9}{10}}(1+C_P^2)^{\frac{9}{10}}}{(F''(1-2\delta))^{\frac{12}{5}}}y_n^{\frac{3}{2}}y_n^{\frac{1}{10}}\\ &= \frac{2^{\frac{3}{2}n+\frac{3}{2}}\|\nabla J\|_{L^1(B_r)}^3\widehat{C}^{\frac{9}{10}}(1+C_P^2)^{\frac{9}{10}}}{(F''(1-2\delta))^{\frac{12}{5}}}y_n^{\frac{8}{5}}.\end{aligned}\tag{4-30}$$

In conclusion, we end up with

$$y_{n+1} \le \frac{2^{\frac{9}{2}n+\frac{9}{2}}\|\nabla J\|_{L^1(B_r)}^3\widehat{C}^{\frac{9}{10}}(1+C_P^2)^{\frac{9}{10}}}{\delta^3(F''(1-2\delta))^{\frac{12}{5}}}y_n^{\frac{8}{5}} \quad \text{for all } n\ge 0.\tag{4-31}$$

Thus we can apply Lemma 3.8. In particular, we have

$$b = 2^{\frac{9}{2}} > 1,\qquad C = \frac{2^{\frac{9}{2}}\|\nabla J\|_{L^1(B_r)}^3\widehat{C}^{\frac{9}{10}}(1+C_P^2)^{\frac{9}{10}}}{\delta^3(F''(1-2\delta))^{\frac{12}{5}}} > 0,\qquad \varepsilon = \frac{3}{5},$$

to get that $y_n \to 0$, as long as

$$y_0 \le C^{-\frac{5}{3}}b^{-\frac{25}{9}},$$

that is,

$$y_0 \le \frac{2^{-20}\delta^5(F''(1-2\delta))^4}{\|\nabla J\|_{L^1(B_r)}^5\widehat{C}^{\frac{3}{2}}(1+C_P^2)^{\frac{3}{2}}}.\tag{4-32}$$

We are left with a last estimate: thanks to (3-7), we know that $\|F'(\phi)\|_{L^\infty(\tau/2,\infty;L^1(\Omega))} \le C(\tau)$ and F' is monotone in a neighborhood of $+1$, so that we infer

$$y_0 = \int_{I_0}\int_{A_0(s)} 1\,\mathrm{d}x\,\mathrm{d}s \le \int_{I_0}\int_{\{x\in\Omega:\ \phi(x,s)\ge 1-2\delta\}} 1\,\mathrm{d}x\,\mathrm{d}s \le \int_{I_0}\int_{A_0(s)}\frac{|F'(\phi)|}{F'(1-2\delta)}\,\mathrm{d}x\,\mathrm{d}s \le \frac{3C(\tau)\tilde{\tau}}{F'(1-2\delta)}.$$

Therefore, if we ensure that

$$\frac{3C(\tau)\tilde{\tau}}{F'(1-2\delta)} \le \frac{2^{-20}\delta^5(F''(1-2\delta))^4}{\|\nabla J\|^5_{L^1(B_r)}\widehat{C}^{\frac{3}{2}}(1+C_P^2)^{\frac{3}{2}}},$$

then (4-32) holds. Having fixed $\tilde{\tau}$ in (4-13) such that

$$\tilde{\tau} = \frac{2^{-20}\delta^5(F''(1-2\delta))^4 F'(1-2\delta)}{3C(\tau)\|\nabla J\|^5_{L^1(B_r)}\widehat{C}^{\frac{3}{2}}(1+C_P^2)^{\frac{3}{2}}}, \tag{4-33}$$

we obtain the result. Notice that δ is fixed, so $\tilde{\tau} > 0$ is not infinitesimal, but it depends on ϕ_0 in a nontrivial way (thus not only on the initial energy) through C_P.

In the end, passing to the limit in y_n as $n \to \infty$, we have obtained that

$$\|(\phi-(1-\delta))^+\|_{L^\infty(\Omega\times(T-\tilde{\tau},T))} = 0,$$

since, as $n \to \infty$,

$$y_n \to \left|\left\{(x,t) \in \Omega \times [T-\tilde{\tau},T] : \phi(x,t) \ge 1-\delta\right\}\right| = 0.$$

We now repeat exactly the same argument for the case $(\phi-(-1+\delta))^-$ (using $\phi_n(t) = (\phi(t)+k_n)^-$). Notice that also for this second case we have the same constant C_P (see (4-12)). Moreover, the argument is exactly the same due to assumption (4-2), which implies that

$$\frac{1}{F''(-1+2\delta)} = O(\delta) \quad \text{and} \quad \frac{1}{|F'(-1+2\delta)|} = O\left(\frac{1}{|\ln(\delta)|}\right),$$

for δ sufficiently small. We can then choose the minima between the δ and $\tilde{\tau}$ obtained in the two cases, to get in the end that there exists a couple $\delta > 0$, $\tilde{\tau} > 0$ such that

$$-1+\delta \le \phi(x,t) \le 1-\delta, \quad \text{a.e. in } \Omega \times (T-\tilde{\tau},T). \tag{4-34}$$

Finally, notice that, due to the choice of T, we have $T-\tilde{\tau} = 2\tilde{\tau}+\frac{\tau}{2} \le \tau$; therefore we can repeat the same procedure on the interval $(T, T+\tilde{\tau})$ (this means that the new starting time will be $t_{-1} = T-2\tilde{\tau} \ge \frac{\tau}{2}$) and so on, reaching eventually the entire interval $[\tau,+\infty)$. Clearly δ and $\tilde{\tau}$ are always the same, since the constant C_P is uniform over the entire interval $[\tau,+\infty)$ and the time horizon T does not enter in any of the estimates. The proof is thus concluded.

Remark 4.7. We point out that the same proof holds for the case of convective nonlocal Cahn–Hilliard equation

$$\begin{cases} \partial_t\phi + \boldsymbol{u}\cdot\nabla\phi - \Delta\mu = 0 & \text{in } \Omega\times(0,T), \\ \mu = F'(\phi) - J*\phi & \text{in } \Omega\times(0,T), \\ \partial_{\boldsymbol{n}}\mu = 0 & \text{on } \partial\Omega\times(0,T), \\ \phi(\cdot,0) = \phi_0 & \text{in } \Omega, \end{cases} \tag{4-35}$$

where $\boldsymbol{u}$ is a sufficiently regular divergence free vector field, such that $\boldsymbol{u}\cdot\boldsymbol{n} = 0$ on $\partial\Omega\times(0,T)$. Indeed, Theorem 3.3 can be mostly extended also to this case (see, e.g., [Gal et al. 2017, Section 6], in which a related system, the nonlocal Cahn–Hilliard–Navier–Stokes system, is analyzed). Moreover, in the proof of

Theorem 4.3 the term $\boldsymbol{u}\cdot\nabla\phi$ does not appear, since in (4-21) we should get an additional $(\boldsymbol{u}\cdot\nabla\phi, \phi_n\eta_n^2)$, which is zero thanks to the assumptions on $\boldsymbol{u}$. Therefore, the separation property could a priori be obtained also in the couplings of the nonlocal Cahn–Hilliard equation with some hydrodynamic models, like Navier–Stokes equations (see, e.g., [Abels and Terasawa 2020] or [Gal et al. 2017, Section 6] for some examples of such models).

Remark 4.8. One might think that the proof of Theorem 4.3 could be adapted also to the conserved Allen–Cahn equation

$$\begin{cases} \partial_t\phi+\mu-\bar{\mu}=0 & \text{in } \Omega\times(0,T),\\ \mu=\Psi'(\phi)-\Delta\phi & \text{in } \Omega\times(0,T),\\ \partial_{\boldsymbol{n}}\phi=0 & \text{on } \partial\Omega\times(0,T),\\ \phi(\,\cdot\,,0)=\phi_0 & \text{in } \Omega, \end{cases} \tag{4-36}$$

where Ψ is defined in (1-2). Indeed, this has been obtained in [Grasselli and Poiatti 2024] in the case of multicomponent conserved Allen–Cahn equation in two and three dimensions, and it is valid also for (4-36). In the proof one loses the term

$$\int_{t_{n-1}}^{t}\int_{A_n(s)} F''(\phi)\nabla\phi\cdot\nabla\phi_n\eta_n^2\,\mathrm{d}x\,\mathrm{d}s,$$

which is substituted by $\int_{t_{n-1}}^{t}\int_{A_n(s)} F'(\phi)\phi_n\eta_n^2\,\mathrm{d}x\,\mathrm{d}s \geq F'(1-2\delta)\int_{t_{n-1}}^{t}\int_{A_n(s)}\phi_n\eta_n^2\,\mathrm{d}x\,\mathrm{d}s$: the presence of the first derivative of F instead of the second derivative, since $F'(1-2\delta)\to+\infty$ as $\delta\to 0^+$, is still enough to carry out the De Giorgi's iteration scheme, by heavily exploiting estimate (4-26). We also mention the fact that in two-dimensional bounded domains the instantaneous strict separation property for (4-36) was proven before in [Giorgini et al. 2022], by a completely different argument.

Remark 4.9. Assumption (H_3) shows that the strict separation property also holds for more general and singular potentials F than the logarithmic one (1-2). Furthermore, by slightly adapting the proof of Theorem 4.3, one can show that the same property also holds for more general double well potentials than F. For instance, one could deal with a chemical potential $\mu=\Psi'(\phi)+(J*1)\phi-J*\phi$, with Ψ defined in (1-2) and obtain an analogous result. Notice that in this new setting the nonlocal term $J*\phi$ is related to diffusion effects (see [Gal et al. 2023a]). Also, in the case of nonconstant mobility $M(\phi)$, the proof should work well as long as it is nondegenerate (i.e., bounded below by a strictly positive constant) and the existence of strong solutions is given. In conclusion, another possible extension could be in the case of dynamic boundary conditions (see, e.g., [Knopf and Signori 2021]): first one needs to assess the existence of strong solutions and the instantaneous regularization of weak solutions, and then apply a De Giorgi's iteration scheme, which seems harder due to the presence of boundary terms which have to be carefully handled.

4.2. ***Proof of Corollary 4.5.*** Observe that, due to Remark 4.6, we only need to prove that the unique solution ϕ departing from ϕ_0 is strictly separated from the pure phases in a neighborhood of the initial time. To this aim we perform again a De Giorgi's iteration scheme, in this case without the use of a cutoff function of the form (4-19). Indeed, the necessity of the cutoff function is merely to eliminate the presence of the initial datum in the estimates, but in our case, up to choosing $\delta \le \delta_0/2$, this problem does not appear any more, as we shall see. Again the Step 1 of the proof of Theorem 4.3 is still valid, and we adopt the same notation. Clearly, thanks to Remark 3.5, we can choose $\tau = 0$, so that again

$$\|\phi_\rho(t)\| \le C_P \|\nabla\phi_\rho(t)\|, \quad \|\tilde{\phi}_\rho(t)\| \le C_P \|\nabla\tilde{\phi}_\rho(t)\|, \quad \text{for almost any } t \ge 0 \text{ and for any } \rho \in [1-2\hat{\delta}, 1]. \tag{4-37}$$

We then start from Step 2. Let us fix δ sufficiently small such that $\delta \le \min\{\hat{\delta}, \delta_0/2\}$, so that (4-37) holds for any $\rho \in [1-2\delta, 1]$. Set then $\tilde{\tau} > 0$ such that (4-43) below holds. As in Theorem 4.3, we define the same sequence (4-16), but we do not need to consider any sequence of times, since we will always use the same, fixed, interval $I := [0, \tilde{\tau}]$. Then we define again

$$\phi_n(x, t) := (\phi - k_n)^+, \tag{4-38}$$

and, for any $n \ge 0$, we introduce the set

$$A_n(t) := \{x \in \Omega : \phi(x, t) - k_n \ge 0\} \quad \text{for all } t \in I,$$

so that

$$A_{n+1}(t) \subseteq A_n(t) \quad \text{for all } n \ge 0 \text{ and } t \in I.$$

We thus set

$$y_n = \int_I \int_{A_n(s)} 1 \, \mathrm{d}x \, \mathrm{d}s \quad \text{for all } n \ge 0.$$

Now, for any $n \ge 0$ we consider the test function $w = \phi_n$, and integrate over $[0, t]$, $t \le \tilde{\tau}$. Then we have, as in Theorem 4.3,

$$\frac{1}{2}\|\phi_n(t)\|^2 + \int_0^t \int_{A_n(s)} F''(\phi)\nabla\phi \cdot \nabla\phi_n \, \mathrm{d}x \, \mathrm{d}s = \int_0^t \int_{A_n(s)} (\nabla J * \phi) \cdot \nabla\phi_n \, \mathrm{d}x \, \mathrm{d}s + \frac{1}{2}\|\phi_n(0)\|^2.$$

Note that, due to the choice of $\delta \le \delta_0/2$, thanks to the strict separation of the initial datum, we immediately infer that $\|\phi_n(0)\| = 0$ for any $n \ge 0$. Following the same arguments as in the proof of Theorem 4.3, we obtain

$$\frac{1}{2}\|\phi_n(t)\|^2 + \frac{1}{2}F''(1-2\delta)\int_0^t \|\nabla\phi_n(s)\|^2 \, \mathrm{d}s \le \frac{\|\nabla J\|^2_{L^1(B_r)}}{2F''(1-2\delta)} y_n$$

for any $t \in [0, \tilde{\tau}]$. Observe that we do not see the presence of the term related to $1/\tilde{\tau}$ (estimated in (4-27)), since it is a consequence of the use of the cutoff function (4-19). This implies

$$\max_{t \in I} \|\phi_n(t)\|^2 \le Z_n, \qquad F''(1-2\delta)\int_I \|\nabla\phi_n\|^2 \, \mathrm{d}s \le Z_n, \tag{4-39}$$

where

$$Z_n := \frac{\|\nabla J\|^2_{L^1(B_r)}}{F''(1-2\delta)} y_n.$$

Observe that, for any $t \in I$ and for almost any $x \in A_{n+1}(t)$, we get

$$\phi_n(x, t) = \underbrace{\phi(x, t) - \left[1 - \delta - \frac{\delta}{2^{n+1}}\right]}_{\phi_{n+1}(x,t) \ge 0} + \delta\left[\frac{1}{2^n} - \frac{1}{2^{n+1}}\right] \ge \frac{\delta}{2^{n+1}},$$

which implies

$$\int_I \int_\Omega |\phi_n|^3 \,\mathrm{d}x\,\mathrm{d}s \geq \int_I \int_{A_{n+1}(s)} |\phi_n|^3 \,\mathrm{d}x\,\mathrm{d}s \geq \left(\frac{\delta}{2^{n+1}}\right)^3 \int_I \int_{A_{n+1}(s)} \mathrm{d}x\,\mathrm{d}s = \left(\frac{\delta}{2^{n+1}}\right)^3 y_{n+1}.$$

Then we have, as in (4-29),

$$\left(\frac{\delta}{2^{n+1}}\right)^3 y_{n+1} \leq \left(\int_I \int_\Omega |\phi_n|^{\frac{10}{3}} \,\mathrm{d}x\,\mathrm{d}s\right)^{\frac{9}{10}} \left(\int_I \int_{A_n(s)} \mathrm{d}x\,\mathrm{d}s\right)^{\frac{1}{10}}. \tag{4-40}$$

Again thanks to (2-3) and (4-37), we have

$$\int_I \int_\Omega |\phi_n|^{\frac{10}{3}} \,\mathrm{d}x\,\mathrm{d}s \leq \widehat{C}(1+C_P^2) \int_I \|\nabla \phi_n\|^2 \|\phi_n\|^{\frac{4}{3}} \,\mathrm{d}s,$$

so that, by (4-39),

$$\begin{aligned}
&\int_I \int_\Omega |\phi_n|^{\frac{10}{3}} \,\mathrm{d}x\,\mathrm{d}s \\
&\quad \leq \widehat{C}(1+C_P^2) \int_I \|\nabla \phi_n\|^2 \|\phi_n\|^{\frac{4}{3}} \,\mathrm{d}s \leq \widehat{C}(1+C_P^2) \max_{t \in I} \|\phi_n\|^{\frac{4}{3}} \int_I \|\nabla \phi_n\|^2 \,\mathrm{d}s \\
&\quad \leq \frac{\widehat{C}(1+C_P^2)}{F''(1-2\delta)} Z_n^{\frac{2}{3}} F''(1-2\delta) \int_I \|\nabla \phi_n\|^2 \,\mathrm{d}s \leq \frac{\widehat{C}(1+C_P^2)}{F''(1-2\delta)} Z_n^{\frac{5}{3}} \leq \frac{\|\nabla J\|_{L^1(B_r)}^{\frac{10}{3}} \widehat{C}(1+C_P^2)}{(F''(1-2\delta))^{\frac{8}{3}}} y_n^{\frac{5}{3}}.
\end{aligned}$$

Therefore, we immediately infer from (4-40) that

$$\begin{aligned}
\left(\frac{\delta}{2^{n+1}}\right)^3 y_{n+1} &\leq \left(\int_I \int_\Omega |\phi_n|^{\frac{10}{3}} \,\mathrm{d}x\,\mathrm{d}s\right)^{\frac{9}{10}} \left(\int_I \int_{A_n(s)} \mathrm{d}x\,\mathrm{d}s\right)^{\frac{1}{10}} \\
&\leq \frac{\|\nabla J\|_{L^1(B_r)}^3 \widehat{C}^{\frac{9}{10}} (1+C_P^2)^{\frac{9}{10}}}{(F''(1-2\delta))^{\frac{12}{5}}} y_n^{\frac{3}{2}} y_n^{\frac{1}{10}} = \frac{\|\nabla J\|_{L^1(B_r)}^3 \widehat{C}^{\frac{9}{10}} (1+C_P^2)^{\frac{9}{10}}}{(F''(1-2\delta))^{\frac{12}{5}}} y_n^{\frac{8}{5}}.
\end{aligned} \tag{4-41}$$

In conclusion, we end up with

$$y_{n+1} \leq \frac{2^{3n+3} \|\nabla J\|_{L^1(B_r)}^3 \widehat{C}^{\frac{9}{10}} (1+C_P^2)^{\frac{9}{10}}}{\delta^3 (F''(1-2\delta))^{\frac{12}{5}}} y_n^{\frac{8}{5}} \quad \text{for all } n \geq 0,$$

and we can apply Lemma 3.8. In particular, we have

$$b = 2^3 > 1, \qquad C = \frac{2^3 \|\nabla J\|_{L^1(B_r)}^3 \widehat{C}^{\frac{9}{10}} (1+C_P^2)^{\frac{9}{10}}}{\delta^3 (F''(1-2\delta))^{\frac{12}{5}}} > 0, \qquad \varepsilon = \frac{3}{5},$$

to get that $y_n \to 0$, as long as

$$y_0 \leq C^{-\frac{5}{3}} b^{-\frac{25}{9}},$$

i.e.,

$$y_0 \leq \frac{2^{-\frac{40}{3}}\delta^5(F''(1-2\delta))^4}{\|\nabla J\|^5_{L^1(B_r)}\widehat{C}^{\frac{3}{2}}(1+C_P^2)^{\frac{3}{2}}}. \tag{4-42}$$

In conclusion, since we know by (3-7) and Remark 3.5 that $\|F'(\phi)\|_{L^\infty(0,\infty;L^1(\Omega))} \leq C$, we infer

$$y_0 = \int_I \int_{A_0(s)} 1 \,\mathrm{d}x\,\mathrm{d}s \leq \int_I \int_{A_0(s)} \frac{|F'(\phi)|}{F'(1-2\delta)}\,\mathrm{d}x\,\mathrm{d}s \leq \frac{C\tilde{\tau}}{F'(1-2\delta)}.$$

Having fixed $\tilde{\tau}$ so that

$$\tilde{\tau} = \frac{2^{-\frac{40}{3}}\delta^5(F''(1-2\delta))^4 F'(1-2\delta)}{C\|\nabla J\|^5_{L^1(B_r)}\widehat{C}^{\frac{3}{2}}(1+C_P^2)^{\frac{3}{2}}}, \tag{4-43}$$

we have

$$\frac{C\tilde{\tau}}{F'(1-2\delta)} \leq \frac{2^{-\frac{40}{3}}\delta^5(F''(1-2\delta))^4}{\|\nabla J\|^5_{L^1(B_r)}\widehat{C}^{\frac{3}{2}}(1+C_P^2)^{\frac{3}{2}}},$$

so that (4-42) holds. In the end, passing to the limit in y_n as $n \to \infty$, we have obtained that

$$\|(\phi-(1-\delta))^+\|_{L^\infty(\Omega\times(0,\tilde{\tau}))} = 0.$$

We now repeat exactly the same argument for the case $(\phi-(-1+\delta))^-$ (using $\phi_n(t) = (\phi(t)+k_n)^-$), to get in the end that there exist $\delta > 0$, $\tilde{\tau} > 0$ such that

$$-1+\delta \leq \phi(x,t) \leq 1-\delta, \quad \text{a.e. in } \Omega\times(0,\tilde{\tau}). \tag{4-44}$$

Notice that $\tilde{\tau}$ can be explicitly computed as a function of the parameters of the problem and the initial datum (see (4-43)). The proof is then concluded, recalling Remark 4.6 with $T_1 = \tilde{\tau}$.

5. Some consequences of the strict separation property

In this section we collect some results which are straightforward consequences of the strict separation property proven in Theorem 4.3.

5.1. ***Regularization in finite time.*** First we show that the weak solution given by Theorem 3.3 actually regularizes more. Indeed, we have a first immediate consequence:

Corollary 5.1. *Under the same assumptions of Theorem 4.3, for any* $\tau > 0$, *there exists a constant* $C = C(\tau) > 0$ *such that*

$$\|F'(\phi(t))\|_{L^\infty(\Omega)} + \|\mu(t)\|_{L^\infty(\Omega)} \leq C \quad \textit{for all } t \geq \tau.$$

Proof. The proof is immediate, since by the strict separation property we deduce $\|F'(\phi(t))\|_{L^\infty(\Omega)} \leq C$ for any $t \geq \tau$ and then by comparison we get the L^∞- control on μ. □

Furthermore, we can also obtain the Hölder regularity of the weak solutions:

Corollary 5.2. *Under the same assumptions of Theorem 4.3, for any $\tau > 0$, there exists $C = C(\tau) > 0$ and $\kappa = \kappa(\tau, \delta) \in (0, 1)$ such that*

$$|\phi(x_1, t_1) - \phi(x_2, t_2)| \leq C\big(|x_1 - x_2|^\kappa + |t_1 - t_2|^{\frac{\kappa}{2}}\big), \tag{5-1}$$

$$|\mu(x_1, t_1) - \mu(x_2, t_2)| \leq C\big(|x_1 - x_2|^\kappa + |t_1 - t_2|^{\frac{\kappa}{2}}\big) \tag{5-2}$$

for all $(x_1, t_1), (x_2, t_2) \in \Omega_t$, where $\Omega_t = [t, t+1] \times \overline{\Omega}$ and $t \geq \tau$.

Proof. We can argue as in [Gal and Grasselli 2014, Lemma 2.11]. In particular, we rewrite the system (1-7) in the form

$$\partial_t \phi = \operatorname{div}(a(x, \phi, \nabla\phi)), \quad (a(x, \phi, \nabla\phi) \cdot \boldsymbol{n})_{|\partial\Omega} = 0,$$

with

$$a(x, \phi, \nabla\phi) := F''(\phi)\nabla\phi - \nabla J * \phi.$$

Since by (H_1) we have $J \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^3)$, $F''(s) \geq \alpha$ for any $s \in (-1, 1)$ by (H_2) and $\|\nabla J * \phi\|_{L^\infty(\Omega)} \leq \|\nabla J\|_{L^1(B_r)}$ by (4-24), by Young's inequality we get

$$\begin{aligned} a(x, \phi, \nabla\phi) \cdot \nabla\phi &= F''(\phi)|\nabla\phi|^2 - (\nabla J * \phi) \cdot \nabla\phi \\ &\geq \alpha|\nabla\phi|^2 - \|\nabla J\|_{L^1(B_r)}|\nabla\phi| \\ &\geq \frac{\alpha}{2}|\nabla\phi|^2 - \frac{1}{2\alpha}\|\nabla J\|^2_{L^1(B_r)}, \end{aligned}$$

and, similarly, by Corollary 5.1 and (4-24),

$$|a(x, \phi, \nabla\phi)| \leq \|F''(\phi)\|_{L^\infty(\Omega)}|\nabla\phi| + \|\nabla J * \phi\|_{L^\infty(\Omega)} \leq C_1|\nabla\phi| + \|\nabla J\|_{L^1(B_r)}$$

for some positive constant C_1 depending on τ, δ. Therefore we infer the desired estimate (5-1) applying [Dung 2000, Corollary 4.2]. Then, by the regularity of F, we immediately deduce the same result for μ, concluding the proof. □

In order to obtain higher-order spatial regularity for the phase variable ϕ, we need to strengthen the assumptions on the interaction kernel J. In particular, we assume

(H_4) Either $J \in W^{2,1}(\mathcal{B}_R)$, where $\mathcal{B}_R := \{x \in \mathbb{R}^3 : |x| < R\}$, with $R \sim \operatorname{diam}(\Omega)$ such that $\overline{\Omega} \subset \mathcal{B}_R$ and $x - \Omega \subset B_R$ for any $x \in \Omega$, or J is admissible in the sense of [Bedrossian et al. 2011, Definition 1].

Remark 5.3. As noticed in [Gal et al. 2017, Remark 5.9], we observe that Newtonian and second-order Bessel potentials satisfy assumption (H_4), namely they are admissible in the sense of [Bedrossian et al. 2011, Definition 1].

Lemma 5.4. *Under the same assumptions of Theorem 4.3, assuming also that J satisfies (H_4) and $F \in C^3(-1, 1)$, for any $\tau > 0$ there exists $C = C(\tau) > 0$ such that*

$$\|\phi\|_{L^{4/3}(t,t+1;H^2(\Omega))} \leq C \quad \textit{for all } t \geq \tau. \tag{5-3}$$

Proof. We first observe that, since we can apply Theorem 3.3, by (3-7)–(3-8) we deduce that

$$\|\nabla\phi\|_{L^{8/3}(t,t+1;L^4(\Omega))} + \|\mu\|_{L^2(t,t+1;V_2)} \leq C(\tau) \quad \text{for all } t \geq \tau, \tag{5-4}$$

for some positive constant $C(\tau)$. Then, as in [Frigeri et al. 2016, Theorem 5], we proceed formally (these computations could be justified in a suitable approximating scheme, see, e.g., [Frigeri et al. 2016, Theorem 5, Step 3]) defining

$$\partial^2_{ij}\phi := \frac{\partial^2\phi}{\partial x_i \partial x_j} \quad \text{for } i, j = 1, 2, 3.$$

We now apply ∂^2_{ij} to the equation for the chemical potential μ and integrate on Ω, to infer

$$\int_\Omega \partial^2_{ij}\mu\partial^2_{ij}\phi \,\mathrm{d}x = \int_\Omega F''(\phi)(\partial^2_{ij}\phi)^2 \,\mathrm{d}x - \int_\Omega \partial_i(\partial_j J*\phi)\partial^2_{ij}\phi \,\mathrm{d}x + \int_\Omega F'''(\phi)\partial_i\phi\partial_j\phi\partial^2_{ij}\phi, \quad i, j = 1, 2, 3.$$

We now recall assumption (H$_4$), so that, by [Bedrossian et al. 2011, Lemma 2],

$$\|\partial_i(\partial_j J * \phi)\|_{L^2(\Omega)} \leq C\|\phi\|_{L^2(\Omega)} \leq C(\tau).$$

Therefore, by Cauchy–Schwartz and Young's inequalities, we infer, recalling that $F''(s) \geq \alpha$ for any $s \in (-1, 1)$, and exploiting the separation property of Theorem 4.3, for any $t \geq \tau$,

$$\frac{\alpha}{2}\|\partial^2_{ij}\phi\|^2 \leq C(1 + \|\partial^2_{ij}\mu\|^2 + \int_\Omega |\partial_i\phi|^2|\partial_j\phi|^2 \,\mathrm{d}x) \leq C\big(1 + \|\mu\|^2_{H^2(\Omega)} + \|\nabla\phi\|^4_{L^4(\Omega)}\big), \quad i, j = 1, 2, 3,$$

which implies (5-3), thanks to (5-4). □

5.2. *Convergence to equilibrium.* We conclude the results of our paper by showing that the strict separation property is essential to study the longtime behavior of the single trajectory. In particular, we can follow [Della Porta et al. 2018, Section 6.2]: for the sake of completeness we give here a sketch of the proofs. We employ the typical strategy based on the Lyapunov property of the associated system (see (3-4)) and the well known Łojasiewicz–Simon inequality. Let us consider the set of admissible initial data

$$\mathcal{H}_m := \{\phi \in L^\infty(\Omega) : \|\phi\|_{L^\infty(\Omega)} \leq 1, \quad |\bar{\phi}| = m\},$$

with $m \in [0, 1)$, and fix an initial datum $\phi_0 \in \mathcal{H}_m$. Let then ϕ be the unique weak global-in-time solution departing from ϕ_0, whose existence and uniqueness is ensured by Theorem 3.3. We introduce the ω-limit set associated to ϕ_0, i.e.,

$$\omega(\phi_0) = \{\tilde{\phi} \in \mathcal{H}_m : \exists t_n \to \infty \text{ such that } \phi(t_n) \to \tilde{\phi} \text{ in } H\}.$$

By (3-6), ϕ is uniformly bounded in V, which is compactly embedded in H. Therefore, by standard results related to the intersection of nonempty, compact (in H), connected and nested sets, we infer that $\omega(\phi_0)$ is nonempty, compact and connected in $\mathcal{H}_m$. We now characterize the set $\omega(\phi_0)$, showing that it is composed by equilibrium points (i.e., stationary solutions) associated to (1-7), which are defined as:

Definition 5.5. ϕ_∞ is an equilibrium point to problem (1-7) if $\phi_\infty \in \mathcal{H}_m \cap V$ satisfies the stationary nonlocal Cahn–Hilliard equation

$$F'(\phi_\infty) - J * \phi_\infty = \mu_\infty \quad \text{in } \Omega, \tag{5-5}$$

where $\mu_\infty \in \mathbb{R}$ is a real constant.

As noticed also in [Della Porta et al. 2018], the existence of a (not necessarily unique, see, e.g., [Bates and Chmaj 1999]) solution to (5-5) can be proven by means of a fixed point argument. Moreover, as shown in [Della Porta et al. 2018, Lemma 6.1], any $\phi_\infty \in V \cap \mathcal{H}_m$ satisfying (5-5) is strictly separated from the pure phases, i.e., there exists $\delta > 0$ such that

$$\|\phi_\infty\|_{L^\infty(\Omega)} \le 1 - \delta.$$

If we now introduce the set of all the stationary points of the nonlocal Cahn–Hilliard equation,

$$\mathcal{S} := \{\phi_\infty \in \mathcal{H}_m \cap V : \phi_\infty \text{ satisfies (5-5)}\},$$

we can easily prove that $\omega(\phi_0) \subset \mathcal{S}$. Indeed, let us consider a sequence $t_n \to \infty$ such that $\phi(t_n) \to \tilde{\phi}$ in H, $\tilde{\phi} \in \omega(\phi_0)$. We then define the sequence of trajectories $\phi_n(t) := \phi(t + t_n)$ and $\mu_n(t) := \mu(t + t_n)$. Thanks to (3-6), we get, up to a nonrelabeled subsequence, that $\phi_n \overset{*}{\rightharpoonup} \phi^*$ in $L^\infty(0, \infty; V)$. Passing to the limit in the equations for ϕ_n, exploiting the results of Theorem 3.3, we infer that also ϕ^* satisfies (3-2)–(3-3) (we denote the corresponding chemical potential by μ^*), with initial datum $\phi^*(0) = \tilde{\phi}$. This last consideration follows from the fact that $\phi_n(0) = \phi(t_n) \to \tilde{\phi}$ strongly in H. Moreover, we clearly have $\lim_{n\to\infty} \mathcal{E}(\phi_n(t)) = \mathcal{E}(\phi^*(t))$ for all $t \ge 0$. By the energy identity (3-4), we infer that the energy $\mathcal{E}(\phi(\cdot))$ is nonincreasing in time, thus there exists E_∞ such that $\lim_{t\to\infty} \mathcal{E}(\phi(t)) = E_\infty$. This means that this convergence also holds for the subsequence $\{t + t_n\}_n$, thus

$$\mathcal{E}(\phi^*(t)) = \lim_{n\to\infty} \mathcal{E}(\phi_n(t)) = \lim_{n\to\infty} \mathcal{E}(\phi(t + t_n)) = E_\infty,$$

entailing that $\mathcal{E}(\phi^*(\cdot))$ is constant in time. Passing then to the limit in (3-4), which is valid for each ϕ_n, we obtain

$$E_\infty + \int_s^t \|\nabla \mu^*(\tau)\|^2 \, d\tau \le E_\infty \quad \text{for all } 0 \le s \le t < \infty,$$

implying $\nabla \mu^* = 0$ almost everywhere in Ω, and thus, by comparison in (3-2), also $\partial_t \phi^* = 0$ almost everywhere in Ω, for almost every $t \ge 0$. Therefore, we infer that ϕ^* is constant in time, namely $\phi^*(t) = \tilde{\phi}$ for all $t \ge 0$. Thus $\tilde{\phi}$ satisfies (5-5) with some constant $\mu_\infty \in \mathbb{R}$, and then $\tilde{\phi} \in \mathcal{S}$, implying, being $\tilde{\phi} \in \omega(\phi_0)$ arbitrary, $\omega(\phi_0) \subset \mathcal{S}$. Notice that in this way we have shown that, for any $\phi_\infty \in \omega(\phi_0)$,

$$\mathcal{E}(\phi_\infty) = E_\infty = \lim_{s\to\infty} \mathcal{E}(\phi(s)) = \inf_{s\ge 0} \mathcal{E}(\phi(s)) \le \mathcal{E}(\phi(t)) \quad \text{for all } t \ge 0. \tag{5-6}$$

We can then conclude by showing that $\omega(\phi_0)$ is a singleton. For the sake of clarity we present here the main tool, which is the Łojasiewicz–Simon inequality (see, e.g., [Della Porta et al. 2018, Proposition 6.2] or [Gajewski and Griepentrog 2006]):

Proposition 5.6. *Let $P_0 : H \to H_0$ be the projector operator. Assume that F satisfies* (H$_2$) *and is real analytic in $(-1, 1)$, $\phi \in V \cap L^\infty(\Omega)$ is such that $-1+\gamma \le \phi(x) \le 1-\gamma$, for any $x \in \overline{\Omega}$, for some $\gamma \in (0, 1)$ and $\phi_\infty \in \mathcal{S}$. Then there exists $\theta \in \left(0, \frac{1}{2}\right)$, $\eta > 0$ and a positive constant C such that*

$$|\mathcal{E}(\phi) - \mathcal{E}(\phi_\infty)|^{1-\theta} \le C \| P_0(F'(\phi) - J * \phi)\|_*, \tag{5-7}$$

whenever $\|\phi - \phi_\infty\| \le \eta$.

Remark 5.7. We observe that the logarithmic potential F is indeed real analytic in $(-1, 1)$, thus the assumption of the foregoing proposition is satisfied.

Theorem 5.8. *Under the same assumptions as in Theorem 4.3, suppose additionally that F is real analytic in $(-1, 1)$. Then the weak solution ϕ, departing from the initial datum $\phi_0 \in \mathcal{H}_m$ converges to a single equilibrium point ϕ_∞ (depending on ϕ_0) and $\omega(\phi_0) = \{\phi_\infty\}$. In particular we have*

$$\lim_{t\to\infty} \|\phi(t) - \phi_\infty\| = 0. \tag{5-8}$$

Proof. Thanks to (5-6), we infer that $\mathcal{E}(\phi(t)) \ge \mathcal{E}(\phi_\infty)$, $\mathcal{E}(\phi(t)) \to \mathcal{E}(\phi_\infty)$, as $t \to \infty$, for any $\phi_\infty \in \omega(\phi_0)$. Without loss of generality we can assume $\mathcal{E}(\phi(t)) > \mathcal{E}(\phi_\infty)$ for all $t \ge 0$. Indeed, if there exists $\bar{t} > 0$ such that $\mathcal{E}(\phi(\bar{t})) = \mathcal{E}(\phi_\infty)$, then clearly $\phi(t) = \phi(\bar{t})$ for any $t \ge \bar{t}$ and the claim follows, since then $\phi(t) = \phi_\infty$ for any $t \ge \bar{t}$. Let us now fix $\theta \in \left(0, \frac{1}{2}\right)$ and $\eta > 0$ given in Proposition 5.6, where we have chosen γ equal to the value of δ given in Theorem 4.3. By a contradiction argument as in [Frigeri et al. 2013, Theorem 4] it is possible to show that there exists $t^* > 0$ such that $\|\phi(t) - \phi_\infty\| \le \eta$, for all $t \ge t^*$. Therefore, since the solution ϕ enjoys the separation property (by Theorem 4.3) and thanks to the choice of γ, by Proposition 5.6 we get, for any $t \ge t^*$,

$$(\mathcal{E}(\phi) - \mathcal{E}(\phi_\infty))^{1-\theta} \le \|P_0(F'(\phi) - J * \phi)\|_* \le C\|P_0\mu\| \le \widehat{C}\|\nabla\mu\|,$$

where $\widehat{C} > 0$ depends on C and on the Poincaré–Wirtinger constant. Therefore, by means of the energy identity (3-4), we deduce, for any $t \ge t^*$,

$$-\frac{d}{dt}(\mathcal{E}(\phi) - \mathcal{E}(\phi_\infty))^\theta = -\theta(\mathcal{E}(\phi) - \mathcal{E}(\phi_\infty))^{\theta-1}\frac{d}{dt}\mathcal{E}(\phi) \ge \frac{\theta\|\nabla\mu\|^2}{\widehat{C}\|\nabla\mu\|} \ge \widetilde{C}\|\nabla\mu\|,$$

where $\widetilde{C} > 0$ is a positive constant independent of t. An integration over $(t^*, +\infty)$, for t^* sufficiently large, implies that $\nabla\mu \in L^1(t^*, \infty; \boldsymbol{H})$. By comparison, we deduce that also $\partial_t\phi \in L^1(t^*, \infty; V')$, so that

$$\phi(t) = \phi(t^*) + \int_{t^*}^t \partial_t\phi(\tau)\, d\tau \xrightarrow{t\to+\infty} \tilde{\phi} \quad \text{in } V',$$

for some $\tilde{\phi} \in V'$. Then we infer that $\phi(t)$ converges in V' as $t \to \infty$. By uniqueness of the limit in V', we can then conclude that $\omega(\phi_0)$ is a singleton, i.e., $\omega(\phi_0) = \{\tilde{\phi}\}$. From now on we will denote $\tilde{\phi}$ by ϕ_∞, since any $\phi_\infty \in \omega(\phi_0)$ coincides with $\tilde{\phi}$. Thanks to (3-6), we then get (5-8) by interpolation:

$$\|\phi(t) - \phi_\infty\| \le C\|\phi(t) - \phi_\infty\|_V^{1/2}\|\phi(t) - \phi_\infty\|_{V'}^{1/2} \le C\|\phi(t) - \phi_\infty\|_{V'}^{1/2} \xrightarrow{t\to+\infty} 0,$$

concluding the proof. ∎

Acknowledgments

The author thanks the referees for their appropriate comments and useful remarks. The author is also grateful to Andrea Giorgini and Maurizio Grasselli for several helpful comments on a preliminary version of this article and to Giorgio Meretti for the fruitful discussion about the lemma on the Poincaré-type inequality. Moreover, the author has been partially funded by MIUR-PRIN Grant 2020F3NCPX "Mathematics for Industry 4.0 (Math4I4)".

References

[Abels and Terasawa 2020] H. Abels and Y. Terasawa, "Weak solutions for a diffuse interface model for two-phase flows of incompressible fluids with different densities and nonlocal free energies", *Math. Methods Appl. Sci.* **43**:6 (2020), 3200–3219. MR Zbl

[Abels and Wilke 2007] H. Abels and M. Wilke, "Convergence to equilibrium for the Cahn–Hilliard equation with a logarithmic free energy", *Nonlinear Anal.* **67**:11 (2007), 3176–3193. MR Zbl

[Allen and Cahn 1979] S. M. Allen and J. W. Cahn, "A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening", *Acta Metallurg.* **27**:6 (1979), 1085–1095.

[Bates and Chmaj 1999] P. W. Bates and A. Chmaj, "An integrodifferential model for phase transitions: stationary solutions in higher space dimensions", *J. Stat. Phys.* **95**:5-6 (1999), 1119–1139. MR Zbl

[Bedrossian et al. 2011] J. Bedrossian, N. Rodríguez, and A. L. Bertozzi, "Local and global well-posedness for aggregation equations and Patlak–Keller–Segel models with degenerate diffusion", *Nonlinearity* **24**:6 (2011), 1683–1714. MR Zbl

[Bertozzi et al. 2007] A. L. Bertozzi, S. Esedoḡlu, and A. Gillette, "Inpainting of binary images using the Cahn–Hilliard equation", *IEEE Trans. Image Process.* **16**:1 (2007), 285–291. MR Zbl

[Boyer and Fabrie 2013] F. Boyer and P. Fabrie, *Mathematical tools for the study of the incompressible Navier–Stokes equations and related models*, Appl. Math. Sci. **183**, Springer, 2013. MR Zbl

[Brezis 2011] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Springer, 2011. MR Zbl

[Cahn and Hilliard 1958] J. W. Cahn and J. E. Hilliard, "Free energy of a nonuniform system, I: Interfacial free energy", *J. Chem. Phys.* **28**:2 (1958), 258–267. Zbl

[Cherfils et al. 2011] L. Cherfils, A. Miranville, and S. Zelik, "The Cahn–Hilliard equation with logarithmic potentials", *Milan J. Math.* **79**:2 (2011), 561–596. MR Zbl

[Colli et al. 2007] P. Colli, P. Krejčí, E. Rocca, and J. Sprekels, "Nonlinear evolution inclusions arising from phase change models", *Czechoslovak Math. J.* **57(132)**:4 (2007), 1067–1098. MR Zbl

[Davoli et al. 2021a] E. Davoli, L. Scarpa, and L. Trussardi, "Nonlocal-to-local convergence of Cahn–Hilliard equations: Neumann boundary conditions and viscosity terms", *Arch. Ration. Mech. Anal.* **239**:1 (2021), 117–149. MR Zbl

[Davoli et al. 2021b] E. Davoli, L. Scarpa, and L. Trussardi, "Local asymptotics for nonlocal convective Cahn–Hilliard equations with $W^{1,1}$ kernel and singular potential", *J. Differential Equations* **289** (2021), 35–58. MR Zbl

[Debussche and Dettori 1995] A. Debussche and L. Dettori, "On the Cahn–Hilliard equation with a logarithmic free energy", *Nonlinear Anal.* **24**:10 (1995), 1491–1514. MR Zbl

[Della Porta and Grasselli 2015] F. Della Porta and M. Grasselli, "Convective nonlocal Cahn–Hilliard equations with reaction terms", *Discrete Contin. Dyn. Syst. Ser. B* **20**:5 (2015), 1529–1553. MR Zbl

[Della Porta et al. 2018] F. Della Porta, A. Giorgini, and M. Grasselli, "The nonlocal Cahn–Hilliard–Hele-Shaw system with logarithmic potential", *Nonlinearity* **31**:10 (2018), 4851–4881. MR Zbl

[DiBenedetto 1993] E. DiBenedetto, *Degenerate parabolic equations*, Springer, 1993. MR Zbl

[Dolgin 2018] E. Dolgin, "What lava lamps and vinaigrette can teach us about cell biology", *Nature* **555** (2018), 300–302.

[Dung 2000] L. Dung, "Remarks on Hölder continuity for parabolic equations and convergence to global attractors", *Nonlinear Anal.* **41**:7-8 (2000), 921–941. MR Zbl

[Elliott and Garcke 1996] C. M. Elliott and H. Garcke, "On the Cahn–Hilliard equation with degenerate mobility", *SIAM J. Math. Anal.* **27**:2 (1996), 404–423. MR Zbl

[Elliott and Luckhaus 1991] C. M. Elliott and L. Luckhaus, "A generalized equation for phase separation of a multi-component mixture with interfacial free energy", preprint 195, Univ. Bonn, 1991.

[Emmerich 2003] H. Emmerich, *The diffuse interface approach in materials science*, Lecture Notes in Phys. Monogr. **73**, Springer, 2003. Zbl

[Fife 2000] P. C. Fife, "Models for phase separation and their mathematics", *Electron. J. Differential Equations* **2000** (2000), art. id. 48. MR Zbl

[Frigeri and Grasselli 2012a] S. Frigeri and M. Grasselli, "Nonlocal Cahn–Hilliard–Navier–Stokes systems with singular potentials", *Dyn. Partial Differ. Equ.* **9**:4 (2012), 273–304. MR Zbl

[Frigeri and Grasselli 2012b] S. Frigeri and M. Grasselli, "Global and trajectory attractors for a nonlocal Cahn–Hilliard–Navier–Stokes system", *J. Dynam. Differential Equations* **24**:4 (2012), 827–856. MR Zbl

[Frigeri et al. 2013] S. Frigeri, M. Grasselli, and P. Krejčí, "Strong solutions for two-dimensional nonlocal Cahn–Hilliard–Navier–Stokes systems", *J. Differential Equations* **255**:9 (2013), 2587–2614. MR Zbl

[Frigeri et al. 2016] S. Frigeri, C. G. Gal, and M. Grasselli, "On nonlocal Cahn–Hilliard–Navier–Stokes systems in two dimensions", *J. Nonlinear Sci.* **26**:4 (2016), 847–893. MR Zbl

[Gajewski and Griepentrog 2006] H. Gajewski and J. A. Griepentrog, "A descent method for the free energy of multicomponent systems", *Discrete Contin. Dyn. Syst.* **15**:2 (2006), 505–528. MR Zbl

[Gajewski and Zacharias 2003] H. Gajewski and K. Zacharias, "On a nonlocal phase separation model", *J. Math. Anal. Appl.* **286**:1 (2003), 11–31. MR Zbl

[Gal and Grasselli 2014] C. G. Gal and M. Grasselli, "Longtime behavior of nonlocal Cahn–Hilliard equations", *Discrete Contin. Dyn. Syst.* **34**:1 (2014), 145–179. MR Zbl

[Gal and Shomberg 2022] C. G. Gal and J. L. Shomberg, "Cahn–Hilliard equations governed by weakly nonlocal conservation laws and weakly nonlocal particle interactions", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **39**:5 (2022), 1179–1234. MR Zbl

[Gal et al. 2017] C. G. Gal, A. Giorgini, and M. Grasselli, "The nonlocal Cahn–Hilliard equation with singular potential: well-posedness, regularity and strict separation property", *J. Differential Equations* **263**:9 (2017), 5253–5297. MR Zbl

[Gal et al. 2023a] C. G. Gal, A. Giorgini, and M. Grasselli, "The separation property for 2D Cahn–Hilliard equations: local, nonlocal and fractional energy cases", *Discrete Contin. Dyn. Syst.* **43**:6 (2023), 2270–2304. MR Zbl

[Gal et al. 2023b] C. G. Gal, A. Giorgini, M. Grasselli, and A. Poiatti, "Global well-posedness and convergence to equilibrium for the Abels–Garcke–Grün model with nonlocal free energy", *J. Math. Pures Appl.* **178** (2023), 46–109. Zbl

[Giacomin and Lebowitz 1996] G. Giacomin and J. L. Lebowitz, "Exact macroscopic description of phase segregation in model alloys with long range interactions", *Phys. Rev. Lett.* **76**:7 (1996), 1094–1097.

[Giacomin and Lebowitz 1997] G. Giacomin and J. L. Lebowitz, "Phase segregation dynamics in particle systems with long range interactions, I: Macroscopic limits", *J. Stat. Phys.* **87**:1-2 (1997), 37–61. MR Zbl

[Giacomin and Lebowitz 1998] G. Giacomin and J. L. Lebowitz, "Phase segregation dynamics in particle systems with long range interactions, II: Interface motion", *SIAM J. Appl. Math.* **58**:6 (1998), 1707–1729. MR Zbl

[Giorgini 2024] A. Giorgini, "On the separation property and the global attractor for the nonlocal Cahn–Hilliard equation in three dimensions", *J. Evol. Equ.* **24**:2 (2024), art. id. 21. MR Zbl

[Giorgini et al. 2017] A. Giorgini, M. Grasselli, and A. Miranville, "The Cahn–Hilliard–Oono equation with singular potential", *Math. Models Methods Appl. Sci.* **27**:13 (2017), 2485–2510. MR Zbl

[Giorgini et al. 2022] A. Giorgini, M. Grasselli, and H. Wu, "On the mass-conserving Allen–Cahn approximation for incompressible binary fluids", *J. Funct. Anal.* **283**:9 (2022), art. id. 109631. MR Zbl

[Grasselli and Poiatti 2024] M. Grasselli and A. Poiatti, "Multi-component conserved Allen–Cahn equations", *Interfaces Free Bound.* **26**:4 (2024), 489–541. MR Zbl

[Grasselli et al. 2023] M. Grasselli, L. Scarpa, and A. Signori, "On a phase field model for RNA-protein dynamics", *SIAM J. Math. Anal.* **55**:1 (2023), 405–457. MR Zbl

[He and Wu 2021] J. He and H. Wu, "Global well-posedness of a Navier–Stokes–Cahn–Hilliard system with chemotaxis and singular potential in 2D", *J. Differential Equations* **297** (2021), 47–80. MR Zbl

[Kenmochi et al. 1995] N. Kenmochi, M. Niezgódka, and I. Pawlow, "Subdifferential operator approach to the Cahn–Hilliard equation with constraint", *J. Differential Equations* **117**:2 (1995), 320–356. MR Zbl

[Knopf and Signori 2021] P. Knopf and A. Signori, "On the nonlocal Cahn–Hilliard equation with nonlocal dynamic boundary condition and boundary penalization", *J. Differential Equations* **280** (2021), 236–291. MR Zbl

[Krejčí et al. 2007] P. Krejčí, E. Rocca, and J. Sprekels, "A nonlocal phase-field model with nonconstant specific heat", *Interfaces Free Bound.* **9**:2 (2007), 285–306. MR Zbl

[Ladyženskaja et al. 1968] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Uralceva, *Linear and quasilinear equations of parabolic type*, Transl. Math. Monogr. **23**, Amer. Math. Soc., Providence, RI, 1968. MR

[Londen and Petzeltová 2011] S.-O. Londen and H. Petzeltová, "Regularity and separation from potential barriers for a non-local phase-field system", *J. Math. Anal. Appl.* **379**:2 (2011), 724–735. MR Zbl

[Londen and Petzeltová 2018] S.-O. Londen and H. Petzeltová, "Regularity and separation from potential barriers for the Cahn–Hilliard equation with singular potential", *J. Evol. Equ.* **18**:3 (2018), 1381–1393. MR Zbl

[Lowengrub and Truskinovsky 1998] J. Lowengrub and L. Truskinovsky, "Quasi-incompressible Cahn–Hilliard fluids and topological transitions", *Proc. A* **454**:1978 (1998), 2617–2654. MR Zbl

[Miranville and Zelik 2004] A. Miranville and S. Zelik, "Robust exponential attractors for Cahn–Hilliard type equations with singular potentials", *Math. Methods Appl. Sci.* **27**:5 (2004), 545–582. MR Zbl

[Nagai et al. 1997] T. Nagai, T. Senba, and K. Yoshida, "Application of the Trudinger–Moser inequality to a parabolic system of chemotaxis", *Funkcial. Ekvac.* **40**:3 (1997), 411–433. MR Zbl

[Poiatti and Signori 2024] A. Poiatti and A. Signori, "Regularity results and optimal velocity control of the convective nonlocal Cahn–Hilliard equation in 3D", *ESAIM Control Optim. Calc. Var.* **30** (2024), art. id. 21. MR Zbl

[van der Waals 1982] J. D. van der Waals, "Thermodynamische theorie der capillariteit in de onderstelling van continue dichtheidsverandering", *Verhandel. Konik. Akad. Weten. Amsterdam* **1**:8 (1982), 1–56. Translated in *J. Stat. Phys* **20**:2 (1979), 197–244. Zbl

Received 7 Aug 2022. Revised 3 May 2023. Accepted 20 Aug 2023.

ANDREA POIATTI: andrea.poiatti@polimi.it
Dipartimento di Matematica, Politecnico di Milano, Milano, Italy

ANALYSIS AND PDE
Vol. 18 (2025), No. 1, pp. 141–169
DOI: 10.2140/apde.2025.18.141

THE KATO SQUARE ROOT PROBLEM FOR WEIGHTED PARABOLIC OPERATORS

ALIREZA ATAEI, MORITZ EGERT AND KAJ NYSTRÖM

We give a simplified and direct proof of the Kato square root estimate for parabolic operators with elliptic part in divergence form and coefficients possibly depending on space and time in a merely measurable way. The argument relies on the nowadays classical reduction to a quadratic estimate and a Carleson-type inequality. The precise organization of the estimates is different from earlier works. In particular, we succeed in separating space and time variables almost completely despite the nonautonomous character of the operator. Hence, we can allow for degenerate ellipticity dictated by a spatial A_2-weight, which has not been treated before in this context.

1. Introduction and main result

In the variables $(x,t) \in \mathbb{R}^n \times \mathbb{R} =: \mathbb{R}^{n+1}$, we consider parabolic operators of the form

$$\mathcal{H}u := \partial_t u - w^{-1} \operatorname{div}_x(A \nabla_x u), \tag{1-1}$$

where the weight $w = w(x)$ is time-independent and belongs to the spatial Muckenhoupt class $A_2(\mathbb{R}^n, \mathrm{d}x)$, and the coefficient matrix $A = A(x,t)$ is measurable with complex entries and possibly depends on all variables. Degeneracy is dictated by the same weight w in the sense that $w^{-1}A$ satisfies the classical uniform ellipticity condition (Section 2.3).

Weighted parabolic operators as in (1-1) occur in various contexts and applications, including the study of fractional powers [Litsgård and Nyström 2023] and heat kernels of Schrödinger equations with singular potential [Ishige et al. 2017]. For contributions to the study of local properties of solutions to $\mathcal{H}u = 0$ and Gaussian estimates, we refer to [Chiarenza and Serapioni 1985; Cruz-Uribe and Rios 2008].

The purpose of this paper is to establish the Kato (square root) estimate for $\mathcal{H}$, that is, to prove Theorem 1.1 stated below. We write $\mathrm{L}^2_\mu = \mathrm{L}^2(\mathbb{R}^{n+1}, \mathrm{d}w\,\mathrm{d}t)$, $\mathrm{d}\mu = \mathrm{d}w\,\mathrm{d}t = w(x)\,\mathrm{d}x\,\mathrm{d}t$, for the natural weighted Lebesgue space associated with $\mathcal{H}$, and E_μ for the corresponding first-order parabolic Sobolev space of functions u such that the spatial gradient $\nabla_x u$, as well as the half-order time derivative $D_t^{1/2}u$, is in L^2_μ. For the sake of the introduction, an intuitive interpretation of these objects suffices. We turn to rigorous definitions in Section 3.

MSC2020: primary 35K10, 42B25, 42B37; secondary 26A33, 47B44.
Keywords: Kato square root problem, second-order parabolic operators, Muckenhoupt weights, square function estimates, fractional derivatives.

Theorem 1.1. *The operator $\mathcal{H}$ can be defined as an m-accretive operator in L^2_μ associated with an accretive sesquilinear form with domain E_μ. The domain of its unique m-accretive square root is the same as the form domain, that is $\mathsf{D}(\sqrt{\mathcal{H}}) = \mathrm{E}_\mu$, and*

$$\|\sqrt{\mathcal{H}}u\|_{\mathrm{L}^2_\mu} \sim \|\nabla_x u\|_{\mathrm{L}^2_\mu} + \|D_t^{1/2}u\|_{\mathrm{L}^2_\mu}, \quad u \in \mathrm{E}_\mu,$$

holds with an implicit constant that depends on the dimension, the ellipticity parameters of A and the A_2-constant for w.

The time derivative ∂_t is a skew-adjoint operator, and hence there are no lower bounds for the formal pairing $\operatorname{Re}\langle \mathcal{H}u, u\rangle$ that contain derivatives in t. However, when the time variable describes the full real line, parabolic operators admit some "hidden coercivity" that can be revealed through the Hilbert transform H_t in the t-variable. Indeed, splitting $\partial_t = D_t^{1/2} H_t D_t^{1/2}$, the sesquilinear form associated with (1-1) over L^2_μ is given by

$$B(u, v) := \iint_{\mathbb{R}^{n+1}} w^{-1} A\nabla_x u \cdot \overline{\nabla_x v} + H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} v} \,\mathrm{d}w\,\mathrm{d}t, \quad u, v \in \mathrm{E}_\mu, \tag{1-2}$$

and lower bounds including both time and space derivatives become available when taking $v = (1 + \delta H_t)u$ with $\delta > 0$ small. This observation is originally due to Kaplan [1966] and has been rediscovered several times ever since; see [Dier and Zacher 2017; Hofmann and Lewis 2005; Nyström 2016] for example. M-accretivity of $\mathcal{H}$ essentially follows from this observation, but to the best of our knowledge an explicit statement, in the unweighted case $w = 1$, only appeared much later in [Auscher and Egert 2016]. For the reader's convenience, we reproduce the full argument in our weighted setting in Section 4. Being m-accretive, $\mathcal{H}$ admits a sectorial functional calculus and in particular a (unique) m-accretive square root $\sqrt{\mathcal{H}}$; see [Haase 2006; Kato 1966] for background. This is how our main result should be understood.

The pursuit of the solution of the Kato problem for unweighted elliptic operators (finally completed in [Auscher et al. 2002]) introduced new techniques that proved extremely viable for extensions and applications to other problems in harmonic analysis and partial differential equations [Amenta and Auscher 2018; Alfonseca et al. 2011; Auscher and Axelsson 2011; Auscher and Mourgoglou 2019; Auscher and Rosén 2012; Auscher et al. 2018; Castro et al. 2016; Escauriaza and Hofmann 2018; Hofmann et al. 2015; 2019; 2022; Nyström 2017]. For this reason, Kato-type estimates for different operators are desirable, and the results of this paper most surely have important implications for, and applications to, boundary value problems for weighted second-order parabolic operators.

Let us mention that the case of A_2-weighted elliptic operators was settled in [Cruz-Uribe and Rios 2015], see also [Cruz-Uribe et al. 2018] for an extension, and rediscovered in the more general framework of first-order Dirac operators in [Auscher et al. 2015]. The third author was first to develop the underlying harmonic analysis in the unweighted parabolic setting in [Nyström 2016], and in the same paper he proved the square function estimates that are essentially equivalent to Theorem 1.1 when $w \equiv 1$ and when the coefficients A are t-independent. Using a framework of parabolic Dirac operators, Auscher, together with the second and third authors, obtained the unweighted parabolic case when coefficients depend measurably on x *and* t [Auscher et al. 2020]. Our Theorem 1.1 completes this succession of results but there is more to it and that makes, as we shall discuss next, the present paper interesting even in the unweighted case.

Under the assumption $A = A(x)$ in [Nyström 2016], the operator $\mathcal{H}$ is an autonomous parabolic operator, and, in retrospect, the main result of that paper could have been obtained by interpolation from maximal regularity of the Cauchy problem for (1-1); see [Ouhabaz 2021]. (In fact, this argument requires only smoothness of order $\frac{1}{2}$ for the coefficients in the t-variable.) However, many of the techniques in [Nyström 2016], such as the parabolic off-diagonal estimates and the construction of Tb-type test functions, had already been strong enough for proving the parabolic Kato estimate in the presence of measurable t-dependence, and our proof of Theorem 1.1 shows exactly how, thereby making our result novel in at least two ways:

- We generalize all previous findings in the parabolic setting by combining measurable dependence of the coefficients on all variables with A_2-weighted degeneracy in space.
- We avoid the Dirac operator framework in [Auscher et al. 2020]. The resulting "second-order" approach for parabolic operators with time-dependent measurable coefficients has not appeared in the literature before, and, when restricted to the unweighted case $w \equiv 1$, it provides a significant simplification of the proof of [Auscher et al. 2020, Theorem 2.6].

Our ambition is to present an almost self-contained argument using only a minimal number of tools. We do not attempt to generalize all further results in [Auscher et al. 2020] to the weighted setting, which should be done by developing a parabolic weighted Dirac operator framework.

As is customary in the field, see [Auscher et al. 2002; Cruz-Uribe and Rios 2015; Hofmann et al. 2022; Nyström 2016], the first reduction in the proof of Theorem 1.1 is to use the bounded H^∞-calculus for m-accretive operators and a duality argument in order to reduce the matter to the one-sided quadratic estimate

$$\int_0^\infty \|\lambda\mathcal{H}(1+\lambda^2\mathcal{H})^{-1}u\|_{\mathrm{L}^2_\mu}^2 \frac{\mathrm{d}\lambda}{\lambda} \lesssim \|\nabla_x u\|_{\mathrm{L}^2_\mu}^2 + \|D_t^{1/2}u\|_{\mathrm{L}^2_\mu}^2, \quad u \in \mathrm{E}_\mu. \tag{1-3}$$

In contrast to the elliptic setting, this reduction does not follow immediately from classical results à la Kato and Lions [Lions 1962], since the sesquilinear form B in (1-2) is *not* closed due to the lack of lower bounds by half-order time derivatives. Some more care is needed but we settle the issue in Section 6. The quadratic estimate (1-3) is then achieved by slightly refining the techniques in [Nyström 2016], and the argument relies on (weighted) Littlewood–Paley theory in L^2 (Section 5), which eventually reduces matters to a Carleson measure estimate that can be proved through a Tb-procedure (Section 8).

It came as a surprise to us that, even though coefficients may depend measurably on all variables, the proof of (1-3) can be arranged in a way that almost completely separates time and space variables. This observation incarnates in three different stages of the proof:

- At the level of Littlewood–Paley theory, it suffices to use weighted elliptic theory in x and classical Fourier analysis in t. The required weighted theory has already been developed in detail by Cruz-Uribe and Rios [2012; 2015] in order to solve the weighted elliptic Kato problem.

- At the level of off-diagonal bounds (averaged "kernel" bounds, see Section 4.4), we only need estimates for operators involving differentiation in space. These estimates can be deduced directly from the equation and come with parabolic scaling. The much more involved off-diagonal decay and Poincaré inequalities for nonlocal derivatives such as $D_t^{1/2}$, which were fundamental novelties in [Auscher et al. 2020], can be avoided.
- At the level of the Tb-argument, the test functions can be constructed based on a product structure, which makes the argument more straightforward compared to the system of functions used in [Auscher et al. 2020].

These three observations have in common that we can regroup derivatives of resolvents of $\mathcal{H}$ in such a way that fine harmonic analysis estimates need only apply to the spatial parts, whereas the t-derivatives appear in blocks that are amenable for simple resolvent estimates in L^2_μ-norm. We shall indicate the most striking examples of this principle along with the proof of the Carleson measure estimate in the final section.

The next section contains some preliminary notation and conventions. The rest of the paper follows the outline above.

2. Preliminaries and basic assumptions

Given $(x, t) \in \mathbb{R}^n \times \mathbb{R}$, we let $\|(x, t)\| := \max\{|x|, |t|^{1/2}\}$. We call $\|(x, t)\|$ the parabolic norm of (x, t). Given a half-open cube $Q = \big(x - \frac{1}{2}r, x + \frac{1}{2}r\big]^n \subset \mathbb{R}^n$ parallel to the coordinate axes with sidelength r and an interval $I = \big(t - \frac{1}{2}r^2, t + \frac{1}{2}r^2\big]$, we call $\Delta := Q \times I \subset \mathbb{R}^{n+1}$ a parabolic cube of size r. Occasionally, we write $\Delta_r(x, t) = Q_r(x) \times I_r(t)$ and $r = \ell(\Delta)$ to indicate the center and size directly. A dyadic parabolic cube of size 2^j is by definition centered in $(2^j\mathbb{Z})^n \times (4^j\mathbb{Z})$. For every $c > 0$, and given Δ, we define $c\Delta$ as the parabolic cube with the same center as Δ and size $c\ell(\Delta)$.

2.1. ***Assumptions and notation concerning the weight.*** For general background and the results cited here, we refer to [Stein 1993, Chapter V]. The weight $w = w(x)$ is a real-valued function belonging to the Muckenhoupt class $A_2(\mathbb{R}^n, \mathrm{d}x)$, that is,

$$[w]_{A_2} := \sup_Q \bigg(⨍_Q w \,\mathrm{d}x \bigg) \bigg(⨍_Q w^{-1} \,\mathrm{d}x \bigg) < \infty, \tag{2-1}$$

where the supremum is taken with respect to all cubes $Q \subset \mathbb{R}^n$. We introduce the measure $\mathrm{d}w(x) := w(x)\,\mathrm{d}x$ on $\mathbb{R}^n$, and we write $w(E) = \int_E \mathrm{d}w$ for all Lebesgue measurable sets $E \subset \mathbb{R}^n$. For averages, we use the notation

$$(g)_{E,w} := ⨍_E g(x)\,\mathrm{d}w(x) := \frac{1}{w(E)} \int_E g(x) w(x)\,\mathrm{d}x$$

if $w(E) \in (0, \infty)$ and g is locally integrable on $\mathbb{R}^n$ with respect to $\mathrm{d}w(x)$. It follows from (2-1) that there are constants $\eta \in (0, 1)$ and $\beta > 0$, depending only on n and $[w]_{A_2}$, such that

$$\beta^{-1} \bigg(\frac{|E|}{|Q|} \bigg)^{1/(2\eta)} \le \frac{w(E)}{w(Q)} \le \beta \bigg(\frac{|E|}{|Q|} \bigg)^{2\eta} \tag{2-2}$$

whenever $E \subset Q$ is measurable, and where $|\cdot|$ denotes Lebesgue measure in $\mathbb{R}^n$. In particular, there exists a constant D depending only on $[w]_{A_2}$ and n, called the doubling constant for w, such that

$$w(2Q) \leq Dw(Q) \quad \text{for all cubes } Q \subset \mathbb{R}^n. \tag{2-3}$$

The measures

$$\begin{aligned} \mathrm{d}\mu = \mathrm{d}\mu(x,t) &:= w(x)\,\mathrm{d}x\,\mathrm{d}t, \\ \mathrm{d}\mu^{-1} = \mathrm{d}\mu^{-1}(x,t) &:= w(x)^{-1}\,\mathrm{d}x\,\mathrm{d}t \end{aligned} \tag{2-4}$$

are defined on $\mathbb{R}^{n+1}$. Naturally, μ and μ^{-1} can be seen as measures on $\mathbb{R}^{n+1}$ defined by $A_2(\mathbb{R}^{n+1}, \mathrm{d}x\,\mathrm{d}t)$ weights, and in the context of these measures we use similar notation as above. The doubling constant for μ with respect to parabolic scaling is $4D$.

2.2. *Maximal functions.* We introduce the maximal operators in the individual variables

$$\begin{aligned} \mathcal{M}^{(1)}(g_1)(x) &:= \sup_{r>0} \fint_{Q_r(x)} |g_1|\,\mathrm{d}x, \\ \mathcal{M}^{(2)}(g_2)(t) &:= \sup_{r>0} \fint_{I_r(t)} |g_2|\,\mathrm{d}t \end{aligned}$$

for all locally integrable functions g_1 and g_2 on $\mathbb{R}^n$ and $\mathbb{R}$, respectively. The operator $\mathcal{M}^{(1)}$ is bounded on the weighted space $\mathrm{L}^2(\mathbb{R}^n, \mathrm{d}w)$ with a bound depending on $[w]_{A_2}$ and n [Stein 1993, Theorem 1, p. 201]. Both maximal operators can be naturally extended to $\mathrm{L}^2(\mathbb{R}^{n+1}, \mathrm{d}\mu)$ by keeping one of the variables fixed, and they are bounded in this setting.

2.3. *Assumptions on the coefficients.* The matrix-valued function

$$A = A(x,t) = \{A_{i,j}(x,t)\}_{i,j=1}^n$$

is assumed to have complex measurable entries $A_{i,j}$ that satisfy

$$c_1|\xi|^2 w(x) \leq \mathrm{Re}(A(x,t)\xi \cdot \bar{\xi}), \quad |A(x,t)\xi \cdot \zeta| \leq c_2 w(x)|\xi||\zeta| \tag{2-5}$$

for some $c_1, c_2 \in (0,\infty)$ and for all $\xi, \zeta \in \mathbb{C}^n$, $(x,t) \in \mathbb{R}^{n+1}$. Here, $u \cdot v = u_1v_1 + \cdots + u_nv_n$, and $\bar{u}$ denotes the complex conjugate of u so that $u \cdot \bar{v}$ is the standard inner product on $\mathbb{C}^n$. We refer to c_1, c_2 as the ellipticity constants of A. Assumption (2-5) is equivalent to saying that $w^{-1}A$ satisfies the classical uniform ellipticity condition.

2.4. *Floating constants.* We refer to n and the constants $[w]_{A_2}$, c_1, c_2, appearing in (2-1) and (2-5), as structural constants. For $A, B \in (0,\infty)$, the notation $A \lesssim B$ means that $A \leq cB$ for some c depending at most on the structural constants. The notation $A \gtrsim B$ and $A \sim B$ should be interpreted similarly.

3. Weighted function spaces

In this section we give a brief account of the relevant weighted function spaces. We let $\mathrm{L}^2_w = \mathrm{L}^2(\mathbb{R}^n, \mathrm{d}w)$ be the Hilbert spaces of square integrable functions with respect to $\mathrm{d}w$. Its norm is denoted by $\|\cdot\|_{2,w}$, its inner product by $\langle \cdot , \cdot \rangle_{2,w}$, and the operator norm of linear operators on that space by $\| \cdot \|_{2\to 2,w}$. Thanks to the A_2-condition, we have

$$\mathrm{L}^2_w \subset \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^n, \mathrm{d}x), \tag{3-1}$$

and the class $\mathrm{C}^\infty_0(\mathbb{R}^n)$ of smooth and compactly supported test functions is dense in L^2_w via the usual truncation and convolution procedure [Kilpeläinen 1994, Section 1]. The same notation and properties apply to L^2_μ in $\mathbb{R}^{n+1}$.

Definition 3.1 (elliptic weighted Sobolev space). We write $\mathrm{H}^1_w := \mathrm{H}^1_w(\mathbb{R}^n)$ for the space of all $f \in \mathrm{L}^2_w$ for which the distributional gradient $\nabla_x f$ is (componentwise) in L^2_w, and we equip the space with the norm $\| \cdot \|_{\mathrm{H}^1_w} := (\| \cdot \|^2_{2,w} + \|\nabla_x \cdot \|^2_{2,w})^{1/2}$.

By construction H^1_w is a Hilbert space, and standard truncation and convolution techniques yield that $\mathrm{C}^\infty_0(\mathbb{R}^n)$ is dense in H^1_w; see [Kilpeläinen 1994, Theorem 2.5].

In order to define parabolic function spaces, we use the Fourier transform $\mathcal{F}$ in the time variable, keeping in mind that if $f \in \mathrm{L}^2(\mathbb{R}^{n+1}, \mathrm{d}\mu)$, then $f(x, \cdot) \in \mathrm{L}^2(\mathbb{R}, \mathrm{d}t)$ for a.e. $x \in \mathbb{R}^n$. The corresponding Fourier variable will be denoted by τ. Then,

$$H_t f := \mathcal{F}^{-1}(\mathrm{i}\,\mathrm{sgn}(\tau)\mathcal{F} f)$$

is our Hilbert transform. If $|\tau|^{1/2}\mathcal{F} f \in \mathrm{L}^2_\mu$, then we define the half-order time derivative

$$D^{1/2}_t f := \mathcal{F}^{-1}(|\tau|^{1/2}\mathcal{F} f),$$

and this is what we mean when we write $D^{1/2}_t f \in \mathrm{L}^2_\mu$. Using a classical formula for fractional Laplacians for a.e. fixed $x \in \mathbb{R}^n$, see [Di Nezza et al. 2012] for example, we obtain

$$\|D^{1/2}_t f\|^2_{2,\mu} = \frac{2}{\pi} \int_{\mathbb{R}^n}\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{|f(x,t) - f(x,s)|^2}{|t-s|^2} \,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}w(x), \tag{3-2}$$

with the right-hand side being finite precisely when $D^{1/2}_t f \in \mathrm{L}^2_\mu$.

Definition 3.2 (parabolic energy space). We write $\mathrm{E}_\mu := \mathrm{E}_\mu(\mathbb{R}^{n+1})$ for the space of all $f \in \mathrm{L}^2_\mu$ for which $\nabla_x f$, $D^{1/2}_t f \in \mathrm{L}^2{}_\mu$, and we equip the space with the norm

$$\| \cdot \|_{\mathrm{E}_\mu} := (\| \cdot \|^2_{2,\mu} + \|\nabla_x \cdot \|^2_{2,\mu} + \|D^{1/2}_t \cdot \|^2_{2,\mu})^{1/2}.$$

For $f \in \mathrm{E}_\mu$, we will refer to the vector $\mathbb{D} f := (\nabla_x f, D^{1/2}_t f)$ as the parabolic gradient of f.

Again, E_μ is a Hilbert space. Note that, in the unweighted setting of [Nyström 2016], the notation $\mathbb{D}$ has a slightly different meaning.

Lemma 3.3. *The following statements are true*:

(i) *The space* $\mathrm{C}^\infty_0(\mathbb{R}^{n+1})$ *is dense in* $\mathrm{E}_\mu(\mathbb{R}^{n+1})$.

(ii) *Multiplication by* $\mathrm{C}^1(\mathbb{R}^{n+1})$*-functions is bounded on* $\mathrm{E}_\mu(\mathbb{R}^{n+1})$.

Proof. We begin with (i). If $f \in \mathrm{E}_\mu$, then convolutions with smooth mollifiers, separately in x and t, provide smooth approximations in E_μ. For the convolution in space, this argument uses the A_2-condition on w as mentioned above. Hence, it suffices to approximate f by compactly supported functions in E_μ. To this end, we can follow the standard pattern of smooth truncation: We pick a sequence $(\eta_j)_j \subset \mathrm{C}_0^\infty(\mathbb{R}^{n+1})$ such that $\eta_j \to 1$ pointwise a.e. as $j \to \infty$, $\|\eta_j\|_\infty + j\|\nabla_x \eta_j\|_\infty + j\|\partial_t \eta_j\|_\infty \le c$ uniformly in j, and then we set $f_j := \eta_j f$. By dominated convergence, we obtain $f_j \to f$ and $\nabla_x f_j \to \nabla_x f$ in L^2_μ as $j \to \infty$. For the half-order derivative, we use (3-2) with $f_j - f$ in place of f. We first bound the integrand in (3-2) by

$$\begin{aligned}
&\frac{|(f_j - f)(x,t) - (f_j - f)(x,s)|^2}{|t-s|^2} \\
&\quad \le 2\frac{|(\eta_j - 1)(x,t) - (\eta_j - 1)(x,s)|^2}{|t-s|^2}|f(x,t)|^2 + 2\frac{|f(x,t) - f(x,s)|^2}{|t-s|^2}|(\eta_j - 1)(x,s)|^2 \\
&\quad \le 2\min\left\{c^2, \frac{4(c+1)^2}{|t-s|^2}\right\}|f(x,t)|^2 + 2(c+1)^2\frac{|f(x,t) - f(x,s)|^2}{|t-s|^2}.
\end{aligned} \tag{3-3}$$

The right-hand side is independent of j and integrable with respect to $\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}w(x)$. Since the middle term tends to 0 a.e. as $j \to \infty$, we conclude

$$\|D_t^{1/2}(f_j - f)\|_{2,\mu} \to 0$$

by dominated convergence. This completes the proof of (i).

As for (ii), we note that if $\eta \in \mathrm{C}^1(\mathbb{R}^{n+1})$ and $f \in \mathrm{E}_\mu$, then

$$\begin{aligned}
\|\eta f\|_{2,\mu} &\le \|\eta\|_\infty \|f\|_{2,\mu}, \\
\|\nabla_x(\eta f)\|_{2,\mu} &\le \|\eta\|_\infty \|\nabla_x f\|_{2,\mu} + \|\nabla_x \eta\|_\infty \|f\|_{2,\mu}, \\
\|D_t^{1/2}(\eta f)\|_{2,\mu} &\le \sqrt{8}\|\eta\|_\infty^{1/2}\|\partial_t \eta\|_\infty^{1/2}\|f\|_{2,\mu} + \|\eta\|_\infty \|D_t^{1/2} f\|_{2,\mu},
\end{aligned}$$

where the third line follows by the same splitting as in (3-3), but with η in place of $1 - \eta_j$. $\square$

Lemma 3.3 (i) implies the chain of continuous and dense embeddings

$$\mathrm{E}_\mu \subset \mathrm{L}^2_\mu \simeq (\mathrm{L}^2_\mu)^* \subset (\mathrm{E}_\mu)^*, \tag{3-4}$$

where we use the upper star to denote (anti)-dual spaces. We have bounded operators

$$\begin{aligned}
D_t^{1/2} &: \mathrm{E}_\mu \to \mathrm{L}^2_\mu, \\
\nabla_x &: \mathrm{E}_\mu \to (\mathrm{L}^2_\mu)^n,
\end{aligned} \tag{3-5}$$

and we denote their adjoints with respect to (3-4) by

$$\begin{aligned}
D_t^{1/2} &: \mathrm{L}^2_\mu \to (\mathrm{E}_\mu)^*, \\
w^{-1}\operatorname{div}_x w &: (\mathrm{L}^2_\mu)^n \to (\mathrm{E}_\mu)^*.
\end{aligned} \tag{3-6}$$

Note carefully that $w^{-1}\operatorname{div}_x w$ is only a suggestive notation reflecting the formal action of this operator. In general, there is no guarantee that this operator splits into a composition of its three building blocks.

