Analysis & PDE

Volume 13 No. 3 2020

msp

Analysis & PDE

msp.org/apde

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

The subscription price for 2020 is US $340/year for the electronic version, and $550/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 13, No. 3, 2020
dx.doi.org/10.2140/apde.2020.13.627

ON THE GAP BETWEEN THE GAMMA-LIMIT AND THE POINTWISE LIMIT FOR A NONLOCAL APPROXIMATION OF THE TOTAL VARIATION

CLARA ANTONUCCI, MASSIMO GOBBINO AND NICOLA PICENNI

We consider the approximation of the total variation of a function by the family of nonlocal and nonconvex functionals introduced by H. Brezis and H.-M. Nguyen in a recent paper. The approximating functionals are defined through double integrals in which every pair of points contributes according to some interaction law.

We answer two open questions concerning the dependence of the Gamma-limit on the interaction law. In the first result, we show that the Gamma-limit depends on the full shape of the interaction law and not only on the values in a neighborhood of the origin. In the second result, we show that there do exist interaction laws for which the Gamma-limit coincides with the pointwise limit on smooth functions.

The key argument is that for some special classes of interaction laws the computation of the Gamma-limit can be reduced to studying the asymptotic behavior of suitable multivariable minimum problems.

1. Introduction

In a recent paper H. Brezis and H.-M. Nguyen [2018] introduced the family of nonlocal functionals

$$\Lambda_\delta(\varphi, u, \Omega) := \iint_{\Omega^2} \varphi\left(\frac{|u(y)-u(x)|}{\delta}\right) \frac{\delta}{|y-x|^{d+1}}\, dx\, dy, \tag{1-1}$$

where d is a positive integer, $\Omega \subseteq \mathbb{R}^d$ is an open set, $\delta > 0$ is a real parameter, $u : \Omega \to \mathbb{R}$ is a measurable function, and $\varphi : [0, +\infty) \to [0, +\infty)$ is a measurable function satisfying suitable properties; see also the note [Brezis and Nguyen 2017] or the conference [Brezis 2016] for a nice presentation of the topic. The function φ, whose presence is motivated by problems in image processing, see [Brezis 2015; Brezis and Nguyen 2018], describes the extent to which a pair $(x, y) \in \Omega^2$ contributes to the double integral (1-1). For this reason, in the sequel we call φ the "interaction law".

Following [Brezis and Nguyen 2018], we restrict ourselves to a special set of functions.

Definition 1.1 (admissible interaction laws). Let $\mathcal{A}$ denote the set of all functions $\varphi : [0, +\infty) \to [0, +\infty)$ not identically equal to 0 such that

(i) φ is nondecreasing and lower semicontinuous on $[0, +\infty)$ and is continuous except at a finite number of points in $(0, +\infty)$,

(ii) there exist a nonnegative real number a and a positive real number b such that

$$\varphi(t) \le \min\{at^2, b\} \quad \text{for all } t \in [0, +\infty).$$

MSC2010: 26B30, 46E35.
Keywords: Gamma-convergence, total variation, bounded-variation functions, nonlocal functional, nonconvex functional.

These conditions guarantee that when $\varphi \in \mathcal{A}$, the right-hand side of (1-1) is finite at least for every u of class C^1 with compact support, and the resulting functional Λ_δ is lower semicontinuous with respect to the convergence in $L^1(\Omega)$.

The basic example is the case where $\varphi(t)$ coincides with

$$\varphi_1(t) := \begin{cases} 0 & \text{if } t \in [0,1], \\ 1 & \text{if } t > 1. \end{cases} \tag{1-2}$$

In [Brezis and Nguyen 2018] also a normalization condition is included in the definition of $\mathcal{A}$. In this paper we do not impose any normalization condition, but instead we set

$$N(\varphi) := \int_0^{+\infty} \frac{\varphi(t)}{t^2}\, dt \quad \text{for all } \varphi \in \mathcal{A}, \tag{1-3}$$

and we exploit this constant as a scale factor when computing limits and Gamma-limits.

Previous literature. The asymptotic behavior of the family Λ_δ was investigated in a series of papers, starting with the model case $\varphi = \varphi_1$; see [Bourgain and Nguyen 2006; Nguyen 2006; 2007; 2008; 2011]. The idea is that $\Lambda_\delta(\varphi, u, \Omega)$ is asymptotically proportional to the functional

$$\Lambda_0(u, \Omega) := \begin{cases} \text{total variation of } u \text{ in } \Omega & \text{if } u \in \mathrm{BV}(\Omega), \\ +\infty & \text{if } u \in L^1(\Omega) \setminus \mathrm{BV}(\Omega). \end{cases}$$

In order to state the precise results, let $\mathbb{S}^{d-1} := \{\sigma \in \mathbb{R}^d : |\sigma| = 1\}$ denote the unit sphere in $\mathbb{R}^d$, and let us consider the geometric constant

$$G_d := \int_{\mathbb{S}^{d-1}} |\langle v, \sigma \rangle|\, d\sigma, \tag{1-4}$$

where v is any element of $\mathbb{S}^{d-1}$ (of course the value of G_d does not depend on the choice of v), and the integration is intended with respect to the $(d-1)$-dimensional Hausdorff measure.

The main convergence results obtained in [Brezis and Nguyen 2018] can be summed up as follows.

• *Pointwise convergence.* For every $\varphi \in \mathcal{A}$ it turns out that

$$\lim_{\delta \to 0^+} \Lambda_\delta(\varphi, u, \mathbb{R}^d) = G_d \cdot N(\varphi) \cdot \Lambda_0(u, \mathbb{R}^d) \quad \text{for all } u \in C^1_c(\mathbb{R}^d),$$

where G_d is the geometric constant defined in (1-4), and $N(\varphi)$ is the scale factor defined in (1-3).

On the other hand, there do exist functions $u \in W^{1,1}(\mathbb{R}^d)$ for which the left-hand side is infinite (while of course the right-hand side is finite). A precise characterization of equality cases is still unknown.

• *Gamma-convergence.* For every $\varphi \in \mathcal{A}$, there exists a constant $K_d(\varphi)$, depending a priori also on the space dimension, such that

$$\Gamma\!-\!\lim_{\delta \to 0^+} \Lambda_\delta(\varphi, u, \mathbb{R}^d) = G_d \cdot N(\varphi) \cdot K_d(\varphi) \cdot \Lambda_0(u, \mathbb{R}^d) \quad \text{for all } u \in L^1(\mathbb{R}^d), \tag{1-5}$$

where the Gamma-limit is intended with respect to the usual metric of $L^1(\mathbb{R}^d)$ (but the result would be the same with respect to the convergence in measure).

Assuming that $K_d(\varphi)$ does not depend on d, as one reasonably expects, we could interpret (1-5) by saying that the Gamma-limit depends on the space dimension through the geometric constant G_d, on the size of φ through the scale factor $N(\varphi)$, and on the shape of φ through the "shape factor" $K_d(\varphi)$.

The behavior under rescaling clarifies the different nature of the scale and the shape factors. If we replace $\varphi(t)$ by $\hat{\varphi}(t) := \alpha\varphi(\beta t)$ for some positive constants α and β, a change of variables shows that $N(\hat{\varphi}) = \alpha\beta N(\varphi)$, and the same scaling affects the left-hand side of (1-5). It follows that $K_d(\varphi) = K_d(\hat{\varphi})$, namely the shape factor is invariant by both horizontal and vertical rescaling.

Very little was known about $K_d(\varphi)$. In [Nguyen 2007; 2011] it was proved that $K_d(\varphi_1) \le \log 2$, where φ_1 is the model interaction law defined in (1-2). In [Brezis and Nguyen 2018] it was proved that

$$K_d(\varphi_1) \le K_d(\varphi) \le 1 \quad \text{for all } \varphi \in \mathcal{A}.$$

In order to shed some light on $K_d(\varphi)$, whose appearance in the Gamma-limit was defined in [Brezis and Nguyen 2018] as "mysterious and somewhat counterintuitive", some open questions were explicitly stated. The first one addresses the dependence of $K_d(\varphi)$ on the full shape of φ.

Question 1 (see [Brezis and Nguyen 2018, Open Problem 4]). Assume that two functions φ and ψ in $\mathcal{A}$ satisfy $N(\varphi) = N(\psi)$, and

$$\varphi \ge \psi \text{ near } 0 \quad (\text{resp. } \varphi = \psi \text{ near } 0).$$

Is it true that

$$K_d(\varphi) \ge K_d(\psi) \quad (\text{resp. } K_d(\varphi) = K_d(\psi))?$$

A positive answer to this question would imply that, once that the scale factor has been fixed, only the behavior of the interaction law in a neighborhood of the origin is relevant to the Gamma-limit. This intuition is supported by the observation that the kernel in (1-1) is divergent on the diagonal $y = x$, and therefore short-range interactions should be more relevant in the computation of $K_d(\varphi)$.

The second question addresses the necessity and the width of the gap between the pointwise limit and the Gamma-limit.

Question 2 (see [Brezis and Nguyen 2018, Open Problem 2]). Is it true that $K_d(\varphi) < 1$ for every $\varphi \in \mathcal{A}$ (and every space dimension d)? Or even better: is it true that

$$\sup\{K_d(\varphi) : \varphi \in \mathcal{A}\} < 1? \tag{1-6}$$

A positive answer to this question, especially in the stronger form (1-6), would imply that the counter-intuitive gap is structural.

Our results. In this paper we show that the answer to both questions is negative. To this end, for every positive integer k we consider the interaction law $\varphi_k : [0, +\infty) \to [0, +\infty)$ defined as

$$\varphi_k(t) := \varphi_1\left(\frac{t}{k}\right) = \begin{cases} 0 & \text{if } t \in [0, k], \\ 1 & \text{if } t > k, \end{cases} \tag{1-7}$$

and then we introduce the following special classes of interaction laws.

Definition 1.2 (special interaction laws). Let $\mathcal{A}$ denote the set of all admissible interaction laws defined in Definition 1.1. We consider the following special subclasses.

- Let $\mathcal{A}_0$ denote the set of interaction laws that vanish in [0, 1], namely

$$\mathcal{A}_0 := \{\varphi \in \mathcal{A} : \varphi(t) = 0 \text{ for all } t \in [0, 1]\}. \tag{1-8}$$

- Let $\mathcal{PCA}$ denote the set of interaction laws that can be written in the form

$$\varphi(t) = \sum_{k=1}^{m} \lambda_k \varphi_k(t) \quad \text{for all } t \geq 0 \tag{1-9}$$

for some positive integer m and some nonnegative real numbers $\lambda_1, \ldots, \lambda_m$ (not all equal to 0).

- Let $\mathcal{PCA}_2$ denote the set of interaction laws of the form (1-9) whose coefficients are equal in packages of powers of 2, namely

$$\lambda_2 = \lambda_3, \quad \lambda_4 = \cdots = \lambda_7, \quad \lambda_8 = \cdots = \lambda_{15},$$

and so on. More precisely, every $\varphi \in \mathcal{PCA}_2$ can be written in the form

$$\varphi(t) := \sum_{j=1}^{m} \left(a_j \sum_{k=2^{j-1}}^{2^j - 1} \varphi_k(t) \right) \quad \text{for all } t \geq 0 \tag{1-10}$$

for some positive integer m, and some nonnegative real numbers $a_1, \ldots, a_m$ (not all equal to 0).

We observe that

$$\mathcal{A} \supseteq \mathcal{A}_0 \supseteq \mathcal{PCA} \supseteq \mathcal{PCA}_2,$$

and all inclusions are strict.

Our first result is the following.

Theorem 1.3 (piecewise constant interaction laws). *Let $K_d(\varphi)$ be the shape factor of an interaction law as defined by* (1-5).

Then in every space dimension d it turns out that

$$\sup\{K_d(\varphi) : \varphi \in \mathcal{PCA}_2\} = 1.$$

A closer look at the proof reveals that the supremum is realized for example by the interaction laws of the form (1-10) with $a_1 = \ldots = a_m = 1$, in the limit as $m \to +\infty$. We point out that these achieving interaction laws are also of the form (1-9) with $\lambda_k = 1$ and m a power of 2.

Theorem 1.3 above provides a negative answer to Question 1, as well as a negative answer to Question 2 in the stronger form (1-6). In particular, this means that the full shape of the interaction law comes into play in the computation of the Gamma-limit, which therefore takes into account both short-range and long-range interactions. At the beginning of Section 6 we present also an example with strict inequalities, namely with $\varphi > \psi$ near the origin, but $K_d(\varphi) < K_d(\psi)$.

Our second result gives a stronger negative answer to Question 2, even when restricted to the smaller class $\mathcal{A}_0$.

Theorem 1.4 (piecewise affine dyadic interaction laws). *Let $K_d(\varphi)$ be the shape factor of an interaction law as defined by* (1-5).

Then the following statements hold true:

(1) *In every space dimension d it turns out that*

$$\max\{K_d(\varphi) : \varphi \in \mathcal{A}_0\} = 1.$$

(2) *More precisely, let $f : \mathbb{Z} \to [0, +\infty)$ be any nondecreasing and bounded function* (*not identically equal to* 0) *such that*

$$\limsup_{n \to +\infty} f(-n) \cdot 4^n < +\infty. \tag{1-11}$$

Let us consider the function $\zeta : [0, +\infty) \to [0, +\infty)$ such that

- $\zeta(0) = 0$,
- $\zeta(2^z) = f(z)$ *for every $z \in \mathbb{Z}$,*
- ζ *is affine in the interval $[2^z, 2^{z+1}]$ for every $z \in \mathbb{Z}$.*

Then it turns out that $\zeta \in \mathcal{A}$, and $K_d(\zeta) = 1$ in every space dimension d.

The second statement of Theorem 1.4 above shows in particular that there are large classes of interaction laws for which the Gamma-limit of (1-1) coincides with the pointwise limit for smooth functions.

Overview of the technique. The proof of these results follows the same strategy that in [Antonucci et al. 2020] led us to show that actually $K_d(\varphi_1) = \log 2$ (see also [Antonucci et al. 2018] for an informal summary of our approach). Since we only need estimates from below for the shape factor, we can limit ourselves to estimating from below the Gamma-liminf. The main steps are the following.

From local to global bounds: In Theorem B we reduce the problem in any dimension to intervals of the real line, namely to showing that for δ small enough we can estimate from below $\Lambda_\delta(\varphi, u, (a, b))$ in terms of the oscillation of u in (a, b).

Reduction to multivariable minimum problems: In Proposition 4.3 we show that, when $\varphi \in \mathcal{PCA}$, we can assume that u is a nondecreasing step function with finite image contained in $\delta\mathbb{Z}$. Up to vertical translations, any such function depends only on the lengths $\ell_1, \dots, \ell_n$ of the steps, where $n \sim \delta^{-1}$. In this way we reduce ourselves to studying the minimum of a δ-dependent multivariable function.

Telescopic effect: The minimum problems found in Proposition 4.3 can be very complicated. Nevertheless, in the special case where $\varphi \in \mathcal{PCA}_2$, a telescopic effect implies that the configurations with $\ell_1 = \cdots = \ell_n$ are asymptotically optimal (see Proposition 3.3). At this point, we have explicit estimates from below for the Gamma-liminf, which in turn yield explicit estimates from below for shape factors, thus leading to the proof of Theorem 1.3.

Approximation from below of special interaction laws: The more general interaction laws of Theorem 1.4 can be approximated from below by sequences of interaction laws that, up to horizontal rescaling, are in $\mathcal{PCA}_2$. Again this provides estimates from below for the Gamma-liminf, and hence for shape factors, that lead to the proof of Theorem 1.4.

Structure of the paper. This paper is organized as follows. In Section 2 we recall the technical results from [Antonucci et al. 2020] that are needed in the sequel. In Section 3 we investigate the asymptotic behavior of suitable sequences of minimum problems for functions of a finite number of variables. In Section 4 we prove Proposition 4.3, which establishes the connection between the Gamma-limit of (1-1) and the minimum problems of the previous section. In Section 5 we prove Theorems 1.3 and 1.4. In Section 6 we speculate about some possible generalizations and future directions.

2. Preliminary results

In this section we collect, for the convenience of the reader, the results from [Antonucci et al. 2020] that are crucial to this paper. To begin with, we recall the definitions of truncation, vertical δ-segmentation and nondecreasing rearrangement.

Definition 2.1 (truncation). Let $\mathbb{X}$ be any set, let $w:\mathbb{X}\to\mathbb{R}$ be any function, and let $A<B$ be two real numbers. The truncation of w between A and B is the function $T_{A,B}w:\mathbb{X}\to\mathbb{R}$ defined by

$$T_{A,B}w(x) := \begin{cases} A & \text{if } w(x) < A, \\ w(x) & \text{if } A \le w(x) \le B, \\ B & \text{if } w(x) > B. \end{cases}$$

Definition 2.2 (vertical δ-segmentation). Let $\mathbb{X}$ be any set, let $w:\mathbb{X}\to\mathbb{R}$ be any function, and let δ be a positive real number. The vertical δ-segmentation of w is the function $S_\delta w:\mathbb{X}\to\mathbb{R}$ defined by

$$S_\delta w(x) := \delta \left\lfloor \frac{w(x)}{\delta} \right\rfloor \quad \text{for all } x \in \mathbb{X}.$$

The function $S_\delta w$ takes its values in $\delta\mathbb{Z}$, and it is uniquely characterized by the fact that $S_\delta w(x) = k\delta$ for some $k\in\mathbb{Z}$ if and only if $k\delta \le w(x) < (k+1)\delta$.

Definition 2.3 (nondecreasing rearrangement). Let $(a,b)\subseteq\mathbb{R}$ be an interval, and let $w:(a,b)\to\mathbb{R}$ be a function whose image is a finite set. The nondecreasing rearrangement of w is the function $Mw:(a,b)\to\mathbb{R}$ defined by

$$Mw(x) := \min\{y\in\mathbb{R} : \operatorname{meas}\{z\in(a,b) : w(z)\le y\} \ge x-a\} \quad \text{for all } x\in\mathbb{R}. \tag{2-1}$$

As expected, the function Mw is nondecreasing and satisfies

$$\operatorname{meas}\{x\in(a,b) : Mw(x)=y\} = \operatorname{meas}\{x\in(a,b) : w(x)=y\} \quad \text{for all } y\in\mathbb{R}.$$

Semidiscrete aggregation problem. Let us recall the main result of [Antonucci et al. 2020, Section 2], where a class of discrete and semidiscrete aggregation problems was considered. Following the terminology introduced therein, the basic ingredients are an interval $(a,b)\subseteq\mathbb{R}$, an integer number $k\ge 1$, and a nonincreasing function $c:(0,b-a)\to\mathbb{R}$, possibly unbounded in a neighborhood of the origin, called the hostility function. A semidiscrete arrangement is any measurable function $u:(a,b)\to\mathbb{Z}$ with finite image.

For any such function u, we define the total k-hostility as

$$\mathcal{F}_k(c, u) = \iint_{(a,b)^2} \varphi_k(|u(y) - u(x)|) \cdot c(|y - x|)\, dx\, dy, \tag{2-2}$$

where $\varphi_k(t)$ is the interaction law defined by (1-7).

The key result proved in [Antonucci et al. 2020, Theorem 2.4] (and equivalent to some rearrangement inequalities found independently in [Taylor 1973; Garsia and Rodemich 1974] in a different context) is that the nondecreasing rearrangement does not increase the total hostility.