4. The parabolic operator

We continue by introducing the formal parabolic operator in (1-1) rigorously as an unbounded operator in the Hilbert space L^2_μ associated with a sesquilinear form.

Denoting by H_t the Hilbert transform in the t-variable and by $D_t^{1/2}$ the half-order time derivative as defined in Section 3, we can factorize

$$\partial_t = D_t^{1/2} H_t D_t^{1/2}.$$

By (3-5) and (3-6), we have $\partial_t : \mathrm{E}_\mu \to (\mathrm{E}_\mu)^*$. We define $\mathcal{H}$ as a bounded operator $\mathrm{E}_\mu \to (\mathrm{E}_\mu)^*$ via

$$(\mathcal{H}u)(v) := B(u, v) := \iint_{\mathbb{R}^{n+1}} w^{-1} A\nabla_x u \cdot \overline{\nabla_x v} + H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} v} \,\mathrm{d}\mu, \quad u, v \in \mathrm{E}_\mu. \tag{4-1}$$

In view of (3-4), it makes sense to consider the maximal restriction of $\mathcal{H}$ to an operator in L^2_μ, called the part of $\mathcal{H}$ in L^2_μ, with domain

$$\mathsf{D}(\mathcal{H}) := \{u \in \mathrm{E}_\mu(\mathbb{R}^{n+1}) : \mathcal{H}u \in \mathrm{L}^2_\mu(\mathbb{R}^{n+1})\}. \tag{4-2}$$

If $u \in \mathsf{D}(\mathcal{H})$, we have, for all $v \in \mathrm{E}_\mu$, that

$$(\mathcal{H}u)(v) = \iint_{\mathbb{R}^{n+1}} \mathcal{H}u \cdot \bar{v} \,\mathrm{d}\mu,$$

and a formal integration by parts in (4-1) reveals that it is indeed justified to say that the part of $\mathcal{H}$ in L^2_μ gives meaning to the formal expression in (1-1). More precisely, in terms of (3-5) and (3-6), we have that $\mathcal{H} : \mathrm{E}_\mu \to (\mathrm{E}_\mu)^*$ acts as the composition of operators

$$\mathcal{H} = D_t^{1/2} H_t D_t^{1/2} - (w^{-1} \operatorname{div}_x w)(w^{-1} A\nabla_x). \tag{4-3}$$

4.1. *Hidden coercivity.* The following lemma relies on the hidden coercivity (proved by Kaplan [1966]) of the parabolic sesquilinear form B in (4-1) that can be revealed through the Hilbert transform.

Lemma 4.1. *Let $\sigma \in \mathbb{C}$ with $\operatorname{Re}\sigma > 0$. For each $f \in (\mathrm{E}_\mu)^*$, there exists a unique $u \in \mathrm{E}_\mu$ such that $(\sigma + \mathcal{H})u = f$. Moreover,*

$$\|u\|_{\mathrm{E}_\mu} \le \sqrt{2} \max\left\{\frac{c_2 + 1}{c_1}, \frac{|\operatorname{Im}\sigma| + 1}{\operatorname{Re}\sigma}\right\} \|f\|_{(\mathrm{E}_\mu)^*}.$$

Proof. By Plancherel's theorem, the Hilbert transform H_t is isometric on E_μ. Hence, we can define a "twisted" sesquilinear form $B_{\delta,\sigma} : \mathrm{E}_\mu \times \mathrm{E}_\mu \to \mathbb{C}$ via

$$\begin{aligned} B_{\delta,\sigma}(u, v) := \iint_{\mathbb{R}^{n+1}} \big(\sigma u \cdot \overline{(1+\delta H_t)v} + w^{-1} A\nabla_x u \cdot \overline{\nabla_x (1+\delta H_t) v} \\ + H_t D_t^{1/2} u \cdot \overline{D_t^{1/2}(1+\delta H_t) v}\big) \,\mathrm{d}\mu, \end{aligned} \tag{4-4}$$

where $\delta \in (0, 1)$ is to be chosen. Clearly $B_{\delta,\sigma}$ is bounded. Since H_t is skew-adjoint, we have

$$\operatorname{Re} \iint_{\mathbb{R}^{n+1}} H_t v \cdot \bar{v} \,\mathrm{d}\mu = 0, \quad v \in \mathrm{L}^2_\mu. \tag{4-5}$$

Expanding $B_{\delta,\sigma}(u, u)$ and using the above along with the weighted ellipticity of the coefficients A, we find

$$\operatorname{Re} B_{\delta,\sigma}(u, u) \geq \delta \|D_t^{1/2} u\|_{2,\mu}^2 + (c_1 - c_2\delta)\|\nabla_x u\|_{2,\mu}^2 + (\operatorname{Re}\sigma - \delta|\operatorname{Im}\sigma|)\|u\|_{2,\mu}^2. \tag{4-6}$$

Choosing $\delta = \min\{c_1/(c_2 + 1), \operatorname{Re}\sigma/(|\operatorname{Im}\sigma| + 1)\}$, the factors in front of the second and third term in the last display become no less than δ. Hence, we obtain the coercivity estimate

$$\operatorname{Re} B_{\delta,\sigma}(u, u) \geq \min\left\{\frac{c_1}{c_2 + 1}, \frac{\operatorname{Re}\sigma}{|\operatorname{Im}\sigma| + 1}\right\}\|u\|_{\mathrm{E}_\mu}^2, \quad v \in \mathrm{E}_\mu. \tag{4-7}$$

The Lax–Milgram lemma yields, for each $f \in (\mathrm{E}_\mu)^*$, a unique $u \in \mathrm{E}_\mu$ satisfying the estimate claimed in the lemma such that

$$B_{\delta,\sigma}(u, v) = f((1 + \delta H_t)v), \quad v \in \mathrm{E}_\mu.$$

(Note that the additional factor $\sqrt{2}$ is an upper bound for the norm of $1 + \delta H_t$ on E_μ.) Plancherel's theorem yields that $1 + \delta H_t$ is an isomorphism on E_μ for all $\delta \in \mathbb{R}$. Thus,

$$\iint_{\mathbb{R}^{n+1}} \sigma u \cdot \bar{v} + w^{-1} A \nabla_x u \cdot \overline{\nabla_x v} + H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} v} \,\mathrm{d}\mu = f(v), \quad v \in \mathrm{E}_\mu,$$

that is, $(\sigma + \mathcal{H})u = f$ as required. □

The proof above fails for $\delta = 0$ since $\operatorname{Re} B(\cdot\,,\cdot)$ does not control $\|D_t^{1/2} \cdot\|_{2,\mu}$ from above. As a consequence, B itself is *not* a closed sesquilinear form in the sense of Kato [1966] or, equivalently, $(\|\cdot\|_{2,\mu}^2 + \operatorname{Re} B(\cdot\,,\cdot))^{1/2}$ does *not* define an equivalent norm on E_μ. In [Auscher and Egert 2016, Lemma 4], it has been (essentially) shown that a parabolic analog of Kato's first representation theorem holds nonetheless. For convenience, we include the short argument with some minor improvements in the next section.

4.2. *M-accretivity.* Recall that an operator $\mathcal{H}$ in a Hilbert space such as L^2_μ is called *m-accretive* if it is closed and densely defined, with resolvent estimates

$$\|(\sigma + \mathcal{H})^{-1}\|_{2\to 2,\mu} \leq (\operatorname{Re}\sigma)^{-1}, \quad \sigma \in \mathbb{C}, \ \operatorname{Re}\sigma > 0.$$

Proposition 4.2. *The part of $\mathcal{H}$ in L^2_μ is m-accretive and $\mathsf{D}(\mathcal{H})$ is dense in E_μ.*

Proof. Fix $\sigma \in \mathbb{C}$ with $\operatorname{Re}\sigma > 0$. Lemma 4.1 yields that $\sigma + \mathcal{H} : \mathsf{D}(\mathcal{H}) \to \mathrm{L}^2_\mu$ is bijective. Given $f \in \mathrm{L}^2_\mu$, we set $u := (\sigma + \mathcal{H})^{-1} f$ and use ellipticity of the coefficients A and (4-5) to deduce

$$\begin{aligned}\operatorname{Re}\sigma \|u\|_{2,\mu}^2 &\leq \operatorname{Re} \iint_{\mathbb{R}^{n+1}} \sigma u \cdot \bar{u} + w^{-1} A \nabla_x u \cdot \overline{\nabla_x u} + H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} u} \,\mathrm{d}\mu \\ &= \operatorname{Re} \iint_{\mathbb{R}^{n+1}} f \cdot \bar{u} \,\mathrm{d}\mu \leq \|f\|_{2,\mu} \|u\|_{2,\mu}.\end{aligned}$$

This gives the required resolvent bound $\|u\|_{2,\mu} \leq (\operatorname{Re}\sigma)^{-1}\|f\|_{2,\mu}$. Moreover, the part of $\mathcal{H}$ in L^2_μ is closed since it has a nonempty resolvent set, and the resolvent estimates for $\sigma > 0$ imply a dense domain [Haase 2006, Proposition 2.1.1]. This proves m-accretivity.

In order to prove that $\mathsf{D}(\mathcal{H})$ is dense in E_μ, we use the sesquilinear form $B_{\delta,1}$ in (4-4) with $\delta > 0$ chosen as in the proof of that lemma. Suppose $v \in \mathrm{E}_\mu$ is orthogonal to $\mathsf{D}(\mathcal{H})$ in E_μ. By the Lax–Milgram lemma, there is $w \in \mathrm{E}_\mu$ such that $\langle u, v\rangle_{\mathrm{E}_\mu} = B_{\delta,1}(u, w)$ for all $u \in \mathrm{E}_\mu$. For $u \in \mathsf{D}(H)$, this identity becomes $0 = \langle (1+\mathcal{H})u, (1+\delta H_t)w\rangle_{2,\mu}$, and since $1+\mathcal{H} : \mathsf{D}(\mathcal{H}) \to \mathrm{L}^2_\mu$ is bijective, we conclude that $(1+\delta H_t)w = 0$. Thus, we have $w = 0$ and therefore also $v = 0$. □

The adjoint $\mathcal{H}^*$ of $\mathcal{H}$ (seen as either a bounded operator $\mathrm{E}_\mu \to (\mathrm{E}_\mu)^*$ or an unbounded operator in L^2_μ) has the same properties as $\mathcal{H}$. Indeed, it can be checked by the very definition that it is associated with the sesquilinear form

$$B^*(u, v) = \overline{B(v, u)}$$

and that it formally corresponds to the backward-in-time operator

$$-\partial_t - w^{-1}(x)\operatorname{div}_x(A^*(x, t)\nabla_x).$$

Here A^* is the conjugate transpose of A.

4.3. *Resolvent estimates.* Using Proposition 4.2, we see that, for $\lambda > 0$, the resolvent operators

$$\begin{aligned}\mathcal{E}_\lambda &:= (I+\lambda^2\mathcal{H})^{-1},\\ \mathcal{E}^*_\lambda &:= (I+\lambda^2\mathcal{H}^*)^{-1}\end{aligned} \tag{4-8}$$

are well defined as bounded operators $\mathrm{L}^2_\mu \to \mathrm{L}^2_\mu$ and $(\mathrm{E}_\mu)^* \to \mathrm{E}_\mu$. Moreover, they are adjoints of each other.

Lemma 4.3. *The following resolvent estimates hold uniformly for all $\lambda > 0$, all $f \in \mathrm{L}^2_\mu$ and all $\boldsymbol{f} \in (\mathrm{L}^2_\mu)^n$:*

(i) $\|\mathcal{E}_\lambda f\|_{2,\mu} + \|\lambda\mathbb{D}\mathcal{E}_\lambda f\|_{2,\mu} \lesssim \|f\|_{2,\mu}$,

(ii) $\|\lambda\mathcal{E}_\lambda D_t^{1/2} f\|_{2,\mu} + \|\lambda^2\mathbb{D}\mathcal{E}_\lambda D_t^{1/2} f\|_{2,\mu} \lesssim \|f\|_{2,\mu}$,

(iii) $\|\lambda\mathcal{E}_\lambda w^{-1}\operatorname{div}_x(w\boldsymbol{f})\|_{2,\mu} + \|\lambda^2\mathbb{D}\mathcal{E}_\lambda w^{-1}\operatorname{div}_x(w\boldsymbol{f})\|_{2,\mu} \lesssim \|\boldsymbol{f}\|_{2,\mu}$.

The same estimates hold with $\mathcal{E}_\lambda$ replaced by $\mathcal{E}^*_\lambda$.

Proof. We first prove (i). Setting $u := (\lambda^{-2}+\mathcal{H})^{-1}f$, we have $\mathcal{E}_\lambda f = \lambda^{-2}u$, and by m-accretivity we obtain

$$\|\mathcal{E}_\lambda f\|_{2,\mu} \le \|f\|_{2,\mu}.$$

Next, we use the twisted sesquilinear form $B_{\delta,\sigma}$ as in (4-4) with parameter $\sigma = \lambda^{-2}$, so that by construction

$$B_{\delta,\sigma}(u, u) = \langle f, (1+\delta H_t)u\rangle_{2,\mu}. \tag{4-9}$$

With this choice for σ, we pick $\delta = c_1/(2c_2)$, use (4-6) on the left, and Cauchy–Schwarz on the right, in order to obtain

$$\|\mathbb{D}u\|^2_{2,\mu} \lesssim \|f\|_{2,\mu}\|u\|_{2,\mu} \le \lambda^2\|f\|^2_{2,\mu}.$$

This is the required uniform bound for $\lambda\mathbb{D}\mathcal{E}_\lambda f$. Since $\mathcal{H}$ is of the same type as $\mathcal{H}^*$ from the point of view of sesquilinear forms, the same estimates also hold for $\mathcal{E}^*_\lambda$ in place of $\mathcal{E}_\lambda$.

Next, we note that the estimates for the leftmost terms in (ii) and (iii) follow by duality from (i) applied to $\mathcal{E}_\lambda^*$.

In order to estimate the second term on the left in (ii), we set $u := (\lambda^{-2} + \mathcal{H})^{-1} D_t^{1/2} f$. Since $D_t^{1/2} f$ is now regarded as an element in $(\mathrm{E}_\mu)^*$, we get $\langle f, D_t^{1/2}(1 + \delta H_t)u\rangle_{2,\mu}$ on the right-hand side in (4-9), and from this we conclude

$$\|\mathbb{D}u\|_{2,\mu}^2 \lesssim \|f\|_{2,\mu} \|\mathbb{D}u\|_{2,\mu},$$

as required. The remaining term in (iii) is estimated in the same way upon replacing $D_t^{1/2} f$ by $w^{-1} \operatorname{div}_x(w\boldsymbol{f})$. □

4.4. ***Off-diagonal estimates.*** Given measurable subsets E and F of $\mathbb{R}^{n+1}$, we let

$$d(E, F) := \inf\{\|(x - y, t - s)\| : (x, t) \in E,\ (y, s) \in F\}$$

denote their parabolic distance. Lemma 4.4 below is an improvement of the uniform bounds in Lemma 4.3. We only state and prove Lemma 4.4 for the families of operators that will require it later. However, let us stress that such estimates are not to be expected in the presence of the nonlocal operator $D_t^{1/2}$, and one of the insights in [Auscher et al. 2020] was that in this case a nonlocal version of off-diagonal bounds should be used.

Lemma 4.4. *Assume that E and F are measurable subsets of $\mathbb{R}^{n+1}$, and let $d := d(E, F)$. Then, there exists a constant $c \in (0, \infty)$, depending only on the structural constants, such that*

(i) $\displaystyle\iint_F |\mathcal{E}_\lambda f|^2 + |\lambda \nabla_x \mathcal{E}_\lambda f|^2 \,\mathrm{d}\mu \lesssim e^{-d/(c\lambda)} \iint_E |f|^2 \,\mathrm{d}\mu,$

(ii) $\displaystyle\iint_F |\lambda \mathcal{E}_\lambda w^{-1} \operatorname{div}_x(w\boldsymbol{f})|^2 \,\mathrm{d}\mu \lesssim e^{-d/(c\lambda)} \iint_E |\boldsymbol{f}|^2 \,\mathrm{d}\mu$

for all $\lambda > 0$ and all $f \in \mathrm{L}^2_\mu$, $\boldsymbol{f} \in (\mathrm{L}^2_\mu)^n$ with support in E. The same statements are true when $\mathcal{E}_\lambda$ is replaced by $\mathcal{E}_\lambda^*$.

Proof. As in the proof of Lemma 4.3, it suffices to treat $\mathcal{E}_\lambda$. Based on Lemma 4.3, we see that it suffices to obtain the exponential estimate for $0 < \lambda \le \alpha d$, where for now $\alpha \in (0, 1)$ is a degree of freedom that will be determined later and which will only depend on the structural constants.

Let $u := \mathcal{E}_\lambda f$, and recall that

$$\iint_{\mathbb{R}^{n+1}} u\bar{v} + \lambda^2 w^{-1} A\nabla_x u \cdot \overline{\nabla_x v} + \lambda^2 H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} v} \,\mathrm{d}\mu = \iint_{\mathbb{R}^{n+1}} f \cdot \bar{v} \,\mathrm{d}\mu \tag{4-10}$$

for all $v \in \mathrm{E}_\mu$. We can pick a real-valued $\tilde{\eta} \in \mathrm{C}^\infty(\mathbb{R}^{n+1})$ such that $\tilde{\eta} = 1$ on F, $\tilde{\eta} = 0$ on E, and such that

$$d|\nabla_x \tilde{\eta}| + d^2 |\partial_t \tilde{\eta}| \le c$$

for some constant c only depending on n. The different scaling in the two terms is due to the definition of the parabolic distance. Next, we let

$$v := u\eta^2 \quad \text{with } \eta := e^{(\alpha d/\lambda)\tilde{\eta}} - 1. \tag{4-11}$$

For this choice of v, we rewrite the real part in (4-10) of the pairing containing half-order derivatives as follows. According to Lemma 3.3, there exists a sequence $\{u_i\} \subset \mathrm{C}_0^\infty(\mathbb{R}^{n+1})$ such that $u_i \to u$ in E_μ as $i \to \infty$. By the same lemma, $\eta^2 u_i \to \eta^2 u$ in E_μ, and therefore

$$\begin{aligned}
\operatorname{Re} \iint_{\mathbb{R}^{n+1}} H_t D_t^{1/2} u \cdot \overline{D_t^{1/2} v} \, \mathrm{d}\mu &= \operatorname{Re} \lim_{i\to\infty} \iint_{\mathbb{R}^{n+1}} H_t D_t^{1/2} u_i \cdot \overline{D_t^{1/2}(u_i \eta^2)} \, \mathrm{d}t \, \mathrm{d}w \\
&= \lim_{i\to\infty} \operatorname{Re} \iint_{\mathbb{R}^{n+1}} \partial_t u_i \cdot \overline{u_i \eta^2} \, \mathrm{d}t \, \mathrm{d}w \\
&= \frac{1}{2} \lim_{i\to\infty} \iint_{\mathbb{R}^{n+1}} \partial_t |u_i|^2 \cdot \eta^2 \, \mathrm{d}t \, \mathrm{d}w \\
&= \frac{1}{2} \lim_{i\to\infty} - \iint_{\mathbb{R}^{n+1}} |u_i|^2 \cdot \partial_t(\eta^2) \, \mathrm{d}t \, \mathrm{d}w \\
&= -\frac{1}{2} \iint_{\mathbb{R}^{n+1}} |u|^2 \cdot \partial_t(\eta^2) \, \mathrm{d}\mu.
\end{aligned}$$

Going back to (4-10) and using that $\eta = 0$ on E, we conclude that

$$\operatorname{Re} \iint_{\mathbb{R}^{n+1}} |u|^2 \eta^2 \, \mathrm{d}\mu + \lambda^2 w^{-1} A \nabla_x u \cdot \overline{\nabla_x(u\eta^2)} - \tfrac{1}{2}\lambda^2 |u|^2 \partial_t(\eta^2) \, \mathrm{d}\mu = 0.$$

Using this identity and ellipticity, we deduce

$$\begin{aligned}
&\iint_{\mathbb{R}^{n+1}} |u|^2 \eta^2 \, \mathrm{d}\mu + c_1 \lambda^2 \iint_{\mathbb{R}^{n+1}} |\nabla_x u|^2 \eta^2 \, \mathrm{d}\mu \\
&\quad \le \lambda^2 \iint_{\mathbb{R}^{n+1}} |u|^2 |\eta| \, |\partial_t \eta| \, \mathrm{d}\mu + 2c_2 \lambda^2 \iint_{\mathbb{R}^{n+1}} |u| \, |\nabla_x u| \, |\eta| \, |\nabla_x \eta| \, \mathrm{d}\mu \\
&\quad \le \frac{1}{2} \iint_{\mathbb{R}^{n+1}} |u|^2 \eta^2 \, \mathrm{d}\mu + \frac{1}{2}\lambda^4 \iint_{\mathbb{R}^{n+1}} |u|^2 |\partial_t \eta|^2 \, \mathrm{d}\mu + \frac{1}{2} c_1 \lambda^2 \iint_{\mathbb{R}^{n+1}} |\nabla_x u|^2 \eta^2 \, \mathrm{d}\mu \\
&\qquad + 2\frac{c_2^2}{c_1} \lambda^2 \iint_{\mathbb{R}^{n+1}} |u|^2 |\nabla_x \eta|^2 \, \mathrm{d}\mu.
\end{aligned}$$

In conclusion,

$$\iint_{\mathbb{R}^{n+1}} |u|^2 \eta^2 \, \mathrm{d}\mu + c_1 \lambda^2 \iint_{\mathbb{R}^{n+1}} |\nabla_x u|^2 \eta^2 \, \mathrm{d}\mu \le \iint_{\mathbb{R}^{n+1}} |u|^2 \Big(\lambda^4 |\partial_t \eta|^2 + 4\frac{c_2^2}{c_1} \lambda^2 |\nabla_x \eta|^2 \Big) \, \mathrm{d}\mu.$$

By the definition of η in (4-11) and since $\lambda \le \alpha d \le d$, we see that

$$|\partial_t \eta|^2 \le \frac{\alpha^2 d^2}{\lambda^2} |\eta + 1|^2 \frac{c^2}{d^4} \le c^2 \alpha^2 \lambda^{-4} |\eta + 1|^2$$

and

$$|\nabla_x \eta|^2 \le \frac{\alpha^2 d^2}{\lambda^2} |\eta + 1|^2 \frac{c^2}{d^2} = c^2 \alpha^2 \lambda^{-2} |\eta + 1|^2.$$

Thus, we get

$$\iint_{\mathbb{R}^{n+1}} |u|^2 \eta^2 \, \mathrm{d}\mu + c_1 \lambda^2 \iint_{\mathbb{R}^{n+1}} |\nabla_x u|^2 \eta^2 \, \mathrm{d}\mu \lesssim \alpha^2 \iint_{\mathbb{R}^{n+1}} |u|^2 |\eta + 1|^2 \, \mathrm{d}\mu.$$

At this point, we make our choice of α. Indeed, using the bound $|\eta+1|^2 \le 2(\eta^2+1)$, we choose α small enough to be able to absorb the part coming from η into the left-hand side. The conclusion is that

$$\iint_{\mathbb{R}^{n+1}} |u|^2\eta^2 \, \mathrm{d}\mu + \lambda^2 \iint_{\mathbb{R}^{n+1}} |\nabla_x u|^2 \eta^2 \, \mathrm{d}\mu \lesssim \iint_{\mathbb{R}^{n+1}} |u|^2 \, \mathrm{d}\mu.$$

On the right-hand side, we can use Lemma 4.3 (i), and, on the left-hand side, we exploit that on F we have

$$\eta = e^{\alpha d/\lambda} - 1 \ge \tfrac{1}{2} e^{\alpha d/\lambda}$$

since we are assuming $\lambda \le \alpha d$. Consequently,

$$e^{2\alpha d/\lambda} \iint_F |u|^2 \, \mathrm{d}\mu + e^{2\alpha d/\lambda} \iint_F |\lambda \nabla_x u|^2 \, \mathrm{d}\mu \lesssim \iint_E |f|^2 \, \mathrm{d}\mu,$$

which proves (i).

The inequality in (ii) follows by a duality argument, using (i) for $\mathcal{E}_\lambda^*$ and interchanging the roles of E and F. In fact,

$$\begin{aligned}\iint_F |\lambda \mathcal{E}_\lambda w^{-1} \operatorname{div}_x(w\boldsymbol{f})|^2 \, \mathrm{d}\mu &= \sup_g \biggl(\iint_{\mathbb{R}^{n+1}} \lambda \mathcal{E}_\lambda w^{-1} \operatorname{div}_x(w\boldsymbol{f}) \cdot \bar{g} \, \mathrm{d}\mu \biggr)^2 \\ &= \sup_g \biggl(\iint_E -\boldsymbol{f} \cdot \overline{\lambda \nabla_x \mathcal{E}_\lambda^* g} \, \mathrm{d}\mu \biggr)^2,\end{aligned}$$

where the supremum is taken with respect to all $g \in \mathrm{L}^2_\mu$, with support in F, such that $\|g\|_{2,\mu} = 1$. We can now complete the proof by applying the Cauchy–Schwarz inequality and (i) of the lemma but for $\mathcal{E}_\lambda^*$. □

5. Weighted Littlewood–Paley theory in the parabolic setting

We could develop a weighted parabolic Littlewood–Paley theory following the approach for singular integrals on spaces of homogeneous type [David et al. 1985]. However, since our weight w is time independent, we have decided to present a down-to-earth approach by combining weighted elliptic theory known in the field [Cruz-Uribe and Rios 2008; 2012; García-Cuerva and Rubio de Francia 1985] with Fourier analysis on the real line. Most of our estimates here are formulated using the square function norm

$$|||\cdot|||_{2,\mu} := \biggl(\int_0^\infty \iint_{\mathbb{R}^{n+1}} |\cdot|^2 \frac{\mathrm{d}\mu \, \mathrm{d}\lambda}{\lambda} \biggr)^{1/2}. \tag{5-1}$$

For the rest of the paper, $\mathcal{P} \in \mathrm{C}_0^\infty(\mathbb{R}^{n+1})$ is a fixed real-valued function in product form

$$\mathcal{P}(x,t) = \mathcal{P}^{(1)}(x)\mathcal{P}^{(2)}(t),$$

where $\mathcal{P}^{(1)}$ and $\mathcal{P}^{(2)}$ are both radial, nonnegative, and have integral 1. For all $x \in \mathbb{R}^n$, $t \in \mathbb{R}$, we set

$$\begin{aligned}\mathcal{P}_\lambda^{(1)}(x) &:= \lambda^{-n}\mathcal{P}^{(1)}(x/\lambda), \\ \mathcal{P}_\lambda^{(2)}(t) &:= \lambda^{-2}\mathcal{P}^{(2)}(t/\lambda^2), \\ \mathcal{P}_\lambda(x,t) &:= \mathcal{P}_\lambda^{(1)}(x)\mathcal{P}_\lambda^{(2)}(t) = \lambda^{-n-2}\mathcal{P}^{(1)}(x/\lambda)\mathcal{P}^{(2)}(t/\lambda^2)\end{aligned}$$

whenever $\lambda > 0$. With a slight abuse of notation, we let $\mathcal{P}_\lambda$ also denote the associated convolution operator

$$\mathcal{P}_\lambda f(x,t) = \mathcal{P}_\lambda * f(x,t) = \iint_{\mathbb{R}^{n+1}} \mathcal{P}_\lambda(x-y,t-s) f(y,s)\,\mathrm{d}y\,\mathrm{d}s,$$

and likewise for $\mathcal{P}_\lambda^{(1)}$ and $\mathcal{P}_\lambda^{(2)}$. We note that

$$\begin{aligned} |\mathcal{P}_\lambda^{(1)} f(x,t)| &\le \mathcal{M}^{(1)}(f(\cdot\,,t))(x), \\ |\mathcal{P}_\lambda^{(2)} f(x,t)| &\le \mathcal{M}^{(2)}(f(x,\cdot\,))(t), \\ |\mathcal{P}_\lambda f(x,t)| &\le \mathcal{M}^{(1)}(\mathcal{M}^{(2)} f(\cdot\,,t))(x) \end{aligned} \tag{5-2}$$

almost everywhere, for every $f \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}^{n+1})$; see [Stein 1993, Section II.2.1]. In particular, these pointwise bounds hold for $f \in \mathrm{L}^2_\mu$. The boundedness of the maximal operators in L^2_μ implies

$$\sup_{\lambda>0} \|\mathcal{P}_\lambda\|_{2\to 2,\mu} \lesssim 1;$$

see Section 2.2.

Lemma 5.1. *For all $f \in \mathrm{L}^2_\mu(\mathbb{R}^{n+1})$,*

$$|||\lambda \nabla_x \mathcal{P}_\lambda f|||_{2,\mu} + |||\lambda^2 \partial_t \mathcal{P}_\lambda f|||_{2,\mu} + |||\lambda D_t^{1/2} \mathcal{P}_\lambda f|||_{2,\mu} \lesssim \|f\|_{2,\mu}.$$

Proof. Here, we write out in detail how the product structure of $\mathcal{P}_\lambda$ can be used to prove parabolic estimates in $\mathbb{R}^{n+1}$ through weighted elliptic theory and classical Fourier analysis. This motif will appear in all proofs of this section. Let $\hat{g}$ denote the Fourier transform in time of a function g on $\mathbb{R}^{n+1}$.

By uniform boundedness of $\mathcal{P}_\lambda^{(1)}$ in L^2_μ and Plancherel's theorem, we have

$$\begin{aligned} |||\lambda D_t^{1/2} \mathcal{P}_\lambda f|||_{2,\mu}^2 &= \int_0^\infty \iint_{\mathbb{R}^{n+1}} |\mathcal{P}_\lambda^{(1)} \lambda D_t^{1/2} \mathcal{P}_\lambda^{(2)} f|^2 \,\frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda} \\ &\lesssim \int_{\mathbb{R}^n} \int_0^\infty \int_{\mathbb{R}} |\lambda D_t^{1/2} \mathcal{P}_\lambda^{(2)} f|^2 \,\frac{\mathrm{d}t\,\mathrm{d}\lambda}{\lambda}\,\mathrm{d}w \\ &= \int_{\mathbb{R}^n} \int_0^\infty \int_{\mathbb{R}} \big|\lambda |\tau|^{1/2} \widehat{\mathcal{P}^{(2)}}(\lambda^2 \tau) \hat{f}(x,\tau)\big|^2 \,\frac{\mathrm{d}\tau\,\mathrm{d}\lambda}{\lambda}\,\mathrm{d}w(x). \end{aligned}$$

The integral in λ is finite and independent of τ since $\widehat{\mathcal{P}^{(2)}}$ is a radial Schwartz function. Applying Plancherel's theorem backwards, we get the desired bound by $\|f\|_{2,\mu}^2$. The same argument yields the bound for $|||\lambda^2 \partial_t \mathcal{P}_\lambda f|||_{2,\mu}$.

Finally, in order to bound $\lambda \nabla_x \mathcal{P}_\lambda f$, we use uniform boundedness of $\mathcal{P}_\lambda^{(2)}$ to get

$$|||\lambda \nabla_x \mathcal{P}_\lambda f|||_{2,\mu}^2 = \int_0^\infty \iint_{\mathbb{R}^{n+1}} |\mathcal{P}_\lambda^{(2)} \lambda \nabla_x \mathcal{P}_\lambda^{(1)} f|^2 \,\frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda} \lesssim \int_{\mathbb{R}} \int_0^\infty \int_{\mathbb{R}^n} |\lambda \nabla_x \mathcal{P}_\lambda^{(1)} f|^2 \,\frac{\mathrm{d}w\,\mathrm{d}\lambda}{\lambda}\,\mathrm{d}t.$$

For fixed t, we now need weighted elliptic Littlewood–Paley theory. The operator $\lambda \nabla_x \mathcal{P}_\lambda^{(1)}$ acts by convolution with Ψ_λ, where $\Psi = \nabla_x \mathcal{P}^{(1)}$ has integral 0. Thus, we can use, e.g., [Cruz-Uribe and Rios 2012, Lemma 4.6] to control the integral in $\mathrm{d}w\,\mathrm{d}\lambda$ by $\|f(\cdot\,,t)\|_{2,w}^2$, and the proof is complete. ☐

Lemma 5.2. *For all* $f \in \mathrm{E}_\mu$,

$$|||\lambda^{-1}(I - \mathcal{P}_\lambda) f|||_{2,\mu} \lesssim \|\mathbb{D} f\|_{2,\mu}.$$

Proof. We first claim

$$|||\lambda^{-1}(I - \mathcal{P}_\lambda^{(1)}) f|||_{2,\mu} \lesssim \|\nabla_x f\|_{2,\mu}, \quad |||\lambda^{-1}(I - \mathcal{P}_\lambda^{(2)}) f|||_{2,\mu} \lesssim \|D_t^{1/2} f\|_{2,\mu}. \tag{5-3}$$

As in the proof of Lemma 5.1, this can be proved using Plancherel's theorem in t for the second term and weighted Littlewood–Paley theory with t fixed for the first term. The required weighted result is [Cruz-Uribe and Rios 2015, Proposition 2.3] (originally [Cruz-Uribe and Rios 2012, Proposition 4.7]) and the application to the concrete operator considered here is detailed in the lines following equation (4.3) in the same paper.

In order to complete the proof of the lemma, we simply write

$$(I - \mathcal{P}_\lambda) = \mathcal{P}_\lambda^{(2)}(1 - \mathcal{P}_\lambda^{(1)}) + (1 - \mathcal{P}_\lambda^{(2)}).$$

The result follows from (5-3) and the uniform boundedness of $\mathcal{P}_\lambda^{(2)}$ in L^2_μ. □

In the following we write $\Delta = Q \times I$ for parabolic cubes in $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$.

Definition 5.3. We define $\mathcal{A}_\lambda^{(1)}$, $\mathcal{A}_\lambda^{(2)}$ and $\mathcal{A}_\lambda$ to be the dyadic averaging operators in x, t and (x,t) with respect to parabolic scaling, that is, if $\Delta = Q \times I$ is the dyadic parabolic cube with $\frac{1}{2}\ell(\Delta) < \lambda \le \ell(\Delta)$ containing (x,t), then

$$\mathcal{A}_\lambda^{(1)} f(x,t) := \fint_Q f \,\mathrm{d}y,$$
$$\mathcal{A}_\lambda^{(2)} f(x,t) := \fint_I f \,\mathrm{d}s,$$
$$\mathcal{A}_\lambda f(x,t) := \fint\!\!\fint_\Delta f \,\mathrm{d}y\,\mathrm{d}s = \mathcal{A}_\lambda^{(1)} \mathcal{A}_\lambda^{(2)} f(x,t).$$

It follows from the bounds for the maximal operators in Section 2.2 and doubling that the dyadic averaging operators are bounded on L^2_μ, uniformly in λ.

Lemma 5.4. *Let* $\mathcal{P}_\lambda$ *and* $\mathcal{A}_\lambda$ *be as above. Then, for all* $f \in \mathrm{L}^2_\mu(\mathbb{R}^{n+1})$,

$$|||(\mathcal{A}_\lambda - \mathcal{P}_\lambda) f|||_{2,\mu} \lesssim \|f\|_{2,\mu}.$$

Proof. We follow our (general) strategy and write

$$\mathcal{A}_\lambda - \mathcal{P}_\lambda = \mathcal{A}_\lambda^{(2)}(\mathcal{A}_\lambda^{(1)} - \mathcal{P}_\lambda^{(1)}) + \mathcal{P}_\lambda^{(1)}(\mathcal{A}_\lambda^{(2)} - \mathcal{P}_\lambda^{(2)}),$$

where we have also used that $\mathcal{A}_\lambda^{(2)}$ and $\mathcal{P}_\lambda^{(1)}$ commute since they act in different variables. Since these operators are uniformly bounded on L^2_μ with respect to λ, we get

$$|||(\mathcal{A}_\lambda - \mathcal{P}_\lambda) f|||_{2,\mu} \lesssim \int_{\mathbb{R}} \int_0^\infty \|(\mathcal{A}_\lambda^{(1)} - \mathcal{P}_\lambda^{(1)}) f(\cdot,t)\|^2_{2,w} \frac{\mathrm{d}\lambda}{\lambda}\,\mathrm{d}t$$
$$+ \int_{\mathbb{R}^n} \int_0^\infty \|(\mathcal{A}_\lambda^{(2)} - \mathcal{P}_\lambda^{(2)}) f(x,\cdot)\|^2_{2,\mathrm{d}t} \frac{\mathrm{d}\lambda}{\lambda}\,\mathrm{d}w(x).$$

For the first term on the right, we can rely on the weighted elliptic version of the lemma [Cruz-Uribe and Rios 2012, Lemma 5.2]. For the second term on the right, we can make a change of variables $\lambda' = \lambda^2$ and use the unweighted one-dimensional version of the lemma, which of course follows from the same reference or the classical proof in [Auscher and Tchamitchian 1998, Appendix C, (4)]. □

6. Reduction to a quadratic estimate

The purpose of this short section is to reduce our main result, Theorem 1.1, to the quadratic estimate

$$|||\lambda \mathcal{H} \mathcal{E}_\lambda f|||_{2,\mu} \lesssim \|\nabla_x f\|_{2,\mu} + \|H_t D_t^{1/2} f\|_{2,\mu}, \qquad f \in \mathrm{E}_\mu. \tag{6-1}$$

Recall that $\mathcal{E}_\lambda = (1+\lambda^2 \mathcal{H})^{-1}$. Since the sesquilinear form associated with $\mathcal{H}$ is not closed, see Section 4, classical results à la Lions [1962] as in the elliptic case do not apply, and here we give full details of this reduction.

At this point, we require some essentials from functional calculus. We give references along the way, and we refer the reader to [Haase 2006; McIntosh 1986] for further background. Since $\mathcal{H}$ is m-accretive (Proposition 4.2), it has a unique m-accretive square root $\sqrt{\mathcal{H}}$ defined by the functional calculus for sectorial operators, and the same is true for the adjoint $\mathcal{H}^*$ with $\sqrt{\mathcal{H}^*} = (\sqrt{\mathcal{H}})^*$.

In order to see the reduction alluded to above, we start out with the Calderón reproducing formula in [Haase 2006, Theorem 5.2.6], and we write

$$\sqrt{\mathcal{H}} f = \frac{16}{\pi} \int_0^\infty \lambda^3 \mathcal{H}^2 (1+\lambda^2 \mathcal{H})^{-3} f \frac{\mathrm{d}\lambda}{\lambda}, \tag{6-2}$$

where $f \in \mathsf{D}(\sqrt{\mathcal{H}})$ and the integral is understood as an improper Riemann integral in L^2_μ. Testing this identity against $g \in \mathrm{L}^2_\mu$ and applying Cauchy–Schwarz, we obtain

$$|\langle \sqrt{\mathcal{H}} f, g\rangle_{2,\mu}| \le \frac{16}{\pi} |||\lambda \mathcal{H}(1+\lambda^2\mathcal{H})^{-1} f|||_{2,\mu} \times |||\lambda^2 \mathcal{H}^*(1+\lambda^2 \mathcal{H}^*)^{-2} g|||_{2,\mu}.$$

The second term is controlled by a structural constant times $\|g\|_{2,\mu}$ since $\mathcal{H}^*$ is m-accretive in L^2_μ — more precisely, this follows from von Neumann's inequality [Haase 2006, Theorem 7.1.7] and the characterization of the emerging functional calculus due to McIntosh [Haase 2006, Theorem 7.3.1]. Taking the supremum over all g yields

$$\|\sqrt{\mathcal{H}} f\|_{2,\mu} \lesssim |||\lambda \mathcal{H}(I+\lambda^2 \mathcal{H})^{-1} f|||_{2,\mu}.$$

Let us now suppose that (6-1) holds. Then, we obtain

$$\|\sqrt{\mathcal{H}} f\|_{2,\mu} \lesssim \|\nabla_x f\|_{2,\mu} + \|H_t D_t^{1/2} f\|_{2,\mu}, \tag{6-3}$$

when f is in $\mathrm{E}_\mu \cap \mathsf{D}(\sqrt{\mathcal{H}}) \supset \mathsf{D}(\mathcal{H})$. However, since this space is dense in E_μ, by Proposition 4.2, and as $\sqrt{\mathcal{H}}$ is closed, the estimate extends to all $f \in \mathrm{E}_\mu$. Next, we note that $\mathcal{H}^*$ is similar to an operator in the same class as $\mathcal{H}$ under conjugation with the "time reversal" $f(t,x) \mapsto f(-t,x)$ and conjugation of A. Hence, we also have

$$\|\sqrt{\mathcal{H}^*} g\|_{2,\mu} \lesssim \|\nabla_x g\|_{2,\mu} + \|H_t D_t^{1/2} g\|_{2,\mu} \tag{6-4}$$

whenever $g \in \mathrm{E}_\mu$. Using (4-6) with $\sigma = 0$ and δ small enough depending on the structural constants, we obtain, for all $f \in \mathsf{D}(\mathcal{H})$, that

$$\delta \|\nabla_x f\|_{2,\mu}^2 + \delta \|D_t^{1/2} f\|_{2,\mu}^2 \le |\langle \mathcal{H} f, (1+\delta H_t) f\rangle_{2,\mu}| \\ \le \|\sqrt{\mathcal{H}} f\|_{2,\mu} \|\sqrt{\mathcal{H}^*}(1+\delta H_t) f\|.$$

Now, (6-4) with $g := (1+\delta H_t) f \in \mathrm{E}_\mu$ implies

$$\|\nabla_x f\|_{2,\mu} + \|D_t^{1/2} f\|_{2,\mu} \lesssim \|\sqrt{\mathcal{H}} f\|_{2,\mu}. \tag{6-5}$$

Since $\mathsf{D}(\mathcal{H})$ is dense in $\mathsf{D}(\sqrt{\mathcal{H}})$ for the graph norm [Haase 2006, Proposition 3.1.1(h)], the estimate extends to all $f \in \mathsf{D}(\sqrt{\mathcal{H}})$.

In conclusion, we have seen that (6-1) implies the statement of Theorem 1.1 through the estimates (6-3) and (6-5). Therefore, the rest of the paper is devoted to the task of proving (6-1).

7. Principal part approximation

In order to prove the square function estimate (6-1), we will eventually split $\mathcal{H}$ into its elliptic and parabolic parts and perform the "hard" analysis only on the elliptic part. This will lead us to the operators

$$\mathcal{U}_\lambda := \lambda \mathcal{E}_\lambda w^{-1} \operatorname{div}_x w, \quad \lambda > 0. \tag{7-1}$$

These operators appeared in Lemma 4.4 on off-diagonal estimates and in particular they are uniformly bounded on $(\mathrm{L}^2_\mu)^n$. Here, we continue their analysis.

Given a cube $Q = Q_r(x) \subset \mathbb{R}^n$ and an interval $I = I_r(t)$, we let $\Delta := Q \times I$ and set

$$C_k(\Delta) = C_k(Q \times I) := 2^{k+1}\Delta \setminus 2^k \Delta, \quad k = 1, 2, \dots, \\ C_0(\Delta) := 2\Delta.$$

In the following, we denote the characteristic function of a set E by 1_E. We use off-diagonal estimates to define $\mathcal{U}_\lambda$ on $(\mathrm{L}^\infty)^n$.

Definition 7.1. For $\boldsymbol{b} \in (\mathrm{L}^\infty)^n$, we define

$$\mathcal{U}_\lambda \boldsymbol{b} =: \lim_{k\to\infty} \mathcal{U}_\lambda(\boldsymbol{b} 1_{2^k\Delta}), \tag{7-2}$$

with convergence locally in $(\mathrm{L}^2_\mu)^n$, where on the right Δ is any parabolic cube.

Definition 7.1 is meaningful and independent of the choice of Δ as we shall see next. To start, if Δ' is any parabolic cube, then for $m > l$ large enough to guarantee that $\Delta' \subset 2^{l-1}\Delta$, applying Lemma 4.4 with $E = C_j(\Delta)$ and $F = \Delta'$ for $j = l, \dots, m-1$ yields

$$\|\mathcal{U}_\lambda(\boldsymbol{b} 1_{2^m\Delta \setminus 2^l \Delta})\|_{\mathrm{L}^2_\mu(\Delta')} \le \sum_{j=l}^{m-1} \|\mathcal{U}_\lambda(\boldsymbol{b} 1_{C_j(\Delta)})\|_{\mathrm{L}^2_\mu(\Delta')} \\ \lesssim \mu(\Delta)^{1/2} \|\boldsymbol{b}\|_\infty \sum_{j=l}^{m-1} e^{-\ell(\Delta) 2^{j-1}/c\lambda} (4D)^{j+1}.$$

Recall that $4D$ is the doubling constant for μ; see (2-3). The right-hand side converges to 0 as $m, l \to \infty$. In conclusion, $\{\mathcal{U}_\lambda(\boldsymbol{b}1_{2^l\Delta})\}_l$ is a Cauchy sequence locally in $(\mathrm{L}^2_\mu)^n$. By the same argument, Definition 7.1 is independent of the particular choice Δ. Taking $\Delta' = \Delta$ and $l = 1$, we get

$$\|\mathcal{U}_\lambda \boldsymbol{b}\|_{\mathrm{L}^2_\mu(\Delta)} \le \|\mathcal{U}_\lambda(\boldsymbol{b}1_{2\Delta})\|_{\mathrm{L}^2_\mu(\Delta)} + \Big\|\lim_{m\to\infty} \mathcal{U}_\lambda(\boldsymbol{b}1_{2^m\Delta\setminus 2\Delta})\Big\|_{\mathrm{L}^2_\mu(\Delta)}$$
$$\lesssim \mu(\Delta)^{1/2}\|\boldsymbol{b}\|_\infty\bigg(1 + \sum_{j=1}^{\infty} e^{-\ell(\Delta)2^{j-1}/c\lambda}(4D)^{j+1}\bigg). \tag{7-3}$$

Lemma 7.2. *Let* $\boldsymbol{b} \in (\mathrm{L}^\infty)^n$ *and* $f \in \mathrm{L}^2_\mu$. *Then,*

$$\|(\mathcal{U}_\lambda \boldsymbol{b}) \cdot \mathcal{A}_\lambda f\|_{2,\mu} \lesssim \|\boldsymbol{b}\|_\infty \|f\|_{2,\mu}.$$

Proof. If $\Delta \subset \mathbb{R}^{n+1}$ is a parabolic cube such that $\frac{1}{2}\ell(\Delta) < \lambda \le \ell(\Delta)$, then by (7-3) we have

$$\iint_\Delta |\mathcal{U}_\lambda \boldsymbol{b}|^2 \,\mathrm{d}\mu \lesssim \mu(\Delta)\|\boldsymbol{b}\|_\infty^2.$$

Since $\mathcal{A}_\lambda f$ is constant on each such Δ, we obtain

$$\iint_\Delta |(\mathcal{U}_\lambda \boldsymbol{b}) \cdot \mathcal{A}_\lambda f|^2 \,\mathrm{d}\mu \le \iint_\Delta |\mathcal{U}_\lambda \boldsymbol{b}|^2 \,\mathrm{d}\mu \cdot \fint\!\!\fint_\Delta |\mathcal{A}_\lambda f|^2 \,\mathrm{d}\mu \lesssim \|\boldsymbol{b}\|_\infty^2 \iint_\Delta |\mathcal{A}_\lambda f|^2 \,\mathrm{d}\mu.$$

The claim follows by summing in Δ and using that $\mathcal{A}_\lambda$ is uniformly bounded on L^2_μ with respect to λ; see Section 5. □

Writing $A = (A_1, \dots, A_n)$ with A_j the j-th column of A, we can use Definition 7.1 to define the action of $\mathcal{U}_\lambda$ on the bounded matrix-valued function $w^{-1}A$ by

$$(\mathcal{U}_\lambda w^{-1}A) := \mathcal{U}_\lambda(w^{-1}A) := (\mathcal{U}_\lambda(w^{-1}A_1), \dots, \mathcal{U}_\lambda(w^{-1}A_n)).$$

We will approximate $\mathcal{U}_\lambda w^{-1}A$ by operators that act via multiplication on the maximal dyadic cubes of size at most λ. To be precise, we will consider

$$\mathcal{R}_\lambda f := \mathcal{U}_\lambda(w^{-1}Af) - (\mathcal{U}_\lambda w^{-1}A) \cdot \mathcal{A}_\lambda f. \tag{7-4}$$

This is nowadays called the "principal part approximation". Using Lemmas 4.3 and 7.2, we see that the $\mathcal{R}_\lambda$ are uniformly bounded on L^2_μ for $\lambda > 0$. Moreover, we prove the following bound.

Proposition 7.3. *Let* $f \in \mathrm{L}^2_\mu \cap \mathrm{C}^\infty$. *Then,*

$$\|\mathcal{R}_\lambda f\|_{2,\mu} \lesssim \|\lambda \nabla_x f\|_{2,\mu} + \|\lambda^2 \partial_t f\|_{2,\mu}.$$

For the proof, we need the following weighted Poincaré-type inequality. In the following we abbreviate $(f)_\Delta = (f)_{\Delta,\mathrm{d}x\,\mathrm{d}t}$.

Lemma 7.4. *Let* $f \in \mathrm{C}^\infty$. *Then, for all parabolic cubes* Δ *and all nonnegative integers* k,

$$\iint_{C_k(\Delta)} |(f - (f)_\Delta)|^2 \,\mathrm{d}\mu \le c2^{k(n+2)} \iint_{2^{k+1}\Delta} \ell(\Delta)^2|\nabla_x f|^2 + \ell(\Delta)^4|\partial_t f|^2 \,\mathrm{d}\mu,$$

where c *depends only on* n *and* $[w]_{A_2}$.

Proof. Let $\Delta = Q \times I$ be a parabolic cube. We set $g := (f)_{Q,\mathrm{d}x}$, which is a function of t, and we split

$$f - (f)_\Delta = (f - (f)_{Q,\mathrm{d}x}) + (g - (g)_{I,\mathrm{d}t}).$$

To the first term we can apply the weighted Poincaré inequality in the x-variable from (the proof of) [Heinonen et al. 1993, Theorem 15.26] and to the second term the standard Poincaré inequality in the t-variable. The result is

$$\left(\iint_\Delta |(f - (f)_\Delta)|^2 \,\mathrm{d}\mu\right)^{1/2} \leq c\left(\iint_\Delta \ell(\Delta)^2 |\nabla_x f|^2 + \ell(\Delta)^4 |\partial_t f|^2 \,\mathrm{d}\mu\right)^{1/2}.$$

Note that in [Heinonen et al. 1993] balls are used instead of cubes, but doubling allows us to switch between one and the other. For the general result it suffices to write

$$f - (f)_\Delta = (f - (f)_{2^{k+1}\Delta}) + ((f)_{2^{k+1}\Delta} - (f)_{2^k\Delta}) + \cdots + ((f)_{2\Delta} - (f)_\Delta)$$

and to use the estimate above on the cubes $2^{k+1}\Delta$ and then on $2^{k+1}\Delta, \ldots, 2\Delta$. □

Proof of Proposition 7.3. We note that if $(x, t) \in \mathbb{R}^{n+1}$ and $\lambda > 0$, then

$$\mathcal{R}_\lambda f(x,t) = \mathcal{U}_\lambda(w^{-1} A(f - (f)_\Delta))(x,t),$$

where Δ is the unique dyadic parabolic cube with $\frac{1}{2}\ell(\Delta) < \lambda \leq \ell(\Delta)$ that contains (x, t). Thus,

$$\begin{aligned}\|\mathcal{R}_\lambda f\|_{2,\mu}^2 &= \sum_\Delta \iint_\Delta |\mathcal{U}_\lambda(w^{-1} A(f - (f)_\Delta))|^2 \,\mathrm{d}\mu \\ &\leq \sum_\Delta \left(\sum_{k=0}^\infty \left(\iint_\Delta |\mathcal{U}_\lambda(w^{-1} A \cdot 1_{C_k(\Delta)}(f - (f)_\Delta))|^2 \,\mathrm{d}\mu\right)^{1/2}\right)^2,\end{aligned}$$

and therefore

$$\begin{aligned}\|\mathcal{R}_\lambda f\|_{2,\mu}^2 &\lesssim \sum_\Delta \left(\sum_{k=0}^\infty e^{-2^k/c} \left(\iint_{C_k(\Delta)} |(f - (f)_\Delta)|^2 \,\mathrm{d}\mu\right)^{1/2}\right)^2 \\ &\lesssim \sum_\Delta \sum_{k=0}^\infty e^{-2^k/c} \iint_{C_k(\Delta)} |(f - (f)_\Delta)|^2 \,\mathrm{d}\mu \\ &\lesssim \sum_\Delta \sum_{k=0}^\infty e^{-2^k/c} 2^{k(n+2)} \iint_{2^{k+1}\Delta} \lambda^2 |\nabla_x f|^2 + \lambda^4 |\partial_t f|^2 \,\mathrm{d}\mu \\ &\leq \left(\sum_{k=0}^\infty e^{-2^k/c} 2^{(2k+1)(n+2)}\right) \iint_{\mathbb{R}^{n+1}} \lambda^2 |\nabla_x f|^2 + \lambda^4 |\partial_t f|^2 \,\mathrm{d}\mu,\end{aligned}$$

where we used, in succession, the off-diagonal estimates, Cauchy–Schwarz inequality, Lemma 7.4, and the fact that each point in $\mathbb{R}^{n+1}$ is contained in exactly $2^{(k+1)(n+2)}$ of the cubes $2^{k+1}\Delta$. The sum in k is still finite, and the proof is complete. □

8. Proof of Theorem 1.1

After the reduction in Section 6, it remains to prove the quadratic estimate (6-1) that we now write in the form

$$|||\lambda\mathcal{E}_\lambda\mathcal{H}f|||_{2,\mu} \lesssim \|\mathbb{D}f\|_{2,\mu}, \quad f\in \mathrm{E}_\mu. \tag{8-1}$$

In the following we will use the operators $\mathcal{P}_\lambda$, $\mathcal{A}_\lambda$, $\mathcal{U}_\lambda$, $\mathcal{R}_\lambda$ that have been introduced in Sections 4, 5 and 7. Collecting the estimates from these sections, we can at this stage prove the following.

Proposition 8.1. *Let $f\in\mathrm{E}_\mu$. Then,*

$$|||(\lambda\mathcal{E}_\lambda\mathcal{H}+(\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda\nabla_x)f|||_{2,\mu}\lesssim\|\mathbb{D}f\|_{2,\mu}.$$

Proof. We begin by writing

$$\lambda\mathcal{E}_\lambda\mathcal{H}f=\lambda\mathcal{E}_\lambda\mathcal{H}\mathcal{P}_\lambda f+\lambda\mathcal{H}\mathcal{E}_\lambda(I-\mathcal{P}_\lambda)f. \tag{8-2}$$

Using the identity

$$\lambda\mathcal{H}\mathcal{E}_\lambda=\lambda^{-1}(I-\mathcal{E}_\lambda),$$

the uniform L^2_μ-boundedness of $\mathcal{E}_\lambda$, and Lemma 5.2, we see that

$$|||\lambda\mathcal{H}\mathcal{E}_\lambda(I-\mathcal{P}_\lambda)f|||_{2,\mu}\lesssim|||\lambda^{-1}(I-\mathcal{P}_\lambda)f|||_{2,\mu}\lesssim\|\mathbb{D}f\|_{2,\mu}.$$

Next, we use (4-3) to write

$$\lambda\mathcal{E}_\lambda\mathcal{H}\mathcal{P}_\lambda f=-\mathcal{U}_\lambda w^{-1}A\nabla_x\mathcal{P}_\lambda f+\lambda\mathcal{E}_\lambda D_t^{1/2}H_tD_t^{1/2}\mathcal{P}_\lambda f. \tag{8-3}$$

Using Lemma 4.3 (i) and then Lemma 5.1, we see that

$$\begin{aligned}|||\lambda\mathcal{E}_\lambda D_t^{1/2}H_tD_t^{1/2}\mathcal{P}_\lambda f|||_{2,\mu}&=|||\lambda\mathcal{E}_\lambda D_t^{1/2}\mathcal{P}_\lambda D_t^{1/2}H_tf|||_{2,\mu}\\&\lesssim|||\lambda D_t^{1/2}\mathcal{P}_\lambda D_t^{1/2}H_tf|||_{2,\mu}\\&\lesssim\|D_t^{1/2}f\|_{2,\mu}.\end{aligned} \tag{8-4}$$

Finally, we bring the principal part approximation into play. We use $\mathcal{U}_\lambda$ and $\mathcal{R}_\lambda$ to write

$$\begin{aligned}\mathcal{U}_\lambda w^{-1}A\nabla_x\mathcal{P}_\lambda f&=\mathcal{U}_\lambda w^{-1}A\mathcal{P}_\lambda\nabla_xf\\&=\mathcal{R}_\lambda\mathcal{P}_\lambda\nabla_xf+(\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda(\mathcal{P}_\lambda-\mathcal{A}_\lambda)\nabla_xf+(\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda\nabla_xf,\end{aligned} \tag{8-5}$$

where we have also used that $(\mathcal{A}_\lambda)^2=\mathcal{A}_\lambda$ for the last term. Applying Proposition 7.3 and Lemma 5.1, we have

$$|||\mathcal{R}_\lambda\mathcal{P}_\lambda\nabla_xf|||_{2,\mu}\lesssim|||\lambda\nabla_x\mathcal{P}_\lambda\nabla_xf|||_{2,\mu}+|||\lambda^2\partial_t\mathcal{P}_\lambda\nabla_xf|||_{2,\mu}\lesssim\|\mathbb{D}f\|_{2,\mu}.$$

Also, by Lemmas 5.4 and 7.2, we have

$$|||(\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda(\mathcal{P}_\lambda-\mathcal{A}_\lambda)\nabla_xf|||_{2,\mu}\lesssim|||(\mathcal{A}_\lambda-\mathcal{P}_\lambda)\nabla_xf|||_{2,\mu}\lesssim\|\nabla_xf\|_{2,\mu}.$$

Looking back at the successive splittings in (8-2), (8-3) and (8-5), we see that the only term that has not been treated in the square function norm is $(\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda\nabla_xf$. This proves the claim. □

To conclude the square function estimate for the final term $(\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda\nabla_x f$, we establish Lemma 8.3 below. The lemma states that

$$|\mathcal{U}_\lambda w^{-1}A|^2\,\frac{\mathrm{d}\mu\,\mathrm{d}\lambda}{\lambda}$$

is a Carleson measure and that we have good control of the constants. Hence,

$$|||(\mathcal{U}_\lambda w^{-1}A)\cdot\mathcal{A}_\lambda\nabla_x f|||_{2,\mu}\lesssim\|\nabla_x f\|_{2,\mu}$$

follows by Carleson's inequality for parabolic cubes; see Lemma 8.2. This completes the proof of the estimate in (8-1), and hence the proof of Theorem 1.1 modulo Lemma 8.3. The reader should observe that, in our proof of (8-1), we have split off the time derivative ∂_t from $\mathcal{H}$ and we have controlled the part coming from ∂_t by a standard Littlewood–Paley estimate in (8-4).

For convenience, we include a proof of the version of Carleson's inequality that is used above. We adapt the elegant dyadic argument found in [Morris 2012, Theorem 4.3].

Lemma 8.2. *Let ν be a Borel measure on $\mathbb{R}^{n+1}\times\mathbb{R}^+$ that satisfies*

$$\|\nu\|_{\mathcal{C}}:=\sup_{\Delta}\frac{\nu(\Delta\times(0,\ell(\Delta)])}{\mu(\Delta)}<\infty,$$

where the supremum is taken over all dyadic parabolic cubes $\Delta\subset\mathbb{R}^{n+1}$. Then there is a constant c that only depends on n and $[w]_{A_2}$ such that, for every $f\in\mathrm{L}^2_\mu$,

$$\int_0^\infty\iint_{\mathbb{R}^{n+1}}|\mathcal{A}_\lambda f(x,t)|^2\,\mathrm{d}\nu(x,t,\lambda)\le c\|\nu\|_{\mathcal{C}}\iint_{\mathbb{R}^{n+1}}|f|^2\,\mathrm{d}\mu.$$

Proof. For $i\in\mathbb{Z}$, let $\{\Delta_i^j\}_j$ be the partition of $\mathbb{R}^{n+1}$ into dyadic parabolic cubes such that $\ell(\Delta_i^j)=2^i$. We have

$$\int_0^\infty\iint_{\mathbb{R}^{n+1}}|\mathcal{A}_\lambda f(x,t)|^2\,\mathrm{d}\nu(x,t,\lambda)=\sum_{i=-\infty}^{\infty}\sum_j\left|\fint\!\!\fint_{\Delta_i^j}f\,\mathrm{d}y\,\mathrm{d}s\right|^2\nu(\Delta_i^j\times(2^{i-1},2^i])=\sum_{i=-\infty}^{\infty}\sum_j|f_i^j|^2\nu_i^j,$$

where we have introduced $\nu_i^j:=\nu(\Delta_i^j\times(2^{i-1},2^i])$ and $f_i^j:=\fint\!\!\fint_{\Delta_i^j}f\,\mathrm{d}y\,\mathrm{d}s$. For $r>0$, let $\{\Delta_k(r)\}_k$ be the collection of maximal dyadic parabolic cubes Δ_i^j such that $|f_i^j|>r$. Note that these cubes are pairwise disjoint and contained in $\{\mathcal{M}^{(1)}\mathcal{M}^{(2)}f>r\}$. Hence,

$$\begin{aligned}\sum_{i=-\infty}^{\infty}\sum_j|f_i^j|^2\nu_i^j&=\int_0^\infty 2r\sum_{i=-\infty}^{\infty}\sum_j 1_{\{|f_i^j|>r\}}\nu_i^j\,\mathrm{d}r\le\int_0^\infty 2r\sum_k\sum_{\Delta\subset\Delta_k(r)}\nu\big(\Delta\times\big(\tfrac12\ell(\Delta),\ell(\Delta)\big]\big)\,\mathrm{d}r\\&=\int_0^\infty 2r\sum_k\nu(\Delta_k(r)\times(0,\ell(\Delta_k(r))])\,\mathrm{d}r\\&\le\|\nu\|_{\mathcal{C}}\int_0^\infty 2r\sum_k\mu(\Delta_k(r))\,\mathrm{d}r\le\|\nu\|_{\mathcal{C}}\int_0^\infty 2r\mu(\{\mathcal{M}^{(1)}\mathcal{M}^{(2)}f>r\})\,\mathrm{d}r\\&=\|\nu\|_{\mathcal{C}}\|\mathcal{M}^{(1)}\mathcal{M}^{(2)}f\|_{2,\mu}^2.\end{aligned}$$

Now, the claim follows from the Hardy–Littlewood–Muckenhoupt inequality. □

The rest of the section is devoted to the proof of the following lemma.

Lemma 8.3. *For all dyadic parabolic cubes $\Delta = Q \times I \subset \mathbb{R}^{n+1}$,*

$$\int_0^{\ell(\Delta)} \iint_\Delta |\mathcal{U}_\lambda w^{-1} A|^2 \frac{\mathrm{d}\mu\, \mathrm{d}\lambda}{\lambda} \lesssim \mu(\Delta).$$

The proof of Lemma 8.3 is based on the use of appropriate local Tb-type test functions.

8.1. ***Construction of appropriate local Tb-type test functions.*** Let $\zeta \in \mathbb{C}^n$ with $|\zeta| = 1$, and let ζ_i denote the i-th component of ζ for $1 \le i \le n$. We let χ and η be smooth functions on $\mathbb{R}^n$ and $\mathbb{R}$, respectively, whose values are in $[0, 1]$. The function χ is equal to 1 on $\left[-\frac{1}{2}, \frac{1}{2}\right]^n$ and has support in $(-1, 1)^n$, and η is equal to 1 on $\left[-\frac{1}{2}, \frac{1}{2}\right]$ with support in $(-1, 1)$. We fix a parabolic dyadic cube Δ and denote its center by (x_Δ, t_Δ). We first introduce

$$\chi_\Delta(x, t) := \chi\left(\frac{x - x_\Delta}{\ell(\Delta)}\right)\eta\left(\frac{t - t_\Delta}{\ell(\Delta)^2}\right).$$

Based on ζ and χ_Δ, we introduce

$$L_\Delta^\zeta(x, t) := \chi_\Delta(x, t)(\Phi_\Delta(x) \cdot \bar{\zeta}), \qquad \Phi_\Delta(x) := (x - x_\Delta).$$

Clearly, $L_\Delta^\zeta \in \mathrm{E}_\mu$. Using the function L_Δ^ζ and $0 < \epsilon \ll 1$, we define the test function

$$f_{\Delta,\epsilon}^\zeta := \mathcal{E}_{\epsilon\ell(\Delta)} L_\Delta^\zeta = (I + (\epsilon\ell(\Delta))^2 \mathcal{H})^{-1} L_\Delta^\zeta. \tag{8-6}$$

Lemma 8.4. *Let $\zeta \in \mathbb{C}^n$ with $|\zeta| = 1$, and let $0 < \epsilon \ll 1$ be a degree of freedom. Given a parabolic dyadic cube Δ, define $f_{\Delta,\epsilon}^\zeta$ as in (8-6). Then,*

(i) $\|f_{\Delta,\epsilon}^\zeta - L_\Delta^\zeta\|_{2,\mu}^2 \lesssim (\epsilon\ell(\Delta))^2 \mu(\Delta)$,

(ii) $\|\mathbb{D}(f_{\Delta,\epsilon}^\zeta - L_\Delta^\zeta)\|_{2,\mu}^2 \lesssim \mu(\Delta)$,

(iii) $\|\mathbb{D} f_{\Delta,\epsilon}^\zeta\|_{2,\mu}^2 \lesssim \mu(\Delta)$.

Proof. Note that

$$\begin{aligned} f_{\Delta,\epsilon}^\zeta - L_\Delta^\zeta &= -(\epsilon\ell(\Delta))^2 \mathcal{E}_{\epsilon\ell(\Delta)} \mathcal{H} L_\Delta^\zeta \\ &= -(\epsilon\ell(\Delta))^2 \mathcal{E}_{\epsilon\ell(\Delta)} D_t^{1/2} H_t D_t^{1/2} L_\Delta^\zeta + (\epsilon\ell(\Delta))^2 \mathcal{E}_{\epsilon\ell(\Delta)} w^{-1} \operatorname{div}_x w(w^{-1} A \nabla_x L_\Delta^\zeta). \end{aligned}$$

Hence, using the uniform L^2_μ-boundedness of $(\epsilon\ell(\Delta))\mathcal{E}_{\epsilon\ell(\Delta)} D_t^{1/2}$ and $(\epsilon\ell(\Delta))\mathcal{E}_{\epsilon\ell(\Delta)} w^{-1} \operatorname{div}_x w$, see Lemma 4.3, we get

$$\iint_{\mathbb{R}^{n+1}} |f_{\Delta,\epsilon}^\zeta - L_\Delta^\zeta|^2 \,\mathrm{d}\mu \lesssim \iint_{\mathbb{R}^{n+1}} |(\epsilon\ell(\Delta))\mathbb{D} L_\Delta^\zeta|^2 \,\mathrm{d}\mu.$$

Furthermore,

$$\iint_{\mathbb{R}^{n+1}} |\mathbb{D} L_\Delta^\zeta|^2 \,\mathrm{d}\mu = \iint_{\mathbb{R}^{n+1}} |\nabla_x L_\Delta^\zeta|^2 \,\mathrm{d}\mu + \iint_{\mathbb{R}^{n+1}} |D_t^{1/2} L_\Delta^\zeta|^2 \,\mathrm{d}\mu \lesssim \mu(\Delta) \tag{8-7}$$

by the construction of L_Δ^ζ (to estimate $D_t^{1/2} L_\Delta^\zeta$ we use the homogeneity of the Fourier symbol). Similarly, we deduce that

$$\iint_{\mathbb{R}^{n+1}} |\mathbb{D}(f_{\Delta,\epsilon}^\zeta - L_\Delta^\zeta)|^2 \,\mathrm{d}\mu \lesssim \mu(\Delta).$$

This proves (i) and (ii). To prove (iii), we simply use (ii) and (8-7). ∎

Lemma 8.5. *Given a parabolic dyadic cube $\Delta = Q \times I$, let $f^{\zeta}_{\Delta,\epsilon}$ be defined as in (8-6). There exist $\epsilon \in (0,1)$, depending only on the structural constants, and a finite set W of unit vectors in $\mathbb{C}^n$, whose cardinality depends on ϵ and n, such that*

$$\sup_{\Delta} \frac{1}{|\Delta|} \int_0^{\ell(\Delta)} \iint_{\Delta} |\mathcal{U}_\lambda w^{-1} A|^2 \, \frac{\mathrm{d}\mu \, \mathrm{d}\lambda}{\lambda} \lesssim \sum_{\zeta \in W} \sup_{\Delta} \frac{1}{|\Delta|} \int_0^{\ell(\Delta)} \iint_{\Delta} |(\mathcal{U}_\lambda w^{-1} A) \cdot \mathcal{A}_\lambda \nabla_x f^{\zeta}_{\Delta,\epsilon}|^2 \, \frac{\mathrm{d}\mu \, \mathrm{d}\lambda}{\lambda},$$

where the supremum is taken over all dyadic parabolic cubes $\Delta \subset \mathbb{R}^{n+1}$.

Proof. Consider a degree of freedom $\epsilon > 0$. Given a unit vector ζ in $\mathbb{C}^n$, we introduce the cone

$$C^{\epsilon}_{\zeta} := \{u \in \mathbb{C}^n : |u - (u \cdot \bar{\zeta})\zeta| \le \epsilon |u \cdot \bar{\zeta}|\}.$$

We note that we can cover $\mathbb{C}^n$ by a finite number of such cones $\{C^{\epsilon}_{\zeta}\}$. The number of cones that are needed depends on ϵ and n. In the following, we fix one C^{ϵ}_{ζ}. We let

$$\gamma^{\epsilon}_{\lambda,\zeta}(x,t) := 1_{C^{\epsilon}_{\zeta}}(\mathcal{U}_\lambda w^{-1} A(x,t)) \cdot \mathcal{U}_\lambda w^{-1} A(x,t)$$

and consider a fixed dyadic parabolic cube $\Delta = Q \times I \subset \mathbb{R}^{n+1}$.

Step 1: Estimate of the test function along $\bar{\zeta}$. We first estimate

$$\iint_{\Delta} (1 - \nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta) \, \mathrm{d}x \, \mathrm{d}t. \tag{8-8}$$

To start the estimate, we write

$$1 - \nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta = \nabla_x g^{\zeta}_{\Delta,\epsilon} \cdot \zeta + (1 - \nabla_x L^{\zeta}_{\Delta} \cdot \zeta),$$

where $g^{\zeta}_{\Delta,\epsilon} := L^{\zeta}_{\Delta} - f^{\zeta}_{\Delta,\epsilon}$. By construction, we have $\nabla_x L^{\zeta}_{\Delta}(x,t) = \bar{\zeta}$ whenever $(x,t) \in \Delta$. Hence,

$$\iint_{\Delta} (1 - \nabla_x L^{\zeta}_{\Delta} \cdot \zeta) \, \mathrm{d}x \, \mathrm{d}t = 0.$$

We have to estimate the contribution to the integral in (8-8) coming from $\nabla_x g^{\zeta}_{\Delta,\epsilon} \cdot \zeta$. To do this, let $s \in (0,1)$ yet to be chosen, and let $\varphi \colon \mathbb{R}^{n+1} \to [0,1]$ be a smooth function which is 1 on $\Delta_s := (1-s)Q \times (1-s^2)I$, supported on Δ, and satisfies $\|\nabla_x \varphi\|_\infty \le c(s\ell(\Delta))^{-1}$, $\|\partial_t \varphi\|_\infty \le c(s\ell(\Delta))^{-2}$ for a dimensional constant $c > 0$. Using φ, we see that

$$\iint_{\Delta} \nabla_x g^{\zeta}_{\Delta,\epsilon} \cdot \zeta \, \mathrm{d}x \, \mathrm{d}t = \iint_{\Delta} (1-\varphi) \nabla_x g^{\zeta}_{\Delta,\epsilon} \cdot \zeta \, \mathrm{d}x \, \mathrm{d}t + \iint_{\Delta} \varphi \nabla_x g^{\zeta}_{\Delta,\epsilon} \cdot \zeta \, \mathrm{d}x \, \mathrm{d}t =: \mathrm{I} + \mathrm{II}.$$

Using the Cauchy–Schwarz inequality, Lemma 8.4 (ii) and (2-2) for the A_2-weight $\mu^{-1}(x,t) = w^{-1}(x)$, we obtain

$$\begin{aligned} |\mathrm{I}| &\le \left(\iint_{\Delta} |1-\varphi|^2 \, \mathrm{d}\mu^{-1} \right)^{1/2} \left(\iint_{\Delta} |\nabla_x g^{\zeta}_{\Delta,\epsilon}|^2 \, \mathrm{d}\mu \right)^{1/2} \\ &\lesssim \mu^{-1}(\Delta \setminus \Delta_s)^{1/2} \mu(\Delta)^{1/2} \lesssim s^{\eta} \mu^{-1}(\Delta)^{1/2} \mu(\Delta)^{1/2} \le s^{\eta} [w]_{A_2} |\Delta|. \end{aligned}$$

To estimate II, we integrate by parts to get

$$\mathrm{II} = -\iint_{\mathbb{R}^{n+1}} g^{\zeta}_{\Delta,\epsilon} \nabla_x \varphi \cdot \zeta \,\mathrm{d}x\,\mathrm{d}t,$$

and using the Cauchy–Schwarz inequality and Lemma 8.4 (i), we obtain similarly

$$\begin{aligned}|\mathrm{II}| &\le \Big(\iint_{\mathbb{R}^{n+1}} |\nabla_x \varphi|^2 \,\mathrm{d}\mu^{-1}\Big)^{1/2} \Big(\iint_{\mathbb{R}^{n+1}} |g^{\zeta}_{\Delta,\epsilon}|^2 \,\mathrm{d}\mu\Big)^{1/2} \\ &\lesssim (s\ell(\Delta))^{-1} \mu(\Delta)^{1/2} \epsilon \ell(\Delta) \mu^{-1}(\Delta)^{1/2} \le \epsilon s^{-1} [w]_{A_2} |\Delta|.\end{aligned}$$

We now choose $s = \epsilon^{1/(\eta+1)}$, so that the estimates for I and II come with the same power of ϵ. Putting the estimates together, we obtain, for the integral in (8-8), that

$$\frac{1}{|\Delta|} \Big| \iint_{\Delta} 1 - \nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta \,\mathrm{d}x\,\mathrm{d}t \Big| \lesssim \epsilon^{\eta/(\eta+1)}. \tag{8-9}$$

Using Lemma 8.4 (iii) and the Cauchy–Schwarz inequality, we also see that

$$\frac{1}{|\Delta|} \iint_{\Delta} |\nabla_x f^{\zeta}_{\Delta,\epsilon}| \,\mathrm{d}x\,\mathrm{d}t \le \frac{1}{|\Delta|} \Big(\iint_{\Delta} |\nabla_x f^{\zeta}_{\Delta,\epsilon}|^2 \,\mathrm{d}\mu\Big)^{1/2} \mu^{-1}(\Delta)^{1/2} \lesssim 1. \tag{8-10}$$

Step 2: Choice of ϵ. Using the estimates in the last two displays, we see, if ϵ is chosen small enough, that

$$\frac{1}{|\Delta|} \iint_{\Delta} \mathrm{Re}(\nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta) \,\mathrm{d}x\,\mathrm{d}t \ge \tfrac{7}{8}$$

and

$$\frac{1}{|\Delta|} \iint_{\Delta} |\nabla_x f^{\zeta}_{\Delta,\epsilon}| \,\mathrm{d}x\,\mathrm{d}t \le c$$

for some large constant c depending only on the structural constants. We now perform a stopping-time decomposition as in [Auscher et al. 2002] to select a collection $\mathcal{S}'_{\zeta} = \{\Delta'\}$ of dyadic parabolic subcubes of Δ, which are maximal with respect to the property that either

$$\frac{1}{|\Delta'|} \iint_{\Delta'} \mathrm{Re}(\nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta) \,\mathrm{d}x\,\mathrm{d}t \le \tfrac{3}{4} \tag{8-11}$$

or

$$\frac{1}{|\Delta'|} \iint_{\Delta'} |\nabla_x f^{\zeta}_{\Delta,\epsilon}| \,\mathrm{d}x\,\mathrm{d}t \ge (4\epsilon)^{-2} \tag{8-12}$$

holds. In other words, we parabolically dyadically subdivide Δ and stop the first time either (8-11) or (8-12) hold. Then, $\mathcal{S}'_{\zeta} = \{\Delta'\}$ is a disjoint set of the parabolic dyadic subcubes of Δ. Let $\mathcal{S}''_{\zeta} = \{\Delta''\}$ be the collection of all the parabolic dyadic subcubes of Δ not contained in any $\Delta' \in \mathcal{S}'_{\zeta}$. Then, each $\Delta'' \in \mathcal{S}''_{\zeta}$ satisfies

$$\begin{aligned}\frac{1}{|\Delta''|} \iint_{\Delta''} \mathrm{Re}(\nabla_x f^{\zeta}_{\Delta,\epsilon} \cdot \zeta) \,\mathrm{d}x\,\mathrm{d}t &\ge \tfrac{3}{4}, \\ \frac{1}{|\Delta''|} \iint_{\Delta''} |\nabla_x f^{\zeta}_{\Delta,\epsilon}| \,\mathrm{d}x\,\mathrm{d}t &\le (4\epsilon)^{-2}.\end{aligned} \tag{8-13}$$

At this stage, we claim that, by the same type of argument as in the proof of statement (i) in Proposition 5.7 in [Auscher et al. 2002], there exists $\epsilon \in (0, 1)$ even smaller and depending only on the structural constants and $\eta' = \eta'(\epsilon) \in (0, 1)$ such that

$$\Bigl| \bigcup_{\Delta' \in \mathcal{S}'_\zeta} \Delta' \Bigr| \le (1 - \eta')|\Delta|. \tag{8-14}$$

In particular, from now on ϵ is fixed. For completeness and the convenience of the reader, we include a proof here.

Let E_1 and E_2 be the unions of all parabolic cubes in $\mathcal{S}'_\zeta$ which satisfy (8-11) and (8-12), respectively. Then,

$$\Bigl| \bigcup_{\Delta' \in \mathcal{S}'_\zeta} \Delta' \Bigr| \le |E_1| + |E_2|.$$

For $|E_2|$, we have

$$|E_2| \le (4\epsilon)^2 \sum_{\Delta' \in \mathcal{S}'_\zeta} \iint_{\Delta'} |\nabla_x f^\zeta_{\Delta,\epsilon}| \,\mathrm{d}x\,\mathrm{d}t \le (4\epsilon)^2 \iint_{\Delta} |\nabla_x f^\zeta_{\Delta,\epsilon}| \,\mathrm{d}x\,\mathrm{d}t \le (4\epsilon)^2 c |\Delta|,$$

where we used (8-10) in the last step. To control $|E_1|$, we let $h := 1 - \mathrm{Re}(\nabla_x f^\zeta_{\Delta,\epsilon} \cdot \zeta)$ and write

$$|E_1| \le 4 \sum_{\Delta'} \iint_{\Delta'} h \,\mathrm{d}x\,\mathrm{d}t = 4 \iint_{\Delta} h \,\mathrm{d}x\,\mathrm{d}t - 4 \iint_{\Delta \setminus E_1} h \,\mathrm{d}x\,\mathrm{d}t, \tag{8-15}$$

where the sum is taken over all parabolic subcubes of E_1. By (8-9), the first term on the right is controlled by $\epsilon^{\eta/(\eta+1)}|\Delta|$ times a constant depending on the structural constants. Using in succession the Cauchy–Schwarz inequality, Lemma 8.4 (iii), the A_2-property and Young's inequality, the second term on the right is controlled by

$$\begin{aligned} 4|\Delta \setminus E_1| + 4\mu^{-1}(\Delta \setminus E_1)^{1/2} \Bigl(\iint_{\Delta} |\nabla_x f^\zeta_{\Delta,\epsilon}|^2 \,\mathrm{d}\mu \Bigr)^{1/2} &\le 4|\Delta \setminus E_1| + 4\tilde{c}\mu^{-1}(\Delta \setminus E_1)^{1/2} \mu(\Delta)^{1/2} \\ &\le 4|\Delta \setminus E_1| + 4\tilde{c}|\Delta \setminus E_1|^{\eta} |\Delta|^{1-\eta} \\ &\le (4 + \tilde{c}\epsilon^{-1/\eta})|\Delta \setminus E_1| + \tilde{c}\epsilon^{1-\eta}|\Delta|, \end{aligned}$$

where $\tilde{c}$ depends on the structural constants and changes from line to line. Going back to (8-15) and rearranging terms, we find

$$|E_1| \le \frac{4 + \tilde{c}\epsilon^{-1/\eta} + \tilde{c}(\epsilon^{\eta/(\eta+1)} + \epsilon^{1-\eta})}{5 + \tilde{c}\epsilon^{-1/\eta}} |\Delta|,$$

and, taking ϵ small enough, we conclude (8-14).

Since μ is an A_2-weight, we obtain from (8-14) — and upon taking η' smaller depending on the structural constants and ϵ — that

$$\mu\Bigl(\bigcup_{\Delta' \in \mathcal{S}'_\zeta} \Delta' \Bigr) \le (1 - \eta')\mu(\Delta); \tag{8-16}$$

see, for example, [Stein 1993, p. 196] for this A_∞-property of A_2-weights.

Step 3: Reintroducing the averaging operator. Given Δ, we consider $\Delta'' \in \mathcal{S}''_\zeta$ as above. Set

$$v := \frac{1}{\mu(\Delta'')} \iint_{\Delta''} \nabla_x f^\zeta_{\Delta,\epsilon} \, dx \, dt \in \mathbb{C}^n. \tag{8-17}$$

If $(x,t) \in \Delta''$ and $\frac{1}{2}\ell(\Delta'') < \lambda \le \ell(\Delta'')$, then $v = (\mathcal{A}_\lambda \nabla_x f^\zeta_{\Delta,\epsilon})(x,t)$. Assume that

$$u := (\mathcal{U}_\lambda w^{-1} A)(x,t) \in C^\epsilon_\zeta.$$

The pair of vectors (u, v) satisfies the estimates in (8-13). Thus, we can apply [Auscher et al. 2002, Lemma 5.10] with $w := \zeta$ and conclude that $|u| \le 4|u \cdot v|$; that is,

$$|\gamma^\epsilon_{\lambda,\zeta}(x,t)| \le 4|(\mathcal{U}_\lambda w^{-1} A(x,t)) \cdot (\mathcal{A}_\lambda \nabla_x f^\zeta_{\Delta,\epsilon})(x,t)|. \tag{8-18}$$

We next observe that, by construction, the Carleson box $\Delta \times (0, \ell(\Delta)]$ can be partitioned into Carleson boxes $\Delta' \times (0, \ell(\Delta')]$, with $\Delta' \in \mathcal{S}'_\zeta$, and Whitney boxes $\Delta'' \times \left(\frac{1}{2}\ell(\Delta''), \ell(\Delta'')\right]$, with $\Delta'' \in \mathcal{S}''_\zeta$. In particular,

$$\frac{1}{\mu(\Delta)} \int_0^{\ell(\Delta)} \iint_\Delta |\gamma^\epsilon_{\lambda,\zeta}(x,t)|^2 \frac{d\mu \, d\lambda}{\lambda} =: \mathrm{I} + \mathrm{II},$$

where

$$\mathrm{I} := \frac{1}{\mu(\Delta)} \sum_{\Delta' \in \mathcal{S}'_\zeta} \int_0^{\ell(\Delta')} \iint_{\Delta'} |\gamma^\epsilon_{\lambda,\zeta}(x,t)|^2 \frac{d\mu \, d\lambda}{\lambda},$$

$$\mathrm{II} := \frac{1}{\mu(\Delta)} \sum_{\Delta'' \in \mathcal{S}''_\zeta} \int_{\ell(\Delta'')/2}^{\ell(\Delta'')} \iint_{\Delta''} |\gamma^\epsilon_{\lambda,\zeta}(x,t)|^2 \frac{d\mu \, d\lambda}{\lambda}.$$

Using (8-16), we obtain

$$\mathrm{I} \le \frac{1}{\mu(\Delta)} \sum_{\Delta' \in \mathcal{S}'_\zeta} A^\epsilon_\zeta \mu(\Delta') \le (1-\eta') A^\epsilon_\zeta,$$

where

$$A^\epsilon_\zeta := \sup_{\tilde{\Delta}} \frac{1}{\mu(\tilde{\Delta})} \int_0^{\ell(\tilde{\Delta})} \iint_{\tilde{\Delta}} |\gamma^\epsilon_{\lambda,\zeta}(x,t)|^2 \frac{d\mu \, d\lambda}{\lambda},$$

and where the supremum is taken over all dyadic parabolic subcubes $\tilde{\Delta} \subset \Delta$. By (8-18), we have

$$\mathrm{II} \le \frac{16}{\mu(\Delta)} \int_0^{\ell(\Delta)} \iint_\Delta |(\mathcal{U}_\lambda w^{-1} A)(x,t) \cdot (\mathcal{A}_\lambda \nabla_x f^\zeta_{\Delta,\epsilon})(x,t)|^2 \frac{d\mu \, d\lambda}{\lambda}.$$

Since these estimates hold for all dyadic parabolic cubes, in particular those which are subcubes of Δ, we conclude that

$$A^\epsilon_\zeta \le (1-\eta') A^\epsilon_\zeta + \sup_{\tilde{\Delta}} \frac{16}{\mu(\tilde{\Delta})} \int_0^{\ell(\tilde{\Delta})} \iint_{\tilde{\Delta}} |(\mathcal{U}_\lambda w^{-1} A)(x,t) \cdot (\mathcal{A}_\lambda \nabla_x f^\zeta_{\tilde{\Delta},\epsilon})(x,t)|^2 \frac{d\mu \, d\lambda}{\lambda}.$$

Summing with respect to $\zeta \in W$ completes the proof of Lemma 8.5 under the a priori assumption that A^ϵ_ζ is qualitatively finite, since it can then be absorbed into the left-hand side.

Step 4: Removing the a priori assumption. The a priori assumption that A^{ϵ}_{ζ} is qualitatively finite can be removed by setting $\gamma^{\epsilon}_{\lambda,\zeta}(x,t)$ to 0 for λ small and large, repeating the argument from (8-18) on and passing to the limit at the end. For the truncated $\gamma^{\epsilon}_{\lambda,\zeta}(x,t)$, we get $A^{\epsilon}_{\zeta} < \infty$ from (7-3). Indeed, for $0 < \delta < 1$ small, we have

$$A^{\epsilon}_{\zeta} \leq \int_{\delta}^{\delta^{-1}} \left(\sup_{\tilde{\Delta}} \frac{1}{\mu(\tilde{\Delta})} \iint_{\tilde{\Delta}} |\mathcal{U}_{\lambda} w^{-1} A|^2 \,\mathrm{d}\mu \right) \frac{\mathrm{d}\lambda}{\lambda} \leq \int_{\delta}^{\delta^{-1}} C \frac{\mathrm{d}\lambda}{\lambda} < \infty,$$

where C depends on $\ell(\Delta)$ and δ. This completes the argument. □

8.2. ***The Carleson measure estimate: proof of Lemma 8.3.*** Thanks to Lemma 8.5, it suffices to prove

$$\int_0^{\ell(\Delta)} \iint_{\Delta} |(\mathcal{U}_{\lambda} w^{-1} A) \cdot \mathcal{A}_{\lambda} \nabla_x f^{\zeta}_{\Delta,\epsilon}|^2 \frac{\mathrm{d}\mu \,\mathrm{d}\lambda}{\lambda} \lesssim \mu(\Delta). \tag{8-19}$$

The left-hand side in (8-19) is bounded by

$$|||(\lambda \mathcal{E}_{\lambda} \mathcal{H} + (\mathcal{U}_{\lambda} w^{-1} A) \cdot \mathcal{A}_{\lambda} \nabla_x) f^{\zeta}_{\Delta,\epsilon}|||^2_{2,\mu} + \int_0^{\ell(\Delta)} \iint_{\Delta} |\lambda \mathcal{E}_{\lambda} \mathcal{H} f^{\zeta}_{\Delta,\epsilon}|^2 \frac{\mathrm{d}\mu \,\mathrm{d}\lambda}{\lambda} =: \mathrm{I} + \mathrm{II}.$$

By Proposition 8.1 and Lemma 8.4, we have

$$\mathrm{I} \lesssim \|\mathbb{D} f^{\zeta}_{\Delta,\epsilon}\|^2_{2,\mu} \lesssim \mu(\Delta).$$

As for II, we obtain from (8-6) that

$$\mathcal{H} f^{\zeta}_{\Delta,\epsilon} = \frac{(L^{\zeta}_{\Delta} - f^{\zeta}_{\Delta,\epsilon})}{(\epsilon \ell(\Delta))^2}.$$

Using the L^2_{μ}-boundedness of $\mathcal{E}_{\lambda}$, see Lemma 4.3, and then Lemma 8.4, we obtain

$$\begin{aligned} \mathrm{II} &\lesssim \int_0^{\ell(\Delta)} \|\lambda (\epsilon \ell(\Delta))^{-2} (L^{\zeta}_{\Delta} - f^{\zeta}_{\Delta,\epsilon})\|^2_{2,\mu} \frac{\mathrm{d}\lambda}{\lambda} \\ &= \frac{1}{2\epsilon^4 \ell(\Delta)^2} \|L^{\zeta}_{\Delta} - f^{\zeta}_{\Delta,\epsilon}\|^2_{2,\mu} \lesssim \epsilon^{-2} \mu(\Delta). \end{aligned}$$

This completes the proof of (8-19), and hence the proof of Theorem 1.1.

References

[Alfonseca et al. 2011] M. A. Alfonseca, P. Auscher, A. Axelsson, S. Hofmann, and S. Kim, "Analyticity of layer potentials and L^2 solvability of boundary value problems for divergence form elliptic equations with complex L^{∞} coefficients", *Adv. Math.* **226**:5 (2011), 4533–4606. MR Zbl

[Amenta and Auscher 2018] A. Amenta and P. Auscher, *Elliptic boundary value problems with fractional regularity data: the first order approach*, CRM Monogr. Ser. **37**, Amer. Math. Soc., Providence, RI, 2018. MR Zbl

[Auscher and Axelsson 2011] P. Auscher and A. Axelsson, "Weighted maximal regularity estimates and solvability of non-smooth elliptic systems, I", *Invent. Math.* **184**:1 (2011), 47–115. MR Zbl

[Auscher and Egert 2016] P. Auscher and M. Egert, "On non-autonomous maximal regularity for elliptic operators in divergence form", *Arch. Math.* (*Basel*) **107**:3 (2016), 271–284. MR Zbl

[Auscher and Mourgoglou 2019] P. Auscher and M. Mourgoglou, "Representation and uniqueness for boundary value elliptic problems via first order systems", *Rev. Mat. Iberoam.* **35**:1 (2019), 241–315. MR Zbl

[Auscher and Rosén 2012] P. Auscher and A. Rosén, "Weighted maximal regularity estimates and solvability of nonsmooth elliptic systems, II", *Anal. PDE* **5**:5 (2012), 983–1061. MR Zbl

[Auscher and Tchamitchian 1998] P. Auscher and P. Tchamitchian, *Square root problem for divergence operators and related topics*, Astérisque **249**, Soc. Math. France, Paris, 1998. MR Zbl

[Auscher et al. 2002] P. Auscher, S. Hofmann, M. Lacey, A. McIntosh, and P. Tchamitchian, "The solution of the Kato square root problem for second order elliptic operators on $\mathbb{R}^n$", *Ann. of Math.* (2) **156**:2 (2002), 633–654. MR Zbl

[Auscher et al. 2015] P. Auscher, A. Rosén, and D. Rule, "Boundary value problems for degenerate elliptic equations and systems", *Ann. Sci. École Norm. Sup.* (4) **48**:4 (2015), 951–1000. MR Zbl

[Auscher et al. 2018] P. Auscher, M. Egert, and K. Nyström, "The Dirichlet problem for second order parabolic operators in divergence form", *J. Éc. Polytech. Math.* **5** (2018), 407–441. MR Zbl

[Auscher et al. 2020] P. Auscher, M. Egert, and K. Nyström, "L^2 well-posedness of boundary value problems for parabolic systems with measurable coefficients", *J. Eur. Math. Soc.* **22**:9 (2020), 2943–3058. MR Zbl

[Castro et al. 2016] A. J. Castro, K. Nyström, and O. Sande, "Boundedness of single layer potentials associated to divergence form parabolic equations with complex coefficients", *Calc. Var. Partial Differential Equations* **55**:5 (2016), art. id. 124. MR Zbl

[Chiarenza and Serapioni 1985] F. Chiarenza and R. Serapioni, "A remark on a Harnack inequality for degenerate parabolic equations", *Rend. Sem. Mat. Univ. Padova* **73** (1985), 179–190. MR Zbl

[Cruz-Uribe and Rios 2008] D. Cruz-Uribe and C. Rios, "Gaussian bounds for degenerate parabolic equations", *J. Funct. Anal.* **255**:2 (2008), 283–312. Correction in **267**:9 (2014), 3507–3513. MR Zbl

[Cruz-Uribe and Rios 2012] D. Cruz-Uribe and C. Rios, "The solution of the Kato problem for degenerate elliptic operators with Gaussian bounds", *Trans. Amer. Math. Soc.* **364**:7 (2012), 3449–3478. MR Zbl

[Cruz-Uribe and Rios 2015] D. Cruz-Uribe and C. Rios, "The Kato problem for operators with weighted ellipticity", *Trans. Amer. Math. Soc.* **367**:7 (2015), 4727–4756. MR Zbl

[Cruz-Uribe et al. 2018] D. Cruz-Uribe, J. M. Martell, and C. Rios, "On the Kato problem and extensions for degenerate elliptic operators", *Anal. PDE* **11**:3 (2018), 609–660. MR Zbl

[David et al. 1985] G. David, J.-L. Journé, and S. Semmes, "Opérateurs de Calderón–Zygmund, fonctions para-accrétives et interpolation", *Rev. Mat. Iberoam.* **1**:4 (1985), 1–56. 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

[Dier and Zacher 2017] D. Dier and R. Zacher, "Non-autonomous maximal regularity in Hilbert spaces", *J. Evol. Equ.* **17**:3 (2017), 883–907. MR Zbl

[Escauriaza and Hofmann 2018] L. Escauriaza and S. Hofmann, "Kato square root problem with unbounded leading coefficients", *Proc. Amer. Math. Soc.* **146**:12 (2018), 5295–5310. MR Zbl

[García-Cuerva and Rubio de Francia 1985] J. García-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland Math. Stud. **116**, North-Holland, Amsterdam, 1985. MR Zbl

[Haase 2006] M. Haase, *The functional calculus for sectorial operators*, Operator Theory: Adv. Appl. **169**, Birkhäuser, Basel, 2006. MR Zbl

[Heinonen et al. 1993] J. Heinonen, T. Kilpeläinen, and O. Martio, *Nonlinear potential theory of degenerate elliptic equations*, Oxford Univ. Press, 1993. MR Zbl

[Hofmann and Lewis 2005] S. Hofmann and J. L. Lewis, "The L^p Neumann problem for the heat equation in non-cylindrical domains", *J. Funct. Anal.* **220**:1 (2005), 1–54. MR Zbl

[Hofmann et al. 2015] S. Hofmann, C. Kenig, S. Mayboroda, and J. Pipher, "Square function/non-tangential maximal function estimates and the Dirichlet problem for non-symmetric elliptic operators", *J. Amer. Math. Soc.* **28**:2 (2015), 483–529. MR Zbl

[Hofmann et al. 2019] S. Hofmann, P. Le, and A. J. Morris, "Carleson measure estimates and the Dirichlet problem for degenerate elliptic equations", *Anal. PDE* **12**:8 (2019), 2095–2146. MR Zbl

[Hofmann et al. 2022] S. Hofmann, L. Li, S. Mayboroda, and J. Pipher, "The Dirichlet problem for elliptic operators having a BMO anti-symmetric part", *Math. Ann.* **382**:1-2 (2022), 103–168. MR Zbl

[Ishige et al. 2017] K. Ishige, Y. Kabeya, and E. M. Ouhabaz, "The heat kernel of a Schrödinger operator with inverse square potential", *Proc. Lond. Math. Soc.* (3) **115**:2 (2017), 381–410. MR Zbl

[Kaplan 1966] S. Kaplan, "Abstract boundary value problems for linear parabolic equations", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (3) **20**:2 (1966), 395–419. MR Zbl

[Kato 1966] T. Kato, *Perturbation theory for linear operators*, Grundl. Math. Wissen. **132**, Springer, 1966. MR Zbl

[Kilpeläinen 1994] T. Kilpeläinen, "Weighted Sobolev spaces and capacity", *Ann. Acad. Sci. Fenn. Ser. A I Math.* **19**:1 (1994), 95–113. MR Zbl

[Lions 1962] J.-L. Lions, "Espaces d'interpolation et domaines de puissances fractionnaires d'opérateurs", *J. Math. Soc. Japan* **14** (1962), 233–241. MR Zbl

[Litsgård and Nyström 2023] M. Litsgård and K. Nyström, "On local regularity estimates for fractional powers of parabolic operators with time-dependent measurable coefficients", *J. Evol. Equ.* **23**:1 (2023), art. id. 3. MR Zbl

[McIntosh 1986] A. McIntosh, "Operators which have an H_∞ functional calculus", pp. 210–231 in *Miniconference on operator theory and partial differential equations* (North Ryde, Australia, 1986), edited by B. Jefferies et al., Proc. Centre Math. Anal. Austral. Nat. Univ. **14**, Austral. Nat. Univ., Canberra, 1986. MR Zbl

[Morris 2012] A. J. Morris, "The Kato square root problem on submanifolds", *J. Lond. Math. Soc.* (2) **86**:3 (2012), 879–910. MR Zbl

[Nyström 2016] K. Nyström, "Square function estimates and the Kato problem for second order parabolic operators in $\mathbb{R}^{n+1}$", *Adv. Math.* **293** (2016), 1–36. MR Zbl

[Nyström 2017] K. Nyström, "L^2 solvability of boundary value problems for divergence form parabolic equations with complex coefficients", *J. Differential Equations* **262**:3 (2017), 2808–2939. MR Zbl

[Ouhabaz 2021] E. M. Ouhabaz, "The square root of a parabolic operator", *J. Fourier Anal. Appl.* **27**:3 (2021), art. id. 59. MR Zbl

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Math. Ser. **43**, Princeton Univ. Press, 1993. MR Zbl

Received 22 Sep 2022. Revised 30 Aug 2023. Accepted 19 Sep 2023.

ALIREZA ATAEI: alireza.ataei@math.uu.se
Department of Mathematics, Uppsala University, Uppsala, Sweden

MORITZ EGERT: egert@mathematik.tu-darmstadt.de
Fachbereich Mathematik, TU Darmstadt, Darmstadt, Germany

KAJ NYSTRÖM: kaj.nystrom@math.uu.se
Department of Mathematics, Uppsala University, Uppsala, Sweden

ANALYSIS AND PDE
Vol. 18 (2025), No. 1, pp. 171–198
DOI: 10.2140/apde.2025.18.171

SMALL SCALE FORMATION FOR THE 2-DIMENSIONAL BOUSSINESQ EQUATION

ALEXANDER KISELEV, JAEMIN PARK AND YAO YAO

We study the 2-dimensional incompressible Boussinesq equations without thermal diffusion, and aim to construct rigorous examples of small scale formations as time goes to infinity. In the viscous case, we construct examples of global smooth solutions satisfying $\sup_{\tau\in[0,t]}\|\nabla\rho(\tau)\|_{L^2} \gtrsim t^\alpha$ for some $\alpha > 0$. For the inviscid equation in the strip, we construct examples satisfying $\|\omega(t)\|_{L^\infty} \gtrsim t^3$ and $\sup_{\tau\in[0,t]}\|\nabla\rho(\tau)\|_{L^\infty} \gtrsim t^2$ during the existence of a smooth solution. These growth results hold for a broad class of initial data, where we only require certain symmetry and sign conditions. As an application, we also construct solutions to the 3-dimensional axisymmetric Euler equation whose velocity has infinite-in-time growth.

1. Introduction

The incompressible Boussinesq equations describe the motion of incompressible fluid under the influence of gravitational forces [Gill and Adrian 1982; Majda 2003; Pedlosky 1979]. Let us denote by $\rho(x,t)$ the density of the fluid (it can also represent the temperature, depending on the physical context) and $u(x,t)$ the velocity field. Throughout this paper, we consider the 2-dimensional incompressible Boussinesq equations in the absence of density/thermal diffusivity:

$$\begin{aligned} &\rho_t + u\cdot\nabla\rho = 0,\\ &u_t + u\cdot\nabla u = -\nabla p - \rho e_2 + \nu\Delta u, \quad x\in\Omega,\ \ t>0,\\ &\qquad\nabla\cdot u = 0, \end{aligned} \tag{1-1}$$

where the initial condition is $u(\,\cdot\,,0) = u_0$ and $\rho(\,\cdot\,,0) = \rho_0$. Here $e_2 := (0,1)^T$, and $\nu \ge 0$ is the viscosity coefficient. We assume the spatial domain Ω is one of the following: the whole space $\mathbb{R}^2$, the torus $\mathbb{T}^2 := (-\pi,\pi]^2$, or the strip $\mathbb{T}\times[0,\pi]$ that is periodic in x_1. When Ω is the strip, we impose the no-slip boundary condition $u|_{\partial\Omega} = 0$ if $\nu > 0$, and the no-flow boundary condition $u\cdot n|_{\partial\Omega} = 0$ if $\nu = 0$.

In the past decade, much progress has been made on the analysis of (1-1) in both the viscous case $\nu > 0$ and inviscid case $\nu = 0$. Below we briefly review the relevant literature and state our main results in each case.

1.1. ***The viscous case $\nu > 0$.*** If the equation for ρ has an additional thermal diffusion term $\kappa\Delta\rho$, global regularity of solutions is well known (see, e.g., [Temam 1988]) and follows from the classical methods for Navier–Stokes equations. In the absence of thermal diffusion, the first global-in-time regularity results were obtained by Hou and Li [2005] in the space $(u,\rho)\in H^m(\mathbb{R}^2)\times H^{m-1}(\mathbb{R}^2)$ for $m\ge 3$, and by

MSC2020: 35Q35, 76B03, 76D03.
Keywords: small scale creation, 2-dimensional Boussinesq system.

Chae [2006] in the space $H^m(\mathbb{R}^2) \times H^m(\mathbb{R}^2)$ for $m \geq 3$. When $\Omega \subset \mathbb{R}^2$ is a bounded domain, Lai, Pan, and Zhao [Lai et al. 2011] proved global well-posedness of solutions in $H^3(\Omega) \times H^3(\Omega)$ with the no-slip boundary condition, and showed that the kinetic energy is uniformly bounded in time. The function space was improved by Hu, Kukavica, and Ziane [Hu et al. 2013] to $(u, \rho) \in H^m(\Omega) \times H^{m-1}(\Omega)$ for $m \geq 2$, where Ω is either a bounded domain, $\mathbb{R}^2$, or $\mathbb{T}^2$. In spaces with lower regularity, global well-posedness of weak solutions was obtained in [Abidi and Hmidi 2007; Danchin and Paicu 2011; Hmidi and Keraani 2007; Larios et al. 2013]. For the temperature patch problem, Gancedo and García-Juárez [2017; 2020] proved global regularity in two dimensions and local regularity in three dimensions.

Regarding upper bounds of the global-in-time solutions, for a bounded domain, Ju [2017] obtained that $\|\rho\|_{H^1(\Omega)} \lesssim e^{Ct^2}$. The e^{Ct^2} bound was improved to an exponential bound e^{Ct} in [Kukavica and Wang 2020] for $\Omega = \mathbb{T}^2$ or a bounded domain, and a super-exponential bound $e^{Ct^{(1+\beta)}}$ for some constant $\beta \approx 0.29$ for $\Omega = \mathbb{R}^2$. When $\Omega = \mathbb{T}^2$, they also obtained the uniform-in-time bound $\|u\|_{W^{2,p}(\mathbb{T}^2)} \leq C(p)$ for all $p \in [2, \infty)$. In recent work by Kukavica, Massatt, and Ziane [Kukavica et al. 2023], when Ω is a bounded domain, the upper bound of the norm of ρ has been improved to $\|\rho\|_{H^2(\Omega)} \leq C_\epsilon e^{\epsilon t}$ for all $\epsilon > 0$, and they also showed $\|u\|_{H^3} \leq C_\epsilon e^{\epsilon t}$ for all $\epsilon > 0$.

We would like to point out that all these results deal with *upper bounds* of solutions, and it is a natural question whether certain norms of solutions *can* actually grow to infinity as $t \to \infty$. When $\nu > 0$ and $\Omega = \mathbb{R}^2$, Brandolese and Schonbek [2012] proved that when the initial data ρ_0 does not have mean zero, $\|u(t)\|_{L^2(\mathbb{R}^2)}$ must grow to infinity like $(1+t)^{1/4}$. Here the growth mechanism is due to potential energy converting into kinetic energy, and does not necessarily imply growth in higher derivatives of u or ρ. To the best of our knowledge, there has been no example in the literature showing that $\|\rho(t)\|_{\dot{H}^m}$ or $\|u(t)\|_{\dot{H}^m}$ can actually grow to infinity as $t \to \infty$ for some $m \geq 1$. The goal of this paper is exactly to construct such examples in $\mathbb{R}^2$ and $\mathbb{T}^2$, where $\|\rho(t)\|_{\dot{H}^m} \to \infty$ as $t \to \infty$ for all $m \geq 1$. Since $\|\rho(t)\|_{L^2}$ is preserved in time, growth of $\|\rho(t)\|_{\dot{H}^m}$ implies that ρ has some small scale formation as $t \to \infty$.

In the viscous case, we set the spatial domain to be either $\mathbb{R}^2$ or $\mathbb{T}^2$, and assume that the initial data (ρ_0, u_0) satisfies the following assumptions (here we write $u_0 = (u_{01}, u_{02})^T$). See Figure 1 for an illustration of the assumptions on ρ_0.

(A1) $\rho_0, u_0 \in C^\infty(\Omega)$. If $\Omega = \mathbb{R}^2$, assume in addition that $\rho_0, u_0 \in C_c^\infty(\mathbb{R}^2)$.

(A2) ρ_0 and u_{02} are odd in x_2, and u_{01} is even in x_2. If $\Omega = \mathbb{T}^2$, assume in addition that ρ_0 and u_{02} are even in x_1, u_{01} is odd in x_1, and $\rho_0 = 0$ on the x_2-axis.[1]

(A3) ρ_0 is not identically zero, and $\rho_0 \geq 0$ for $x_2 \geq 0$.

As we show in Section 2.1, under these assumptions, both the potential energy $E_P(t) := \int_\Omega \rho(x,t) x_2 \, dx$ and kinetic energy $E_K(t) = \frac{1}{2}\|u(t)\|_{L^2(\Omega)}^2$ of the solution remain bounded for all times, and the total energy is decreasing in time. We prove that, for all $s \geq 1$, the Sobolev norm $\|\rho(t)\|_{\dot{H}^s}$ grows to infinity at least algebraically in t.

[1]Note that if the $\rho_0 = 0$ on the x_2-axis assumption is removed, the initial data would include some steady states with horizontally stratified density, which clearly would not lead to any growth.

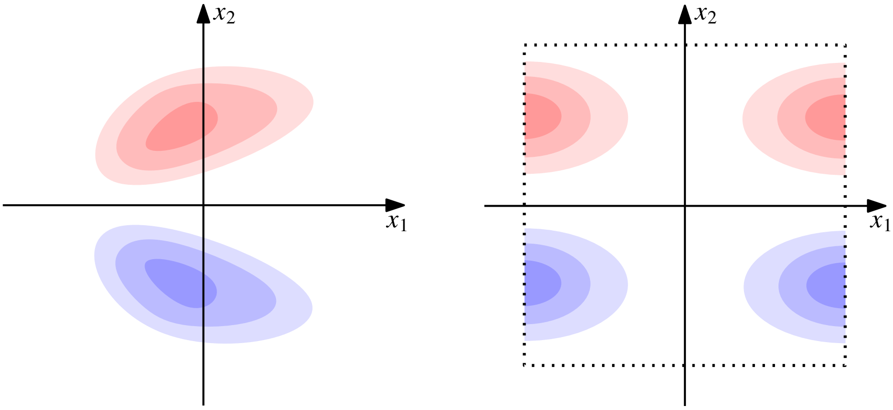

Figure 1. Illustration of the symmetry and sign assumptions on ρ_0 in the plane $\mathbb{R}^2$ (left) and torus $\mathbb{T}^2$ (right) for the viscous Boussinesq equations. Here red denotes positive ρ_0 and blue denotes negative ρ_0.

Theorem 1.1. *Assume* $\nu > 0$, *and let* $\Omega = \mathbb{R}^2$ *or* $\mathbb{T}^2$. *For any initial data* (ρ_0, u_0) *satisfying* (A1)–(A3), *the global-in-time smooth solution* (ρ, u) *to* (1-1) *satisfies the following*:

- *If* $\Omega = \mathbb{R}^2$, *we have*

$$\limsup_{t\to\infty} t^{-s/10} \|\rho(t)\|_{\dot{H}^s(\Omega)} = +\infty \quad \textit{for all } s \geq 1. \tag{1-2}$$

- *If* $\Omega = \mathbb{T}^2$, *we have*

$$\limsup_{t\to\infty} t^{-s(2s-1)/(8s-2)} \|\rho(t)\|_{\dot{H}^s(\Omega)} = +\infty \quad \textit{for all } s \geq 1. \tag{1-3}$$

Remark 1.2. It is a natural question whether these growth rates are sharp. While the powers are likely nonsharp, we point out that $\|\rho(t)\|_{H^1}$ cannot have exponential growth under the assumptions (A1)–(A3). Namely, following arguments similar to [Kukavica and Wang 2020], we show in Proposition 2.4 that, under the assumptions (A1)–(A3), $\|\rho(t)\|_{H^1}$ has a refined subexponential upper bound

$$\|\rho(t)\|_{H^1(\Omega)} \lesssim \exp(Ct^\alpha) \quad \text{for all } t > 0$$

for some constant $\alpha \in (0, 1)$. Therefore in this setting, the fastest possible growth rate of $\|\rho(t)\|_{H^1(\Omega)}$ is somewhere between algebraic and subexponential.

The proof of Theorem 1.1 is motivated by a recent result on small scale formation in solutions to incompressible porous media (IPM) equation by the first and third author [Kiselev and Yao 2023]. The main idea there was to use the monotonicity of the potential energy $E_P(t) = \int \rho(x,t)x_2\,dx$: on the one hand, for solutions with certain symmetries, $E_P(t)$ is bounded below with $E_P'(t) = -\|\partial_1\rho(t)\|^2_{\dot{H}^{-1}}$, thus the integral $\int_0^\infty \|\partial_1\rho(t)\|^2_{\dot{H}^{-1}}\,dt$ is finite; on the other hand, under certain symmetries, one can show that $\|\partial_1\rho(t)\|^2_{\dot{H}^{-1}}$ can only be small if $\|\rho(t)\|_{H^s} \gg 1$ for some $s > 0$, leading to growth of ρ in Sobolev norms.

The IPM and Boussinesq equations are related in the sense that, in both equations, the density ρ is transported by an incompressible u, where $u = -\nabla p - \rho e_2$ in IPM, whereas $Du/Dt = -\nabla p - \rho e_2 + \nu\Delta u$ in Boussinesq equations. Since the velocity in Boussinesq equations has one more time derivative than

IPM, we formally expect that $E_P''(t)$ should be related to $-\|\partial_1\rho(t)\|_{\dot H^{-1}}^2$. While this turns out to be true, the situation is more delicate for the Boussinesq equations because $E_P''(t)$ also contains other terms coming from the pressure and viscosity terms. By carefully controlling these additional terms, we prove that if $\|\rho(t)\|_{H^s}$ grows too slowly for $s \ge 1$, $E_P'(t)$ would become unbounded below, contradicting the uniform-in-time bound of energy.

1.2. ***The inviscid case $\nu = 0$.*** For the inviscid Boussinesq equations in two dimensions, it is well known that the system (1-1) can be rewritten into an equivalent system for the density ρ and the vorticity $\omega = \partial_1 u_2 - \partial_2 u_1$:

$$\begin{aligned} \rho_t + u\cdot\nabla\rho &= 0, \\ \omega_t + u\cdot\nabla\omega &= -\partial_1\rho, \end{aligned} \tag{1-4}$$

where the velocity u can be recovered from the vorticity ω from the Biot–Savart law $u = \nabla^\perp(-\Delta)^{-1}\omega$. While local well-posedness results are available in a variety of functional spaces for $\Omega = \mathbb{R}^2,\ \mathbb{T}^2$, or a bounded domain [Chae and Nam 1997; Chae et al. 1999; Danchin 2013], whether smooth initial data in $\mathbb{T}^2$ or $\mathbb{R}^2$ with finite energy can develop a finite-time singularity is an outstanding open question in fluid dynamics. Note that smooth, infinite-energy initial data can lead to a finite-time blowup, as shown in [Sarria and Wu 2015].

In the presence of boundary, there have been many exciting developments regarding finite-time singularity formation of solutions in the past few years. Luo and Hou [2014] provided numerical evidence for finite-time blowup in smooth solutions of the 3-dimensional axisymmetric Euler equation in a cylinder. When the domain has a corner, Elgindi and Jeong [2020] proved that blow-up can happen for inviscid Boussinesq equations with smooth initial data. When $\Omega = \mathbb{R}^2_+$ is the upper half-plane, Chen and Hou [2021] proved that solutions with $C^{1,\alpha}$ velocity and density can have a nearly self-similar finite-time blowup. Recently, for smooth initial data, Wang, Lai, Gómez-Serrano, and Buckmaster [Wang et al. 2023] used physics-informed neural networks to construct an approximate self-similar blow-up solution numerically. In a very recent preprint, Chen and Hou [2022] put forward an argument combining impressive analytical tools and computer assisted estimates to show that smooth initial data can lead to a stable nearly self-similar blowup.

Note that the inviscid Boussinesq equations (1-4) become the 2-dimensional Euler equation when $\rho \equiv 0$, where it is well known that $\|\nabla\omega(t)\|_{L^\infty}$ can have infinite-in-time growth [Denisov 2009; 2015; Kiselev and Šverák 2014; Nadirashvili 1991; Zlatoš 2015]. Therefore we will only focus on proving infinite-in-time growth of either $\nabla\rho$ (since ρ itself is preserved along the trajectory, one can at most obtain growth results for $\nabla\rho$) or L^p norms of ω itself not involving any derivatives (where such growth is not possible for the 2-dimensional Euler equation since $\|\omega\|_{L^p}$ is preserved in time).

Our first result is set up in the periodic domain $\Omega = \mathbb{T}^2$. We show that, for all smooth initial data (ρ_0, ω_0) in $\mathbb{T}^2$ under some symmetry assumptions, as long as ρ_0 takes values of different sign along the two line segments $\{0\}\times[0,\pi]$ and $\{\pi\}\times[0,\pi]$ (see the left figure of Figure 2 for an illustration), $\|\nabla\rho(t)\|_{L^\infty}$ must grow to infinity at least algebraically in time for all time during the existence of a smooth solution.

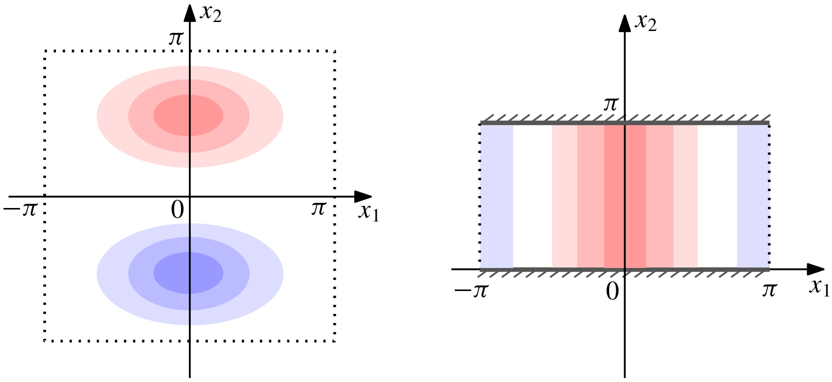

Figure 2. Illustration of the symmetry and sign assumptions on ρ_0 in the torus $\mathbb{T}^2$ (left) and the strip $\mathbb{T}\times[0,\pi]$ (right) for the inviscid Boussinesq equation. Here red denotes positive ρ_0 and blue denotes negative ρ_0.

Theorem 1.3. *Let $\rho_0 \in C^\infty(\mathbb{T}^2)$ be odd in x_2 and even in x_1, and $\omega_0 \in C^\infty(\mathbb{T}^2)$ be odd in both x_1 and x_2. Assume $\rho_0 \geq 0$ on $\{0\}\times[0,\pi]$ with $k_0 := \sup_{x_2\in[0,\pi]} \rho_0(0,x_2) > 0$, and $\rho_0 \leq 0$ on $\{\pi\}\times[0,\pi]$. Then there exists some constant $c(\rho_0,\omega_0) > 0$ such that the corresponding solution (ρ,ω) to (1-4) satisfies*

$$\sup_{\tau\in[0,t]} \|\nabla\rho(\tau)\|_{L^\infty(\mathbb{T}^2)} > c(\rho_0,\omega_0)t^{1/2} \quad \text{for all } t\in[0,T), \tag{1-5}$$

where T is the lifespan of the smooth solution (ρ,ω).

Next we consider the inviscid Boussinesq equation in the strip $\mathbb{T}\times[0,\pi]$. Here the presence of boundary allows us to obtain a faster growth rate in $\|\nabla\rho(t)\|_{L^\infty}$: we prove that the growth is at least like t^2 in the strip (as compared to $t^{1/2}$ in Theorem 1.3). We are also able to obtain a superlinear lower bound for $\|\omega(t)\|_{L^p}$ (for $p=\infty$ it grows like t^3) and a linear lower bound for $\|u(t)\|_{L^\infty}$. Although these algebraic lower bounds are far from finite-time blowup, they hold for a broad class of initial data: no assumption on ω_0 is needed other than being odd in x_1, and ρ_0 only needs to be even in x_1 and satisfy some sign conditions along two line segments (see the right figure of Figure 2 for an illustration). The proofs are soft but might provide an insight into the behavior of smooth solutions during their lifespan.

Theorem 1.4. *Let $\Omega = \mathbb{T}\times[0,\pi]$. Let $\rho_0 \in C^\infty(\Omega)$ be even in x_1 and $\omega_0 \in C^\infty(\Omega)$ be odd in x_1. Assume that there exists $k_0 > 0$ such that $\rho_0 \geq k_0 > 0$ on $\{0\}\times[0,\pi]$ and $\rho_0 \leq 0$ on $\{\pi\}\times[0,\pi]$. Then there exist some constants $T_0(\rho_0,\omega_0) \geq 0$ and $c(\rho_0,\omega_0) > 0$ such that the corresponding solution (ρ,ω) to (1-4) satisfies*

$$\|\omega(t)\|_{L^p(\Omega)} \geq ct^{3-2/p} \quad \text{for all } p\in[1,\infty],\ \ t\in[T_0,T), \tag{1-6}$$

$$\|u(t)\|_{L^\infty(\Omega)} \geq ct \qquad \text{for all } t\in[T_0,T), \tag{1-7}$$

and

$$\sup_{\tau\in[0,t]} \|\nabla\rho(\tau)\|_{L^\infty(\Omega)} > ct^2 \quad \text{for all } t\in[0,T), \tag{1-8}$$

where T is the lifespan of the smooth solution (ρ,ω). In particular, if $\int_{[0,\pi]\times[0,\pi]} \omega_0\,dx \geq 0$, then $T_0 = 0$ in all the estimates above.

Remark 1.5. In the estimates for $\|\omega(t)\|_{L^p(\Omega)}$ and $\|u(t)\|_{L^\infty(\Omega)}$ above, it is necessary to have a "waiting time" T_0 depending on the initial data. This is because, for any $t_1 > 0$, there exists some initial data satisfying the assumption of Theorem 1.4 with $\omega(\cdot, t_1) \equiv 0$. (To see this, one can start with $\omega(\cdot, t_1) \equiv 0$ and go backwards in time.) That being said, it can be easily seen from the proof that, if $\int_{[0,\pi]\times[0,\pi]} \omega_0 \, dx \geq 0$, no waiting time is needed.

Remark 1.6. If the symmetry assumptions on ρ_0 and ω_0 are dropped, we still have $\|\omega(t)\|_{L^1(\Omega)} \gtrsim t$ for $t \gg 1$. This infinite-in-time growth implies that, given any steady state ω_s for the 2-dimensional Euler equation on the strip, we have $(0, \omega_s)$ is a nonlinearly unstable steady state for the inviscid Boussinesq equation. See Remark 3.3 for more discussions.

Remark 1.7. Note that the growth result in Theorem 1.4 also holds for the rectangular domain $[-\pi, \pi] \times [0, \pi]$, since the symmetries imposed on the initial data automatically implies $u \cdot n = 0$ on all boundaries of $[-\pi, \pi] \times [0, \pi]$ for all time. However, the proof of Theorem 1.4 does not apply to domains with smooth boundary. That being said, for any bounded domain that is symmetric about both the x_1 and x_2 axis and has a smooth boundary, one can proceed similarly as in Theorem 1.3 (and Lemma 3.1) to obtain the same growth of $\|\nabla\rho\|_{L^\infty}$ as in Theorem 1.3. We leave the details of the argument to interested readers.

For both Theorems 1.3 and 1.4, the proof is based on an interplay between various monotone and conservative quantities. Under the symmetry assumptions, one can easily check that the sign assumptions $\rho \geq 0$ on $\{0\} \times [0, \pi]$ and $\rho \leq 0$ on $\{\pi\} \times [0, \pi]$ remain true for all times. This allows us to make the elementary but important observation that the vorticity integral $\int_{[0,\pi]\times[0,\pi]} \omega(x, t) \, dx$ is monotone increasing for all times. More precisely, for the strip, the growth is linear for all times during the existence of a smooth solution, whereas in $\mathbb{T}^2$ we relate the growth with $\|\nabla\rho(t)\|_{L^\infty}$. Another key ingredient is the relation between the vorticity integral and kinetic energy: since the kinetic energy has a uniform-in-time bound, we prove that if the vorticity integral is large, the L^p norm of vorticity must be much larger. For a strip, this allows us to upgrade the linear growth of $\|\omega(t)\|_{L^1}$ to superlinear growth for $\|\omega(t)\|_{L^p}$ for $p \in (1, \infty]$.

1.3. ***Infinite-in-time growth for the 3-dimensional axisymmetric Euler equation.*** The question whether the incompressible Euler equation in $\mathbb{R}^3$ can have a finite-time blowup from smooth initial data of finite energy is an outstanding open problem in nonlinear PDE and fluid dynamics. As we mentioned earlier, for the 3-dimensional axisymmetric Euler equation, when the equation is set up in a cylinder with boundary, Luo and Hou [2014] gave convincing numerical evidence that smooth initial data can lead to a finite-time singularity formation on the boundary. Recent numerical evidence by Hou and Huang [2022; 2023] and Hou [2022] suggests that the blowup can also happen in the interior of domain, but apparently not in self-similar fashion. The first rigorous blow-up result for finite-energy solutions was established in domains with corners by Elgindi and Jeong [2019]. For initial data in $C^{1,\alpha}$ in $\mathbb{R}^3$, Elgindi [2021] showed that such initial data can lead to a self-similar blowup. Very recently, using the connection between 3-dimensional axisymmetric Euler and Boussinesq equations, Chen and Hou [2022] set up a computer-assisted argument that smooth solutions to 3-dimensional axisymmetric Euler equation can form a stable nearly self-similar blowup. The singularity formation happens for initial data in a small neighborhood of a profile that is selected carefully with computer assistance.

In addition to the blow-up v.s. global-in-time regularity question, it is also interesting to investigate whether Sobolev norms of solutions to the 3-dimensional Euler equation can have infinite-in-time growth for broader classes of initial data. Choi and Jeong [2023] constructed smooth compactly supported initial data in $\mathbb{R}^3$ with $\|\nabla^2\omega(t)\|_{L^\infty}$ growing algebraically for all times, and $\|\omega(t)\|_{L^\infty}$ growing exponentially for finite (but arbitrarily long) time. It is also well known that the "two-and-a-half dimensional" solutions (i.e., where u only depends on x, y, not z) can lead to infinite-in-time linear growth of ω; see [Bardos and Titi 2007, Remark 3.1] for example. See the excellent survey [Drivas and Elgindi 2023] for more results on growth and singularity formation for 2-dimensional and 3-dimensional Euler equations.

It is well known that, away from the axis of symmetry, the 3-dimensional axisymmetric Euler equation is closely related to the inviscid 2-dimensional Boussinesq equations; see [Majda and Bertozzi 2002, Section 5.4.1]. To see this connection, recall that the 3-dimensional axisymmetric Euler equation can be reduced to the system

$$D_t(ru^\theta)=0, \quad D_t\left(\frac{\omega^\theta}{r}\right)=\frac{\partial_z(ru^\theta)^2}{r^4}, \tag{1-9}$$

where u^θ and ω^θ only depend on r, z, t, and $D_t:=\partial_t+u^r\partial_r+u^z\partial_z$ is the material derivative. Heuristically speaking, ru^θ plays the role of ρ in the Boussinesq equation, whereas ω^θ/r plays the role of ω in the Boussinesq equation. Here (u^r, u^z) can be recovered from ω^θ/r by the Biot–Savart law

$$(u^r, u^z)=\frac{1}{r}(-\partial_z\psi, \partial_r\psi), \quad \text{where } -\frac{1}{r}\partial_r\left(\frac{1}{r}\partial_r\psi\right)-\frac{1}{r^2}\partial_z^2\psi=\frac{\omega^\theta}{r}. \tag{1-10}$$

We note that the analog of Theorem 1.4 holds for the 3-dimensional axisymmetric Euler equation. We set the spatial domain to be a (not rotating) Taylor–Couette tank

$$\Omega=\{(r,\theta,z): r\in[\pi,2\pi],\ \theta\in\mathbb{T},\ z\in\mathbb{T}\}, \tag{1-11}$$

with no-penetration boundary condition at $r=\pi, 2\pi$ and periodic boundary conditions in z. Our assumptions and results are as follows.

Theorem 1.8. *Consider the* 3*-dimensional axisymmetric Euler equation* (1-9)–(1-10) *set on the domain* Ω *in* (1-11). *Let* $u_0^\theta\in C^\infty(\Omega)$ *be even in* z *and* $\omega_0^\theta\in C^\infty(\Omega)$ *be odd in* z. *Assume that there exists* $k_0>0$ *such that* $u_0^\theta\ge k_0>0$ *on* $z=\pi$ *and* $|u_0^\theta|\le\frac{1}{8}k_0$ *on* $z=0$. *Then there exist some constants* $T_0(u_0)\ge 0$ *and* $c(u_0)>0$ *such that the corresponding solution satisfies*

$$\|\omega^\theta(t)\|_{L^p(\Omega)}\ge ct^{3-2/p} \quad \text{for all } p\in[1,\infty],\quad t\in[T_0,T) \tag{1-12}$$

and

$$\|u(t)\|_{L^\infty(\Omega)}\ge ct \quad \text{for all } t\in[T_0,T), \tag{1-13}$$

where T *is the lifespan of the smooth solution. In particular, if* $\int_0^\pi\int_\pi^{2\pi}\omega_0^\theta\,dr\,dz\ge 0$, *then* $T_0=0$ *in both estimates above.*

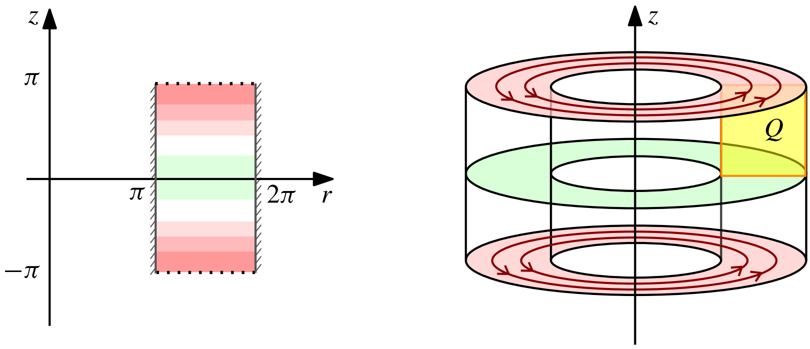

Figure 3. Illustration of the domain and assumptions on u_0^θ for the 3-dimensional axisymmetric Euler equation. The left figure illustrates u_0^θ on the rz plane, and the right figure shows the 3-dimensional setting. Here red denotes positive u_0^θ (and deeper color means larger magnitude), and green denotes u_0^θ with a smaller magnitude (whose sign can be positive or negative). With such initial data, we will show that the "secondary flow" within the yellow square Q grows to infinity as $t \to \infty$.

See Figure 3 for an illustration of the domain and initial data. Note that our setting is almost the same as the Hou–Luo scenario [Luo and Hou 2014], except that we replace the cylinder by an annular cylinder. While our growth estimates are far from a finite-time blowup, they hold for a broad class of initial data: in addition to some symmetry assumptions on u_0^θ and ω_0^θ, all we need is u_0^θ being uniformly positive on $z = \pi$ and having small magnitude on $z = 0$. The proof is a simple argument analogous to Theorem 1.4 for Boussinesq equations, where the key idea is the interplay between the monotonicity of a vorticity integral and the boundedness of kinetic energy.

After the completion of this manuscript, we became aware of work by Serre [1991; 1999], where he studied the 3-dimensional axisymmetric Euler equation in the same domain as in our setting and obtained linear growth of vorticity.

2. Small scale formation for viscous Boussinesq equation

In this section, we aim to prove Theorem 1.1. To begin with, we discuss some properties on the solution (ρ, u) when the initial data satisfies (A1)–(A3). Under the assumption (A1), it is well known that $\rho(\cdot, t)$ and $u(\cdot, t)$ remain in $C^\infty(\Omega)$. And if $\Omega = \mathbb{R}^2$, we have $\rho(\cdot, t) \in C_c^\infty(\mathbb{R}^2)$ and $u(\cdot, t) \in H^k(\mathbb{R}^2)$ for all $k \in \mathbb{N}$ and $t \geq 0$; see, e.g., [Chae 2006; Hou and Li 2005].

Note that the symmetry in (A2) holds true for all times thanks to the uniqueness of solutions. If $\Omega = \mathbb{T}^2$, the additional symmetry in x_1 leads to $u_1(\cdot, t) = 0$ on the x_2-axis for all times, thus $\rho(0, x_2, t) = 0$ for all $x_2 \in \mathbb{T}$ and $t \geq 0$.

The symmetry in x_2 in (A2) also gives $u_2(\cdot, t) = 0$ on the x_1-axis for all times, and combining it with (A3) gives $\rho(x, t) \geq 0$ for $x_2 \geq 0$ and all $t \geq 0$.

We also note that, due to the incompressibility of u, all L^p norms of ρ are conserved in time; that is,

$$\|\rho(t, \cdot)\|_{L^p(\Omega)} = \|\rho_0\|_{L^p(\Omega)} \quad \text{for all } t \geq 0, \quad p \in [1, \infty]. \tag{2-1}$$

2.1. ***Evolution of the potential and kinetic energy.*** Let us define the *potential energy* and *kinetic energy* of the solution as, respectively,

$$E_P(t) := \int_\Omega \rho(x,t)x_2\,dx \quad \text{and} \quad E_K(t) := \frac{1}{2}\int_\Omega |u(x,t)|^2\,dx. \tag{2-2}$$

As we will see, the evolution of these energies plays a crucial role in the proof of Theorem 1.1. The rate of change of E_P can be easily computed as

$$E_P'(t) = \int_\Omega \rho_t x_2\,dx = \int_\Omega -u\cdot(\nabla\rho)x_2\,dx = \int_\Omega \rho u_2\,dx, \tag{2-3}$$

where the last equality follows from the divergence theorem and $\nabla\cdot u = 0$, and note that the boundary integral in the divergence theorem is zero: in $\mathbb{R}^2$ it follows from $\rho(\cdot,t)$ having compact support, and in $\mathbb{T}^2$ it follows from the symmetries in (A2).

Similarly, one can compute the rate of change of the kinetic energy E_K as

$$E_K'(t) = -\int_\Omega \rho u_2\,dx - \nu\int_\Omega |\nabla u|^2\,dx.$$

Combining the two equations, the total energy $E_P(t)+E_K(t)$ is nonincreasing in time, and more precisely we have

$$E_P(t)+E_K(t)+\nu\int_0^t\int_\Omega |\nabla u(x,s)|^2\,dx\,ds = E_K(0)+E_P(0) \quad \text{for all } t\ge 0. \tag{2-4}$$

From our discussion above, $\rho(\cdot,t)$ remains odd in x_2 for all $t\ge 0$, and the property (A3) holds for all $t\ge 0$. Thus $E_P(t)$ is positive for all times. Combining this with (2-4) gives

$$0\le E_P(t)\le E_P(0)+E_K(0) \quad \text{and} \quad 0\le E_K(t)\le E_P(0)+E_K(0) \qquad \text{for all } t\ge 0. \tag{2-5}$$

In addition, using that $E_P(t)\ge 0$ and $E_K(t)\ge 0$ for all $t\ge 0$, we can send $t\to\infty$ in (2-4) to obtain

$$\nu\int_0^\infty \|\nabla u(t)\|_{L^2(\Omega)}^2\,dt \le E_P(0)+E_K(0). \tag{2-6}$$

In the next lemma we compute the second derivative of E_P, which will be used later.

Lemma 2.1. *Let (ρ,u) be a solution to* (1-1) *with initial data (ρ_0,u_0) satisfying* (A1)–(A3)*. Then the potential energy E_P defined in* (2-2) *satisfies*

$$E_P''(t) = A(t)+B(t)-\delta(t) \quad \textit{for all } t\ge 0, \tag{2-7}$$

where

$$A(t) := \sum_{i,j=1}^2 \int_\Omega ((-\Delta)^{-1}\partial_2\rho)\partial_i u_j \partial_j u_i\,dx, \quad B(t) := \nu\int_\Omega \rho\Delta u_2\,dx, \quad \textit{and} \quad \delta(t) := \|\partial_1\rho\|_{\dot H^{-1}(\Omega)}^2. \tag{2-8}$$

Proof. Differentiating (2-3) in time, we get

$$E_P''(t) = \int_\Omega -u\cdot\nabla(\rho u_2)+\rho(-\partial_2 p-\rho+\nu\Delta u_2)\,dx = \int_\Omega \rho(-\partial_2 p-\rho+\nu\Delta u_2)\,dx, \tag{2-9}$$

where the second equality follows from the incompressibility of u and the fact that the boundary integral is zero as we apply the divergence theorem: for $\Omega = \mathbb{R}^2$ it follows from $\rho(\cdot, t)$ having compact support, whereas for $\Omega = \mathbb{T}^2$ we are using $u \cdot n = 0$ on the boundary of $[-\pi, \pi]^2$ due to our symmetry assumptions in (A2). Comparing (2-9) with our goal (2-7), it suffices to show that

$$\int_\Omega \rho(-\partial_2 p - \rho)\,dx = A(t) - \delta(t). \tag{2-10}$$

To do so, we take divergence in the equation for u in (1-1). Using the incompressibility of u, we get $\nabla \cdot (u \cdot \nabla u) = -\Delta p - \partial_2 \rho$, and hence

$$p = (-\Delta)^{-1} \nabla \cdot (u \cdot \nabla u) + (-\Delta)^{-1} \partial_2 \rho,$$

where $(-\Delta)^{-1}$ is the inverse Laplacian in Ω (which is either $\mathbb{R}^2$ or $\mathbb{T}^2$) defined in the standard way using Fourier transform (for $\Omega = \mathbb{R}^2$) or Fourier series (for $\Omega = \mathbb{T}^2$). Therefore it follows that

$$\begin{aligned} -\partial_2 p - \rho &= -\partial_2 (-\Delta)^{-1} \nabla \cdot (u \cdot \nabla u) - (-\Delta)^{-1} \partial_{22} \rho - \rho \\ &= -\sum_{i,j=1}^{2} \partial_2 (-\Delta)^{-1} (\partial_i u_j \partial_j u_i) + (-\Delta)^{-1} \partial_{11} \rho. \end{aligned}$$

This immediately yields that

$$\begin{aligned} \int_\Omega \rho(-\partial_2 p - \rho)\,dx &= -\sum_{i,j=1}^{2} \int_\Omega \rho \partial_2 (-\Delta)^{-1} (\partial_i u_j \partial_j u_i)\,dx + \int_\Omega \rho (-\Delta)^{-1} \partial_{11} \rho\,dx \\ &= A(t) - \delta(t), \end{aligned}$$

where the second equality follows from integration by parts. This finishes the proof. □

The relation between $\delta(t)$ and $\|\rho(t)\|_{\dot{H}^s(\Omega)}$ has been investigated in [Kiselev and Yao 2023]. Below we state the results from that paper and give a slightly improved estimate for the $\Omega = \mathbb{R}^2$ case.[2] For the sake of completeness, we give a proof in the Appendix. In the statement of the lemma we replace $\rho(t)$ by μ to emphasize that the estimate does not depend on the equation that $\rho(t)$ satisfies.

Lemma 2.2. (a) *Assume $\Omega = \mathbb{R}^2$. Consider all $\mu \in C_c^\infty(\mathbb{R}^2)$ that are odd in x_2 and not identically zero. For all such μ, there exists $c_1(s, \|\mu\|_{L^1}, \|\mu\|_{L^2}) > 0$ such that*

$$\|\mu\|_{\dot{H}^s(\mathbb{R}^2)} \ge c_1 (\|\partial_1 \mu\|^2_{\dot{H}^{-1}(\mathbb{R}^2)})^{-s/4} \quad \textit{for all } s > 0. \tag{2-11}$$

(b) *Assume $\Omega = \mathbb{T}^2$. Consider all $\mu \in C^\infty(\mathbb{T}^2)$ that are not identically zero, odd in x_2, even in x_1, with $\mu = 0$ on the x_2-axis, and $\mu \ge 0$ in $\mathbb{T} \times [0, \pi]$. For all such μ, there exists $c_2\big(s, \int_{\mathbb{T}\times[0,\pi]} \mu^{1/3}\,dx\big) > 0$ such that*

$$\|\mu\|_{\dot{H}^s(\mathbb{T}^2)} \ge c_2 (\|\partial_1 \mu\|^2_{\dot{H}^{-1}(\mathbb{T}^2)})^{-s+1/2} \quad \textit{for all } s > \tfrac{1}{2}. \tag{2-12}$$

[2] In [Kiselev and Yao 2023], the estimate corresponding to (2-11) is [Kiselev and Yao 2023, (3.4)], where an extra condition $\|\partial_1 \mu\|^2_{\dot{H}^{-1}} < \frac{1}{4}\|\mu\|^2_{L^2}$ was imposed. In this lemma we give a slightly improved estimate where this assumption is dropped.

2.2. ***Infinite-in-time growth of Sobolev norms.*** Using Lemma 2.1, for any $t_2 > t_1 \geq 0$, integrating E_P'' from t_1 to t_2 we get

$$E_P'(t_2) - E_P'(t_1) = \int_{t_1}^{t_2} A(t)\,dt + \int_{t_1}^{t_2} B(t)\,dt - \int_{t_1}^{t_2} \delta(t)\,dt. \tag{2-13}$$

In the next lemma we estimate the two integrals $\int_{t_1}^{t_2} A(t)\,dt$ and $\int_{t_1}^{t_2} B(t)\,dt$ on the right-hand side.

Lemma 2.3. *Assume $\nu > 0$. Let (ρ, u) be a solution to (1-1) with initial data (ρ_0, u_0) satisfying (A1)–(A3). Then, for all $t_2 > t_1 \geq 0$, $A(t)$ defined in (2-8) satisfies*

$$\int_{t_1}^{t_2} |A(t)|\,dt \leq C(\rho_0) \int_{t_1}^{t_2} \|\nabla u(t)\|_{L^2(\Omega)}^2\,dt. \tag{2-14}$$

Furthermore, for all $s \geq 1$ and $t_2 > t_1 \geq 0$, $B(t)$ defined in (2-8) satisfies

$$\int_{t_1}^{t_2} |B(t)|\,dt \leq C(s, \rho_0)\nu \left(\int_{t_1}^{t_2} \|\nabla u(t)\|_{L^2(\Omega)}^2\,dt \right)^{1/2} \left(\int_{t_1}^{t_2} \|\rho(t)\|_{\dot{H}^s(\Omega)}^{2/s}\,dt \right)^{1/2}. \tag{2-15}$$

Proof. Let us show (2-14) first. Let $f := (-\Delta)^{-1}\partial_2 \rho$; we claim that

$$\|f(\cdot, t)\|_{L^\infty(\Omega)} \leq C(\rho_0) \quad \text{for all } t \geq 0. \tag{2-16}$$

Once this is proved, it follows that

$$\int_{t_1}^{t_2} |A(t)|\,dt \leq \int_{t_1}^{t_2} \|f\|_{L^\infty(\Omega)} \|\nabla u\|_{L^2(\Omega)}^2\,dt \leq C(\rho_0) \int_{t_1}^{t_2} \|\nabla u\|_{L^2(\Omega)}^2\,dt.$$

To estimate $\|f\|_{L^\infty(\Omega)}$, we recall the following Hardy–Littlewood–Sobolev inequality for $\Omega = \mathbb{R}^2$ or $\mathbb{T}^2$: (when $\Omega = \mathbb{T}^2$, the function g needs to satisfy an additional assumption $\int_\Omega g(x)\,dx = 0$)

$$\|(-\Delta)^{-\alpha/2} g\|_{L^q(\Omega)} \leq C(\alpha, p, q)\|g\|_{L^p(\Omega)} \quad \text{for } 0 < \alpha < 2,\ \ 1 < p < q < \infty,\ \text{and } \frac{1}{q} = \frac{1}{p} - \frac{\alpha}{2}.$$

We choose $\alpha = 1$, $q = 4$, $p = \frac{4}{3}$, and $g = (-\Delta)^{1/2} f(\cdot, t)$ (note that $g = (-\Delta)^{-1/2}\partial_2\rho$ indeed has mean zero when $\Omega = \mathbb{T}^2$). Then the above inequality becomes

$$\|f(\cdot, t)\|_{L^4(\Omega)} \leq C\|(-\Delta)^{1/2} f\|_{L^{4/3}(\Omega)} = C\|(-\Delta)^{-1/2}\partial_2\rho\|_{L^{4/3}(\Omega)} \leq C\|\rho\|_{L^{4/3}(\Omega)} \leq C(\rho_0),$$

and we also have

$$\|(-\Delta)^{1/2} f(\cdot, t)\|_{L^4(\Omega)} = \|(-\Delta)^{-1/2}\partial_2\rho\|_{L^4(\Omega)} \leq C\|\rho\|_{L^4(\Omega)} \leq C(\rho_0).$$

In the above two estimates, the second-to-last inequality in both equations is due to the Riesz transform being bounded in $L^p(\Omega)$ for $1 < p < \infty$, and the last inequality in both equations comes from (2-1). Combining these estimates together, we have

$$\|f(\cdot, t)\|_{W^{1,4}(\Omega)} \leq C(\rho_0) \quad \text{for all } t \geq 0.$$

Then the boundedness of f follows immediately from Morrey's inequality $W^{1,4}(\Omega) \subset C^{0,1/2}(\Omega)$ for both $\Omega = \mathbb{R}^2$ and $\mathbb{T}^2$. This leads to $\|f(\cdot, t)\|_{L^\infty(\Omega)} \leq C\|f(\cdot, t)\|_{W^{1,4}(\Omega)} \leq C(\rho_0)$ for all $t \geq 0$, which proves (2-16).

Now we turn to the estimate for $B(t)$. Applying the divergence theorem to the definition of $B(t)$ from (2-8), we see that

$$\int_{t_1}^{t_2} |B(t)|\,dt = \nu \int_{t_1}^{t_2} \left| \int_\Omega \nabla\rho\cdot\nabla u_2\,dx \right| dt \le \nu \left(\int_{t_1}^{t_2} \|u(t)\|^2_{\dot{H}^1(\Omega)}\,dt \right)^{1/2} \left(\int_{t_1}^{t_2} \|\rho(t)\|^2_{\dot{H}^1(\Omega)}\,dt \right)^{1/2}, \quad (2\text{-}17)$$

where we used the Cauchy–Schwarz inequality in the last step. Using the Gagliardo–Nirenberg interpolation inequality, we obtain

$$\begin{aligned}\int_{t_1}^{t_2} |B(t)|\,dt &\le \nu \left(\int_{t_1}^{t_2} \|\nabla u(t)\|^2_{L^2(\Omega)}\,dt \right)^{1/2} \left(\int_{t_1}^{t_2} C(s) \|\rho(t)\|^{2(1-1/s)}_{L^2} \|\rho(t)\|^{2/s}_{\dot{H}^s(\Omega)}\,dt \right)^{1/2} \\ &\le C(s,\rho_0)\nu \left(\int_{t_1}^{t_2} \|\nabla u(t)\|^2_{L^2(\Omega)}\,dt \right)^{1/2} \left(\int_{t_1}^{t_2} \|\rho(t)\|^{2/s}_{\dot{H}^s(\Omega)}\,dt \right)^{1/2},\end{aligned}$$

where the last inequality follows from (2-1). This finishes the proof of (2-15). □

Now we are ready to prove Theorem 1.1.

Proof of Theorem 1.1. The main idea of the proof is to estimate all terms in (2-13) for $t_1 = T$ and $t_2 = 2T$ for $T \gg 1$, and obtain a contradiction if $\sup_{t\in[T,2T]} \|\rho(t)\|_{\dot{H}^s}$ grows slower than a certain power of T.

First, to bound the left-hand side of (2-13), note that (2-3) and the Cauchy–Schwarz inequality yield

$$|E_P'(t)| \le \|\rho(t)\|_{L^2} \|u(t)\|_{L^2} \le \|\rho_0\|_{L^2} \sqrt{2E_K(t)} \le C(\rho_0, u_0) < \infty \quad \text{for all } t \ge 0, \quad (2\text{-}18)$$

where the second inequality follows from (2-1) and the definition of E_K in (2-2), and the third inequality follows from (2-5). Thus

$$|E_P'(2T) - E_P'(T)| \le C_0(\rho_0, u_0) < \infty \quad \text{for all } T > 0. \quad (2\text{-}19)$$

Plugging the estimates (2-19) and (2-14) into the identity (2-13), we have

$$\int_T^{2T} \delta(t)\,dt \le C_0(\rho_0,u_0) + C_1(\rho_0) \int_T^{2T} \|\nabla u(t)\|^2_{L^2(\Omega)}\,dt + \int_T^{2T} |B(t)|\,dt \quad \text{for all } T > 0. \quad (2\text{-}20)$$

Next we will bound the two integrals on the right-hand side from above, and $\int_T^{2T} \delta(t)\,dt$ from below. Let us define

$$\eta(T) := \int_T^{2T} \|\nabla u(t)\|^2_{L^2(\Omega)}\,dt \quad \text{and} \quad M_s(T) := \sup_{t\in[T,2T]} \|\rho(t)\|_{\dot{H}^s(\Omega)}.$$

Combining (2-4) and (2-5) yields

$$\int_0^\infty \|\nabla u(t)\|^2_{L^2(\Omega)}\,dt \le \nu^{-1} C(\rho_0, u_0) < \infty,$$

where we also used the assumption $\nu > 0$. This implies

$$\lim_{T\to\infty} \eta(T) = 0. \quad (2\text{-}21)$$

To bound $\int_T^{2T} |B(t)|\,dt$, we use (2-15) and the definitions of $\eta(T)$ and $M_s(T)$ to get

$$\begin{aligned}\int_T^{2T} |B(t)|\,dt &\le C(s,\rho_0)\nu\left(\int_T^{2T} \|\nabla u\|^2_{L^2(\Omega)}\,dt\right)^{1/2}\left(\int_T^{2T} \|\rho\|^{2/s}_{\dot H^s(\Omega)}\,dt\right)^{1/2}\\ &\le C_2(s,\rho_0,\nu)\eta(T)^{1/2}M_s(T)^{1/s}T^{1/2} \qquad \text{for all } s\ge 1,\quad T>0. \end{aligned}\tag{2-22}$$

Next we will bound the integral $\int_T^{2T} \delta(t)\,dt$ from below. If $\Omega=\mathbb{R}^2$, assumption (A2) allows us to apply Lemma 2.2 (a) to $\rho(\,\cdot\,,t)$ (and note that its L^1 and L^2 norms are preserved in time), so there exists $c_3(s,\rho_0)>0$ such that

$$\|\rho(t)\|_{\dot H^s(\mathbb{R}^2)} \ge c_3(s,\rho_0)\delta(t)^{-s/4} \quad \text{for all } s>0,\quad t>0. \tag{2-23}$$

And if $\Omega=\mathbb{T}^2$, using assumptions (A2) and (A3) $\big($note that these imply that $\int_{\mathbb{T}\times[0,\pi]}\rho(x,t)^{1/3}\,dx$ is preserved in time$\big)$, by Lemma 2.2 (b), there exists $c_4(s,\rho_0)>0$ such that

$$\|\rho(t)\|_{\dot H^s(\mathbb{T}^2)} \ge c_4(s,\rho_0)\delta(t)^{-(s-1/2)} \quad \text{for all } s>\tfrac12,\quad t>0. \tag{2-24}$$

Let us rewrite (2-23) and (2-24) above in a unified manner for the two cases $\Omega=\mathbb{R}^2$ and $\mathbb{T}^2$, so we do not need to repeat similar proofs twice. For Ω either being $\mathbb{R}^2$ or $\mathbb{T}^2$, let us define

$$\alpha_\Omega := \begin{cases}\frac14 s & \Omega=\mathbb{R}^2,\\ s-\frac12 & \Omega=\mathbb{T}^2,\end{cases}\qquad \underline{s}_\Omega := \begin{cases}0 & \Omega=\mathbb{R}^2,\\ \frac12 & \Omega=\mathbb{T}^2,\end{cases}\qquad c_\Omega(s,\rho_0) := \begin{cases}c_3(s,\rho_0) & \Omega=\mathbb{R}^2,\\ c_4(s,\rho_0) & \Omega=\mathbb{T}^2.\end{cases} \tag{2-25}$$

With this notation, (2-23) and (2-24) become

$$\|\rho(t)\|_{\dot H^s(\Omega)} \ge c_\Omega(s,\rho_0)\delta(t)^{-\alpha_\Omega} \quad \text{for all } s>\underline{s}_\Omega,\quad t>0. \tag{2-26}$$

Combining (2-26) with the definition of M_s gives

$$\int_T^{2T} \delta(t)\,dt \ge \int_T^{2T} c_\Omega^{1/\alpha_\Omega}\|\rho(t)\|^{-1/\alpha_\Omega}_{\dot H^s(\Omega)}\,dt \ge c_\Omega^{1/\alpha_\Omega} M_s(T)^{-1/\alpha_\Omega}T \quad \text{for all } s>\underline{s}_\Omega,\quad T>0. \tag{2-27}$$

Applying the bounds (2-22) and (2-27) and the definition of $\eta(T)$ to the inequality (2-20) (and noting that $\underline{s}_\Omega<1$), we have

$$c_5 M_s(T)^{-1/\alpha_\Omega}T \le C_0 + C_1\eta(T) + C_2\eta(T)^{1/2}M_s(T)^{1/s}T^{1/2} \quad \text{for all } s\ge 1,\quad T>0,$$

where $c_5 := c_\Omega(s,\rho_0)^{1/\alpha_\Omega}$, $C_0 := C_0(\rho_0,u_0)$, $C_1 := C_1(\rho_0)$, and $C_2 := C_2(s,\rho_0,\nu)$ — note that they are all strictly positive and do not depend on T. Rearranging the terms, the inequality is equivalent to

$$(c_5 - C_2\eta(T)^{1/2}T^{-1/2}M_s(T)^{1/s+1/\alpha_\Omega})M_s(T)^{-1/\alpha_\Omega}T \le C_0 + C_1\eta(T) \quad \text{for all } s\ge 1,\quad T>0. \tag{2-28}$$

We claim that this implies

$$\limsup_{T\to\infty} T^{-1/2}M_s(T)^{1/s+1/\alpha_\Omega} = +\infty \quad \text{for all } s\ge 1. \tag{2-29}$$

Towards a contradiction, assume

$$A := \limsup_{T\to\infty} T^{-1/2}M_s(T)^{1/s+1/\alpha_\Omega} < \infty \quad \text{for some } s\ge 1.$$

Combining this assumption with (2-21) gives

$$\limsup_{T\to\infty} \eta(T)^{1/2} T^{-1/2} M_s(T)^{1/s+1/\alpha_\Omega} = \Bigl(\limsup_{T\to\infty} T^{-1/2} M_s(T)^{1/s+1/\alpha_\Omega}\Bigr)\Bigl(\lim_{T\to\infty} \eta(T)^{1/2}\Bigr) = 0,$$

so the parenthesis in (2-28) converges to c_5 as $T \to \infty$. For the remaining term on the left-hand of (2-28), we have

$$\begin{aligned}\liminf_{T\to\infty} M_s(T)^{-1/\alpha_\Omega} T &= \liminf_{T\to\infty} (T^{-1/2} M_s(T)^{1/s+1/\alpha_\Omega})^{-s/(s+\alpha_\Omega)} T^{(s+2\alpha_\Omega)/(2(s+\alpha_\Omega))} \\ &= \liminf_{T\to\infty} A^{-s/(s+\alpha_\Omega)} T^{(s+2\alpha_\Omega)/(2(s+\alpha_\Omega))} = +\infty. \end{aligned} \tag{2-30}$$

The above discussion yields that the liminf of the left-hand side of (2-28) is $+\infty$. This contradicts (2-21), which says the right-hand side of (2-28) goes to $C_0 < \infty$ as $T \to \infty$. This finishes the proof of the claim (2-29).