Theorem A (total hostility minimization). *Let $(a, b) \subseteq \mathbb{R}$ be an interval, let $k \geq 1$ be an integer, and let $c : (0, b - a) \to \mathbb{R}$ be a nonincreasing function. Let $u : (a, b) \to \mathbb{Z}$ be a measurable function with finite image, let $Mu : (a, b) \to \mathbb{Z}$ be its nondecreasing rearrangement defined by* (2-1), *and let $\mathcal{F}_k(c, u)$ be the functional defined by* (2-2).

Then it turns out that

$$\mathcal{F}_k(c, u) \geq \mathcal{F}_k(c, Mu).$$

Localization technique. One of the main points in [Antonucci et al. 2020] was obtaining a localized version of the Gamma-liminf inequality, namely an asymptotic estimate from below for $\Lambda_\delta(\varphi, u_\delta, (a, b))$ in terms of the oscillation of u_δ in (a, b). After such an estimate has been established, a quite classical path, see for example [Gobbino 1998; Gobbino and Mora 2001], independent of the presence of the interaction law φ, leads to an estimate from below for the Gamma-liminf of Λ_δ in any space dimension, and hence to an estimate from below for the shape factor of φ. The precise statement is the following.

Theorem B (from local to global bounds). *Let $\varphi \in \mathcal{A}$ be an interaction law, with scale factor $N(\varphi)$ defined by* (1-3) *and shape factor $K_d(\varphi)$ defined by* (1-5).

Let us assume that there exists a constant K_0 such that, for every interval $(a, b) \subseteq \mathbb{R}$ and every family $\{u_\delta\}_{\delta>0} \subseteq L^1((a, b))$, it happens that

$$\liminf_{\delta \to 0^+} \Lambda_\delta(\varphi, u_\delta, (a, b)) \geq K_0 \cdot \liminf_{\delta \to 0^+} \operatorname{osc}(u_\delta, (a, b)), \tag{2-3}$$

where $\operatorname{osc}(u_\delta, (a, b))$ denotes the essential oscillation of u_δ in (a, b).

Then for every positive integer d it turns out that

$$\Gamma\text{–}\liminf_{\delta \to 0^+} \Lambda_\delta(\varphi, u, \mathbb{R}^d) \geq G_d \cdot \tfrac{1}{2} K_0 \cdot \Lambda_0(u, \mathbb{R}^d) \quad \textit{for all } u \in L^1(\mathbb{R}^d). \tag{2-4}$$

The proof of Theorem B has two distinct steps.

The first one, for which we refer to [Antonucci et al. 2020, Section 3.2], exploits a piecewise affine approximation in order to deduce (2-4) in dimension 1 from the local estimate (2-3). We remark that $G_1 = 2$, and therefore in dimension 1 the geometric constant cancels the denominator, so that (2-3) is exactly the localized version of (2-4).

The second step, for which we refer to [Antonucci et al. 2020, Section 4], relies on an integral-geometric representation of both the total variation and the double integral (1-1). This representation leads from (2-4) in dimension 1 to the analogous inequality in any space dimension.

3. A family of multivariable minimum problems

In this section we introduce some notation, and we prove asymptotic estimates for a family of inequalities involving multivariable functions. Roughly speaking, these multivariable functions represent the functional Λ_δ computed on a piecewise constant function with a finite number of steps of lengths $\ell_1, \dots, \ell_n$. The main result of this section, namely estimate (3-10), is going to play a crucial role when combined with the result of Proposition 4.3, where we estimate from below the asymptotic cost of oscillations.

Let n be a positive integer, and let $(\ell_1, \dots, \ell_n)$ be an n-tuple of nonnegative real numbers. For every positive integer $k \le n$ we consider all possible sums of k consecutive terms

$$S_{i,k}(\ell_1, \dots, \ell_n) := \sum_{h=0}^{k-1} \ell_{i+h} \quad \text{for all } i \in \{1, \dots, n-k+1\}, \tag{3-1}$$

and we define the set

$$D_{n,k} := \{(\ell_1, \dots, \ell_n) \in [0, +\infty)^n : S_{i,k}(\ell_1, \dots, \ell_n) > 0 \text{ for all } i \in \{1, \dots, n-k+1\}\} \tag{3-2}$$

of all n-tuples of nonnegative real numbers without k consecutive components equal to 0. When $n \ge k+1$, for every $(\ell_1, \dots, \ell_n) \in D_{n,k}$ we can set

$$L_k(\ell_1, \dots, \ell_n) := \sum_{i=1}^{n-k} \log \frac{[S_{i,k+1}(\ell_1, \dots, \ell_n)]^2}{S_{i,k}(\ell_1, \dots, \ell_n) \cdot S_{i+1,k}(\ell_1, \dots, \ell_n)}. \tag{3-3}$$

Given any interaction law $\varphi \in \mathcal{PCA}$, we call $\mu(\varphi)$ the smallest integer k such that $\lambda_k \neq 0$ in the representation (1-9), and for every $n \ge m+1$ we consider the homogeneous function $P_{n,\varphi} : D_{n,\mu(\varphi)} \to \mathbb{R}$ defined by

$$P_{n,\varphi}(\ell_1, \dots, \ell_n) := \sum_{k=\mu(\varphi)}^{m} \lambda_k L_k(\ell_1, \dots, \ell_n) \quad \text{for all } (\ell_1, \dots, \ell_n) \in D_{n,\mu(\varphi)}, \tag{3-4}$$

and its infimum

$$\begin{aligned} I_n(\varphi) &:= \inf\{P_{n,\varphi}(\ell_1, \dots, \ell_n) : (\ell_1, \dots, \ell_n) \in D_{n,\mu(\varphi)}\} \\ &= \inf\{P_{n,\varphi}(\ell_1, \dots, \ell_n) : (\ell_1, \dots, \ell_n) \in (0, +\infty)^n\}. \end{aligned} \tag{3-5}$$

Let us consider for example the interaction law φ_1 defined in (1-2). In this case $\mu(\varphi_1) = 1$, the set $D_{n,1}$ is just $(0, +\infty)^n$, and the function in (3-4) has the form

$$P_{n,\varphi_1}(\ell_1, \dots, \ell_n) = L_1(\ell_1, \dots, \ell_n) = \log \frac{(\ell_1+\ell_2)^2}{\ell_1\ell_2} + \log \frac{(\ell_2+\ell_3)^2}{\ell_2\ell_3} + \cdots + \log \frac{(\ell_{n-1}+\ell_n)^2}{\ell_{n-1}\ell_n}.$$

All the fractions inside the logarithms are greater than or equal to 4, and hence $I_n(\varphi_1) = (n-1)\log 4$, with the minimum realized when all the variables are equal. This computation was the final step in the proof of the crucial estimate in [Antonucci et al. 2020]. In Section 4 of the present paper we show that the asymptotic behavior of $I_n(\varphi)$ plays a fundamental role in estimating from below the shape factor of any interaction law $\varphi \in \mathcal{PCA}$.

Unfortunately, things are not that simple for larger values of k. For example, when the interaction law is φ_3, the function $L_3(\ell_1, \ldots, \ell_n)$ is given by

$$\log \frac{(\ell_1+\ell_2+\ell_3+\ell_4)^2}{(\ell_1+\ell_2+\ell_3)(\ell_2+\ell_3+\ell_4)} + \log \frac{(\ell_2+\ell_3+\ell_4+\ell_5)^2}{(\ell_2+\ell_3+\ell_4)(\ell_3+\ell_4+\ell_5)} + \cdots,$$

and now the minimum is not realized when all the variables are equal (for example the periodic pattern $1, 0, 0, 1, 0, 0, \ldots$ is better).

Of course any interaction law of the form φ_k can be dealt with as a rescaling of φ_1, but nevertheless the appearance of different patterns in the minimization process seems to suggest that things get worse and worse when we take linear combinations of the form (3-4). Fortunately this is not always the case. Indeed, when we expand linear combinations of this form, the numerators of the terms of L_k can partially cancel with the denominators of the terms of L_{k+1}, leading to the following result.

Lemma 3.1 (telescopic effect). *Let a, b, and n be positive integers such that $a \le b \le n-1$. Let $S_{i,k}$ and L_k be the functions of n variables defined in* (3-1) *and* (3-3).

Then for every $(\ell_1, \ldots, \ell_n) \in D_{n,a}$ it turns out that

$$\sum_{j=a}^{b} L_j(\ell_1, \ldots, \ell_n) \ge \sum_{i=1}^{n-b} \log \frac{[S_{i,b+1}(\ell_1, \ldots, \ell_n)]^2}{S_{i,a}(\ell_1, \ldots, \ell_n) \cdot S_{i+(b-a)+1,a}(\ell_1, \ldots, \ell_n)}. \tag{3-6}$$

Proof. To begin with, we observe that (3-6) is an equality when $b = a$. Therefore, in the sequel we assume that $b \ge a+1$. For the sake of shortness, throughout this proof we omit the explicit dependence on the variables $\ell_1, \ldots, \ell_n$. The left-hand side of (3-6) can be written in the form

$$\sum_{j=a}^{b} L_j = 2\Sigma_1 - \Sigma_2 - \Sigma_3, \tag{3-7}$$

where

$$\Sigma_1 := \sum_{j=a}^{b} \sum_{i=1}^{n-j} \log S_{i,j+1}, \quad \Sigma_2 := \sum_{j=a}^{b} \sum_{i=1}^{n-j} \log S_{i,j}, \quad \Sigma_3 := \sum_{j=a}^{b} \sum_{i=1}^{n-j} \log S_{i+1,j}.$$

With some algebra (shift of indices and separation of the terms corresponding to the first or last value of some index) we can rewrite the three sums as

$$\Sigma_1 = \sum_{i=1}^{n-b} \log S_{i,b+1} + \sum_{j=a+1}^{b} \log S_{1,j} + \sum_{j=a+1}^{b} \log S_{n-j+1,j} + \sum_{j=a+1}^{b} \sum_{i=2}^{n-j} \log S_{i,j},$$

$$\Sigma_2 = \sum_{i=1}^{n-a} \log S_{i,a} + \sum_{j=a+1}^{b} \log S_{1,j} + \sum_{j=a+1}^{b} \sum_{i=2}^{n-j} \log S_{i,j},$$

$$\Sigma_3 = \sum_{i=2}^{n-a+1} \log S_{i,a} + \sum_{j=a+1}^{b} \log S_{n-j+1,j} + \sum_{j=a+1}^{b} \sum_{i=2}^{n-j} \log S_{i,j}.$$

When we plug these three equalities into (3-7), all double sums cancel, and also the second sums in Σ_2 and Σ_3 cancel with a part of the second and third terms in Σ_1. We end up with

$$2\Sigma_1-\Sigma_2-\Sigma_3=2\sum_{i=1}^{n-b}\log S_{i,b+1}+\sum_{j=a+1}^{b}\log S_{1,j}+\sum_{j=a+1}^{b}\log S_{n-j+1,j}-\sum_{i=1}^{n-a}\log S_{i,a}-\sum_{i=2}^{n-a+1}\log S_{i,a}. \quad (3\text{-}8)$$

Let us reorganize these terms. In the second and third sums we change the indices and we rewrite them as

$$\sum_{j=a+1}^{b}\log S_{1,j}=\sum_{k=1}^{b-a}\log S_{1,k+a},$$

$$\sum_{j=a+1}^{b}\log S_{n-j+1,j}=\sum_{k=n-b+1}^{n-a}\log S_{k,n+1-k}.$$

In the fourth sum we split the terms as

$$\sum_{i=1}^{n-a}\log S_{i,a}=\sum_{i=1}^{n-b}\log S_{i,a}+\sum_{k=n-b+1}^{n-a}\log S_{k,a}.$$

In the fifth sum we split the terms, and then we shift one index in order to rewrite the sum as

$$\sum_{i=2}^{n-a+1}\log S_{i,a}=\sum_{i=2}^{b-a+1}\log S_{i,a}+\sum_{i=b-a+2}^{n-a+1}\log S_{i,a}=\sum_{k=1}^{b-a}\log S_{k+1,a}+\sum_{i=1}^{n-b}\log S_{i+(b-a)+1,a}.$$

Plugging all these equalities into (3-8) we find that

$$2\Sigma_1-\Sigma_2-\Sigma_3=\sum_{i=1}^{n-b}\log\frac{[S_{i,b+1}]^2}{S_{i,a}\cdot S_{i+(b-a)+1,a}}+\sum_{k=1}^{b-a}\log\frac{S_{1,k+a}}{S_{k+1,a}}+\sum_{k=n-b+1}^{n-a}\log\frac{S_{k,n+1-k}}{S_{k,a}}.$$

In the sums of the last line, all terms are nonnegative because in all the fractions the numerators are greater than or equal to the corresponding denominators. Recalling (3-7), it follows that

$$\sum_{j=a}^{b}L_j=2\Sigma_1-\Sigma_2-\Sigma_3\ge\sum_{i=1}^{n-b}\log\frac{[S_{i,b+1}]^2}{S_{i,a}\cdot S_{i+(b-a)+1,a}},$$

which completes the proof of (3-6). □

Corollary 3.2. *Let us consider the situation described in Lemma 3.1 in the special case where* $a=2^{m-1}$ *and* $b=2^m-1$ *for some positive integer* m.

Then for every $n\ge 2^m$ *it turns out that*

$$\sum_{k=2^{m-1}}^{2^m-1}L_k(\ell_1,\dots,\ell_n)\ge(n-2^m+1)\cdot 2\log 2\quad \textit{for all } (\ell_1,\dots,\ell_n)\in D_{n,a}. \quad (3\text{-}9)$$

Proof. In this special case it turns out that

$$S_{i,b+1}=S_{i,a}+S_{i+(b-a)+1,a}\quad \text{for all } i\le n-2^m+1.$$

Therefore, from the inequality between arithmetic mean and geometric mean, we deduce that all the fractions in the right-hand side of (3-6) are greater than or equal to 4, and this is enough to establish (3-9). □

From Corollary 3.2 we deduce a lower bound for the asymptotic behavior of $I_n(\varphi)$ for interaction laws $\varphi \in \mathcal{PCA}_2$.

Proposition 3.3 (interaction laws with package structure). *Let m be a positive integer, let $a_1, \dots, a_m$ be nonnegative real numbers (not all equal to 0), and let $\varphi \in \mathcal{PCA}_2$ be defined as in* (1-10).

For every integer $n \geq 2^m$, let $P_{n,\varphi}$ be the homogeneous function defined by (3-4), *and let $I_n(\varphi)$ be its infimum as in* (3-5).

Then it turns out that

$$\liminf_{n\to+\infty} \frac{I_n(\varphi)}{n} \geq 2\log 2 \cdot \sum_{j=1}^{m} a_j. \tag{3-10}$$

Proof. Let $m_0(\varphi)$ denote the smallest integer k such that $a_k > 0$. From Corollary 3.2 we know

$$\sum_{k=2^{j-1}}^{2^j-1} L_k(\ell_1, \dots, \ell_n) \geq (n - 2^j + 1) \cdot 2\log 2$$

for every $j \in \{m_0(\varphi), \dots, m\}$, and therefore

$$P_{n,\varphi}(\ell_1, \dots, \ell_n) \geq \sum_{j=m_0(\varphi)}^{m} a_j \cdot (n - 2^j + 1) \cdot 2\log 2 \geq (n - 2^m) \cdot 2\log 2 \cdot \sum_{j=1}^{m} a_j$$

for every admissible choice of $\ell_1, \dots, \ell_n$.

Dividing by n, and letting $n \to +\infty$, we obtain (3-10). □

4. Asymptotic cost of oscillations

In this section we clarify the connection between the Gamma-limit of the family (1-1) and the multivariable functions of Section 3. In analogy with [Antonucci et al. 2020], the question we address is the following. Let us assume that a function $u_\delta(x)$ oscillates between two values A and B in some interval (a, b). Does this provide an estimate from below for $\Lambda_\delta(\varphi, u_\delta, (a, b))$, at least when δ is small enough? A quantitative answer is provided by Proposition 4.3 and Corollary 4.4, and this answer is connected to the Gamma-limit of the family (1-1) by Theorem B, as we clarify in Proposition 4.5.

To begin with, we show that three simplifying operations can be performed on u_δ without changing its oscillation between A and B or increasing its energy.

Lemma 4.1 (truncation, segmentation, rearrangement). *Let $a < b$ and $A < B$ be real numbers, let $u : (a, b) \to \mathbb{R}$ be a measurable function, and let $\varphi \in \mathcal{PCA}$.*

Then for every $\delta > 0$ it turns out that

$$\Lambda_\delta(\varphi, u, (a, b)) \geq \Lambda_\delta(\varphi, M S_\delta T_{A,B} u, (a, b)), \tag{4-1}$$

where $T_{A,B}$, S_δ, and M are the operators of truncation, vertical δ-segmentation, and nondecreasing rearrangement defined at the beginning of Section 2.

Proof. Since Λ_δ is linear with respect to φ, it is enough to show inequality (4-1) when $\varphi = \varphi_k$ for some positive integer k, in which case

$$\Lambda_\delta(\varphi_k, u, (a,b)) = \iint_{I_k(\delta,u,(a,b))} \frac{\delta}{(y-x)^2}\,dx\,dy,$$

where

$$I_k(\delta, u, (a,b)) := \{(x,y) \in (a,b)^2 : |u(y)-u(x)| > k\delta\}.$$

Let us examine the effects on Λ_δ of the three operations performed on u. The arguments are the same as in the first part of the proof of [Antonucci et al. 2020, Proposition 3.4], where however only the case of φ_1 was considered.

Truncation: For every x and y in (a,b) it turns out that

$$|T_{A,B}u(y) - T_{A,B}u(x)| > k\delta \quad \Longrightarrow \quad |u(y)-u(x)| > k\delta.$$

This implies $I_k(\delta, T_{A,B}u, (a,b)) \subseteq I_k(\delta, u, (a,b))$, and therefore

$$\Lambda_\delta(\varphi_k, u, (a,b)) \geq \Lambda_\delta(\varphi_k, T_{A,B}u, (a,b)). \tag{4-2}$$

Vertical δ-segmentation: For every x and y in (a,b) it turns out that

$$|S_\delta u(y) - S_\delta u(x)| > k\delta \quad \Longrightarrow \quad |S_\delta u(y) - S_\delta u(x)| \geq (k+1)\delta \quad \Longrightarrow \quad |u(y)-u(x)| > k\delta.$$

As before this implies

$$\Lambda_\delta(\varphi_k, T_{A,B}u, (a,b)) \geq \Lambda_\delta(\varphi_k, S_\delta T_{A,B}u, (a,b)). \tag{4-3}$$

Nondecreasing rearrangement: We claim that

$$\Lambda_\delta(\varphi_k, S_\delta T_{A,B}u, (a,b)) \geq \Lambda_\delta(\varphi_k, MS_\delta T_{A,B}u, (a,b)). \tag{4-4}$$

This inequality, together with (4-2) and (4-3), completes the proof of (4-1).