Finally, using the definition of M_s, we have that (2-29) is equivalent to

$$\limsup_{t\to\infty} t^{-1/2} \|\rho(t)\|_{\dot H^s}^{1/s+1/\alpha_\Omega} = +\infty.$$

Recalling the definition of α_Ω from (2-25), we see that the desired estimates (1-2) and (1-3) follow immediately. □

Although it is unclear whether the algebraic rates are sharp, in the next proposition we show that, under the assumptions (A1)–(A3), $\|\rho(t)\|_{H^1(\Omega)}$ can at most have subexponential growth.

Proposition 2.4. *Let $\Omega = \mathbb{R}^2$ or $\mathbb{T}^2$. For any initial data (ρ_0, u_0) satisfying (A1)–(A3), $\|\rho(t)\|_{H^1(\Omega)}$ satisfies the subexponential bound*

$$\|\rho(t)\|_{H^1(\Omega)} \lesssim \exp(Ct^\alpha) \quad \textit{for all } t > 0,$$

for some constant $\alpha \in (0, 1)$.

Proof. The proposition can be proved by making a slight modification to [Kukavica and Wang 2020, Theorem 3.1]. For the sake of completeness, we will provide a sketch of the proof. For both $\Omega = \mathbb{T}^2$ and $\mathbb{R}^2$, standard energy estimates give that $\|\nabla\rho(t)\|_{L^2(\Omega)}$ satisfies the estimate

$$\frac{d}{dt}\|\nabla\rho(t)\|_{L^2(\Omega)} \le \|\nabla u(t)\|_{L^\infty} \|\nabla\rho(t)\|_{L^2(\Omega)},$$

which leads to

$$\|\nabla\rho(t)\|_{L^2(\Omega)} \lesssim \exp\biggl(\int_0^t \|\nabla u(s)\|_{L^\infty(\Omega)}\, ds\biggr) \|\nabla\rho_0\|_{L^2(\Omega)}. \tag{2-31}$$

Recall that (2-6) gives

$$\int_0^\infty \|\nabla u(t)\|_{L^2}^2\, dt \le C(\nu, \rho_0, u_0). \tag{2-32}$$

Here the time integrability of $\|\nabla u(t)\|_{L^2}^2$ follows from the symmetry assumptions in our setting, and it allows us to obtain a refined upper bound compared to [Kukavica and Wang 2020, Theorem 3.1]. Namely, combining (2-32) with the Gagliardo–Nirenberg inequality

$$\|\nabla u\|_{L^\infty(\Omega)} \le \|\nabla u\|_{L^2(\Omega)}^{(p-2)/(2p-2)} \|\nabla^2 u\|_{L^p(\Omega)}^{p/(2p-2)} \quad \text{for } p > 2 \tag{2-33}$$

and Hölder's inequality, the exponent in (2-31) can be bounded above by

$$\int_0^t \|\nabla u(s)\|_{L^\infty(\Omega)}\,ds \le C(p,\nu,\rho_0,u_0)\left(\int_0^t \|\nabla^2 u\|_{L^p(\Omega)}^{(2p)/(3p-2)}\,ds\right)^{(3p-2)/(4p-4)}. \tag{2-34}$$

When $\Omega=\mathbb{T}^2$, by [Kukavica and Wang 2020, Theorem 2.1], $\|u(t)\|_{W^{2,p}} < C(p,\nu,\rho_0,u_0)$ for all $p<\infty$. So one can choose $p \gg 1$ to obtain the subexponential upper bound

$$\|\nabla\rho(t)\|_{L^2} \le C(\rho_0)\exp(C(\epsilon,\nu,\rho_0,u_0)t^{3/4+\epsilon}) \quad \text{for any } \epsilon>0,\ \ t>0. \tag{2-35}$$

Next we move on to the $\Omega=\mathbb{R}^2$ case. In this case, it suffices to prove $\|\nabla^2 u(t)\|_{L^p} < C(p,\nu,\rho_0,u_0)$ for all $p<\infty$ under our symmetry setting. Once this is shown, an identical argument as (2-31)–(2-35) again leads to the subexponential growth, since all these estimates also hold for $\mathbb{R}^2$.

To begin with, we show that $\|\omega(t)\|_{L^2}$ is uniformly bounded in time under our symmetry assumptions. Noting from (1-1) that ω satisfies $\omega_t + u\cdot\nabla\omega = \nu\Delta\omega - \partial_1\rho$, we can obtain a standard energy inequality

$$\begin{aligned}\frac{d}{dt}\|\omega(t)\|_{L^2(\mathbb{R}^2)}^2 + \nu\|\nabla\omega(t)\|_{L^2(\mathbb{R}^2)}^2 &= 2\int_{\mathbb{R}^2}\rho(t,x)\partial_1\omega(t,x)\,dx \\ &\le \tfrac{1}{2}\nu\|\nabla\omega(t)\|_{L^2(\mathbb{R}^2)}^2 + C(\nu)\|\rho(t)\|_{L^2(\mathbb{R}^2)}.\end{aligned}$$

Since $\|\rho(t)\|_{L^2}$ is conserved, the above estimate leads to $(d/dt)\|\omega(t)\|_{L^2(\mathbb{R}^2)}^2 \le C(\nu,\rho_0)$. Combining this with (2-32) (and recall $\|\omega\|_{L^2} = \|\nabla u\|_{L^2}$), we have

$$\|\omega(t)\|_{L^2(\mathbb{R}^2)} < C(\nu,\rho_0,u_0) \quad \text{for all } t\ge 0. \tag{2-36}$$

Following the notation from [Kukavica and Wang 2020], let us define $\zeta = \omega - \partial_1(I-\Delta)^{-1}\rho$ to be the modified vorticity. Since one has $\|\partial_1(I-\Delta)^{-1}\rho\|_{W^{1,p}} \le C(p,\rho_0)$ for all $1<p<\infty$, it implies

$$\|\zeta-\omega\|_{L^p} \le C(p,\rho_0) \quad \text{and} \quad \|\nabla\zeta-\nabla\omega\|_{L^p} \le C(p,\rho_0). \tag{2-37}$$

Combining (2-36) and (2-37) gives a uniform-in-time bound $\|\zeta(t)\|_{L^2} < C(\nu,\rho_0,u_0)$. Now, let us define $\psi_p(t) := \int_{\mathbb{R}^2}|\nabla\zeta(t)|^p$ for $p\ge 2$. Using (1-1), one can express the equation for ζ as [Kukavica and Wang 2020, (2.21)]

$$\zeta_t + u\cdot\nabla\zeta = \nu\Delta\zeta + F, \qquad F := [\partial_1(I-\Delta)^{-1}, u\cdot\nabla]\rho - ((I-\Delta)^{-1}\Delta - I)\partial_1\rho.$$

A straightforward calculation yields that $\psi_2'(t) = 2\int_{\mathbb{R}^2}\nabla\zeta\cdot\nabla F\,dx - 2\nu\int_{\mathbb{R}^2}|\nabla^2\zeta|^2\,dx$. Using the interpolation inequality

$$\|\nabla^2\zeta\|_{L^2} \ge \frac{\|\nabla\zeta\|_{L^2}^2}{C\|\zeta\|_{L^2}},$$

we obtain

$$\psi_2'(t) + \frac{\psi_2^2}{C\|\zeta\|_{L^2}^2} \le 2\int_{\mathbb{R}^2}\nabla\zeta\cdot\nabla F\,dx.$$

To obtain an estimate of the right-hand side, a more careful analysis is required, and the same argument as in [Kukavica and Wang 2020, (3.2)] gives that

$$\psi_2'(t) + \frac{\psi_2^2}{C\|\zeta\|_{L^2}^2} \le C\psi_2 + C.$$

Thus the above uniform-in-time bound for $\|\zeta\|_{L^2}^2$ gives a uniform-in-time bound for $\psi_2(t)$. For any $2 \le p < \infty$, [Kukavica and Wang 2020, (3.3)] gives

$$\psi_{2p}'(t) + \frac{\psi_{2p}^2}{C\psi_p^2} \le Cp^2\psi_{2p} + Cp^5\psi_{2p}^{(p-1)/p}.$$

One can use induction (for $p = 2, 4, 8, \dots$) to obtain a uniform-in-time bound $\psi_p(t) \le C(p, \nu, \rho_0, u_0)$, and combining this bound with (2-37) gives

$$\|\nabla^2 u(t)\|_{L^p} \le C(p)\|\nabla\omega(t)\|_{L^p} \le C(p)(\|\nabla\zeta(t)\|_{L^p} + C(p, \rho_0)) \le C(p, \nu, \rho_0, u_0).$$

Finally, choosing an arbitrarily large $p \gg 1$ and plugging the above uniform-in-time estimate into (2-34), we again have the subexponential upper bound (2-35) for $\Omega = \mathbb{R}^2$. □

3. Infinite-in-time growth for inviscid Boussinesq and 3-dimensional Euler

3.1. ***Vorticity lemma for flows with fixed kinetic energy.*** Before proving the main theorems, let us start with a simple observation: it says that for any vector field u in a square $Q = [0, \pi]^2$ with a fixed kinetic energy, if its vorticity integral $A := \int_Q \omega\, dx$ is big, then, for $1 < p \le \infty$, $\|\omega\|_{L^p}$ must be even bigger, at least of order $A^{3-2/p}$.

Lemma 3.1. *Let $Q := [0, \pi]^2$. For any vector field $u \in C^\infty(Q)$, let $\omega := \partial_1 u_2 - \partial_2 u_1$. Let us define*

$$E_0 := \int_Q |u|^2\, dx \quad \text{and} \quad A := \int_Q \omega(x)\, dx.$$

Then we have the following lower bound for $\|\omega\|_{L^p(Q)}$:

$$\|\omega\|_{L^p(Q)} \ge c_0 \max\{E_0^{-1+1/p}|A|^{3-2/p}, |A|\} \quad \text{for all } p \in [1, \infty], \tag{3-1}$$

where $c_0 = (128\pi^2)^{-1} > 0$ is a universal constant.

Proof. Without loss of generality, assume $A > 0$. (If $A < 0$, we can prove the estimate for $-u$, whose vorticity integral would be positive.) By Green's theorem, we have

$$\int_{\partial Q} |u(x)|\, ds \ge \int_{\partial Q} u(x) \cdot dl = \int_Q \omega(x)\, dx = A,$$

where the integral in ds denotes the (scalar) line integral with respect to arclength, and the integral in dl denotes the (vector) line integral counterclockwise along ∂Q.

For any $r \in \left[0, \frac{\pi}{2}\right)$, let us define

$$Q_r := [r, \pi - r] \times [r, \pi - r].$$

Note that $Q_0 = Q$ and Q_r shrinks to a point as $r \nearrow \frac{\pi}{2}$. Let us define

$$r_0 := \inf\left\{ r \in \left[0, \tfrac{\pi}{2}\right) : \int_{\partial Q_r} |u(x)|\, ds = \tfrac{1}{2} A \right\}.$$

Since

$$\int_{\partial Q_0} |u(x)|\, ds > A \qquad \text{and} \qquad \int_{\partial Q_r} |u(x)|\, ds \to 0 \quad \text{as } r \nearrow \tfrac{\pi}{2},$$

the above definition leads to a well-defined $r_0 \in \left(0, \frac{\pi}{2}\right)$, and in addition we have

$$\int_{\partial Q_r} |u(x)|\, ds > \tfrac{1}{2} A \quad \text{for all } r \in [0, r_0).$$

Next we claim that

$$r_0 < 16\pi E_0 A^{-2}. \tag{3-2}$$

To show this, note that, for all $0 < r < r_0$, we can apply the Cauchy–Schwarz inequality on ∂Q_r (and use $|\partial Q_r| < 4\pi$) to obtain

$$\int_{\partial Q_r} |u|^2\, ds \geq \frac{1}{4\pi} \left(\int_{\partial Q_r} |u|\, ds \right)^2 > \frac{A^2}{16\pi}.$$

Integrating the above inequality for $r \in (0, r_0)$ over the direction transversal to ∂Q_r $\big($and noting that $\bigcup_{r \in (0, r_0)} \partial Q_r = Q \setminus Q_{r_0}\big)$, we obtain

$$E_0 \geq \int_{Q \setminus Q_{r_0}} |u|^2\, dx = \int_0^{r_0} \int_{\partial Q_r} |u|^2\, ds\, dr > \frac{A^2 r_0}{16\pi},$$

which yields the claim (3-2). Note that (3-2) implies

$$|Q \setminus Q_{r_0}| = \int_0^{r_0} |\partial Q_r|\, dr \leq \min\{4\pi r_0, \pi^2\} \leq \min\{64\pi^2 E_0 A^{-2}, \pi^2\}. \tag{3-3}$$

By Green's theorem and the definition of r_0,

$$\int_{Q \setminus Q_{r_0}} \omega\, dx = \int_{\partial Q} u \cdot dl - \int_{\partial Q_{r_0}} u \cdot dl \geq A - \tfrac{1}{2} A = \tfrac{1}{2} A. \tag{3-4}$$

Finally, we apply Hölder's inequality to bound $\|\omega\|_{L^p(Q)}$ from below for $p \in [1, \infty]$:

$$\|\omega\|_{L^p(Q)} \geq \|\omega\|_{L^p(Q \setminus Q_{r_0})} \geq \left(\int_{Q \setminus Q_{r_0}} \omega\, dx \right) |Q \setminus Q_{r_0}|^{-1+1/p} \quad \text{for all } p \in [1, \infty].$$

Applying the estimates (3-4) and (3-3) in the above inequality finishes the proof of (3-1) with a universal constant $c_0 = (128\pi^2)^{-1}$. □

3.2. ***Infinite-in-time growth for inviscid Boussinesq equations.*** Now we are ready to prove the infinite-in-time growth results. Let us start with Theorem 1.3 for $\Omega = \mathbb{T}^2$.

Proof of Theorem 1.3. Using the Biot–Savart law $u = \nabla^\perp(-\Delta)^{-1}\omega$, one can easily check that, in $\mathbb{T}^2 = (-\pi, \pi]^2$, the even-odd symmetry of ρ and odd-odd symmetry of ω is preserved for all times. This implies the odd-even symmetry of u_1 and even-odd symmetry of u_2 hold for all times. In particular, defining

$$Q := [0, \pi] \times [0, \pi],$$

we have $u \cdot n = 0$ on ∂Q for all times.

For any $x \in \mathbb{T}^2$ and $t \geq 0$, let $\Phi_t(x)$ be the flow map defined by

$$\partial_t \Phi_t(x) = u(\Phi_t(x), t), \qquad \Phi_0(x) = x.$$

Using $u \cdot n = 0$ on ∂Q for all times (and $u = 0$ at the four corners of ∂Q), for any $x \in \partial Q$, $\Phi_t(x)$ remains on the same side of ∂Q for all times during the existence of a smooth solution. Combining this with the fact that ρ is preserved along the flow map, the assumptions on ρ_0 implies

$$\rho(0, x_2, t) \geq 0 \quad \text{and} \quad \rho(\pi, x_2, t) \leq 0 \qquad \text{for all } x_2 \in [0, \pi], \quad t \geq 0. \tag{3-5}$$

Note that the odd-in-x_2 symmetry of ρ_0 yields $\rho_0(0, 0) = \rho_0(0, \pi) = 0$, so the supremum in $k_0 := \sup_{x_2 \in [0,\pi]} \rho_0(0, x_2) > 0$ is achieved at some $\rho(0, a)$ for $a \in (0, \pi)$. In addition, by continuity of ρ_0, there exists some $b \in (0, a)$ such that $\rho_0(0, b) = \frac{1}{2}k_0$ and $\rho_0 \geq \frac{1}{2}k_0$ on $\{0\} \times [b, a]$.

Since $u \cdot n = 0$ on ∂Q for all times, $\Phi_t(0, a)$ and $\Phi_t(0, b)$ remain on the line segment $\{0\} \times (0, \pi)$ for all times. Define

$$h(t) := |\Phi_t(0, b) - \Phi_t(0, a)|, \tag{3-6}$$

which is strictly positive as long as u remains smooth. Note $\rho(\Phi_t(0, a), t) = k_0$ and $\rho(\Phi_t(0, b), t) = \frac{1}{2}k_0$ for all times. This implies

$$\|\nabla\rho(t)\|_{L^\infty(Q)} \geq \frac{|\rho(\Phi_t(0, b), t) - \rho(\Phi_t(0, a), t)|}{|\Phi_t(0, b) - \Phi_t(0, a)|} \geq \tfrac{1}{2}k_0 h(t)^{-1} \tag{3-7}$$

for all times during the existence of a smooth solution.

Next let us define

$$A(t) := \int_Q \omega(x, t)\, dx;$$

we make a simple but useful observation about the monotonicity of $A(t)$. Using the symmetries and the facts $\nabla \cdot u = 0$ in Q and $u \cdot n = 0$ on ∂Q, we find

$$\begin{aligned} A'(t) &= -\int_Q u(x, t) \cdot \nabla\omega(x, t)\, dx - \int_Q \partial_{x_1}\rho(x, t)\, dx \\ &= \int_0^\pi \rho(0, x_2, t)\, dx_2 - \int_0^\pi \rho(\pi, x_2, t)\, dx_2 \geq \tfrac{1}{2}k_0 h(t), \end{aligned} \tag{3-8}$$

where the inequality follows from (3-5), the definition of $h(t)$, and the fact that $\rho(\cdot\,,t) \geq \frac{1}{2}k_0$ on the line segment connecting $\Phi_t(0,a)$ and $\Phi_t(0,b)$. We now integrate (3-8) in $[0,t]$ and apply (3-7). This yields

$$A(t) \geq \tfrac{1}{4}k_0^2 \int_0^t \|\nabla\rho(\tau)\|_{L^\infty(Q)}^{-1}\,d\tau + A(0). \tag{3-9}$$

In order to apply Lemma 3.1, we need to bound $\|u(t)\|_{L^2(Q)}^2$ from above. From the same calculation in Section 2.1, the sum of the kinetic and potential energies is conserved in $\mathbb{T}^2$, and hence it is also conserved in Q due to the symmetries

$$\frac{1}{2}\int_Q |u(x,t)|^2\,dx + \int_Q x_2\rho(x,t)\,dx = \frac{1}{2}\int_Q |u_0(x)|^2\,dx + \int_Q x_2\rho_0(x)\,dx.$$

Since ρ is advected by the flow, $\|\rho(t)\|_{L^1(Q)}$ is conserved in time, so $\left|\int_Q x_2\rho(x,t)\,dx\right| \leq \pi\|\rho_0\|_{L^1(Q)}$ for all times. This implies

$$\int_Q |u(x,t)|^2\,dx \leq \int_Q |u_0(x)|^2\,dx + 4\pi\|\rho_0\|_{L^1(Q)} =: E_0(\rho_0,u_0)$$

for all times. Now we can apply Lemma 3.1 with $p=+\infty$ to conclude

$$\|\omega(t)\|_{L^\infty} \geq c_0E_0^{-1}A(t)^3 \geq c_0E_0^{-1}\left(\tfrac{1}{4}k_0^2\int_0^t \|\nabla\rho(\tau)\|_{L^\infty}^{-1}\,d\tau + A(0)\right)^3, \tag{3-10}$$

where we used (3-9) in the last step. Note that $A(0)$ may be positive or negative.

On the other hand, the Lagrangian form of the evolution equation for vorticity

$$\frac{d}{dt}\omega(\Phi_t(x),t) = -\partial_{x_1}\rho(\Phi_t(x),t)$$

implies that

$$\|\omega(t)\|_{L^\infty} \leq \int_0^t \|\nabla\rho(\tau)\|_{L^\infty}\,d\tau + \|\omega_0\|_{L^\infty}. \tag{3-11}$$

Combining (3-10) and (3-11), we arrive at

$$\int_0^t \|\nabla\rho(\tau)\|_{L^\infty}\,d\tau + \|\omega_0\|_{L^\infty} \geq c_0E_0^{-1}\left(\tfrac{1}{4}k_0^2\int_0^t \|\nabla\rho(\tau)\|_{L^\infty}^{-1}\,d\tau + A_0\right)^3. \tag{3-12}$$

Let us define

$$F(t) := \int_0^t \|\nabla\rho(\tau)\|_{L^\infty}\,d\tau.$$

Since the Cauchy–Schwarz inequality yields

$$\int_0^t \|\nabla\rho(\tau)\|_{L^\infty}^{-1}\,d\tau \geq t^2\left(\int_0^t \|\nabla\rho(\tau)\|_{L^\infty}\,d\tau\right)^{-1} \geq t^2F(t)^{-1} \quad \text{for all } t>0,$$

plugging it into (3-12) gives an inequality relating $F(t)$ to itself:

$$F(t) \geq c_0E_0^{-1}\left(\tfrac{1}{4}k_0^2t^2F(t)^{-1} + A_0\right)^3 - \|\omega_0\|_{L^\infty}. \tag{3-13}$$

Our goal is to show that there exists some $c_1(\rho_0,\omega_0) > 0$ such that

$$F(t) \geq c_1(\rho_0,\omega_0)t^{3/2} \quad \text{for all } t \geq 1. \tag{3-14}$$

Towards a contradiction, suppose (3-14) does not hold at some $t_1 \ge 1$, so $t_1^2 F(t_1)^{-1} \ge c_1^{-1} t_1^{1/2}$. Since $t_1 \ge 1$, one can choose c_1 sufficiently small (only depending on initial data) such that the right-hand side of (3-13) is bounded below by $4^{-4} c_0 E_0^{-1} k_0^6 c_1^{-3} t_1^{3/2}$. On the other hand, the left-hand side is bounded above by $c_1 t_1^{3/2}$. Thus we obtain a contradiction if we further require $c_1 < 4^{-1}(c_0 E_0^{-1} k_0^6)^{1/4}$.

Finally, note that (3-14) directly implies $\sup_{\tau\in[0,t]} \|\nabla\rho(\tau)\|_{L^\infty} \ge c_1(\rho_0, \omega_0) t^{1/2}$ for all $t \ge 1$. For $t \in (0,1)$, recall that the definition of k_0 and the fact $\rho(0,0,t) = 0$ yield

$$\|\nabla\rho(t)\|_{L^\infty} \ge \tfrac{1}{\pi} k_0 \ge \left(\tfrac{1}{\pi} k_0\right) t^{1/2} \quad \text{for } t \in (0,1).$$

Combining these two estimates finishes the proof. □

Remark 3.2. Theorem 1.3 does not give us an infinite-in-time growth result for $\omega(\cdot, t)$. All we have is the following conditional growth estimate coming from (3-10): if $\limsup_{t\to\infty} t^{-1}\|\nabla\rho(t)\|_{L^\infty} < \infty$, this must imply $\lim_{t\to\infty} \|\omega(t)\|_{L^\infty} = \infty$.

Proof of Theorem 1.4. The proof is similar to the previous one, and in fact it is easier due to the uniform positivity of ρ_0 on $\{0\} \times [0,\pi]$. Using the Biot–Savart law, one can check that the even-in-x_1 symmetry of ρ and odd-in-x_1 symmetry of ω is preserved for all times. Defining $Q := [0,\pi] \times [0,\pi]$, the symmetries and the boundary condition yield that $u \cdot n = 0$ on ∂Q for all times. In particular, this implies

$$\rho(0, x_2, t) \ge k_0 > 0 \quad \text{and} \quad \rho(\pi, x_2, t) \le 0 \qquad \text{for all } x_2 \in [0,\pi],\ \ t \ge 0, \tag{3-15}$$

during the existence of a smooth solution.

Again, let us define $A(t) := \int_Q \omega(x,t)\,dx$. A calculation similar to the previous proof shows that in this case

$$A'(t) \ge \int_0^\pi \rho(0, x_2, t)\,dx_2 - \int_0^\pi \rho(\pi, x_2, t)\,dx_2 \ge k_0\pi,$$

where the last inequality follows from (3-15). This gives us a lower bound

$$A(t) \ge k_0\pi t + A(0) \quad \text{for all } t \ge 0. \tag{3-16}$$

An identical argument as in the proof of Theorem 1.3 gives

$$\int_\Omega |u(x,t)|^2\,dx \le E_0(\rho_0, u_0)$$

uniformly in time; thus we can apply Lemma 3.1 to obtain

$$\|\omega\|_{L^p(Q)} \ge c_0 E_0^{-1+1/p} |A(t)|^{3-2/p} \quad \text{for all } p \in [1,\infty]. \tag{3-17}$$

Also, note that Green's theorem yields

$$A(t) = \int_{\partial Q} u \cdot dl \le 4\pi \|u(t)\|_{L^\infty}. \tag{3-18}$$

Regarding the growth of $\nabla\rho$, note that (3-11) still holds in a strip, so

$$\sup_{\tau\in[0,t]} \|\nabla\rho(\tau)\|_{L^\infty} \ge t^{-1}(\|\omega(t)\|_{L^\infty} - \|\omega_0\|_{L^\infty}) \quad \text{for all } t > 0. \tag{3-19}$$

Below we discuss two cases.

Case 1: $A(0) \geq 0$. In this case (3-16) gives

$$A(t) \geq k_0 \pi t \quad \text{for all } t > 0.$$

We then apply (3-17) and (3-18) to obtain lower bounds for $\|\omega(t)\|_{L^p(Q)}$ and $\|u(t)\|_{L^\infty}$:

$$\|\omega(t)\|_{L^p(Q)} \geq c_1(\rho_0, \omega_0) t^{3-2/p} \quad \text{for all } p \in [1, +\infty], \quad t \geq 0, \tag{3-20}$$

$$\|u(t)\|_{L^\infty(Q)} \geq \tfrac{1}{4} k_0 t \quad \text{for all } t \geq 0. \tag{3-21}$$

Regarding the growth of $\nabla \rho$, we apply (3-20) with $p = +\infty$ and combine it with (3-19) to obtain

$$\sup_{\tau \in [0,t]} \|\nabla \rho(\tau)\|_{L^\infty} \geq t^{-1}(c_1(\rho_0, \omega_0) t^3 - \|\omega_0\|_{L^\infty}),$$

which implies

$$\sup_{\tau \in [0,t]} \|\nabla \rho(\tau)\|_{L^\infty} \geq c_1(\rho_0, \omega_0) t^2 \quad \text{for all } t \geq \left(\frac{\|\omega_0\|_{L^\infty}}{c_1(\rho_0, \omega_0)} \right)^{1/3}.$$

Combining this large time estimate with the trivial lower bound $\|\nabla \rho(t)\|_{L^\infty} \geq \frac{1}{\pi} k_0$ for all times, there exists some $c_2(\rho_0, \omega_0) > 0$ such that

$$\sup_{\tau \in [0,t]} \|\nabla \rho(\tau)\|_{L^\infty} \geq c_2(\rho_0, \omega_0) t^2 \quad \text{for all } t \geq 0. \tag{3-22}$$

Case 2: $A_0 < 0$. In this case the right-hand side of (3-16) becomes positive for $t > |A_0|/(k_0 \pi)$. In addition, we have

$$A(t) \geq \tfrac{1}{2} k_0 \pi t \quad \text{for all } t \geq T_0 =: \frac{2|A_0|}{k_0 \pi}.$$

Once we obtain this (positive) linear lower bound for $t \geq T_0$, we can argue as in Case 1 to obtain lower bounds for $\|\omega(t)\|_{L^p(Q)}$, $\|u(t)\|_{L^\infty}$, and $\sup_{\tau \in [0,t]} \|\nabla \rho(\tau)\|_{L^\infty}$ for all $t \geq T_0$. In addition, combining the lower bound for $\|\nabla \rho(t)\|_{L^\infty}$ for $t \geq T_0$ with the trivial lower bound $\|\nabla \rho(t)\|_{L^\infty} \geq k_0/\pi$ for all times, we again have (3-22) with a smaller coefficient $c(\rho_0, \omega_0) > 0$ that only depends on the initial data. □

Remark 3.3. If the assumptions on symmetries of ρ_0 and ω_0 are dropped, the following simple argument still gives $\|\omega(t)\|_{L^1} \gtrsim t$ for $t \gg 1$. Let $Q_t := \{\Phi_t(x) : x \in [0, \pi] \times [0, \pi]\}$, and denote by

$$\Gamma_t^1 := \{\Phi_t(x) : x \in \{0\} \times [0, \pi]\} \quad \text{and} \quad \Gamma_t^2 := \{\Phi_t(x) : x \in \{\pi\} \times [0, \pi]\}$$

the left and right boundary of Q_t. (Since $u \cdot n = 0$ on $\partial\Omega$, the top and bottom boundaries of Q_t remain on $\partial\Omega$ for all times.) In addition, since ρ is preserved along the flow, at each t we have $\rho(\cdot, t)|_{\Gamma_t^1} \geq k_0 > 0$ and $\rho(\cdot, t)|_{\Gamma_t^2} \leq 0$. Thus a computation similar to (3-8) in the moving domain Q_t gives

$$\frac{d}{dt} \int_{Q_t} \omega(x, t)\, dx = \int_{Q_t} -\partial_{x_1} \rho(t)\, dx \geq k_0 \pi \quad \text{for all } t \geq 0.$$

Therefore, as long as the solution (ρ, ω) remains smooth, we have

$$\|\omega(t)\|_{L^1} \geq \int_{Q_t} \omega(x, t)\, dx \geq k_0 \pi t - \|\omega_0\|_{L^1} \quad \text{for all } t \geq 0. \tag{3-23}$$

However, since Q_t is in general largely deformed from a square for $t \gg 1$, we are not able to apply Lemma 3.1 to obtain faster growth rate for higher L^p norms.

Note that given any steady state ω_s of the 2-dimensional Euler equation on the strip Ω, $(0, \omega_s)$ is automatically a steady state of the inviscid Boussinesq equations (1-4). Thus the infinite-in-time growth estimate (3-23) directly implies that any such steady state (with zero density) is nonlinearly unstable, in the sense that, for any $0 < k_0 \ll 1$, an arbitrarily small perturbation $\rho_0 = k_0 \cos(x_1)$, $\omega_0 = \omega_s$ leads to $\lim_{t\to\infty} \|\omega(t)\|_{L^1} = \infty$. See [Bedrossian et al. 2023; Castro et al. 2019; Deng et al. 2021; Doering et al. 2018; Masmoudi et al. 2022; Tao et al. 2020; Zillinger 2023] for more results on stability/instability of steady states of the inviscid or viscous Boussinesq equations.

3.3. ***Application to 3-dimensional axisymmetric Euler equation.*** In this subsection we will prove Theorem 1.8, whose proof is a close analog of Theorem 1.4.

Proof of Theorem 1.8. Using the Biot–Savart law, one can easily check that ω^θ remains odd in z and u^θ remains even in z for all times while the solution stays smooth. Combining these symmetries with the Biot–Savart law (1-10) gives $u^z = 0$ for $z = 0$ and $z = \pi$ for all times. For a point x on the rz-plane, let us define the flow-map $\Phi_t(x) : [\pi, 2\pi] \times \mathbb{T} \to [\pi, 2\pi] \times \mathbb{T}$, given by

$$\frac{d}{dt}\Phi_t(x) = (u_r(\Phi_t(x), t), u_z(\Phi_t(x), t)).$$

Since $u_z = 0$ on $z = \pi$, for any $x \in [\pi, 2\pi] \times \{\pi\}$, we have $\Phi_t(x)$ remains on $[\pi, 2\pi] \times \{\pi\}$. From the first equation in (1-9), we have ru^θ is conserved along the trajectory. Thus, for any point (r, π) with $r \in [\pi, 2\pi]$, we have

$$ru^\theta(r, \pi, t) \geq \pi u_0^\theta(\Phi_t^{-1}(r, \pi), 0) \geq \pi k_0,$$

where the last inequality follows from the assumption $u_0^\theta \geq k_0 > 0$ on $z = \pi$ and the fact that $\Phi_t^{-1}(r, \pi) \in [\pi, 2\pi] \times \{\pi\}$. This implies

$$u^\theta(r, \pi, t) \geq \tfrac{1}{2}k_0 > 0 \quad \text{for all } r \in [\pi, 2\pi], \quad t \geq 0. \tag{3-24}$$

Applying a similar argument for $z = 0$, the assumption $|u_0^\theta| < \frac{1}{8}k_0$ on $z = 0$ leads to

$$|u^\theta(r, 0, t)| \leq \tfrac{1}{4}k_0 \quad \text{for all } r \in [\pi, 2\pi], \quad t \geq 0. \tag{3-25}$$

Defining $Q := [\pi, 2\pi] \times [0, \pi]$ to be a square on the rz-plane, the above symmetry results give $(u^r, u^z) \cdot n = 0$ on ∂Q for all times. Using this boundary condition as well as the divergence-free property of (ru^r, ru^z) in (r, z) (which follows from (1-10)), we apply the divergence theorem to obtain

$$\begin{aligned}
\frac{d}{dt}\int_Q \omega^\theta(r, z, t)\, dr\, dz &= \int_Q (ru_r, ru_z) \cdot \nabla_{r,z}\left(\frac{\omega^\theta}{r}\right) + \frac{\partial_z (u^\theta)^2}{r}\, dr\, dz \\
&= \int_Q \frac{\partial_z (u^\theta)^2}{r}\, dr\, dz \\
&= \int_\pi^{2\pi} \frac{1}{r}(u^\theta(r, \pi, t)^2 - u^\theta(r, 0, t)^2)\, dr \\
&\geq \ln 2 \tfrac{3}{16}k_0^2 \geq \tfrac{1}{10}k_0^2
\end{aligned}$$

for all times during the existence of a smooth solution, where the last inequality follows from (3-24) and (3-25). This directly implies

$$A(t) := \int_Q \omega^\theta(r,z,t)\,dr\,dz \geq \tfrac{1}{10}k_0^2 t + \int_Q \omega_0^\theta\,dr\,dz.$$

In particular, if $\int_Q \omega_0^\theta\,dr\,dz \geq 0$, this implies

$$A(t) \geq \tfrac{1}{10}k_0^2 t \quad \text{for all } t \geq 0, \tag{3-26}$$

and if $\int_Q \omega_0^\theta\,dr\,dz < 0$, we have

$$A(t) \geq \tfrac{1}{20}k_0^2 t \quad \text{for all } t \geq T_0 =: 20k_0^{-2}\left|\int_Q \omega_0^\theta\,dr\,dz\right|. \tag{3-27}$$

Another ingredient we need is the energy conservation. It is well known that the kinetic energy is conserved for the 3-dimensional Euler equation, i.e., $\int_\Omega |u(x,t)|^2\,dx = \int_\Omega |u_0|^2\,dx$. Since Ω has an inner boundary with positive radius π, this implies, in the domain Q in the rz plane, we also have

$$\int_Q (u^r(r,z,t)^2 + u^z(r,z,t)^2)\,dr\,dz \leq E_0(u_0).$$

Recall that ω^θ and (u^r, u^z) are related by $\omega^\theta = \partial_r u^z - \partial_z u^r$. Thus we can apply Lemma 3.1 to conclude that

$$\|\omega^\theta(t)\|_{L^p(Q)} \geq c_0 E_0^{-1+1/p}|A(t)|^{3-2/p} \quad \text{for all } p \in [1,\infty], \quad t \geq 0,$$

which directly leads to (1-12) once we plug estimates (3-26) and (3-27) of $A(t)$ into the above equation.

Finally, applying Green's theorem in Q, we have

$$A(t) = \int_Q \omega^\theta\,drdz = \int_Q (\partial_r u^z - \partial_z u^r)\,dr\,dz = \int_{\partial Q} u\cdot dl \leq 4\pi\|u(t)\|_{L^\infty}.$$

Combining this with the estimates (3-26) and (3-27) directly gives (1-13), finishing the proof. □

Appendix: Proof of Lemma 2.2

In the appendix we prove Lemma 2.2. The proof is almost the same as in [Kiselev and Yao 2023] other than a small improvement in part (a). We sketch a proof for both parts below for the sake of completeness.

Proof of Lemma 2.2 (a). Here the proof mostly follows [Kiselev and Yao 2023, (3.4)], except that we make a small improvement dropping the assumption $\|\partial_1\mu\|_{\dot{H}^{-1}}^2 < \frac{1}{4}\|\mu\|_{L^2}^2$ in that paper. Let us define

$$\delta := \|\partial_1\mu\|_{\dot{H}^{-1}(\mathbb{R}^2)}^2, \qquad A := \|\mu\|_{L^2(\mathbb{R}^2)}^2.$$

Clearly,

$$\delta = \int_{\mathbb{R}^2} \frac{\xi_1^2}{|\xi|^2}|\hat{\mu}|^2\,d\xi \leq A.$$

Let us discuss the following two cases.

Case 1: $\delta < \frac{1}{4}A$. In this case let us define

$$D_\delta := \left\{(\xi_1, \xi_2) : \frac{|\xi_1|}{|\xi|} \geq \sqrt{\frac{2\delta}{A}}\right\}.$$

By definition of D_δ, we have

$$\delta \geq \int_{D_\delta} \frac{\xi_1^2}{|\xi|^2} |\hat{\mu}(\xi)|^2 \, d\xi \geq \frac{2\delta}{A} \int_{D_\delta} |\hat{\mu}|^2 \, d\xi.$$

This gives $\int_{D_\delta} |\hat{\mu}|^2 \, d\xi \leq \frac{1}{2}A$, and thus $\int_{D_\delta^c} |\hat{\mu}|^2 \, d\xi \geq \frac{1}{2}A$. Note that D_δ^c can be expressed in polar coordinates as

$$D_\delta^c = \left\{(r\cos\theta, r\sin\theta) : r \geq 0, |\cos\theta| < \sqrt{\frac{2\delta}{A}}\right\}.$$

Since $\mu \in C_c^\infty(\mathbb{R}^2)$, we have

$$\|\hat{\mu}\|_{L^\infty(\mathbb{R}^2)} \leq (2\pi)^{-1} \|\mu\|_{L^1(\mathbb{R}^2)} =: B.$$

Let $h_\delta > 0$ be such that $|D_\delta^c \cap \{|\xi_2| < h_\delta\}| = (4B^2)^{-1}A$, which we will estimate later. Such a definition gives

$$\int_{D_\delta^c \cap \{|\xi_2| \geq h_\delta\}} |\hat{\mu}|^2 \, d\xi = \int_{D_\delta^c} |\hat{\mu}|^2 \, d\xi - \int_{D_\delta^c \cap \{|\xi_2| < h_\delta\}} |\hat{\mu}|^2 \, d\xi \geq \tfrac{1}{2}A - (4B^2)^{-1}AB^2 = \tfrac{1}{4}A,$$

which implies

$$\|\mu\|_{\dot{H}^s(\mathbb{R}^2)}^2 \geq \int_{\mathbb{R}^2} |\xi_2|^{2s} |\hat{\mu}|^2 \, d\xi \geq h_\delta^{2s} \int_{D_\delta^c \cap \{|\xi_2| \geq h_\delta\}} |\hat{\mu}|^2 \, d\xi \geq \tfrac{1}{4}Ah_\delta^{2s}. \tag{A-1}$$

To estimate h_δ, let us define $\theta_0 := \cos^{-1}(\sqrt{2\delta/A})$. Since $D_\delta^c \cap \{|\xi_2| < h_\delta\}$ consists of two identical triangles with height h_δ and base $2h_\delta \cot\theta_0$, we have

$$(4B^2)^{-1}A = |D_\delta^c \cap \{|\xi_2| < h_\delta\}| = 2h_\delta^2 \cot\theta_0 \leq 4\sqrt{\delta}A^{-1/2}h_\delta^2,$$

where the inequality follows from $\cos\theta_0 = \sqrt{2\delta/A}$ and $\sin\theta_0 = \sqrt{1 - 2\delta/A} \geq 1/\sqrt{2}$, due the assumption $\delta < \frac{1}{4}A$ in Case 1. Therefore $h_\delta \geq (4B)^{-1}A^{3/4}\delta^{-1/4}$. Plugging it into (A-1) yields

$$\|\mu\|_{\dot{H}^s(\mathbb{R}^2)} \geq \tfrac{1}{2}\sqrt{A}h_\delta^s \geq c(s, A, B)\delta^{-s/4},$$

finishing the proof of Lemma 2.2 in Case 1.

Case 2: $\delta \geq \frac{1}{4}A$. As in Case 1, let us define $\|\hat{\mu}\|_{L^\infty(\mathbb{R}^2)} \leq (2\pi)^{-1}\|\mu\|_{L^1(\mathbb{R}^2)} =: B$. Let $r_0 := (A/(2\pi B^2))^{1/2}$. Such a definition leads to

$$\int_{B(0,r_0)} |\hat{\mu}|^2 \, d\xi \leq \pi r_0^2 \|\hat{\mu}\|_{L^\infty(\mathbb{R}^2)}^2 \leq \tfrac{1}{2}A,$$

and thus

$$\|\mu\|_{\dot{H}^s(\mathbb{R}^2)}^2 \geq \int_{B(0,r_0)^c} |\xi|^{2s} |\hat{\mu}|^2 \, d\xi \geq r_0^{2s} \tfrac{1}{2}A \geq c(s, A, B)\delta^{-s/4},$$

where the last inequality follows from the assumption $\delta \geq \frac{1}{4}A$ in Case 2. This finishes the proof of part (a). □

Proof of Lemma 2.2 (*b*). This part is equivalent to the last (unnumbered) equation in the proof of Theorem 1.2 in [Kiselev and Yao 2023]. We sketch a proof below for completeness, and also to clarify the dependence of $c_2\big(s, \int_{\mathbb{T}\times[0,\pi]} \mu^{1/3}\,dx\big)$ in (2-12).

For any $k = (k_1, k_2) \in \mathbb{Z}^2$, the Fourier coefficient $\hat{\mu}(k_1, k_2)$ can be written as

$$\begin{aligned}\hat{\mu}(k_1,k_2) &= \frac{1}{(2\pi)^2}\int_{\mathbb{T}} e^{-ik_1x_1}\int_{\mathbb{T}} e^{-ik_2x_2}\mu(x_1,x_2)\,dx_2\,dx_1\\ &= \frac{1}{(2\pi)^2}\int_{\mathbb{T}} e^{-ik_1x_1}(-2i)\underbrace{\int_0^{\pi}\sin(k_2x_2)\mu(x_1,x_2)\,dx_2}_{=:g(x_1,k_2)}\,dx_1,\end{aligned} \tag{A-2}$$

where the last identity is due to μ being odd in x_2. With $g(x_1, k_2)$ defined in the last line of (A-2), when setting $k_2 = 1$, we claim that $g(x_1, 1)$ satisfies the following properties:

(a) $g(x_1, 1)$ is even in x_1 and nonnegative for all $x_1 \in \mathbb{T}$.

(b) $g(0, 1) = 0$.

(c) $\int_{\mathbb{T}} g(x_1, 1)\,dx_1 \ge c\big(\int_D \mu(x)^{1/3}\,dx\big)^3$ for some universal constant $c > 0$.

Here property (a) follows from the facts that μ is even in x_1 and nonnegative on $D := [0, \pi]^2$. Property (b) follows from $\mu(0,\cdot) \equiv 0$. For property (c), note that

$$\int_{\mathbb{T}} g(x_1,1)\,dx_1 = 2\int_0^{\pi} g(x_1,1)\,dx_1 = 2\int_D \sin(x_2)\mu(x)\,dx.$$

Combining Hölder's inequality with the fact that $\sin(x_2)\mu(x) \ge 0$ in D, we have

$$\int_D \sin(x_2)\mu(x)\,dx \ge \bigg(\int_D \sin(x_2)^{-1/2}\,dx\bigg)^{-2}\bigg(\int_D \mu(x)^{1/3}\,dx\bigg)^3 \ge c_0\bigg(\int_D \mu(x)^{1/3}\,dx\bigg)^3$$

for some universal constant $c_0 > 0$. This proves property (c).

For any $k_1 \in \mathbb{Z}$, let $\hat{g}(k_1)$ be the Fourier coefficient of $g(\cdot, 1)$; that is,

$$\hat{g}(k_1) := \frac{1}{2\pi}\int_{\mathbb{T}} e^{-ik_1x_1} g(x_1,1)\,dx. \tag{A-3}$$

Denote by $\bar{g} := \frac{1}{2\pi}\int_{\mathbb{T}} g(x_1, 1)\,dx_1$ the average of $g(\cdot, 1)$. Applying the definition of $\hat{g}$ to (A-2) gives

$$\hat{\mu}(k_1,1) = \frac{-2i}{2\pi}\hat{g}(k_1) \quad \text{for any } k_1 \in \mathbb{Z}. \tag{A-4}$$

This allows us to bound $\delta := \|\partial_1\mu\|^2_{\dot{H}^{-1}(\mathbb{T}^2)}$ from below as

$$\begin{aligned}\delta &\ge (2\pi)^2\sum_{k_1\in\mathbb{Z}\setminus\{0\}}\frac{k_1^2}{k_1^2+1}|\hat{\mu}(k_1,1)|^2\\ &\ge 2\sum_{k_1\in\mathbb{Z}\setminus\{0\}}|\hat{g}(k_1)|^2 = \frac{1}{\pi}\int_{\mathbb{T}}|g(x_1,1)-\bar{g}|^2\,dx_1.\end{aligned} \tag{A-5}$$

By property (c), $\bar{g} \geq c\left(\int_D \mu(x)^{1/3}\,dx\right)^3 > 0$. Applying [Kiselev and Yao 2023, Lemma 3.3] to $g(x_1, 1)$ yields

$$\|g(\cdot\,,1)\|_{\dot{H}^s(\mathbb{T})} \geq c\left(s, \int_D \mu(x)^{1/3}\,dx\right)\delta^{-s+1/2} \quad \text{for all } s > \tfrac{1}{2}. \tag{A-6}$$

Note that

$$\|g(\cdot\,,1)\|^2_{\dot{H}^s(\mathbb{T})} = 2\pi^3 \sum_{k_1 \neq 0} |k_1|^{2s} |\hat{\mu}(k_1, 1)|^2 \leq \frac{\pi}{\sqrt{2}} \|\partial_1 \mu\|^2_{\dot{H}^{s-1}(\mathbb{T}^2)} \leq \frac{\pi}{\sqrt{2}} \|\mu\|^2_{\dot{H}^s(\mathbb{T}^2)}, \tag{A-7}$$

where the first inequality follows by the assumption $s > \frac{1}{2}$. Finally, combining inequalities (A-6) and (A-7) gives (2-12). □

Acknowledgements

Kiselev was partially supported by NSF-DMS grants 2006372 and 2306726. Park was partially supported by the SNF grant 212573-FLUTURA and the Ambizione fellowship project PZ00P2-216083. Yao was partially supported by the NUS startup grant A-0008382-00-00 and MOE Tier 1 grant A-0008491-00-00.

References

[Abidi and Hmidi 2007] H. Abidi and T. Hmidi, "On the global well-posedness for Boussinesq system", *J. Differential Equations* **233**:1 (2007), 199–220. MR Zbl

[Bardos and Titi 2007] C. Bardos and E. S. Titi, "Euler equations for an ideal incompressible fluid", *Uspekhi Mat. Nauk* **62**:3(375) (2007), 5–46. In Russian; translated in *Russian Math. Surv.* **62**:3 (2007), 409–451. MR Zbl

[Bedrossian et al. 2023] J. Bedrossian, R. Bianchini, M. C. Zelati, and M. Dolce, "Nonlinear inviscid damping and shear-buoyancy instability in the two-dimensional Boussinesq equations", *Comm. Pure Appl. Math.* **76**:12 (2023), 3551–4136. Zbl

[Brandolese and Schonbek 2012] L. Brandolese and M. E. Schonbek, "Large time decay and growth for solutions of a viscous Boussinesq system", *Trans. Amer. Math. Soc.* **364**:10 (2012), 5057–5090. MR Zbl

[Castro et al. 2019] Á. Castro, D. Córdoba, and D. Lear, "On the asymptotic stability of stratified solutions for the 2D Boussinesq equations with a velocity damping term", *Math. Models Methods Appl. Sci.* **29**:7 (2019), 1227–1277. MR Zbl

[Chae 2006] D. Chae, "Global regularity for the 2D Boussinesq equations with partial viscosity terms", *Adv. Math.* **203**:2 (2006), 497–513. MR Zbl

[Chae and Nam 1997] D. Chae and H.-S. Nam, "Local existence and blow-up criterion for the Boussinesq equations", *Proc. Roy. Soc. Edinburgh Sect. A* **127**:5 (1997), 935–946. MR Zbl

[Chae et al. 1999] D. Chae, S.-K. Kim, and H.-S. Nam, "Local existence and blow-up criterion of Hölder continuous solutions of the Boussinesq equations", *Nagoya Math. J.* **155** (1999), 55–80. MR Zbl

[Chen and Hou 2021] J. Chen and T. Y. Hou, "Finite time blowup of 2D Boussinesq and 3D Euler equations with $C^{1,\alpha}$ velocity and boundary", *Comm. Math. Phys.* **383**:3 (2021), 1559–1667. Correction in **399**:1 (2023), 573–575. MR Zbl

[Chen and Hou 2022] J. Chen and T. Y. Hou, "Stable nearly self-similar blowup of the 2D Boussinesq and 3D Euler equations with smooth data, I: Analysis", preprint, 2022. Zbl arXiv 2210.07191

[Choi and Jeong 2023] K. Choi and I.-J. Jeong, "Filamentation near Hill's vortex", *Comm. Partial Differential Equations* **48**:1 (2023), 54–85. MR Zbl

[Danchin 2013] R. Danchin, "Remarks on the lifespan of the solutions to some models of incompressible fluid mechanics", *Proc. Amer. Math. Soc.* **141**:6 (2013), 1979–1993. MR Zbl

[Danchin and Paicu 2011] R. Danchin and M. Paicu, "Global existence results for the anisotropic Boussinesq system in dimension two", *Math. Models Methods Appl. Sci.* **21**:3 (2011), 421–457. MR Zbl

[Deng et al. 2021] W. Deng, J. Wu, and P. Zhang, "Stability of Couette flow for 2D Boussinesq system with vertical dissipation", *J. Funct. Anal.* **281**:12 (2021), art. id. 109255. MR Zbl

[Denisov 2009] S. A. Denisov, "Infinite superlinear growth of the gradient for the two-dimensional Euler equation", *Discrete Contin. Dyn. Syst.* **23**:3 (2009), 755–764. MR Zbl

[Denisov 2015] S. A. Denisov, "Double exponential growth of the vorticity gradient for the two-dimensional Euler equation", *Proc. Amer. Math. Soc.* **143**:3 (2015), 1199–1210. MR Zbl

[Doering et al. 2018] C. R. Doering, J. Wu, K. Zhao, and X. Zheng, "Long time behavior of the two-dimensional Boussinesq equations without buoyancy diffusion", *Phys. D* **376/377** (2018), 144–159. MR Zbl

[Drivas and Elgindi 2023] T. D. Drivas and T. M. Elgindi, "Singularity formation in the incompressible Euler equation in finite and infinite time", *EMS Surv. Math. Sci.* **10**:1 (2023), 1–100. MR Zbl

[Elgindi 2021] T. Elgindi, "Finite-time singularity formation for $C^{1,\alpha}$ solutions to the incompressible Euler equations on $\mathbb{R}^3$", *Ann. of Math.* (2) **194**:3 (2021), 647–727. MR Zbl

[Elgindi and Jeong 2019] T. M. Elgindi and I.-J. Jeong, "Finite-time singularity formation for strong solutions to the axi-symmetric 3D Euler equations", *Ann. PDE* **5**:2 (2019), art. id. 16. MR Zbl

[Elgindi and Jeong 2020] T. M. Elgindi and I.-J. Jeong, "Finite-time singularity formation for strong solutions to the Boussinesq system", *Ann. PDE* **6**:1 (2020), art. id. 5. MR Zbl

[Gancedo and García-Juárez 2017] F. Gancedo and E. García-Juárez, "Global regularity for 2D Boussinesq temperature patches with no diffusion", *Ann. PDE* **3**:2 (2017), art. id. 14. MR Zbl

[Gancedo and García-Juárez 2020] F. Gancedo and E. García-Juárez, "Regularity results for viscous 3D Boussinesq temperature fronts", *Comm. Math. Phys.* **376**:3 (2020), 1705–1736. MR Zbl

[Gill and Adrian 1982] A. E. Gill and E. Adrian, *Atmosphere-ocean dynamics*, Int. Geophys. Ser. **30**, Academic Press, New York, 1982. Zbl

[Hmidi and Keraani 2007] T. Hmidi and S. Keraani, "On the global well-posedness of the two-dimensional Boussinesq system with a zero diffusivity", *Adv. Differential Equations* **12**:4 (2007), 461–480. MR Zbl

[Hou 2022] T. Y. Hou, "Potential singularity of the 3D Euler equations in the interior domain", *Found. Comput. Math.* (online publication September 2022). Zbl

[Hou and Huang 2022] T. Y. Hou and D. Huang, "A potential two-scale traveling wave singularity for 3D incompressible Euler equations", *Phys. D* **435** (2022), art. id. 133257. MR Zbl

[Hou and Huang 2023] T. Y. Hou and D. Huang, "Potential singularity formation of incompressible axisymmetric Euler equations with degenerate viscosity coefficients", *Multiscale Model. Simul.* **21**:1 (2023), 218–268. MR Zbl

[Hou and Li 2005] T. Y. Hou and C. Li, "Global well-posedness of the viscous Boussinesq equations", *Discrete Contin. Dyn. Syst.* **12**:1 (2005), 1–12. MR Zbl

[Hu et al. 2013] W. Hu, I. Kukavica, and M. Ziane, "On the regularity for the Boussinesq equations in a bounded domain", *J. Math. Phys.* **54**:8 (2013), art. id. 081507. MR Zbl

[Ju 2017] N. Ju, "Global regularity and long-time behavior of the solutions to the 2D Boussinesq equations without diffusivity in a bounded domain", *J. Math. Fluid Mech.* **19**:1 (2017), 105–121. MR Zbl

[Kiselev and Šverák 2014] A. Kiselev and V. Šverák, "Small scale creation for solutions of the incompressible two-dimensional Euler equation", *Ann. of Math.* (2) **180**:3 (2014), 1205–1220. MR Zbl

[Kiselev and Yao 2023] A. Kiselev and Y. Yao, "Small scale formations in the incompressible porous media equation", *Arch. Ration. Mech. Anal.* **247**:1 (2023), art. id. 1. MR Zbl

[Kukavica and Wang 2020] I. Kukavica and W. Wang, "Long time behavior of solutions to the 2D Boussinesq equations with zero diffusivity", *J. Dynam. Differential Equations* **32**:4 (2020), 2061–2077. MR Zbl

[Kukavica et al. 2023] I. Kukavica, D. Massatt, and M. Ziane, "Asymptotic properties of the Boussinesq equations with Dirichlet boundary conditions", *Discrete Contin. Dyn. Syst.* **43**:8 (2023), 3060–3081. MR Zbl

[Lai et al. 2011] M.-J. Lai, R. Pan, and K. Zhao, "Initial boundary value problem for two-dimensional viscous Boussinesq equations", *Arch. Ration. Mech. Anal.* **199**:3 (2011), 739–760. MR Zbl

[Larios et al. 2013] A. Larios, E. Lunasin, and E. S. Titi, "Global well-posedness for the 2D Boussinesq system with anisotropic viscosity and without heat diffusion", *J. Differential Equations* **255**:9 (2013), 2636–2654. MR Zbl

[Luo and Hou 2014] G. Luo and T. Y. Hou, "Potentially singular solutions of the 3D axisymmetric Euler equations", *Proc. Nat. Acad. Sci.* **111**:36 (2014), 12968–12973. Zbl

[Majda 2003] A. Majda, *Introduction to PDEs and waves for the atmosphere and ocean*, Courant Lecture Notes in Math. **9**, Courant Inst. Math. Sci., New York, 2003. MR Zbl

[Majda and Bertozzi 2002] A. J. Majda and A. L. Bertozzi, *Vorticity and incompressible flow*, Cambridge Texts in Appl. Math. **27**, Cambridge Univ. Press, 2002. MR Zbl

[Masmoudi et al. 2022] N. Masmoudi, B. Said-Houari, and W. Zhao, "Stability of the Couette flow for a 2D Boussinesq system without thermal diffusivity", *Arch. Ration. Mech. Anal.* **245**:2 (2022), 645–752. MR Zbl

[Nadirashvili 1991] N. S. Nadirashvili, "Wandering solutions of the two-dimensional Euler equation", *Funktsional. Anal. i Prilozhen.* **25**:3 (1991), 70–71. In Russian; translated in *Funct. Anal. Appl.* **25**:3 (1992), 220–221. MR Zbl

[Pedlosky 1979] J. Pedlosky, *Geophysical fluid dynamics*, Springer, 1979. Zbl

[Sarria and Wu 2015] A. Sarria and J. Wu, "Blowup in stagnation-point form solutions of the inviscid 2d Boussinesq equations", *J. Differential Equations* **259**:8 (2015), 3559–3576. MR Zbl

[Serre 1991] D. Serre, "Écoulement parfait incompressible à rotationnel croissant avec le temps", *C. R. Acad. Sci. Paris Sér. I Math.* **312**:3 (1991), 315–318. MR Zbl

[Serre 1999] D. Serre, "La croissance de la vorticité dans les écoulements parfaits incompressibles", *C. R. Acad. Sci. Paris Sér. I Math.* **328**:6 (1999), 549–552. MR Zbl

[Tao et al. 2020] L. Tao, J. Wu, K. Zhao, and X. Zheng, "Stability near hydrostatic equilibrium to the 2D Boussinesq equations without thermal diffusion", *Arch. Ration. Mech. Anal.* **237**:2 (2020), 585–630. MR Zbl

[Temam 1988] R. Temam, *Infinite-dimensional dynamical systems in mechanics and physics*, Appl. Math. Sci. **68**, Springer, 1988. MR Zbl

[Wang et al. 2023] Y. Wang, C.-Y. Lai, J. Gómez-Serrano, and T. Buckmaster, "Asymptotic self-similar blow-up profile for three-dimensional axisymmetric Euler equations using neural networks", *Phys. Rev. Lett.* **130**:24 (2023), art. id. 244002. MR

[Zillinger 2023] C. Zillinger, "On stability estimates for the inviscid Boussinesq equations", *J. Nonlinear Sci.* **33**:6 (2023), art. id. 106. MR Zbl

[Zlatoš 2015] A. Zlatoš, "Exponential growth of the vorticity gradient for the Euler equation on the torus", *Adv. Math.* **268** (2015), 396–403. MR Zbl

Received 1 Dec 2022. Revised 6 Jun 2023. Accepted 18 Sep 2023.

ALEXANDER KISELEV: kiselev@math.duke.edu
Department of Mathematics, Duke University, Durham, NC, United States

JAEMIN PARK: jpark776@yonsei.ac.kr
Departement Mathematik und Informatik, University of Basel, Basel, Switzerland
Current address: *Department of Mathematics, Yonsei University, Seoul, South Korea*

YAO YAO: yaoyao@nus.edu.sg
Department of Mathematics, National University of Singapore, Singapore

ANALYSIS AND PDE
Vol. 18 (2025), No. 1, pp. 199–264
DOI: 10.2140/apde.2025.18.199

C^∞ PARTIAL REGULARITY OF THE SINGULAR SET IN THE OBSTACLE PROBLEM

FEDERICO FRANCESCHINI AND WIKTORIA ZATOŃ

We show that the singular set Σ in the classical obstacle problem can be locally covered by a C^∞ hypersurface, up to an "exceptional" set E, which has Hausdorff dimension at most $n-2$ (countable in the $n=2$ case). Outside this exceptional set, the solution admits a polynomial expansion of arbitrarily large order. We also prove that $\Sigma \setminus E$ is extremely unstable with respect to monotone perturbations of the boundary datum. We apply this result to the planar Hele-Shaw flow, showing that the free boundary can have singular points for at most countable many times.

1. Introduction

1.1. ***The classical obstacle problem.*** The classical obstacle problem consists in studying the solutions of the variational problem

$$\min\left\{\frac{1}{2}\int_{B_1} |\nabla w|^2 : w \geq \varphi \text{ in } B_1 \subseteq \mathbb{R}^n,\ w = g \text{ on } \partial B_1\right\},$$

where $g : \partial B_1 \to \mathbb{R}$ and $\varphi : B_1 \to \mathbb{R}$ are given, with $\varphi < g$ on ∂B_1. In two dimensions an intuitive interpretation of this problem is the following: The graph of the minimizer w represents the shape of a thin membrane stretched over $\bar{B}_1$ and fixed on ∂B_1 along the profile g. The hypograph of φ represents a solid "obstacle" above which the membrane must lie, possibly touching it. The Dirichlet energy, finally, corresponds to the linearization of the surface energy of the membrane, which is assumed proportional to the area of the graph of w.

MSC2020: 35R35.
Keywords: obstacle problem, singular set, higher-regularity.

It is well known (see for example the exposition in [Fernández-Real and Ros-Oton 2022, Chapter 5]) that there exists a unique optimal shape v and that it enjoys $C^{1,1}_{\text{loc}}$ regularity, provided that φ is smooth enough. Furthermore, $u := v - \varphi$ solves the Euler–Lagrange equation

$$\Delta u = -\Delta\varphi\chi_{\{u>0\}} \quad \text{in } B_1.$$

One of the most challenging problems is to understand the a priori unknown interface $\partial\{u > 0\}$, called the "free boundary". Unless $\Delta\varphi < 0$, simple examples show that the free boundary can be any closed set; hence it is standard to assume that φ is superharmonic. Thus, we deal with solutions of

$$\begin{cases} \Delta u = f\chi_{\{u>0\}} & \text{in } B_1, \\ u \geq 0 & \text{in } B_1, \end{cases} \tag{1-1}$$

where $f \in C^\infty(B_1)$ is given and positive.

By classical works of Caffarelli [1977; 1998], the free boundary $\partial\{u > 0\}$ splits into a regular and a singular part:

$$\partial\{u > 0\} = \text{Reg}(u) \cup \Sigma(u).$$

Points in these sets can be characterized, for example, by density considerations:

$$\begin{aligned} x_\circ \in \text{Reg}(u) \quad &\iff \quad |\{u = 0\} \cap B_r(x_\circ)| = \tfrac{1}{2}|B_r| + o(r^n), \\ x_\circ \in \Sigma(u) \quad &\iff \quad |\{u = 0\} \cap B_r(x_\circ)| = o(r^n). \end{aligned}$$

Caffarelli showed that $\text{Reg}(u)$ is relatively open in the free boundary and, locally, is a C^1 hypersurface (smoothness and analyticity were proved later in [Kinderlehrer and Nirenberg 1977]). On the other hand, $\Sigma(u)$ can always be covered, locally, by a single C^1 hypersurface (see [Caffarelli 1998, Theorem 8]). Thus, when we speak about "regularity of $\Sigma(u)$" we are actually speaking about the regularity of the manifold covering it. In this paper we will improve the smoothness of this hypersurface. We remark that $\Sigma(u)$ can display a very wild structure, as long as it is contained inside a single hypersurface; see example (1-2).

1.2. ***The singular set: important examples and known results.*** The following simple example shows that $\Sigma(u)$ can be rather wild. Furthermore, it can have Hausdorff dimension equal to $n-1$, thus it can be as "large" as $\text{Reg}(u)$. Consider the function

$$u(x) := x_n^2 + h(x') \quad \text{for } (x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}, \tag{1-2}$$

where $h \in C^\infty(\mathbb{R}^{n-1})$ is nonnegative. Possibly multiplying h by a small factor, u solves (1-1) (with some f depending on h) and

$$\Sigma(u) = \{x_n = 0\} \cap \{h = 0\}, \qquad \text{Reg}(u) = \varnothing.$$

Hence $\Sigma(u)$ can be any closed subset of $\{x_n = 0\}$. See the second point below for even wilder examples due to Schaeffer [1977], where the contact set has nonempty interior. In this paper we show that, locally, $\Sigma(u)$ is always contained in a C^∞ hypersurface (the hyperplane $\{x_n = 0\}$, in this example), except for an

$(n-2)$-dimensional piece (the empty set, in this example). Before turning to the statements, let us try to give an overview of what is known about $\Sigma(u)$.

- Concerning the pointwise structure, $\Sigma(u)$ consists precisely of those points $x_\circ \in \partial\{u > 0\}$ such that

$$r^{-2}u(x_\circ + rx) \to p_{2,x_\circ}(x) \quad \text{as } r \downarrow 0,$$

where $p_{2,x_\circ}$ is a convex and 2-homogeneous polynomial with $\Delta p_{2,x_\circ} = f(x_\circ)$. Thus, zooming in on $x_\circ$, one sees the contact set $\{u = 0\}$ clustering around the linear space $\{p_{2,x_\circ} = 0\}$. When $\dim\{p_{2,x_\circ} = 0\} = m$ for some integer $m \le n-1$, this suggests that $\Sigma(u)$ displays, qualitatively, an m-dimensional structure at $x_\circ$.

- Concerning the local structure, when $n = 2$ and f is real analytic, Sakai [1993] gave a complete characterization of the possible shapes of the free boundary around a singular point. In brief, $\Sigma(u)$ is locally either an analytic curve, or an isolated point, around which $\partial\{u > 0\}$ is the union of at most two analytic arcs. In particular, $\Sigma(u)$ has codimension 1 inside the free boundary. We remark that his approach relies on complex analysis techniques [Sakai 1991]. On the other hand, Schaeffer [1977] constructed examples of rather wild free boundaries in the case where f is "just" C^∞. He showed that

$$\operatorname{int}(\{u = 0\} \cap \{x_n = 0\}) \quad \text{and} \quad \{u = 0\} \cap \{x_n = 0\}$$

can be any nested couple of relatively open and closed subsets of $\{x_n = 0\}$ (the interior part is taken with respect to the relative topology). In particular, the contact set might form infinitely many cusps, which in turn produce an arbitrarily closed subset of $\{x_n = 0\}$ as the singular set. This shows that the sharpness of Sakai's results is due to analytic rigidity.

- Despite the counterexamples of Schaeffer, it is still possible to obtain refined statements about the shape of $\Sigma(u)$. Caffarelli [1977; 1998] showed that $\Sigma(u)$ is locally contained in a C^1 hypersurface. More precisely, if we partition $\Sigma(u)$ into the strata

$$\Sigma_m := \{x_\circ \in \Sigma(u) : \dim\ker A_{x_\circ} = m\} \quad \text{for } m = 0, \dots, n-1,$$

then each Σ_m is locally contained in an m-dimensional C^1 manifold. Building on Weiss [1999] and Monneau [2003], Figalli and Serra [2019] extended these results by showing that $\Sigma(u) \setminus E$ can be covered by $C^{1,1}$ manifolds, where the excluded set E has Hausdorff dimension at most $n-2$.

- By the implicit function theorem, this type of covering result is tightly linked with the fact that u admits a polynomial expansion around singular points. From this perspective, Caffarelli showed that, at each $x_\circ \in \Sigma(u)$,

$$u(x) = p_{2,x_\circ}(x - x_\circ) + \sigma(x - x_\circ)|x - x_\circ|^2,$$

with an abstract dimensional modulus of continuity σ. This was improved in [Colombo et al. 2018], showing that $\sigma(x - x_\circ) \le C\big|\log|x - x_\circ|\big|^{-\varepsilon_\circ}$ for some dimensional $C, \varepsilon_\circ > 0$. Figalli and Serra [2019] proved that $\sigma(x - x_\circ) \le C|x - x_\circ|^{\alpha_\circ}$ when $x_\circ \in \Sigma_{n-1}$, and also that $\sigma(x - x_\circ) \le C|x - x_\circ|$ provided $x_\circ \in \Sigma(u) \setminus E$, where $\dim_{\mathcal{H}} E \le n-2$. Similar results were recovered in [Savin and Yu 2023] with independent methods. This analysis was pushed further by Figalli, Ros-Oton and Serra [Figalli et al. 2020] in the framework of "generic" regularity. They showed that, for all $\varepsilon > 0$ small, there is a set $E \subseteq \Sigma(u)$ with Hausdorff dimension at most $n-2$ such that, if $x_\circ \in \Sigma(u) \setminus E$, then

$$u(x) = P_{4,x_\circ}(x) + O(|x - x_\circ|^{5-\varepsilon}) \tag{1-3}$$

for some polynomial $P_{4,x_\circ}$ with $\Delta P_{4,x_\circ} = 1$.

1.3. *Main results of this paper.* Concerning expansion (1-3), the approach in [Figalli et al. 2020] was blocked at order 5, and new ideas were needed to go further, as we explain in detail in Section 1.4.2 below. In fact, the natural question whether such an expansion could be pushed to some order $k \geq 5$ (at most points) was explicitly raised in [Figalli 2018, p. 22]. The main contribution of this paper is providing a positive answer to this question: we prove that u admits a C^∞ polynomial expansion at most points in Σ.

Theorem 1.1. *Let $n \geq 2$, $mu > 0$ and $f \in C^\infty(B_1)$ be given, with $f \geq \mu$. Let $u \in C^{1,1}_{\mathrm{loc}}(B_1)$ be a solution of the obstacle problem (1-1), and let Σ be its singular set. Then there exists a closed set $\Sigma^\infty \subseteq \Sigma$ such that*

(i) $\dim_{\mathcal{H}}(\Sigma \setminus \Sigma^\infty) \leq n-2$ (*countable if* $n = 2$),

(ii) *locally, Σ^∞ is contained in one $(n-1)$-dimensional C^∞ manifold.*

Moreover, at every point $x \in \Sigma^\infty$, the solution u has a polynomial expansion of arbitrarily large order. That is, for every $x \in \Sigma^\infty$ and $k \in \mathbb{N}$, there exists a unique polynomial $P_{k,x}$ with $\deg P_{k,x} \leq k$ *such that the expansion*

$$|u(x+h) - P_{k,x}(h)| \leq C|h|^{k+1} \quad \textit{for all } |h| \leq \tfrac{1}{2}(1-|x|) \tag{1-4}$$

holds with a constant C depending only on n, k, μ, $\|f\|_{C^{k+1}}$, $1-|x|$. We further have that $\Delta P_{k+2,x} = f_{k,x}$, where $f_{k,x}$ is the k-th Taylor polynomial of f centered at x. Finally, the map $\Sigma^\infty \ni x \mapsto (P_{k,x})_{k\in\mathbb{N}}$ is smooth in the sense of Whitney.[1]

We remark that there are solutions in dimension 2 with $f \equiv 1$ for which $0 \in \Sigma(u)$, but where expansion (1-4) does not hold at 0 for $k \geq 3$ (e.g., see the cusp-type solutions in [Sakai 1993]). In this sense the dimensional bound in (i) is optimal. Furthermore, example (1-2) shows one cannot hope to show that Σ is a smooth manifold, at least when the right-hand sides f are not analytic — this motivates (ii).

Example (1-2) is also a model situation which our result describes effectively: in this case $\Sigma = \Sigma^\infty$.

In dimension 2, Theorem 1.1 can be also read as an extension of Sakai's result to nonanalytic right-hand sides f (see the proof of Corollary 1.4 for a detailed comparison of the two results).

Our analysis further shows that the set Σ^∞ is extremely unstable with respect to monotone perturbations of the boundary datum. Indeed, following [Figalli et al. 2020], we also prove:

Theorem 1.2. *Let $\{u^t\}_{t\in(-1,1)}$ be a family of solutions to (1-1), with f independent from t, which is "uniformly monotone" in the sense that, for every $t \in (-1,1)$ and any compact set $K \subseteq \partial B_1 \cap \{u^t > 0\}$, there exists $c = c(t,K) > 0$ such that*

$$\min_{x\in K}(u^{t+h}(x) - u^t(x)) \geq ch \quad \textit{for all } -1 < t < t+h < 1. \tag{1-5}$$

[1] We denote by $\dim_{\mathcal{H}}$ the Hausdorff dimension, and we refer to Whitney's definition of smoothness on a closed set [Whitney 1934, Section 3].

With the notation of Theorem 1.1, define the singular sets

$$\mathbf{\Sigma} := \{(x_\circ, t_\circ) : x_\circ \in \Sigma(u^{t_\circ})\},$$
$$\mathbf{\Sigma}^\infty := \{(x_\circ, t_\circ) : x_\circ \in \Sigma^\infty(u^{t_\circ})\}.$$

Then, denoting the standard projections by $\pi_t : B_1 \times (-1, 1) \to (-1, 1)$ *and* $\pi_x : B_1 \times (-1, 1) \to B_1$, *we have that* $\mathbf{\Sigma}$ *is a graph over* $\pi_x(\mathbf{\Sigma})$, *and*

(i) $\dim_{\mathcal{H}}(\pi_t(\mathbf{\Sigma}^\infty)) = 0$ *in any dimension* $n \geq 2$,

(ii) $\dim_{\mathcal{H}}(\pi_t(\mathbf{\Sigma})) = 0$ *in dimension* $n = 2$,

(iii) $\dim_{\mathcal{H}}(\pi_x(\mathbf{\Sigma} \setminus \mathbf{\Sigma}^\infty)) \leq n - 2$ (*countable if* $n = 2$).

Remark 1.3. The Hausdorff dimension bound in (i) can actually be improved to zero Minkowski dimension (see [Mattila 1995, Chapter 5] for the definition).

Combining this result with Sakai's classification we also get an improvement of [Figalli et al. 2020, Theorem 1.2] concerning the generic regularity of the free boundary in the planar Hele-Shaw flow.

Corollary 1.4. *Let* $O \subseteq \mathbb{R}^2$ *be an open and bounded set with Lipschitz boundary, and let* $\Omega := \mathbb{R}^2 \setminus \overline{O}$. *For each* $t > 0$, *let* u^t *be a weak solution of*

$$\begin{cases} \Delta u^t = \chi_{\{u^t > 0\}} & \text{in } \Omega, \\ u^t = t & \text{on } \partial\Omega, \\ u^t \geq 0 & \text{in } \Omega. \end{cases} \tag{1-6}$$

Then the set of $t \in (0, \infty)$ *such that* $\Sigma(u^t) \neq \varnothing$ *is countable.*

1.4. *On the proofs of the main results.* Let us now explain the main ideas in the proof of Theorem 1.1, the general outline of the argument being inspired by [Figalli et al. 2020]. As pointed out, the key feature is the Taylor expansion (1-4). Furthermore, we can work in the top-dimensional stratum Σ_{n-1}, as the lower strata Σ_m, $m \in \{1, \ldots, n-2\}$, have Hausdorff dimension at most m thanks to Caffarelli's covering result.

We will perform a blow-up analysis on the functions $u - \mathcal{P}_k$, where the $\mathcal{P}_k$ are suitable polynomials. The core of this blow-up analysis is an Almgren-type monotonicity formula for $u - \mathcal{P}_k$.

1.4.1. *Polynomial ansatz.* Similarly to [Figalli et al. 2023], we construct $\mathcal{P}_k$, the prototypical k-th Taylor polynomial of u at $0 \in \Sigma_{n-1}$. These polynomials should be approximate solutions of (1-1), that is to say

$$\begin{cases} \Delta \mathcal{P}_k = f + O(|x|^{k-1}) & \text{in } B_1, \\ \mathcal{P}_k \geq -O(|x|^{k+2}) & \text{in } B_1. \end{cases}$$

Furthermore, they should have an $(n-1)$-dimensional zero set (we are in the top-dimensional stratum Σ_{n-1}). Together with nonnegativity, this suggests that $\mathcal{P}_k$ is almost a square:

$$\mathcal{P}_k = (\text{polynomial})^2 + O(|x|^{k+2}).$$

The coefficients of $\mathcal{P}_k$ can be chosen with some freedom, which can be used to modulate the shape of the hypersurface $\{\mathcal{P}_k = 0\}$ around 0. Notice that, by Caffarelli's analysis, $\mathcal{P}_0 = \mathcal{P}_1 = 0$ and (in suitable coordinates) $\mathcal{P}_2 = \frac{1}{2}x_n^2$. For example, we will see that $\mathcal{P}_3$ needs to have a more complicated form:

$$\mathcal{P}_3 = \left(x_n + \frac{p_3}{x_n} + x_n R_2\right)^2 + O(|x|^{k+2}),$$

where p_3 is any 3-homogeneous harmonic polynomial odd in x_n and $R_2 = R_2[p_3]$ is a 2-homogeneous polynomial determined uniquely by p_3.

1.4.2. *Almgren monotonicity formula.* In [Figalli and Serra 2019], it was first noticed that the Almgren frequency function of $u - p_2$, that is

$$r \mapsto \frac{\|r\nabla(u-p_2)(r\,\cdot)\|_{L^2(B_1)}}{\|(u-p_2)(r\,\cdot)\|_{L^2(\partial B_1)}} =: \phi(r, u-p_2), \tag{1-7}$$

is increasing. This property is known to be very powerful and paved the way for most of the results of [Figalli and Serra 2019]. Similarly, the key observation that allows us to prove expansion (1-4) is that, for all $k \geq 2$, (a version of) the Almgren frequency function is (almost) monotone on functions of the form $u - \mathcal{P}_k$. Proving this crucial fact for all k is one of the main contributions of this paper. The case $k = 3$ was already obtained in [Figalli et al. 2020],[2] but their approach was blocked and the general case cannot be obtained by tuning their argument; let us explain why.

By known computations, for every function w,

$$\frac{d}{dr}\phi(r, w) \geq \frac{2}{r}\frac{\int_{B_1}(\phi(r,w)w_r - x\cdot\nabla w_r)\Delta w_r}{\|w_r\|^2_{L^2(\partial B_1)}}. \tag{1-8}$$

In [Figalli and Serra 2019], it was shown that the right-hand side is nonnegative if $w = u - \mathcal{P}_2$, but in order to carry out our arguments it is enough to prove that the negative part of the right-hand side is integrable around $r = 0$ when $w = u - \mathcal{P}_k$. In trying to do so, the term $w_r \Delta w_r$ has (up to some errors) the right sign, while the main difficulties come from the term $(x \cdot \nabla w_r)\Delta w_r$.

In order to control this last term in the case $k = 3$, the following crucial Lipschitz estimate was proved in [Figalli et al. 2020, Lemma 4.7]:

$$r\|\nabla(u-\mathcal{P}_3)\|_{L^\infty(B_r)} \leq C(\|(u-\mathcal{P}_3)(r\,\cdot)\|_{L^2(B_2)} + r^5) \tag{1-9}$$

for some constant C independent of u, $\mathcal{P}_3$, and $r \in \left(0, \frac{1}{4}\right)$. In order to prove (1-9), one needs to take incremental quotients of $u - \mathcal{P}_3$ along the flow of some circular vector fields $\{X_j\}$, which have the property that $X_j X_j \mathcal{P}_3 = O(|x|^3)$ (see [Figalli et al. 2020, Lemma 4.6]). Now, exploiting that the Laplacian is invariant under rotations (that such vector fields generate), one can prove (1-9). There is little hope to make this ingenious argument work for a general $u - \mathcal{P}_k$, since the Laplacian does not commute with more general diffeomorphisms, which, on the other hand, would be needed to ensure $X_j X_j \mathcal{P}_k = O(|x|^k)$.

[2] Actually, as a consequence of the method of proof, the authors also immediately obtain the same monotonicity for $u - \mathcal{P}_3 - P$ for some homogeneous harmonic polynomial of order 4. This then leads the expansion up to order $5 - \varepsilon$, but to continue on what is missing is Almgren's frequency monotonicity for $u - \mathcal{P}_4$, where $\mathcal{P}_4 \neq \mathcal{P}_3 + P$.

With a completely different proof, in Section 3.2, we will prove the following weaker Lipschitz estimate, which — nontrivially — is still enough to get the integrability of the right-hand side of (1-8). For all $k \geq 2$, we have

$$r\|\nabla(u-\mathcal{P}_k)\|_{L^\infty(B_r)} \leq C(\|(u-\mathcal{P}_k)(r\,\cdot\,)\|_{L^2(B_\theta)} + r^{k+2})^{1-\beta}, \tag{1-10}$$

where $\beta > 0$ can be chosen arbitrarily small, $C, \theta > 1$ are constants independent of u, and $r \in (0, 1/\theta)$. This allows us to prove the monotonicity of a suitable modification of Almgren's frequency function introduced in [Figalli et al. 2020]. More precisely, the function

$$r \mapsto \phi^\gamma(r, u-\mathcal{P}_k) + Cr^\varepsilon, \quad \text{where } \phi^\gamma(r, w) := \frac{\|\nabla w_r\|^2_{L^2(B_1)} + \gamma r^{2\gamma}}{\|w_r\|^2_{L^2(\partial B_1)} + r^{2\gamma}}, \tag{1-11}$$

is increasing. Here $w_r := w(r\,\cdot\,)$, C, ε are positive constants and $\gamma > k+1$ is the truncation parameter.

1.4.3. *Blow-up analysis.* The monotonicity formula allows us to pursue a blow-up analysis to every order. Similarly to [Figalli and Serra 2019], we classify the possible blowups, that is, we study the possible limits of the normalized sequence

$$\tilde{v}_{r_\ell} := \frac{v_{r_\ell}}{\|v_\ell\|^2_{L^2(\partial B_1)}} \quad \text{as } r_\ell \downarrow 0, \quad \text{where } v_r := (u-\mathcal{P}_k)(r\,\cdot\,).$$

We show that $\tilde{v}_{r_\ell} \to q$, where q is a nontrivial global solution of a certain PDE: the Signorini problem. Furthermore, q is homogeneous of degree $\phi^\gamma(0^+, v)$. The blowup can be performed at each point of Σ_{n-1} of course, but q could be a nonpolynomial function or even have nonintegral homogeneity. At the points where this happens, there cannot be a Taylor polynomial of order $k+1$, so the expansion (1-4) must stop. These points must be shown to be "rare".

1.4.4. *Dimension reduction.* We show that $\phi^\gamma(0^+, u-\mathcal{P}_k) = k+1$ outside of a set of dimension at most $n-2$ for a suitable choice of $\mathcal{P}_k$, and the blowup q is a harmonic polynomial vanishing on $\{p_2 = 0\}$ (a particular class of solutions of the Signorini problem). This allows us to determine the next ansatz $\mathcal{P}_{k+1}$ in terms of $\mathcal{P}_k$ and q and prove the Taylor expansion up to order $k+1$ with remainder $o(r^{k+1})$. The various dimension reduction techniques are inspired by [Figalli and Serra 2019], but new barrier arguments are introduced to deal with the points with even frequency.

1.5. ***Structure of the paper.*** In Section 2 we fix the notation and recall some basic results on the obstacle problem and the Signorini problem. In Section 3.1 we give the construction of the polynomial ansatz $\mathcal{P}_k$. In Section 3.2 we prove our Lipschitz estimate (1-10). In Section 4 we prove the almost-monotonicity of the truncated frequency $\phi^\gamma(\,\cdot\,, u-\mathcal{P}_k)$. In Section 5 we perform and classify the blowups. In Section 6 we perform the dimension reductions distinguishing various cases; the proof of Theorem 1.1 is given in Section 6.4. In Section 7 we give the proof of the instability result Theorem 1.2.

In Appendix A we reprove a result [Figalli and Serra 2019, Remark 2.14] for a general right-hand side. In Appendix B we explain, line by line, which modifications are needed for a smooth right-hand side in the previous proofs. Finally, Appendix C contains two technical lemmas.

2. Preliminaries

2.1. ***Notation.*** We work in $\mathbb{R}^n$ endowed with its Euclidean structure and assume $n \geq 2$. We will often perform blowups: given a function $v : B_1 \to \mathbb{R}$, we set $v_r := v(r\,\cdot)$ which is defined in $B_{1/r}$; the parameter $r > 0$ is thought to be small. We remark that $\nabla v_r = r(\nabla v)_r$. We will sometimes write $X \lesssim_{a,b} Y$, meaning that $X \leq CY$ for some constant $C > 0$ which depends only on a and b.

2.2. ***Known results.*** Fix $\mu > 0$ and a function $f \in C^\infty(B_1)$ such that $f \geq \mu$ in B_1. We will denote by u any solution of

$$\begin{cases} \Delta u = f\chi_{\{u>0\}} & \text{in } B_1, \\ u \geq 0 & \text{in } B_1, \\ 0 \in \partial\{u > 0\}. \end{cases} \tag{2-1}$$

The last condition is added for normalization purposes, as we want to stay away from ∂B_1. We recall some basic properties of the solution u, relying on the classical theory by Caffarelli [1977; 1998], see also [Figalli et al. 2020, Section 3] for a summary of the known results. There exists

$$C = C(n, \mu, \|f\|_{L^\infty(B_1)}) > 0$$

such that

$$\|u\|_{C^{1,1}(B_{1/2})} \leq C \quad \text{and} \quad \sup_{B_{1/2}} u \geq \frac{1}{C}. \tag{2-2}$$

Thus we will assume throughout the paper that $u \in C^{1,1}_{\mathrm{loc}}(B_1)$. We remark that the problem has a natural scaling; in fact, for any $r > 0$, we have that $r^{-2}u_r$ solves

$$\begin{cases} \Delta(r^{-2}u_r) = f_r\chi_{\{u_r>0\}} & \text{in } B_{1/r}, \\ u_r \geq 0 & \text{in } B_{1/r}, \\ 0 \in \partial\{u_r > 0\}. \end{cases}$$

The free boundary $\partial\{u > 0\}$ consists of regular points (in the neighborhood of which $\partial\{u > 0\}$ is an analytic hypersurface) and singular points $\Sigma \subseteq \partial\{u > 0\}$ (at which the volume density of $\{u = 0\}$ is 0). It is well known that the singular points are characterized by the condition that the blowup

$$p_{2,x_\circ}(x) := \lim_{r\to 0} r^{-2}u(x_\circ + rx)$$

exists and is a convex 2-homogeneous polynomial with $\Delta p_{2,x_\circ} \equiv f(x_\circ)$. When $x_\circ = 0$, we denote the blowup simply by p_2. The singular set Σ stratifies according to $\dim\{p_{2,x_\circ} = 0\}$. The strata

$$\Sigma_m := \{x_\circ \in \Sigma : \dim\{p_{2,x_\circ} = 0\} = m\} \quad \text{for } m = 0, \dots, n-1$$

are locally contained in m-dimensional C^1 manifolds. As we want to prove a statement "up to sets of codimension 2", we will be mostly interested in the top-dimensional stratum Σ_{n-1}.

The following lemma is crucial for our analysis. It shows that, in Σ_{n-1}, the rate of convergence of u to its blowup is more than quadratic. It was proved in [Figalli and Serra 2019, Remark 2.14] for $f \equiv 1$; for completeness we give the proof for a general $f \in C^\delta(B_1)$ in Appendix A.

Lemma 2.1. *Assume that* $0 \in \Sigma_{n-1}$ *and* $r^{-2}u_r \to p_2$. *Then there are* $C, \alpha_\circ > 0$ *such that*

$$\sup_{B_r} |u - p_2| \le C r^{2+2\alpha_\circ} \quad \textit{for all } r \in \left(0, \tfrac{1}{2}\right). \tag{2-3}$$

In particular, we have

$$\{u_r = 0\} \cap B_1 \subseteq \{x : \operatorname{dist}(x, \{p_2 = 0\}) \le C r^{\alpha_\circ}\} \quad \textit{for all } r \in (0, 1). \tag{2-4}$$

The constants C *and* $\alpha_\circ$ *depend only on* n, μ, δ, $\|f\|_{C^\delta(B_1)}$, *where* $0 < \delta \le 1$ *can be freely chosen.*

Notice that (2-4) immediately follows from (2-3) because $\operatorname{dist}(\,\cdot\,, \{p_2 = 0\})^2 = p_2$ since $0 \in \Sigma_{n-1}$.

2.3. ***Truncated frequency function.*** We will make extensive use of the following functionals. For $w \in C^{1,1}_{\text{loc}}(B_1)$, $r \in (0, 1)$ and a parameter $\gamma \ge 0$, let us define the nondimensional quantities

$$D(r, w) := r^{2-n} \int_{B_r} |\nabla w|^2 = \int_{B_1} |\nabla w_r|^2, \quad H(r, w) := r^{1-n} \int_{\partial B_r} w^2 = \int_{\partial B_1} w_r^2 \tag{2-5}$$

and the truncated frequency function

$$\phi^\gamma(r, w) := \frac{D(r, w) + \gamma r^{2\gamma}}{H(r, w) + r^{2\gamma}}, \tag{2-6}$$

which has been introduced in [Figalli et al. 2020]. By [Figalli et al. 2020, Lemma 2.3], the following formula is valid for all $w \in C^{1,1}_{\text{loc}}(B_1)$ and $r \in (0, 1)$:

$$\frac{d}{dr}\phi^\gamma(r, w) \ge \frac{2}{r} \frac{\left(r^{2-n} \int_{B_r} w \Delta w\right)^2 + E^\gamma(r, w)}{(H(r, w) + r^{2\gamma})^2},$$

where

$$E^\gamma(r, w) := \left(r^{2-n} \int_{B_r} w \Delta w\right)(D(r, w) + \gamma r^{2\gamma}) - \left(r^{2-n} \int_{B_r} (x \cdot \nabla w) \Delta w\right)(H(r, w) + r^{2\gamma}).$$

Thus we have

$$\frac{d}{dr}\phi^\gamma(r, w) \ge \frac{2}{r} \frac{\int_{B_1} (\phi^\gamma(r, w) w_r - x \cdot \nabla w_r) \Delta w_r}{H(r, w) + r^{2\gamma}}. \tag{2-7}$$

We recall from [Figalli et al. 2020] the following result which says, roughly speaking, that the value of $\phi^\gamma(\,\cdot\,, v)$ corresponds to the power at which $H(\,\cdot\,, v)$ grows. This lemma will be used extensively to pass from L^2 norms over spheres to L^2 norms over thick shells.

Lemma 2.2 [Figalli et al. 2020, Lemma 4.1, Remark 4.2]. *Let* $R \in (0, 1)$, *and let* $w : B_R \to [0, \infty)$ *be a* $C^{1,1}$ *function. Assume that, for some* $\varepsilon \in (0, 1)$ *and a constant* $C_\circ > 0$, *we have*

$$\frac{d}{dr}(\phi^\gamma(r, w) + C_\circ r^\varepsilon) \ge \frac{2}{r} \frac{\left(r^{2-n} \int_{B_r} w \Delta w\right)^2}{(H(r, w) + r^{2\gamma})^2} \quad \textit{for all } r \in (0, R).$$

Then the following hold:

(a) *Suppose that* $0 < \underline{\lambda} \le \phi^\gamma(r, w) \le \bar{\lambda}$ *for all* $r \in (0, R)$. *Then, for any given* $\delta > 0$, *we have*

$$\frac{1}{C_\delta}\left(\frac{R}{r}\right)^{2\underline{\lambda}-\delta} \le \frac{H(R, w) + R^{2\gamma}}{H(r, w) + r^{2\gamma}} \le C_\delta\left(\frac{R}{r}\right)^{2\bar{\lambda}+\delta} \quad \text{for all } r \in (0, R),$$

where C_δ *depends on* $n, \gamma, \varepsilon, \bar{\lambda}, C_\circ, \delta$.

(b) *Assume in addition that*

$$\frac{r^{2-n}\int_{B_r} w\Delta w}{H(r, w) + r^{2\gamma}} \ge -C_\circ r^\varepsilon \quad \text{for all } r \in (0, R).$$

Then, for $\lambda_* := \phi^\gamma(0^+, w)$, *we have* $\lambda_* \le \gamma$ *and*

$$e^{-C_\circ/\varepsilon^2}\left(\frac{R}{r}\right)^{2\lambda_*} \le \frac{H(R, w) + R^{2\gamma}}{H(r, w) + r^{2\gamma}}.$$

2.4. ***The Signorini problem.*** The Signorini problem, called also the thin obstacle problem, consists of the following system of PDEs

$$\begin{cases} \Delta q \le 0 \text{ and } q\Delta q = 0 & \text{in } \mathbb{R}^n, \\ \Delta q = 0 & \text{in } \mathbb{R}^n \setminus L, \\ q \ge 0 & \text{on } L, \end{cases} \tag{2-8}$$

where $L \subseteq \mathbb{R}^n$ is a hyperplane and q is at least continuous. Recall that the following regularity is known (see [Athanasopoulos and Caffarelli 2004]) for weak solutions: if $L = \{x_n = 0\}$ then $q \in C^{1,1/2}_{\text{loc}}(\{x_n \ge 0\})$.

For each solution q we will consider its *singular set*,[3] defined by

$$\Sigma(q) := \{x \in L : q = |\nabla q| = 0\}. \tag{2-9}$$

We will be interested in homogeneous solutions, so, for every $\lambda \ge 0$ and every hyperplane $L \subseteq \mathbb{R}^n$, let us define

$$\mathcal{S}_\lambda(L) := \{q \in W^{1,2}_{\text{loc}} \cap C^0_{\text{loc}}(\mathbb{R}^n) : q \text{ is } \lambda\text{-homogeneous and solves (2-8)}\}. \tag{2-10}$$

We will use the following characterization of homogeneous solutions.

Lemma 2.3. *Let* $q \in W^{1,2}_{\text{loc}} \cap C^0_{\text{loc}}(\mathbb{R}^n)$ *be a weak solution of* (2-8), *and let* $\lambda \ge 0$. *Then* q *is* λ*-homogeneous if and only if*

$$\frac{D(r, q)}{H(r, q)} = \lambda \quad \text{for all } r > 0.$$

Proof. Setting to zero the derivative with respect to r of the left-hand side, one formally gets

$$q(x) = \lambda x \cdot \nabla q(x).$$

One can make the computation rigorous using the $C^{1,\alpha}$ regularity of q; see [Fernández-Real 2022]. □

[3] We warn the reader that our singular set has nothing to do with the "singular points" of the free boundary of q. Our terminology is instead consistent with [Naber and Valtorta 2017].

Every $q \in \mathcal{S}(\nu^\perp)$, with ν a unit vector, can be split into its even and odd part with respect to L, namely

$$q^{\text{even}}(x) := \tfrac{1}{2}(q(x) + q(x - 2(x \cdot \nu)\nu)), \quad q^{\text{odd}}(x) := \tfrac{1}{2}(q(x) - q(x - 2(x \cdot \nu)\nu)), \tag{2-11}$$

so that $q = q^{\text{even}} + q^{\text{odd}}$. It is easy to show that q^{even} and q^{odd} solve (2-8) separately, thus it is natural to define

$$\mathcal{S}_\lambda^{\text{even}}(L) := \{q \in \mathcal{S}_\lambda(L) : q \text{ is even with respect to } L\},$$
$$\mathcal{S}_\lambda^{\text{odd}}(L) := \{q \in \mathcal{S}_\lambda(L) : q \text{ is odd with respect to } L\}.$$

When it is not relevant, we will drop the dependence on L. We gather information on these sets next.

Proposition 2.4. *For every $m \in \mathbb{N}$, the following hold*:

(i) *Every element of $\mathcal{S}_{2m+1}(L)$ vanishes on the obstacle L.*

(ii) *$\mathcal{S}_\lambda^{\text{odd}}(L)$ consists exactly of those λ-homogeneous harmonic polynomials that vanish on L, thus it's empty when $\lambda \notin \mathbb{N}$.*

(iii) *$\mathcal{S}_{2m}^{\text{even}}(L)$ consists exactly of those $2m$-homogeneous harmonic polynomials that are nonnegative on L.*

(iv) *If $q \in \mathcal{S}_{2m+1}^{\text{even}}(e_n^\perp)$, then $q(x) = -|x_n|(q_0(x') + x_n^2 q_1(x))$, where q_0 and q_1 are polynomials such that $q_0 \geq 0$ and $\Delta(-x_n q_0(x') + x_n^3 q_1(x)) = 0$.*

(v) *The* (*real*) *values of λ for which $\mathcal{S}_\lambda^{\text{even}}(L)$ is not empty are known only in dimension $n = 2$, in which case we have $\lambda \in \mathbb{N} \cup \{2m + \frac{3}{2} : m \in \mathbb{N}\}$.*

Proof. For (i) see [Figalli et al. 2020, Lemma 5.1]. To show (ii), notice that q is harmonic in a half-space and coincides with its odd reflection, thus it is harmonic everywhere. The third point is proven in [Garofalo and Petrosyan 2009, Lemma 1.3.4]. For (iv) see [Figalli et al. 2020, Appendix B]. The last point follows by separating variables and explicitly solving the resulting ODE. □

Remark 2.5. Using Proposition 2.4, it is easy to check that if $n = 2$ and $\lambda \in \mathbb{N}$, $\lambda \geq 2$, then $\Sigma(q^{\text{even}}) = \{0\}$.

3. Lipschitz estimates

For the sake of readability, we deal first with the case $f \equiv 1$ and $\mu = 1$. A list of notational changes needed to address a general f is given in Appendix B. This section is devoted to the derivation of the Lipschitz estimate (1-10), which contains the most original part of this work.

3.1. *Polynomial ansatz.* We denote by V_j the vector space of homogeneous polynomials of degree $j \geq 1$. We introduce the projection map $\pi_j : \mathbb{R}[x] \to V_j$, which sends a polynomial to its j-homogeneous part, and the map $\pi_{\leq j}$, which truncates it at degree j. We define for $k \geq 2$ the set $\boldsymbol{P}_k \subseteq V_2 \times \cdots \times V_k$ by saying that $(p_2, \ldots, p_k) \in \boldsymbol{P}_k$ if and only if

(i) p_j is a j-homogeneous polynomial for each $2 \leq j \leq k$,

(ii) $p_2 \geq 0$, $\Delta p_2 = 1$ and $\dim\{p_2 = 0\} = n - 1$,

(iii) $\Delta p_j = 0$ and p_j vanish on $\{p_2 = 0\}$ for each $3 \leq j \leq k$.

Notice that any p_2 satisfying (ii) must be of the form $p_2(x) = \frac{1}{2}(\nu\cdot x)^2$ for some unit vector ν; of course ν is not unique as we can always choose $-\nu$, but that is the only freedom we have. Furthermore, for every $(p_2, \dots, p_k) \in V_2 \times \cdots \times V_k$, we set

$$|(p_2, \dots, p_k)| := \sum_{2\le j\le k} \|p_j\|_{L^2(\partial B_1)}.$$

Lemma 3.1. *Let $k \ge 2$ and $(p_2, \dots, p_k) \in \boldsymbol{P}_k$ be given. Then there exists a unique collection of polynomials*

$$(R_1, \dots, R_{k-1}) \in V_1 \times \cdots \times V_{k-1}$$

such that the following holds. If $p_2(x) = \frac{1}{2}(\nu\cdot x)^2$ for some unit vector ν and we define

$$\mathcal{A}_{k,\nu}(x) := (\nu\cdot x) + \sum_{j=1}^{k-1}(\nu\cdot x)R_j(x) + \sum_{j=3}^{k}\frac{p_j(x)}{(\nu\cdot x)},$$

then $\Delta\big(\frac{1}{2}\mathcal{A}_{k,\nu}^2\big) = 1 + O(|x|^k)$. Furthermore, each R_j is determined only by $(p_2, \dots, p_{j+1})$ and does not depend on which ν we choose, so that $\mathcal{A}_{k,-\nu} = -\mathcal{A}_{k,\nu}$. Finally, $\frac{1}{2}\mathcal{A}_{2,\nu}^2 = p_2$.

Proof. We prove the full statement by induction on k, beginning with $k = 2$. We compute

$$\Delta\big(\tfrac{1}{2}\mathcal{A}_{2,\nu}^2\big) = 1 + \Delta(2p_2R_1) + O(|x|^2),$$

thus R_1 must solve $\Delta(p_2R_1) = 0$; this is true if and only if $R_1 = 0$, as we can see with the following general argument. For $m \ge 1$, we consider the linear map $\delta_m : V_m \to V_m$ given by $\delta_m(q) := \Delta(p_2q)$. We claim that δ_m is a isomorphism. Indeed, if $\delta_m(q) = 0$, then the polynomial p_2q is a harmonic function that vanishes on the hyperplane $\{p_2 = 0\}$ along with its normal derivative, thus $p_2q \equiv 0$ (by reflection, p_2q is both even and odd with respect to $\{p_2 = 0\}$, thus $q \equiv 0$). As V_m has finite dimension, the map δ_m is invertible. In particular, $R_1 = \delta_1^{-1}(0) = 0$, regardless of p_2 and ν.

Assume that the full statement is proved up to some $k \ge 2$. Fix $(p_2, \dots, p_{k+1}) \in \boldsymbol{P}_{k+1}$ and ν, and for simplicity set $x_\nu := (\nu\cdot x)$. Notice that, for every $(R_1, \dots, R_k)$, we have $\mathcal{A}_{k+1,\nu} = \mathcal{A}_{k,\nu} + x_\nu R_k + p_{k+1}/x_\nu$. A direct computation again gives

$$\Delta\big(\tfrac{1}{2}\mathcal{A}_{k+1,\nu}^2\big) = \pi_{\le k-1}\big(\Delta\big(\tfrac{1}{2}\mathcal{A}_{k,\nu}^2\big)\big) + \pi_k\big(\Delta\big(\tfrac{1}{2}\mathcal{A}_{k,\nu}^2\big)\big) + \Delta p_{k+1} + \Delta\frac{p_3p_{k+1}}{2p_2} + \Delta(2p_2R_k) + O(|x|^{k+1}).$$

Hence we have $\Delta\frac{1}{2}\mathcal{A}_{k+1,\nu}^2 = 1 + O(|x|^{k+1})$ if and only if

$$\begin{cases}\Delta\frac{1}{2}\mathcal{A}_{k,\nu}^2 = 1 + O(|x|^k),\\ \pi_k\big(\Delta\frac{1}{2}\mathcal{A}_{k,\nu}^2\big) + \Delta p_3p_{k+1}/(2p_2) + \Delta 2p_2R_k = 0,\end{cases}$$

by inductive assumption the first equation determines uniquely $(R_1, \dots, R_{k-1})$ and thus $\mathcal{A}_{k,\nu}$. The second equation then determines uniquely R_k, indeed, as in the base step, we set

$$R_k := -\tfrac{1}{2}\delta_k^{-1}\bigg(\pi_k\Delta\big(\tfrac{1}{2}\mathcal{A}_{k,\nu}^2\big) + \Delta\frac{p_3p_{k+1}}{2p_2}\bigg).$$

Finally, by inductive assumption $\frac{1}{2}\mathcal{A}_{k,\nu}^2 = \frac{1}{2}\mathcal{A}_{k,-\nu}^2$, thus it is manifest that R_k does not depend on the choice of ν. □

This lemma shows that we can construct a function $\frac{1}{2}\mathcal{A}_k^2 : \boldsymbol{P}_k \to \mathbb{R}[x]$ defined by

$$\mathcal{A}_k^2 : (p_2, \dots, p_k) \mapsto \mathcal{A}_{k,\nu}^2, \quad \text{where } p_2(x) = (\nu \cdot x)^2. \tag{3-1}$$

Definition 3.2. Given $k \geq 2$, we define $\mathcal{P}_k : \boldsymbol{P}_k \to \mathbb{R}[x]$ by

$$\mathcal{P}_k(p_2, \dots, p_k) := \pi_{\leq k+1}\left(\tfrac{1}{2}\mathcal{A}_k^2\right), \tag{3-2}$$

where $\mathcal{A}_k^2 = \mathcal{A}_{k,\pm\nu}^2$ are constructed from $(p_2, \dots, p_k)$ as in Lemma 3.1.

We now give some simple properties of $\frac{1}{2}\mathcal{A}_k^2$ and $\mathcal{P}_k$; given a unit vector e, we write $\partial_e u = e \cdot \nabla u$.

Proposition 3.3. *Let $k \geq 2$, $(p_2, \dots, p_k) \in \boldsymbol{P}_k$ and $\tau > 0$ be such that $|(p_2, \dots, p_k)| \leq \tau$. Choose some unit vector ν for which $p_2(x) = \frac{1}{2}(\nu \cdot x)^2$. Then the polynomials $\frac{1}{2}\mathcal{A}_k^2(p_2, \dots, p_k)$ and $\mathcal{P}_k(p_2, \dots, p_k)$ satisfy*:

(i) $\Delta \mathcal{P}_k = 1$ *and* $\partial_e\left(\frac{1}{2}\mathcal{A}_k^2\right) = \partial_e \mathcal{P}_k + O(|x|^{k+1})$, *for any unit vector e.*

(ii) *We have* $\mathcal{P}_k(p_2, \dots, p_k) = \mathcal{P}_{k-1}(p_2, \dots, p_{k-1}) + p_k + O(|x|^{k+1})$.

(iii) *For all $|x| \leq r_0$, we have $\frac{1}{2} \leq |\partial_\nu \mathcal{A}_{k,\nu}(x)| \leq 2$, and thus*

$$\tfrac{1}{2}|\mathcal{A}_k(x)| \leq \left|\partial_\nu\left(\tfrac{1}{2}\mathcal{A}_k^2(x)\right)\right| \leq 2|\mathcal{A}_k(x)|,$$

where $r_0 = r_0(n, k, \tau) \in (0, 1)$.

(iv) *If u is a solution as in (2-1), $0 \in \Sigma_{n-1}$ and $r^{-2}u(r\,\cdot) \to p_2$, then by (2-3) we have, for all $0 < r < \frac{1}{2}$,*

$$\sup_{B_r \cap \{u=0\}} |\partial_\nu \mathcal{P}_k| \leq C r^{1+\alpha_\circ}$$

for some constant $C = C(n, k, \tau) > 0$.

Proof. Point (i) follows immediately from $\Delta\left(\frac{1}{2}\mathcal{A}_k^2\right) = 1 + O(|x|^k)$ and the fact that $\mathcal{P}_k - \frac{1}{2}\mathcal{A}_k^2 = O(|x|^{k+2})$. The second point can be easily checked by direct computation using the structure of the polynomial $\mathcal{A}_{k,\nu}$. For (iii), we compute $\partial_\nu \mathcal{A}_{k,\nu} = \partial_\nu(\nu \cdot x + O(|x|^2)) = 1 + O(|x|)$. Lastly, for the fourth point, note that by construction $\mathcal{P}_k = p_2 + O(|x|^3)$ and apply Lemma 2.1. □

3.2. *Regularity estimates near the free boundary.* We now turn to the Lipschitz estimate on functions of the form $u - \mathcal{P}_k$.

Proposition 3.4. *Let u be a solution of the obstacle problem (2-1) with $f \equiv 1$, $\mu = 1$, and suppose $0 \in \Sigma_{n-1}$. Let $k \geq 2$ be an integer, $\tau > 0$ and $\beta \in (0, \alpha_\circ^2/(k+2))$. Let $(p_2, \dots, p_k) \in \boldsymbol{P}_k$ be such that $|(p_2, \dots, p_k)| \leq \tau$. Suppose that*

$$r^{-2}u(r\,\cdot) \to p_2 = \tfrac{1}{2}x_n^2,$$

and set

$$v := u - \mathcal{P}_k,$$

where $\mathcal{P}_k = \mathcal{P}_k(p_2, \dots, p_k)$ is the polynomial ansatz from Definition 3.2. Then:

(i) *there are* $C > 0$ *and* $r_0 \in (0, 1)$, *depending on* $n,\ k,\ \tau$, *such that, for each* $j = 1, \dots, n-1$ *and* $0 < r < r_0$, *we have*

$$\|\partial_j v_r\|_{L^\infty(B_1)} \le C(\|v_{\theta r}\|_{L^2(B_2 \setminus B_{1/2})} + r^{k+2}), \tag{3-3}$$

where $\theta = \theta(n, k) > 1$,

(ii) *there are* $C > 0$ *and* $r_0 \in (0, 1)$, *depending on* $n,\ k,\ \tau,\ \beta$, *such that, for every* $0 < r < r_0$, *we have*

$$\|\partial_n v_r\|_{L^\infty(B_1)} \le C(\|v_{\theta r}\|_{L^2(B_2 \setminus B_{1/2})} + r^{k+2})^{1-\beta}, \tag{3-4}$$

where $\theta = \theta(n, k, \beta) > 1$.

As this result will be crucial let us explain briefly its proof. As $p_2 = \frac{1}{2}x_n^2$, we split $\mathbb{R}^n$ into the "tangential" directions $e_1, \dots, e_{n-1}$ and the "normal" direction e_n. The idea is to study the quantity

$$\sup_{B_r \cap \{u=0\}} |\partial_n \mathcal{P}_k|$$

for small r.

Geometrically this quantity tells us how much the zero set of $\mathcal{P}_k$ is a good approximation of the contact set $\{u = 0\}$ around the origin; see Proposition 3.3.

Analytically, $\sup_{B_r \cap \{u=0\}} |\partial_n \mathcal{P}_k|$ is crucial because it will be the "pivot" linking Lipschitz estimates along "tangential" directions with the one along the "normal" direction; let us see how.

First, taking difference quotients along tangential directions, we will prove that when $j \ne n$, we have,

$$r\|\partial_j v\|_{L^\infty(B_r)} \lesssim_{n,k,\tau} r^2 \cdot \sup_{\{u=0\} \cap B_r} |\partial_n \mathcal{P}_k| + \|v_r\|_{L^2(B_2 \setminus B_{1/2})} + r^{k+2}.$$

The fact that we have r^2 (and not r!) in front of $\sup_{\{u=0\} \cap B_r} |\partial_n \mathcal{P}_k|$ is the crucial gain, peculiar to the tangential derivatives.

Now we have to bound $\sup_{\{u=0\} \cap B_r} |\partial_n \mathcal{P}_k|$ from above. As this is much more complex, we just give the heuristics. First, notice that, as u is C^1, we have $\partial_n \mathcal{P}_k \equiv \partial_n v$ in $\{u = 0\}$. As v is harmonic in $\Omega := B_r \cap \{u > 0\}$, elliptic regularity suggests that we should be able to control $\|\partial_n v\|_{C^{0,\beta}(\Omega)}$ with $\|v\|_{C^{1,\beta}(\partial\Omega)} + \|v\|_{L^\infty(\Omega)}$. Furthermore, by Lemma 2.1, $\partial\Omega$ is close to the hyperplane $\{x_n = 0\}$, so we expect that the main contribution in $\|v\|_{C^{1,\beta}(\partial\Omega)}$ comes from the tangential derivatives $\|\partial_j v\|_{C^{0,\beta}(\partial\Omega)}$ for $j \ne n$. If we could take $\beta = 0$ and knew that $\partial\Omega$ was regular enough, this argument would give a bound on $\sup_{B_r} |\partial_n \mathcal{P}_k|$ in terms of $\|\partial_j v\|_{L^\infty(B_r)}$. But this is too much to ask: Schauder estimates break down at the Lipschitz scale and $\partial\Omega$ could be wild, this is why we lose a power β on the right-hand side of (3-4). The first issue will be fixed choosing β small and interpolating, the second will require us to construct a different set $\Omega \subseteq \{u > 0\}$ through a geometric barrier argument.

We start with a preliminary L^2-L^∞ estimate.

Lemma 3.5. *In the setting of Proposition 3.4, there exists a constant* $C = C(n, k, \tau)$ *such that*

$$\|v_r\|_{L^\infty(B_1)} \le C\|v_r\|_{L^2(B_2 \setminus B_{1/2})} + Cr^{k+2} \tag{3-5}$$

for all $0 < r < \frac{1}{2}$.

Proof. Recalling $\Delta\mathcal{P}_k = 1$, we have $\Delta v = \Delta(u - \mathcal{P}_k) = -\chi_{\{u=0\}} \le 0$. Using the mean-value inequality for superharmonic functions, for some $z \in \partial B_r$, we have

$$\min_{\bar{B}_r} v = v(z) \ge \fint_{B_{1/2}(z)} v_r \gtrsim_n -\|v_r\|_{L^2(B_2\setminus B_{1/2})}.$$

This provides the estimate from below. Inside $\{u=0\}\cap B_r$, we have

$$v = -\tfrac{1}{2}\mathcal{A}_k^2 + O(|x|^{k+2}) \le Cr^{k+2}$$

for some $C = C(n,k,\tau)$. Then we "glue" the functions Cr^{k+2} and v, which is harmonic in $B_r\cap\{u>0\}$, so that

$$V := \max\{Cr^{k+2}, v\} \quad \text{is subharmonic in } B_r.$$

Thus to estimate from above v on $B_r\cap\{u>0\}$, we just use the mean-value property on V as above: this gives the upper bound up to corrections of size Cr^{k+2}. □

With a similar technique, we bound the first derivatives of v with $\sup_{\{u=0\}\cap B_r}|\partial_n\mathcal{P}_k|$.

Lemma 3.6. *In the setting of Proposition 3.4, there exists a constant $C = C(n,k,\tau)$, such that, for each $j = 1,\dots,n-1$ and $0 < r < \frac{1}{2}$, we have*

$$\|\partial_j v_r\|_{L^\infty(B_1)} \le Cr\cdot \sup_{\{u=0\}\cap B_r}|r\partial_n\mathcal{P}_k| + C\|v_r\|_{L^2(B_2\setminus B_{1/2})} + Cr^{k+2},$$
$$\|\partial_n v_r\|_{L^\infty(B_1)} \le C \sup_{\{u=0\}\cap B_r}|r\partial_n\mathcal{P}_k| + C\|v_r\|_{L^2(B_2\setminus B_{1/2})} + Cr^{k+2}.$$

We remark that $\|\partial_\ell v_r\|_{L^\infty(B_1)} = r\|\partial_\ell v\|_{L^\infty(B_r)}$ for all ℓ and $r>0$.

Proof. We address first the case $j \ne n$. By construction, we have $\Delta\mathcal{P}_k = 1$, so $\Delta v = -\chi_{\{u=0\}}$. Hence, $\partial_j v$ is harmonic in $B_r\cap\{u>0\}$. On the other hand, in $B_r\cap\{u=0\}$, we have $\partial_j v = -\partial_j\mathcal{P}_k$, so by Proposition 3.3

$$|\partial_j\mathcal{P}_k| \le |\partial_j\mathcal{A}_k||\mathcal{A}_k| + O(|x|^{k+1}) \lesssim_{n,k,\tau} |x||\partial_n\mathcal{P}_k(x)| + |x|^{k+1}.$$

Here we crucially used that $|\partial_j\mathcal{A}_k| \lesssim_{n,k,\tau} |x|$ as $\mathcal{A}_k(x) = x_n + O(x^2)$. Hence,

$$\sup_{B_r\cap\{u=0\}}|\partial_j v| \le Cr^{k+1} + Cr\cdot\sup_{B_r\cap\{u=0\}}|\partial_n\mathcal{P}_k| =: K$$

for some $C = C(n,k,\tau)$. In order to estimate $\partial_j v$ on $B_r\cap\{u>0\}$, we truncate it at levels K and $-K$ to obtain that

$$f := \max\{K, \partial_j v\} \quad\text{and}\quad g := \min\{-K, \partial_j v\}$$

are, respectively, subharmonic and superharmonic in B_r. Choose $x \in \partial B_r$ a maximum point of f in $\bar{B}_r$ and use the mean-value property:

$$\sup_{B_r\cap\{u>0\}}\partial_j v \le \sup_{B_r} f = f(x) \le \fint_{B_{1/2}(x)} f_r \lesssim_n K + \frac{1}{r}\|\partial_j v_r\|_{L^1(B_{3/2}\setminus B_{1/2})},$$

where we used that $B_{r/2}(x) \subseteq B_{3r/2} \setminus B_{r/2}$. By standard elliptic estimates we find

$$\|\nabla v_r\|_{L^1(B_{3/2}\setminus B_{1/2})} \lesssim_n \|\Delta v_r\|_{L^1(B_2\setminus B_{1/2})} + \|v_r\|_{L^2(B_2\setminus B_{1/2})}.$$

Recalling $\Delta v_r = -r^2 \chi_{\{u_r=0\}} \le 0$, we integrate by parts with some smooth cut-off function

$$\chi_{B_2\setminus B_{1/2}} \le \psi \le \chi_{B_3\setminus B_{1/3}},$$

with $\|\psi\|_{C^2} \lesssim_n 1$, and we have

$$\int_{B_2\setminus B_{1/2}} |\Delta v_r| = -\int_{B_2\setminus B_{1/2}} \Delta v_r \le -\int_{B_3\setminus B_{1/3}} \Delta v_r \psi \lesssim_n \|v_r\|_{L^2(B_3\setminus B_{1/3})}.$$

The same computation with g instead of f provides an analogous estimate from below on $\partial_j v$. In conclusion we proved that, in every case, in B_r we have

$$|\partial_j v| \lesssim_n K + \frac{1}{r}\|v_r\|_{L^2(B_3\setminus B_{1/3})}.$$

Multiplying by r on both sides we find the first estimate.

The second estimate is proven with the same reasoning, just replacing K with

$$K' := Cr^{k+1} + C \cdot \sup_{B_r\cap\{u=0\}} |\partial_n \mathcal{P}_k|,$$

without the "extra r". □

We now use global Schauder estimates to bound from above the term $\sup_{\{u=0\}\cap B_r} |r\partial_n \mathcal{P}_k|$.

Lemma 3.7. *In the setting of Proposition 3.4, for any $\beta \in (0, \alpha_\circ^2/(k+2))$, there exists $C = C(n, k, \tau, \beta)$ such that, for all $0 < r < \frac{1}{2}$, we have*

$$\sup_{\{u=0\}\cap B_{r/4}} |r\partial_n \mathcal{P}_k| \le Cr^{\beta/(\alpha_\circ-\beta)}\|\partial_n v_r\|_{L^\infty(B_1)} + Cr^{\beta+\beta/\alpha_\circ}\|\nabla_{x'} v_r\|^{1-\beta/\alpha_\circ}_{L^\infty(B_1)} + C\|v_r\|_{L^\infty(B_1)} + Cr^{k+2}. \quad (3\text{-}6)$$

We recall that the dimensional constant $\alpha_\circ > 0$ has been defined in Lemma 2.1.

Proof. We will split the coordinates $x = (x', x_n)$ and denote by B'_r the intersection of B_r and $\{x_n = 0\}$. First of all, we recall from Lemma 2.1 that

$$\sup_{\{u=0\}\cap B_r} |\partial_n \mathcal{P}_k| = \sup_{\{u=0\}\cap B_r} |x_n + O(|x|^2)| \le C_\circ r^{1+\alpha_\circ} \quad \text{for all } 0 < r < 1 \quad (3\text{-}7)$$

for some $\alpha_\circ, C_\circ$ depending only on n and k. It is enough to prove the claim for $r \in (0, r_0)$, for some r_0 whose size will be constrained during the proof in terms of n, k, τ (recall that $\tau \ge |(p_2, \dots, p_3)|$). We will prove in detail only the upper bound, as the lower bound is derived in the same way, with a "symmetric" argument.

We choose a point $z \in \bar{B}_{r/4} \cap \{u = 0\}$ such that

$$\sup_{\{u=0\}\cap B_{r/4}} r\partial_n \mathcal{P}_k = r\partial_n \mathcal{P}_k(z).$$

Step 1. If r_0 is small enough, then, for all $r \in (0, r_0)$, we can find an open set Ω satisfying the following:

(i) Ω is a smooth domain inside B_{r_0}, that is $\Omega \cap B_{r_0} = \{x_n > \gamma(x')\} \cap B_{r_0}$ for some $\gamma \in C^\infty(B'_{r_0})$. Furthermore we have the following estimates on γ:

$$\|\gamma\|_{L^\infty(B'_r)} \le Cr^{1+\alpha_\circ}, \quad \|\nabla'\gamma\|_{L^\infty(B'_r)} \le Cr^{\alpha_\circ}, \quad [\nabla'\gamma]_{C^{\alpha_\circ}(B'_r)} \le C \tag{3-8}$$

for some $C = C(n, k, \tau)$.

(ii) $\Omega \subseteq \{u > 0\}$ and there exists $z^* \in \partial\Omega \cap \{u = 0\} \cap B_{r/2}$.

We emphasize that Ω, γ and z^* may depend on r, but the constant in (3-8) does not.

Proof of Step 1. We fix some $r \in (0, r_0)$. For all $b \in \mathbb{R}$ and $L = L(n, k, \tau)$ to be determined, we define the domains

$$\Omega(b) := \{x \in B_r : \partial_n \mathcal{P}_k(x) > \partial_n \mathcal{P}_k(z) + Lr^{\alpha_\circ - 1}|x' - z'|^2 + b\}.$$

Roughly speaking, $\Omega(b)$ looks like a perturbed paraboloid with vertex at $(z', z_n + b)$, provided $r_0 \lesssim_{n,k,\tau} 1$. Now, starting with b large, we decrease it until $\Omega(b)$ touches the contact set in B_r. That is, define

$$\Omega := \Omega(b^*), \quad \text{where } b^* := \inf\{b \in \mathbb{R} : \Omega(b) \cap B_r \subseteq \{u > 0\}\}.$$

We start checking that b^* is well defined. Thanks to (3-7), if there exists $x \in \{u = 0\} \cap B_r \cap \Omega(b)$ then

$$Lr^{\alpha_\circ - 1}|x' - z'|^2 + b < |\partial_n \mathcal{P}_k(x)| + |\partial_n \mathcal{P}_k(z)| \le 2C_\circ r^{1+\alpha_\circ}. \tag{3-9}$$

This shows that b^* is well defined and $b^* \le 2C_\circ r^{1+\alpha_\circ}$; in fact $\{u = 0\} \cap B_r \cap \Omega(b)$ must be empty for larger b. We also notice that $b \ge 0$, as $z \in \{u = 0\} \cap B_r \cap \Omega(b)$ for all $b < 0$.

We now prove (ii). Take $b < b^*$. By definition there exists $x_b \in \{u = 0\} \cap \Omega(b)$, and inequality (3-9) shows that $|x'_b - z'| \le \frac{1}{8}r$ for an appropriate choice of $L \gtrsim C_\circ$. Using the triangular inequality and (3-7) to estimate $|(x_b)_n - z_n|$, we find, for r_0 small, that

$$x_b \in \{u = 0\} \cap B_r \cap \Omega(b) \quad \Longrightarrow \quad x_b \in B_{r/2} \text{ and } b \le 2C_\circ r^{1+\alpha_\circ}. \tag{3-10}$$

We take $z^* \in \bar{B}_{r/2}$ to be any accumulation point of x_b as $b \uparrow b^*$. We clearly have $z^* \in \partial\Omega \cap \{u = 0\} \cap \bar{B}_{r/2}$.

Let us now prove (i). Consider the map $\Phi : x \mapsto (x', \partial_n \mathcal{P}_k(x))$. As $\partial_n \mathcal{P}_k(x) = x_n + O(x^2)$, we have that Φ is a diffeomorphism from B_{r_0}, provided r_0 is small. Let Ψ denote the n-th component of its inverse which is defined in some ball B_{R_0}. Clearly we have that $\Psi(y) = y_n + O(|y|^2)$, thus

$$|\nabla'\Psi(y)| \lesssim_{n,k,\tau} |y| \text{ and } \partial_n \Psi(y) \ge \tfrac{1}{2} \quad \text{for all } y \in B_{R_0}. \tag{3-11}$$

Therefore we conclude that $x \in \Omega$ if, and only if, $x \in B_r$ and

$$x_n = \Psi(x', \partial_n \mathcal{P}_k(x)) > \Psi(x', \partial_n \mathcal{P}_k(z) + Lr^{\alpha_\circ - 1}|x' - z'|^2 + b^*) := \gamma(x').$$

We have that $\gamma \in C^\infty(B'_{r_0})$ is well defined as, for r_0 small,

$$(x', \partial_n \mathcal{P}_k(z) + Lr^{\alpha_\circ - 1}|x' - z'|^2 + b^*) \in B_{R_0}.$$

This proves that Ω is a smooth domain. We remark that, while γ depends on r, Ψ only depends on r_0. This observation along with (3-11) easily gives the estimates on γ with constants independent of r. For example, we estimate, for all $x' \in B'_r$,

$$\begin{aligned}|\nabla'\gamma(x')| &\le |\nabla'\Psi(x', \partial_n\mathcal{P}_k(z) + Lr^{\alpha_\circ-1}|x'-z'|^2 + b^*)| \\ &\qquad + 2Lr^{\alpha_\circ-1}|x'-z'|\,|\partial_n\Psi(x', \partial_n\mathcal{P}_k(z) + Lr^{\alpha_\circ-1}|x'-z'|^2 + b^*)| \\ &\lesssim_{n,\tau} |x'| + |\partial_n\mathcal{P}_k(z) + Lr^{\alpha_\circ-1}|x'-z'|^2 + b^*| + Lr^{\alpha_\circ-1}\|\partial_n\Psi\|_{L^\infty(B_{R_0})}|x'-z'|,\end{aligned}$$

and all the terms on the right-hand side are of order $r^{\alpha_\circ}$ or higher thanks to (3-9). The other estimates can be proven identically, using (3-9) and the fact that $\Psi(0,0) = |\nabla'\Psi(0,0)| = 0$.

Step 2. For each $\beta \in (0, \alpha_\circ]$, there is $C = C(n,k,\tau,\beta)$ such that

$$\frac{1}{C}\sup_{\{u=0\}\cap B_{r/4}} r\partial_n\mathcal{P}_k \le [\nabla'(v\circ\Gamma)_r]_{C^\beta(B'_{3/4})} + \|v_r\|_{L^\infty(B_1)},$$

where $\Gamma : B'_r \to \mathbb{R}^n$ is the graph of γ, that is, $\Gamma(x') = (x', \gamma(x'))$.

Proof of Step 2. We observe that, by definition of z^*,

$$\partial_n\mathcal{P}_k(z) \le \partial_n\mathcal{P}_k(z^*),$$

so we have

$$\sup_{\{u=0\}\cap B_{r/4}} r\partial_n\mathcal{P}_k = r\partial_n\mathcal{P}_k(z) \le r\partial_n\mathcal{P}_k(z^*) \le \|\partial_n v_r\|_{L^\infty(\partial\widetilde{\Omega}\cap B_{1/2})},$$

where we used the rescaled domain $\widetilde{\Omega} := \Omega/r$, whose boundary is the graph of $\tilde{\gamma} = \gamma(r\,\cdot)/r$. We now employ global Schauder estimates in $\widetilde{\Omega}$ (see, e.g., [Gilbarg and Trudinger 1983, Theorem 8.33]) to control the right-hand side:

$$\|\partial_n v_r\|_{L^\infty(\partial\widetilde{\Omega}\cap B_{1/2})} \lesssim_{n,\beta} [\nabla'(v\circ\Gamma)_r]_{C^\beta(B'_{3/4})} + \|\Delta v_r\|_{L^\infty(\widetilde{\Omega})} + \|v_r\|_{L^\infty(B_1)}. \tag{3-12}$$

Note that $\widetilde{\Omega}$ lies in $\{u_r > 0\}$, where $\Delta v_r \equiv 0$. We remark that the previous estimate holds with a constant which depends on $\|\nabla\tilde{\gamma}\|_{C^{\alpha_\circ}(B_1)}$, which is bounded independently by r from (3-8):

$$\|\nabla\tilde{\gamma}\|_{L^\infty(B'_1)} = \|\nabla\gamma\|_{L^\infty(B'_r)} \le Cr^{\alpha_\circ} \quad \text{and} \quad [\nabla\tilde{\gamma}]_{C^{\alpha_\circ}(B'_1)} = r^{\alpha_\circ}\|\nabla\gamma\|_{C^{\alpha_\circ}(B'_r)} \le Cr^{\alpha_\circ}.$$

Furthermore, we also used that, thanks to (3-8), the graph of $\tilde{\gamma}$ is contained in the strip $\left\{|y_n| \le \frac{1}{10}\right\} \cap B_1$, provided that r_0 is small.

Step 3. There is $C = C(n,k,\tau)$ such that

$$\begin{aligned}\frac{1}{C}\|\nabla'(v\circ\Gamma)_r\|_{L^\infty(B'_{3/4})} &\le \|\nabla' v_r\|_{L^\infty(B_1)} + r^{\alpha_\circ}\|\partial_n v_r\|_{L^\infty(B_1)}, \\ \frac{1}{C}[\nabla'(v\circ\Gamma)_r]_{C^{\alpha_\circ}(B'_{3/4})} &\le r^{1+\alpha_\circ}.\end{aligned}$$

Proof of Step 3. Let $\widetilde{\Gamma}$ denote the graph of $\tilde{\gamma}$, that is, $\widetilde{\Gamma}(y') = (y', \tilde{\gamma}(y'))$. We have

$$\nabla'(v\circ\Gamma)_r = \nabla'(v_r\circ\widetilde{\Gamma}) = \nabla' v_r + (\partial_n v_r\circ\widetilde{\Gamma})\nabla'\tilde{\gamma},$$

so the first bound follows taking the $L^\infty(B'_{3/4})$ norms, using $\widetilde{\Gamma}(B'_{3/4}) \subseteq B'_1$ and $\|\nabla\tilde{\gamma}\|_{L^\infty(B'_1)} \le Cr^{\alpha_\circ}$. For the second bound we use the optimal regularity of u (see (2-2)):

$$\begin{aligned}[\nabla'(v\circ\Gamma)_r]_{C^{\alpha_\circ}(B'_{3/4})} &= r^{1+\alpha_\circ}[(\nabla' v\circ\Gamma)(\nabla'\Gamma)]_{C^{\alpha_\circ}(B'_{3r/4})}\\ &\le r^{1+\alpha_\circ}(\|\nabla v\|_{L^\infty(B_r)}[\nabla'\gamma]_{C^{\alpha_\circ}(B'_r)} + [\nabla v]_{C^{0,1}(B'_r)}[\Gamma]_{C^{\alpha_\circ}(B'_r)}\|\nabla'\Gamma\|_{L^\infty(B'_r)})\\ &\le r^{1+\alpha_\circ}\|\gamma\|_{C^{\alpha_\circ}(B'_r)}\|v\|_{C^{1,1}(B_{1/2})}\\ &\lesssim_{n,k,\tau} r^{1+\alpha_\circ}.\end{aligned}$$

Step 4. There is $C = C(n,k,\tau,\beta)$ such that

$$\frac{1}{C}[\nabla'(v\circ\Gamma)_r]_{C^\beta(B'_{3/4})} \le r^{\alpha_\circ^2/\beta} + r^{\beta/(\alpha_\circ-\beta)}\|\partial_n v_r\|_{L^\infty(B_1)} + r^{\beta+\beta/\alpha_\circ}\|\nabla' v_r\|^{1-\beta/\alpha_\circ}_{L^\infty(B_1)}. \tag{3-13}$$

Proof of Step 4. The idea is to bound the C^β norm in (3-12) with an interpolation of the L^∞ and the $C^{\alpha_\circ}$ norms, which we bounded in Step 3. This gives

$$\begin{aligned}[\nabla'(v\circ\Gamma)_r]_{C^\beta(B'_{3/4})} &\le [\nabla'(v\circ\Gamma)_r]^{\beta/\alpha_\circ}_{C^{\alpha_\circ}(B'_1)}[\nabla'(v\circ\Gamma)_r]^{1-\beta/\alpha_\circ}_{L^\infty(B'_1)}\\ &\le Cr^{\beta+\beta/\alpha_\circ}(\|\nabla' v_r\|_{L^\infty(B_1)} + r^{\alpha_\circ}\|\partial_n v_r\|_{L^\infty(B_1)})^{1-\beta/\alpha_\circ}\\ &\le Cr^{\alpha_\circ}(r^{\beta/(\alpha_\circ-\beta)}\|\partial_n v_r\|_{L^\infty(B_1)})^{1-\beta/\alpha_\circ} + Cr^{\beta+\beta/\alpha_\circ}\|\nabla' v_r\|^{1-\beta/\alpha_\circ}_{L^\infty(B_1)},\end{aligned}$$

where in the last line we used the subadditivity of $t \mapsto t^{1-\beta/\alpha_\circ}$. Finally we obtain (3-13) using Young's inequality on the first term, with exponent $1/p = \beta/\alpha_\circ$:

$$r^{\alpha_\circ}(r^{\beta/(\alpha_\circ-\beta)}\|\partial_n v_r\|_{L^\infty(B_1)})^{1-\beta/\alpha_\circ} \lesssim_{\alpha_\circ,\beta} r^{\alpha_\circ^2/\beta} + r^{\beta/(\alpha_\circ-\beta)}\|\partial_n v_r\|_{L^\infty(B_1)}.$$

Combining Steps 2 and 4 and recalling that $\alpha_\circ^2/\beta > k+2$, we obtain (3-6) for all $r \in (0, r_0)$, as the constants do not depend on r. □

With the help of the two previous lemmas, we can now prove our main Lipschitz estimate by using $\sup_{B_r\cap\{u=0\}}|\partial_n \mathcal{P}_k|$ as "pivot".

Proof of Proposition 3.4. We first address (ii), the estimate in the normal direction. All constants will depend on n, k, τ, β. Linking Lemma 3.7 with the second estimate of Lemma 3.6, we get

$$\frac{1}{C}\|\partial_n v_r\|_{L^\infty(B_{1/4})} \le r^{\beta/(\alpha_\circ-\beta)}\|\partial_n v_r\|_{L^\infty(B_1)} + r^{\beta+\beta/\alpha_\circ}\|\nabla_{x'} v_r\|^{1-\beta/\alpha_\circ}_{L^\infty(B_1)} + \Lambda(r),$$

where we set for brevity

$$\Lambda(r) := \|v\|_{L^\infty(B_r)} + r^{k+2}.$$

Notice that Λ is increasing in r. Now we use trivial bounds and the tangential estimate of Lemma 3.6 to deal with the central terms on the right-hand side:

$$r^{\beta+\beta/\alpha_\circ}\|\nabla_{x'} v_r\|^{1-\beta/\alpha_\circ}_{L^\infty(B_1)} \le Cr^{\beta+\beta/\alpha_\circ}(r\|\partial_n v_r\|_{L^\infty(B_1)} + \Lambda(2r))^{1-\beta/\alpha_\circ}.$$

Then, using Young's inequality with $1/p = \beta/\alpha_\circ$ and $r^{\alpha_\circ/\beta} \le r^{k+2}$, we obtain

$$r^{\beta+\beta/\alpha_\circ}(r\|\partial_n v_r\|_{L^\infty(B_1)})^{1-\beta/\alpha_\circ} = r(r^{\alpha_\circ\beta/(\alpha_\circ-\beta)}\|\partial_n v_r\|_{L^\infty(B_1)})^{1-\beta/\alpha_\circ} \lesssim_{\alpha_\circ,\beta} r^{\alpha_\circ\beta/(\alpha_\circ-\beta)}\|\partial_n v_r\|_{L^\infty(B_1)} + r^{k+2}.$$

Thus, by subadditivity of $t \mapsto t^{1-\beta/\alpha_\circ}$ and enlarging the constants, we finally arrive at

$$\|\partial_n v_r\|_{L^\infty(B_{1/4})} \le Cr^{\alpha_\circ\beta/(\alpha_\circ-\beta)}\|\partial_n v_r\|_{L^\infty(B_1)} + C\Lambda(2r)^{1-\beta/\alpha_\circ}. \tag{3-14}$$

We conclude by iteration of this inequality. Let us set

$$f(r) := \|\partial_n v_r\|_{L^\infty(B_1)} = r\|\partial_n v\|_{L^\infty(B_r)} \quad \text{and} \quad \delta := \frac{\alpha_\circ\beta}{\alpha_\circ-\beta}.$$

Then (3-14) reads as

$$f\left(\tfrac14 r\right) \le Cr^\delta f(r) + C\Lambda(2r)^{1-\beta/\alpha_\circ} \quad \text{for all } 0 < r < \tfrac14.$$

Since f and Λ are increasing functions, we can iterate this inequality $N \sim k/\delta$ times (N does not depend on r), and it becomes

$$f(r) \le C_N\Lambda(4^N r)^{1-\beta/\alpha_\circ} + C_N f\left(\tfrac14\right)r^{k+2} \quad \text{for all } 0 < r < 4^{-N},$$

and $f\left(\tfrac14\right)$ is again bounded by a dimensional constant, by optimal regularity. Finally, we use Lemma 3.5 to replace $\|v_{4^N r}\|_{L^\infty(B_1)}$ with $\|v_{4^N r}\|_{L^2(B_2\setminus B_{1/2})} + r^{k+2}$. We have proved that

$$\|\partial_n v_r\|_{L^\infty(B_1)} \le C(\|v_{\theta r}\|_{L^2(B_2\setminus B_{1/2})} + r^{k+2})^{1-\beta/\alpha_\circ},$$

with $C = C(n, k, \tau, \beta)$ and $\theta(n, k, \beta) = 4^N$, for all $r \in (0, r_0)$. Since $\beta \in (0, \alpha_\circ/(k+2))$ we proved (3-4).

We turn to the proof of the tangential estimate (3-3). Combining the previous step with Lemma 3.6, we have, for $C > 0$ and $r \in (0, r_0)$,

$$\frac{1}{C}\|\partial_j v_r\|_{L^\infty(B_1)} \le r(\|v_{\theta r}\|_{L^2(B_2\setminus B_{1/2})} + r^{k+2})^{1-\beta/\alpha_\circ} + \|v_r\|_{L^2(B_2\setminus B_{1/2})} + r^{k+2} \le (r^{1-\beta(k+2)/\alpha_\circ} + 1)(\|v_{\theta r}\|_{L^2(B_2\setminus B_{1/2})} + r^{k+2}).$$

We used that θ is large and Lemma 3.5 to bound

$$\|v_r\|_{L^2(B_2\setminus B_{1/2})} \le \|v_r\|_{L^\infty(B_2)} \le \|v_{\theta r}\|_{L^2(B_1)} \lesssim_{n,k,\theta} \|v_{\theta r}\|_{L^2(B_2\setminus B_{1/2})}.$$

Note that with the choice $\beta := \alpha_\circ^2/(k+4)$, we get rid of the β-dependence in the constants and obtain the claimed estimate. □

4. Monotonicity of the truncated frequency

In this section we show that the truncated frequency function $\phi^\gamma(r, u - \mathcal{P}_k)$ from (2-6) is almost monotone for $\gamma < k+2$, regardless of the $p_3, \dots, p_k$. All the proofs in this section do not change for a generic smooth f, as we use the inequalities of the previous section as "black boxes".

The core of the monotonicity is the following computational lemma.

Lemma 4.1. *Let $k \geq 2$, $\tau > 0$, and let $u : B_1 \to \mathbb{R}$ be a solution of the obstacle problem (2-1), with $f \equiv 1$ and $\mu = 1$. Assume $r^{-2}u(r\,\cdot) \to p_2$, and take $(p_2, \ldots, p_k) \in \boldsymbol{P}_k$ such that $|(p_2, \ldots, p_k)| \leq \tau$. Consider $v := u - \mathcal{P}_k$, where $\mathcal{P}_k = \mathcal{P}_k(p_2, \ldots, p_k)$ is constructed as in Definition 3.2. For each $\gamma \in [0, k+2)$ and $\beta \in (0, \alpha_\circ/(k+2))$, set*

$$\epsilon := \min\{\alpha_\circ - \beta(k+2), k+2-\gamma\} > 0,$$

where the dimensional constant $\alpha_\circ$ is the one of Lemma 2.1. Then there exists $r_0 = r_0(k, n, \tau, \beta) \in (0, 1)$ such that, for all $0 < r < r_0$,

$$\frac{d}{dr}\phi^\gamma(r, v) \geq -Cr^{\epsilon-1}(g^\gamma(r, v) + 1)(\phi^\gamma(r) + 1) \tag{4-1}$$

and

$$\frac{\int_{B_r} |v_r \Delta v_r|}{H(r, v) + r^{2\gamma}} \leq Cr^\epsilon(g^\gamma(r, v) + 1), \tag{4-2}$$

where we set

$$g^\gamma(r, v) := \frac{\|v_{\theta r}\|^2_{L^2(B_2 \setminus B_{1/2})}}{H(r, v) + r^{2\gamma}}.$$

Here $\theta = \theta(k, \beta)$ and $C = C(n, k, \tau, \beta)$ are constants.

Proof. Throughout the proof, C will be a constant depending on n, k, τ, β. Up to a rotation of the coordinate axes, we may assume $p_2 = \frac{1}{2}x_n^2$. We begin with recalling the estimate on the derivative of ϕ^γ, compare (2-7),

$$\frac{d}{dr}\phi^\gamma(r, v) \geq \frac{2}{r}\frac{\phi^\gamma(r, v)\int_{B_1} v_r \Delta v_r - \int_{B_1}(x \cdot \nabla v_r)\Delta v_r}{H(r, v) + r^{2\gamma}}.$$

Using $\operatorname{supp} \Delta v_r \subseteq \{u_r = 0\}$, we reduce (4-1) to a bound from below on the quantity

$$\frac{2}{r}\frac{\|\phi^\gamma(r)v_r - x \cdot \nabla v_r\|_{L^\infty(B_1 \cap \{u_r = 0\})}\int_{B_1 \cap \{u_r=0\}} |\Delta v_r|}{H(r, v) + r^{2\gamma}}.$$

For $x \in B_1 \cap \{u_r = 0\}$ and $r < r_0(n, k, \tau, \beta)$, we have

(i) $|x_n| \leq Cr^{\alpha_\circ}$, see Lemma 2.1;

(ii) $|v_r(x)| \leq Cr|\partial_n(\mathcal{P}_k)_r(x)| + Cr^{k+2}$, see Proposition 3.3 (ii);

(iii) $|\partial_j v_r(x)| \leq Cr|\partial_n(\mathcal{P}_k)_r(x)| + Cr^{k+2}$ for all $j \neq n$, see Lemma 3.6;

(iv) $r|\partial_n \mathcal{P}_k(x)| \leq C(\|v_{\theta r}\|_{L^2(B_2 \setminus B_{1/2})} + r^{k+2})^{1-\beta}$, see Proposition 3.4.

Putting together all these bounds we get, for all $x \in B_1 \cap \{u_r = 0\}$,

$$\begin{aligned}
|\phi^\gamma(r)v_r - x \cdot \nabla v_r| &\leq (\phi^\gamma(r) + 1)(|v_r| + |x_n|\,|\partial_n(\mathcal{P}_k)_r| + |\nabla_{x'} v_r|) \\
&\leq C(\phi^\gamma(r) + 1)(r^{\alpha_\circ}\|\partial_n(\mathcal{P}_k)_r\|_{L^\infty(B_1 \cap \{u_r=0\})} + r^{k+2}) \\
&\leq C(\phi^\gamma(r) + 1)r^{\alpha_\circ}(\|v_{\theta r}\|_{L^2(B_3 \setminus B_{1/2})} + r^{k+2})^{1-\beta} + C(\phi^\gamma(r) + 1)r^{k+2} \\
&\leq C(\phi^\gamma(r) + 1)r^\epsilon(\|v_{\theta r}\|_{L^2(B_3 \setminus B_{1/2})} + r^{k+2}) + C(\phi^\gamma(r) + 1)r^{k+2},
\end{aligned}$$

where in the last line we argued

$$(\|v_{\theta r}\|_{L^2(B_3\setminus B_{1/2})}+r^{k+2})^{1-\beta}\le r^{-\beta(k+2)}(\|v_{\theta r}\|_{L^2(B_3\setminus B_{1/2})}+r^{k+2}).$$

Thus, using $(H(r,v)+r^{2\gamma})^{-1/2}\le r^{-\gamma}$ and $r^{k+2-\gamma}\le r^{\epsilon}$, we get

$$\frac{\|\phi^\gamma(r)v_r-x\cdot\nabla v_r\|_{L^\infty(B_1\cap\{u_r=0\})}}{(H(r,v)+r^{2\gamma})^{1/2}}\le r^{\epsilon}(\phi^\gamma(r)+1)(g^\gamma(r)^{1/2}+1).$$

Now, using (ii) and (iv) to estimate $|\nabla v_r(x)|$ for $x\in B_1\cap\{u_r=0\}$ as above, we get

$$\frac{\|v_r\|_{L^\infty(B_1\cap\{u_r=0\})}}{(H(r,v)+r^{2\gamma})^{1/2}}\le r^{\epsilon}(g^\gamma(r)^{1/2}+1).$$

Thanks to this observation, to prove both (4-1) and (4-2) we only need to show

$$\frac{\int_{B_1\cap\{u_r=0\}}|\Delta v_r|}{(H(r,v)+r^{2\gamma})^{1/2}}\le Cg^\gamma(r)^{1/2}.$$

As $\Delta v_r=-r^2\chi_{\{u_r=0\}}$, we can integrate by parts with some regular cut-off $\chi_{B_1}\le\psi\le\chi_{B_2}$:

$$\int_{B_1\cap\{u_r=0\}}|\Delta v_r|=-\int_{B_1}\Delta v_r\le-\int_{B_2}\Delta v_r\psi\lesssim_n\|v_r\|_{L^1(B_2\setminus B_1)},$$

which concludes the proof. □

We now make a specific choice of β and derive from these preliminary bounds the monotonicity of the truncated frequency.

Proposition 4.2. *Let $k\ge 2$, $\tau>0$, and let u be a solution of the obstacle problem (2-1) with $f\equiv 1$ and $\mu=1$. Assume $r^{-2}u(r\,\cdot)\to p_2$, and take $(p_2,\dots,p_k)\in\boldsymbol{P}_k$ such that $|(p_2,\dots,p_k)|\le\tau$. Consider $v:=u-\mathcal{P}_k$, where $\mathcal{P}_k=\mathcal{P}_k(p_2,\dots,p_k)$ is constructed as in Definition 3.2, and, for each $\gamma\in(0,k+2)$, set*

$$\varepsilon(\gamma):=\min\bigl\{\tfrac12\alpha_\circ;\,k+2-\gamma\bigr\},$$

where the dimensional constant $\alpha_\circ$ is the one of Lemma 2.1. Then, there exist $C(n,k,\tau,\gamma)>0$ and $r_0(n,k,\tau)\in(0,1)$ such that, for all $0<r<r_0$, we have

$$\frac{d}{dr}\phi^\gamma(r)\ge -Cr^{\varepsilon-1},\quad \phi^\gamma(r)\le C\quad\text{and}\quad \frac{\int_{B_r}|v_r\Delta v_r|}{H(r,v)+r^{2\gamma}}\le Cr^{\varepsilon}. \tag{4-3}$$

In particular, $\phi^\gamma(0^+,v)=\lim_{r\downarrow 0}\phi^\gamma(r,v)$ exists and $\phi^\gamma(0^+,v)\le\gamma$.

Proof. Fix $\gamma_\circ\in(0,k+2)$, and let $\varepsilon_\circ=\varepsilon(\gamma_\circ)$. Let $r_0=r_0(n,k,\tau,\beta=\alpha_\circ/(2(k+2)))$ be as in Lemma 4.1. By Lemma 4.1 and estimate (4-1), we only need to show that the functions $g^{\gamma_\circ}(\cdot)$ and $\phi^{\gamma_\circ}(\cdot)$ are uniformly bounded in the interval $(0,r_0]$. We are going to prove it by increasing iteratively the parameter γ, exactly as in the proof of [Figalli et al. 2020, Lemma 4.3]. Throughout the proof, C_γ will denote a general constant depending on $n,\ k,\ \tau,\ \gamma$, and similarly for $C_{\gamma,\delta}$.

First notice that by (2-2) the functions $\phi^0(\,\cdot\,, v)$ and $g^0(\,\cdot\,)$ are uniformly bounded in $[0, r_0]$. Fix any $\gamma \in [0, \gamma_\circ]$. The core of the proof is the observation

$$\begin{cases} \phi^\gamma + g^\gamma \le C_\gamma & \text{in } (0, r_0], \\ 0 < 5\delta \le \varepsilon_\circ, \end{cases} \implies \phi^{\gamma+\delta} + g^{\gamma+\delta} \le C_{\gamma,\delta} \quad \text{in } (0, r_0]. \tag{4-4}$$

We iterate this observation to reach the conclusion. Define the sequence $\gamma_0 = 0,\ \gamma_{j+1} := \gamma_j + \frac{1}{5}\varepsilon_\circ$, where $j \ge 0$. With a finite number of iterations we get closer than $\frac{1}{5}\varepsilon_\circ$ to $\gamma_\circ$, and applying (4-4) once more with an appropriate δ we get to $\gamma_\circ$.

We prove (4-4) for a generic γ. Keeping in mind $r, \delta \le 1$, we estimate

$$\phi^{\gamma+\delta}(r) = \frac{D(r, v) + (\gamma+\delta) r^{2\gamma+2\delta}}{H(r, v) + r^{2\gamma+2\delta}} \le \frac{1}{r^{2\delta}} \frac{D(r, v) + \gamma r^{2\gamma}}{H(r, v) + r^{2\gamma}} + 1 \le \frac{C_\gamma}{r^{2\delta}}$$

and

$$g^{\gamma+\delta}(r) = \frac{\|v_{\theta r}\|^2_{L^2(B_3\setminus B_{1/2})}}{H(r, v) + r^{2\gamma+2\delta}} \le \frac{1}{r^{2\delta}} \frac{\|v_{\theta r}\|^2_{L^2(B_2\setminus B_{1/2})}}{H(r, v) + r^{2\gamma}} \le \frac{C_\gamma}{r^{2\delta}}.$$

Now we apply Lemma 4.1 with $\beta := \alpha_\circ/(2(k+2))$ and $\gamma \to \gamma + \delta$:

$$\frac{d}{dr}\phi^{\gamma+\delta}(r) \ge -C r^{\epsilon-1}(g^{\gamma+\delta}(r)+1)(\phi^{\gamma+\delta}(r)+1) \ge -C_\gamma r^{\varepsilon_\circ - 4\delta - 1} \ge -C_\gamma r^{\delta-1},$$

where we used $\epsilon(\beta, \gamma+\delta) \ge \varepsilon_\circ$ and the smallness of δ. Integrating this inequality and using $C^{1,1}$ estimates, we obtain

$$\phi^{\gamma+\delta}(r) \le \phi^{\gamma+\delta}(r_0) + C_\gamma r_0^\delta/\delta \lesssim_{n,k,\tau} C_{\gamma,\delta}$$

for all $r \in (0, r_0]$. The uniform boundedness of $g^{\gamma+\delta}$ is now a consequence of Lemma 2.2 (a), as we can take $\bar\lambda = C_{\gamma,\delta}$. Indeed, for any $0 < 2\theta r < r_0$ and any $R \in \left(\frac{1}{2} r\theta, 2\theta r\right)$, we have

$$\frac{H(R, v) + R^{2\gamma+2\delta}}{H(r, v) + r^{2\gamma+2\delta}} \lesssim_{n,k,\tau} \left(\frac{R}{r}\right)^{C_{\gamma,\delta}} \le C_{\gamma,\delta}.$$

Thus, for all $r \in (0, r_0/(2\theta)]$, we have

$$g^{\gamma+\delta}(r) = \frac{\|v_{\theta r}\|^2_{L^2(B_2\setminus B_{1/2})}}{H(r, v) + r^{2\gamma+2\delta}} \lesssim_{n,k,\tau} \fint_{\theta r/2}^{2\theta r} \frac{H(R, v)\, dR}{H(r, v) + r^{2\gamma+2\delta}} \le C_{\gamma,\delta}.$$

Finally $g^{\gamma+\delta}$ is clearly bounded in $[r_0/(2\theta), r_0]$ by (2-2). This concludes the proof. □

The following is an immediate corollary of Proposition 4.2.

Corollary 4.3. *With the same notation from Proposition 4.2, we have that, for all $r \in (0, r_0)$,*

$$\frac{d}{dr}\log(r^{-2\lambda}(H(r, u - \mathcal{P}_k) + r^{2\gamma})) \ge -C r^{\varepsilon-1}, \tag{4-5}$$

provided $\lambda \le \phi^\gamma(0^+, v)$. Here $r_0,\ C,\ \varepsilon$ are the same positive numbers from Proposition 4.2.

Proof. Computing the derivative and using Proposition 4.2, we obtain

$$\frac{d}{dr}\log(r^{-2\lambda}(H(r,v)+r^{2\gamma})) = \frac{2}{r}(\phi^{\gamma}(r,v)-\lambda)+\frac{2}{r}\frac{\int_{B_1} v_r\Delta v_r}{H(r,v)+r^{2\gamma}}$$
$$\geq -C\int_0^r s^{\varepsilon-1}\,ds - Cr^{\varepsilon-1} \geq -Cr^{\varepsilon-1}. \qquad \square$$

5. The sets $\Sigma^{k\text{-th}}$ and higher-order blowups

The estimates obtained in the previous two sections did not assume any relationship between u and $\mathcal{P}_k$, beside the crucial fact that p_2 was the blowup of u at 0. In this section, instead, we define the set of points at which u admits a polynomial expansion of order k. This expansion will identify the polynomial $\mathcal{P}_k$ up to order k, leaving some freedom for the $k+1$ terms (compare Proposition 3.3 (ii)).

The results of this section never use explicitly the simplifying assumption $f \equiv 1$; they use it implicitly, though, employing the results of the previous two sections. Hence for a generic smooth f, all the results of this section apply without change.