In order to prove (4-4), we rely on the theory of semidiscrete arrangements. To this end, we consider the semidiscrete arrangement $v_\delta : (a,b) \to \mathbb{Z}$ defined by

$$v_\delta(x) := \frac{1}{\delta} S_\delta T_{A,B}u(x) \quad \text{for all } x \in (a,b) \tag{4-5}$$

(we recall that $S_\delta T_{A,B}u$ takes its values in $\delta\mathbb{Z}$, and hence $v_\delta(x)$ is integer-valued) and the hostility function $c : (0, b-a) \to \mathbb{R}$ defined by $c(\sigma) := \delta\sigma^{-2}$. We observe that

$$MS_\delta T_{A,B}u(x) = \delta M v_\delta(x) \quad \text{for all } x \in (a,b),$$

where Mv_δ is the nondecreasing rearrangement of v_δ. From (4-5) and (2-2) it turns out that

$$\Lambda_\delta(\varphi_k, S_\delta T_{A,B}u, (a,b)) = \delta\Lambda_1(\varphi_k, v_\delta, (a,b)) = \mathcal{F}_k(c, v_\delta),$$

and similarly

$$\Lambda_\delta(\varphi_k, MS_\delta T_{A,B}u, (a,b)) = \delta\Lambda_1(\varphi_k, Mv_\delta, (a,b)) = \mathcal{F}_k(c, Mv_\delta),$$

so that now (4-4) follows from Theorem A. □

As a second simplifying step, we show that the double integral over $(a,b)^2$ can be replaced, in the computation of the liminf, by a double integral over an infinite strip, which is easier to handle. To this end, we introduce the family of functionals

$$\hat{\Lambda}_\delta(\varphi, u, (c,d)) := \int_c^d dx \int_{-\infty}^{+\infty} \varphi\left(\frac{|u(y)-u(x)|}{\delta}\right)\frac{\delta}{(y-x)^2}\,dy,$$

and we prove the following result.

Lemma 4.2 (extension to a vertical strip). *Let $a<c<d<b$ be real numbers, and let $\varphi:[0,+\infty)\to[0,+\infty)$ be a bounded measurable function. For every $\delta>0$, let $u_\delta:(a,b)\to\mathbb{R}$ be a measurable function. Let us extend u_δ to the whole real line by setting $u_\delta(x)=0$ for every $x\notin(a,b)$.*

Then it turns out that

$$\liminf_{\delta\to 0^+}\Lambda_\delta(\varphi,u_\delta,(a,b)) \ge \liminf_{\delta\to 0^+}\hat{\Lambda}_\delta(\varphi,u_\delta,(c,d)). \tag{4-6}$$

Proof. Let us set for shortness

$$f_\delta(x,y) := \varphi\left(\frac{|u_\delta(y)-u_\delta(x)|}{\delta}\right)\frac{\delta}{(y-x)^2} \quad \text{for all } (x,y)\in(a,b)\times\mathbb{R}.$$

Since φ is nonnegative and $(c,d)\subseteq(a,b)$, for every $\delta>0$ it turns out that

$$\begin{aligned}\Lambda_\delta(\varphi,u_\delta,(a,b)) &\ge \int_c^d dx\int_a^b f_\delta(x,y)\,dy\\ &= \hat{\Lambda}_\delta(\varphi,u_\delta,(c,d)) - \int_c^d dx\int_{\mathbb{R}\setminus[a,b]} f_\delta(x,y)\,dy.\end{aligned} \tag{4-7}$$

From the boundedness of φ it follows that

$$\int_c^d dx\int_b^{+\infty} f_\delta(x,y)\,dy \le \delta\|\varphi\|_\infty\int_c^d dx\int_b^{+\infty}\frac{1}{(y-x)^2}\,dy.$$

Since $d<b$, the double integral in the right-hand side is convergent, and hence

$$\lim_{\delta\to 0^+}\int_c^d dx\int_b^{+\infty} f_\delta(x,y)\,dy = 0. \tag{4-8}$$

In an analogous way we obtain

$$\lim_{\delta\to 0^+}\int_c^d dx\int_{-\infty}^{a} f_\delta(x,y)\,dy = 0. \tag{4-9}$$

At this point, (4-6) follows from (4-7), (4-8), and (4-9). □

We are now ready to state and prove the main result of this section.

Proposition 4.3 (limit cost of vertical oscillations). *Let $(a,b)\subseteq\mathbb{R}$ be an interval, let $\{u_\delta\}_{\delta>0}\subseteq L^1((a,b))$ be a family of functions, let $\varphi\in\mathcal{PCA}$ be a piecewise constant interaction law, let $P_{n,\varphi}$ be the multivariable function defined by* (3-4), *and let $I_n(\varphi)$ be its infimum as in* (3-5).

Let us assume that there exist two real numbers $A \leq B$ such that

$$\liminf_{\delta\to 0^+} \operatorname{meas}\{x \in (a,b) : u_\delta(x) \leq A+\varepsilon\} > 0 \quad \textit{for all } \varepsilon > 0, \tag{4-10}$$

$$\liminf_{\delta\to 0^+} \operatorname{meas}\{x \in (a,b) : u_\delta(x) \geq B-\varepsilon\} > 0 \quad \textit{for all } \varepsilon > 0. \tag{4-11}$$

Then it turns out that

$$\liminf_{\delta\to 0^+} \Lambda_\delta(\varphi, u_\delta, (a,b)) \geq (B-A) \cdot \liminf_{n\to+\infty} \frac{I_n(\varphi)}{n}. \tag{4-12}$$

Proof. To begin with, we observe that (4-12) is trivial if $A = B$, or if the left-hand side is infinite. Up to restricting ourselves to a sequence $\delta_k \to 0^+$, we can also assume that the liminf is actually a limit. Therefore, in the sequel we assume that the left-hand side of (4-12) is uniformly bounded from above and that $A < B$.

Let us fix $\varepsilon > 0$ such that $4\varepsilon < B - A$. Due to assumptions (4-10) and (4-11), there exist $\eta > 0$ and $\delta_0 > 0$ such that

$$\operatorname{meas}\{x \in (a,b) : u_\delta(x) \leq A+\varepsilon\} \geq \eta \quad \text{for all } \delta \in (0,\delta_0), \tag{4-13}$$

$$\operatorname{meas}\{x \in (a,b) : u_\delta(x) \geq B-\varepsilon\} \geq \eta \quad \text{for all } \delta \in (0,\delta_0). \tag{4-14}$$

Let us consider the modified family $\hat{u}_\delta := MS_\delta T_{A,B} u_\delta$ as in Lemma 4.1. From (4-13) and (4-14) it follows that the nondecreasing function $\hat{u}_\delta$ satisfies

$$\hat{u}_\delta(x) \leq A + 2\varepsilon \quad \text{for all } x \in (a, a+\eta), \text{ for all } \delta \in (0,\delta_1), \tag{4-15}$$

$$\hat{u}_\delta(x) \geq B - 2\varepsilon \quad \text{for all } x \in (b-\eta, b), \text{ for all } \delta \in (0,\delta_1), \tag{4-16}$$

where $\delta_1 := \min\{\varepsilon, \delta_0\}$. Moreover, from Lemmas 4.1 and 4.2 it follows that

$$\liminf_{\delta\to 0^+} \Lambda_\delta(\varphi, u_\delta, (a,b)) \geq \liminf_{\delta\to 0^+} \Lambda_\delta(\varphi, \hat{u}_\delta, (a,b)) \geq \liminf_{\delta\to 0^+} \hat{\Lambda}_\delta(\varphi, \hat{u}_\delta, (a+\eta, b-\eta)),$$

where in the computation of the latter we imagine that $\hat{u}_\delta$ has been extended to the whole real line by setting it equal to 0 (or any other value) outside (a,b).

In order to compute the last liminf, we need a deeper description of the structure of $\hat{u}_\delta$. We know that $\hat{u}_\delta$ is nondecreasing and that its image is contained in $\delta\mathbb{Z}$. Let δm_0 denote the value of $\hat{u}_\delta$ in a right neighborhood of a, let us set $x_0 = a$, and for every positive integer i let us define

$$x_i := \sup\{x \in (a,b) : \hat{u}_\delta(x) < (m_0 + i)\delta\}.$$

The sequence x_i is nondecreasing, and $x_i = b$ for every large enough index i. If $x_{i+1} > x_i$ for some index i, then it turns out that

$$\hat{u}_\delta(x) = (m_0 + i)\delta \quad \text{for all } x \in (x_i, x_{i+1}).$$

If $x_{i+1} = x_i$ for some index i, this means that

$$\operatorname{meas}\{x \in (a,b) : \hat{u}_\delta(x) = (m_0 + i)\delta\} = 0.$$

Let α and β be the two indices (which of course do depend on δ) such that

$$\alpha := \min\{i \in \mathbb{N} : x_i \geq a+\eta\} \quad \text{and} \quad \beta := \max\{i \in \mathbb{N} : x_i \leq b-\eta\}.$$

Let us consider now the interaction law φ, which we assumed of the form (1-9), and let $\mu(\varphi)$ denote the smallest index $k \leq m$ such that $\lambda_k > 0$. To begin with, we show that $x_{i+\mu(\varphi)} > x_i$ for every index i with $\alpha \leq i \leq \beta$. Indeed, if this is not the case, then it turns out that

$$\hat{u}_\delta(y) - \hat{u}_\delta(x) \geq (\mu(\varphi)+1)\delta \quad \text{for all } (x, y) \in (a, x_i) \times (x_i, b),$$

and in particular

$$\Lambda_\delta(\varphi, \hat{u}_\delta, (a, b)) \geq \lambda_{\mu(\varphi)} \Lambda_\delta(\varphi_{\mu(\varphi)}, \hat{u}_\delta, (a, b)) \geq \lambda_{\mu(\varphi)} \int_a^{x_i} dx \int_{x_i}^b \frac{\delta}{(y-x)^2}\, dy,$$

which is absurd because the left-hand side is uniformly bounded from above, while the double integral in the right-hand side is divergent.

Let us consider now an integer $k \in \{\mu(\varphi), \dots, m\}$, and for every $x \in (a, b)$ let us set

$$H_{k,+}(x) := \int_x^{+\infty} \varphi_k\left(\frac{|\hat{u}_\delta(y) - \hat{u}_\delta(x)|}{\delta}\right) \frac{\delta}{(y-x)^2}\, dy,$$
$$H_{k,-}(x) := \int_{-\infty}^{x} \varphi_k\left(\frac{|\hat{u}_\delta(y) - \hat{u}_\delta(x)|}{\delta}\right) \frac{\delta}{(y-x)^2}\, dy.$$

With this notation it turns out that

$$\begin{aligned} \hat{\Lambda}_\delta(\varphi_k, \hat{u}_\delta, (a+\eta, b-\eta)) &= \int_{a+\eta}^{b-\eta} dx \int_{-\infty}^{+\infty} \varphi_k\left(\frac{|\hat{u}_\delta(y) - \hat{u}_\delta(x)|}{\delta}\right) \frac{\delta}{(y-x)^2}\, dy \\ &= \int_{a+\eta}^{b-\eta} (H_{k,+}(x) + H_{k,-}(x))\, dx \\ &\geq \int_{x_\alpha}^{x_{\beta-k}} H_{k,+}(x)\, dx + \int_{x_{\alpha+k}}^{x_\beta} H_{k,-}(x)\, dx. \end{aligned} \tag{4-17}$$

Let us compute the last two integrals separately. For every index $i \in \{\alpha+1, \dots, \beta-k\}$ it turns out that

$$H_{k,+}(x) = \int_{x_{i+k}}^{+\infty} \frac{\delta}{(y-x)^2}\, dy = \frac{\delta}{x_{i+k}-x} \quad \text{for all } x \in (x_{i-1}, x_i).$$

The previous equality assumes that $x_{i-1} < x_i$, but actually it is true for trivial reasons also if $x_{i-1} = x_i$. It follows that

$$\int_{x_{i-1}}^{x_i} H_{k,+}(x)\, dx = \delta \log \frac{x_{i+k} - x_{i-1}}{x_{i+k} - x_i}$$

for every $i \in \{\alpha+1, \dots, \beta-k\}$, and therefore

$$\int_{x_\alpha}^{x_{\beta-k}} H_{k,+}(x)\, dx = \sum_{i=\alpha+1}^{\beta-k} \int_{x_{i-1}}^{x_i} H_{k,+}(x)\, dx = \delta \sum_{i=\alpha+1}^{\beta-k} \log \frac{x_{i+k} - x_{i-1}}{x_{i+k} - x_i}. \tag{4-18}$$

In an analogous way, for every index $i \in \{\alpha+k+1, \dots, \beta\}$ it turns out that

$$H_{k,-}(x) = \int_{-\infty}^{x_{i-k-1}} \frac{\delta}{(y-x)^2}\,dy = \frac{\delta}{x - x_{i-k-1}} \quad \text{for all } x \in (x_{i-1}, x_i),$$

so that, with a shift of indices, we obtain

$$\int_{x_{i+k-1}}^{x_{i+k}} H_{k,-}(x)\,dx = \delta \log \frac{x_{i+k} - x_{i-1}}{x_{i+k-1} - x_{i-1}}$$

for every $i \in \{\alpha+1, \dots, \beta-k\}$, and therefore

$$\int_{x_{\alpha+k}}^{x_\beta} H_{k,-}(x)\,dx = \sum_{i=\alpha+1}^{\beta-k} \int_{x_{i+k-1}}^{x_{i+k}} H_{k,-}(x)\,dx = \delta \sum_{i=\alpha+1}^{\beta-k} \log \frac{x_{i+k} - x_{i-1}}{x_{i+k-1} - x_{i-1}}. \tag{4-19}$$

Plugging (4-18) and (4-19) into (4-17), we find that

$$\hat{\Lambda}_\delta(\varphi_k, \hat{u}_\delta, (a+\eta, b-\eta)) \geq \delta \sum_{i=\alpha+1}^{\beta-k} \log \frac{(x_{i+k} - x_{i-1})^2}{(x_{i+k-1} - x_{i-1})(x_{i+k} - x_i)}.$$

Setting $\ell_i := x_{\alpha+i} - x_{\alpha+i-1}$ for every $i \in \{1, \dots, \beta-\alpha\}$, we can write the last inequality in the form

$$\begin{aligned}\hat{\Lambda}_\delta(\varphi_k, \hat{u}_\delta, (a+\eta, b-\eta)) &\geq \delta \sum_{i=1}^{\beta-\alpha-k} \log \frac{(\ell_i + \dots + \ell_{i+k})^2}{(\ell_i + \dots + \ell_{i+k-1})(\ell_{i+1} + \dots + \ell_{i+k})} \\ &= \delta L_k(\ell_1, \dots, \ell_{\beta-\alpha}),\end{aligned}$$

where L_k is the multivariable function defined in (3-3). We observe that the denominators do not vanish because $k \geq \mu(\varphi)$, and we have already proved that $x_{i+\mu(\varphi)} > x_i$, which is equivalent to saying that the list $(\ell_1, \dots, \ell_{\beta-\alpha})$ contains no k consecutive terms that vanish. Since $\hat{\Lambda}_\delta$ is linear with respect to φ, we deduce

$$\begin{aligned}\hat{\Lambda}_\delta(\varphi, \hat{u}_\delta, (a+\eta, b-\eta)) &\geq \delta \sum_{k=\mu(\varphi)}^{m} \lambda_k L_k(\ell_1, \dots, \ell_{\beta-\alpha}) \\ &= \delta P_{\beta-\alpha,\varphi}(\ell_1, \dots, \ell_{\beta-\alpha}) \geq \delta I_{\beta-\alpha}(\varphi).\end{aligned}$$

Letting $\delta \to 0^+$, and observing that $\beta - \alpha \to +\infty$, we conclude that

$$\begin{aligned}\liminf_{\delta\to 0^+} \hat{\Lambda}_\delta(\varphi, \hat{u}_\delta, (a+\eta, b-\eta)) &\geq \liminf_{\delta\to 0^+} \delta I_{\beta-\alpha}(\varphi) \\ &\geq \liminf_{\delta\to 0^+} \delta(\beta-\alpha) \cdot \liminf_{\delta\to 0^+} \frac{I_{\beta-\alpha}(\varphi)}{\beta-\alpha} \\ &\geq \liminf_{\delta\to 0^+} \delta(\beta-\alpha) \cdot \liminf_{n\to+\infty} \frac{I_n(\varphi)}{n}.\end{aligned} \tag{4-20}$$

It remains to compute the liminf of $\delta(\beta-\alpha)$. To this end, from (4-15) and the minimality of α we deduce

$$A + 2\varepsilon \geq \hat{u}_\delta(x) = (m_0 + \alpha - 1)\delta \quad \text{for all } x \in (x_{\alpha-1}, x_\alpha).$$

Similarly, from (4-16) and the maximality of β we deduce

$$B - 2\varepsilon \leq \hat{u}_\delta(x) = (m_0 + \beta)\delta \quad \text{for all } x \in (x_\beta, x_{\beta+1}).$$

It follows that $(\beta - \alpha)\delta \geq B - A - 4\varepsilon - \delta$, and therefore from (4-20) we conclude that

$$\liminf_{\delta \to 0^+} \hat{\Lambda}_\delta(\varphi, \hat{u}_\delta, (a+\eta, b-\eta)) \geq (B - A - 4\varepsilon) \cdot \liminf_{n \to +\infty} \frac{I_n(\varphi)}{n}.$$

Letting $\varepsilon \to 0^+$, we finally deduce (4-12). □

As observed in [Antonucci et al. 2020], we can rewrite Proposition 4.3 as a relation between the liminf of the energy and the liminf of oscillations as follows.

Corollary 4.4. *Let (a, b), u_δ, φ, and $I_n(\varphi)$ be as in Proposition 4.3. For every $\delta > 0$, let $\operatorname{osc}(u_\delta, (a, b))$ denote the essential oscillation of u_δ in (a, b).*

Then it turns out that

$$\liminf_{\delta \to 0^+} \Lambda_\delta(\varphi, u_\delta, (a, b)) \geq \Big(\liminf_{\delta \to 0^+} \operatorname{osc}(u_\delta, (a, b))\Big) \cdot \liminf_{n \to +\infty} \frac{I_n(\varphi)}{n}.$$

Proof. Let i_δ and s_δ denote the essential infimum and the essential supremum of $u_\delta(x)$ in (a, b), respectively. Let us assume that i_δ and s_δ are real numbers (otherwise an analogous argument works with standard minor changes). Let us set $w_\delta(x) := u_\delta(x) - i_\delta$, and let us observe that

$$\Lambda_\delta(\varphi, u_\delta, (a, b)) = \Lambda_\delta(\varphi, w_\delta, (a, b)) \quad \text{for all } \delta > 0$$

and that w_δ satisfies (4-10) and (4-11) with $A := 0$ and

$$B := \liminf_{\delta \to 0^+} (s_\delta - i_\delta) = \liminf_{\delta \to 0^+} \operatorname{osc}(u_\delta, (a, b)).$$

At this point the conclusion follows from Proposition 4.3. □

Combining Theorem B and Corollary 4.4, we obtain the following result, which connects the Gamma-liminf of the family (1-1) to the multivariable minimum problems of Section 3.