Recall that $\boldsymbol{P}_k$ was defined at the beginning of Section 3.1.

Definition 5.1. Let $u : B_1 \to [0,\infty)$ solve (2-1) and $x_\circ \in \Sigma_{n-1}$. We say the following:

(i) $\Sigma^{2\text{nd}} := \Sigma_{n-1}$.

(ii) $x_\circ \in \Sigma^{3\text{rd}}$ if

$$r^{-3}(u(x_\circ + r\,\cdot\,) - p_{2,x_\circ}(r\,\cdot\,)) \to p_{3,x_\circ} \quad \text{in } W^{1,2}_{\text{loc}}(\mathbb{R}^n)\cap C^0_{\text{loc}}(\mathbb{R}^n)$$

as $r \downarrow 0$, where $p_{3,x_\circ}$ is some 3-homogeneous harmonic polynomial vanishing on $\{p_{2,x_\circ}=0\}$. Thus we have $(p_{2,x_\circ}, p_{3,x_\circ}) \in \boldsymbol{P}_3$.

(iii) $x_\circ \in \Sigma^{k\text{-th}}$ for some integer $k \geq 4$, if $x_\circ \in \Sigma^{(k-1)\text{-th}}$ and

$$r^{-k}(u(x_\circ + r\,\cdot\,) - \mathcal{P}_{k-1,x_\circ}(r\,\cdot\,)) \to p_{k,x_\circ} \quad \text{in } W^{1,2}_{\text{loc}}(\mathbb{R}^n)\cap C^0_{\text{loc}}(\mathbb{R}^n)$$

as $r \downarrow 0$, where $\mathcal{P}_{k-1,x_\circ} = \mathcal{P}_{k-1}(p_{2,x_\circ},\dots,p_{k-1,x_\circ})$ is the polynomial ansatz from Definition 3.2 and the limit $p_{k,x_\circ}$ is some k-homogeneous harmonic polynomial vanishing on $\{p_{2,x_\circ}=0\}$. Thus we have $(p_{2,x_\circ},\dots,p_{k,x_\circ}) \in \boldsymbol{P}_k$.

When $x_\circ = 0$, we simply drop $x_\circ$ from $p_{j,x_\circ}$.

Remark 5.2. Let $u : B_1 \to [0,\infty)$ solve (2-1) and suppose $0 \in \Sigma^{k\text{-th}}$ for some $k \geq 2$. Then

(i) $\phi^{\gamma}(0^+, u-\mathcal{P}_k)$ exists for all $\gamma \in [k, k+2)$, see Proposition 4.2;

(ii) for every $\delta > 0$, we have $r^{2\phi^{\gamma}(0^+,u-\mathcal{P}_k)+\delta} \ll H(r,u-\mathcal{P}_k) = o(r^{2k})$ as $r \downarrow 0$. The lower bound follows by Lemma 2.2, the upper bound by continuity of the embedding $W^{1,2}(B_1) \to L^2(\partial B_1)$.

With help of Corollary 4.3, we prove a Monneau-type monotonicity formula; see [Monneau 2003]. The argument is an adaptation of [Figalli and Serra 2019, Lemma 4.1].

Proposition 5.3 (Monneau-type monotonicity). *Let $u : B_1 \to [0,\infty)$ solve (2-1). Suppose $0 \in \Sigma^{k\text{-th}}$ for some $k \geq 3$. Let q be any polynomial such that $(p_{2,0},\dots,p_{k-1,0},q) \in \boldsymbol{P}_k$, with $|(p_{2,0},\dots,p_{k-1,0},q)| \leq \tau$ for some number $\tau > 0$. Set*

$$w := u - \mathcal{P}_k(p_{2,0},\dots,p_{k-1,0},q).$$

Then there exists $r_0 = r_0(n,k,\tau)$ such that, for all $r \in (0,r_0)$,

$$r^{-2k} H(r,w) \leq C \quad \text{and} \quad \frac{d}{dr}(r^{-2k} H(r,w)) \geq -C r^{\varepsilon-1}$$

for some constants $C = C(n,k,\tau)$ and $\varepsilon(n,k) > 0$.

Proof. For the sake of the proof, let us fix $\gamma = k+1+\frac{1}{2}$; all constants are allowed to depend on n, k, γ, even if not explicitly stated, and can change value from line to line. We begin by showing $\phi^\gamma(0^+,w) \geq k$. By construction of the $\mathcal{P}_k$ (see Proposition 3.3), we have

$$w = u - \mathcal{P}_{k,0} - p_{k,0} + q + O(x^{k+1}),$$

so using $0 \in \Sigma^{k\text{-th}}$ we find

$$H(r,w)^{1/2} \leq H(r,u-\mathcal{P}_k)^{1/2} + H(r,p_k)^{1/2} + H(r,q)^{1/2} + H(r,O(|x|^{k+1}))^{1/2} = o(r^k) + O(r^k).$$

If $\phi^\gamma(0^+,w) < k$ were true, then Lemma 2.2 would give, for $r \ll 1$,

$$r^{2\phi^\gamma(0^+,w)+\delta} \ll H(r,w) + r^{2\gamma} \lesssim r^{2k},$$

which would be a contradiction for $\delta > 0$ small. Hence, $\phi^\gamma(0^+,w) \geq k$, and we can apply Corollary 4.3 with $\gamma = k+1+\frac{1}{2}$ to find that the function

$$f(r) := \log(r^{-2k}(H(r,w) + r^{2\gamma})) + C r^{\varepsilon}$$

is increasing in $(0,r_0)$ for appropriate r_0, C, ε depending on n, k, τ. Using Equation (2-2) we have that $\|w\|_{L^\infty(B_1)} \leq C$, and so $f(r_0) \leq C$. Thus, by monotonicity of f,

$$r^{-2k}(H(r,w) + r^{2\gamma}) \leq C$$

for $r \in (0,r_0)$. Inserting again this estimate in Corollary 4.3, we get

$$\frac{d}{dr}(r^{-2k}(H(r,w)+r^{2\gamma})) \geq -C r^{\varepsilon-1}(r^{-2k}(H(r,w)+r^{2\gamma})) \geq -C r^{\varepsilon-1},$$

and as $(d/dr) r^{2(\gamma-k)} \ll r^{\varepsilon-1}$ we conclude. □

The following result proves the continuity of the map $x \mapsto \mathcal{P}_{k,x}$, defined on $\Sigma^{k\text{-th}}$, for $k \geq 2$. Our argument is a direct adaptation of [Figalli and Serra 2019, Proposition 4.5] for the case $k = 3$.

Proposition 5.4. *Let $u : B_1 \to [0,\infty)$ solve (2-1) and $k \geq 2$. Then the map $\Sigma^{k\text{-th}} \ni x \mapsto \mathcal{P}_{k,x}$ is continuous. Furthermore, there exists a constant $\tau(n,k)$ such that*

$$\sup\{|(p_{2,x},\dots,p_{k,x})| : x \in \Sigma^{k\text{-th}} \cap B_{1/2}\} \leq \tau(n,k). \tag{5-1}$$

Proof. We first prove the bound on $\tau(n,k)$. Since $x \in \Sigma^{k\text{-th}}$ implies $x \in \Sigma^{j\text{-th}}$ for $j \le k$, we proceed by induction on k. The inductive step follows from Proposition 5.3 applied to the functions $u(x+\cdot)$ and $q=0$, which allows us to deduce that $p_{k+1,x}$ is bounded in terms of n, k and $\tau(n,k-1)$. We can take as the base step $k=2$, for which $|(p_{2,x})| \le \frac{1}{2}$.

Let us prove continuity at 0. Again, we proceed inductively and suppose that the statement is true for $k-1$ (see [Figalli and Serra 2019] for $k-1=2$). Let $(x_\ell)_{\ell\in\mathbb{N}} \subseteq \Sigma^{k\text{-th}} \cap B_{1/2}$ with $x_\ell \to 0$, and choose a sequence of rotations $(R_\ell)_{\ell\in\mathbb{N}} \subseteq \mathrm{SO}(n)$ mapping $\{p_{2,x_\ell}=0\}$ to $\{p_{2,0}=0\}$ for each ℓ and satisfying $R_\ell \to \mathrm{id}$. We apply Proposition 5.3 to the functions $u(x_\ell+\cdot)$ and the polynomials $q_\ell := p_{k,0}\circ R_\ell$, and since

$$u(x_\ell+y) - \mathcal{P}_{k,x_\ell}(p_{2,x_\ell},\dots,p_{k-1,x_\ell},q) = u(x_\ell+y) - \mathcal{P}_{k-1,x_\ell}(y) + q_\ell(y) + O(|y|^{k+1}),$$

we find that the function

$$r \mapsto \int_{\partial B_1} \left| \frac{u(x_\ell + r\,\cdot) - \mathcal{P}_{k-1,x_\ell}(r\,\cdot)}{r^k} - p_{k,0}\circ R_\ell + \frac{O(|rx|^{k+1})}{r^k} \right|^2 d\sigma + Cr^\varepsilon \tag{5-2}$$

is increasing in $(0,r_0)$ for all $\ell \in \mathbb{N}$, for some r_0 and C uniform in ℓ. Using this information for the constant sequence $x_\ell = 0$, we find that, for any $\delta > 0$, there is $r_\delta < \min\{r_0,\delta\}$ such that

$$\int_{\partial B_1} \left| \frac{u(r_\delta\,\cdot) - \mathcal{P}_{k-1,0}(r_\delta\,\cdot)}{r_\delta^k} - p_{k,0} + \frac{O(|r_\delta x|^{k+1})}{r_\delta^k} \right|^2 d\sigma \le \delta. \tag{5-3}$$

Using (5-2) we estimate, for each ℓ,

$$\begin{aligned}\int_{\partial B_1} |p_{k,x_\ell} - p_{k,0}\circ R_\ell|^2 &= \lim_{r\downarrow 0} \int_{\partial B_1} \left| \frac{u(x_\ell + r\,\cdot) - \mathcal{P}_{k-1,x_\ell}(r\,\cdot)}{r^k} - p_{k,0}\circ R_\ell \right|^2 d\sigma \\ &\le \int_{\partial B_1} \left| \frac{u(x_\ell + r_\delta\,\cdot) - \mathcal{P}_{k-1,x_\ell}(r_\delta\,\cdot)}{r_\delta^k} - p_{k,0}\circ R_\ell + O(r_\delta) \right|^2 d\sigma + Cr_\delta^\varepsilon.\end{aligned}$$

As $\mathcal{P}_{k-1,x_\ell} \to \mathcal{P}_{k-1,0}$ by inductive assumption, taking the upper limit in ℓ on both sides and using (5-3), we find

$$\limsup_{\ell} \int_{\partial B_1} |p_{k,x_\ell} - p_{k,0}|^2 \le \delta,$$

and letting $\delta \downarrow 0$ we conclude. □

The following definition is useful to quantify the rate at which $\mathcal{P}_k$ approximates u.

Definition 5.5 (frequency). For $u : B_1 \to [0,\infty)$ a solution to (2-1) and $k \ge 2$, define the k-th frequency $\lambda_k : \Sigma^{k\text{-th}} \to [k, k+2]$ by

$$\lambda_k(x) := \sup\{\phi^\gamma(0^+, u(x+\cdot) - \mathcal{P}_{k,x}) : \gamma \in [k, k+2)\}.$$

At $x=0$, we write $\lambda_k := \lambda_k(0)$.

We comment that in the definition above we indeed have $\lambda_k(\Sigma^{k\text{-th}}) \subseteq [k,k+2]$. First, the fact that $\phi^\gamma(0^+, u-\mathcal{P}_k) \ge k$ holds for every $\gamma \in [k,k+2)$ was observed in Proposition 5.3. Second, we always have $\phi^\gamma(0^+, u-\mathcal{P}_k) \le \gamma$, as observed in Proposition 4.2. The following lemma shows that indeed ϕ^γ is a truncation of the frequency, that is $\phi^\gamma(0^+, u-\mathcal{P}_k) = \min\{\lambda_k, \gamma\}$.

Lemma 5.6. *For $u : B_1 \to [0,\infty)$ a solution to (2-1) and $k \ge 2$, consider $x_\circ \in \Sigma^{k\text{-}th}$. Then, for all $\gamma \in (\lambda_k, k+2)$,*

$$\lambda_k(x_\circ) = \phi^\gamma(0^+, u(x_\circ + \cdot\,) - \mathcal{P}_{k,x_\circ}) = \lim_{r\downarrow 0} \phi(r, u(x_\circ + \cdot\,) - \mathcal{P}_{k,x_\circ}),$$

where $\phi(r, v) := D(r, v)/H(r, v)$ is the (nontruncated) Almgren frequency function.

Proof. For simplicity, let $x_\circ = 0$ and set $v := u - \mathcal{P}_k$. By Lemma 2.2, for each $\delta > 0$, there is a constant c_δ and a radius r_δ such that $C_\delta r^{2\lambda_k+\delta} \ll H(r, u - \mathcal{P}_k) + r^{2\gamma}$ for every $0 < r < r_\delta$. Hence, after picking $0 < \delta < \frac{1}{10}(\gamma - \lambda_k)$, we find

$$\phi^\gamma(0^+, v) = \lim_{r\downarrow 0} \frac{\phi(r, v) + o(1)}{1 + o(1)} = \lim_{r\downarrow 0} \phi(r, v) =: \tilde\lambda,$$

where $\tilde\lambda$ does not depend on the choice of $\gamma \in (\lambda_k, k+2)$. On the one hand, $\tilde\lambda \le \gamma$ holds for any such γ, implying $\tilde\lambda \le \lambda_k$; see Proposition 4.2. On the other hand, $\lambda_k \ge \phi^\gamma(0^+, v) = \tilde\lambda$, by definition. □

We now give a more flexible characterization of $\Sigma^{k\text{-th}}$.

Lemma 5.7. *For every solution $u : B_1 \to [0,\infty)$ to (2-1) and $k \ge 2$, we have*

$$\Sigma^{k\text{-}th} \equiv \widetilde\Sigma^{k\text{-}th} := \Big\{x \in \Sigma^{(k-1)\text{-}th} : \exists (q_2, \dots, q_k) \in \boldsymbol{P}_k, \\ \exists r_\ell \downarrow 0 \text{ such that } r_\ell^{-k}(u - \mathcal{P}_{k-1}(q_2, \dots, q_{k-1}))_{r_\ell} \rightharpoonup q_k \text{ in } W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)\Big\}.$$

Proof. We just need to show that $\widetilde\Sigma^{k\text{-th}} \subseteq \Sigma^{k\text{-th}}$, because the other inclusion follows by definition. Let $0 \in \widetilde\Sigma^{k\text{-th}}$. We know that

$$r_\ell^{-k}(u - \mathcal{P}_{k-1}(q_2, \dots, q_{k-1}))_{r_\ell} \rightharpoonup q_k \quad \text{in } W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)$$

for a certain $(q_2, \dots, q_k) \in \boldsymbol{P}_k$ and a certain sequence $r_\ell \downarrow 0$.

We first show that necessarily $q_j = p_{j,0}$ for all $2 \le j \le k-1$. To prove this we reason inductively and exploit the fact that $0 \in \Sigma^{j\text{-th}}$; in particular, we have the uniform convergence

$$\lim_{r\to 0} r^{-j}(u - \mathcal{P}_j(p_{2,0}, \dots, p_{j,0}))_r = 0.$$

Suppose $(p_{2,0}, \dots, p_{j-1,0}) = (q_2, \dots, q_{j-1})$ holds for some $j \ge 2$. By Proposition 3.3 (ii), we have

$$\begin{aligned}
0 &= \lim_\ell \frac{(u - \mathcal{P}_k(q_2, \dots, q_j))_{r_\ell}}{r_\ell^j} \\
&= \lim_\ell \frac{(u - \mathcal{P}_j(p_{2,0}, \dots, p_{j,0}))_{r_\ell}}{r_\ell^j} + \frac{(\mathcal{P}_j(p_{2,0}, \dots, p_{j,0}) - \mathcal{P}_k(q_2, \dots, q_k))_{r_\ell}}{r_\ell^j} \\
&= 0 + \lim_\ell \frac{(\mathcal{P}_{j-1}(p_{2,0}, \dots, p_{j-1,0}) + p_{j,0} - \mathcal{P}_{j-1}(q_2, \dots, q_{j-1}) - q_j + O(|x|^{j+1}))_{r_\ell}}{r_\ell^j} \\
&= p_{j,0} - q_j,
\end{aligned}$$

in $L^2(\partial B_1)$. This completes the inductive step. The same computation gives also the base step $q_2 = p_{2,0}$.

Now let $\mathcal{P}_k = \mathcal{P}_k(p_2, \dots, p_{k-1}, q_k)$. Then $\mathcal{P}_k = \mathcal{P}_{k-1} + q_k + P$ for some $(k+1)$-homogeneous harmonic polynomial P (see (ii) in Proposition 3.3). We set $v = u - \mathcal{P}_k$ and notice that by assumption $r_\ell^{-k} v_{r_\ell} \rightharpoonup 0$

in $W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)$. We will show that this convergence is strong, locally uniform and happens along the full range $r \downarrow 0$, which in turn implies that $0 \in \Sigma^{k\text{-th}}$.

Take r_0 as in Proposition 4.2 and fix some $\gamma \in (k+1, k+2)$. As in (ii) in Remark 5.2, we obtain $H(r_\ell, v) = o(r_\ell^{2k})$, which as in Proposition 5.3 implies $\lambda := \lim_{r\to 0} \phi^\gamma(r, v) \geq k$. Since by Proposition 4.2 $\phi^\gamma(\cdot\,, v)$ is bounded by C_γ in $(0, r_0)$, we have

$$r^{-2k} \int_{B_R} |\nabla v_r|^2 \leq r^{-2k}\phi^\gamma(Rr, v)(H(Rr, v) + (Rr)^{2\gamma}) \leq C_\gamma r^{-2k}(H(Rr, v) + (Rr)^{2\gamma}),$$

provided $Rr < r_0$. We now exploit that the logarithm of the right-hand side is almost monotone in r thanks to Corollary 4.3 and get

$$\begin{aligned}\limsup_{r\downarrow 0} \log\Bigl(r^{-2k} \int_{B_R} |\nabla v_r|^2\Bigr) &\leq \log C_\gamma + \lim_{s\downarrow 0} \log(s^{-k}(H(s, v) + s^{2\gamma})) \\ &= \log C_\gamma + \lim_{\ell\to\infty} \log(r_\ell^{-k}(H(r_\ell, v) + r_\ell^{2\gamma})) = -\infty,\end{aligned}$$

thus $\lim_{r\downarrow 0} \|r^{-k}\nabla v_k\|_{L^2(B_R)} = 0$ for all fixed $R > 0$. The proof of local uniform convergence is very similar: namely, using Lemma 3.5 and then Lemma 2.2, we have

$$\|v_r\|_{L^\infty(B_R)} \leq C\|v_{Rr}\|_{L^2(B_2\setminus B_{1/2})} + C(Rr)^{k+2} \leq CH(Rr, v)^{1/2} + C(Rr)^{k+2},$$

provided Rr is small, thus we can divide by r^k and argue as before exploiting the log-monotonicity. □

With the same kind of reasoning we can prove the following basic lemma.

Lemma 5.8. *Let $u : B_1 \to [0, \infty)$ be a solution to (2-1) and $k \geq 2$, and suppose $0 \in \Sigma^{k\text{-th}}$ with $\lambda_k > k+1$. Then $0 \in \Sigma^{(k+1)\text{-th}}$ and $p_{k+1} = 0$.*

Proof. Set $v := u - \mathcal{P}_k$, and pick any $\gamma \in (\lambda_k, k+2)$, so that $\phi^\gamma(0^+, v) > k+1$. Arguing as in the proof of Lemma 5.7, we find

$$r^{-2(k+1)} \int_{B_R} |\nabla v_r|^2 \lesssim r^{-2(k+1)}(H(Rr, v) + (Rr)^{2\gamma}),$$

provided $Rr < r_0 \ll 1$. On the other hand, as $\phi^\gamma(0^+, v) > k+1$, we have $\phi^\gamma(r, u - \mathcal{P}_k) > k+1$ for $r \ll 1$. Thus, with Lemma 2.2 we deduce $H(r, u - \mathcal{P}_k) = o(r^{2(k+1)})$. Taking the above estimate into account, we conclude $r^{-2(k+1)}v_r \to 0$ in $W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)$, thus by Lemma 5.7, we have $0 \in \Sigma^{(k+1)\text{-th}}$. □

Finally, we study the blowups of $u - \mathcal{P}_k$ when Lemma 5.8 does not apply. Our argument is an adaptation of [Figalli and Serra 2019, Proposition 2.10]. We will study the sequence of functions $\tilde{v}_r := H(r, u - \mathcal{P}_k)^{-1/2}(u - \mathcal{P}_k)(r\,\cdot)$ as $r \downarrow 0$. Any limit of $\tilde{v}_r$ will be a λ_k-homogeneous solution of a certain PDE (the Signorini problem (5-4)) but not necessarily a polynomial.

Proposition 5.9. *Let $0 \in \Sigma^{k\text{-th}}$ with $\lambda_k \leq k+1$. Let $(r_\ell)_{\ell\in\mathbb{N}}$ be an infinitesimal sequence, and let $x_\ell \in \Sigma^{k\text{-th}} \cap B_{r_\ell}$. For every ℓ, set $v_{x_\ell} := u(x_\ell + \cdot\,) - \mathcal{P}_{k,x_\ell}$, and suppose that $\lambda_k(x_\ell) \to \lambda_k$. Consider the sequence*

$$\tilde{v}_{r_\ell,x_\ell} := \frac{v_{x_\ell}(r_\ell\,\cdot)}{H(r_\ell, v_{x_\ell})^{1/2}}.$$

Then:

(i) $(\tilde{v}_{r_\ell,x_\ell})_{\ell\in\mathbb{N}}$ *is bounded in* $W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)$ *and* $C^{0,1/(n+1)}_{\mathrm{loc}}(\mathbb{R}^n)$.

(ii) *If* $\tilde{v}_{r_\ell,x_\ell} \rightharpoonup q \in W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)$, *then the convergence is in fact strong and q must be a nontrivial λ_k-homogeneous solution of the Signorini problem with obstacle $\{p_2=0\}$, that is*

$$\begin{cases} \Delta q \le 0 \text{ and } q\Delta q = 0 & \text{in } \mathbb{R}^n, \\ \Delta q = 0 & \text{in } \mathbb{R}^n \setminus \{p_2=0\}, \\ q \ge 0 & \text{on } \{p_2=0\}. \end{cases} \tag{5-4}$$

Finally, if $\lambda_k < k+1$, then q is even with respect to the thin obstacle.

Proof. For the sake of readability, we set $v_\ell := v_{x_\ell}$ and $\tilde{v}_\ell := \tilde{v}_{r_\ell,x_\ell}$. Furthermore, we will omit the dependence of the constants on n and k, and set $\delta := \frac{1}{100}\varepsilon$, where $\varepsilon(n,k)$ is the same as in Proposition 5.3.

Without loss of generality, assume ℓ is large enough that $x_\ell \in B_{1/2}$ and $\lambda_k(x_\ell) \le \lambda_k + \delta$. Within this proof we fix $\gamma := k+1+\frac{3}{4}$, so by Lemma 5.6 we have that $\lambda_k(x_\ell) = \phi^\gamma(0^+, v_\ell)$ for all ℓ.

Step 1. We claim that there are $\varepsilon, r_0 \in \left(0,\frac{1}{2}\right)$ and $C_0, c_0 > 0$, all independent of ℓ and ℓ_0, such that if we define

$$f_{x_\ell}(r) := \phi^\gamma(r, v_\ell) + C_0 r^\varepsilon \quad \text{and} \quad h_{x_\ell}(r) := r^{-2k} H(r, v_\ell) + C_0 r^\varepsilon,$$

then:

- f_{x_ℓ} and h_{x_ℓ} are continuous and increasing on $[0, r_0]$ and converge uniformly in this interval to f_0 and h_0, respectively. Furthermore, $h_{x_\ell}(0^+) = 0$ identically.
- We have $f_{x_\ell}(r) \le \lambda_k + 2\delta$ and $H(r, v_\ell) \ge c_0 r^{2\lambda_k+5\delta}$ for all $r \in [0, r_0]$ and all $\ell > \ell_0$.

The fact that, for each ℓ, both f_{x_ℓ} and h_{x_ℓ} are increasing is a consequence of Proposition 4.2 and Proposition 5.3, respectively. By Proposition 5.4 we can choose $\tau := \tau(n,k)$ such that r_0 and C_0 can be taken uniform in ℓ. Since $x_\ell \in \Sigma^{k\text{-th}}$, we already observed in Remark 5.2 that $h_{x_\ell}(0^+) = 0$. Furthermore, by assumption, $f_{x_\ell}(0^+) = \lambda_k(x_\ell) \to \lambda_k = f_0(0^+)$, thus $f_{x_\ell} \to f_0$ and $h_{x_\ell} \to h_0$ pointwise. As they are monotone and the limit functions are continuous, the convergence must be uniform, and thus (a) is proved. We turn to (b): possibly taking a smaller r_0, we have that $f_0 \le \lambda_k + \delta$ in $[0, r_0]$, and thus by uniform convergence there is ℓ_0 such that $f_{x_\ell} \le \lambda_k + 2\delta$. Now the last statement follows if we apply Lemma 2.2 with $w = v_\ell,\ R = r_0,\ \bar{\lambda} = \lambda_k + \delta,\ \delta = \frac{1}{100}\varepsilon$.

We will use many times through the proof that

$$c_0 r^{2\gamma} \le H(r, v_\ell) r^{1/2} \quad \text{in } [0, r_0], \tag{5-5}$$

uniformly in $\ell > \ell_0$. This is a direct consequence of (b) and the fact that $2\gamma - \frac{1}{2} > 2\lambda_k + 5\delta$.

Step 2. We prove (i). Fix some $R > 1$ and for $Rr_\ell < r_0$ write

$$D(R, \tilde{v}_\ell) \le \phi^\gamma(Rr_\ell, v_\ell)\frac{H(Rr_\ell, v_\ell) + (Rr_\ell)^{2\gamma}}{H(r_\ell, v_\ell)} \le C_R \frac{H(Rr_\ell, v_\ell)}{H(r_\ell, v_\ell)} + C_R \frac{r_\ell^{2\gamma}}{H(r_\ell, v_\ell)} \le C_R + o(1),$$

where we used (5-5) and Lemma 2.2 to find the second and the first addendum, respectively. Thus $\|\nabla \tilde{v}_\ell\|_{L^2(B_R)}$ is bounded in ℓ.

Now we want to combine the uniform L^2 bound on $\nabla \tilde{v}_\ell$ and the Lipschitz estimate on $\nabla_{x'} \tilde{v}_r$ to produce uniform Hölder bounds. Fix some ℓ and choose coordinates so that $p_{2,x_\ell} = \frac{1}{2}x_n^2$. By Proposition 3.4, we have, for $2\theta R r_\ell < r_0$ and $j \neq n$,

$$\|\partial_j \tilde{v}_\ell\|_{L^\infty(B_R)} \le C\|\tilde{v}_\ell\|_{L^2(B_{2\theta R}\setminus B_{\theta R/2})} + C\frac{r^{k+2}}{H(r_\ell, v_\ell)^{1/2}}.$$

The second term is $o(1)$ in ℓ again by (5-5), while for the first we employ again Lemma 2.2 to find

$$\|v_\ell(r_\ell\,\cdot)\|^2_{L^2(B_{2\theta R}\setminus B_{\theta R/2})} = \int_{\theta R/2}^{2\theta R} H(sr_\ell, v_\ell)s^{n-1}\,ds \le C_R H(r_\ell, v_\ell). \tag{5-6}$$

Hence, $\|\partial_j \tilde{v}_\ell\|_{L^\infty(B_R)} \le C_R$ for all $j \neq n$. By Lemma C.1, this gives the Hölder bound

$$[\tilde{v}_\ell]_{C^{0,1/(n+1)}(B_{R/2})} \le C_R(\|\nabla_{x'}\tilde{v}_\ell\|_{L^\infty(B_R)} + \|\nabla \tilde{v}_\ell\|_{L^2(B_R)}) \le C_R.$$

This concludes the proof of (i).

Step 3. We turn to (ii) and prove that q solves (5-4). Since $\Delta v_\ell = -\chi_{\{u(x_\ell+\,\cdot\,)=0\}} \le 0$, we have that $\Delta q \le 0$ weakly in $\mathbb{R}^n$. Furthermore, integrating by parts with some cut-off function $\chi_{B_R} \le \psi \le \chi_{B_{2R}}$ leads to

$$\int_{B_R} |\Delta \tilde{v}_\ell| \le -\int_{\mathbb{R}^n} \Delta \tilde{v}_\ell \psi \le C_R \|\tilde{v}_\ell\|_{L^2(B_{2R}\setminus B_R)} \le C_R,$$

where in the last step we argued as in (5-6). Hence by compactness $\Delta \tilde{v}_\ell \overset{*}{\rightharpoonup} \Delta q$ in $C_c(\mathbb{R}^n)^*$. On the other hand, by (i), $\tilde{v}_\ell \to q$ locally uniformly, and so

$$\tilde{v}_\ell \Delta \tilde{v}_\ell \overset{*}{\rightharpoonup} q\Delta q \quad \text{in } C_c(\mathbb{R}^n)^*.$$

We now apply Proposition 4.2 to v_ℓ with our particular choice of γ and recall that by (4-3) we have, for $Rr_\ell < r_0$,

$$\int_{B_R} |\tilde{v}_\ell \Delta \tilde{v}_\ell| \frac{H(r_\ell, v_\ell)}{H(Rr_\ell, v_\ell) + (Rr_\ell)^{2\gamma}} \le C_R r_\ell^{\varepsilon} = o(1).$$

Notice that the constants are independent of ℓ as, by Proposition 5.4, we can choose a uniform $\tau = \tau(n,k)$ for all $x_\ell \in \Sigma^{k\text{-th}} \cap B_{1/2}$. Sending $\ell \uparrow \infty$, we get $q\Delta q = 0$. In order to show $\Delta q = 0$ outside $\{p_{2,0} = 0\}$, we exploit once again Lemma 2.1 to find

$$B_{R/2} \cap \operatorname{supp}(\Delta \tilde{v}_\ell) \subseteq \{\operatorname{dist}(\{p_{2,x_\ell} = 0\}, \cdot\,) \le C_R r_\ell^{\alpha_\circ}\},$$

and as $p_{2,x_\ell} \to p_{2,0}$ and R can be taken arbitrarily large, we deduce that $\operatorname{supp}\Delta q \subseteq \{p_{2,0} = 0\}$. It remains to show that q is nonnegative on the thin obstacle. Up to a rotation we can assume $p_{2,0} = \frac{1}{2}x_n^2$. Pick $x_* \in \{x_n = 0\}$ and consider some sequence $(y_\ell)_{\ell\in\mathbb{N}}$ such that

$$y_\ell \in \{\mathcal{A}_{k,x_\ell}(r_\ell\,\cdot\,) = 0\}, \qquad y_\ell \to x_*.$$

Thus, by locally uniform convergence and (5-5),

$$q(x_*) = \lim_\ell \tilde{v}_{r_\ell}(y_\ell) = \lim_\ell \frac{u(r_\ell y_\ell) - \frac{1}{2}\mathcal{A}^2_{k,x_\ell}(r_\ell y_\ell) + O(r_\ell^{k+2})}{H(r_\ell, v_\ell)^{1/2}} \ge 0.$$

To construct such a sequence set $y_\ell := \Phi_\ell(x_*)$, where $\Phi_\ell \in C^\infty(B_{R_\ell})$ are the inverse functions of $\Psi_\ell : x \mapsto (x', r_\ell^{-1}(\mathcal{A}_{k,x_\ell})_{r_\ell})$. Notice that $\Psi_\ell \to \mathrm{id}$ in C^1_{loc} and that $R_\ell \uparrow +\infty$ as $\ell \to \infty$. So $\Phi_\ell \to \mathrm{id}$ and $x_* \in B_{R_\ell}$ eventually, thus $y_\ell \to x_*$. Therefore

$$q \geq 0 \quad \text{on} \quad \{x_n = 0\}.$$

Hence we proved that q is a global solution of the Signorini problem (5-4).

Step 4. We show that $\tilde{v}_\ell \to q$ in $W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)$ and that q is λ_k-homogeneous. Fix any $\eta \in C^\infty_c(\mathbb{R}^n)$ and exploit as before that $\|\tilde{v}_{r_\ell} \Delta \tilde{v}_{r_\ell}\|_{L^1(B_R)} \to 0$ and integrate by parts in $\mathbb{R}^n$:

$$\begin{aligned}\int |\nabla(\eta \tilde{v}_\ell)|^2 = -\int \eta \tilde{v}_\ell \Delta(\eta \tilde{v}_\ell) &= -\int (\eta \tilde{v}_\ell^2 \Delta \eta + 2\eta \tilde{v}_\ell \nabla \eta \cdot \nabla \tilde{v}_\ell + \eta^2 \tilde{v}_\ell \Delta \tilde{v}_\ell) \\ &\leq -\int (\eta \Delta \eta \tilde{v}_\ell^2 + 2\eta \tilde{v}_\ell \nabla \eta \cdot \nabla \tilde{v}_\ell) + C(\eta) \|\tilde{v}_\ell \Delta \tilde{v}_\ell\|_{L^1(B_R)}.\end{aligned}$$

Taking the upper limit and using $\nabla \tilde{v}_\ell \rightharpoonup \nabla q$ in $L^2_{\mathrm{loc}}(\mathbb{R}^n)$ and $\tilde{v}_\ell \to q$ in $C^0_{\mathrm{loc}}(\mathbb{R}^n)$, we get

$$\limsup_\ell \int |\nabla(\eta \tilde{v}_\ell)|^2 \leq -\int (\eta \Delta \eta q^2 + 2\eta q \nabla \eta \cdot \nabla q + \eta^2 q \Delta q) = \int |\nabla(\eta q)|^2,$$

where we used $q\Delta q = 0$. By weak lower semicontinuity we always have the converse inequality, thus $\nabla(\eta \tilde{v}_\ell) \to \nabla(\eta q)$ strongly in $L^2(\mathbb{R}^n)$. This in particular gives, for every $R > 0$,

$$\phi(R, q) = \lim_\ell \phi(R, \tilde{v}_\ell) = \lim_\ell \phi(R r_\ell, v_\ell) = \lim_\ell \phi^\gamma(R r_\ell, v_\ell),$$

where in the last line we used (5-5). On the other hand, (a) in Step 1 implies

$$\lim_\ell \phi^\gamma(R r_\ell, v_\ell) = \lim_\ell f_{x_\ell}(R r_\ell) = f_0(0^+) = \lambda_k,$$

thus $\phi(R, q) \equiv \lambda_k$ for all $R > 0$. As a standard consequence, we have that q is λ_k-homogeneous; see [Athanasopoulos et al. 2008].

Step 5. We finally prove that, for $\lambda_k < k+1$, we have $q^{\mathrm{odd}} = 0$. Notice that by Proposition 2.4 q^{odd} is harmonic and thus has integral homogeneity; hence the only nontrivial case is when $\lambda_k = k$. We need to show that q is orthogonal in $L^2(\partial B_1)$ to every k-homogeneous harmonic polynomial P vanishing on $\{p_{2,0} = 0\}$. Fix such a P and apply Proposition 5.3 with

$$w_\ell := u(x_\ell + \cdot\,) - \mathcal{P}_k(p_{2,x_\ell}, \ldots, p_{k-1,x_\ell}, p_{k,x_\ell} - P \circ R_\ell),$$

where R_ℓ are rotations sending $\{p_{2,x_\ell=0}\}$ to $\{p_{2,0} = 0\}$ and $R_\ell \to \mathrm{id}$. Thus, with constants uniform in ℓ (P is fixed),

$$\begin{aligned}r^{-2k} H(r, w_\ell) + C r^\varepsilon &= C r^\varepsilon + \int_{\partial B_1} \left(\frac{v_\ell(r\,\cdot)}{r^k} + P \circ R_\ell + O\left(\frac{|x|^{k+1}}{r^k} \right) \right)^2 \\ &\geq \lim_{r \to 0} r^{-2k} H(r, w_\ell) + C r^\varepsilon = \int_{\partial B_1} P^2.\end{aligned}$$

Now divide by the sequence $\varepsilon_\ell := (H(r_\ell, v_\ell) r_\ell^{-2k})^{1/2}$, which by Step 1 (b) satisfies (recall $\lambda_k = k$)

$$r_\ell^{5\delta} \le \varepsilon_\ell^2 \le h_{x_\ell}(r_\ell) = o(1).$$

We compute the squares and rearrange the terms to get

$$\int_{\partial B_1} \tilde{v}_\ell^2 \varepsilon_\ell + 2 \int_{\partial B_1} \tilde{v}_\ell P \circ R_\ell \ge -C \frac{r_\ell^\varepsilon}{\varepsilon_\ell} \ge -C r_\ell^{\varepsilon - 5\delta/2}.$$

Since $\delta \le \frac{1}{100}\varepsilon$, we can send $\ell \to \infty$ and get

$$\int_{\partial B_1} q P \ge 0.$$

The conclusion follows by linearity in P. □

Remark 5.10. An important application of Proposition 5.9 is when the sequence x_ℓ is identically equal to 0.

Remark 5.11. Step 1 of the proof shows that the function $\Sigma^{k\text{-th}} \ni x \mapsto \phi^\gamma(u(x + \cdot) - \mathcal{P}_{k,x}, 0^+)$ is upper semicontinuous. In fact, with the same notations we have, for each $r < r_0$,

$$\limsup_\ell \phi^\gamma(u(x_\ell + \cdot) - \mathcal{P}_{k,x_\ell}, 0^+) \le \limsup_\ell \phi^\gamma(u(x_\ell + \cdot) - \mathcal{P}_{k,x_\ell}, r) + C r^\varepsilon = \phi^\gamma(u - \mathcal{P}_k, r) + C r^\varepsilon,$$

and the conclusion follows letting $r \downarrow 0$.

Proposition 5.9 shows that in order to pursue our analysis further we need to have some basic knowledge about homogeneous solutions of the Signorini Problem (5-4). In the next chapter we will use extensively the results reported in Section 2.4.

6. Estimating the size of the sets $\Sigma^{k\text{-th}} \setminus \Sigma^{(k+1)\text{-th}}$

Throughout this section u will be a solution of (2-1) with $f \equiv 1$ and $\mu = 1$. We will show that, for all $k \ge 2$,

$$\dim_{\mathcal{H}}(\Sigma^{k\text{-th}} \setminus \Sigma^{(k+1)\text{-th}}) \text{ is less than or equal to } n-2 \text{ and is countable if } n = 2.$$

In the last subsection we will show how this constrains the geometry of Σ. We remark that, by Caffarelli's analysis, $\Sigma \setminus \Sigma^{2\text{nd}}$ has locally finite $\mathcal{H}^{n-2}$ measure (see, e.g., [Caffarelli 1998, Theorem 8 (c)]).

In this chapter we repeatedly use the facts and notation concerning the Signorini problem recalled and/or established in Section 2.4. In particular, we will use $\mathcal{S}_k$, $\mathcal{S}^{\text{even}}$, q^{even}, q^{odd}, $\Sigma(q), \ldots$.

We need to understand the nature of points in $\Sigma^{k\text{-th}} \setminus \Sigma^{(k+1)\text{-th}}$. Therefore, suppose $0 \in \Sigma^{k\text{-th}}$ and $0 \notin \Sigma^{(k+1)\text{-th}}$. We necessarily have $\lambda_k \le k+1$; see Lemma 5.8. Notice that, with the notation of Proposition 5.9 and Lemma 2.2, we have

$$\frac{(u - \mathcal{P}_k)(r\,\cdot)}{r^{k+1}} = \frac{H(r, u - \mathcal{P}_k)^{1/2}}{r^{k+1}} \frac{(u - \mathcal{P}_k)(r\,\cdot)}{H(r, u - \mathcal{P}_k)^{1/2}} = O(r^{\lambda_k - (k+1)}) \tilde{v}_{r,0}.$$

As every accumulation point of $\tilde v_{r,0}$ equals some nonzero $q \in \mathcal{S}_{k+1}(\{p_2 = 0\})$ (see Proposition 5.9), in order to conclude $0 \notin \Sigma^{(k+1)\text{-th}}$ (see the flexible definition of this set from Lemma 5.7), either we must have $\lambda_k < k+1$ or $\lambda_k = k+1$ and every accumulation point q satisfies $q^{\text{even}} \not\equiv 0$.

These observations inspire the following trichotomy. If $x \in \Sigma^{k\text{-th}} \setminus \Sigma^{(k+1)\text{-th}}$ then exactly one of the following happens:

(1) $\lambda_k(x) = k$,

(2) $\lambda_k(x) \in (k, k+1)$,

(3) $\lambda_k(x) = k+1$, but every accumulation point of $r^{-(k+1)}(u(x+\cdot) - \mathcal{P}_{k,x})(r\,\cdot)$ has a nonzero even part.

We rephrase these cases with a notation closer to that adopted in [Figalli and Serra 2019; Figalli et al. 2020]. Namely, for each $k \ge 2$, define

$$\Sigma^{>k} := \Sigma^{k\text{-th}} \cap \{\lambda_k > k\}, \qquad \Sigma^{\ge k+1} := \Sigma^{k\text{-th}} \cap \{\lambda_k \ge k+1\}.$$

So we have the descending chain of inclusions

$$\Sigma_{n-1} = \Sigma^{2\text{nd}} = \Sigma^{>2} \supseteq \cdots \supseteq \Sigma^{k\text{-th}} \supseteq \Sigma^{>k} \supseteq \Sigma^{\ge k+1} \supseteq \Sigma^{(k+1)\text{-th}} \supseteq \cdots \supseteq \bigcap_{j\ge 2} \Sigma^{j\text{-th}} =: \Sigma^{\infty}.$$

With this notation case (1) corresponds to the set $\Sigma^{k\text{-th}} \setminus \Sigma^{>k}$, case (2) to $\Sigma^{>k} \setminus \Sigma^{\ge k+1}$ and case (3) to $\Sigma^{\ge k+1} \setminus \Sigma^{(k+1)\text{-th}}$. In the next subsections we will address separately each case.

We point out that, in cases (1) and (3), the parity of k will play a role in our arguments. This is related to the different shape of the functions in $\mathcal{S}_k^{\text{even}}$ according to the parity of k (see Proposition 2.4).

6.1. ***The size of*** $\boldsymbol{\Sigma^{k\text{-th}} \setminus \Sigma^{>k}}$**.** We start by showing that when k is even, the set $\Sigma^{k\text{-th}} \setminus \Sigma^{>k}$ is in fact empty. This is a simple consequence of the following monotonicity formula, which is an extension of [Figalli et al. 2020, Lemma 4.14] to higher values of k.

Lemma 6.1. *Let u solve (2-1), and let $0 \in \Sigma^{k\text{-th}}$ for some $k \ge 2$. Let $v := u - \mathcal{P}_k$, and let P be any k-homogeneous harmonic polynomial such that $P \ge 0$ on $\{p_2 = 0\}$. Then there exists $\varepsilon, r_0 > 0$ depending on n and k such that, for all $r \in (0, r_0)$,*

$$\frac{d}{dr}\left(r^{-k}\int_{\partial B_1} v_r P\right) \le C r^{\varepsilon - 1}$$

for some constant C depending only on n, k, $\|P\|_{L^2(\partial B_1)}$.

Proof. The proof is identical to in [Figalli et al. 2020, Lemma 4.14]; we give it nevertheless for the reader's convenience. Integration by parts and the fact that P is harmonic lead to

$$\frac{d}{dr}\int_{\partial B_1} v_r P = \frac{1}{r}\left(\int_{\partial B_1} v_r \partial_\nu P + \int_{B_1} \Delta v_r P\right) = \frac{1}{r}\left(k\int_{\partial B_1} v_r q + \int_{B_1} \Delta v_r P\right),$$

where we used the homogeneity of P to deduce that $\partial_\nu P = kP$ on ∂B_1. As $\Delta v_r = -r^2 \chi_{\{u_r = 0\}}$, we can rewrite this as

$$\frac{d}{dr}\left(r^{-k}\int_{\partial B_1} v_r P\right) = -\frac{1}{r^{k-1}}\int_{B_1 \cap \{u_r = 0\}} P.$$

We have

$$r^{-k}\|v_r\|_{L^2(B_1)} = \|\tilde{v}_r\|_{L^2(B_1)}(r^{-2k}H(r,v))^{1/2} \le C(n,k)$$

for $r \lesssim 1$ sufficiently small, thanks to Proposition 5.9 (i) and Proposition 5.3. Combining this estimate with the Lipschitz bounds from Proposition 3.4 (ii), with $\beta \in (0, 1/(k+2))$ to be chosen, we find

$$\{u_r = 0\} \cap B_1 \subseteq \{x \in B_1 : r|\partial_n \mathcal{P}_k|(rx) = r^2|x_n + O(|x|^2)| \le Cr^{k(1-\beta)}\},$$

with some constant C depending on n and k only (as we can choose $\tau = \tau(n,k)$ from Proposition 5.4). This shows that $|\{u_r = 0\} \cap B_1| \le Cr^{k(1-\beta)-2}$. On the other hand, by the maximum principle, we have $-P \le C|x_n|$ in B_1. Hence, using Lemma 2.1, we obtain

$$-\int_{B_1\cap\{u_r=0\}} P \le Cr^{\alpha_\circ}|\{u_r = 0\} \cap B_1| \le Cr^{k+\alpha_\circ-k\beta-2},$$

and the lemma follows choosing $\beta = \alpha_\circ/(2k)$. □

As a simple corollary we get our claim.

Corollary 6.2. *For every even integer $k \ge 2$, we have $\Sigma^{k\text{-th}} \setminus \Sigma^{>k} = \varnothing$.*

Proof. Let us assume, on the contrary, that $0 \in \Sigma^{k\text{-th}} \setminus \Sigma^{>k}$, that is $\lambda_k = k$. Then, by Proposition 5.9, any accumulation point q of $\tilde{v}_r = (u - \mathcal{P}_k)_r / H(r, u - \mathcal{P}_k)^{1/2}$ lies in $\mathcal{S}_k^{\text{even}}(\{p_2 = 0\}) \setminus \{0\}$. Furthermore, by Proposition 2.4, any such q satisfies the assumptions of Lemma 6.1. As $0 \in \Sigma^{k\text{-th}}$, we moreover have $v_r/r^k \to 0$, thus after combining this with Lemma 6.1 with $P = q$, we find

$$r^{-k}\int_{\partial B_1} v_r q \le Cr^{\varepsilon}$$

for small $r \le 1$. Dividing by $H(r, u - \mathcal{P}_k)^{1/2}$ leads to

$$\int_{\partial B_1} \tilde{v}_{r_\ell} q \le C\frac{r_\ell^{k+\varepsilon}}{H(r_\ell, u - \mathcal{P}_k)^{1/2}}.$$

Thanks to (ii) in Remark 5.2, we deduce that the right-hand side vanishes as $\ell \uparrow \infty$, implying $\int_{\partial B_1} q^2 \le 0$ and contradicting $\|q\|_{L^2(\partial B_1)} = 1$. □

Let us now consider an odd k. We point out that, for $k = 3$, we still have $\Sigma^{3\text{rd}} \setminus \Sigma^{\ge 3} = \varnothing$, but the proof is more refined; see [Figalli et al. 2020, Proposition 5.8]. We will instead rely on a more robust argument which will be also employed later to deal with the case $\lambda_k = k+1$ (see Lemma 6.10). The main step is contained in the following lemma, based on a barrier argument.

Lemma 6.3. *Let $k \ge 3$ be odd. For all $x \in \Sigma^{k\text{-th}} \setminus \Sigma^{>k}$ and $\varepsilon > 0$, there is $\varrho = \varrho(\varepsilon, x) > 0$ such that, for each $0 < r < \varrho$, there exists $q \in \mathcal{S}_k^{\text{even}}(\{p_{2,x} = 0\})$ such that*

$$\Sigma(u) \cap B_r(x) \subseteq \Sigma(q) + B_{\varepsilon r}(x). \tag{6-1}$$

Recall that $\Sigma(q) := \{q = |\nabla q| = 0\} \cap \{p_{2,x} = 0\}$ was defined in (2-9).

Proof. Up to an isometry, suppose $x = 0$ and $p_2 = \frac{1}{2}x_n^2$. We argue by contradiction and (rescaling the space) suppose that there are $\varepsilon_\circ > 0$ and $r_\ell \downarrow 0$ such that

$$\Sigma(u_{r_\ell}) \cap \{y \in B_1 : \operatorname{dist}(y, \Sigma(q)) > \varepsilon_\circ\} \neq \varnothing \quad \text{for all } q \in \mathcal{S}_k^{\text{even}}.$$

Thanks to Proposition 5.9, we can extract a subsequence (that we do not rename) such that

$$\frac{(u - \mathcal{P}_k)_{r_\ell}}{H(r_\ell, u - \mathcal{P}_k)^{1/2}} \to \bar{q} \in \mathcal{S}_k^{\text{even}} \setminus \{0\} \quad \text{in } C^0_{\text{loc}}(\mathbb{R}^n). \tag{6-2}$$

Thus, there are $y_\ell \in B_1$ such that

$$y_\ell \in \Sigma(u_{r_\ell}), \quad \operatorname{dist}(y_\ell, \Sigma(\bar{q})) \geq \varepsilon_\circ, \quad y_\ell \to y_\infty \in \{x_n = 0\}.$$

By Proposition 2.4 we can write $\bar{q}(x) = -|x_n|(q_0(x') + x_n^2 q_1(x))$ for some polynomials q_0 and q_1, with $q_0 \geq 0$. For brevity, we set $h_\ell := H(r_\ell, u - \mathcal{P}_k)^{1/2}$ and remark that $r_\ell^{k+\delta} \leq C_0 h_\ell \ll r_\ell^k$ for some constants C_0 and δ (see Remark 5.2).

As $y_\ell \to y_\infty$, in order to reach a contradiction, it suffices to show that, for some small radius $R > 0$ and all ℓ large,

$$\Sigma(u_{r_\ell}) \cap B_R(y_\infty) = \varnothing. \tag{6-3}$$

The rest of the proof is devoted to showing (6-3) for a suitable R independent of ℓ.

We start by choosing a radius ρ as follows. As $y_\infty \in \{x_n = 0\} \setminus \Sigma(\bar{q}) \subseteq \{q_0 > 0\}$, we can find some small $\rho, m \in (0, 1)$ such that

$$q(x) \leq -m|x_n| \quad \text{for } x \in B_{4\rho}(y_\infty), \tag{6-4}$$

and we can also require that $\rho \leq \frac{1}{100}m$.

Let us introduce some notation. Define the set of points that admit a barrier as

$$Z_\ell := \{z \in B_\rho(y_\infty) : \exists\, \phi_{z,\ell} \text{ of class } C^2 \text{ in a neighborhood of } B_\rho(z) \text{ solving (6-5)}\},$$

where

$$\begin{cases} \phi_{z,\ell}(z) = 0, & \\ \phi_{z,\ell} \geq 0 & \text{in } B_\rho(z), \\ \Delta\phi_{z,\ell} < r_\ell^2 & \text{in } \overline{B_\rho(z)}, \\ u(r_\ell\,\cdot\,) < \phi_{z,\ell} & \text{on } \partial B_\rho(z). \end{cases} \tag{6-5}$$

We also set

$$\Gamma_\ell := \{\mathcal{A}_k(r_\ell\,\cdot\,) = 0\} \cap B_2.$$

For ℓ large in terms of n, k and τ, we have that Γ_ℓ is a smooth hypersurface inside B_2, which converges to the hyperplane $\{x_n = 0\}$, say, in the C^2 norm. Furthermore, combining Lemma 2.1 and the fact that $\mathcal{A}_k(x) = x_n + O(|x|^2)$, we get

$$(\{u_{r_\ell} = 0\} \cup \Gamma_\ell) \cap B_2 \subseteq \{|x_n| \leq C_1 r_\ell^{\alpha_\circ}\} \cap B_2 \tag{6-6}$$

for some constant $C_1 = C_1(n, k) > 0$.

Claim. There is ℓ_0 such that, for $\ell > \ell_0$, the following hold:

(i) All points in $B_\rho(y_\infty)$ belonging to the hypersurface Γ_ℓ admit a barrier, that is, $\Gamma_\ell \cap B_\rho(y_\infty) \subseteq Z_\ell$.

(ii) The function u vanishes on Z_ℓ. Moreover, Z_ℓ is open and contained in the interior of the contact set $\{u(r_\ell\,\cdot\,) = 0\}$.

(iii) There is a dimensional constant $N(n) > 0$ such that $(\Sigma(u_{r_\ell}) \cap B_{\rho/2N}(y_\infty)) \setminus \Gamma_\ell = \varnothing$.

The combination of these three claims will give (6-3) with $R = \rho/(2N)$. In fact, as there are no singular points in the interior of the singular set, (i) and (ii) give $\Sigma(u_{r_\ell}) \cap B_{\rho/2N}(y_\infty) \cap \Gamma_\ell = \varnothing$. We conclude with (iii). □

Proof of the Claim. We begin by proving (ii). First, for any $z \in Z_\ell$ and any ξ close to z, we can define

$$\phi_{\xi,\ell}(x) := \phi_{z,\ell}(x + (z - \xi)).$$

By continuity of translations, $\phi_{\xi,\ell}$ solves (6-5) on $B_\rho(\xi)$ for $|\xi - z|$ small enough; hence Z_ℓ is open. We now apply the comparison principle, using the last two properties of the barrier in (6-5) to find $u|_{Z_\ell} \equiv 0$. Notice that, for $z \in Z_\ell$, two cases arise: either for all $c > 0$ we have $u(r_\ell\,\cdot\,) < \phi_{z,\ell} + c$ on $B_\rho(z)$, or there exist the largest $c_* > 0$ such that $u(r_\ell x_*) = \phi_{z,\ell}(x_*) + c_*$ for some $x_* \in \overline{B_\rho(z)}$. In the first case, evaluating at z and sending $c \downarrow 0$, we get $u(r_\ell z) = 0$. In the second case, we notice that by (6-5) we must have $x_* \notin \partial B_\rho(z)$, thus we get

$$r_\ell^2 \chi_{\{u(r_\ell\cdot)>0\}}(x_*) = \Delta u_{r_\ell}(x_*) \le \Delta\phi_{z,\ell}(x_*) < r_\ell^2 \quad \Rightarrow \quad x_* \notin \{u(r_\ell\,\cdot\,) > 0\}.$$

Then $0 = u(r_\ell x_*) = \phi_{z,\ell}(x_*) + c_* \ge c_* > 0$, a contradiction.

Next, we turn to the proof of (iii). First recall that there exists a dimensional modulus of continuity (see [Caffarelli 1998, Theorem 8 and Corollary 11]) such that, for all $x \in \Sigma(u)$, we have

$$\|u(x + \,\cdot\,) - p_{2,x}\|_{L^\infty(B_r)} \le r^2\omega(r).$$

Suppose by contradiction that we can find

$$y_\ell^* \in (\Sigma(u_{r_\ell}) \cap B_{\rho/2N}(y_\infty)) \setminus \Gamma_\ell$$

for arbitrarily large ℓ. We set for brevity $p_\ell := p_{2,r_\ell y_\ell^*}$. As p_ℓ is convex and $\Delta p_\ell \ge 1$, for every such ℓ, we choose a unit vector e_ℓ such that

$$p_\ell(x) \ge \frac{1}{2n}(e_\ell \cdot x)^2.$$

Now by item (ii), we have, for all $z \in B_{\rho/N}(y_\infty) \cap \Gamma_\ell$,

$$\begin{aligned} 0 = u(r_\ell z) = u_{r_\ell}(y_\ell^* + (z - y_\ell^*)) &\ge p_\ell(z - y_\ell^*) - |z - y_\ell^*|^2\omega(|z - y_\ell^*|) \\ &\ge \frac{1}{2n}(e_\ell \cdot (z - y_\ell^*))^2 - |z - y_\ell^*|^2\omega(|z - y_\ell^*|). \end{aligned}$$

From this inequality we reach a contradiction: On one hand we have $|z - y_\ell^*| \le 1/N$, so we can require N to be large enough that $\omega(|z - y_\ell^*|) \le 1/(100n)$. On the other hand we can choose z in such a way that the nonzero vector $z - y_\ell^*$ is almost aligned with e_ℓ, thus we get a contradiction dividing by $|z - y_\ell^*|^2$.

In order to prove (i) we need to construct a barrier for all $z \in B_\rho(y_\infty) \cap \Gamma_\ell$. Set

$$\phi_{z,\ell}(x) := \left(1 - \frac{h_\ell}{r_\ell^2}\right)\frac{1}{2}\mathcal{A}_k^2(r_\ell x) + \frac{h_\ell}{4n}|x' - z'|^2.$$

We have to check that $\phi_{z,\ell}$ indeed satisfies (6-5) for ℓ large. The first two equations in (6-5) are clearly fulfilled for any ℓ. For the third condition, let us compute, for $x \in B_\rho(z)$,

$$\Delta\phi_{z,\ell} = r_\ell^2 - h_\ell + C_2 r_\ell^{k+2} + \frac{2(n-1)}{4n}h_\ell \le r_\ell^2 - \tfrac{1}{2}h_\ell + C_0 C_2 r_\ell h_\ell,$$

where we used, for some $C_2 = C_2(n,k)$, that $\Delta\frac{1}{2}\mathcal{A}_k^2 \le 1 + C_2|x|^k$ and $r_\ell^{k+1} \le C_0 h_\ell$. Hence, $\Delta\phi_{z,\ell} < r_\ell^2$ as soon as ℓ is large enough.

We turn to the last condition of (6-5). For any fixed $\eta \le \rho^2/(100n)$, we have, by uniform convergence of (6-2) for ℓ large,

$$u_{r_\ell} \le \tfrac{1}{2}\mathcal{A}_k^2(r_\ell\,\cdot\,) + h_\ell\bar{q} + h_\ell\eta$$

in B_2. As for some constant $C_3 = C_3(n,k) > 0$, we have, in B_2,

$$\tfrac{1}{2}\mathcal{A}_k^2(x) \le \tfrac{1}{2}x_n^2 + C_3|x|^3,$$

and recalling the choice of ρ from (6-4), we get, for $x \in \overline{B_\rho(z)}$,

$$\begin{aligned} u(r_\ell x) &\le \tfrac{1}{2}\mathcal{A}_k^2(r_\ell x) - h_\ell m|x_n| + h_\ell\eta \\ &\le \left(1 - \frac{h_\ell}{r_\ell^2}\right)\frac{1}{2}\mathcal{A}_k^2(r_\ell x) + \frac{h_\ell}{2}x_n^2 - h_\ell m|x_n| + C_3 h_\ell r_\ell|x|^3 + h_\ell\eta \\ &= \phi_{z,\ell}(x) + h_\ell\left(\frac{1}{2}x_n^2 - m|x_n| + C_3 r_\ell|x|^3 + \eta - \frac{1}{4n}|x' - z'|^2\right). \end{aligned}$$

We show that, whenever $x \in \partial B_\rho(z)$, the term between parentheses is negative. Using (6-6) and the fact that $|x| \le 2$, we get, for all $x \in \partial B_\rho(z)$,

$$\begin{aligned} \frac{1}{2}x_n^2 - m|x_n| + C_3 r_\ell|x|^3 + \eta - \frac{1}{4n}|x' - z'|^2 &= \frac{1}{4n}|x_n - z_n|^2 + \frac{1}{2}x_n^2 - m|x_n| + C_3 r_\ell|x|^3 + \eta - \frac{\rho^2}{4n} \\ &\le x_n^2 - m|x_n| + C_1^2 r_\ell^{2\alpha_\circ} + 8C_3 r_\ell + \eta - \frac{\rho^2}{4n}. \end{aligned}$$

We claim that this quantity is negative as soon as

$$r_\ell^{\alpha_\circ} \le \min\left\{\frac{\rho}{C_1}, \frac{\eta}{8C_3}, \frac{\eta}{C_1^2}, 1\right\}.$$

In fact by (6-6), we have, uniformly in z,

$$|x_n| \le |x_n - z_n| + |z_n| \le \rho + C_1 r_\ell^{\alpha_\circ} \le 2\rho \le \frac{2m}{100},$$

thus $x_n^2 - m|x_n| \le 0$ and

$$C_1 r_\ell^2 + 8C_3 r_\ell + \eta - \frac{\rho^2}{4n} \le 3\eta - \frac{\rho^2}{4n} \le \frac{3\rho^2}{100n} - \frac{\rho^2}{4n} < 0.$$

□

Exploiting some recent volume estimates for the tubular neighborhood of the critical set of harmonic functions from [Naber and Valtorta 2017], we can now deduce the following.

Lemma 6.4. *Given $\beta_1 > n-2$ and $k \geq 3$ odd, there exists an $\hat{\varepsilon} = \hat{\varepsilon}(n, \beta_1)$ small such that the following holds. Let $E \subseteq \mathbb{R}^n$ be any set satisfying*

$$E \subseteq B_r(x) \cap (\Sigma(q) + B_{\hat{\varepsilon}r}(x))$$

for some $r \in (0, 1)$, $x \in \mathbb{R}^n$ and $q \in \mathcal{S}_k^{\text{even}}(L) \setminus \{0\}$ for some hyperplane L. Then E can be covered with $\lfloor \gamma^{-\beta_1} \rfloor$ balls of radius γr centered at points of E for $\gamma = \frac{1}{5}\hat{\varepsilon}$.

Proof. By translation and scaling we can recover the general case from the case $r = 1$, $x = 0$. Let $\hat{\varepsilon} \in (0, 1)$ be a parameter to be fixed later, and take q as in the statement, recalling that q vanishes on L (see Proposition 2.4). For simplicity we assume $L = \{x_n = 0\}$ and consider Q, the odd (with respect to L) extension of $q|_{\{x_n>0\}}$ to $\mathbb{R}^n$. Q is harmonic, and it is easily checked that $\Sigma(q) \subseteq \{Q = |\nabla Q| = 0\} =: \Sigma(Q)$; hence a fortiori

$$E \subseteq B_1 \cap \{\text{dist}(\,\cdot\,, \Sigma(Q)) \leq \hat{\varepsilon}\}.$$

As Q is harmonic and nonzero, we can apply the volume estimates in [Naber and Valtorta 2017, Theorem 1.1] to find

$$\mathcal{H}^n(B_2 \cap \{\text{dist}(\,\cdot\,, \Sigma(Q)) \leq t\}) \leq C(n) t^2$$

for all $t \in (0, 1)$. Now, consider a covering of E of the form $\{B_{\hat{\varepsilon}}(x)\}_{x \in E}$. By Vitali's covering lemma, there exists a disjoint subcollection $\{B_{\hat{\varepsilon}}(x_i)\}_{i \in I}$ such that

$$E \subseteq \bigcup_{x \in E} \overline{B_{\hat{\varepsilon}}(x)} \subseteq \bigcup_{i \in I} B_{5\hat{\varepsilon}}(x_i).$$

We need to estimate the cardinality of I. Denoting by ω_n the volume of the unit ball in $\mathbb{R}^n$ and using that $B_{\hat{\varepsilon}}(x_i) \subseteq (E + B_{\hat{\varepsilon}} \subseteq \Sigma(Q) + B_{2\hat{\varepsilon}}) \cap B_2$, we have

$$\omega_n \hat{\varepsilon}^n \# I = \mathcal{H}^n\Bigl(\bigcup_{i \in I} \overline{B_{\hat{\varepsilon}}(x_i)}\Bigr) \leq \mathcal{H}^n(\{\text{dist}(\,\cdot\,, \Sigma(Q)) \leq 2\hat{\varepsilon}\}) \leq C(n)\hat{\varepsilon}^2,$$

and thus $\# I \leq C(n)\hat{\varepsilon}^{2-n}$. As $\beta_1 > n-2$, choosing $\hat{\varepsilon}(n, \beta_1)$ small enough, we find $\# I \leq \bigl(\frac{1}{5}\hat{\varepsilon}\bigr)^{-\beta_1}$, which finishes the proof. □

We employ Lemma 6.4 to get a Reifenberg-type result. We need to incorporate the lower-semicontinuous function τ into the statement, as we will use this result in the next section.

Proposition 6.5 [Figalli et al. 2020, Proposition 7.5]. *Let $\tau : E \to \mathbb{R}$ be a lower-semicontinuous function and $E \subseteq \mathbb{R}^n$ be a measurable set with the following property. For any $\varepsilon > 0$ and $x \in E$, there exists $\varrho = \varrho(x, \varepsilon) > 0$ such that, for all $r \in (0, \varrho)$, there exist a hyperplane L, an odd integer $k \geq 3$ and $q \in \mathcal{S}_k^{\text{even}}(L) \setminus \{0\}$ such that*

$$E \cap \overline{B_r(x)} \cap \tau^{-1}([\tau(x), +\infty)) \subseteq \Sigma(q) + \overline{B_{\varepsilon r}(x)}.$$

Then $\dim_{\mathcal{H}}(E) \leq n-2$.

Proof. The result follows by iterating Lemma 6.4, and we skip the details. See for example the proof of Propositions 7.3 or 7.5 in [Figalli et al. 2020]. □

This finally gives the desired dimensional estimate.

Corollary 6.6. *Let $k \geq 3$ be odd. Then $\dim_{\mathcal{H}}(\Sigma^{k\text{-}th} \setminus \Sigma^{>k}) \leq n-2$. Furthermore, if $n=2$, then $\Sigma^{k\text{-}th} \setminus \Sigma^{>k}$ is discrete in Σ.*

Proof. Recall that if $n=2$, then $\Sigma(q)=\{0\}$ for every $q \in \mathcal{S}_k^{\text{even}} \setminus \{0\}$. Pick any $x \in \Sigma^{k\text{-th}} \setminus \Sigma^{>k}$ and apply Lemma 6.3 with $\varepsilon := \frac{1}{2}$. Then, for all $r < \varrho\left(x, \frac{1}{2}\right)$, we have

$$\Sigma \cap (B_r(x) \setminus \overline{B_{r/2}(x)}) = \varnothing.$$

This clearly implies that $\Sigma \cap B_\varrho(x) = \{x\}$, thus x is isolated in Σ.

For the case $n \geq 3$ we apply Proposition 6.5 to $E := \Sigma^{k\text{-th}} \setminus \Sigma^{>k}$ and the constant function $\tau \equiv 1$. The hypothesis are satisfied thanks to Lemma 6.3. □

6.2. *The size of* $\Sigma^{>k} \setminus \Sigma^{\geq k+1}$. The key idea behind this dimensional reduction is that at an accumulation point of $\Sigma^{>k} \setminus \Sigma^{\geq k+1}$, the blowup gains a translation symmetry along the direction of the approaching sequence. This observation corresponds to Lemmas 6.8 or 6.9 in [Figalli et al. 2020].

Lemma 6.7. *Let $0 \in \Sigma^{>k} \setminus \Sigma^{\geq k+1}$ for some $k \geq 2$. Suppose there exists an infinitesimal sequence $r_\ell \downarrow 0$ and points $x_\ell \in \Sigma^{k\text{-}th} \cap B_{r_\ell}$, $x_\ell \neq 0$ such that $\lambda_k(x_\ell) \to \lambda_k$. Assume further that, as $\ell \to \infty$, we have*

(i) $x_\ell / r_\ell \to y_\infty \in \bar{B}_1$,

(ii) $\tilde{v}_{r_\ell} = (u - \mathcal{P}_k)_{r_\ell} / H(r_\ell, u - \mathcal{P}_k)^{1/2} \to q$ *in $C^0_{\text{loc}}(\mathbb{R}^n)$ for some $q \in \mathcal{S}_{\lambda_k}(\{p_2 = 0\}) \setminus \{0\}$.*

Then $y_\infty \in \{p_2 = 0\}$ and $q = q(y_\infty + \cdot\,)$.

Proof. Consider a sequence $(x_\ell)_{\ell \in \mathbb{N}} \subseteq \Sigma^{k\text{-th}} \cap B_{r_\ell}$ as in the statement of the lemma. We begin by recalling that $y_\infty \in \{p_2 = 0\}$ because $r_\ell^{-2} u(r_\ell \cdot) \to p_2$ uniformly in B_2. Then we apply Proposition 5.9 with varying centers $(x_\ell)_{\ell \in \mathbb{N}}$, and after passing to a subsequence (denoted again with r_ℓ) we have

$$\tilde{v}_{r_\ell, x_\ell} := \frac{u(x_\ell + r_\ell \cdot\,) - \mathcal{P}_{k, x_\ell}(r_\ell \cdot\,)}{\|(u(x_\ell + \cdot\,) - \mathcal{P}_{k, x_\ell})_{r_\ell}\|_{L^2(\partial B_1)}} \to Q$$

in $C^0_{\text{loc}}(\mathbb{R}^n)$ for some $Q \in \mathcal{S}_{\lambda_k}(\{p_2 = 0\}) \setminus \{0\}$. On the other hand, by uniform convergence,

$$q(y_\infty + \cdot\,) = \lim_\ell \tilde{v}_{r_\ell}\left(\frac{x_\ell}{r_\ell} + \cdot\right) = \lim_\ell \frac{u(x_\ell + r_\ell \cdot\,) - \mathcal{P}_k(x_\ell + r_\ell \cdot\,)}{H(r_\ell, u - \mathcal{P}_k)^{1/2}}.$$

So putting everything together we can write

$$\tilde{v}_{r_\ell}\left(\frac{x_\ell}{r_\ell} + \cdot\right) = \tilde{v}_{r_\ell, x_\ell} \cdot I_\ell + J_\ell, \tag{6-7}$$

where

$$I_\ell := \frac{H(r_\ell, u(x_\ell + \cdot\,) - \mathcal{P}_{k, x_\ell}(\cdot\,))^{1/2}}{H(r_\ell, u - \mathcal{P}_k)^{1/2}} = \frac{\|u(x_\ell + r_\ell \cdot\,) - \mathcal{P}_{k, x_\ell}(r_\ell \cdot\,)\|_{L^2(\partial B_1)}}{\|u(r_\ell \cdot\,) - \mathcal{P}_k(r_\ell \cdot\,)\|_{L^2(\partial B_1)}}$$

is a numerical sequence and

$$J_\ell := \frac{\mathcal{P}_{k,x_\ell}(r_\ell\,\cdot\,) - \mathcal{P}_k(x_\ell + r_\ell\,\cdot\,)}{H(r_\ell, u - \mathcal{P}_k)^{1/2}}$$

is a sequence of harmonic polynomials of degree at most $k+1$. Now two cases arise:

$$\text{either} \quad \sup_\ell I_\ell < \infty \quad \text{or} \quad I_{\ell_m} \uparrow \infty \quad \text{for some subsequence} \quad \ell_m \to \infty.$$

Let us begin with the first case. Up to a subsequence that we do not rename, we have $I_\ell \to \alpha$ for some $\alpha \geq 0$. Equation (6-7) then implies that $J_\ell \to J$ locally uniformly to some harmonic polynomial J of degree at most $k+1$. Thus, sending $\ell \uparrow \infty$ in (6-7), we obtain

$$q(y_\infty + \cdot\,) = \alpha Q + J. \tag{6-8}$$

We exploit homogeneity: for large $R > 0$, we have

$$R^{\lambda_k} q\Big(\frac{y_\infty}{R} + \cdot\Big) = R^{\lambda_k}\alpha Q + J(R\,\cdot\,),$$

so $\lim_{R\uparrow\infty} R^{-\lambda_k} J(Rx)$ exists for every $x \in \mathbb{R}^n$. As λ_k is not an integer and J is a polynomial, the only possibility is that $\lim_{R\uparrow\infty} R^{-\lambda_k} J(Rx) = 0$ for all x, and so $\deg J \leq k$. Hence the last identity reads

$$q = \alpha Q.$$

Inserting this back in (6-8), we find $q(y_\infty + \cdot\,) = q + J$. Now, using again that q is homogeneous, for any $R > 0$, we have

$$R\Big(q\Big(\frac{y_\infty}{R} + \cdot\Big) - q(\,\cdot\,)\Big) = R^{1-\lambda_k} J(R\,\cdot\,).$$

Sending $R \to \infty$, the left-hand side converges to $y_\infty \cdot \nabla q$, but as before the right-hand side can only converge to 0, so $y_\infty \cdot \nabla q = 0$.

The second case is simpler. We divide (6-7) by I_{ℓ_m} and find, after passing to a subsequence of ℓ_m, that

$$0 = Q + \tilde{J}$$

for some harmonic polynomial $\tilde{J}$ of degree at most $k+1$. This is a contradiction, since $Q \neq 0$ is a $\lambda_k \in (k, k+1)$ homogeneous function and hence not a polynomial. This finishes the proof. ☐

Lemma 6.7 triggers a Federer-type dimension reduction, exactly as in [Figalli and Serra 2019].

Proposition 6.8 [Figalli et al. 2020, Proposition 7.3]. *Let $E \subseteq \mathbb{R}^n$, $f : E \to \mathbb{R}$ and $m \in \{1, \dots, n\}$. Assume that, for any $\varepsilon > 0$ and $x \in E$, there exists $\varrho = \varrho(x, \varepsilon) > 0$ such that, for all $r \in (0, \varrho)$, we have*

$$E \cap \overline{B_r(x)} \cap f^{-1}([f(x) - \varrho, f(x) + \varrho]) \subseteq \Pi_{x,r} + B_{\varepsilon r}$$

for some m-dimensional plane $\Pi_{x,r}$ passing through x (possibly depending on r). Then $\dim_{\mathcal{H}}(E) \leq m$.

We combine Lemma 6.7 and Proposition 6.8 to prove the dimensional estimate.

Proposition 6.9. *For every $k \geq 2$, we have $\dim_{\mathcal{H}}(\Sigma^{>k} \setminus \Sigma^{\geq k+1}) \leq n-2$. Moreover, if $n = 2$, then $\Sigma^{>k} \setminus \Sigma^{\geq k+1}$ consists of isolated points if k is odd and is empty if k is even.*

Proof. We want to apply Proposition 6.8 with $E := \Sigma^{>k} \setminus \Sigma^{\geq k+1}$, $m = n-2$ and the function f given on E by $x \mapsto \lambda_k(x) \in (k, k+1)$. It suffices to show that, for all $x \in E$ and for all $\varepsilon > 0$, there exist $\varrho = \varrho(x, \varepsilon) > 0$ and an $(n-2)$-dimensional plane $\Pi_{x,r}$ passing through x such that

$$E \cap B_r(x_\circ) \cap \lambda_k^{-1}([\lambda_k(x) - \varrho, \lambda_k(x) + \varrho]) \subseteq \{x : \operatorname{dist}(x, \Pi_{x,r}) \leq \varepsilon r\} \quad \text{for all } r \in (0, \varrho).$$

We argue by contradiction. Assume that, for $x = 0$ and some $\varepsilon_\circ > 0$, the above does not hold. Then we make the following simple geometric claim. For each ℓ there exists $r_\ell \in (0, 2^{-\ell})$ and $n-1$ points $x_\ell^{(1)}, \ldots, x_\ell^{(n-1)}$ in $E \cap B_{r_\ell}$ such that

$$|x_\ell^{(1)} \wedge \cdots \wedge x_\ell^{(n-1)}| \geq \delta r_\ell^{n-1}, \quad |\lambda_k(x_\ell^{(j)}) - \lambda_k| \leq 2^{-\ell}$$

for all $j \in \{1, \ldots, n-1\}$ and for some $\delta = \delta(n, \varepsilon_\circ) \in (0, 1)$. In particular, $\{x_\ell^{(1)}, \ldots, x_\ell^{(n-1)}\}$ span a hyperplane and, for each fixed j, the sequence $(x_\ell^{(j)})_{\ell \in \mathbb{N}}$ lies in $E \subseteq \Sigma^{k\text{-th}}$, with

$$\lambda_k(x_\ell^{(j)}) \to \lambda_k.$$

We extract a finite number of subsequences to ensure $x_\ell^{(j)}/r_\ell \to y_\infty^{(j)}$ for each j. Exploiting the lower bound on the exterior product, we again have that

$$\dim \operatorname{span}\{y_\infty^{(1)}, \ldots, y_\infty^{(n-1)}\} = n-1.$$

Now we apply Proposition 5.9 to each $(x_\ell^{(j)})_{\ell \in \mathbb{N}}$ and get

$$\tilde{v}_{r_\ell} = \frac{(u - \mathcal{P}_k)_{r_\ell}}{H(r_\ell, u - \mathcal{P}_k)^{1/2}} \to q \quad \text{in } C^0_{\text{loc}}$$

for some $q \in \mathcal{S}_{\lambda_k}$. Notice that, taking at each time a subsequence, q can be taken the same for all j's. By Lemma 6.7, we conclude that q is translation-invariant in the directions $y_\infty^{(j)}$ for all $1 \leq j \leq n-1$; hence q is a 1-dimensional homogeneous solution to the obstacle problem vanishing at the origin. Thus, after a rotation of coordinates, we must have $q(x) = -A|x_n| + Bx_n$ for some constants $A \geq 0$ and $B \in \mathbb{R}$, which contradicts $\lambda_k > 1$.