Proposition 4.5. *For every positive integer d and every interaction law $\varphi \in \mathcal{PCA}$ it turns out that*

$$\Gamma\!-\!\liminf_{\delta \to 0^+} \Lambda_\delta(\varphi, u, \mathbb{R}^d) \geq G_d \cdot \frac{1}{2} \liminf_{n \to +\infty} \frac{I_\varphi(n)}{n} \cdot \Lambda_0(u, \mathbb{R}^d) \quad \textit{for all } u \in L^1(\mathbb{R}^d).$$

5. Proofs of our main results

Proof of Theorem 1.3. Let us consider, for every positive integer m, the interaction law defined as

$$\psi_m(t) := \sum_{k=1}^{2^m - 1} \varphi_k(t) \quad \text{for all } t \geq 0. \tag{5-1}$$

This interaction law can be written in the form (1-10) with $a_1 = \cdots = a_m = 1$, and hence $\psi_m \in \mathcal{PCA}_2$. As a consequence, from Proposition 3.3 we deduce

$$\liminf_{n \to +\infty} \frac{I_n(\psi_m)}{n} \geq m \cdot 2 \log 2,$$

and therefore from Proposition 4.5 we obtain

$$\Gamma\text{–}\liminf_{\delta\to 0^+} \Lambda_\delta(\psi_m, u, \mathbb{R}^d) \geq G_d \cdot m \log 2 \cdot \Lambda_0(u, \mathbb{R}^d). \tag{5-2}$$

On the other hand, from (1-5) we know

$$\Gamma\text{–}\lim_{\delta\to 0^+} \Lambda_\delta(\psi_m, u, \mathbb{R}^d) = G_d \cdot N(\psi_m) \cdot K_d(\psi_m) \cdot \Lambda_0(u, \mathbb{R}^d), \tag{5-3}$$

and with some simple calculus we find that

$$N(\psi_m) = \int_0^{+\infty} \frac{\psi_m(t)}{t^2}\,dt = \sum_{k=1}^{2^m-1} \int_0^{+\infty} \frac{\varphi_k(t)}{t^2}\,dt = \sum_{k=1}^{2^m-1} \int_k^{+\infty} \frac{1}{t^2}\,dt = \sum_{k=1}^{2^m-1} \frac{1}{k}. \tag{5-4}$$

Comparing (5-2) and (5-3) we obtain

$$K_d(\psi_m) \geq \frac{m \log 2}{N(\psi_m)} \quad \text{for all } m \geq 1. \tag{5-5}$$

Now from (5-4) we know that $N(\psi_m) \sim m \log 2$ as $m \to +\infty$, and therefore we conclude that $K_d(\psi_m) \to 1$ as $m \to +\infty$, independently of the space dimension. □

Proof of Theorem 1.4(1). Let us consider the interaction law $\theta \in \mathcal{A}_0$ defined by

$$\theta(t) := \begin{cases} 0 & \text{if } t \in [0, 1], \\ t-1 & \text{if } t \in [1, 2], \\ 1 & \text{if } t \geq 2. \end{cases} \tag{5-6}$$

We claim that the shape factor of θ is 1 in any space dimension.

For every positive integer m we consider the interaction law

$$\theta_m(t) := \sum_{k=2^{m-1}}^{2^m-1} \varphi_k(t),$$

and the rescaled function

$$\hat{\theta}_m(t) := \frac{1}{2^{m-1}} \theta_m((2^{m-1}-1)t).$$

To begin with, we show that

$$\theta(t) \geq \hat{\theta}_m(t) \quad \text{for all } t \geq 0, \text{ for all } m \geq 1. \tag{5-7}$$

To this end, we distinguish three cases.

Case 1: If $t \in [0, 1]$, then $(2^{m-1}-1)t \leq 2^{m-1}-1$, and hence $\varphi_k((2^{m-1}-1)t) = 0$ for every $k \geq 2^{m-1}$. It follows that $\hat{\theta}_m(t) = 0$, and hence (5-7) is trivial.

Case 2: If $t \geq 2$, then

$$\hat{\theta}_m(t) \leq \frac{1}{2^{m-1}} \sum_{k=2^{m-1}}^{2^m-1} 1 = 1 = \theta(t),$$

and therefore (5-7) is again satisfied.

Case 3: If $t \in (1, 2)$, let us choose $i \in \{2^{m-1}, \dots, 2^m - 1\}$ such that

$$\frac{i}{2^{m-1}} < t \leq \frac{i+1}{2^{m-1}}.$$

Since

$$(2^{m-1} - 1)t \leq (2^{m-1} - 1) \cdot \frac{i+1}{2^{m-1}} = i - \frac{i+1-2^{m-1}}{2^{m-1}} \leq i,$$

we deduce

$$\varphi_k((2^{m-1} - 1)t) = 0 \quad \text{for all } k \geq i,$$

and therefore

$$\hat{\theta}_m(t) \leq \frac{1}{2^{m-1}} \sum_{k=2^{m-1}}^{i-1} \varphi_k(t) \leq \frac{1}{2^{m-1}}(i - 2^{m-1}) = \frac{i}{2^{m-1}} - 1 \leq t - 1 = \theta(t),$$

which proves (5-7) also in this case.

From inequality (5-7), and the rescaling properties of Λ_δ with respect to the interaction law, we deduce

$$\begin{aligned} \Gamma\!-\liminf_{\delta\to 0^+} \Lambda_\delta(\theta, u, \mathbb{R}^d) &\geq \Gamma\!-\liminf_{\delta\to 0^+} \Lambda_\delta(\hat{\theta}_m, u, \mathbb{R}^d) \\ &= \frac{2^{m-1}-1}{2^{m-1}} \cdot \Gamma\!-\liminf_{\delta\to 0^+} \Lambda_\delta(\theta_m, u, \mathbb{R}^d). \end{aligned} \tag{5-8}$$

Now we observe that the interaction law $\theta_m(t)$ can be written in the form (1-10) with $a_1 = \dots = a_{m-1} = 0$ and $a_m = 1$, and hence $\theta_m \in \mathcal{PCA}_2$. As a consequence, from Proposition 3.3 we deduce

$$\liminf_{n\to+\infty} \frac{I_n(\theta_m)}{n} \geq 2\log 2,$$

and therefore from Proposition 4.5 we obtain

$$\Gamma\!-\liminf_{\delta\to 0^+} \Lambda_\delta(\theta_m, u, \mathbb{R}^d) \geq G_d \cdot \log 2 \cdot \Lambda_0(u, \mathbb{R}^d).$$

Plugging this estimate into (5-8), and letting $m \to +\infty$, we deduce

$$\Gamma\!-\liminf_{\delta\to 0^+} \Lambda_\delta(\theta, u, \mathbb{R}^d) \geq G_d \cdot \log 2 \cdot \Lambda_0(u, \mathbb{R}^d). \tag{5-9}$$

On the other hand, from (1-5) we know

$$\Gamma\!-\lim_{\delta\to 0^+} \Lambda_\delta(\theta, u, \mathbb{R}^d) = G_d \cdot N(\theta) \cdot K_d(\theta) \cdot \Lambda_0(u, \mathbb{R}^d), \tag{5-10}$$

and with some simple calculus we find that

$$N(\theta) = \int_1^2 \frac{t-1}{t^2}\, dt + \int_2^{+\infty} \frac{1}{t^2}\, dt = \log 2.$$

Comparing (5-9) and (5-10) we conclude that $K_d(\theta) = 1$ in any space dimension.

Proof of Theorem 1.4(2). The function ζ is continuous because $f(z) \to 0$ as $z \to -\infty$. It is also bounded and monotone due to the corresponding assumptions on $f(z)$. Finally, assumption (1-11) implies the existence of a constant a such that $\zeta(t) \le at^2$ for every $t \ge 0$. This proves that $\zeta \in \mathcal{A}$.

In order to compute scale and shape factor of ζ, we observe that it can be written in the form

$$\zeta(t) = \sum_{z=-\infty}^{+\infty} (f(z+1) - f(z)) \cdot \theta(2^{-z}t) \quad \text{for all } t \ge 0,$$

where θ is the interaction law defined in (5-6). Due to the additivity and to the rescaling properties of Λ_δ with respect to the interaction law, from this representation it follows that

$$\begin{aligned}\Gamma\text{--}\lim_{\delta\to 0^+} \Lambda_\delta(\zeta, u, \mathbb{R}^d) &\ge \sum_{z=-\infty}^{+\infty} (f(z+1) - f(z))2^{-z} \cdot \Gamma\text{--}\liminf_{\delta\to 0^+} \Lambda_\delta(\theta, u, \mathbb{R}^d) \\ &= \sum_{z=-\infty}^{+\infty} (f(z+1) - f(z))2^{-z} \cdot G_d \cdot N(\theta) \cdot \Lambda_0(u, \mathbb{R}^d), \end{aligned} \tag{5-11}$$

where in the last equality we have exploited that $K_d(\theta) = 1$.

On the other hand, from (1-5) we know

$$\Gamma\text{--}\lim_{\delta\to 0^+} \Lambda_\delta(\zeta, u, \mathbb{R}^d) = G_d \cdot N(\zeta) \cdot K_d(\zeta) \cdot \Lambda_0(u, \mathbb{R}^d). \tag{5-12}$$

Since

$$N(\zeta) = N(\theta) \cdot \sum_{z=-\infty}^{+\infty} (f(z+1) - f(z))2^{-z},$$

comparing (5-11) and (5-12) we conclude that $K_d(\zeta) = 1$ in any space dimension. □

6. Final remarks

In this section we present some variants of our main results, and we speculate about some possible future extensions of the theory developed in this paper.

A counterexample to the short-range question with strict inequalities. Let us consider the interaction laws

$$\varphi_\varepsilon(t) := \begin{cases} c_{1,\varepsilon} \cdot \varepsilon t^2 & \text{if } t \in [0, 1], \\ c_{1,\varepsilon} & \text{if } t > 1, \end{cases} \qquad \psi(t) := c_2\psi_2(t),$$

where $\psi_2(t)$ is defined by (5-1) with $m = 2$, and the constants $c_{1,\varepsilon}$ and c_2 are chosen in such a way that $N(\varphi_\varepsilon) = N(\psi) = 1$.

From (5-5) and (5-4) with $m = 2$ it follows that $K_d(\psi) \ge (12/11)\log 2$. On the other hand, it is possible (but not completely trivial) to show that $K_d(\varphi_\varepsilon) \to \log 2$ as $\varepsilon \to 0^+$.

Therefore, when ε is small enough, this is an example of two interaction laws φ_ε and ψ with equal scale factor, satisfying $\varphi_\varepsilon(t) > \psi(t)$ for every $t \in (0, 1]$, but nevertheless $K_d(\varphi_\varepsilon) < K_d(\psi)$ in every space dimension. This provides a counterexample to Question 1 with strict inequalities.

True Gamma-limits and smooth recovery families. In this paper we limited ourselves to providing estimates from below for the Gamma-liminf, since they are enough to establish both Theorems 1.3 and 1.4. On the other hand, with little further effort we could prove that actually the lower bound coincides with the Gamma-limit.

This is evident in the case of the interaction laws with shape factor equal to 1, for example all those provided by statement (2) of Theorem 1.4, because for them the pointwise limit coincides on smooth functions with the estimate from below for the Gamma-liminf. Therefore, for all these interaction laws we now know both the Gamma-limit with exact values of the constants in any space dimension and the existence of smooth recovery families.

As for the interaction laws $\psi_m(t)$ defined by (5-1), again we can show that the Gamma-limit coincides with the lower bound we obtained for the Gamma-liminf, namely

$$\Gamma\text{–}\lim_{\delta\to 0^+} \Lambda_\delta(\psi_m, u, \mathbb{R}^d) = G_d \cdot m \log 2 \cdot \Lambda_0(u, \mathbb{R}^d). \tag{6-1}$$

In order to prove this result, one should follow the path we pursued in [Antonucci et al. 2020]. The main idea is that in any space dimension the family $S_\delta u$ of vertical δ-segmentations of u is a recovery family when u is piecewise C^1 or piecewise affine with compact support, and those classes are dense in energy for the right-hand side of (6-1). Since vertical δ-segmentations of piecewise affine functions with compact support are step functions with level sets that are finite unions of polytopes, it is enough to further approximate them in order to produce recovery sequences made by functions of class C^∞ with compact support. We refer to [Antonucci et al. 2020] for the details. Therefore, also in the case of the interaction laws $\psi_m(t)$, we end up with a Gamma-convergence result with both exact values of the constants in any space dimension, and existence of smooth recovery families.

The same argument should work for all interaction laws in $\mathcal{PCA}_2$, and more generally for all interaction laws $\varphi \in \mathcal{PCA}$ for which $I_n(\varphi)$ is realized asymptotically when all the variables are equal.

Toward a general formula for the Gamma-limit. We suspect that the lower bound in Proposition 4.3 might be optimal for every $\varphi \in \mathcal{PCA}$, and that the liminf in the right-hand side of (4-12) is actually a limit. Thanks to Proposition 4.5, this would imply

$$\Gamma\text{–}\lim_{\delta\to 0^+} \Lambda_\delta(\varphi, u, \mathbb{R}^d) = G_d \cdot \frac{1}{2} \lim_{n\to+\infty} \frac{I_n(\varphi)}{n} \cdot \Lambda_0(u, \mathbb{R}^d) \quad \text{for all } u \in L^1(\mathbb{R}^d), \tag{6-2}$$

for every $\varphi \in \mathcal{PCA}$. In order to prove this result, the main difficulty seems to be the construction of recovery sequences, which in general can no longer be obtained simply by vertical δ-segmentation. On the contrary, the construction should now take into account the pattern that realizes the infimum $I_n(\varphi)$.

A representation of the form (6-2), if true, would be important because any interaction law can be approximated from below by piecewise constant interaction laws with steps of equal horizontal length (as we did in the proof of statement (1) of Theorem 1.4), and these laws are rescalings of laws in $\mathcal{PCA}$.

This kind of representation would be even more important if it were true that the Gamma-limit of $\Lambda_\delta(\varphi, u, \mathbb{R}^d)$ is the supremum of the Gamma-limits of $\Lambda_\delta(\psi, u, \mathbb{R}^d)$ as ψ varies in the set of all piecewise constant interaction laws, with steps of equal horizontal length that are less than or equal to φ. A confirmation of this conjecture would open the way for answering several questions raised in [Brezis and Nguyen 2018]: a simplified proof of the Gamma-convergence result in full generality, a less implicit formula for shape factors, and existence of smooth recovery families.

Characterization of interaction laws without gap. Concerning the gap between the pointwise limit and the Gamma-limit, the challenge is now characterizing all interaction laws with shape factor equal to 1. Let us summarize what we know for the time being on this specific issue.

• Continuity does not guarantee lack of the gap. Among continuous interaction laws, we have both examples without gap (all interaction laws provided by Theorem 1.4) and interaction laws with gap. Indeed, it is possible to show that the piecewise affine interaction law that is equal to 0 for $t \in [0, 1-\varepsilon]$ and equal to 1 for $t \geq 1$ has a shape factor that tends to $\log 2$ as $\varepsilon \to 0^+$.

Conversely, we have no example of discontinuous interaction law without gap.

• It is not a matter of vanishing in a neighborhood of the origin. Among the interactions laws in $\mathcal{A}_0$ we have both examples without gap (the interaction law θ defined in (5-6)) and examples with gap (the model interaction law φ_1). Among the interaction laws that are positive for every $t > 0$ we have both examples without gap (defined as in statement (2) of Theorem 1.4) and examples with gap (the interaction law $\varphi_\varepsilon(t)$ defined at the beginning of this section).

• The shape factor is concave when restricted to interaction laws with equal scale factor. As a consequence, any convex combination of interaction laws with the same scale factor, and shape factor equal to 1, has again shape factor equal to 1. Considering that now we know many interaction laws with shape factor equal to 1, this leads us to guess that the set of interaction laws with shape factor equal to 1 might be quite large.

More general exponents. It should not be difficult to extend the results of this paper to the more general family of functionals

$$\Lambda_{\delta,p}(\varphi, u, \Omega) := \iint_{\Omega^2} \varphi\left(\frac{|u(y)-u(x)|}{\delta}\right) \frac{\delta}{|y-x|^{d+p}}\, dx\, dy,$$

where $p > 1$ is a real number. This case was investigated in [Brezis and Nguyen 2020]. The Gamma-limit turns out to be a multiple of the L^p-norm of the gradient of u, and the exact constant was found in [Antonucci et al. 2020] in the case $\varphi = \varphi_1$. When extending the results of this paper, the presence of the general exponent $p > 1$ requires probably only a change in definition (3-3), which now should be replaced by something like

$$\begin{aligned} L_{k,p} &:= \int_{S_{i,k}}^{S_{i,k+1}} \frac{1}{\sigma^p}\, d\sigma + \int_{S_{i+1,k}}^{S_{i,k+1}} \frac{1}{\sigma^p}\, d\sigma \\ &= \frac{1}{p-1} \sum_{i=1}^{n-k} \left(-\frac{2}{[S_{i,k+1}]^{p-1}} + \frac{1}{[S_{i,k}]^{p-1}} + \frac{1}{[S_{i+1,k}]^{p-1}} \right). \end{aligned}$$

Again, the special structure of the terms of the sum should guarantee the telescopic effect as in Lemma 3.1.

Acknowledgments

We would like to thank H. Brezis for his positive feedback on the manuscript [Antonucci et al. 2020], and for encouraging us to investigate the gap between the Gamma-limit and the pointwise limit.

References

[Antonucci et al. 2018] C. Antonucci, M. Gobbino, M. Migliorini, and N. Picenni, "On the shape factor of interaction laws for a non-local approximation of the Sobolev norm and the total variation", *C. R. Math. Acad. Sci. Paris* **356**:8 (2018), 859–864. MR Zbl

[Antonucci et al. 2020] C. Antonucci, M. Gobbino, M. Migliorini, and N. Picenni, "Optimal constants for a nonlocal approximation of Sobolev norms and total variation", *Anal. PDE* **13**:2 (2020), 595–625.

[Bourgain and Nguyen 2006] J. Bourgain and H.-M. Nguyen, "A new characterization of Sobolev spaces", *C. R. Math. Acad. Sci. Paris* **343**:2 (2006), 75–80. MR Zbl

[Brezis 2015] H. Brezis, "New approximations of the total variation and filters in imaging", *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* **26**:2 (2015), 223–240. MR Zbl

[Brezis 2016] H. Brezis, "Another triumph for De Giorgi's Gamma convergence", recorded lecture from "A mathematical tribute to Ennio De Giorgi" conference, 2016, available at https://www.youtube.com/watch?v=1Y6fvZX1fx8.

[Brezis and Nguyen 2017] H. Brezis and H.-M. Nguyen, "Non-convex, non-local functionals converging to the total variation", *C. R. Math. Acad. Sci. Paris* **355**:1 (2017), 24–27. MR Zbl

[Brezis and Nguyen 2018] H. Brezis and H.-M. Nguyen, "Non-local functionals related to the total variation and connections with image processing", *Ann. PDE* **4**:1 (2018), art. id. 9. MR Zbl

[Brezis and Nguyen 2020] H. Brezis and H.-M. Nguyen, "Non-local, non-convex functionals converging to Sobolev norms", *Nonlinear Anal.* **191** (2020), art. id. 111626.

[Garsia and Rodemich 1974] A. M. Garsia and E. Rodemich, "Monotonicity of certain functionals under rearrangement", *Ann. Inst. Fourier* (*Grenoble*) **24**:2 (1974), 67–116. MR Zbl

[Gobbino 1998] M. Gobbino, "Finite difference approximation of the Mumford–Shah functional", *Comm. Pure Appl. Math.* **51**:2 (1998), 197–228. MR Zbl

[Gobbino and Mora 2001] M. Gobbino and M. G. Mora, "Finite-difference approximation of free-discontinuity problems", *Proc. Roy. Soc. Edinburgh Sect. A* **131**:3 (2001), 567–595. MR Zbl

[Nguyen 2006] H.-M. Nguyen, "Some new characterizations of Sobolev spaces", *J. Funct. Anal.* **237**:2 (2006), 689–720. MR Zbl

[Nguyen 2007] H.-M. Nguyen, "Γ-convergence and Sobolev norms", *C. R. Math. Acad. Sci. Paris* **345**:12 (2007), 679–684. MR Zbl

[Nguyen 2008] H.-M. Nguyen, "Further characterizations of Sobolev spaces", *J. Eur. Math. Soc.* **10**:1 (2008), 191–229. MR Zbl

[Nguyen 2011] H.-M. Nguyen, "Γ-convergence, Sobolev norms, and BV functions", *Duke Math. J.* **157**:3 (2011), 495–533. MR Zbl

[Taylor 1973] H. Taylor, "Rearrangements of incidence tables", *J. Combinatorial Theory Ser. A* **14** (1973), 30–36. MR Zbl

Received 29 May 2018. Revised 23 Dec 2018. Accepted 7 Mar 2019.

CLARA ANTONUCCI: clara.antonucci@sns.it
Scuola Normale Superiore, Pisa, Italy

MASSIMO GOBBINO: massimo.gobbino@unipi.it
Università degli Studi di Pisa, Pisa, Italy

NICOLA PICENNI: nicola.picenni@sns.it
Scuola Normale Superiore, Pisa, Italy

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 13, No. 3, 2020
dx.doi.org/10.2140/apde.2020.13.651

EXTERNAL BOUNDARY CONTROL OF THE MOTION OF A RIGID BODY IMMERSED IN A PERFECT TWO-DIMENSIONAL FLUID

OLIVIER GLASS, JÓZSEF J. KOLUMBÁN AND FRANCK SUEUR

We consider the motion of a rigid body immersed in a two-dimensional irrotational perfect incompressible fluid. The fluid is governed by the Euler equation, while the trajectory of the solid is given by Newton's equation, the force term corresponding to the fluid pressure on the body's boundary only. The system is assumed to be confined in a bounded domain with an impermeable condition on a part of the external boundary. The issue considered here is the following: is there an appropriate boundary condition on the remaining part of the external boundary (allowing some fluid going in and out the domain) such that the immersed rigid body is driven from some given initial position and velocity to some final position (in the same connected component of the set of possible positions as the initial position) and velocity in a given positive time, without touching the external boundary? In this paper we provide a positive answer to this question thanks to an impulsive control strategy. To that purpose we make use of a reformulation of the solid equation into an ODE of geodesic form, with some force terms due to the circulation around the body, as used by Glass, Munnier and Sueur (*Invent. Math.* **214**:1 (2018), 171–287), and some extra terms here due to the external boundary control.

1. Introduction and main result

1A. ***The model without control.*** A simple model of fluid-solid evolution is that of a single rigid body surrounded by a perfect incompressible fluid. Let us describe this system. We consider a two-dimensional bounded, open, smooth and simply connected[1] domain $\Omega \subset \mathbb{R}^2$. The domain Ω is composed of two disjoint parts: the open part $\mathcal{F}(t)$ filled with fluid and the closed part $\mathcal{S}(t)$ representing the solid; see

MSC2010: 76B75, 93C15, 93C20.
Keywords: fluid-solid interaction, impulsive control, geodesics, coupled ODE/PDE system, fluid mechanics, Euler equation, control problem, external boundary control.

[1] The condition of simple connectedness is actually not essential and one could generalize the present result to the case where Ω is merely open and connected at the price of long but straightforward modifications.

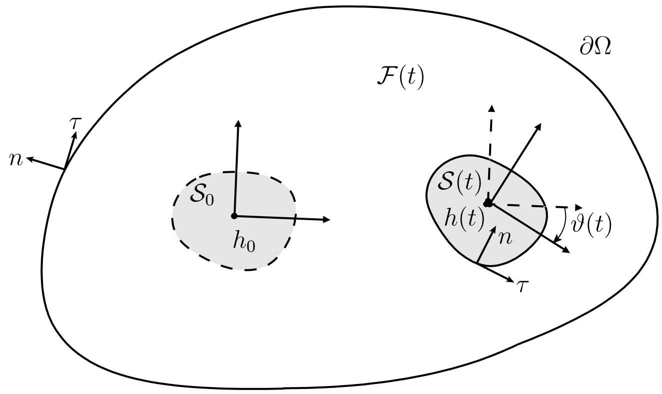

Figure 1. The domains Ω, $\mathcal{S}(t)$ and $\mathcal{F}(t) = \Omega \setminus \mathcal{S}(t)$.

Figure 1. These parts depend on time t. Furthermore, we assume that $\mathcal{S}(t)$ is also smooth and simply connected. On the fluid part $\mathcal{F}(t)$, the velocity field $u : \{(t,x) : t \in [0,T],\ x \in \overline{\mathcal{F}(t)}\} \to \mathbb{R}^2$ and the pressure field $\pi : \{(t,x) : t \in [0,T],\ x \in \overline{\mathcal{F}(t)}\} \to \mathbb{R}$ satisfy the incompressible Euler equation:

$$\frac{\partial u}{\partial t} + (u \cdot \nabla)u + \nabla\pi = 0 \quad \text{and} \quad \operatorname{div} u = 0 \quad \text{for } t \in [0,T] \text{ and } x \in \mathcal{F}(t). \tag{1-1}$$

We consider impermeability boundary conditions, namely, on the solid boundary, the normal velocity coincides with the solid normal velocity

$$u \cdot n = u_S \cdot n \quad \text{on } \partial\mathcal{S}(t), \tag{1-2}$$

where u_S denotes the solid velocity described below, while on the outer part of the boundary we have

$$u \cdot n = 0 \quad \text{on } \partial\Omega, \tag{1-3}$$

where n is the unit outward normal vector on $\partial\mathcal{F}(t)$. The solid $\mathcal{S}(t)$ is obtained by a rigid movement from $\mathcal{S}(0)$, and one can describe its position by the center of mass, $h(t)$, and the angle variable with respect to the initial position, $\vartheta(t)$. Consequently, we have

$$\mathcal{S}(t) = h(t) + R(\vartheta(t))(\mathcal{S}_0 - h_0), \tag{1-4}$$

where h_0 is the center of mass at initial time, and

$$R(\vartheta) = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}.$$

Moreover the solid velocity is hence given by

$$u_S(t,x) = h'(t) + \vartheta'(t)(x - h(t))^{\perp}, \tag{1-5}$$

where for $x = (x_1, x_2)$ we define $x^{\perp} = (-x_2, x_1)$.

The solid evolves according to Newton's law, and is influenced by the fluid's pressure on the boundary:

$$mh''(t) = \int_{\partial\mathcal{S}(t)} \pi n \, d\sigma \quad \text{and} \quad \mathcal{J}\vartheta''(t) = \int_{\partial\mathcal{S}(t)} \pi (x - h(t))^{\perp} \cdot n \, d\sigma. \tag{1-6}$$

Here the constants $m > 0$ and $\mathcal{J} > 0$ denote respectively the mass and the moment of inertia of the body, where the fluid is supposed to be homogeneous of density 1, without loss of generality. Furthermore, the circulation around the body is constant in time, that is,

$$\int_{\partial\mathcal{S}(t)} u(t) \cdot \tau \, d\sigma = \int_{\partial\mathcal{S}_0} u_0 \cdot \tau \, d\sigma = \gamma \in \mathbb{R} \quad \text{for all } t \geq 0, \tag{1-7}$$

due to Kelvin's theorem, where τ denotes the unit counterclockwise tangent vector.

The Cauchy problem for this system with initial data

$$\begin{gathered} u|_{t=0} = u_0 \quad \text{for } x \in \mathcal{F}(0), \\ h(0) = h_0, \quad h'(0) = h'_0, \quad \vartheta(0) = 0, \quad \vartheta'(0) = \vartheta'_0 \end{gathered} \tag{1-8}$$

is now well-understood; see, e.g., [Glass et al. 2014; Glass and Sueur 2015; Houot et al. 2010; Ortega et al. 2005; 2007]. Furthermore, the three-dimensional case has also been studied in [Glass et al. 2012; Rosier and Rosier 2009]. Note in passing that it is our convention used throughout the paper that $\vartheta(0) = 0$.

In this paper, we will furthermore assume that the fluid is irrotational at the initial time, that is $\operatorname{curl} u_0 = 0$ in $\mathcal{F}(0)$, which implies that it stays irrotational at all times, due to Helmholtz's third theorem, i.e.,

$$\operatorname{curl} u = 0 \quad \text{for } x \in \mathcal{F}(t), \text{ for all } t \geq 0. \tag{1-9}$$

1B. ***The control problem and the main result.*** We are now in position to state our main result.

Our goal is to investigate the possibility of controlling the solid by means of a boundary control acting on the fluid. Consider Σ a nonempty, open part of the outer boundary $\partial\Omega$. Suppose that one can choose some nonhomogeneous boundary conditions on Σ. One natural possibility is due to Yudovich [1962], which consists of prescribing on the one hand the normal velocity on Σ, i.e., choosing some function $g \in C_0^\infty([0, T] \times \Sigma)$ with $\int_\Sigma g = 0$ and imposing that

$$u(t, x) \cdot n(x) = g(t, x) \quad \text{on } [0, T] \times \Sigma, \tag{1-10}$$

while on the rest of the boundary we have the usual impermeability condition

$$u \cdot n = 0 \quad \text{on } [0, T] \times (\partial\Omega \setminus \Sigma), \tag{1-11}$$

and on the other hand the vorticity on the set Σ^- of points of $[0, T] \times \Sigma$ where the velocity field points inside Ω. Note that Σ^- is deduced immediately from g.

Since we are interested in the vorticity-free case, we will actually consider here a null control in vorticity, that is,

$$\operatorname{curl} u(t, x) = 0 \quad \text{on } \Sigma^- = \{(t, x) \in [0, T] \times \Sigma : u(t, x) \cdot n(x) < 0\}. \tag{1-12}$$

Condition (1-12) enforces the validity of (1-9) as in the uncontrolled setting despite the fact that some fluid is entering the domain.

The general question of this paper is how to control the solid's movement by using the above boundary control (that is, the function g). In particular we raise the question of driving the solid from a given position and a given velocity to some other prescribed position and velocity. Note that we cannot expect to control the fluid velocity in the situation described above: for instance, Kelvin's theorem gives an invariant of the dynamics, regardless of the control.

Throughout this paper we will only consider solid trajectories which stay away from the boundary. Therefore we introduce

$$\mathcal{Q} = \{q := (h, \vartheta) \in \Omega \times \mathbb{R} : d(h + R(\vartheta)(\mathcal{S}_0 - h_0), \partial\Omega) > 0\}.$$

Furthermore, let us from here on set

$$\mathcal{D}_T := \{(t, x) : t \in [0, T],\ x \in \overline{\mathcal{F}(t)}\},$$

where we have omitted from the notation the dependence on $\mathcal{F}(\cdot)$, and therefore on the unknown $(h, \vartheta)(\cdot)$.

The main result of this paper is the following statement.

Theorem 1. *Let $T > 0$. Consider $\mathcal{S}_0 \subset \Omega$ bounded, closed, simply connected with smooth boundary, which is not a disk, and $u_0 \in C^\infty(\overline{\mathcal{F}(0)}; \mathbb{R}^2)$, $\gamma \in \mathbb{R}$, $q_0 = (h_0, 0)$, $q_1 = (h_1, \vartheta_1) \in \mathcal{Q}$, $h_0', h_1' \in \mathbb{R}^2$, $\vartheta_0', \vartheta_1' \in \mathbb{R}$ such that $(h_0, 0)$ and (h_1, ϑ_1) belong to the same connected component of $\mathcal{Q}$ and*

$$\operatorname{div} u_0 = \operatorname{curl} u_0 = 0 \quad \text{in } \mathcal{F}(0), \qquad u_0 \cdot n = 0 \quad \text{on } \partial\Omega,$$
$$u_0 \cdot n = (h_0' + \vartheta_0'(x - h_0)^\perp) \cdot n \quad \text{on } \partial\mathcal{S}_0, \qquad \int_{\partial\mathcal{S}_0} u_0 \cdot \tau \, d\sigma = \gamma.$$

(See Figure 2.) Then there exists a control $g \in C_0^\infty((0, T) \times \Sigma)$ and a solution $(h, \vartheta, u) \in C^\infty([0, T]; \mathcal{Q}) \times C^\infty(\mathcal{D}_T; \mathbb{R}^2)$ to (1-1), (1-2), (1-6), (1-7), (1-8), (1-9), (1-10), (1-11), *which satisfies $(h, h', \vartheta, \vartheta')(T) = (h_1, h_1', \vartheta_1, \vartheta_1')$.*

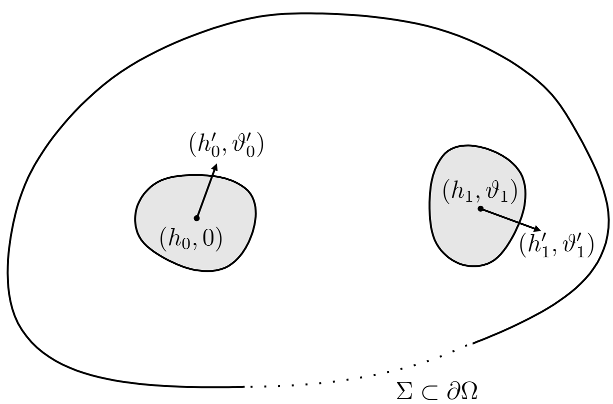

Figure 2. The initial and final positions and velocities in the control problem.

Remark 2. In Theorem 1 the control g can be chosen with an arbitrary small total flux through Σ^-; that is, for any $T > 0$, for any $\nu > 0$, there exists a control g and a solution (h, ϑ, u) satisfying the properties of Theorem 1 and such that moreover

$$\left| \int_0^T \int_{\Sigma^-} u \cdot n \, d\sigma \, dt \right| < \nu.$$

See Section 5D for more explanation. Let us mention that such a small flux condition cannot be guaranteed in the results [Coron 1993; Glass 2000; 2001] regarding the controllability of the Euler equations.

When $\mathcal{S}_0$ is a disk, the second equation in (1-6) becomes degenerate, so it needs to be treated separately. For instance, in the case of a homogeneous disk, i.e., when the center of mass coincides with the center of the disk and we have $(x - h(t))^\perp \cdot n = 0$ for any $x \in \partial\mathcal{S}(t)$, $t \geq 0$, we cannot control ϑ. However, we have a similar result for controlling the center of mass h.

Theorem 3. *Let $T > 0$. Given a homogeneous disk $\mathcal{S}_0 \subset \Omega$, $u_0 \in C^\infty(\overline{\mathcal{F}(0)}; \mathbb{R}^2)$, $\gamma \in \mathbb{R}$, $h_0, h_1 \in \Omega$, $h_0', h_1' \in \mathbb{R}^2$ such that $(h_0, 0)$ and $(h_1, 0)$ are in the same connected component of $\mathcal{Q}$, and*

$$\operatorname{div} u_0 = \operatorname{curl} u_0 = 0 \quad \text{in } \mathcal{F}(0), \qquad u_0 \cdot n = 0 \quad \text{on } \partial\Omega,$$
$$u_0 \cdot n = h_0' \cdot n \quad \text{on } \partial\mathcal{S}_0, \qquad \int_{\partial\mathcal{S}_0} u_0 \cdot \tau \, d\sigma = \gamma,$$

there exists $g \in C_0^\infty((0, T) \times \Sigma)$ and a solution (h, u) in $C^\infty([0, T]; \Omega) \times C^\infty(\mathcal{D}_T; \mathbb{R}^2)$ of (1-1), (1-2), (1-6), (1-7), (1-9), (1-10), (1-11), (1-12) *with initial data (h_0, h_0', u_0), which satisfies $(h, h')(T) = (h_1, h_1')$.*

The proof is similar to that of Theorem 1, with the added consideration that $(x - h(t))^\perp \cdot n = 0$ for any $x \in \partial\mathcal{S}(t)$, $t \geq 0$. We therefore omit the proof. In the case where the disk is nonhomogeneous the analysis is technically more intricate already in the uncontrolled setting, see [Glass et al. 2018], and we will omit this case in this paper.

References. Let us mention a few results of boundary controllability of a fluid alone, that is without any moving body. The problem is then finding a boundary control which steers the fluid velocity from u_0 to some prescribed state u_1. For the incompressible Euler equations small-time global exact boundary controllability has been obtained in [Coron 1993; Glass 2000] in the two-dimensional, respectively three-dimensional case. This result has been recently extended to the case of the incompressible Navier–Stokes equation with Navier slip-with-friction boundary conditions in [Coron et al. 2020]; see also [Coron et al. 2017] for a gentle exposition. Note that the proof there relies on the previous results for the Euler equations by means of a rapid and strong control which drives the system in a high Reynolds regime. This strategy was initiated in [Coron 1996], where an interior controllability result was already established. For "viscous fluid + rigid body" control systems (with Dirichlet boundary conditions), local controllability results have already been obtained in both two and three dimensions; see, e.g., [Boulakia and Guerrero 2013; Boulakia and Osses 2008; Imanuvilov and Takahashi 2007]. These results rely on Carleman estimates on the linearized equation, and consequently on the parabolic character of the fluid equation.

A different type of fluid-solid control result can be found in [Glass and Rosier 2013], where the fluid is governed by the two-dimensional Euler equation. However in this paper the control is located on the solid's boundary which makes the situation quite different.

Actually, the results of Theorems 1 and 3 can rather be seen as some extensions to the case of an immersed body of the results [Glass and Horsin 2010; 2012; 2016] concerning Lagrangian controllability of the incompressible Euler and Stokes equations, where the control takes the same form as here.

1C. ***Generalizations and open problems.*** First, as we mentioned before, using the techniques of this paper, the result could be straightforwardly generalized for domains that are not simply connected. One could also manage in the same way the control of several solids (the reader may in particular see that the argument using Runge's theorem in Section 7 is local around the solid).

We would also like to underline that the absence of vorticity is not central here. This may surprise the reader acquainted with the Euler equation, but actually following the arguments of [Coron 1993; 1996], one knows how to control the full model when one can control the irrotational one. This is by the way the technique that we use to take care of the circulation γ (see in particular Section 3). But the presence of vorticity makes a lot of complications from the point of view of the initial boundary problem, in particular for what concerns the uniqueness issue; see [Yudovich 1962]. To avoid these unnecessary technical complications, we restrain ourselves to the irrotational problem. But the full problem could undoubtedly be treated in the same way.

Furthermore, one might ask the question of whether or not it is possible to control with a reduced number of controls, i.e., to only look for controls g which take the form of a linear combination of some a priori given controls $\{g_i\}_{i=1,\dots,I}$, which may depend on the geometry, but not the initial or final data of the control problem. We consider that our methods can be adapted to prove such a result, in particular since in Section 3 we prove that Theorem 1 follows from a simpler result, Theorem 13, where the solid displacement, the solid velocities and the circulation are small. It then suffices to discretize the control with respect to the parameters $(h_0, h_0', \vartheta_0, \vartheta_0')$, $(h_1, h_1', \vartheta_1, \vartheta_1')$ and γ. This does not pose a problem since our control is actually constructed continuously with respect to these parameters, so one may apply a compactness argument. However, the set of controls $\{g_i\}_{i=1,\dots,I}$ will depend on the parameter $\delta > 0$ from Theorem 13, used to restrict the set of admissible positions $\mathcal{Q}$ to the set $\mathcal{Q}_\delta$ defined in (3-1). This subtlety is due to the fact that the closure of $\mathcal{Q}$ also contains points where the solid touches the outer boundary, while this is no longer the case with $\mathcal{Q}_\delta$ for a given fixed $\delta > 0$, and we use this for the compactness argument mentioned above.