Let us sketch the geometric argument needed to construct such $\{x_\ell^{(1)}, \ldots, x_\ell^{(n-1)}\}$. Fixing ℓ, we pick any $(n-2)$-plane Π_0 and any $x^{(1)} \in (E \cap B_{r_\ell}) \setminus (\Pi_0 + B_{\varepsilon_\circ})$. Then we choose any plane Π_1 containing $x^{(1)}$ and any $x^{(2)}$ in $(E \cap B_{r_\ell}) \setminus (\Pi_1 + B_{\varepsilon_\circ})$. We can go on in this way and construct the whole set $\{x^{(1)}, \ldots, x^{(n-1)}\}$. Finally, we notice that, by compactness,

$$\delta := \min\{|z_1 \wedge \cdots \wedge z_{n-1}| : z_j \in \mathbb{R}^n, \forall j \ \operatorname{dist}(z_j, \operatorname{span}\{z_i : i \neq j, 1 \leq i \leq n-1\}) \geq \varepsilon_\circ\} > 0.$$

We are left with the case $n = 2$. Recall that if q is a λ-homogeneous solution to the thin obstacle problem, then $\lambda \in \mathbb{N}_+ \cup \left\{2m - \frac{1}{2} : m \in \mathbb{N}_+\right\}$ (see Proposition 2.4). Thus, having in mind Proposition 5.9, we find that $\Sigma^{>k} \setminus \Sigma^{\geq k+1}$ is empty for k even. If k is odd, we find $\lambda_k(x_\circ) = k + \frac{1}{2}$ for every $x_\circ \in \Sigma^{>k} \setminus \Sigma^{\geq k+1}$. If this set was not discrete, we could apply Lemma 6.7 and reach a contradiction, obtaining a one-dimensional and $\left(k + \frac{1}{2}\right)$-homogeneous solution of the thin obstacle problem. □

6.3. ***The size of*** $\Sigma^{\geq k+1}\setminus\Sigma^{(k+1)\text{-}th}$.

Lemma 6.10. *Let* $k\geq 2$, $x\in\Sigma^{\geq k+1}\setminus\Sigma^{(k+1)\text{-}th}$ *and* $\varepsilon>0$. *Then there exists* $\varrho=\varrho(\varepsilon,x)>0$ *such that, for each* $r\in(0,\varrho)$, *there is* $q\in\mathcal{S}^{\text{even}}_{k+1}(\{p_{2,x}=0\})\setminus\{0\}$ *such that*

$$\Sigma(u)\cap B_r(x)\subseteq\Sigma(q)+B_{\varepsilon r}(x). \tag{6-9}$$

Recall that $\Sigma(q)=\{q=|\nabla q|=0\}\cap\{p_{2,x}=0\}$ *was defined in* (2-9).

Proof of the case $(k+1)$ *even.* Up to an isometry, we can assume $x=0$ and $p_2=\frac{1}{2}x_n^2$. We argue by contradiction and rescale everything: we find $\varepsilon_\circ>0$ and a sequence $r_\ell\downarrow 0$ such that

$$y_\ell\in\Sigma(u(r_\ell\,\cdot\,))\cap B_1\quad\text{and}\quad\operatorname{dist}(y_\ell,\Sigma(q'))\geq\varepsilon_\circ$$

for all $q'\in\mathcal{S}^{\text{even}}_{k+1}(\{x_n=0\})$. Up to taking subsequences, we can assume $y_\ell\to y_\infty\in\{x_n=0\}$, and by Proposition 5.9,

$$\frac{(u-\mathcal{P}_k)(r_\ell\,\cdot\,)}{r_\ell^{k+1}}\to\bar{q}\in\mathcal{S}_{k+1}(\{x_n=0\})$$

in $C^0_{\text{loc}}(\mathbb{R}^n)$. Since $0\notin\Sigma^{(k+1)\text{-th}}$, we have $\bar{q}^{\text{even}}\neq 0$. Rearranging the terms we can equivalently write

$$w_\ell:=\frac{\big(u-\frac{1}{2}\mathcal{A}^2_{k+1}(p_{2,0},\dots,p_{k,0},\bar{q}^{\text{odd}})\big)(r_\ell\,\cdot\,)}{r_\ell^{k+1}}\to\bar{q}^{\text{even}}\quad\text{in } C^0_{\text{loc}}(\mathbb{R}^n).$$

Now recall that, $k+1$ being even, we have $\Sigma(\bar{q}^{\text{even}})=\{\bar{q}^{\text{even}}=0\}$, thus $\eta:=\bar{q}^{\text{even}}(y_\infty)>0$, as y_∞ lies on the thin obstacle. So we can find a small radius $\delta>0$ and a large ℓ_0 such that, for all $\ell>\ell_0$, we have

$$\inf_{B_\delta(y_\infty)}w_\ell\geq\inf_{B_\delta(y_\infty)}\bar{q}^{\text{even}}-\|w_\ell-\bar{q}^{\text{even}}\|_{L^\infty(B_2)}\geq\tfrac{1}{2}\eta-\tfrac{1}{4}\eta=\tfrac{1}{4}\eta.$$

However, eventually we will have $y_\ell\in B_\delta(y_\infty)$, and this is a contradiction as

$$0\geq-\frac{\mathcal{A}^2_{k+1}(r_\ell y_\ell)}{2r_\ell^{k+1}}=w_\ell(y_\ell)\geq\inf_{B_\delta(y_\infty)}w_\ell\geq\tfrac{1}{4}\eta.$$

We remark that we only used that $y_\ell\in\{u_{r_\ell}=0\}$, not that the y_ℓ were singular points. □

Proof of the case $(k+1)$ *odd.*. Arguing by contradiction as in the even case, we find $\varepsilon_\circ>0$ and a sequence $r_\ell\downarrow 0$ such that, for each ℓ,

$$y_\ell\in\Sigma(u(r_\ell\,\cdot\,))\cap B_1\quad\text{and}\quad\operatorname{dist}(y_\ell,\Sigma(q))\geq\varepsilon_\circ\quad\text{for all } q\in\mathcal{S}^{\text{even}}_{k+1}(\{x_n=0\})\setminus\{0\}.$$

We can also assume that $y_\ell\to y_\infty\in\{x_n=0\}$ and

$$w_\ell:=\frac{\big(u-\frac{1}{2}\mathcal{A}^2_{k}(p_{2,0},\dots,p_{k,0},\bar{q}^{\text{odd}})\big)(r_\ell\,\cdot\,)}{r_\ell^{k+1}}\to\bar{q}^{\text{even}}\quad\text{in } C^0_{\text{loc}}(\mathbb{R}^n).$$

From this point the proof is conducted analogously to the proof of Lemma 6.3; it suffices to replace k with $k+1$ and h_ℓ with r_ℓ^{k+1}. □

Remark 6.11. In the case $k+1$ even the proof actually gives a stronger result: as we only used that $y_\ell \in \{u=0\}$, we can replace $\Sigma(u)$ with the full contact set. In other words we can replace (6-9) with

$$\{u=0\} \cap B_r(x) \subseteq \Sigma(q) + B_{\varepsilon r}(x). \tag{6-10}$$

We can conclude now exactly as in the case $\lambda_k = k$ for k odd.

Corollary 6.12. *Suppose $k \geq 2$. Then $\dim_{\mathcal{H}}(\Sigma^{\geq k+1} \setminus \Sigma^{(k+1)\text{-}th}) \leq n-2$. Furthermore, if $n=2$, then $\Sigma^{\geq k+1} \setminus \Sigma^{(k+1)\text{-}th}$ is discrete in the full Σ.*

Proof. Recall that if $n=2$ then $\Sigma(q)=\{0\}$ for every $q \in \mathcal{S}_{k+1}^{\text{even}}$. Pick $x \in \Sigma^{\geq k+1} \setminus \Sigma^{(k+1)\text{-th}}$ and apply Lemma 6.10 with $\varepsilon := \frac{1}{2}$. This gives that, for all $r < \varrho\big(x, \frac{1}{2}\big)$, we have

$$\Sigma(u) \cap (B_r(x) \setminus \overline{B_{r/2}(x)}) = \varnothing.$$

This clearly gives $\Sigma(u) \cap B_\varrho(x) = \{x\}$, thus x is isolated in $\Sigma(u)$.

For $n \geq 3$, we argue as in Corollary 6.6; namely, we apply Proposition 6.5 to $E := \Sigma^{\geq k} \setminus \Sigma^{(k+1)\text{-th}}$. The assumptions are satisfied thanks to Lemma 6.10. □

In particular we notice that in dimension $n=2$ this forces the sets $\Sigma^{k\text{-th}}$ to be closed.

Corollary 6.13. *If $n=2$, then the sets $\Sigma^{k\text{-}th}$ are closed for all $k \geq 2$.*

Proof. We prove the assertion by induction on k. Let $(x_\ell)_{\ell\in\mathbb{N}} \subseteq \Sigma^{(k+1)\text{-th}} \setminus \{0\}$ be a sequence with $x_\ell \to 0$. In particular, we have $x_\ell \in \Sigma^{\geq k+1}$. By inductive assumption we can assume $0 \in \Sigma^{k\text{-th}}$, and by the upper semicontinuity of the truncated frequency we get $\lambda_k(0) \geq k+1$ (see Remark 5.11). But by Corollary 6.12 the origin cannot lie in $\Sigma^{\geq k+1} \setminus \Sigma^{(k+1)\text{-th}}$, because it is an accumulation point of the sequence of singular points x_ℓ. Since $\Sigma^{2\text{nd}} = \Sigma_{n-1}$ is closed, by lower semicontinuity of the rank, the proof is finished. □

6.4. *The geometry of* Σ^∞.

Let us put together the results obtained so far. By definition,

$$\Sigma^\infty := \bigcap_{k\geq 2} \Sigma^{k\text{-th}}.$$

In the last three subsections we proved the following.

Proposition 6.14. *We have $\dim_{\mathcal{H}}(\Sigma \setminus \Sigma^\infty) \leq n-2$. If $n=2$, then $\Sigma \setminus \Sigma^\infty$ is countable.*

Proof. By definition we have that

$$\Sigma \setminus \Sigma^\infty = (\Sigma \setminus \Sigma_{n-1}) \cup \bigcup_{j\geq 2} (\Sigma^{j\text{-th}} \setminus \Sigma^{>j}) \cup \bigcup_{j\geq 2} (\Sigma^{>j} \setminus \Sigma^{\geq j+1}) \cup \bigcup_{j\geq 3} (\Sigma^{\geq j} \setminus \Sigma^{j\text{-th}}).$$

But now

- $\dim_{\mathcal{H}}(\Sigma \setminus \Sigma_{n-1}) \leq n-2$ (discrete if $n=2$), by [Caffarelli 1998, Theorem 8];
- $\dim_{\mathcal{H}}(\Sigma^{j\text{-th}} \setminus \Sigma^{>j}) \leq n-2$ (discrete if $n=2$), by Corollaries 6.2 and 6.6;
- $\dim_{\mathcal{H}}(\Sigma^{j\text{-th}} \setminus \Sigma^{>j}) \leq n-2$ (discrete if $n=2$), by Proposition 6.9;
- $\dim_{\mathcal{H}}(\Sigma^{\geq j} \setminus \Sigma^{j\text{-th}}) \leq n-2$ (discrete if $n=2$), by Corollary 6.12. □

At each point of Σ^∞ we have Taylor polynomials of every order, and they vary smoothly in the sense of Whitney. This also gives that Σ^∞ locally is contained in a smooth hypersurface. Let us first phrase a suitable statement.

Theorem 6.15. *Let $E \subseteq \mathbb{R}^n$ be any set, and, for each $k \in \mathbb{N}$, consider a collection of polynomials $\{P_{k,x}\}_{x\in E}$ of degree at most k. Suppose that these polynomials satisfy*

(i) *$P_{k,x} = \pi_{\le k}(P_{k+\ell,x})$ for all $k, \ell \in \mathbb{N}$ and $x \in E$,*

(ii) *for each $k \in \mathbb{N}$, there is a constant $C(k)$ such that, for each multi-index α, $|\alpha| \le k$, we have*

$$|\partial^\alpha P_{k,x}(0) - \partial^\alpha P_{k,y}(x-y)| \le C(k)|x-y|^{k-|\alpha|+1} \quad \text{for all } x, y \in E.$$

Then there exists a function $F \in C^\infty(\mathbb{R}^n)$ such that, for each $x \in E$ and $k \in \mathbb{N}$, we have

$$F(x+h) = P_{k,x}(h) + O(|h|^{k+1}) \quad \text{as } |h| \to 0.$$

Proof. This is just a restatement of Whitney's extension theorem for smooth functions. The interested reader can find in Appendix C how to derive this formulation from the original, namely [Whitney 1934, Theorem I]. □

Lemma 6.16. *Let u be a solution to the obstacle problem (2-1). Then Σ^∞ is closed and locally covered by one smooth manifold of dimension $n-1$.*

Proof. The main idea is to combine the implicit function theorem and Whitney's extension theorem (Theorem 6.15). We will first prove the covering and then the closeness.

As the statement is local we can assume that $0 \in \Sigma^\infty$ and that u solves (2-1) in $B_2(0) \subseteq \mathbb{R}^n$. We want to apply Whitney's extension theorem (Theorem 6.15) with $E := \Sigma^\infty \cap B_1$ and the polynomials

$$P_{k,x} := \pi_{\le k}(\mathcal{P}_{k,x}) \quad \text{for all } x \in \Sigma^\infty \cap B_1, \quad k \ge 0.$$

Assumption (i) holds because $\mathcal{P}_{k+\ell}$ and $\mathcal{P}_k$ agree up to order k (see also Lemma 5.7). We need to show that (ii) holds. It is not restrictive to do it only for some fixed $k \ge 3$. To do so we exploit our previous analysis on $\Sigma^{(k+1)\text{-th}}$. More precisely, combining Lemma 3.5 with the uniform estimate in Proposition 5.4 and growth estimates from Proposition 4.2 and Lemma 2.2, we find $R = R(n,k)$ and $C = C(n,k)$ such that, for all $x \in \Sigma^{(k+1)\text{-th}} \cap B_1$ and $0 < r < R < \frac{1}{2}$, we have

$$\|u(x+\cdot) - P_{k,x}\|_{L^\infty(B_r(0))} \le Cr^{k+1}. \tag{6-11}$$

Thus this must hold, a fortiori, for all $x \in \Sigma^\infty \cap B_1$. Let now $x_1, x_2 \in \Sigma^\infty \cap B_1$ such that $|x_1 - x_2| \le \frac{1}{10}R$. Then, since $B_{2|x_1-x_2|}(x_1) \subseteq B_{4|x_1-x_2|}(x_2)$, by (6-11) applied at x_1 with $r_1 = 2|x_1-x_2|$ and at x_2 with $r_2 = 4|x_1-x_2|$, together with the triangle inequality, we find

$$\|P_{k,x_1}(\cdot - x_1) - P_{k,x_2}(\cdot - x_2)\|_{L^\infty(B_{2|x_1-x_2|}(x_1))} \le C|x_1-x_2|^{k+1}. \tag{6-12}$$

If we consider the polynomial $Q = P_{k,x_1}(\cdot - x_1) - P_{k,x_2}(\cdot - x_2)$, equation (6-12) reads

$$\|Q(x_1 + 2|x_1-x_2|\cdot)\|_{L^\infty(B_1)} \le C|x_1-x_2|^{k+1},$$

hence by the equivalence of norms on the space of polynomials of degree bounded by k, we conclude

$$\|(\partial^\alpha Q)(x_1 + 2|x_1 - x_2|\,\cdot\,)\|_{L^\infty(B_1)} \le C|x_1 - x_2|^{k+1-|\alpha|}$$

for all multi-index $|\alpha| \le k$, with some $C = C(n, k)$. In particular, looking at the center of B_1, we get

$$|\partial^\alpha P_{k,x_1}(x_2 - x_1) - \partial^\alpha P_{k,x_2}(0)| \le C(n, k)|x_1 - x_2|^{k+1-l}, \tag{6-13}$$

and this proves that assumption (ii) holds. By the Whitney extension theorem, there exists a C^∞ function $F : \mathbb{R}^n \to \mathbb{R}$ for which

$$F(x) = P_{k,x_\circ}(x - x_\circ) + O(|x - x_\circ|^{k+1})$$

at every $x_\circ \in \Sigma^\infty \cap \overline{B}_1$.

We now conclude as in [Figalli et al. 2020, Proposition 8.1c)], using the implicit function theorem. As $\Sigma^\infty \subseteq \{\nabla F = 0\}$ and $\nabla^2 F(x_1) = \nabla^2 p_{2,x_1}(0)$ has rank 1, we conclude with the implicit function theorem that $\{\nabla F = 0\}$ is a smooth hypersurface in some neighborhood of x_1.

Let us now prove that Σ^∞ is closed. Suppose $0 \in \overline{\Sigma^\infty}$, so there exists $x_\ell \in \Sigma^\infty$ such that $x_\ell \to 0$. First observe that $0 \in \Sigma$ since the full singular set Σ is closed, and hence p_2 exists. We define by continuity $\mathcal{P}_k := \lim_\ell \mathcal{P}_{k,x_\ell}$ for all $k \ge 3$; this is a well-posed definition as by Lemma 6.16 the map $x \mapsto \mathcal{P}_{k,x}$ is Lipschitz on Σ^∞ and hence admits a unique extension to the closure. Now, by Proposition 5.3, for some constants $C, \varepsilon > 0$ and a radius r_0, both independent of ℓ, we have, for all $r \in (0, r_0)$,

$$\frac{d}{dr}(r^{-2k} H(r, u(x_\ell + \cdot\,) - \mathcal{P}_{k,x_\ell})) \ge -Cr^{\varepsilon-1} \quad \text{and} \quad r^{-2k} H(r, u(x_\ell + \cdot\,) - \mathcal{P}_{k,x_\ell}) \le C.$$

We can pass both these inequalities to the limit $\ell \to \infty$ and apply the same reasoning to Proposition 4.2 to show that the sequence $r^{-k}(u - \mathcal{P})_r$ is uniformly bounded in $W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)$. Hence it is immediate that $0 \in \widetilde{\Sigma}^{(k-1)\text{-th}}$ (see Lemma 5.7); as k was arbitrary we conclude $0 \in \Sigma^\infty$. □

We conclude this section proving Theorem 1.1.

Proof of Theorem 1.1. The bound on the dimension of $\Sigma \setminus \Sigma^\infty$ is given by Proposition 6.14 and the covering by Lemma 6.16. The polynomial expansion has been shown in (6-11) above, we simply take $P_{k,x_\circ} := \pi_{\le k}(\mathcal{P}_{k,x_\circ})$. Although we often assumed $f \equiv 1$ and $\mu = 1$ to simplify the notation, essentially no modifications are needed for a general f. The reader can find a complete account of the modifications needed in the statements and in the proofs in Appendix B. □

In the following section we aim to explain in which sense the set Σ^∞ is unstable and disappears after a slight perturbation of the boundary data in the obstacle problem.

7. Extension to a monotone family of solutions

In this section we aim to prove Theorem 1.2 and Corollary 1.4. For simplicity we take $f \equiv 1$. This allows us to use verbatim some lemmas from [Figalli et al. 2020] and shortens the notation, without affecting the proofs. We list the changes needed in Appendix B.

We remark that in Sections 7.1 and 7.2 we only assume to have a monotone family of solutions, while in Section 7.3 we work under the "uniform monotonicity" assumption (7-8).

7.1. ***Setup and strategy.*** For the rest of the section, we let $u : \bar{B}_1 \times [-1, 1] \to \mathbb{R}$, $u \geq 0$, be a monotone 1-parameter family of solutions of the obstacle problem, namely

$$\begin{cases} \Delta u(\,\cdot\,, t) = \chi_{\{u(\,\cdot\,,t)>0\}}, \\ 0 \leq u(\,\cdot\,, s) \leq u(\,\cdot\,, t) \qquad \text{in } B_1 \text{ for } -1 \leq s \leq t \leq 1. \end{cases} \tag{7-1}$$

We will also use the notation $u^t := u(\,\cdot\,, t)$. We will assume in addition that $u \in C^0(\bar{B}_1 \times [-1, 1])$. We remark that this continuity property in t follows by the maximum principle whenever $u|_{\partial B_1 \times [-1,1]}$ is continuous.

We will often think of t as the time parameter, as intuitively we imagine lifting the boundary datum of a solution of (2-1). However, no equation in t is given.

For each fixed t, we can apply the results of the previous sections, so we introduce further notation for the following subsets of $\bar{B}_1 \times [-1, 1]$:

$$\begin{aligned} \boldsymbol{\Sigma} &:= \{(x_\circ, t_\circ) : x_\circ \in \Sigma(u(\,\cdot\,, t_\circ))\}, \\ \boldsymbol{\Sigma}_{n-1} &:= \{(x_\circ, t_\circ) : x_\circ \in \Sigma_{n-1}(u(\,\cdot\,, t_\circ))\}, \\ \boldsymbol{\Sigma}^{k\text{-th}} &:= \{(x_\circ, t_\circ) : x_\circ \in \Sigma^{k\text{-th}}(u(\,\cdot\,, t_\circ))\}, \qquad k \geq 2, \\ \boldsymbol{\Sigma}^{>k} &:= \{(x_\circ, t_\circ) : x_\circ \in \Sigma^{>k}(u(\,\cdot\,, t_\circ))\}, \qquad k \geq 2, \\ \boldsymbol{\Sigma}^{\geq k+1} &:= \{(x_\circ, t_\circ) : x_\circ \in \Sigma^{\geq k+1}(u(\,\cdot\,, t_\circ))\}, \quad k \geq 2, \\ \boldsymbol{\Sigma}^{\infty} &:= \{(x_\circ, t_\circ) : x_\circ \in \Sigma^{\infty}(u(\,\cdot\,, t_\circ))\}. \end{aligned} \tag{7-2}$$

This setup (up to $k = 4$) has already been considered in [Figalli et al. 2020]. As we use the same notation, we begin recalling two important lemmas from [Figalli et al. 2020] about the set $\boldsymbol{\Sigma}$.

Lemma 7.1 [Figalli et al. 2020, Lemma 6.2]. *Let $u \in C^0(\bar{B}_1 \times [-1, 1])$ solve (7-1). Then*

(i) $\boldsymbol{\Sigma} \cap \bar{B}_\varrho \times [-1, 1]$ *is closed for any $\varrho < 1$, and*

$$\boldsymbol{\Sigma} \cap \bar{B}_\varrho \times [-1, 1] \ni (x_k, t_k) \to (x_\infty, t_\infty) \quad \Rightarrow \quad p_{2,x_k,t_k} \to p_{2,x_\infty,t_\infty}.$$

(ii) *If $(x_\circ, t_1)$ and $(x_\circ, t_2)$ both belong to $\boldsymbol{\Sigma}$ and $t_1 < t_2$, then there exists $r > 0$ such that $u(x, t)$ is independent of t for all $(x, t) \in B_r(x_\circ) \times [t_1, t_2]$.*

The next result concerns the quantitative behavior of the first blowup $p_{2,k} := p_{2,x_k,t_k}$ with respect to the convergence $x_k \to 0$ (here it is assumed $(x_k, t_k) \in \boldsymbol{\Sigma}$ for some sequence of times).

Lemma 7.2 [Figalli et al. 2020, Lemma 6.3]. *Let $u \in C^0(\bar{B}_1 \times [-1, 1])$ solve (7-1), let $(x_k, t_k) \in \boldsymbol{\Sigma}$, $(0, 0) \in \boldsymbol{\Sigma}$ and assume that $x_k \to 0$. If we set $p_2 := p_{2,0,0}$, then we have*

$$\left\| p_{2,k} - p_2\left(\frac{x_k}{|x_k|} + \cdot\right) \right\|_{L^\infty(B_1)} \leq C\omega(2|x_k|) \quad \textit{and} \quad \|p_{2,k} - p_2\|_{L^\infty(B_1)} \leq C\omega(2|x_k|)$$

for some dimensional modulus of continuity ω. In addition,

$$\operatorname{dist}\left(\frac{x_k}{|x_k|}, \{p_2 = 0\}\right) \to 0 \quad \textit{as } k \to \infty.$$

Our strategy follows the one exhibited in [Figalli et al. 2020]. For each $(x_\circ, t_\circ) \in \mathbf{\Sigma}^{\geq k+1}$, we first prove the approximation

$$\|u(x_\circ + \cdot\,, t_\circ) - P_{k,x_\circ,t_\circ}\|_{L^\infty(B_r)} \leq Cr^{k+1}, \tag{7-3}$$

where the polynomial $P_{k,x_\circ,t_\circ}$ of degree at most k is unique and $\Delta P_{k,x_\circ,t_\circ} = 1$. Using the fact that the polynomials above are almost positive together with barrier-type arguments (see Lemma 7.11), we are able to conclude a "cleaning property" in space-time in the following sense. For each $(x_\circ, t_\circ) \in \mathbf{\Sigma}^{\geq k+1}$, there exist $\varrho > 0$ and $C > 0$ (depending on $n,\ k,\ x_\circ$) such that

$$\{(x,t) \in B_\varrho(x_\circ) \times (t_\circ, 1) : t - t_\circ > C|x - x_\circ|^k\} \cap \{u = 0\} = \varnothing. \tag{7-4}$$

This property expresses the instability of $\Sigma^{\geq k+1}(u^t)$ with respect to increments of the t parameter. From here, we will conclude with the next geometric measure theory result, that the set $\pi_t(\mathbf{\Sigma}^\infty)$ has zero Hausdorff dimension.

Proposition 7.3 [Figalli et al. 2020, Corollary 7.8]. *Let $E \subseteq \mathbb{R}^n \times [-1,1]$, let (x,t) denote a point in $\mathbb{R}^n \times [-1,1]$, and let $\pi_x : (x,t) \mapsto x$ and $\pi_t : (x,t) \mapsto t$ be the standard projections. Assume that, for some $\beta \in (0, n]$ and $s > \beta$, we have*:

- $\dim_{\mathcal{H}}(\pi_x(E)) \leq \beta$.
- *For all $(x_\circ, t_\circ) \in E$ and $\varepsilon > 0$, there exists $\varrho = \varrho_{x_\circ,t_\circ,\varepsilon} > 0$ such that*

$$\{(x,t) \in B_\varrho(x_\circ) \times [-1,1] : t - t_\circ > |x - x_\circ|^{s-\varepsilon}\} \cap E = \varnothing.$$

Then $\dim_{\mathcal{H}}(\pi_t(E)) \leq \beta/s$.

To obtain Corollary 1.4 we also need to take care of the points where the expansion (7-3) fails for some k. To this end we need to generalize to one-parameter solutions some of the previous results.

7.2. ***Adaptation of previous sections to family of solutions.*** In this section we establish an analog of Theorem 1.1 for monotone families of solutions. The generalization of the polynomial expansion is obvious in the set $\pi_x(\mathbf{\Sigma}^\infty)$, so the only nontrivial task is to show that $\pi_x(\mathbf{\Sigma} \setminus \mathbf{\Sigma}^\infty)$ has again Hausdorff dimension at most $n-2$. We will show this by repeating the arguments of Section 6. While the arguments of Sections 6.1 and 6.3 adapt immediately by exploiting monotonicity, the arguments of Section 6.2 require a bit more care. Specifically, we have to check Lemma 6.7 for varying times.

We start observing that in Proposition 5.9 one can also consider the varying time parameter.

Proposition 7.4. *Let $u \in C^0(\bar{B}_1 \times [-1,1])$ solve (7-1) and $(0,0) \in \mathbf{\Sigma}^{k\text{-}th}$, with $\lambda_k = \lambda_{k,0}(0) \leq k+1$. Let $(r_\ell)_{\ell\in\mathbb{N}}$ be an infinitesimal sequence, and let $x_\ell \in \Sigma^{k\text{-}th}(u^{t_\ell}) \cap B_{r_\ell}$. For every ℓ, set*

$$v_\ell := u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell},$$

and suppose that $\lambda_{k,t_\ell}(x_\ell) \to \lambda_k$. Consider the sequence

$$\tilde{v}_\ell := \frac{v_\ell(r_\ell\,\cdot)}{H(r_\ell, v_\ell)^{1/2}}.$$

Then:

(i) $(\tilde{v}_\ell)_{\ell\in\mathbb{N}}$ *is bounded in* $W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)$ *and* $C^{0,1/(n+1)}_{\mathrm{loc}}(\mathbb{R}^n)$.

(ii) *If* $\tilde{v}_\ell \rightharpoonup q \in W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n)$, *then the convergence is strong and* q *must be a nontrivial* λ_k*-homogeneous solution of the thin obstacle problem* (5-4) *with obstacle* $\{p_2 = 0\}$, *that is*

$$\begin{cases} \Delta q \le 0 \ \text{ and } \ q\Delta q = 0 & \text{in } \mathbb{R}^n, \\ \Delta q = 0 & \text{in } \mathbb{R}^n \setminus \{p_2 = 0\}, \\ q \ge 0 & \text{on } \{p_2 = 0\}. \end{cases}$$

Finally, *if* $\lambda_k < k+1$ *then* q *is even with respect to the thin obstacle.*

Proof. Given the convergence assumption $\lambda_{k,t_\ell}(x_\ell) \to \lambda_k$ and Lemma 7.2, the proof is almost identical to Proposition 5.9. □

We now turn to the time-dependent version of Lemma 6.7.

Lemma 7.5. *Let* $u \in C^0(\bar{B}_1 \times [-1,1])$ *solve* (7-1), *let* $k \ge 2$ *and suppose* $(0,0) \in \mathbf{\Sigma}^{>k} \setminus \mathbf{\Sigma}^{\ge k+1}$, *that is* $\lambda_k = \lambda_{k,0}(0) \in (k, k+1)$. *Suppose there exists an infinitesimal sequence* $r_\ell \downarrow 0$ *and* $(x_\ell, t_\ell) \in \mathbf{\Sigma}^{k\text{-th}} \cap B_{r_\ell}$ *such that* $\lambda_{k,t_\ell}(x_\ell) \to \lambda_k$. *Assume further that, as* $\ell \uparrow \infty$, *we have*

(i) $x_\ell / r_\ell \to y_\infty \in \bar{B}_1$,

(ii) $\tilde{v}_{r_\ell} = (u_0 - \mathcal{P}_{k,0,0})_{r_\ell} / H(r_\ell, u_0 - \mathcal{P}_{k,0,0})^{1/2} \to q$ *in* $C^0_{\mathrm{loc}}(\mathbb{R}^n)$ *for some* $q \in \mathcal{S}_{\lambda_k}(\{p_{2,0,0} = 0\}) \setminus \{0\}$.

Then $y_\infty \in \{p_{2,0,0} = 0\}$ *and* $q = q(y_\infty + \cdot\,)$.

Proof. Whenever $x = t = 0$, we simplify the notation by dropping the indices, e.g., $p_{2,0,0} = p_2$. Consider a sequence $(x_\ell, t_\ell)_{\ell\in\mathbb{N}} \subseteq \mathbf{\Sigma}^{k\text{-th}} \cap B_{r_\ell}$ as in the statement of the lemma. Note that $y_\infty \in \{p_2 = 0\}$ due to Lemma 7.2. Applying Proposition 7.4 with varying centers $(x_\ell)_{\ell\in\mathbb{N}}$ and respective sequence of times $(t_\ell)_{\ell\in\mathbb{N}}$, we find (after passing to a subsequence)

$$\tilde{v}_\ell := \frac{u(x_\ell + r_\ell\,\cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell}(r_\ell\,\cdot\,)}{\|(u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell})_{r_\ell}\|_{L^2(\partial B_1)}} \to Q$$

in $C^0_{\mathrm{loc}}(\mathbb{R}^n)$ for some $Q \in \mathcal{S}_{\lambda_k}(\{p_2 = 0\}) \setminus \{0\}$. On the other hand, by uniform convergence,

$$q(y_\infty + \cdot\,) = \lim_\ell \frac{u(x_\ell + r_\ell\,\cdot\,) - \mathcal{P}_k(x_\ell + r_\ell\,\cdot\,)}{H(r_\ell, u - \mathcal{P}_k)^{1/2}}.$$

We write

$$\frac{u(x_\ell + r_\ell\,\cdot\,) - \mathcal{P}_k(x_\ell + r_\ell\,\cdot\,)}{H(r_\ell, u - \mathcal{P}_k)^{1/2}} = \frac{u(x_\ell + r_\ell\,\cdot\,) - u(x_\ell + r_\ell\,\cdot\,, t_\ell)}{\|(u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell})_{r_\ell}\|_{L^2(B_1)}} \cdot a_\ell I_\ell + \tilde{v}_\ell \cdot b_\ell I_\ell + J_\ell,$$

where

$$I_\ell := \frac{\|(u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell})_{r_\ell}\|_{L^2(B_1)} + H(r_\ell, u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell})^{1/2}}{H(r_\ell, u - \mathcal{P}_k)^{1/2}}$$

is a numerical sequence and

$$J_\ell := \frac{\mathcal{P}_{k,x_\ell}(r_\ell\,\cdot\,) - \mathcal{P}_k(x_\ell + r_\ell\,\cdot\,, t_\ell)}{H(r_\ell, u - \mathcal{P}_k)^{1/2}}$$

is a sequence of harmonic polynomials of degree at most $k+1$. The numerical sequences

$$a_\ell := \frac{\|(u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell})_{r_\ell}\|_{L^2(B_1)}}{\|(u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell})_{r_\ell}\|_{L^2(B_1)} + H(r_\ell, u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell})^{1/2}}$$

and

$$b_\ell := \frac{H(r_\ell, u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell})^{1/2}}{\|(u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell})_{r_\ell}\|_{L^2(B_1)} + H(r_\ell, u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell})^{1/2}}$$

are both bounded by 1 and hence, up to a subsequence, converge to some a and b, respectively, in $[0, 1]$. Now two cases arise:

(i) $\sup_\ell I_\ell < \infty$.

(ii) $I_{\ell_m} \uparrow \infty$ for some subsequence $\ell_m \to \infty$.

Let us begin with the first case. Up to a subsequence that we do not rename, we have $I_\ell \to \alpha$, and passing to the limit as $\ell \to \infty$ in (7-5) in $L^2(B_1)$ implies that J_ℓ converges to some harmonic polynomial J of degree at most $k+1$. We find that

$$q(y_\infty + \cdot\,) = a\alpha w + b\alpha Q + J \tag{7-5}$$

in L^2 for some function w having a constant sign. Combining this fact and exploiting the homogeneity of q and Q, as in Proposition 5.9, we find

$$q(\cdot\,) \le a\alpha Q(\cdot\,) + R^{-\lambda_k} J(R\,\cdot\,) \quad \text{or} \quad q \ge a\alpha Q + R^{-\lambda_k} J(R\,\cdot\,). \tag{7-6}$$

Next, we remark that J does not have a constant sign, as it is harmonic and vanishes somewhere on the line segment $\overline{-y_\infty 0}$. The last property is seen as follows. First, note that, for large ℓ, we have $H(r_\ell, u - \mathcal{P}_k)^{1/2} \gg r_\ell^{\lambda_k+\delta} \gg r_\ell^{k+1}$. On the other hand, we calculate

$$H(r_\ell, u - \mathcal{P}_k)^{1/2} J_\ell(0) \le C r_\ell^{k+2} \quad \text{and} \quad H(r_\ell, u - \mathcal{P}_k)^{1/2} J_\ell\Big(-\frac{x_\ell}{r_\ell}\Big) \ge -C r_\ell^{k+2}.$$

Hence, there exists a sequence of points $\bar{y}_\ell \in \overline{-(x_\ell/r_\ell)0}$ with $|J_\ell(\bar{y}_\ell)| \le C r_\ell$, and so J vanishes at some point in the line segment $\overline{-y_\infty 0}$.

As J does not have a constant sign, there are directions $x_\pm \in \mathbb{S}^{n-1}$ with $J(Rx_\pm) \to \pm\infty$ as $R \to \infty$. Combining this with (7-6), we thus find $\deg J \le k$ and

$$q \le a\alpha Q \quad \text{or} \quad q \ge a\alpha Q.$$

Thus, in any of the cases, we have found two ordered λ_k-homogeneous solutions to the thin obstacle problem, and they must be equal, see [Figalli et al. 2020, Lemma A.4]. Inserting this back in (7-5), we find $q(y_\infty + \cdot\,) = q + J$. Therefore, for any $R > 0$, we have

$$R\Big(q\Big(\frac{y_\infty}{R} + \cdot\Big) - q(\cdot\,)\Big) \le R^{1-\lambda_k} J(R\,\cdot\,) \quad \text{or} \quad R\Big(q\Big(\frac{y_\infty}{R} + \cdot\Big) - q(\cdot\,)\Big) \ge R^{1-\lambda_k} J(R\,\cdot\,).$$

As the left-hand side is bounded (and converges to $y_\infty \cdot \nabla q$) as $R \to \infty$, we exploit the fact that J does not have a constant sign to find that the k-th coefficients must vanish. And so

$$y_\infty \cdot \nabla q \leq 0 \quad \text{or} \quad y_\infty \cdot \nabla q \geq 0.$$

Reasoning as in Step 3 in the proof of [Figalli et al. 2020, Lemma 6.5], we find that in any of the cases we must have $y_\infty \cdot \nabla q \equiv 0$, as otherwise $y_\infty \cdot \nabla q$ would be a multiple of an eigenfunction to some elliptic problem on a subset on the sphere, which contradicts the high homogeneity of $y_\infty \cdot \nabla q$.

The second case is simpler. We divide (7-5) by I_{ℓ_m} and find, after passing to a subsequence of ℓ_m,

$$0 = aw + bQ + \tilde{J}$$

for some harmonic polynomial $\tilde{J}$ of degree at most $k+1$. This is a contradiction, as the three functions $\tilde{J}$, w and Q are not linearly dependent. Indeed, w has a sign, $Q \neq 0$ is a $\lambda_k \in (k, k+1)$ homogeneous function and $\tilde{J}$ a harmonic polynomial with no constant sign. The fact that $\tilde{J}$ vanishes somewhere can be checked as we checked that J vanishes somewhere, using that

$$I_\ell \geq \frac{H(r_\ell, u(x_\ell + \cdot\,, t_\ell) - \mathcal{P}_{k,x_\ell,t_\ell})^{1/2}}{H(r_\ell, u - \mathcal{P}_k)^{1/2}} \gg \frac{r_\ell^{k+1}}{H(r_\ell, u - \mathcal{P}_k)^{1/2}}. \qquad \square$$

In order to perform the necessary dimension reductions we need some adaptations of Section 6. We start with the following variation of Proposition 6.5, taken from [Figalli et al. 2023].

Proposition 7.6 [Figalli et al. 2023, Proposition 7.6]. *Let $k \geq 2$ and $E \subseteq \mathbb{R}^n \times \mathbb{R}$. Suppose that*

$$\forall (x,t) \in E, \quad \forall \varepsilon > 0, \quad \exists \varrho > 0, \quad \forall r \in (0, \varrho), \quad \exists L \text{ hyperplane}, \quad \exists q \in \mathcal{S}_k^{\text{even}}(L)$$

such that

$$\pi_x(E \cap (\overline{B_r(x)} \times (-\infty, t])) \subseteq \Sigma(q) + \overline{B_{\varepsilon r}(x)}.$$

Then $\dim_{\mathcal{H}}(E) \leq n-2$.

We want to apply this proposition to $E = \mathbf{\Sigma}^{k\text{-th}} \setminus \mathbf{\Sigma}^{(k+1)\text{-th}}$. We check in the next two lemmas that this is possible.

Lemma 7.7. *Let $k \geq 2$ and $(0,0) \in \mathbf{\Sigma}^{k\text{-th}} \setminus \mathbf{\Sigma}^{(k+1)\text{-th}}$, and suppose $\lambda := \lambda_k(0,0)$ is an even integer. Then,*

$$\forall \varepsilon > 0, \quad \exists \varrho > 0, \quad \forall r \in (0, \varrho), \quad \exists q \in \mathcal{S}_\lambda^{\text{even}}(\{p_{2,0,0} = 0\})$$

such that

$$\pi_x(\{u = 0\} \cap (\overline{B_r(x)} \times [0,1])) \subseteq \Sigma(q) + \overline{B_{\varepsilon r}(x)}.$$

Proof. Notice that, by Lemma 6.1, $\lambda = k+1$. Further we have by monotonicity

$$\pi_x(\{u = 0\} \cap (\overline{B_r(x)} \times [0,1])) \subseteq \{u(\cdot\,, 0) = 0\} \cap \overline{B_r(x)}.$$

Now by Lemma 6.10, and taking Remark 6.11 into account, the set $\{u(\cdot\,,0) = 0\} \cap \overline{B_r(x)}$ can be covered with a tubular neighborhood of the singular set of a Signorini solution, provided r is small enough. $\square$

For odd frequencies we use a control "in the past".

Lemma 7.8. *Let $k \geq 2$ and $(0,0) \in \mathbf{\Sigma}^{k\text{-}th} \setminus \mathbf{\Sigma}^{(k+1)\text{-}th}$, and suppose $\lambda := \lambda_k(0,0)$ is an odd integer. Then,*

$$\forall \varepsilon > 0, \quad \exists \varrho > 0, \quad \forall r \in (0, \varrho), \quad \exists q \in \mathcal{S}_\lambda^{\text{even}}(\{p_{2,0,0} = 0\})$$

such that

$$\pi_x(\mathbf{\Sigma} \cap (\overline{B_r(x)} \times [-1, 0])) \subseteq \Sigma(q) + \overline{B_{\varepsilon r}(x)}.$$

Proof. Notice that either $\lambda = k$ or $\lambda = k+1$. In the first case, we reproduce the proof of Lemma 6.3. In the second case, we reproduce the proof of Lemma 6.10 for the odd case. In both cases, it suffices to replace u with $u(\,\cdot\,, 0)$, and the argument for a single solution can be applied. The key point is that, by monotonicity, the barriers $\{\phi_{z,\ell}\}$ will work for all $u(\,\cdot\,, t)$ for $t \leq 0$. Indeed, following the proof with the same notation, one arrives at

$$\{\mathcal{A}_{k,0,0}(r_\ell\,\cdot\,) = 0\} \cap B_\rho(y_\infty) \subseteq Z_\ell \subseteq \operatorname{int}\{u(r_\ell\,\cdot\,, 0) = 0\} \subseteq \operatorname{int}\{u(r_\ell\,\cdot\,, t) = 0\}.$$

Hence the contact set of $u(\,\cdot\,, t)$ is fat around y_∞. This gives $\Sigma(u(r_\ell\,\cdot\,, t)) \cap B_{\rho/N}(y_\infty) = \varnothing$ for some dimensional constant N and for all $t \leq 0$ (see [Caffarelli 1998, Theorem 7] or the proof of [Figalli et al. 2020, Lemma 9.4]). This is the desired contradiction as $y_\ell \to y_\infty$, where $y_\ell \in \Sigma(u(r_\ell\,\cdot\,, t_\ell))$. □

Putting all these results together, we can prove the main theorem of this section. It is an extension of the fifth-order approximation result [Figalli et al. 2020, Theorem 8.7] to every order. For a fixed solution, this is just the content of our main Theorem 1.1.

Theorem 7.9. *Let $u \in C^0(\bar{B}_1 \times [-1, 1])$ solve (7-1). Then $\dim_{\mathcal{H}}(\pi_x(\mathbf{\Sigma} \setminus \mathbf{\Sigma}^\infty)) \leq n-2$, and the set is countable if $n = 2$. Moreover, for every $k \geq 2$, there exist constants $C = C(n,k)$ and $\rho = \rho(n,k)$ such that*

$$\|u(x_\circ + \,\cdot\,, t_\circ) - P_{k,x_\circ,t_\circ}\|_{L^\infty(B_r)} \leq C r^{k+1} \tag{7-7}$$

holds with a unique polynomial $P_{k,x_\circ,t_\circ}$ of degree at most k and $\Delta P_{k,x_\circ,t_\circ} = 1$, for all $0 < r < \rho$ and $(x_\circ, t_\circ) \in \mathbf{\Sigma}^\infty \cap B_{1/2} \times (-1, 1)$.

Proof. We recall from [Figalli et al. 2020, Proposition 8.1] that $\dim_{\mathcal{H}}(\pi_x(\mathbf{\Sigma} \setminus \mathbf{\Sigma}_{n-1})) \leq n-2$ and that $\dim_{\mathcal{H}}(\pi_x(\mathbf{\Sigma} \setminus \mathbf{\Sigma}_{n-1}))$ is countable if $n = 2$. Thus we need to show that, for all $k \geq 2$,

(i) $\dim_{\mathcal{H}}(\pi_x(\mathbf{\Sigma}^{\geq k} \setminus \mathbf{\Sigma}^{k\text{-th}})) \leq n-2$ (countable if $n = 2$),

(ii) $\dim_{\mathcal{H}}(\pi_x(\mathbf{\Sigma}^{k\text{-th}} \setminus \mathbf{\Sigma}^{>k})) \leq n-2$ (countable if $n = 2$),

(iii) $\dim_{\mathcal{H}}(\pi_x(\mathbf{\Sigma}^{>k} \setminus \mathbf{\Sigma}^{\geq k+1})) \leq n-2$ (countable if $n = 2$).

By Lemmas 7.8 and 7.7, to prove (i) and (ii) we can use Proposition 7.6 (or an obvious version of it for future times) with $E = \mathbf{\Sigma}^{\geq k} \setminus \mathbf{\Sigma}^{k\text{-th}}$ and $E = \mathbf{\Sigma}^{k\text{-th}} \setminus \mathbf{\Sigma}^{>k}$, respectively.

We turn to the proof of (iii). We can apply Proposition 6.8 to the set $E = \pi_x(\mathbf{\Sigma}^{>k} \setminus \mathbf{\Sigma}^{\geq k+1})$ using the function $f(x_\circ) := \lambda_{k,\tau(x_\circ)}(x_\circ)$, where $\tau : \pi_x(\mathbf{\Sigma}) \to [-1, 1]$ is defined by

$$\tau(x_\circ) := \min\{t \in [-1, 1] : (x_\circ, t) \in \mathbf{\Sigma}\}.$$

The assumptions of Proposition 6.8 hold for such E: if not we could argue by contradiction and blow up exactly as in Proposition 6.9. The only difference is that we have to use Lemma 7.5 above, instead of Lemma 6.7. □

7.3. ***Cleaning lemmas in the time variable.*** Following [Figalli et al. 2020], in this section we consider any monotone family of solutions $\{u^t\}_{t\in(-1,1)}$ of (2-1) in B_1, which additionally satisfy the following "uniform monotonicity" condition:

$$\begin{aligned}&\text{For every } t\in(-1,1) \text{ and any compact set } K_t\subseteq\partial B_1\cap\{u^t>0\},\\ &\qquad \text{there exists } c_{K_t}>0 \text{ such that } \inf_{x\in K_t}(u^{t'}(x)-u^t(x))\geq c_{K_t}(t'-t) \text{ for all } -1<t<t'<1.\end{aligned}\tag{7-8}$$

This condition rules out the existence of regions that remain stationary as we increase the parameter t. Combining this observation with (iii) in Lemma 7.1, one gets that $\boldsymbol{\Sigma}$ is a graph above B_1 in the sense that

$$x\in\Sigma(u^t)\cap\Sigma(u^s)\quad\Rightarrow\quad s=t.$$

We now turn to the "cleaning lemmas", namely Lemmas 7.10 and 7.11. Using a barrier argument, we show that if u^0 is $O(r^\kappa)$-close to a polynomial ansatz in B_r, then u^t is positive in B_r as soon as $t\sim r^\kappa$: thus the contact set was "cleaned" from B_r. The larger the κ, the faster this cleaning takes place. Then we combine this reasoning with the polynomial expansions given by the $\mathcal{P}_k$.

Lemma 7.10. *Let $u\in C^0(\overline{B}_1\times[-1,1])$ solve (7-1) and satisfy the uniform monotonicity condition (7-8). Assume $(0,0)\in\boldsymbol{\Sigma}$, and let $\mathcal{P}$ be a solution of $\Delta\mathcal{P}=1$ such that*

$$|u(\,\cdot\,,0)-\mathcal{P}|\leq Cr^\kappa\quad\text{in } B_r \text{ for all } r\in\left(0,\tfrac{1}{2}\right)$$

for some $C,\kappa>0$. Then there exist $r_\circ, c>0$ such that

$$u(\,\cdot\,,t)\geq\mathcal{P}+crt-Cr^\kappa\quad\text{in } B_{r/4} \text{ for all } r\in(0,r_\circ).$$

Proof. This is a combination of Lemmas 9.1 and 9.2 in [Figalli et al. 2020]. □

The next result shows that, if $(x_\circ,t_\circ)\in\boldsymbol{\Sigma}^{\geq k+1}$, then the contact set surely disappears from $B_r(x_\circ)$ after $t-t_\circ\sim r^k$ units of time.

Lemma 7.11. *Let $u\in C^0(\overline{B}_1\times[-1,1])$ solve (7-1) and satisfy the uniform monotonicity condition (7-8). Suppose $(0,0)\in\boldsymbol{\Sigma}^{\geq k+1}$ for some $k\geq 2$. Then there exists $r, C_0>0$ depending on n and k such that*

$$\{(x,t)\in B_r\times(0,1):t>C_0|x|^k\}\cap\{u=0\}=\varnothing.$$

Proof. Since $0\in\Sigma^{\geq k+1}(u^t)$, there exists $C(n,k)>0$ such that, for every $r\in\left(0,\tfrac{1}{2}\right)$,

$$|u(\,\cdot\,,0)-\mathcal{P}_k|\leq Cr^{k+1}\quad\text{in } B_r.$$

Moreover, recall from Proposition 3.3 that $\mathcal{P}_k$ is almost positive, in the sense of

$$\mathcal{P}_k\geq -C(n,k)|x|^{k+2}\quad\text{in } B_1.$$

Combining this with Lemma 7.10 with $\mathcal{P}=\mathcal{P}_k$ and $\kappa=k+1$, we get

$$u(\,\cdot\,,t)\geq\mathcal{P}_k+crt-Cr^{k+1}\geq -Cr^{k+1}+crt\quad\text{in } B_{r/4} \text{ for all } r\in(0,r_\circ) \text{ for all } t\geq 0$$

for some $r_\circ, c>0$. Now evaluating this at $(x,t)\in\partial B_r\times(0,1)$, with $t>C_0r^k$, we get $u(x,t)>0$ as soon as r is small enough and C_0 is large enough in terms of c and C. □

We finally prove Theorem 1.2, combining Lemma 7.11 with Proposition 7.3.

Proof of Theorem 1.2. For any $k \geq 2$, we can apply Proposition 7.3 to the set $E = \mathbf{\Sigma}^{\geq k+1}$ with $\beta = n$ and $s = k+1$, as the assumptions are satisfied thanks to Lemma 7.11. Hence, we get

$$\dim_{\mathcal{H}}(\pi_t(\mathbf{\Sigma}^\infty)) \leq \dim_{\mathcal{H}}(\pi_t(\mathbf{\Sigma}^{\geq k+1})) \leq \frac{n}{k+1},$$

and (i) follows letting $k \uparrow \infty$. As noted in Remark 1.3, this bound can be improved to a Minkowski dimension bound by directly applying Lemma 4.2 in [Fernández-Real and Ros-Oton 2021], which is a refinement of Proposition 7.3.

For (ii) it suffices to show that $\pi_t(\mathbf{\Sigma} \setminus \mathbf{\Sigma}^\infty)$ has zero Hausdorff dimension. By Proposition 6.14, the set $\pi_x(\mathbf{\Sigma} \setminus \mathbf{\Sigma}^\infty)$ is countable, provided $n = 2$. On the other hand, by the strict monotonicity condition (7-8), $\mathbf{\Sigma}$ is a graph above the space variables and hence $\mathbf{\Sigma} \setminus \mathbf{\Sigma}^\infty$ is also countable; this finishes the proof. Finally, (iii) is contained in Theorem 7.9. □

We turn to the proof of Corollary 1.4. We remark that, for analytic f, we have at most countable many singular times (combining Theorem 1.2 with [Sakai 1993, Theorem 1.1]). For smooth f, Theorem 1.2 gives that singular times have zero Hausdorff dimension.

Proof of Corollary 1.4. We divide the proof into two steps.

Step 1. The set $\Sigma(u^t) \setminus \Sigma^\infty(u^t)$ is not empty at most for countably many times.

The result follows directly from Theorem 1.2 (iii) provided we show that $\{u^t\}$ satisfies the uniform monotonicity condition (7-8). For completeness we give the argument: fix $t, h > 0$ and $K \Subset \{u^t > 0\}$. For brevity, we work with the assumption that Ω is connected and thus unbounded. Notice that $w := u^{t+h} - u^t$ is harmonic in $\{u^t > 0\}$, which is connected. By Schauder estimates and Lipschitz regularity of $\partial\Omega$, we have that $\operatorname{dist}(\{u^t = 0\}, O) > \delta$ for some $\delta = \delta(n, \partial\Omega, t) > 0$. Hence we can build an open and connected set V with Lipschitz boundary such that

$$\overline{O} \cup K \subseteq V \Subset \{u^t > 0\}.$$

By comparison we have $w \geq h \cdot \phi$, where ϕ solves

$$\begin{cases} \Delta\phi = 0 & \text{in } V \setminus \overline{O}, \\ \phi = 1 & \text{on } \partial O = \partial\Omega, \\ \phi = 0 & \text{in } \partial V. \end{cases}$$

As $\phi > 0$ in $V \setminus \overline{O}$, we have $c := \min_K \phi > 0$, so, for all $h > 0$ and $x \in K$, we have

$$u^{t+h}(x) - u^t(x) \geq h \min_K \phi = ch.$$

We used that V, and hence ϕ, did not depend on h.

Step 2. The set $\Sigma^\infty(u^t)$ is not empty for at most countably many times.

Assume $0 \in \Sigma^\infty(u^0)$. Then we will show that we have an instantaneous cleaning of the zero set, that is: there exists a universal $\delta > 0$ such that $B_\delta \cap \{u^t = 0\} = \varnothing$ for all $t > 0$. In fact, referring to the classification provided in [Sakai 1993, Theorem 1.1], we have that 0 must be a "degenerate" point (case 2a), that is:

$\{u^0 = 0\} \cap B_\delta$ must be an analytic arc (it cannot be an isolated point). In particular, $\Delta u^0 = 1$ in B_δ and $u^t - u^0$ is harmonic and nonnegative in B_δ, thus it is strictly positive in $B_{\delta/2}$ since, by assumption (7-8), it cannot be the zero function.

We explain how to prove that 0 is not a "double point" (case 2b) nor a "cusp" (case 2c). If 0 was a double point, it would be the tangency point of two distinct analytic arcs, but since the expansion of u holds at any order these two arcs should have the same Taylor expansion; hence they are the same arc (so we are in case 2a). If 0 was a cusp point, the cusp should be of the form given in [Sakai 1993, Proposition 4.1]; in particular, up to a rotation, we would have two different functions $\alpha, \beta : [0, \delta) \to \mathbb{R}$ such that

$$\{u^0 = 0\} \cap B_\delta = \{(x, y) : \alpha(x) \le y \le \beta(x),\ x \ge 0\} \cap B_\delta.$$

But by the Lipschitz estimate (1-10), we get, for all $k \ge 2$,

$$|\mathcal{A}_k(x, \alpha(x))| + |\mathcal{A}_k(x, \beta(x)| \lesssim \sup_{\{u^0=0\}\cap B_r} |\partial_n \mathcal{P}_k| \lesssim r^k, \quad x \in [0, \delta).$$

This shows that the graphs of α and β are both tangent to the manifold $\{\mathcal{A}_k = 0\}$ up to order $k-1$. As k was arbitrary, this forces α and β to have the same polynomial expansions. By Proposition 4.1 in [Sakai 1993], this requires that $\alpha \equiv \beta$, a contradiction. □

Appendix A: Proof of Lemma 2.1

We quickly prove Lemma 2.1 for a solution of (2-1) with $f \in C^\delta(B_1)$ for some $\delta \in (0, 1]$. This is just an adaptation of the argument given in [Figalli and Serra 2019].

In this section we will call "universal" any constant depending on n, μ, δ, $\|f\|_{C^\delta(B_1)}$. We also assume that $0 \in \partial\{u > 0\}$ and $0 \in \Sigma(u)$, meaning that there exists a sequence $r_k \downarrow 0$ such that

$$\frac{|\{u = 0\} \cap B_{r_k}|}{|B_{r_k}|} \to 0 \quad \text{as } k \to \infty.$$

Lemma A.1. *There is a universal constant C such that, for all $r \in \left(0, \tfrac{1}{2}\right)$,*

$$r^2 \le C \sup_{\partial B_r} u, \quad \|u\|_{L^\infty(B_r)} \le Cr^2, \quad \|Du\|_{L^\infty(B_r)} \le Cr, \quad \|D^2 u\|_{L^\infty(B_r)} \le C. \tag{A-1}$$

Proof. See [Caffarelli 1998, Theorem 2 and Lemma 5]. □

From this we classify all possible blowups.

Lemma A.2. *Up to subsequences, we have that*

$$r_k^{-2} u(r_k\,\cdot) \rightharpoonup f(0) p_2 \quad \text{in } C^{1,1}_{\text{loc}}(\mathbb{R}^n), \tag{A-2}$$

where p is a 2-homogeneous nonnegative polynomial with $\Delta p_2 = 1$. We denote with $\boldsymbol{P}$ the set of such polynomials.

Proof. Set $v_k := r_k^{-2} u(r_k\,\cdot) \in C^{1,1}(B_{1/r_k})$. By weak* compactness, v_k has a limit point $v \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$ with $v \ge 0,\ v(0) = 0$ and

$$\|\nabla^2 v\|_{L^\infty(\mathbb{R}^n)} \le \liminf_k \|\nabla^2 v_k\|_{L^\infty(B_{1/2r_k})} \le \|\nabla^2 u\|_{L^\infty(B_{1/2})} \le C.$$

Since $0 \in \Sigma(u)$, we also have that $f(r_k\,\cdot)\chi_{\{v_k=0\}} \to f(0)$ in $L^1_{\mathrm{loc}}(\mathbb{R}^n)$. A nonnegative entire function with Laplacian $f(0)$ and bounded Hessian must be in $\boldsymbol{P}$. □

Now we show that the blowups are unique using the Weiss monotonicity formula for the adjusted energy; see [Weiss 1999]. We set

$$W_\lambda(r, v) := r^{-2\lambda}\{D(r, v) - \lambda H(r, v)\}.$$

Lemma A.3. *There is a universal constant C such that, for all $p \in \boldsymbol{P}$ and $r \in (0, 1)$, we have*

$$\frac{d}{dr} W_2(r, u - f(0)p) \ge -Cr^{\delta-1}. \tag{A-3}$$

Proof. Set $v := u - f(0)p$, and directly compute

$$\frac{d}{dr} W_2(r, u - f(0)p) \ge \frac{2}{r^5}\int_{B_1} (2v_r - x\cdot\nabla v_r)\Delta v_r.$$

Notice that $|\Delta v_r + r^2 f_r \chi_{\{u_r=0\}}| \le r^2 \sup_{B_r} |f - f(0)|$. And thus

$$\begin{aligned}\int_{B_1} (2v_r - x\cdot\nabla v_r)\Delta v_r &\ge -r^2 \int_{B_1\cap\{u=0\}} (2v_r - x\cdot\nabla v_r) f_r - C\int_{B_1} |2v_r - x\cdot\nabla v_r| r^{2+\delta} \\ &\ge r^2 \int_{B_1\cap\{u=0\}} \underbrace{(2p_r - x\cdot\nabla p_r)}_{=0} f_r - Cr^{4+\delta} \ge -Cr^{4+\delta}.\end{aligned}$$ □

We deduce uniqueness of blowups and Monneau's almost-monotonicity formula.

Lemma A.4. *For all $p \in \boldsymbol{P}$, we have $W_2(0^+, u - f(0)p) = 0$ and*

$$\frac{d}{dr}(r^{-4} H(r, u - f(0)p)) \ge -Cr^{\delta-1} \quad \textit{for all } r \in (0, 1), \tag{A-4}$$

with C universal. In particular, the blowup is unique at singular points and there exists a universal modulus of continuity $\omega : (0, 1) \to \mathbb{R}$, $\omega(0^+) = 0$, such that

$$r^{-4} H(r, u - f(0)p_2) \le \omega(r) \quad \textit{for all } r \in (0, 1),$$

provided p_2 is the blowup.

Proof. Choose some subsequence $r_k \downarrow 0$ and $p \in \boldsymbol{P}$ such that $r_k^{-2} u_{r_k} \to p$. Then, by Lemma A.3 and (A-2), we have

$$\begin{aligned} W_2(0^+, u - f(0)p) &= \lim_k W_2(r_k, u - f(0)q) = \lim_k D(1, r_k^{-2}u_{r_k} - f(0)q) - 2H(1, r_k^{-2}u_{r_k} - f(0)q) \\ &= \int_{B_1} |\nabla(p - q)|^2 - 2\int_{\partial B_1} (p - q)^2 = 0,\end{aligned}$$

where in the last step we used that p and q are 2-homogeneous and $\Delta p = \Delta q$.

Integrating (A-3) we get $W_2(r, v) \geq -Cr^\delta$, so by direct computation

$$\begin{aligned}\frac{d}{dr}(r^{-4}H(r, u - f(0)p)) &= \frac{2}{r}\left\{W_2(r, v) + \frac{1}{r^4}\int_{B_1} v_r \Delta v_r\right\}\\ &\geq \frac{2}{r}\left\{-Cr^\delta + \int_{B_1\cap\{u_r=0\}} \underbrace{f(0)pf(r\,\cdot\,)}_{\geq 0} - Cr^\delta\right\}\\ &\geq -Cr^{\delta-1}.\end{aligned}$$

This immediately gives uniqueness of the blowups, let us prove the existence of a universal rate of convergence of u to such blowups. Arguing by contradiction, one finds $\epsilon > 0$ and u_k such that

$$r_k^{-4}H(r_k, u_k - f(0)p_{2,k}) \geq \epsilon.$$

Setting $v_k := r_k^{-2}u_k(r_k\,\cdot)$ and arguing as in Lemma A.2, one finds $q \in \boldsymbol{P}$ such that $v_k \to f(0)q$ in $C^1_{\mathrm{loc}}(\mathbb{R}^n)$. Now we get a contradiction using Monneau's monotonicity on u_k and q:

$$\begin{aligned}\epsilon &\leq H(1, v_k - f(0)p_{2,k}) \lesssim H(1, v_k - f(0)q) + H(1, f(0)q - f(0)p_{2,k})\\ &\leq H(1, v_k - f(0)q) + r_k^{-4}H(r_k, u_k - f(0)q) + Cr_k^\delta\\ &\leq 2H(1, v_k - f(0)q) + Cr_k^\delta \to 0\end{aligned}$$

as $k \to \infty$. □

From now on we will denote with $f(0)p_2$ the unique blowup. Let us give several preliminary estimates on the function $v := u - f(0)p_2$.

Lemma A.5. *Take any $p \in \boldsymbol{P}$ and set $v_r := (u - f(0)p)_r$. Then the following estimates hold with universal constants for all $r \in \left(0, \frac{1}{2}\right)$:*

$$\begin{aligned}\Delta v_r &= -r^2 f_r \chi_{\{u_r=0\}} + O(r^{2+\delta}),\\ \|v_r\|_{L^\infty(B_1)} &\lesssim \|v_r\|_{L^2(B_2\setminus B_{1/2})} + r^{2+\delta},\\ r^2|\{u_r = 0\}\cap B_1| &\lesssim \|v_r\|_{L^2(B_2\setminus B_{1/2})} + r^{2+\delta},\\ v_r\Delta v_r &= r^4 f(0) f_r p_2 \chi_{\{u_r=0\}} + v_r O(r^{2+\delta}),\\ \|\nabla v_r\|_{L^2(B_1)} &\lesssim \|v_r\|_{L^2(B_2\setminus B_{1/2})} + r^{2+\delta},\\ [v_r]_{C^{0,\delta/(2\delta+n-1)}} &\lesssim \|v_r\|_{L^2(B_2\setminus B_{1/2})} + r^{2+\delta}, \quad \textit{provided } \dim\{p = 0\} = n - 1.\end{aligned}$$

Proof. The first is a direct computation exploiting Hölder continuity of f.

For the second we notice that

- $\Delta v \leq Cr^\delta$ in B_r and $\Delta v \geq -Cr^\delta$ in $B_r \cap \{u > 0\}$.
- $v \leq 0$ in $B_r \cap \{u = 0\}$.

Hence sub- and superharmonic comparisons give the result as in Lemma 3.5.

For the third, we choose $\chi_{B_1} \le \eta \le \chi_{B_2}$ and compute

$$\begin{aligned}\mu r^2 |\{u_r = 0\} \cap B_1| &\le \int_{B_1} r^2 f_r \chi_{\{u=0\}} \le C r^{2+\delta} - \int_{B_1} \Delta v_r \\ &= \int_{B_1} \underbrace{\Delta\Big(\frac{C r^{2+\delta}|x|^2}{2n} - v_r\Big)}_{\ge 0} \le \int_{B_2} \Delta\Big(\frac{C r^{2+\delta}|x|^2}{2n} - v_r\Big)\eta \\ &\le C_\eta(\|v_r\|_{L^2(B_2\setminus B_1)} + r^{2+\delta}).\end{aligned}$$

The fourth is a direct computation.

Since $v_r \Delta v_r \ge -C r^{2+\delta}|v_r|$, for the fifth we can use the Caccioppoli inequality:

$$\begin{aligned}\int_{B_2} \eta^2 |\nabla v_r|^2 &= -2\int_{B_2} \eta v_r \nabla v_r \cdot \nabla \eta - \int_{B_2} \eta^2 v_r \Delta v_r \\ &\le 4\|\eta \nabla v_r\|_{L^2(B_2)} \|v_r\|_{L^2(B_2\setminus B_1)} + C r^{2+\delta} \int_{B_2} |v_r| \\ &\le \tfrac{1}{2}\|\eta\nabla v_r\|^2_{L^2(B_2)} + C(\|v_r\|^2_{L^2(B_2\setminus B_1)} + r^{2(2+\delta)}),\end{aligned}$$

where η is as above.

For the sixth, assume $p_2 = \frac{1}{2}x_n^2$ and consider, for $0 < t < 1$, $j \ne n$, the function

$$w_\pm(x) := \frac{v_r(x \pm t e_j) - v_r(x)}{t^\delta} = \frac{u_r(x \pm t e_j) - u_r(x)}{t^\delta}.$$

Notice that, with constants uniform in t, we have

$$\|w_\pm\|_{L^2(B_2\setminus B_{1/2})} \lesssim t^{1-\delta}\|\nabla v_r\|_{L^2(B_4\setminus B_{1/4})} \lesssim \|v_r\|_{L^2(B_8\setminus B_{1/8})} + r^{2+\delta}.$$

On the other hand, in $\{u_r > 0\} \cap B_1$ we have $\Delta w_\pm \lesssim r^{2+\delta}$, and in $\{u_r = 0\} \cap B_1$ we have $w_\pm \ge 0$. Thus the function

$$\min\Big\{w_\pm + C r^{2+\delta}\frac{1-|x|^2}{2n}; 0\Big\}$$

is superharmonic in B_1. Using the minimum principle and the previous estimate, we get

$$\min_{\bar{B}_1} w_\pm \ge -C(\|v_r\|_{L^2(B_4\setminus B_{1/4})} + r^{2+\delta}).$$

By the symmetry $w_\pm(x \mp t e_j) = -w_\mp(x)$, we also have the upper bound on a smaller ball. Since all the constants are uniform in t, we conclude, using the following estimate (see Lemma C.1),

$$[f]_{C^{0,\delta/(2\delta+n-1)}(B_1)} \lesssim_{n,\delta} \sum_{j=1}^{n-1} \sup_{x\in B_1, |t|\le 1} \frac{|f(x+te_j) - f(x)|}{|t|^\delta} + \|\partial_n f\|_{L^2(B_2)},$$

valid for every $f \in \mathrm{Lip}(B_2)$. □

Since the blowup is well defined, we can from now on assume to be in the top-dimensional stratum, that is $p_2 = \frac{1}{2}x_n^2$. Arguing as is Section 4, we exploit the truncated frequency ϕ^γ with some $\gamma(\delta) > 2$.

Lemma A.6. *Let $p \in \boldsymbol{P}$ and $\gamma = 2 + \frac{1}{8}\delta$, and set $v := u - f(0)p$. Then there is $\varepsilon = \varepsilon(\delta) > 0$ such that the following inequalities hold for all $r \in (0, 1)$:*

$$\phi^\gamma(r, v) \geq 2 - Cr^\varepsilon, \quad \phi^\gamma(r, v) \leq C, \quad \frac{d}{dr}\phi^\gamma(r, v) \geq -Cr^{\varepsilon-1}, \tag{A-5}$$

with C universal. Furthermore, we also have

$$\frac{\int_{B_1} v_r \Delta v_r}{H(r, v) + r^{2\gamma}} \geq -Cr^\varepsilon. \tag{A-6}$$

Proof. For the first inequality in (A-5), we employ Lemma A.3 and get

$$\begin{aligned}\phi^\gamma(r, v) - 2 &= \frac{D(r) - 2H(r) + (\gamma - 2)r^{2\gamma}}{H(r) + r^{2\gamma}} \\ &\geq \frac{W_2(r)}{r^{-4}H(r) + r^{2\gamma-4}} \geq -Cr^{\delta-2(\gamma-2)},\end{aligned}$$

so we can set $\varepsilon := \frac{3}{4}\delta$. For the second, we need to estimate from below with $-Cr^{\varepsilon-1}$ the term

$$\frac{2}{r(H(r) + r^{2\gamma})}\int_{B_1}(\lambda_r v_r - x \cdot \nabla v_r)\Delta v_r \, dx,$$

where for brevity $\lambda_r := \phi^\gamma(r, v)$ (see Proposition 4.2). Recall that

$$|\Delta v_r + r^2 f_r \chi_{\{u_r=0\}}| \leq Cr^{2+\delta},$$

and estimate each term recalling that p_2 is 2-homogeneous:

$$\begin{aligned}\int_{B_1}(\lambda_r v_r - x \cdot \nabla v_r)\Delta v_r &= -r^2\int_{B_1\cap\{u_r=0\}}(\lambda_r v_r - x \cdot \nabla v_r) f_r - Cr^{2+\delta}\int_{B_1}|\lambda_r v_r - x \cdot \nabla v_r| \\ &\geq r^2\int_{B_1\cap\{u_r=0\}}(\lambda_r p_r - 2p_r) f(0) f_r - Cr^{4+\delta}(\lambda_r + 1) \\ &\geq r^4 \underbrace{(\lambda_r - 2)}_{\geq -Cr^\varepsilon} \underbrace{\int_{B_1\cap\{u_r=0\}} p f(0) f_r}_{\geq 0} - Cr^{4+\delta}(\lambda_r + 1) \\ &\geq -Cr^4(r^\varepsilon + r^\delta(\lambda_r + 1)),\end{aligned}$$

so with crude bounds the frequency solves the ODI

$$\lambda_r' \geq -Cr^{3-2\gamma}(r^\varepsilon + r^\delta(\lambda_r + 1)) \geq -Cr^{3+\varepsilon-2\gamma}(\lambda_r + 1). \tag{A-7}$$

From here we see that $\log(1 + \lambda_r)$ is almost monotone and bounded above by some constant, provided $\gamma < 2 + \frac{1}{2}\varepsilon$. Thus plugging this back into (A-7), we get

$$\lambda_r' \geq -Cr^{3+\varepsilon-2\gamma},$$

which was the claim up to redefining ε. Equation (A-6) follows as in the proof of Lemma A.4 above. □

Hence $\phi^\gamma(0^+, v) \geq 2$ exists for all p, and we want to show that there is a universal number $\alpha_\circ > 0$ such that $\phi^\gamma(0^+, u - f(0)p_2) \geq 2 + 2\alpha_\circ$, provided p_2 is indeed the blowup at 0. Let us show how to conclude from here. Up to universal constants we have the following: By Lemma A.5, we have

$$\|v\|_{L^\infty(B_1)} \lesssim \|v_r\|_{L^2(B_2 \setminus B_{1/2})} + r^{2+\delta}.$$

But $\phi^\gamma \leq C$ in (0, 1), so by Lemma 2.2, we have in turn

$$\|v_r\|^2_{L^2(B_2 \setminus B_{1/2})} \lesssim H\big(\tfrac{1}{2}r\big) + r^{2\gamma}$$

and $\gamma > 2$. Now, since $\phi^\gamma(0^+, v) = 2 + 2\alpha_\circ$, we have, again by Lemma 2.2, that

$$H(r) + r^{2\gamma} \lesssim r^{2(2+2\alpha_\circ)},$$

hence putting everything together we obtain Lemma 2.1:

$$\|v\|_{L^\infty(B_1)} \lesssim r^{2+2\alpha_\circ}.$$

So we are left to show that

$$\lambda_2(0) := \phi^\gamma(0^+, u - f(0)p_2) \geq 2 + 2\alpha_\circ, \tag{A-8}$$

and it is also clear that we can work under the assumption that $\lambda_2(0) \leq 2 + \frac{1}{16}\delta$, otherwise (A-8) holds with $\alpha_\circ = \frac{1}{64}\delta$. The following proposition is crucial and the proof follows the same line of Proposition 5.9 (or also of [Figalli and Serra 2019, Proposition 2.12]). As the only technical complications are settled by the bounds gathered in Lemma A.5, we omit the proof.

Lemma A.7. *Assume $0 \in \Sigma_{n-1}$ and $\lambda_2 \leq 2 + \frac{1}{16}\alpha$. Then the sequence*

$$\tilde{v}_r := \frac{v_r}{\|v_r\|_{L^2(\partial B_1)}}$$

is bounded in $W^{1,2}_{\mathrm{loc}}(\mathbb{R}^n) \cap C^{\delta/(2\delta+n-1)}_{\mathrm{loc}}(\mathbb{R}^n)$. Furthermore, every accumulation point of $\{v_r\}_{r>0}$ solves the Signorini problem (5-4) and is λ_2-homogeneous.

The following combination of Monneau monotonicity and the characterization of blowups will prove (A-8). The proof is in fact very similar to Step 5 in the proof of Proposition 5.9.

Lemma A.8. *There cannot be sequences u_k, f_k, μ_k, δ_k, with $0 \in \partial\{u_\ell > 0\}$ and*

$$\sup_\ell \left(\|f_\ell\|_{C^{\delta_\ell}(B_1)} + \frac{1}{\delta_\ell} + \frac{1}{\mu_\ell} \right) < +\infty,$$

such that $\lambda_2^{(k)} \downarrow 2$, where

$$\lambda_2^{(k)} := \phi^{2+\delta_k/8}(0^+, u_k - f_k(0)p_2^{(k)}).$$

In particular, (A-8) holds for some $\alpha_\circ = \alpha_\circ(n, k, \delta, \|f\|_{C^\delta(B_1)}) \in (0, 1)$.

Proof. **Step 1**. If $\tilde{v}_{r_k} \rightharpoonup q$ in $W^{1,2}_{\mathrm{loc}}$ then, for all $p \in \boldsymbol{P}$, we have

$$\int_{\partial B_1} q(p_2 - p) \geq 0. \tag{A-9}$$

Proof of Step 1. Define $\varepsilon_k^2 := H(r_k, v)$ and notice that, by the growth Lemma 2.2 and the compactness of the trace operator, we have, for k large,

$$r_k^\delta \ll \varepsilon_k \to 0,$$

where we used that $\phi^\gamma(r_k) \le 2 + \frac{1}{100}\delta$ for all k large enough. By Monneau monotonicity (Lemma A.4) applied to p instead of p_2, we have

$$\int_{\partial B_1} (\varepsilon_k \tilde{v}_{r_k} + p_2 - p)^2 + C r_k^\delta \ge \int_{\partial B_1} (p_2 - p)^2,$$

computing the squares and dividing by ε_k we get

$$\varepsilon_k \int_{\partial B_1} \tilde{v}_{r_k}^2 + 2\int_{\partial B_1} \tilde{v}_{r_k}(p_2 - p) + C\frac{r_k^\delta}{\varepsilon_k} \ge 0,$$

and sending $k \uparrow \infty$ we obtain (A-9). We remark that all the constants in these computations are universal.

Step 2. If q is a 2-homogeneous harmonic polynomial such that (A-9) holds for all $p \in \boldsymbol{P}$, then $q \le 0$ on the hyperplane $\{p_2 = 0\}$.

Proof of Step 2. This is exactly [Figalli and Serra 2019, Lemma 2.12].

Step 3. For each $u_k,\ f_k,\ \mu_k,\ \delta_k$ as in the assumptions, Lemma A.6 gives q_k, a $\lambda_2^{(k)}$-homogeneous solution of the Signorini problem with $\|q_k\|_{L^2(\partial B_1)} = 1$. It is easy to see that, by compactness, $q_k \to q$, where q is a 2-homogeneous solution of Signorini with $\|q\|_{L^2(\partial B_1)} = 1$. Thus, q is a harmonic polynomial, nonnegative on the thin obstacle (see Proposition 2.4). But this contradicts Step 1, up to taking a diagonal subsequence. A careful verification that all the bounds are uniform is the same as Step 1 in the proof of Proposition 5.9, and it is not repeated here. □

Appendix B: Adaptations for general right-hand sides

In this section we collect the modification needed to work with a general f and μ.

The main difference is that $u - \mathcal{P}_k$ will not be harmonic in $\{u > 0\} \cap B_r$, but its Laplace operator will be of size $O(r^k)$. This is the size of the error we would have in every estimate. Keeping this in mind, it is clear that all the arguments go through with the same proof, provided we can indeed construct $\mathcal{P}_k$ with the same properties as before. This is not a hard task. We will, for completeness, list also the other modifications needed. Let us remark that all constants that in the case $f \equiv 1$ depend on n and k will now also depend on μ and $\|f\|_{C^k}$.

In the following, we provide a generalization of Section 3.1. We begin with the respective polynomial ansatz, which will additionally depend on the Taylor expansion of f and on the center of expansion. We will denote by $F_{k,x}$ the k-th Taylor polynomial of f at x, that is

$$F_{k,x}(h) := \sum_{|\alpha| \le k} \frac{\partial^\alpha f(x)}{\alpha!} h^\alpha.$$

The sets $\boldsymbol{P}_k$ and V_j are the same as in Section 3.1.

Lemma B.1. *Let $k \geq 2$, $f \in C^{k-1}(B_1)$, $x \in B_1$ and $(p_2, \dots, p_k) \in \boldsymbol{P}_k$ be given. Let ν be any unit vector such that $p_2(h) = \frac{1}{2}(h \cdot \nu)^2$. There exists a unique collection of polynomials*

$$(R_1, \dots, R_{k-1}) \in V_1 \times \cdots \times V_{k-1}$$

such that if we define the polynomial

$$\mathcal{A}_{x,k,\nu}(h) := (\nu \cdot h) + \sum_{j=1}^{k-1} (\nu \cdot h) R_j(y) + \sum_{j=3}^{k} \frac{p_j(h)}{(\nu \cdot h)},$$

then

$$\Delta\big(\tfrac{1}{2} f(x) \mathcal{A}^2_{x,k,\nu}\big)(h) = F_{k-1,x}(h) + O(|h|^k).$$

Furthermore, each R_j is determined (analytically) only by $(p_2, \dots, p_{j+1})$ and the coefficients of $F_{k,x}$. In particular, each R_j does not depend on ν, so $\mathcal{A}_{x,k,-\nu} = -\mathcal{A}_{x,k,\nu}$.

Proof. The proof is almost identical to the proof of Lemma 3.1, the only difference being that we have to take into account the Taylor expansion of f. Let us work out explicitly the case $k = 2$. By a direct computation we find

$$\Delta\big(\tfrac{1}{2} f(x) \mathcal{A}^2_{x,k,\nu}\big)(h) = f(x) + \Delta(2 f(x) p_2 R_1)(h) + O(|h|^2).$$

Thus, the right (and unique) choice for R_1 is

$$R_1 := \frac{1}{2 f(x)} \delta_1^{-1}(F_{1,x}),$$

where the linear isomorphisms $\delta_m : V_m \to V_m$ were introduced in the proof of Lemma 3.1. □

Using Lemma B.1, we can define the polynomial ansatz functions $\mathcal{A}^2_k, \mathcal{P}_k : B_1 \times \boldsymbol{P}_k \to \mathbb{R}[h]$, which now depend explicitly also on the center of expansion x. We set

$$\mathcal{A}_k(x; p_2, \dots, p_{k-1}) := \mathcal{A}^2_{x,k,\nu}, \quad \mathcal{P}_k(x; p_2, \dots, p_{k-1}) := \pi_{\leq k+1}\big(\tfrac{1}{2} f(x) \mathcal{A}^2_{x,k,\nu}\big),$$

and notice that the dependence on f is hidden in the dependence on x. Once again any norm of $\mathcal{P}_k$ is bounded by constants depending on n, k, $\|f\|_{C^{k-1}(B_1)}$ and $|(p_2, \dots, p_k)|$. Furthermore, the function $\mathcal{P}_k(x; \cdot)$ is injective.

With this construction we obtain

$$\Delta\mathcal{P}_k(x; p_2, \dots, p_k)(h) = f(x+h) + O(|h|^k) \quad \text{and} \quad \mathcal{P}_k(x; p_2, \dots, p_k)(h) \geq -C|h|^{k+2}, \tag{B-1}$$

where the big O is a C^k function of x and h. Here comes the only difference with the case in which $f \equiv 1$. When we apply the Laplace operator to the function $v := u(x_\circ + \cdot) - \mathcal{P}_k(x_\circ; p_2, \dots, p_k)$, we get, in $B_r(x_\circ)$,

$$\Delta v = -f(x_\circ + \cdot)\chi_{\{u(x_\circ + \cdot) = 0\}} + O(r^k), \tag{B-2}$$

while in the case $f \equiv 1$ we had $\Delta v = -\chi_{\{u(x_\circ + \cdot)=0\}}$ exactly.

We now state an analog of Proposition 3.3, which contained all crucial properties of the ansatz.

Proposition B.2. *Let $k \geq 2$, $(p_2, \dots, p_k) \in \boldsymbol{P}_k$ and $\tau > 0$ be such that $|(p_2, \dots, p_k)| \leq \tau$. Choose some unit vector ν for which $p_2(x) = \frac{1}{2}(\nu \cdot x)^2$. Let $f \in C^{k-1,1}(B_1)$, and let $|x_\circ| < \frac{1}{2}$. Then the polynomials $\mathcal{A}_k^2(x_\circ; p_2, \dots, p_k)$ and $\mathcal{P}_k(x_\circ; p_2, \dots, p_k)$ satisfy:*

(i) $\Delta \mathcal{P}_k = f(x_\circ + \cdot\,) + O(|\cdot|^k)$ *and* $\partial_e\big(\frac{1}{2}\mathcal{A}_k^2\big) = \partial_e \mathcal{P}_k + O(|\cdot|^{k+1})$ *for any unit vector e.*

(ii) *We have* $\mathcal{P}_k(x_\circ; p_2, \dots, p_k) = \mathcal{P}_{k-1}(x_\circ, p_2, \dots, p_{k-1}) + p_k + O(|\cdot|^{k+1})$.

(iii) *For all* $|h| \leq r_0$, *we have* $\frac{1}{2} \leq |\partial_\nu \mathcal{A}_{x_\circ,k,\nu}(h)| \leq 2$, *and thus*

$$\tfrac{1}{2}|\mathcal{A}_k(h)| \leq \big|\partial_\nu\big(\tfrac{1}{2}\mathcal{A}_k^2(h)\big)\big| \leq 2|\mathcal{A}_k(h)|,$$

where $r_0 = r_0(n, k, \tau, \|f\|_{C^k(B_1)}) \in (0, 1)$.

(iv) *If u is a solution as in* (2-1), $0 \in \Sigma_{n-1}$ *and* $r^{-2}u(r\,\cdot\,) \to p_2$, *then by* (2-3) *we have, for all* $0 < r < \frac{1}{2}$,

$$\sup_{B_r(x_\circ)\cap\{u=0\}} |\partial_\nu \mathcal{P}_k| \leq C r^{1+\alpha_\circ}$$

for some constant $C = C(n, k, \tau, \|f\|_{C^k})$.

Now that the polynomial ansatz has the right formal properties (i.e., (B-1), (B-2) and those collected in Proposition B.2), it is simple to check that the rest of the arguments go through. The rest of this section is a list of the modifications needed to obtain Theorem 1.1 in its full generality.

- Lemmas 3.5 and 3.6 are the same: even if, in $B_r \cap \{u > 0\}$, the function $u - \mathcal{P}_k$ is not harmonic, pointwise we have $\Delta(u - \mathcal{P}_k) = O(r^k)$. The comparison principle we use remains valid in this case.

- The proof of Lemma 3.7 is identical, except for the fact that in Ω our function is not harmonic. This is used only in (3-12), where we have the term $\|\Delta v_r\|_{L^\infty(\tilde{\Omega}\cap B_1)}$, but it can be absorbed into the term Cr^{k+2}.

- In Lemma 4.1, we also pick up an extra term, which, however, is much smaller than the one we are estimating. Indeed we have

$$\int_{B_1} |v_r \Delta v_r| \leq M \int_{B_1\cap\{u=0\}} |v_r| + Cr^k \int_{B_1} |v_r|.$$

The first integral is treated as in Lemma 4.1. Since we have $|v_r| \lesssim r^2$ around a contact point, for the second term we estimate

$$\frac{1}{r}\frac{Cr^{k+2}}{H(r, v) + r^{2\gamma}} \lesssim r^{k+1-\gamma} = r^{\epsilon-1}.$$

As $|x \cdot \nabla v_r| \lesssim r^2$ in B_r, the same reasoning applies to the term $\int_{B_1} |(x \cdot \nabla v_r)\Delta v_r|$.

The rest of Section 4 goes on with exactly the same proofs.

- Section 5 essentially uses the statements of two previous sections as black boxes. The only modification is in the very definition of the sets $\Sigma^{k\text{-th}}$, namely, for $x_\circ \in \Sigma^{k\text{-th}}$, we use the ansatz

$$\mathcal{P}_{k,x_\circ} := \mathcal{P}_k(x_\circ; p_{2,x_\circ}, \dots, p_{k,x_\circ}),$$

which is again continuous in the $x_\circ$ variable. Notice that, to use our argument, we need it to make sense to construct $\mathcal{P}_k$ in $\Sigma^{k\text{-th}}$; hence we require that, at a minimum, $f \in C^k(B_1)$.

• Concerning Section 6, note that we can no longer use that $\Delta \mathcal{P}_k = 0$ in $\{u > 0\}$. Thus in Lemma 6.1 we introduce a small modification, namely in the term

$$\int_{B_1} P \Delta v_r = -\int_{B_1 \cap \{u_r = 0\}} f P + O(r^{k+2}) \int_{B_1} P,$$

but we underline that the extra factor $O(r^{k+2})$ does not affect any subsequent computation.

• In Lemma 6.3, the definition of the barrier function needs to be adapted. Namely (6-5) must be replaced with

$$\begin{cases} \phi_{z,\ell}(z) = 0, & \\ \phi_{z,\ell} \geq 0 & \text{in } B_\rho(z), \\ \Delta\phi_{z,\ell} < r_\ell^2 f(r_\ell \cdot) & \text{in } \overline{B_\rho(z)}, \\ u(r_\ell \cdot) < \phi_{z,\ell} & \text{on } \partial B_\rho(z), \end{cases}$$

so that the proofs of Claim (ii) and Claim (iii) are the same. For Claim (i) we must use the following barrier:

$$\phi_{z,\ell}(x) := \left(1 - \frac{h_\ell}{r_\ell^2}\right) \frac{f(0)}{2} \mathcal{A}_k^2(r_\ell x) + \frac{h_\ell}{4nM} |x' - z'|^2.$$

• Sections 6.1, 6.2, 6.3 do not require further modifications.

• In the proof of Lemma 6.16, the constant of (6-11) depends on $\|\nabla^k f\|_{L^\infty}$.

• As they are based on Section 6, Sections 7.1 and 7.2 do not require any modification.

• In Section 7.3, it is easily checked that Lemma 7.11 works if we assume that $\Delta \mathcal{P} = f + O(r^\kappa)$ instead of $\Delta \mathcal{P} = 1$, and the cleaning works just as before.

Appendix C: Auxiliary lemmas

Lemma C.1. *For every $u \in \operatorname{Lip}(B_2)$, $1 \leq j \leq n$ and $\beta \in (0, 1]$, define*

$$[\delta_j u]_\beta := \sup_{x \in B_1, |t| \leq 1} \frac{|u(x + te_j) - u(x)|}{|t|^\beta}.$$

Then, for all $p > 1$,

$$[u]_{C^{0,\sigma}(B_1)} \leq C\bigg(\sum_{1 \leq j < n} [\delta_j u]_\beta + \|\partial_n u\|_{L^p(B_2)}\bigg),$$

where $C = C(n, \beta, p)$ and $\sigma = \beta(p-1)/(\beta p + n - 1)$.

Proof. By homogeneity we can assume that the right-hand side is 1. Set $h = (0, \dots, 0, r)$, and consider $A_r := B'_{r^\alpha} \times [0, r]$, where $\theta > 0$ is small. By Fubini's theorem, we can find some $z' \in B'_{r^\theta}$ such that

$$\int_0^r |\partial_n u(z', s)|^p \, ds \leq r^{\theta(1-n)} \|\partial_n u\|^p_{L^p(A_r)} \leq r^{\theta(1-n)}.$$

The fundamental theorem of calculus and Hölder's inequality give

$$\begin{aligned}|u(0)-u(h)| &\le |u(0)-u(z',0)|+|u(z',0)-u(z',r)|+|u(z',r)-u(0,r)| \\ &\le 2n|z'|^{\beta}+\int_0^r |\partial_n u(z',s)|\,ds \\ &\lesssim_n r^{\theta\beta}+r^{\theta(1-n)/p}r^{1/p'}.\end{aligned}$$

Since $|h|=r$, u is σ-Hölder continuous for every $\sigma \le \min\{\theta\beta, \theta(1-n)/p+1/p'\}$, and maximizing with respect to the interpolation parameter $\theta>0$ we get the optimal value for σ. □

Let us finally give, for completeness, the proof of our statement of Whitney's theorem for C^∞ functions.

Proof of Theorem 6.15. Let us define the functions $f_\alpha : E \to \mathbb{R}$ for each multi-index $\alpha \in \mathbb{N}^n$ by

$$f_\alpha(x) := \partial^\alpha P_{|\alpha|,x}(\,\cdot\,) \quad (= \partial^\alpha P_{|\alpha|+\ell,x}(0) \text{ for all } \ell \in \mathbb{N}).$$

Assumption (ii) with $k=|\alpha|$ immediately gives

$$|f_\alpha(x)-f_\alpha(y)| = |\partial^\alpha P_{|\alpha|,x}(0)-\partial^\alpha P_{|\alpha|,y}(x-y)| \le C(|\alpha|)|x-y|.$$

Thus each f_α admits a canonical Lipschitz extension to $\bar{E}$, which we don't rename.

For each $x, y, \in \bar{E}$, $m \in \mathbb{N}$ and $|\alpha| \le m$, define the remainder

$$R_{m,\alpha}(x,y) := f_\alpha(x) - \sum_{|\beta|\le m-|\alpha|} \frac{f_{\alpha+\beta}(y)}{\beta!}(x-y)^\beta. \tag{C-1}$$

If we show that each remainder satisfies $|R_{\alpha,m}(x,y)| \le C(m)|x-y|^{|\alpha|-m+1}$ for all $x, y \in E$, then we can extend it by continuity, so that it holds in the full $\overline{E\times E}$, and conclude by applying [Whitney 1934, Theorem I] verbatim. To check this, notice that the left-hand side of assumption (ii) is just (C-1) in disguise:

$$\begin{aligned}R_{m,\alpha}(x,y) &= \partial^\alpha P_{m,x}(0) - \sum_{|\beta|\le m-|\alpha|} \frac{\partial^\beta(\partial^\alpha P_{m,y})(0)}{\beta!}(x-y)^\beta \\ &= \partial^\alpha P_{m,x}(0) - \partial^\alpha P_{m,y}(x-y) = O(|x-y|^{m-|\alpha|+1}),\end{aligned}$$

where we used that polynomials equal their Taylor expansion of sufficiently high degree (and here $\deg \partial^\alpha P_{m,y} \le m-|\alpha|$). This concludes the proof. □

Acknowledgements

The authors would like to thank their supervisors Joaquim Serra and Xavier Ros-Oton for introducing them to the problem and for the continuous guidance and support during the preparation of this work.

Federico Franceschini was supported by the Swiss NSF, Ambizione grant PZ00P2 180042, and by the European Research Council (ERC) under grant agreement no. 721675. Wiktoria Zatoń was supported by the European Research Council (ERC) under grant agreement no. 801867.

References

[Athanasopoulos and Caffarelli 2004] I. Athanasopoulos and L. A. Caffarelli, "Optimal regularity of lower dimensional obstacle problems", *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* (*POMI*) **310** (2004), 49–66. MR Zbl

[Athanasopoulos et al. 2008] I. Athanasopoulos, L. A. Caffarelli, and S. Salsa, "The structure of the free boundary for lower dimensional obstacle problems", *Amer. J. Math.* **130**:2 (2008), 485–498. MR Zbl

[Caffarelli 1977] L. A. Caffarelli, "The regularity of free boundaries in higher dimensions", *Acta Math.* **139**:3-4 (1977), 155–184. MR Zbl

[Caffarelli 1998] L. A. Caffarelli, "The obstacle problem revisited", *J. Fourier Anal. Appl.* **4**:4-5 (1998), 383–402. MR Zbl

[Colombo et al. 2018] M. Colombo, L. Spolaor, and B. Velichkov, "A logarithmic epiperimetric inequality for the obstacle problem", *Geom. Funct. Anal.* **28**:4 (2018), 1029–1061. MR Zbl

[Fernández-Real 2022] X. Fernández-Real, "The thin obstacle problem: a survey", *Publ. Mat.* **66**:1 (2022), 3–55. MR Zbl

[Fernández-Real and Ros-Oton 2021] X. Fernández-Real and X. Ros-Oton, "Free boundary regularity for almost every solution to the Signorini problem", *Arch. Ration. Mech. Anal.* **240**:1 (2021), 419–466. MR Zbl

[Fernández-Real and Ros-Oton 2022] X. Fernández-Real and X. Ros-Oton, *Regularity theory for elliptic PDE*, Zurich Lect. Adv. Math. **28**, Eur. Math. Soc., Berlin, 2022. MR Zbl

[Figalli 2018] A. Figalli, "Free boundary regularity in obstacle problems", exposé 2 in *Journées équations aux dérivées partielles* (Obernai, France, 2018), edited by F. Golse and F. Merle, École Polytechnique, Palaiseau, France, 2018. Zbl

[Figalli and Serra 2019] A. Figalli and J. Serra, "On the fine structure of the free boundary for the classical obstacle problem", *Invent. Math.* **215**:1 (2019), 311–366. MR Zbl

[Figalli et al. 2020] A. Figalli, X. Ros-Oton, and J. Serra, "Generic regularity of free boundaries for the obstacle problem", *Publ. Math. Inst. Hautes Études Sci.* **132** (2020), 181–292. MR Zbl

[Figalli et al. 2023] A. Figalli, X. Ros-Oton, and J. Serra, "The singular set in the Stefan problem", *J. Amer. Math. Soc.* (online publication July 2023).

[Garofalo and Petrosyan 2009] N. Garofalo and A. Petrosyan, "Some new monotonicity formulas and the singular set in the lower dimensional obstacle problem", *Invent. Math.* **177**:2 (2009), 415–461. MR Zbl

[Gilbarg and Trudinger 1983] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, 2nd ed., Grundl. Math. Wissen. **224**, Springer, 1983. Corrected reprint, 1998. MR Zbl

[Kinderlehrer and Nirenberg 1977] D. Kinderlehrer and L. Nirenberg, "Regularity in free boundary problems", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **4**:2 (1977), 373–391. MR Zbl

[Mattila 1995] P. Mattila, *Geometry of sets and measures in Euclidean spaces*: *fractals and rectifiability*, Cambridge Stud. Adv. Math. **44**, Cambridge Univ. Press, 1995. MR Zbl

[Monneau 2003] R. Monneau, "On the number of singularities for the obstacle problem in two dimensions", *J. Geom. Anal.* **13**:2 (2003), 359–389. MR Zbl

[Naber and Valtorta 2017] A. Naber and D. Valtorta, "Volume estimates on the critical sets of solutions to elliptic PDEs", *Comm. Pure Appl. Math.* **70**:10 (2017), 1835–1897. MR Zbl

[Sakai 1991] M. Sakai, "Regularity of a boundary having a Schwarz function", *Acta Math.* **166**:3-4 (1991), 263–297. MR Zbl

[Sakai 1993] M. Sakai, "Regularity of free boundaries in two dimensions", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **20**:3 (1993), 323–339. MR Zbl

[Savin and Yu 2023] O. Savin and H. Yu, "Regularity of the singular set in the fully nonlinear obstacle problem", *J. Eur. Math. Soc.* **25**:2 (2023), 571–610. MR Zbl

[Schaeffer 1977] D. G. Schaeffer, "Some examples of singularities in a free boundary", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **4**:1 (1977), 133–144. MR Zbl

[Weiss 1999] G. S. Weiss, "A homogeneity improvement approach to the obstacle problem", *Invent. Math.* **138**:1 (1999), 23–50. MR Zbl

[Whitney 1934] H. Whitney, "Analytic extensions of differentiable functions defined in closed sets", *Trans. Amer. Math. Soc.* **36**:1 (1934), 63–89. MR Zbl

Received 19 Dec 2022. Revised 18 Jul 2023. Accepted 31 Aug 2023.

FEDERICO FRANCESCHINI: ffederico@ias.edu
Department of Mathematics, ETH Zürich, Zürich, Switzerland
Current address: *School of Mathematics, Institute for Advanced Study, Princeton, NJ, United Stated*

WIKTORIA ZATOŃ: wiktoria.beckmann@gmail.com
Institut für Mathematik, Universität Zürich, Zürich, Switzerland

ANALYSIS AND PDE
Vol. 18 (2025), No. 1, pp. 265–278
DOI: 10.2140/apde.2025.18.265

DIMENSIONS OF FURSTENBERG SETS AND AN EXTENSION OF BOURGAIN'S PROJECTION THEOREM

PABLO SHMERKIN AND HONG WANG

We show that the Hausdorff dimension of (s,t)-Furstenberg sets is at least $s+\frac{1}{2}t+\epsilon$, where $\epsilon>0$ depends only on s and t. This improves the previously best known bound for $2s<t\le 1+\epsilon(s,t)$, in particular providing the first improvement since 1999 to the dimension of classical s-Furstenberg sets for $s<\frac{1}{2}$. We deduce this from a corresponding discretized incidence bound under minimal nonconcentration assumptions that simultaneously extends Bourgain's discretized projection and sum-product theorems. The proofs are based on a recent discretized incidence bound of T. Orponen and the first author and a certain duality between (s,t) and $\left(\frac{1}{2}t, s+\frac{1}{2}t\right)$-Furstenberg sets.

1. Introduction and main results

1.1. ***Dimension of Furstenberg sets.*** Let $s\in(0,1]$. We say that a set $K\subset\mathbb{R}^2$ is an s-Furstenberg set if for almost all directions $\theta\in S^1$ there is a line ℓ_θ in direction θ such that $\dim_{\mathrm{H}}(K\cap\ell_\theta)\ge s$. Motivated by work of Furstenberg and by the Szemerédi–Trotter theorem in incidence geometry, T. Wolff [1999] posed the problem of estimating the smallest possible dimension $\gamma(s)$ of an s-Furstenberg set. Using elementary geometric arguments, Wolff showed that

$$\gamma(s)\ge\max\left(2s, s+\tfrac{1}{2}\right).$$

Note that both bounds coincide for $s=\frac{1}{2}$. J. Bourgain [2003], building up on work of N. Katz and T. Tao [2001], proved that $\gamma\left(\frac{1}{2}\right)>1+\epsilon$ for some small universal constant $\epsilon>0$; this is much deeper. Much more recently, T. Orponen and the first author [Orponen and Shmerkin 2023] established a similar improvement for $s\in\left(\frac{1}{2},1\right)$:

$$\gamma(s)\ge 2s+\epsilon(s),$$

where $\epsilon(s)>0$ for $s\in\left(\frac{1}{2},1\right)$. For the case $s<\frac{1}{2}$, the first author [Shmerkin 2022] recently obtained a similar improvement, with the significant caveat that it involves only the *packing* dimension of s-Furstenberg sets (which can be larger than Hausdorff dimension). In this paper, as a corollary of our main result we obtain the corresponding Hausdorff dimension improvement:

Shmerkin is supported by an NSERC Discovery Grant. Wang is supported by NSF grants DMS-2055544 and DMS-2238818 .
MSC2020: primary 28A80; secondary 28A75, 28A78.
Keywords: Furstenberg sets, Bourgain's projection theorem, projections, incidences, Hausdorff dimension, sum-product, discretized sets.

Theorem 1.1. *For every $s \in (0,1)$ there is $\epsilon(s) > 0$ such that every s-Furstenberg set K satisfies*

$$\dim_{\mathrm{H}}(K) \geq s + \tfrac{1}{2} + \epsilon(s).$$

Recently, there has been much interest in the following generalization of Furstenberg sets: we say that $K \subset \mathbb{R}^2$ is an (s,t)-*Furstenberg set* if there is a family of lines $\mathcal{L}$ with $\dim_{\mathrm{H}}(\mathcal{L}) \geq t$ such that

$$\dim_{\mathrm{H}}(K \cap \ell) \geq s, \quad \ell \in \mathcal{L}.$$

Since lines are a two-dimensional manifold, the Hausdorff dimension of $\mathcal{L}$ is well-defined. Classical s-Furstenberg sets are, of course, $(s,1)$-Furstenberg sets. The central problem, initiated by U. Molter and E. Rela [2012], is to estimate $\gamma(s,t)$, the smallest possible Hausdorff dimension of (s,t)-Furstenberg sets; this can be seen as a continuous analog of the Szemerédi–Trotter incidence bound. Recent works investigating this problem include [Lutz and Stull 2020; Héra et al. 2022; Di Benedetto and Zahl 2021; Dąbrowski et al. 2022; Fu and Ren 2024]. The best currently known bounds are summarized as follows. Suppose first that $t \leq 1 + \epsilon'(s,t)$ (where $\epsilon'(s,t)$ is a small positive parameter). Then (see [Molter and Rela 2012; Lutz and Stull 2020; Héra et al. 2022; Orponen and Shmerkin 2023])

$$\gamma(s,t) \geq \begin{cases} s+t & \text{if } t \leq s, \\ 2s + \epsilon(s,t) & \text{if } s \leq t \leq 2s - \epsilon'(s,t), \\ s + \frac{1}{2}t & \text{if } 2s - \epsilon'(s,t) \leq t. \end{cases}$$

Suppose now that $t \geq 1 + \epsilon'(s,t)$. Then (see [Fu and Ren 2024])

$$\gamma(s,t) \geq \begin{cases} 2s + t - 1 & \text{if } s + t \leq 2, \\ s + 1 & \text{if } s + t \geq 2. \end{cases}$$

The bounds $s+t$, $s+1$ are sharp in the respective regimes, but the other bounds are not expected to be sharp. In this article we obtain a small improvement upon the $s+\frac{1}{2}t$ bound:

Theorem 1.2. *Given $s \in (0,1]$, $t \in (0,2]$, there is $\epsilon(s,t) > 0$ such that the following holds. Let K be an (s,t)-Furstenberg set. Then*

$$\dim_{\mathrm{H}}(K) \geq s + \tfrac{1}{2}t + \epsilon(s,t).$$

Theorem 1.1 is an immediate corollary, taking $t = 1$.

1.2. ***Discretized incidence estimates and a strengthening of Bourgain's discretized projection and sum-product theorems.*** Theorem 1.2 is a consequence of the following discretized incidence estimate. See Section 2.2 for the definition of (δ, s, C)-sets of points and tubes.

Proposition 1.3. *Given $s \in (0,1)$ and $t \in (s,2]$, there are $\epsilon, \eta > 0$ such that the following holds for small enough δ: Let $P \subset B^2(0,1)$ be a $(\delta, t, \delta^{-\epsilon})$-set. For each $p \in P$, let $\mathcal{T}(p)$ be a $(\delta, s, \delta^{-\epsilon})$-set of tubes through p with $|\mathcal{T}(p)| \geq M$. Then the union $\mathcal{T} = \bigcup_{p \in P} \mathcal{T}(p)$ satisfies*

$$|\mathcal{T}| \geq M\delta^{-(t/2+\eta)}. \tag{1-1}$$

By duality between points and lines (see, e.g., [Orponen and Shmerkin 2023, Theorem 3.2] and the discussion afterwards) we obtain the following corollary of Proposition 1.3 which more closely resembles the Furstenberg set problem. In the statement, $|P|_\delta$ stands for the δ-covering number of P (see Section 2.1).

Corollary 1.4. *Given $0 < s < 1$, $s < t$, there are $\epsilon, \eta > 0$ such that the following holds for small enough dyadic δ: Let $\mathcal{L}$ be a $(\delta, t, \delta^{-\epsilon})$-set of lines intersecting $B^2(0, 1)$. For each $\ell \in \mathcal{L}$, let $P(\ell)$ be a $(\delta, s, \delta^{-\epsilon})$-set contained in $\ell^{(\delta)}$. Suppose $|P(\ell)|_\delta \geq M$ for all $\ell \in \mathcal{L}$. Then the union $P = \bigcup_{\ell\in\mathcal{L}} P(\ell)$ satisfies*

$$|P|_\delta \geq M\delta^{-(t/2+\eta)}.$$

Note that $M \geq \delta^{-(s-\epsilon)}$ and hence (up to changing the values of η, ϵ) we also have the conclusion $|\mathcal{T}| \geq \delta^{-(s+t/2+\eta)}$. In turn, by [Héra et al. 2022, Lemma 3.3] this yields Theorem 1.2.

For comparison's sake, and because it plays a crucial role in the proof of Proposition 1.3, we recall [Orponen and Shmerkin 2023, Theorem 3.2].

Theorem 1.5. *Given $s \in (0, 1)$ and $t \in (s, 2]$, there are $\epsilon, \eta > 0$ such that the following holds for all small enough dyadic δ: Let $\mathcal{T}$ be a $(\delta, t, \delta^{-\epsilon})$-set of dyadic δ-tubes. Assume that for every $T \in \mathcal{T}$ there exists a $(\delta, s, \delta^{-\epsilon})$-set $P(T)$ such that $p \in T$ for all $p \in \mathcal{P}(T)$. Then*

$$\Big|\bigcup_{T\in\mathcal{T}} P(T)\Big| \geq \delta^{-2s-\eta}.$$

The nonconcentration assumption on $\mathcal{T}(p)$ in Proposition 1.3 is quite mild, since M can potentially be much larger than δ^{-s} (and a $\delta^{-\epsilon}$ factor is also allowed). What about P? *Some* nonconcentration is needed, as the following standard example shows: if $P = B(x, r)$ and $\mathcal{T}(p)$ is the set of tubes through p with slopes in a fixed (δ, s, C)-set $\mathcal{S}$, then

$$|\mathcal{T}| \sim |\mathcal{S}|_\delta \cdot \frac{r}{\delta} \sim |\mathcal{T}(p)| \cdot |P|_\delta^{1/2}.$$

A similar estimate holds if P is very dense in the union of a small number of r-balls. The next result asserts that under a minimal single-scale nonconcentration assumption on P that rules out this scenario, there is a gain over the "trivial" estimate of Corollary 2.6:

Theorem 1.6. *Given $s, u \in (0, 1)$, there are $\epsilon, \eta > 0$ such that the following holds for small enough dyadic δ: Let $P \subset B^2(0, 1)$ be set such that*

$$|P \cap B(x, \delta|P|_\delta^{1/2})|_\delta \leq \delta^u |P|_\delta, \quad x \in B^2(0, 1). \tag{1-2}$$

For each $p \in P$, let $\mathcal{T}(p)$ be a $(\delta, s, \delta^{-\epsilon})$-set of dyadic tubes through p with $|\mathcal{T}(p)| \geq M$. Then the union $\mathcal{T} = \bigcup_{p\in P} \mathcal{T}(p)$ satisfies

$$|\mathcal{T}| \geq \delta^{-\eta} M |P|_\delta^{1/2}. \tag{1-3}$$

This result extends Proposition 1.3 (however, the proposition is used in the proof) and due to the minimal nonconcentration assumption it provides new information even when $|P|_\delta \gg \delta^{-1}$. Perhaps more significantly, Theorem 1.6 generalizes Bourgain's celebrated discretized projection theorem [2010, Theorem 3] (or even the refined version with single-scale nonconcentration in [Shmerkin 2023, Theorem 1.7]). Roughly speaking, Bourgain's theorem corresponds to the special "product" case in which the slopes of the tubes in $\mathcal{T}(p)$ are (nearly) independent of p; see Section 5 for the details. This connection between Furstenberg sets and projections is well known; see, e.g., [Oberlin 2014] or [Orponen and Shmerkin 2023, §3.2]. Bourgain's discretized projection theorem is used in the proof of Theorem 3.2 of the latter work (recalled as Theorem 1.5 above), which is in turn used to prove Theorem 1.6, so this does not provide a new proof of the projection theorem. Using a well-known argument of G. Elekes, Theorem 1.6 (or rather its dual formulation below) also easily recovers Bourgain's discretized sum-product theorem [Bourgain 2003, Theorem 0.3; Bourgain and Gamburd 2008, Proposition 3.2]. The details are sketched in Section 5. It is worth noting that although Bourgain's discretized sum-product and projection theorems are closely connected to each other, deducing either from the other takes a substantial amount of work, while they are both rather direct corollaries of Theorem 1.6.

By duality between points and lines (again we refer to [Orponen and Shmerkin 2023, Theorem 3.2] for details), we have the following corollary of Theorem 1.6:

Corollary 1.7. *Given $s, u \in (0,1)$, there are $\epsilon, \eta > 0$ such that the following holds for small enough dyadic δ: Let $\mathcal{T} \subset \mathcal{T}^{\delta}$ be a set of dyadic tubes such that*

$$|\{T \in \mathcal{T} : T \subset \boldsymbol{T}\}| \le \delta^{u} |\mathcal{T}|$$

for all $(\delta|\mathcal{T}|^{1/2})$-tubes $\boldsymbol{T}$.

For each $T \in \mathcal{T}$, let $P(T) \subset T$ be a $(\delta, s, \delta^{-\epsilon})$-set with $|P(T)|_{\delta} \ge M$. Then $P = \bigcup_{T \in \mathcal{T}} P(T)$ satisfies

$$|P| \ge \delta^{-\eta} M |T|^{1/2}.$$

1.3. ***Sketch of proof.*** Many "ϵ-improvements" in discretized geometry are obtained by showing that, in the absence of it, the relevant geometric object has a rigid structure that eventually is shown to contradict some previously known bounds (often involving some other "ϵ-improvement"). This is also the approach we take here. By the simple elementary bounds in Lemmas 2.4–2.5, one obtains the improved bound in Proposition 1.3 unless

$$\begin{aligned} |P \cap T|_{\delta} &\approx |P|_{\delta}^{1/2}, \quad T \in \mathcal{T}, \\ |P|_{\delta} &\approx \delta^{-t}. \end{aligned} \tag{1-4}$$

(In this section the notation $\approx$ should be interpreted informally as "up to small powers of δ"). A first ingredient of the proof is showing that (after suitable refinements of P and $\mathcal{T}$) one also gets the desired conclusion unless

$$|P \cap T \cap Q|_{\delta} \approx |P \cap Q|_{\delta}^{1/2} \tag{1-5}$$

for all $T \in \mathcal{T}$, all Δ-squares Q intersecting P and a "dense" set of scales $\delta < \Delta < 1$. To see this, we combine the elementary bounds applied to $P \cap Q$ (and a subsystem of tubes passing through $P \cap Q$) with an induction-on-scales mechanism from [Orponen and Shmerkin 2023], recalled as Proposition 2.7 below.

The relation (1-5) can be shown to imply that either one gets the improved bound we are seeking, or P intersects each tube T in a $\left(\delta, \frac{1}{2}t\right)$-set. By (known) elementary arguments $\mathcal{T}$ is a $\left(\delta, s + \frac{1}{2}t\right)$-set of tubes. Hence P is a discretized $\big(\tfrac{1}{2}t, s+\tfrac{1}{2}t\big)$-Furstenberg set. But it follows from Theorem 1.5 that $|P|_\delta \geq \delta^{-t-\eta}$ for some $\eta = \eta\big(\tfrac{1}{2}t, s+\tfrac{1}{2}t\big) > 0$. This contradicts (1-4), showing the impossibility of the rigid configuration described by (1-5) and hence establishing Proposition 1.3. To our knowledge this dual relationship between (s,t) and $\big(\tfrac{1}{2}t, s+\tfrac{1}{2}t\big)$-Furstenberg sets hadn't been noticed before.

We obtain Theorem 1.6 by applying Proposition 1.3 to each scale in a multiscale decomposition of P into "nontrivial Frostman pieces" that was established in [Shmerkin 2023], and is recalled as Theorem 2.10 below. The scales are combined together by another application of Proposition 2.7.

2. Preliminaries

2.1. ***Notation.*** The notation $A \lesssim B$ or $A = O(B)$ stands for $A \leq C \cdot B$ for some constant $C > 0$, and similarly for $A \gtrsim B$ and $A \sim B$. The δ-covering number of a set X (in a metric space) is defined as the smallest number of δ-balls needed to cover X, and is denoted by $|X|_\delta$. The open r-neighbourhood of a set X is denoted by $X^{(r)}$.

2.2. ***(δ, s)-sets of points and tubes.*** Given $r \in 2^{-\mathbb{N}}$, we denote the family of (half-open) dyadic cubes of side-length in $\mathbb{R}^d$ by $\mathcal{D}_r$. The set of cubes in $\mathcal{D}_r$ intersecting a set X is denoted by $\mathcal{D}_r(X)$.

In this article, we work with the following notion of discretization of sets of dimension s:

Definition 2.1 ((δ, s, C)-set). Let $P \subset \mathbb{R}^d$ be a bounded nonempty set, $d \geq 1$. Let $\delta \in 2^{-\mathbb{N}}$, $0 \leq s \leq d$, and $C > 0$. We say that P is a (δ, s, C)-*set* if

$$|P \cap Q|_\delta \leq C \cdot |P|_\delta \cdot r^s, \quad Q \in \mathcal{D}_r(\mathbb{R}^d),\ \delta \leq r \leq 1. \tag{2-1}$$

If $P \subset \mathcal{D}_\delta$ (so P is a family of dyadic cubes), we will abuse notation by identifying P with $\bigcup P$, so it makes sense to speak of (δ, s, C)-sets of dyadic cubes.

We also need to work with discretized families of tubes:

Definition 2.2 (dyadic δ-tubes). Let $\delta \in 2^{-\mathbb{N}}$. A *dyadic δ-tube* is a set of the form

$$\boldsymbol{D}(p) := \{(x, y) : y = ax + b \text{ for some } (a, b) \in p\},$$

where $p \in \mathcal{D}_\delta([-1, 1) \times \mathbb{R})$. The collection of all dyadic δ-tubes is denoted by

$$\mathcal{T}^\delta := \{\boldsymbol{D}(p) : p \in \mathcal{D}_\delta([-1, 1) \times \mathbb{R})\}.$$

A finite collection of dyadic δ-tubes $\{\boldsymbol{D}(p)\}_{p \in \mathcal{P}}$ is called a (δ, s, C)-set if $\mathcal{P}$ is a (δ, s, C)-set.

We remark that a dyadic δ-tube is not exactly a δ-neighbourhood of some line, but the intersection of

$$T = \boldsymbol{D}([a, a+\delta] \times [b, b+\delta])$$

with some fixed bounded set B satisfies $\ell_{a,b}^{(c\delta)} \cap B \subset T \cap B \subset \ell_{a,b}^{(C\delta)}$, where $\ell_{a,b} = \big\{y = \big(a + \tfrac{1}{2}\delta\big)x + \big(b + \tfrac{1}{2}\delta\big)\big\}$ and c, C depend only on B.

An elementary but important observation is that tubes in $\mathcal{T}^\delta$ that intersect a fixed square $p \in \mathcal{D}_\delta$ are parametrized by their slope in a bilipschitz way. In particular, if $\mathcal{T}(p) \subset \mathcal{T}^\delta$ is a family of tubes intersecting $p \in \mathcal{D}_\delta$, then $\mathcal{T}(p)$ is a (δ, s, C)-set if and only if the slopes of tubes in $\mathcal{T}(p)$ form a (δ, s, C') set for $C' \sim C$; see [Orponen and Shmerkin 2023, Corollary 2.12] for the precise statement.

2.3. *Elementary incidence bounds.* We now collect some elementary incidence bounds; they correspond in various ways to the lower bound $s + \frac{1}{2}t$ for the dimension of (s, t)-Furstenberg sets. We state our bounds in terms of the following notion:

Definition 2.3. Fix $\delta \in 2^{-\mathbb{N}}$, $s \in [0, 1]$, $C > 0$, $M \in \mathbb{N}$. We say that a pair $(P, \mathcal{T}) \subset \mathcal{D}_\delta \times \mathcal{T}^\delta$ is a (δ, s, C, M)-*nice configuration* if for every $p \in P$ there exists a (δ, s, C)-set $\mathcal{T}(p) \subset \mathcal{T}$ with $|\mathcal{T}(p)| = M$ and such that $T \cap p \neq \varnothing$ for all $T \in \mathcal{T}(p)$.

Lemma 2.4. *Let $(P, \mathcal{T})$ be a (δ, s, C, M)-nice configuration. Then for any δ-tube T (not necessarily in $\mathcal{T}$),*

$$|\mathcal{T}| \gtrsim C^{-1/s} \cdot |T \cap P|_\delta \cdot M.$$

Proof. We may assume P is δ-separated. Fix $p \in T \cap P$. Since $\mathcal{T}(p)$ is a (δ, s, C)-set, there is a subset $\mathcal{T}'(p) \subset \mathcal{T}(p)$ such that $|\mathcal{T}'(p)| \geq \frac{1}{2}|\mathcal{T}(p)| = \frac{1}{2}M$ and each tube $T' \in \mathcal{T}'(p)$ makes an angle $\gtrsim C^{-1/s}$ with the direction of T. In turn, this implies that the sets $\mathcal{T}'(p)$, $p \in T \cap P$ have overlap $\lesssim C^{1/s}$. This gives the claim. □

Lemma 2.5. *Let $(P, \mathcal{T})$ be a (δ, s, C, M)-nice configuration. Suppose $|T \cap P| \leq K$ for all $T \in \mathcal{T}$. Then*

$$|\mathcal{T}| \geq K^{-1} \cdot |P| \cdot M.$$

Proof. We have

$$|P| \cdot M = \sum_{p \in P} |\mathcal{T}_p| = \sum_{T \in \mathcal{T}} |\{p : T \in \mathcal{T}_p\}| \leq \sum_{T \in \mathcal{T}} |T \cap P| \leq |\mathcal{T}| \cdot K.$$ □

Corollary 2.6. *Let $(P, \mathcal{T})$ be a (δ, s, C, M)-nice configuration. Then*

$$|\mathcal{T}| \gtrsim C^{-1/s} \cdot |P|^{1/2} \cdot M.$$

Moreover, $\mathcal{T}$ contains a $\left(\delta, s + \frac{1}{2}t, \log(1/\delta)^{O(1)} C^{1/s}\right)$-set of δ-tubes.

Proof. If $|T \cap P| \geq |P|^{1/2}$ for some $T \in \mathcal{T}$, we apply Lemma 2.4, otherwise we apply Lemma 2.5. In any case we get the first claim.

Let $\mathcal{L}$ be the set of lines corresponding to tubes in $\mathcal{T}$ (or, equivalently, the δ-neighbourhood of the central lines of tubes in T). It follows from the first claim and two dyadic pigeonholings that the Hausdorff content of $\mathcal{L}$ satisfies

$$\mathcal{H}^{s+t/2}_\infty(\mathcal{L}) \gtrsim \log(1/\delta)^{-O(1)} C^{-1/s}.$$

See, e.g., the proof of [Héra et al. 2022, Lemma 3.3] or [Orponen et al. 2024, Lemma 3.5, in particular, equation (3.9)]. The conclusion then follows from the discrete version of Frostman's lemma [Fässler and Orponen 2014, Proposition A.1] (which is stated in $\mathbb{R}^3$ but works just as well in the space of lines). □

2.4. ***A multiscale incidence bound.*** We next recall [Orponen and Shmerkin 2023, Proposition 5.2]. Fix two dyadic scales $0 < \delta < \Delta \leq 1$ and families $P_0 \subset \mathcal{D}_\delta$ and $\mathcal{T}_0 \subset \mathcal{T}^\delta$. For $Q \in \mathcal{D}_\Delta$ and $\boldsymbol{T} \in \mathcal{T}^\Delta$, we define

$$P_0 \cap Q = \{p \in P_0 : p \subset Q\} \quad \text{and} \quad \mathcal{T}_0 \cap \boldsymbol{T} := \{T \in \mathcal{T}_0 : T \subset \boldsymbol{T}\}.$$

We also write

$$\mathcal{D}_\Delta(P_0) = \{Q \in \mathcal{D}_\Delta : P_0 \cap Q \neq \varnothing\},$$
$$\mathcal{T}^\Delta(\mathcal{T}_0) = \{\boldsymbol{T} \in \mathcal{T}^\Delta : \mathcal{T}_0 \cap \boldsymbol{T} \neq \varnothing\}.$$

In the next proposition, for $\Delta \in 2^{-\mathbb{N}}$ and $Q \in \mathcal{D}_\Delta$, the map $S_Q : \mathbb{R}^2 \to \mathbb{R}^2$ is the homothety that maps Q to the square $[0, 1)^2$, and $S_Q(P_0) = \{S_Q(p) : p \in P_0\}$. Furthermore, the notation $A \lessapprox_\delta B$ stands for $A \leq \log(1/\delta)^C B$ for a constant $C > 0$, and likewise for $A \approx_\delta B$.

Proposition 2.7 [Orponen and Shmerkin 2023, Proposition 5.2]. *Fix dyadic numbers* $0 < \delta < \Delta \leq 1$. *Let* $(P_0, \mathcal{T}_0)$ *be a* (δ, s, C, M)*-nice configuration. Then there exist sets* $P \subset P_0$ *and* $\mathcal{T}(p) \subset \mathcal{T}_0(p)$, $p \in P$, *such that defining* $\mathcal{T} = \bigcup_{p\in P} \mathcal{T}(p)$ *the following hold*:

(i) $|\mathcal{D}_\Delta(P)| \approx_\delta |\mathcal{D}_\Delta(P_0)|$ *and* $|P \cap Q| \approx_\delta |P_0 \cap Q|$ *for all* $Q \in \mathcal{D}_\Delta(P)$.

(ii) $|\mathcal{T}(p)| \gtrapprox_\delta |\mathcal{T}_0(p)| = M$ *for* $p \in P$.

(iii) $(\mathcal{D}_\Delta(P), \mathcal{T}^\Delta(\mathcal{T}))$ *is* $(\Delta, s, C_\Delta, M_\Delta)$*-nice for some* $C_\Delta \approx_\delta C$ *and* $M_\Delta \geq 1$.

(iv) *For each* $Q \in \mathcal{D}_\Delta(P)$ *there exist* $C_Q \approx_\delta C$, $M_Q \geq 1$, *and a family of tubes* $\mathcal{T}_Q \subset \mathcal{T}^{\delta/\Delta}$ *such that* $(S_Q(P \cap Q), \mathcal{T}_Q)$ *is* $(\delta/\Delta, s, C_Q, M_Q)$*-nice.*

Furthermore, the families $\mathcal{T}_Q$ *can be chosen so that*

$$\frac{|\mathcal{T}_0|}{M} \gtrapprox_\delta \frac{|\mathcal{T}^\Delta(\mathcal{T})|}{M_\Delta} \cdot \left(\max_{Q\in\mathcal{D}_\Delta(P)} \frac{|\mathcal{T}_Q|}{M_Q}\right). \tag{2-2}$$

2.5. ***Uniformization.*** Next, we recall a basic lemma asserting the existence of large uniform subsets. See, e.g., [Orponen and Shmerkin 2023, Lemma 7.3] for the proof.

Definition 2.8. Let $N \geq 1$, and let

$$\delta = \Delta_N < \Delta_{N-1} < \cdots < \Delta_1 \leq \Delta_0 = 1$$

be a sequence of dyadic scales. We say that a set $P \subset [0, 1)^2$ is $(\Delta_j)_{j=1}^N$*-uniform* if there is a sequence $(K_j)_{j=1}^N$ such that $|P \cap Q|_{\Delta_j} = K_j$ for all $j \in \{1, \dots, N\}$ and all $Q \in \mathcal{D}_{\Delta_{j-1}}(P)$.

Lemma 2.9. *Given* $P \subset [0, 1)^2$ *and a sequence* $\delta = \Delta_N < \Delta_{N-1} < \cdots < \Delta_1 \leq \Delta_0 = 1$ *of dyadic numbers*, $N \geq 1$, *there is a* $(\Delta_j)_{j=1}^N$*-uniform set* $P' \subset P$ *such that*

$$|P'|_\delta \geq (4N^{-1}\log(1/\delta))^{-N}|P|_\delta. \tag{2-3}$$

Note that if the number N of scales is independent of δ, then the lower bound on $|P'|_\delta$ can be simplified as

$$|P'|_\delta \geq C_N^{-1} \cdot \log(1/\delta)^{-C_N} \cdot |P|_\delta,$$

with C_N independent of δ.

2.6. ***A multiscale decomposition.*** To conclude this section, we recall the multiscale decomposition into "Frostman pieces" of uniform sets that satisfy a single-scale nonconcentration assumption provided by [Shmerkin 2023, Theorem 4.1].

Theorem 2.10. *For every $u > 0$ and $\epsilon > 0$ there are $\xi = \xi(u) > 0$ and $\tau = \tau(\epsilon) > 0$ such that the following holds for all sufficiently small $\rho \le \rho_0(\epsilon)$ and $n \ge n_0(\rho, \epsilon)$: Let P be a $(\rho^j)_{j=1}^n$-uniform set and write $\delta = \rho^n$. Suppose*

$$|P \cap B(x, \delta|P|_\delta^{1/2})|_\delta \le \delta^u |P|_\delta \quad \textit{for all } x. \tag{2-4}$$

Then there exists a collection of dyadic scales

$$\delta = \Delta_N < \Delta_{N-1} < \cdots < \Delta_1 < \Delta_0 = 1, \quad N \le N_0(\epsilon),$$

each of which is a power of ρ, and numbers $\alpha_0, \ldots, \alpha_{N-1} \in [0, 2]$ such that, defining $\lambda_j = \Delta_{j+1}/\Delta_j$, the following hold:

(i) *For each j and each $Q \in \mathcal{D}_{\Delta_j}(P)$,*

$$|P \cap Q \cap B(x, r\Delta_j)|_{\Delta_{j+1}} \le \lambda_j^{-\epsilon} \cdot r^{\alpha_j} \cdot |P \cap Q|_{\Delta_{j+1}}$$

for all $x \in B^2(0, 1)$ and all $r \in [\lambda_j, 1]$.

(ii) $\sum\{\alpha_j \log(1/\lambda_j) : \lambda_j \le \delta^\tau\} \ge \log|P|_\delta - 2\epsilon \log(1/\delta)$.

(iii) $\sum\{\log(1/\lambda_j) : \alpha_j \in [\xi, d - \xi]$ *and* $\lambda_j \le \delta^\tau\} \ge \xi \log(1/\delta)$.

This is just (a slightly weaker version of) [Shmerkin 2023, Theorem 4.1], although stated using different notation: the measure μ there corresponds to normalized Lebesgue measure on $P(\delta)$ (or the union of δ-squares intersecting δ); ρ corresponds to 2^{-T} in [Shmerkin 2023]; the scales Δ_j correspond to both 2^{-TA_i} and 2^{-TB_i}. The last two claims only concern scale intervals $[\Delta_{j+1}, \Delta_j]$ corresponding to $[2^{-TB_i}, 2^{-TA_i}]$, while for $[\Delta_{j+1}, \Delta_j] = [2^{-TB_{i+1}}, 2^{-TA_i}]$ we simply take $\alpha_j = 0$.

3. Proof of Proposition 1.3

In this section we prove Proposition 1.3. Note that there is no loss of generality in assuming that P is a union of dyadic δ-squares and the tubes in $\mathcal{T}(p)$ are dyadic δ-tubes intersecting p.

The parameter ϵ should be thought of as being much smaller than η (and will be chosen after η). Both ϵ, η will ultimately be chosen in terms of s, t only. We let $N = \lceil \eta^{-1} \rceil$, so that $N = N(s, t)$. Let $\rho = \delta^{1/N}$; without loss of generality, ρ is dyadic.

In the course of this proof, $A \lessapprox B$ stands for $A \le C(s, t)\delta^{-C(s,t)\epsilon} B$. Likewise, a (δ, u)-set is a (δ, u, C)-set for $C \lessapprox 1$.

Replacing P by its $(\rho^j)_{j=1}^N$-uniformization (given by Lemma 2.9), we may assume that P is $(\rho^j)_{j=1}^N$-uniform. Note that the uniformization is still a (δ, t)-set.

We will construct sequences $P_j \subset \mathcal{D}_\delta$, $\mathcal{T}_j \subset \mathcal{T}^\delta$, $j = 0, 1, \ldots, N$, with the following properties:

(a) P_j is $(\rho^j)_{j=1}^N$-uniform.

(b) $P_{j+1} \subset P_j$, $P_0 \subset P$ and $|P_{j+1}| \gtrapprox |P_j|$, $|P_0| \gtrapprox |P|$. In particular, P_N is a (δ, t)-set.

(c) $\mathcal{T}_0(p) = \mathcal{T}(p)$, $\mathcal{T}_{j+1}(p) \subset \mathcal{T}_j(p)$ and $M_{j+1} \sim |\mathcal{T}_{j+1}(p)| \gtrapprox |\mathcal{T}_j(p)|$ for $p \in P_{j+1}$ and some M_{j+1}. In particular, each $\mathcal{T}(p)$, $p \in P_N$, is a (δ, t)-set.

(d) For each j, any $Q \in \mathcal{D}_{\rho^j}(P_j)$ and any δ-tube T,

$$|\mathcal{T}| \gtrapprox M \cdot |P_j|_{\rho^j}^{1/2} \cdot |T \cap P_j \cap Q|_\delta.$$

Recall that $|\mathcal{T}(p)| \geq M$ for each $p \in P$. Pigeonhole a dyadic number M_0 such that $|\mathcal{T}(p)| \sim M_0$ for all $p \in P_0' \subset P$ with $|P_0'| \gtrapprox |P|$. We let P_0 be the $(\rho^j)_{j=1}^N$-uniformization of P_0' given by Lemma 2.9. Then P_0 is a (δ, t)-set, and we take $\mathcal{T}_0(p) = \mathcal{T}(p)$ for $p \in P_0$.

Once $P_j, \mathcal{T}_j$ are defined, we let P_{j+1}', $\mathcal{T}_{j+1}'$ be the objects provided by Proposition 2.7 applied to $(P_j, \mathcal{T}_j)$ at scale $\Delta = \rho^{j+1}$. It follows from Proposition 2.7(i) and the regularity of P_j that $P_{j+1}' \subset P_j$ and $|P_{j+1}'| \gtrapprox P_j$. Pigeonhole a number M_{j+1} such that $|\mathcal{T}_{j+1}'(p)| \sim M_{j+1}$ for all $p \in P_{j+1}'' \subset P_{j+1}'$, where $|P_{j+1}''| \gtrapprox |P_{j+1}'|$. Finally, let P_{j+1} be the $(\rho^j)_{j=1}^N$-uniformization of P_{j+1}'' and $\mathcal{T}_{j+1}(p) = \mathcal{T}_{j+1}'(p)$ for $p \in P_{j+1}$.

Properties (a)–(b) hold by construction. Property (c) follows from Proposition 2.7(ii). To see property (d), let C_Δ, M_Δ, $\mathcal{T}_Q$, C_Q, M_Q be the objects provided by Proposition 2.7(iii)–(iv). By Corollary 2.6,

$$|\mathcal{T}^\Delta(\mathcal{T}_{j+1}')| \gtrapprox M_\Delta \cdot |P_{j+1}|^{1/2},$$

and by Lemma 2.4 and rescaling,

$$|\mathcal{T}_Q| \gtrapprox M_Q \cdot |T \cap P_j \cap Q|_\delta.$$

Putting these facts together with (2-2), we see that property (d) holds.

We pause to observe that $(\mathcal{P}_N, \mathcal{T}_N)$ is a (δ, s, C_N, M_N)-nice configuration, where $C_N \lessapprox 1$ and $M_N \gtrapprox M$.

We now consider several cases. Suppose first that there are $T \in \mathcal{T}_N$, $j \in \{1, \dots, N\}$, and $Q \in \mathcal{D}_{\rho^j}(P_N)$ such that

$$|T \cap P_N \cap Q|_\delta \geq \delta^{-2\eta} \cdot |P_N \cap Q|_\delta^{1/2}.$$

Then, by (d), the uniformity of P_N, and the fact that $P_j \supset P_N$ is a (δ, t)-set, we see that (1-1) holds if ϵ is small enough in terms of s, t, η.

Hence, we assume from now on that

$$|T \cap P_N \cap Q|_\delta \leq \delta^{-2\eta} \cdot |P_N \cap Q|_\delta^{1/2} \tag{3-1}$$

for $j \in \{1, \dots, N\}$, $Q \in \mathcal{D}_{\rho^j}(P_N)$, and $T \in \mathcal{T}_N$.

We consider two further subcases. Suppose first that for at least half of the squares p in P_N, at least half of the tubes $T \in \mathcal{T}_N(p)$ satisfy

$$|T \cap P_N|_\delta \leq \delta^{2\eta} \cdot |P_N|_\delta^{1/2}.$$

Then Lemma 2.5 (applied to suitable restrictions of P_N and $\mathcal{T}_N$) yields (1-1).

We can then assume that

$$|T \cap P_N|_\delta \geq \delta^{2\eta} \cdot |P_N|_\delta^{1/2} \quad \text{for all } T \in \mathcal{T}_N', \tag{3-2}$$

where $\mathcal{T}'_N = \bigcup_{p\in P'_N} \mathcal{T}'(p)$, with $|P'_N| \geq \frac{1}{2}|P_N|$ and $|\mathcal{T}'_N(p)| \geq \frac{1}{2}|\mathcal{T}_N(p)|$. It then follows from (3-1) that, for any $Q \in \mathcal{D}_{\rho^n}(P_N)$ and $T \in \mathcal{T}'_N$,

$$\begin{aligned}|T\cap P_N\cap Q|_\delta &\leq \delta^{-4\eta}(|P_N|_\delta^{-1/2}|P_N\cap Q|_\delta^{1/2})|T\cap P_N|_\delta\\ &\leq \delta^{-5\eta}(\rho^j)^{t/2}|T\cap P_N|_\delta,\end{aligned}$$

using that P_N is a (δ, t)-set and taking ϵ sufficiently small in terms of η, s, t.

Recalling that $N = \lceil \eta^{-1}\rceil$, we deduce that, for any r-ball B_r with $\rho^{j+1} \leq r < \rho^j$,

$$|T\cap P_N\cap B_r|_\delta \lesssim \delta^{-5\eta}(\rho^j)^{t/2}|T\cap P_N|_\delta \leq \delta^{-6\eta}r^{t/2}|T\cap P_N|_\delta,$$

so that $T \cap P_N$ is a $\big(\delta, \frac{1}{2}t, \delta^{-7\eta}\big)$-set for each $T \in \mathcal{T}'$.

Now, taking ϵ small enough in terms of s, t, η, we deduce from Corollary 2.6 that $\mathcal{T}'$ contains a $\big(\delta, s+\frac{1}{2}t, \delta^{-\eta}\big)$-set $\mathcal{T}''$. Then $\{T \cap P_N : T \in \mathcal{T}''\}$ satisfies the assumptions of Theorem 1.5 (with $\frac{1}{2}t$ in place of s and $s+\frac{1}{2}t$ in place of t), provided that η and δ are taken small enough in terms of s, t only. Applying Theorem 1.5, we conclude that $|P_N|_\delta > \delta^{-t-3\eta}$ (again assuming $\eta = \eta(s, t)$ is small enough). The first claim of Corollary 2.6 then yields (1-1).

4. Proof of Theorem 1.6

By iterating Proposition 2.7, we obtain the follow multiscale version.

Corollary 4.1. *Fix $N \geq 2$ and dyadic numbers*

$$0 < \delta = \Delta_N < \Delta_{N-1} < \cdots < \Delta_1 < \Delta_0 = 1.$$

Let $(P_0, \mathcal{T}_0)$ be a (δ, s, C, M)-nice configuration. Then there exists a set $P \subset P_0$ such that the following hold:

(i) $|\mathcal{D}_{\Delta_j}(P)| \approx_\delta |\mathcal{D}_{\Delta_j}(P_0)|$ *and* $|P\cap Q| \approx_\delta |P_0\cap Q|$ *for all* $j \in [1, N]$ *and all* $Q \in \mathcal{D}_{\Delta_j}(P)$.

(ii) *For each* $j \in [0, N-1]$ *and each* $Q \in \mathcal{D}_{\Delta_j}(P)$ *there exist* $C_Q \approx_\delta C$, $M_Q \geq 1$, *and a family of tubes* $\mathcal{T}_Q \subset \mathcal{T}^{\Delta_{j+1}/\Delta_j}$ *such that* $(S_Q(P\cap Q), \mathcal{T}_Q)$ *is* $(\Delta_{j+1}/\Delta_j, s, C_Q, M_Q)$*-nice.*

Furthermore, the families $\mathcal{T}_Q$ can be chosen so that if $Q_j \in \mathcal{D}_{\Delta_j}(P)$, then

$$\frac{|\mathcal{T}_0|}{M} \gtrapprox_\delta \prod_{j=0}^{N-1}\frac{|\mathcal{T}_{Q_j}|}{M_{Q_j}}. \tag{4-1}$$

All the constants implicit in the $\approx_\delta$ notation are allowed to depend on N.

Proof. We proceed by induction in N. The case $N = 2$ follows from Proposition 2.7. Suppose the claim holds for N and let us verify it for $N+1$. First apply Proposition 2.7 with $\delta = \Delta_{N+1}$ and $\Delta = \Delta_N$. Let P', $\mathcal{T}$ be the resulting objects. Property (ii) holds (at the moment for P') for $j = N$ thanks to Proposition 2.7(iv). We then apply the inductive assumption to $(\mathcal{D}_{\Delta_N}(P'), \mathcal{T}^{\Delta_N}(\mathcal{T}))$, which is legitimate by Proposition 2.7(iii). This yields a set $P'' \subset \mathcal{D}_{\Delta_N}$; we define

$$P = \bigcup_{Q\in\mathcal{D}_{\Delta_N}(P'')} P'\cap Q.$$

This ensures that $S_Q(P \cap Q)$, when viewed at scale Δ_{j+1}/Δ_j, equals P'' for $j < N$ and P' for $j = N$, and so property (ii) holds for all $j \in [0, N]$. Finally, (4-1) holds thanks to (2-2) and the inductive assumption. □

Proof of Theorem 1.6. We will eventually take $\eta = \eta(s, u)$ and $\epsilon = \epsilon(\eta, s, u) \ll \eta$. Fix $\rho = \rho(\epsilon) \in (0, 1)$ small enough that the conclusion of Theorem 2.10 holds, and

$$\frac{\log(4 \log(1/\rho))}{\log(1/\rho)} < \epsilon.$$

Then, for $\delta = \rho^n$, the $(\rho^j)_{j=1}^n$-uniformization P' of P given by Lemma 2.9 satisfies $|P'| \geq \delta^\epsilon |P|$. We assume from now on that $\delta = \rho^n$, where n is taken sufficiently large in terms of all other parameters. Take $\epsilon < \min\left(\frac{1}{2}u, \frac{1}{2}\eta\right)$. Then P' satisfies (1-2) with $\frac{1}{2}u$ in place of u, and the conclusion (1-3) for P' implies it for P with $\frac{1}{2}\eta$ in place of η. Hence we assume from now on that P is $(\rho^j)_{j=1}^n$-uniform.

We apply Theorem 2.10 to P. Let $\xi = \xi(u)$, $\tau = \tau(\epsilon)$ be the numbers provided by the theorem, and let $(\Delta_j)_{j=0}^N$ and $(\alpha_j)_{j=1}^N$ be the scales and exponents corresponding to P.

Let $\lambda_j = \Delta_{j+1}/\Delta_j$. We apply Corollary 4.1 to $(P, \mathcal{T})$ and the scales (Δ_j). Let $P' \subset P$ be the resulting set. Since $N \leq N_0(\epsilon)$, the notation $A \lessapprox_\delta B$ in the statement of Corollary 4.1 translates to $A \leq \log(1/\delta)^{C(\epsilon)} B$; in particular, $A \leq \delta^{\epsilon\tau/2} B$ if δ is small enough in terms of ϵ. With these remarks, the $(\Delta_j)_{j=1}^N$-uniformity of P, Corollary 4.1(i) and Theorem 2.10(i) show that if $Q \in \mathcal{D}_{\Delta_j}(P')$, then $S_Q(P' \cap Q)$ is a $(\lambda_j, \alpha_j, \lambda_j^{-\epsilon})$-set whenever $\lambda_j \leq \delta^\tau$.

Let

$$\mathcal{N} = \{j : \lambda_j \leq \delta^\tau,\ \alpha_j \notin [\xi, 2 - \xi]\},$$
$$\mathcal{G} = \{j : \lambda_j \leq \delta^\tau,\ \alpha_j \in [\xi, 2 - \xi]\}.$$

(Here $\mathcal{N}$ stands for "normal" and $\mathcal{G}$ for "good" scales.) It follows from Corollary 4.1(ii) combined with Corollary 2.6 and Proposition 1.3 that, for $Q \in \mathcal{D}_{\Delta_j}(P')$,

$$j \in \mathcal{N} \quad \Longrightarrow \quad |\mathcal{T}_Q| \geq (1/\log\delta)^{C(\epsilon,s)} \cdot M_Q \cdot |S_Q(P \cap Q)|_{\lambda_j}^{1/2},$$
$$j \in \mathcal{G} \quad \Longrightarrow \quad |\mathcal{T}_Q| \geq (1/\log\delta)^{C(\epsilon,s)} \cdot M_Q \cdot \lambda_j^{-(\alpha_j/2 - \eta)},$$

where $\eta = \eta(\xi, s) = \eta(u, s) > 0$. It is indeed possible to take a value of η uniformly over $t \in [\xi, 2 - \xi]$ because the value of η in Proposition 1.3 is robust under perturbations of t. Note that since $\tau = \tau(\epsilon)$, if δ is small enough in terms of ϵ, and $j \in \mathcal{G}$, then $\lambda_j \leq \delta^\tau$ is small enough that Proposition 1.3 is indeed applicable. It follows from Theorem 2.10(i) that

$$|S_Q(P \cap Q)|_{\lambda_j} \geq \lambda_j^{\epsilon - \alpha_j}.$$

Combining these facts with the conclusion (4-1) and the trivial bound $|\mathcal{T}_Q| \geq M_Q$ (applied at scales outside $\mathcal{N} \cup \mathcal{G}$), we obtain

$$\frac{|\mathcal{T}|}{M} \geq \left(\prod_{j\in\mathcal{N}} \lambda_j^{\epsilon/2} \lambda_j^{-\alpha_j/2}\right)\left(\prod_{j\in\mathcal{G}} \lambda_j^{-\eta} \lambda_j^{-\alpha_j/2}\right) \geq \delta^\epsilon \left(\prod_{j\in\mathcal{N}\cup\mathcal{G}} \lambda_j^{-\alpha_j/2}\right)\left(\prod_{j\in\mathcal{G}} \lambda_j^{-\eta}\right) \geq \delta^{3\epsilon} \cdot |P|_\delta^{1/2} \cdot \delta^{-\xi\eta},$$

using Theorem 2.10(ii)–(iii) for the last inequality. Taking $\epsilon < \frac{1}{6}\xi\eta$, this gives the claim with $\frac{1}{2}\xi\eta$ in place of η. □

5. Connection with Bourgain's projection theorem

We conclude by showing how Theorem 1.6 has as corollaries both Bourgain's discretized projection and sum-product theorems. We start with the former. Let $\Pi_s(x, y) = x - sy$.

Theorem 5.1 (Bourgain's discretized projection theorem [2010]; see also [Shmerkin 2023]). *Given $s, u \in (0, 1)$ there are $\epsilon, \eta > 0$ such that the following hold: Let $P \subset B^2(0, 1)$ satisfy*

$$|P \cap B(x, \delta|P|_\delta^{1/2})|_\delta \le \delta^u |P|_\delta, \quad x \in B^2(0, 1).$$

Let $S \subset [1, 2]$ be a $(\delta, s, \delta^{-\epsilon})$-set. Then there is $s \in S$ such that

$$|\Pi_s(P')|_\delta \ge \delta^{-\eta} \cdot |P|_\delta^{1/2} \quad \textit{for all } P' \subset P,\ |P'|_\delta \ge \delta^\epsilon |P|_\delta.$$

Proof. The argument is standard. Suppose the claim does not hold. Hence, for each $s \in S$ there is a set $P_s \subset P$ with $|P_s|_\delta \ge \delta^\epsilon |P|_\delta$ such that

$$|\Pi_s(P_s)|_\delta \le \delta^{-\eta} \cdot |P|^{1/2}.$$

The set $X = \{(p, s) : p \in P_s\}$ has size $\ge \delta^\epsilon |P||S|$; hence there is a set $P' \subset P$ with $|P'|_\delta \gtrsim \delta^\epsilon |P|_\delta$ such that $|S_p| \gtrsim \delta^\epsilon |S|$, where $S_p = \{s : p \in P_s\}$.

Given a tube $T = \boldsymbol{D}([a, a+\delta] \times [b, b+\delta])$ (recall Definition 2.2), we let $\sigma(T) = [a, a+\delta]$ be the corresponding slope interval. For each $p \in P'$ let $\mathcal{T}_p$ be the set of dyadic tubes through p such that $\sigma(T) \cap S_p \ne \varnothing$.

If $\epsilon < \frac{1}{2}u$, then P' still satisfies the single-scale nonconcentration assumption (with $\frac{1}{2}u$ in place of u). Since S_p is a $(\delta, s, O(\delta^{-2\epsilon}))$-set, so is $\mathcal{T}_p$. Also, $|\mathcal{T}_p| \gtrsim \delta^\epsilon |S|$. Hence if $\epsilon > 0$ is small enough in terms of s, u, we can apply Theorem 1.6 to obtain (for $\mathcal{T} = \bigcup_{p \in P'} \mathcal{T}_p$)

$$|\mathcal{T}| \gtrsim \delta^{-\eta'} \cdot \delta^\epsilon |S| \cdot |P'|_\delta^{1/2} \ge \delta^{2\epsilon - \eta'} \cdot |S| \cdot |P|_\delta^{1/2},$$

where $\eta' > 0$ depends on s, u only. Taking $\eta' > 4\epsilon$, this implies that there is $s \in S$ such that there are $\gtrsim \delta^{-\eta'/2} \cdot |P|_\delta^{1/2}$ tubes $T \in \mathcal{T}$ with $s \in \sigma(T)$. All of these tubes intersect P_s by construction. We conclude that

$$|\Pi_s P_s|_\delta \gtrsim \delta^{-\eta'/2} |P|_\delta^{1/2}$$

which is a contradiction if we take $\eta = \frac{1}{3}\eta'$. □

We turn to the discretized sum-product problem. We have the following corollary of Theorem 1.6:

Corollary 5.2. *Given $s, u \in (0, 1)$, there are $\epsilon, \eta > 0$ such that the following holds for all small enough δ: Let $A, B_1, B_2 \subset [1, 2]$ satisfy the following: A is a $(\delta, s, \delta^{-\epsilon})$-set, and B_1, B_2 satisfy the single-scale nonconcentration bound*

$$|B_i \cap [a, a + \delta|B_i|_\delta]|_\delta \le \delta^u |B_i|, \quad a \in [1, 2].$$

Then

$$|A + B_1|_\delta |A \cdot B_2|_\delta \ge \delta^{-\eta} |A|_\delta |B_1|_\delta^{1/2} |B_2|_\delta^{1/2}. \tag{5-1}$$

Taking $A = B_1 = B_2$, one immediately recovers Bourgain's discretized sum-product theorem, even under the weak nonconcentration assumption of [Bourgain and Gamburd 2008, Proposition 3.2].

To prove the corollary, consider (as in [Elekes 1997]) the set $P = (A + B_1) \times (A \cdot B_2)$. For each $(b_1, b_2) \in B_1 \times B_2$, the set P intersects the line $\ell_{b_1,b_2} = \{x + b_1, b_2 x : x \in \mathbb{R}\}$ in an affine copy of A. Using that $(b_1, b_2) \to \ell_{b_1,b_2}$ is bilipschitz, it is routine to verify that this configuration satisfies the assumptions of Corollary 1.7, with $2u$ in place of u (considering for each (b_1, b_2) the δ-dyadic tube that contains $\ell_{b_1,b_2} \cap [0, 2]^2$). See, e.g., [Dąbrowski et al. 2022, §6.3] for details of adapting Elekes' argument to the discretized setting. The conclusion of Corollary 1.7 is then precisely (5-1).

References

[Bourgain 2003] J. Bourgain, "On the Erdős–Volkmann and Katz–Tao ring conjectures", *Geom. Funct. Anal.* **13**:2 (2003), 334–365. MR Zbl

[Bourgain 2010] J. Bourgain, "The discretized sum-product and projection theorems", *J. Anal. Math.* **112** (2010), 193–236. MR Zbl

[Bourgain and Gamburd 2008] J. Bourgain and A. Gamburd, "On the spectral gap for finitely-generated subgroups of SU(2)", *Invent. Math.* **171**:1 (2008), 83–121. MR Zbl

[Di Benedetto and Zahl 2021] D. Di Benedetto and J. Zahl, "New estimates on the size of $(\alpha, 2\alpha)$-Furstenberg sets", preprint, 2021. arXiv 2112.08249

[Dąbrowski et al. 2022] D. Dąbrowski, T. Orponen, and M. Villa, "Integrability of orthogonal projections, and applications to Furstenberg sets", *Adv. Math.* **407** (2022), art. id. 108567. MR Zbl

[Elekes 1997] G. Elekes, "On the number of sums and products", *Acta Arith.* **81**:4 (1997), 365–367. MR Zbl

[Fässler and Orponen 2014] K. Fässler and T. Orponen, "On restricted families of projections in $\mathbb{R}^3$", *Proc. Lond. Math. Soc.* (3) **109**:2 (2014), 353–381. MR Zbl

[Fu and Ren 2024] Y. Fu and K. Ren, "Incidence estimates for α-dimensional tubes and β-dimensional balls in $\mathbb{R}^2$", *J. Fractal Geom.* **11**:1 (2024), 1–30. MR Zbl

[Héra et al. 2022] K. Héra, P. Shmerkin, and A. Yavicoli, "An improved bound for the dimension of $(\alpha, 2\alpha)$-Furstenberg sets", *Rev. Mat. Iberoam.* **38**:1 (2022), 295–322. MR Zbl

[Katz and Tao 2001] N. H. Katz and T. Tao, "Some connections between Falconer's distance set conjecture and sets of Furstenburg type", *New York J. Math.* **7** (2001), 149–187. MR Zbl

[Lutz and Stull 2020] N. Lutz and D. M. Stull, "Bounding the dimension of points on a line", *Inform. Comput.* **275** (2020), art. id. 104601. MR Zbl

[Molter and Rela 2012] U. Molter and E. Rela, "Furstenberg sets for a fractal set of directions", *Proc. Amer. Math. Soc.* **140**:8 (2012), 2753–2765. MR Zbl

[Oberlin 2014] D. M. Oberlin, "Some toy Furstenberg sets and projections of the four-corner Cantor set", *Proc. Amer. Math. Soc.* **142**:4 (2014), 1209–1215. MR Zbl

[Orponen and Shmerkin 2023] T. Orponen and P. Shmerkin, "On the Hausdorff dimension of Furstenberg sets and orthogonal projections in the plane", *Duke Math. J.* **172**:18 (2023), 3559–3632. MR Zbl

[Orponen et al. 2024] T. Orponen, P. Shmerkin, and H. Wang, "Kaufman and Falconer estimates for radial projections and a continuum version of Beck's theorem", *Geom. Funct. Anal.* **34**:1 (2024), 164–201. MR Zbl

[Shmerkin 2022] P. Shmerkin, "On the packing dimension of Furstenberg sets", *J. Anal. Math.* **146**:1 (2022), 351–364. MR Zbl

[Shmerkin 2023] P. Shmerkin, "A non-linear version of Bourgain's projection theorem", *J. Eur. Math. Soc.* **25**:10 (2023), 4155–4204. MR Zbl

[Wolff 1999] T. Wolff, "Recent work connected with the Kakeya problem", pp. 129–162 in *Prospects in mathematics* (Princeton, NJ, 1996), edited by H. Rossi, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

Received 11 Feb 2023. Revised 23 Sep 2023. Accepted 22 Oct 2023.

PABLO SHMERKIN: pshmerkin@math.ubc.ca
Department of Mathematics, University of British Columbia, Vancouver, BC, Canada

HONG WANG: hw3639@nyu.edu
Courant Institute of Mathematical Sciences, New York University, New York, NY, United States

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 18 No. 1 2025