There remain also many open problems.

Considering the recent progress on the controllability in the viscous case, a natural question is whether or not the results in this paper could be adapted to the case where a rigid body is moving in a fluid driven by the incompressible Navier–Stokes equation. In [Kolumbán 2020] we extend the analysis performed here to prove the small-time global controllability of the motion of a rigid body in a viscous incompressible fluid, driven by the incompressible Navier–Stokes equation, in the case where Navier slip-with-friction boundary conditions are prescribed at the interface between the fluid and the solid. However, the case of Dirichlet boundary conditions remains completely open.

Let us mention the following open problem regarding the motion planning of a rigid body immersed in an inviscid incompressible irrotational flow.

Open problem. Let $T > 0$, $(h_0, 0)$ in $\mathcal{Q}$, and ξ in $C^2([0, T]; \mathcal{Q})$, with $\xi(0) = (h_0, 0)$. Let us decompose $\xi'(0)$ into $\xi'(0) = (h_0', \vartheta_0')$. Consider $\mathcal{S}_0 \subset \Omega$ bounded, closed, simply connected with smooth boundary, which is not a disk, $\gamma \in \mathbb{R}$, and $u_0 \in C^\infty(\overline{\mathcal{F}(0)}; \mathbb{R}^2)$ such that

$$\begin{aligned} \operatorname{div} u_0 = \operatorname{curl} u_0 = 0 \quad \text{in } \mathcal{F}(0), \qquad & u_0 \cdot n = 0 \quad \text{on } \partial\Omega, \\ u_0 \cdot n = (h_0' + \vartheta_0'(x - h_0)^\perp) \cdot n \quad \text{on } \partial\mathcal{S}_0, \qquad & \int_{\partial\mathcal{S}_0} u_0 \cdot \tau \, d\sigma = \gamma. \end{aligned}$$

Do there exist $g \in C([0, T] \times \Sigma)$ and a solution $(h, \vartheta, u) \in C^2([0, T]; \mathcal{Q}) \times C^1(\mathcal{D}_T; \mathbb{R}^2)$ to (1-1), (1-2), (1-6), (1-7), (1-8), (1-9), (1-10), (1-11), which satisfies $\xi = (h, \vartheta)$?

Even the approximate motion planning in C^2, i.e., the same statement as above but with

$$\|\xi - (h, \vartheta)\|_{C^2([0,T])} \leq \varepsilon$$

(with $\varepsilon > 0$ arbitrary) instead of $\xi = (h, \vartheta)$, is an open problem.

Furthermore, in this paper we have ignored any possible thermodynamic effect in the model; however, it would be a natural question to ask how our results could be generalized to the case when the fluid is heat-conductive.

1D. ***Plan of the paper and main ideas behind the proof of Theorem 1.*** The paper is organized as follows.

In Section 2 we first recall from [Glass et al. 2018] a reformulation of the Newton equations (1-6) as an ODE in the uncontrolled case and then extend it to the case with control.

To be more precise, setting $q := (h, \vartheta)$ and considering a manifold of admissible positions $\mathcal{Q}$ (to be defined later), the authors proved in [Glass et al. 2018] that there exist a field $\mathcal{M} : \mathcal{Q} \to S^{++}(\mathbb{R}^3)$ of symmetric positive-definite matrices and smooth fields $E, B : \mathcal{Q} \to \mathbb{R}^3$ such that the fluid-solid system is equivalent to the following ODE in q:

$$\mathcal{M}(q)q'' + \langle \Gamma(q), q', q' \rangle = \gamma^2 E(q) + \gamma q' \times B(q),$$

where $\Gamma(q)$ is a bilinear symmetric mapping, given by the so-called Christoffel symbols of the first kind:

$$\Gamma_{i,j}^k = \frac{1}{2}\left(\frac{\partial(\mathcal{M})_{k,j}}{\partial q_i} + \frac{\partial(\mathcal{M})_{k,i}}{\partial q_j} - \frac{\partial(\mathcal{M})_{i,j}}{\partial q_k} \right).$$

In particular, the case with zero circulation represents the fact that the particle q is moving along the geodesics associated with the Riemannian metric induced on $\mathcal{Q}$ by the so-called total inertia matrix $\mathcal{M}$.

We extend the above result to the case with control $g \in C_0^\infty([0, T] \times \Sigma)$ to find that q satisfies the ODE

$$\mathcal{M}(q)q'' + \langle \Gamma(q), q', q' \rangle = \gamma^2 E(q) + \gamma q' \times B(q) + F_1(q, q', \gamma)[\alpha] + F_2(q)[\partial_t \alpha], \tag{1-13}$$

where F_1 and F_2 are regular, and α is defined as the unique smooth solution of the Neumann problem

$$\Delta\alpha = 0 \quad \text{in } \mathcal{F}(t) \qquad \text{and} \qquad \partial_n \alpha = g \mathbb{1}_\Sigma \quad \text{on } \partial\mathcal{F}(t), \tag{1-14}$$

with zero mean.

Note that in both cases above, the fluid velocity u can be recovered by solving some simple elliptic PDEs.

In Section 3 we prove that Theorem 1 can be deduced from a simpler result, namely Theorem 13, where the solid displacement, the initial and final solid velocities and the circulation are assumed to be small.

This is achieved on one hand by using the usual time-rescale properties of the Euler equation in order to pass from arbitrary solid velocities and circulation to small ones. More precisely, if $u(t,\cdot)$ is a solution to the Euler equation on $[0, T]$, then for any $\lambda > 0$

$$u^\lambda(t,\cdot) := \frac{1}{\lambda} u\left(\frac{t}{\lambda},\cdot\right)$$

is a solution to the Euler equation on the time interval $[0, \lambda T]$. The corresponding scaling for the initial and final solid velocities and the circulation associated with u^λ becomes q_0'/λ, q_1'/λ and γ/λ. Hence, if one can find a solution with small initial and final velocities and small circulation on $[0, T]$, one can pass to the arbitrary (or large) case on $[0, \lambda T]$ with $\lambda \in (0, 1)$ small enough, thus obtaining the controllability result in smaller time. There are multiple possibilities for using up the remaining time from λT to T, and we give one in Section 3, relying on the time-reversal properties of the Euler equation.

On the other hand, one may use a compact covering argument to pass from the case when q_0 and q_1 are remote to the case when their distance is small.

In Section 4 we prove that another reduction is possible, as we prove that an approximate controllability result (rather than an exact one), namely Theorem 14, allows us to deduce Theorem 13.

Indeed, if instead of $(q, q')(T) = (q_1, q_1')$ one only has $\|(q, q')(T) - (q_1, q_1')\| \le \eta$ for $\eta > 0$ small enough, then it is possible to pass to exact controllability by using a Brouwer-type topological argument. However, for such a result to be applied, one has to make sure that the map $(q_1, q_1') \mapsto (q, q')(T)$ is well-defined and continuous for (q_1, q_1') in some small enough ball, which we will indeed achieve during our construction.

Section 5 is devoted to the proof of Theorem 14 and is the core of the paper. In order to achieve the aforementioned approximate controllability, we rely on the following strategy.

Suppose we have $\gamma = 0$ (if this is not the case, one can at least expect to be close in some sense to the case without circulation when γ is small enough), and suppose that we can find some appropriate control $g \in C_0^\infty([0, T]; \mathcal{C})$ such that the term $F_1(q, q', 0)[\alpha] + F_2(q)[\partial_t \alpha]$ in (1-13) behaves approximately like $v_0\delta_0(t) + v_1\delta_T(t)$ for any given $v_0, v_1 \in \mathbb{R}^3$, where δ_0 and δ_T denote the Dirac distributions at times $t = 0^+$ and $t = T^-$ respectively.

Then, (1-13) is going to be close (in an appropriate sense) to the formal toy model

$$\mathcal{M}(\tilde{q})\tilde{q}'' + \langle \Gamma(\tilde{q}), \tilde{q}', \tilde{q}' \rangle = v_0\delta_0 + v_1\delta_T, \tag{1-15}$$

and controlling (1-13) (at least approximately) reduces to controlling (1-15) by using the vectors $v_0, v_1 \in \mathbb{R}^3$ as our control. In fact, we consider a control of the form

$$g(t,x) = \beta_0(t)\bar{g}_0(x) + \beta_1(t)\bar{g}_1(x), \tag{1-16}$$

where the functions β_0, β_1 are chosen as square roots of sufficiently close smooth approximations of δ_0, δ_T (since it turns out that F_1 depends quadratically on α, and by consequence also on g, see (1-14)), and with some appropriate functions $\bar{g}_0, \bar{g}_1$.

Let us quickly explain how the controllability of the toy model (1-15) can be established. Given $q_0, q_1 \in \mathcal{Q}$, there exists (at least in the case when q_0 and q_1 are sufficiently close, hence the arguments of Section 3) a geodesic associated with the Riemannian metric induced on $\mathcal{Q}$ by $\mathcal{M}$, which connects q_0 with q_1. More precisely, there exists a unique smooth function $\bar{q}$ satisfying

$$\mathcal{M}(\bar{q})\bar{q}'' + \langle \Gamma(\bar{q}), \bar{q}', \bar{q}' \rangle = 0 \quad \text{on } [0,T], \quad \text{with } \bar{q}(0) = q_0,\ \bar{q}(T) = q_1. \tag{1-17}$$

So, one can arrive at the desired final position q_1, but a priori the final velocity $\bar{q}'(T)$ differs from q_1'; furthermore even the initial velocity $\bar{q}'(0)$ differs from q_0'.

Then, controlling the solution $\tilde{q}$ of (1-15) from (q_0, q_0') to (q_1, q_1') just amounts to setting $v_0 := \mathcal{M}(q_0)(\bar{q}'(0) - q_0')$ and $v_1 := -\mathcal{M}(q_1)(\bar{q}'(T) - q_1')$, which transforms the initial and final velocities $\tilde{q}'(0)$ and $\tilde{q}'(T)$ exactly to the desired velocities in order to achieve controllability.

In Section 6 we prove a proposition that is important for Theorem 14, namely that the whole system will behave like the toy model above, in a certain regime (and in particular for small γ). This relies on some appropriate estimations of the terms F_1, F_2 and some Gronwall-type arguments.

Section 7 explains how one can construct the control by means of complex analysis: it can be considered as the cornerstone of our control strategy. It is here that we construct the spacial parts $\bar{g}_0, \bar{g}_1$ of our control g from (1-16), as functions of v_0, v_1.

2. Reformulation of the solid's equation into an ODE

In this section we establish a reformulation of the Newton equations (1-6) as an ODE for the three degrees of freedom of the rigid body with coefficients obtained by solving some elliptic-type problems on a domain depending on the solid position. Indeed the fluid velocity can be recovered from the solid position and velocity by an elliptic-type problem, so that the fluid state may be seen as solving an auxiliary steady problem, where time only appears as a parameter, instead of the evolution equation (1-1). The Newton equations can therefore be rephrased as a second-order differential equation on the solid position whose coefficients are determined by the auxiliary fluid problem.

Such a reformulation in the case without boundary control was already achieved in [Glass et al. 2018] and we will start by recalling this case in Section 2A; see Proposition 12 below. A crucial fact in the analysis is that in the ODE reformulation the prefactor of the body's accelerations is the sum of the inertia of the solid and of the so-called "added inertia" which is a symmetric positive-semidefinite matrix depending only on the body's shape and position, and which encodes the amount of incompressible fluid that the rigid body has also to accelerate around itself. Remarkably enough in the case without control and where the circulation is 0 it turns out that the solid equations can be recast as a geodesic equation associated with the metric given by the total inertia.

Then we will extend this analysis to the case where there is a control on a part of the external boundary in Section 2B; see Theorem 6. In particular we will establish that the remote influence of the external boundary control translates into two additional force terms in the second-order ODE for the solid position; indeed we will distinguish one force term associated with the control velocity and another one associated with its time derivative.

To simplify notation, we define the positions and velocities $q = (h, \vartheta)$, $q' = (h', \vartheta')$, and

$$\mathcal{S}(q) = h + R(\vartheta)(\mathcal{S}_0 - h_0) \quad \text{and} \quad \mathcal{F}(q) = \Omega \setminus \mathcal{S}(q),$$

since the dependence in time of the domain occupied by the solid comes only from the position q. Furthermore, we set $q(t) = (h(t), \vartheta(t))$.

2A. ***A reminder of the uncontrolled case.*** We first recall that in the case without any control the fluid velocity satisfies (1-2), (1-3), (1-7) and (1-9). Therefore at each time t the fluid velocity u satisfies the div/curl system

$$\begin{cases} \operatorname{div} u = \operatorname{curl} u = 0 & \text{in } \mathcal{F}(q), \\ u \cdot n = 0 & \text{on } \partial\Omega, \\ u \cdot n = (h' + \vartheta'(x-h)^\perp) \cdot n & \text{on } \partial\mathcal{S}(q), \\ \int_{\partial\mathcal{S}(q)} u \cdot \tau \, d\sigma = \gamma, \end{cases} \tag{2-1}$$

where the dependence in time is only due to that of q and q'. Given the solid position q and the right-hand sides, the system (2-1) uniquely determines the fluid velocity u in the space of C^∞ vector fields on the closure of $\mathcal{F}(q)$. Moreover thanks to the linearity of the system with respect to its right-hand sides, its unique solution u can be uniquely decomposed with respect to the following functions which depend only on the solid position $q = (h, \vartheta)$ in $\mathcal{Q}$ and encode the contributions of elementary right-hand sides.

- The Kirchhoff potentials

$$\Phi = (\Phi_1, \Phi_2, \Phi_3)(q, \cdot) \tag{2-2}$$

are defined as the solution of the Neumann problems

$$\begin{gathered} \Delta\Phi_i(q,x) = 0 \quad \text{in } \mathcal{F}(q), \qquad \partial_n\Phi_i(q,x) = 0 \quad \text{on } \partial\Omega, \text{ for } i \in \{1,2,3\}, \\ \partial_n\Phi_i(q,x) = \begin{cases} n_i & \text{on } \partial\mathcal{S}(q), \text{ for } i \in \{1,2\}, \\ (x-h)^\perp \cdot n & \text{on } \partial\mathcal{S}(q), \text{ for } i = 3, \end{cases} \end{gathered} \tag{2-3}$$

where all differential operators are with respect to the variable x.

- The stream function ψ for the circulation term is defined in the following way. First we consider the solution $\tilde{\psi}(q, \cdot)$ of the Dirichlet problem $\Delta\tilde{\psi}(q,x) = 0$ in $\mathcal{F}(q)$, $\tilde{\psi}(q,x) = 0$ on $\partial\Omega$, and $\tilde{\psi}(q,x) = 1$ on $\partial\mathcal{S}(q)$. Then we set

$$\psi(q, \cdot) = -\left(\int_{\partial\mathcal{S}(q)} \partial_n\tilde{\psi}(q,x)\, d\sigma\right)^{-1} \tilde{\psi}(q, \cdot) \tag{2-4}$$

such that we have

$$\int_{\partial\mathcal{S}(q)} \partial_n \psi(q,x)\, d\sigma = -1,$$

noting that the strong maximum principle gives us $\partial_n \tilde{\psi}(q,x) < 0$ on $\partial\mathcal{S}(q)$.

Remark 4. The Kirchhoff potentials Φ and the stream function ψ are C^∞ as functions of q on $\mathcal{Q}$. We will use several times some properties of regularity with respect to the domain of solutions to linear elliptic problems, included for another potential $\mathcal{A}[q, g]$ associated with the control; see Definition 8 below. We will mention throughout the proof the properties which will be used and we refer to [Chambrion and Munnier 2012; Henrot and Pierre 2005; Lohéac and Munnier 2014] for more on this material, which is now standard in fluid-structure interaction.

The following statement is an immediate consequence of the definitions above.

Lemma 5. *For any $q = (h, \vartheta)$ in $\mathcal{Q}$, for any $p = (\ell, \omega)$ in $\mathbb{R}^2 \times \mathbb{R}$ and for any γ, the unique solution u in $C^\infty(\overline{\mathcal{F}(q)})$ to the system*

$$\begin{cases} \operatorname{div} u = \operatorname{curl} u = 0 & \text{in } \in \mathcal{F}(q), \\ u \cdot n = 0 & \text{on } \partial\Omega, \\ u \cdot n = (\ell + \omega(x-h)^\perp) \cdot n & \text{on } \partial\mathcal{S}(q), \\ \int_{\partial\mathcal{S}(q)} u \cdot \tau \, d\sigma = \gamma \end{cases} \tag{2-5}$$

is given by the formula, for x in $\overline{\mathcal{F}(q)}$,

$$u(x) = \nabla(p \cdot \Phi(q,x)) + \gamma \nabla^\perp \psi(q,x). \tag{2-6}$$

Above $p \cdot \Phi(q,x)$ denotes the inner product

$$p \cdot \Phi(q,x) = \sum_{i=1}^{3} p_i \Phi_i(q,x).$$

Let us now address the solid dynamics. The solid motion is driven by the Newton equations (1-6) where the influence of the fluid on the solid appears through the fluid pressure. The pressure can in turn be related to the fluid velocity thanks to the Euler equations (1-1). The contributions to the solid dynamics of the two terms in the right-hand side of the fluid velocity decomposition formula (2-6) are very different. On the one hand the potential part, i.e., the first term in the right-hand side of (2-6), contributes as an added inertia matrix, together with a connection term which ensures a geodesic structure [Munnier 2009], whereas on the other hand the contribution of the term due to the circulation, i.e., the second term in the right-hand side of (2-6), turns out to be a force which reminds us of the Lorentz force in electromagnetism by its structure [Glass et al. 2018]. We therefore introduce the following notation.

- We respectively define the genuine and added mass 3×3 matrices by

$$\mathcal{M}_g = \begin{pmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & \mathcal{J} \end{pmatrix},$$

and, for $q \in \mathcal{Q}$,

$$\mathcal{M}_a(q) = \left(\int_{\mathcal{F}(q)} \nabla \Phi_i(q,x) \cdot \nabla \Phi_j(q,x)\, dx \right)_{1 \leqslant i,j \leqslant 3}.$$

Note that $\mathcal{M}_a$ is a symmetric Gram matrix and is C^∞ on $\mathcal{Q}$.

- We define the symmetric bilinear map $\Gamma(q)$ given by

$$\langle \Gamma(q), p, p \rangle = \left(\sum_{1 \le i,j \le 3} \Gamma^k_{i,j}(q) p_i p_j \right)_{1 \le k \le 3} \in \mathbb{R}^3 \quad \text{for all } p \in \mathbb{R}^3,$$

where, for each $i, j, k \in \{1,2,3\}$, $\Gamma^k_{i,j}$ denotes the Christoffel symbol of the first kind defined on $\mathcal{Q}$ by

$$\Gamma^k_{i,j} = \frac{1}{2}\left(\frac{\partial (\mathcal{M}_a)_{k,j}}{\partial q_i} + \frac{\partial (\mathcal{M}_a)_{k,i}}{\partial q_j} - \frac{\partial (\mathcal{M}_a)_{i,j}}{\partial q_k} \right). \tag{2-7}$$

It can be checked that Γ is of class C^∞ on $\mathcal{Q}$.

- We introduce the following C^∞ vector fields on $\mathcal{Q}$ with values in $\mathbb{R}^3$:

$$E = -\frac{1}{2} \int_{\partial \mathcal{S}(q)} |\partial_n \psi(q,\cdot)|^2 \partial_n \Phi(q,\cdot)\, d\sigma, \tag{2-8}$$

$$B = \int_{\partial \mathcal{S}(q)} \partial_n \psi(q,\cdot)(\partial_n \Phi(q,\cdot) \times \partial_\tau \Phi(q,\cdot))\, d\sigma. \tag{2-9}$$

We recall that the notation Φ was given in (2-2).

The reformulation of the model as an ODE is given in the following result, which was first established in [Munnier 2009] in the case $\gamma = 0$ and in [Glass et al. 2018] in the case $\gamma \in \mathbb{R}$.

Theorem 6. *Given $q = (h, \vartheta) \in C^\infty([0,T]; \mathcal{Q})$, $u \in C^\infty(\mathcal{D}_T; \mathbb{R}^2)$ we have that (q,u) is a solution to* (1-1), (1-2), (1-3), (1-6), (1-7) *and* (1-9) *if and only if q satisfies the ODE on* $[0,T]$

$$(\mathcal{M}_g + \mathcal{M}_a(q))q'' + \langle \Gamma(q), q', q' \rangle = \gamma^2 E(q) + \gamma q' \times B(q), \tag{2-10}$$

and u is the unique solution to the system (2-1). *Moreover the total kinetic energy $\frac{1}{2}(\mathcal{M}_g + \mathcal{M}_a(q))q' \cdot q'$ is conserved in time for smooth solutions of* (2-10), *at least as long as there is no collision.*

Note that in the case where $\gamma = 0$, the ODE (2-10) means that the particle q is moving along the geodesics associated with the Riemannian metric induced on $\mathcal{Q}$ by the matrix field $\mathcal{M}_g + \mathcal{M}_a(q)$. Note that, since $\mathcal{Q}$ is a manifold with boundary and the metric $\mathcal{M}_g + \mathcal{M}_a(q)$ may become singular at the boundary of $\mathcal{Q}$, the Hopf–Rinow theorem does not apply and geodesics may not be global. However we will make use only of local geodesics.

Remark 7. Let us also mention that the whole "inviscid fluid + rigid body" system can be reinterpreted as a geodesic flow on an infinite-dimensional manifold; see [Glass and Sueur 2012]. However the reformulation established by Theorem 6 relies on the finite-dimensional manifold $\mathcal{Q}$ and sheds more light on the dynamics of the rigid body.

Below we provide a sketch of the proof of Theorem 6; this will be useful in Section 2B when extending the analysis to the controlled case.

Proof. Let us focus on the direct part of the proof for sake of clarity but all the subsequent arguments can be arranged in order to ensure the converse part of the statement as well. Using Green's first identity and the properties of the Kirchhoff functions, the Newton equations (1-6) can be rewritten as

$$\mathcal{M}_g q'' = \int_{\mathcal{F}(q)} \nabla\pi \cdot \nabla\Phi(q,x)\,dx. \tag{2-11}$$

Moreover when u is irrotational, (1-1) can be rephrased as

$$\nabla\pi = -\partial_t u - \tfrac{1}{2}\nabla_x |u|^2 \quad \text{for } x \text{ in } \mathcal{F}(q(t)), \tag{2-12}$$

and Lemma 5 shows that for any t in $[0, T]$

$$u(t,\cdot) = \nabla(q'(t)\cdot\Phi(q(t),\cdot)) + \gamma\nabla^{\perp}\psi(q(t),\cdot). \tag{2-13}$$

Substituting (2-13) into (2-12) and then the resulting decomposition of $\nabla\pi$ into (2-11) we get

$$\begin{aligned}\mathcal{M}_g q'' = &-\int_{\mathcal{F}(q)} \left(\partial_t \nabla(q'\cdot\Phi(q,x)) + \frac{\nabla|\nabla(q'\cdot\Phi(q,x))|^2}{2}\right)\cdot\nabla\Phi(q,x)\,dx \\ &-\gamma\int_{\mathcal{F}(q)} \big(\partial_t\nabla^{\perp}\psi(q,x) + \nabla(\nabla(q'\cdot\Phi(q,x))\cdot\nabla^{\perp}\psi(q,x))\big)\cdot\nabla\Phi(q,x)\,dx \\ &-\gamma^2\int_{\mathcal{F}(q)} \frac{\nabla|\nabla\psi(q,x)|^2}{2}\cdot\nabla\Phi(q,x)\,dx.\end{aligned} \tag{2-14}$$

According to Lemmas 32, 33 and 34 in [Glass et al. 2018], the terms in the three lines of the right-hand side above are respectively equal to $-\mathcal{M}_a(q)q'' - \langle\Gamma(q), q', q'\rangle$, $\gamma q' \times B(q)$ and $\gamma^2 E(q)$, so that we easily deduce the ODE (2-10) from (2-14).

The conservation of the kinetic energy $\frac{1}{2}(\mathcal{M}_g + \mathcal{M}_a(q))q'\cdot q'$ is then simply obtained by multiplying the ODE (2-10) by q' and observing that

$$((\mathcal{M}_g + \mathcal{M}_a(q))q'' + \langle\Gamma(q), q', q'\rangle)\cdot q' = \big(\tfrac{1}{2}(\mathcal{M}_g + \mathcal{M}_a(q))q'\cdot q'\big)', \tag{2-15}$$

completing the proof. □

2B. ***Extension to the controlled case.*** We now tackle the case where a control is imposed on the part Σ of the external boundary $\partial\Omega$. At any time this control has to be compatible with the incompressibility of the fluid, meaning that the flux through Σ has to be zero. We therefore introduce the set

$$\mathcal{C} := \left\{ g \in C_0^{\infty}(\Sigma; \mathbb{R}) : \int_{\Sigma} g\,d\sigma = 0 \right\}.$$

The decomposition of the fluid velocity u then involves a new potential term involving the following function.

Definition 8. For any $q \in \mathcal{Q}$ and $g \in \mathcal{C}$ we consider the unique solution $\bar{\alpha} := \mathcal{A}[q, g] \in C^\infty(\overline{\mathcal{F}(q)}; \mathbb{R})$ to the Neumann problem

$$\Delta\bar{\alpha} = 0 \quad \text{in } \mathcal{F}(q) \qquad \text{and} \qquad \partial_n \bar{\alpha} = g \mathbb{1}_\Sigma \quad \text{on } \partial\mathcal{F}(q), \tag{2-16}$$

with zero mean on $\mathcal{F}(q)$.

Let us mention that the zero-mean condition above allows us to determine a unique solution to the Neumann problem but plays no role in the sequel.

Now Lemma 5 can be modified as follows.

Lemma 9. *For any $q = (h, \vartheta)$ in $\mathcal{Q}$, for any $p = (\ell, \omega)$ in $\mathbb{R}^2 \times \mathbb{R}$, for any $\bar{g}$ in $\mathcal{C}$, the unique solution u in $C^\infty(\overline{\mathcal{F}(q)})$ to*

$$\begin{cases} \operatorname{div} u = \operatorname{curl} u = 0 & \text{in } \mathcal{F}(q), \\ u \cdot n = \mathbb{1}_\Sigma \bar{g} & \text{on } \partial\Omega, \\ u \cdot n = (\ell + \omega(x - h)^\perp) \cdot n & \text{on } \partial\mathcal{S}(q), \\ \int_{\partial\mathcal{S}(q)} u \cdot \tau \, d\sigma = \gamma \end{cases}$$

is given by

$$u = \nabla(p \cdot \Phi(q, \cdot)) + \gamma \nabla^\perp \psi(q, \cdot) + \nabla \mathcal{A}[q, \bar{g}]. \tag{2-17}$$

Let us avoid possible confusion by mentioning that the ∇ operator above has to be considered with respect to the space variable x. The function $\mathcal{A}[q, \bar{g}]$ and its time derivative will respectively be involved in the arguments of the following force terms.

Definition 10. We define, for any q in $\mathcal{Q}$, p in $\mathbb{R}^3$, α in $C^\infty(\overline{\mathcal{F}(q)}; \mathbb{R})$ and γ in $\mathbb{R}$, $F_1(q, p, \gamma)[\alpha]$ and $F_2(q)[\alpha]$ in $\mathbb{R}^3$ by

$$F_1(q, p, \gamma)[\alpha] := -\frac{1}{2} \int_{\partial\mathcal{S}(q)} |\nabla\alpha|^2 \, \partial_n \Phi(q, \cdot) \, d\sigma \\ - \int_{\partial\mathcal{S}(q)} \nabla\alpha \cdot (\nabla(p \cdot \Phi(q, \cdot)) + \gamma \nabla^\perp \psi(q, \cdot)) \partial_n \Phi(q, \cdot) \, d\sigma, \tag{2-18}$$

$$F_2(q)[\alpha] := -\int_{\partial\mathcal{S}(q)} \alpha \, \partial_n \Phi(q, \cdot) \, d\sigma. \tag{2-19}$$

Observe that (2-18) and (2-19) only require α and $\nabla\alpha$ to be defined on $\partial\mathcal{S}(q)$. Moreover when these formulas are applied to $\alpha = \mathcal{A}[q, g]$ for some g in $\mathcal{C}$, only the trace of α and the tangential derivative $\partial_\tau \alpha$ on $\partial\mathcal{S}(q)$ are involved, since the normal derivative of α vanishes on $\partial\mathcal{S}(q)$ by definition; see (2-16).

We define our notion of controlled solution of the "fluid + solid" system as follows.

Definition 11. We say that (q, g) in $C^\infty([0, T]; \mathcal{Q}) \times C_0^\infty([0, T]; \mathcal{C})$ is a controlled solution if the following ODE holds true on $[0, T]$:

$$(\mathcal{M}_g + \mathcal{M}_a(q))q'' + \langle \Gamma(q), q', q' \rangle = \gamma^2 E(q) + \gamma q' \times B(q) + F_1(q, q', \gamma)[\alpha] + F_2(q)[\partial_t \alpha], \tag{2-20}$$

where $\alpha(t, \cdot) := \mathcal{A}[q(t), g(t, \cdot)]$.

We have the following result for reformulating the model as an ODE.

Proposition 12. *Given*

$$q \in C^\infty([0,T];\mathcal{Q}), \quad u \in C^\infty(\mathcal{D}_T;\mathbb{R}^2) \quad \textit{and} \quad g \in C_0^\infty([0,T];\mathcal{C}),$$

we have that (q,u) *is a solution to* (1-1), (1-2), (1-6), (1-7), (1-8), (1-9), (1-10), (1-11), (1-12) *if and only if* (q,g) *is a controlled solution and* u *is the unique solution to the unique div/curl-type problem*

$$\begin{cases} \operatorname{div} u = \operatorname{curl} u = 0 & \textit{in } \mathcal{F}(q), \\ u\cdot n = \mathbb{1}_\Sigma g & \textit{on } \partial\Omega, \\ u\cdot n = (h' + \vartheta'(x-h)^\perp)\cdot n & \textit{on } \partial\mathcal{S}(q), \\ \int_{\partial\mathcal{S}(q)} u\cdot\tau\, d\sigma = \gamma, \end{cases}$$

with $q = (h,\vartheta)$.

Proposition 12 therefore extends Theorem 6 to the case with an external boundary control (in particular one recovers Theorem 6 in the case where g is identically vanishing).

Proof. We proceed as in the proof of Theorem 6 recalled above, with some modifications due to the extra term involved in the decomposition of the fluid velocity; compare (2-6) and (2-17). In particular some extra terms appear in the right-hand side of (2-14) after substituting the right-hand side of (2-17) for u in (2-12). Using some integration by parts and the properties of the Kirchhoff functions we obtain integrals on $\partial\mathcal{S}(q)$ whose sum precisely gives $F_1(q,q',\gamma)[\alpha(t,\cdot)] + F_2(q)[\partial_t\alpha(t,\cdot)]$. This allows us to conclude. ☐

3. Reduction to the case where the displacement, velocities and circulation are small

For $\delta > 0$, we introduce the set

$$\mathcal{Q}_\delta = \{q \in \Omega\times\mathbb{R} : d(\mathcal{S}(q),\partial\Omega) > \delta\}. \tag{3-1}$$

The goal of this section is to prove that Theorem 1 can be deduced from the following result. The balls have to be understood for the Euclidean norm (rather than for the metric $\mathcal{M}_g + \mathcal{M}_a(q)$).

Theorem 13. *Given* $\delta > 0$, $\mathcal{S}_0 \subset \Omega$ *bounded, closed, simply connected with smooth boundary, which is not a disk,* q_0 *in* $\mathcal{Q}_\delta$ *and* $T > 0$, *there exists* $r > 0$ *such that for any* q_1 *in* $B(q_0,r)$, *for any* $\gamma \in \mathbb{R}$ *with* $|\gamma| \le r$ *and for any* $q_0', q_1' \in B(0,r)$, *there is a controlled solution* (q,g) *in* $C^\infty([0,T];\mathcal{Q}_\delta)\times C_0^\infty([0,T]\times\Sigma)$ *such that* $(q,q')(0) = (q_0,q_0')$ *and* $(q,q')(T) = (q_1,q_1')$.

Note in particular that for $r > 0$ small enough, $B(q_0,r)$ is included in the connected component of $\mathcal{Q}_\delta$ containing q_0.

Proof of Theorem 1 from Theorem 13. We proceed in two steps: first we use a time-rescaling argument in order to deduce from Theorem 13 a more general result covering the case where the initial and final velocities q_0' and q_1' and the circulation γ are large. This argument is reminiscent of a time-rescaling argument used by J.-M. Coron [1993] for the Euler equation, which has been also used in [Glass and Rosier 2013] in order to pass from the potential case to the case with vorticity. Then we use a compactness argument in order to deal with the case where q_0 and q_1 are remote (but of course in the same connected component of $\mathcal{Q}_\delta$).

The time-rescaling argument relies on the following observation: it follows from (2-20) that (q, g) is a controlled solution on $[0, T]$ with circulation γ if and only if (q^λ, g^λ) is a controlled solution on $[0, \lambda T]$ with circulation γ/λ, where (q^λ, g^λ) is defined by

$$q^\lambda(t) := q\left(\frac{t}{\lambda}\right) \quad \text{and} \quad g^\lambda(t, x) := \frac{1}{\lambda} g\left(\frac{t}{\lambda}, x\right). \tag{3-2}$$

Of course the initial and final conditions

$$(q, q')(0) = (q_0, q_0') \quad \text{and} \quad (q, q')(T) = (q_1, q_1')$$

translate respectively into

$$(q^\lambda, (q^\lambda)')(0) = \left(q_0, \frac{q_0'}{\lambda}\right) \quad \text{and} \quad (q^\lambda, (q^\lambda)')(\lambda T) = \left(q_1, \frac{q_1'}{\lambda}\right). \tag{3-3}$$

Now consider q_0 in $\mathcal{Q}_\delta$ and q_1 in $\bar{B}(q_0, r)$ in the same connected component of $\mathcal{Q}_\delta$ as q_0, with $r > 0$ as in Theorem 13, and q_0', q_1' and γ without size constraint. For λ small enough, $(q_0, \lambda q_0')$, $(q_1, \lambda q_1')$ and $\lambda\gamma$ satisfy the assumptions of Theorem 13. Hence there exists a controlled solution (q, g) on $[0, T]$, achieving $(q, q')(0) = (q_0, \lambda q_0')$ and $(q, q')(T) = (q_1, \lambda q_1')$. On the other hand, the corresponding trajectory q^λ constructed above will satisfy the conclusions of Theorem 1 on $[0, \lambda T]$, in particular that $(q^\lambda, (q^\lambda)')(0) = (q_0, q_0')$ and $(q^\lambda, (q^\lambda)')(\lambda T) = (q_1, q_1')$. Moreover we can assume that it is the case without loss of generality that λ is small, and in particular that $\lambda \leq 1$. Thus the result is obtained but in a shorter time interval.

To get to the desired time interval, using that (2-20) enjoys some invariance properties by translation and time-reversal (up to the change of the sign of γ) it is sufficient to glue together an odd number, say $2N + 1$ with N in $\mathbb{N}^*$, of appropriate controlled solutions each defined on a time interval of length λT with $\lambda = 1/(2N + 1)$, going back and forth between (q_0, q_0') and (q_1, q_1') until time $T = (2N + 1)\lambda T$. Moreover one can see that the gluings are not only C^2 but even C^∞.

We have therefore already proven that Theorem 1 is true in the case where q_1 is close to q_0, or more precisely for any q_0 in $\mathcal{Q}_\delta$ and q_1 in $\bar{B}(q_0, r_{q_0})$.

For the general case where q_0 and q_1 are in the same connected component of $\mathcal{Q}_\delta$ for some $\delta > 0$, without the closeness condition, we use again a gluing process. Consider indeed a smooth curve from q_0 to q_1. For each point q on this curve, there is an $r_q > 0$ such that for any $\tilde{q}$ in $B(q, r_q)$, any q_0', q_1' and any γ, one can connect (q, q_0') to $(\tilde{q}, q_1')$ by a solution of the system for any time $T > 0$. Extract a finite subcover of the curve by the balls $B(q, r_q)$. Therefore we find $N \geq 2$ and $(q_{i/N})_{i=1,\dots,N-1}$ in the same connected component of $\mathcal{Q}_\delta$ as q_0 such that for any $i = 1, \dots, N$, one has that $q_{i/N}$ is in $\bar{B}(q_{(i-1)/N}, r_{q_{(i-1)/N}})$ (note that this includes q_0 and q_1). Therefore, using again the local result obtained above, there exist some controlled solutions from $(q_{(i-1)/N}, 0)$ to $(q_{i/N}, 0)$ (for $i = 1$ and $i = N$ we use $(q_0, q_0'/N)$ and $(q_1, q_1'/N)$ rather than $(q_0, 0)$ and $(q_1, 0)$), each on a time interval of length T associated with circulation γ/N. One deduces by time-rescaling some controlled solutions associated with circulation γ on a time interval of length T/N. Gluing them together leads to the desired controlled solution. □

4. Reduction to an approximate controllability result

The goal of this section is to prove that Theorem 13 can be deduced from the following approximate controllability result thanks to a topological argument already used in [Glass and Rosier 2013]; see Lemma 15 below. Let us mention that a similar argument has also been used for control purposes but in other contexts; see, e.g., [Aronsson 1973; Brunovský and Lobry 1975; Grasse 1981; 1982].

Theorem 14. *Given $\delta > 0$, $\mathcal{S}_0 \subset \Omega$ bounded, closed, simply connected with smooth boundary, which is not a disk, q_0 in $\mathcal{Q}_\delta$ and $T > 0$, there is $\tilde{r} > 0$ such that $B(q_0, \tilde{r})$ is included in the same connected component of $\mathcal{Q}_\delta$ as q_0, and furthermore, for any $\eta > 0$, there exists $r' = r'(\eta) > 0$ such that, for any $\gamma \in \mathbb{R}$ with $|\gamma| \le r'$ and for any q'_0 in $\bar{B}(0, \tilde{r})$, there is a mapping*

$$\mathcal{T} : \bar{B}((q_0, q'_0), \tilde{r}) \to C^\infty([0, T]; \mathcal{Q}_\delta),$$

which with (q_1, q'_1) associates q where (q, g) is a controlled solution associated with the initial data (q_0, q'_0), such that the mapping

$$(q_1, q'_1) \in \bar{B}((q_0, q'_0), \tilde{r}) \mapsto (\mathcal{T}(q_1, q'_1), \mathcal{T}(q_1, q'_1)')(T) \in \mathcal{Q}_\delta \times \mathbb{R}^3$$

is continuous and such that for any (q_1, q'_1) in $\bar{B}((q_0, q'_0), \tilde{r})$

$$\|(\mathcal{T}(q_1, q'_1), \mathcal{T}(q_1, q'_1)')(T) - (q_1, q'_1)\| \leqslant \eta.$$

The proof of Theorem 14 will be given in Section 5. Here we prove that Theorem 13 follows from Theorem 14.

Proof of Theorem 13 from Theorem 14. Let $\delta > 0$, $\mathcal{S}_0 \subset \Omega$ bounded, closed, simply connected with smooth boundary, which is not a disk, q_0 in $\mathcal{Q}_\delta$ and $T > 0$. Let $\tilde{r} > 0$ as in Theorem 14 and $\eta = \tilde{r}/2$. We deduce that for any $\gamma \in \mathbb{R}$ with $|\gamma| \le r' = r'(\tilde{r}/2)$ and q'_0 in $\bar{B}(0, \tilde{r})$, there is a mapping

$$\mathcal{T} : \bar{B}((q_0, q'_0), \tilde{r}) \to C^\infty([0, T]; \mathcal{Q}_\delta)$$

which maps (q_1, q'_1) to q where (q, g) is a controlled solution associated with the initial data (q_0, q'_0) such that for any (q_1, q'_1) in $\bar{B}((q_0, q'_0), \tilde{r})$,

$$\|(\mathcal{T}(q_1, q'_1), \mathcal{T}(q_1, q'_1)')(T) - (q_1, q'_1)\| \leqslant \frac{\tilde{r}}{2}.$$

We define a mapping f from $\bar{B}((q_0, q'_0), \tilde{r})$ to $\mathbb{R}^6$ which maps (q_1, q'_1) to

$$f(q_1, q'_1) := (\mathcal{T}(q_1, q'_1), \mathcal{T}(q_1, q'_1)')(T).$$

Then we apply the following lemma borrowed from [Glass and Rosier 2013, pp. 32–33] to $w_0 = (q_0, q'_0)$ and $\kappa = \tilde{r}$.

Lemma 15. *Let $w_0 \in \mathbb{R}^n$, $\kappa > 0$, $f : \bar{B}(w_0, \kappa) \to \mathbb{R}^n$ a continuous map such that we have $|f(w) - w| \leq \kappa/2$ for any x in $\partial B(w_0, \kappa)$. Then $B(w_0, \kappa/2) \subset f(\bar{B}(w_0, \kappa))$.*

This allows us to conclude the proof of Theorem 13 by setting

$$r = \min\left\{ \frac{\tilde{r}}{2\sqrt{5}}, r'\left(\frac{\tilde{r}}{2}\right) \right\},$$

since the conditions $q_1 \in B(q_0, r)$, $|\gamma| \leq r$ and $q_0', q_1' \in B(0, r)$ imply $|\gamma| \leq r'(\tilde{r}/2)$ and $(q_1, q_1') \in B((q_0, q_0'), \tilde{r}/2)$. ☐

5. Proof of the approximate controllability result Theorem 14

In this section we prove Theorem 14 by exploiting the geodesic feature of the uncontrolled system with zero circulation; see the observation below Theorem 6. To do so, we will use some well-chosen impulsive controls which allow us to modify the velocity q' in a short time interval and put the state of the system on a prescribed geodesic (and use that $|\gamma|$ is small). We mention here [Bressan 1996] for many more examples on the impulsive control strategy.

5A. ***First step.*** We consider $\mathcal{S}_0 \subset \Omega$ as before and consider $\delta > 0$ so that $q_0 \in \mathcal{Q}_\delta$. We let $r_1 > 0$ be small enough so that $B(q_0, r_1) \subset \mathcal{Q}_\delta$. We also let $T > 0$.

The first step consists in considering the geodesics associated with the uncontrolled, potential case ($\gamma = 0$). The following classical result regarding the existence of geodesics can be found for instance in [Marsden and Ratiu 1994, Section 7.5]; see also [Gaines 1969] for the continuity feature.

Lemma 16. *There exists r_2 in $(0, r_1/2)$ such that for any q_1 in $\bar{B}(q_0, r_2)$ there exists a unique C^∞ solution $\bar{q}(t)$ lying in $B(q_0, r_1/2)$ to*

$$(\mathcal{M}_g + \mathcal{M}_a(\bar{q}))\bar{q}'' + \langle \Gamma(\bar{q}), \bar{q}', \bar{q}' \rangle = 0 \quad \text{on } [0, T], \quad \text{with } \bar{q}(0) = q_0,\ \bar{q}(T) = q_1. \tag{5-1}$$

Furthermore the map $q_1 \in \bar{B}(q_0, r_2) \mapsto (c_0, c_1) \in \mathbb{R}^6$ given by $c_0 = \bar{q}'(0)$, $c_1 = \bar{q}'(T)$ is continuous.

Let us fix r_2 as in the lemma before. Let q_0' in $\bar{B}(0, r_2)$ and (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$.

Our goal is to make the system follow approximately such a geodesic $\bar{q}$, which we consider fixed during this section. For the geodesic equation in (5-1), q_0 and q_1 determine the initial and final velocities (which of course differ in general from q_0' and q_1'). But we will see that it is possible to use the penultimate term of (2-20) in order to modify the initial and final velocities of the system. Precisely, the control will be used so that the right-hand side of (2-20) behaves like two Dirac masses at times close to 0 and T, driving the velocity q' from the initial and final velocities to the ones of the geodesic in two short time intervals close to 0 and T.

5B. ***Illustration of the method on a toy model.*** Let us illustrate this strategy on a toy model. We will later on adapt the analysis to the complete model; see Proposition 21.

Let $\beta : \mathbb{R} \to \mathbb{R}$ be a smooth, nonnegative function supported in $[-1, 1]$ such that $\int_{-1}^{1} \beta(t)^2 \, dt = 1$ and, for ε in $(0, 1)$,

$$\beta_\varepsilon(t) := \frac{1}{\sqrt{\varepsilon}} \beta\left(\frac{t-\varepsilon}{\varepsilon}\right),$$

so that[2] $(\beta_\varepsilon^2)_\varepsilon$ is an approximation of the unity when $\varepsilon \to 0^+$.

For a function f defined on $[0, T]$, we will define

$$\|f\|_{T,\varepsilon} := \|f\|_{C^0([0,T])} + \|f\|_{C^1([2\varepsilon, T-2\varepsilon])}. \tag{5-2}$$

Lemma 17. *Let q_0, r_2, q_1, q_0' and q_1' as above. Let*

$$v_0 := (\mathcal{M}_g + \mathcal{M}_a(q_0))(c_0(q_1) - q_0') \quad \text{and} \quad v_1 := -(\mathcal{M}_g + \mathcal{M}_a(q_1))(c_1(q_1) - q_1'). \tag{5-3}$$

Let, for ε in $(0, 1)$, q_ε be the maximal solution to the Cauchy problem

$$(\mathcal{M}_g + \mathcal{M}_a(q_\varepsilon))q_\varepsilon'' + \langle \Gamma(q_\varepsilon), q_\varepsilon', q_\varepsilon' \rangle = \beta_\varepsilon^2(\,\cdot\,)v_0 + \beta_\varepsilon^2(T - \,\cdot\,)v_1, \tag{5-4}$$

with $q_\varepsilon(0) = q_0$ and $q_\varepsilon'(0) = q_0'$. Then for ε small enough, $q_\varepsilon(t)$ lies in $B(q_0, r_1)$ for t in $[0, T]$ and, as $\varepsilon \to 0^+$, $\|q_\varepsilon - \bar{q}\|_{T,\varepsilon} \to 0$ and $(q_\varepsilon, q_\varepsilon')(T) \to (q_1, q_1')$.

Proof. For ε in $(0, 1)$, let us define $T_\varepsilon = \sup\{\hat{T} > 0 : q_\varepsilon(t) \in B(q_0, r_1) \text{ for } t \in (0, \hat{T})\}$. Let us first prove that there exists $\widetilde{T} > 0$ such that for any ε in $(0, 1)$ we have $T_\varepsilon \geq \widetilde{T}$. Using the identity (2-15), we obtain indeed, for any ε in $(0, 1)$, for any $t \in (0, T_\varepsilon)$

$$(\mathcal{M}_g + \mathcal{M}_a(q_\varepsilon(t)))q_\varepsilon'(t) \cdot q_\varepsilon'(t) = (\mathcal{M}_g + \mathcal{M}_a(q_0))q_0' \cdot q_0' + 2\int_0^t (\beta_\varepsilon^2(\,\cdot\,)v_0 + \beta_\varepsilon^2(T - \,\cdot\,)v_1) \cdot q_\varepsilon'.$$

Moreover, relying on Remark 4, we see that there exists $c > 0$ (which depends on δ) such that for any q in $\mathcal{Q}_\delta$, for any p in $\mathbb{R}^3$

$$c|p|^2 \leq (\mathcal{M}_g + \mathcal{M}_a(q))p \cdot p \leq c^{-1}|p|^2. \tag{5-5}$$

Therefore using Gronwall's lemma we obtain that there exists $C > 0$ such that for any ε in $(0, 1)$, for any $t \in (0, T_\varepsilon)$ we have $\sup_{t \in (0, T_\varepsilon)} \|q_\varepsilon'(t)\| \leq C$. Therefore by the mean value theorem for $\widetilde{T} := r_1/2C$, for any ε in $(0, 1)$ one has $T_\varepsilon \geq \widetilde{T}$.

We now prove in the same time that for $\varepsilon > 0$ small enough $T_\varepsilon \geq T$, and the convergence results stated in Lemma 17. In order to exploit the supports of the functions $\beta_\varepsilon(\,\cdot\,)$ and $\beta_\varepsilon(T - \,\cdot\,)$ in the right-hand side of (5-4) we compare the dynamics of q_ε and $\bar{q}$ during the three time intervals $[0, 2\varepsilon]$, $[2\varepsilon, T - 2\varepsilon]$ and $[T - 2\varepsilon, T]$.

For $\varepsilon_1 := \widetilde{T}/2$ and ε in $(0, \varepsilon_1)$, one already has that $T_\varepsilon \geq 2\varepsilon$ and we can therefore simply compare the dynamics of q_ε and $\bar{q}$ on the first interval $[0, 2\varepsilon]$. Indeed using again the mean value theorem we obtain

[2]In the next lemma we are going to make use only of the square function β_ε^2 but we will also have to deal with the function β_ε itself in the sequel; see below Proposition 19.

that $\sup_{t\in[0,2\varepsilon]} |q_\varepsilon - q_0|$ converges to 0 as ε goes to 0. Moreover integrating (5-4) on $[0, 2\varepsilon]$ and taking into account the choice of v_0 in (5-3), we obtain

$$(\mathcal{M}_g + \mathcal{M}_a(q_\varepsilon(2\varepsilon)))q_\varepsilon'(2\varepsilon) \\ = (\mathcal{M}_g + \mathcal{M}_a(q_0))c_0(q_1) - \int_0^{2\varepsilon} (D\mathcal{M}_a(q_\varepsilon)\cdot q_\varepsilon')\cdot q_\varepsilon'\,dt - \int_0^{2\varepsilon} \langle \Gamma(q_\varepsilon), q_\varepsilon', q_\varepsilon'\rangle\,dt. \quad (5\text{-}6)$$

Now, there exists $C > 0$ such that for any q in $\mathcal{Q}_\delta$, for any p in $\mathbb{R}^3$

$$|(D\mathcal{M}_a(q)\cdot p)\cdot p| + |\langle \Gamma(q), p, p\rangle| \le C|p|^2. \quad (5\text{-}7)$$

Combining this and the bound on q_ε' we see that the two terms of the last line of (5-6) above converge to 0 as ε goes to 0. Since $q \mapsto \mathcal{M}_a(q)$ is continuous on $\mathcal{Q}_\delta$ and $q_\varepsilon(2\varepsilon)$ converges to q_0 as $\varepsilon \to 0$, the matrix $\mathcal{M}_a(q_\varepsilon)$ converges to $\mathcal{M}_a(q_0)$ as $\varepsilon \to 0$. Therefore, using that the matrix $\mathcal{M}_g + \mathcal{M}_a(q_0)$ is invertible we deduce that $q_\varepsilon'(2\varepsilon)$ converges to $c_0(q_1)$ as ε goes to 0.

During the time interval $[2\varepsilon, T - 2\varepsilon]$, the right-hand side of (5-4) vanishes and the equation therefore reduces to the geodesic equation in (5-1). Since this equation is invariant by translation in time, one may use the following elementary result on the continuous dependence on the data, with a time shift of 2ε.

Lemma 18. *There exists $\eta > 0$ such that for any $(\tilde{q}_0, \tilde{q}_0')$ in $B((q_0, c_0(q_1)), \eta)$ there exists a unique C^∞ solution $\tilde{q}(t)$ lying in $B(q_0, r_1)$ to*

$$(\mathcal{M}_g + \mathcal{M}_a(\tilde{q}))\tilde{q}'' + \langle \Gamma(\tilde{q}), \tilde{q}', \tilde{q}'\rangle = 0 \quad \text{on } [0, T], \quad \text{with } \tilde{q}(0) = \tilde{q}_0,\ \tilde{q}'(0) = \tilde{q}_0'.$$

Furthermore $\|\tilde{q} - \bar{q}\|_{C^1([0,T])} \to 0$ as $(\tilde{q}_0, \tilde{q}_0') \to (q_0, c_0(q_1))$.

Since $q_\varepsilon(2\varepsilon)$ and $q_\varepsilon'(2\varepsilon)$ respectively converge to q_0 and $c_0(q_1)$, according to Lemma 18 there exists ε_2 in $(0, \varepsilon_1)$ such that for ε in $(0, \varepsilon_2)$, there exists a unique C^∞ solution $\tilde{q}_\varepsilon(t)$ lying in $B(q_0, r_1)$ to

$$(\mathcal{M}_g + \mathcal{M}_a(\tilde{q}_\varepsilon))\tilde{q}_\varepsilon'' + \langle \Gamma(\tilde{q}_\varepsilon), \tilde{q}_\varepsilon', \tilde{q}_\varepsilon'\rangle = 0 \quad \text{on } [0, T],$$

with

$$\tilde{q}_\varepsilon(0) = q_\varepsilon(2\varepsilon), \quad \tilde{q}_\varepsilon'(0) = q_\varepsilon'(2\varepsilon) \quad \text{and} \quad \|\tilde{q}_\varepsilon - \bar{q}\|_{C^1([0,T])} \to 0 \quad \text{as } \varepsilon \to 0.$$

Since the function defined by $\hat{q}_\varepsilon(t) = q_\varepsilon(t + 2\varepsilon)$ also satisfies

$$(\mathcal{M}_g + \mathcal{M}_a(\hat{q}_\varepsilon))\hat{q}_\varepsilon'' + \langle \Gamma(\hat{q}_\varepsilon), \hat{q}_\varepsilon', \hat{q}_\varepsilon'\rangle = 0 \quad \text{on } [0, T - 4\varepsilon], \quad \text{with } \hat{q}_\varepsilon(0) = q_\varepsilon(2\varepsilon),\ \hat{q}_\varepsilon'(0) = q_\varepsilon'(2\varepsilon),$$

by the uniqueness part in the Cauchy–Lipschitz theorem one has that $T_\varepsilon \ge T - 2\varepsilon$ and $\hat{q}_\varepsilon$ and $\tilde{q}_\varepsilon$ coincide on $[0, T - 4\varepsilon]$, so that, shifting back in time, $\|q_\varepsilon - \bar{q}(\cdot - 2\varepsilon)\|_{C^1([2\varepsilon, T-2\varepsilon])} \to 0$ as $\varepsilon \to 0$. Since $\bar{q}$ is smooth, this gives that $\|q_\varepsilon - \bar{q}\|_{C^1([2\varepsilon, T-2\varepsilon])} \to 0$ as $\varepsilon \to 0$.

Finally one deals with the time interval $[T - 2\varepsilon, T]$ in the same way as the first step. In particular, reducing ε one more time if necessary one obtains, by an energy estimate, a Gronwall estimate and the mean value theorem, that $T_\varepsilon \ge T$. Moreover the choice of the vector v_1 in (5-3) allows us to reorient the velocity q_ε' from $c_1(q_1)$ to q_1', whereas the position is not much changed (due to the uniform bound of q_ε' and the mean value theorem) so that the value of q_ε at time T converges to q_1 as ε goes to 0. ☐

5C. ***Back to the complete model.*** Now in order to mimic the right-hand side of (5-4) we are going to use one part of the force term F_1 introduced in Definition 10. Let us therefore introduce some notations for the different contributions of the force term F_1. We define, for any q in $\mathcal{Q}$, p in $\mathbb{R}^3$, α in $C^\infty(\overline{\mathcal{F}(q)};\mathbb{R})$

$$F_{1,a}(q)[\alpha] := -\frac{1}{2}\int_{\partial\mathcal{S}(q)} |\nabla\alpha|^2\, \partial_n\Phi(q,\cdot)\,d\sigma, \tag{5-8}$$

$$F_{1,b}(q,p)[\alpha] := -\int_{\partial\mathcal{S}(q)} \nabla\alpha\cdot\nabla(p\cdot\Phi(q,\cdot))\,\partial_n\Phi(q,\cdot)\,d\sigma, \tag{5-9}$$

$$F_{1,c}(q)[\alpha] := -\int_{\partial\mathcal{S}(q)} \nabla\alpha\cdot\nabla^\perp\psi(q,\cdot)\,\partial_n\Phi(q,\cdot)\,d\sigma, \tag{5-10}$$

so that for any γ in $\mathbb{R}$

$$F_1(q,p,\gamma)[\alpha] = F_{1,a}(q)[\alpha] + F_{1,b}(q,p)[\alpha] + \gamma F_{1,c}(q)[\alpha].$$

The part which will allow us to approximate the right-hand side of (5-4) is $F_{1,a}$. More precisely we are going to see (see Proposition 20) that there exists a control α (chosen below as $\alpha = \mathcal{A}[q, g_\varepsilon]$ with g_ε given by (5-14)) that in the appropriate regime the dynamics of (2-20) behaves like the equation with only $F_{1,a}$ on the right-hand side. Moreover the following lemma, where the time parameter does not appear, proves that the operator $F_{1,a}(q)[\,\cdot\,]$ can actually attain any value v in $\mathbb{R}^3$. Recall that $\delta > 0$ was fixed at the beginning of Section 5A.

Proposition 19. *There exists a continuous mapping* $\bar{g} : \mathcal{Q}_\delta \times \mathbb{R}^3 \to \mathcal{C}$ *such that for any* (q, v) *in* $\mathcal{Q}_\delta \times \mathbb{R}^3$ *the function* $\bar\alpha := \mathcal{A}[q, \bar{g}(q,v)]$ *in* $C^\infty(\overline{\mathcal{F}(q)};\mathbb{R})$ *satisfies*

$$\Delta\bar\alpha = 0 \quad \text{in } \mathcal{F}(q) \qquad \text{and} \qquad \partial_n\bar\alpha = 0 \quad \text{on } \partial\mathcal{F}(q)\setminus\Sigma, \tag{5-11}$$

$$\int_{\partial\mathcal{S}(q)} |\nabla\bar\alpha|^2\,\partial_n\Phi(q,\cdot)\,d\sigma = v, \tag{5-12}$$

$$\int_{\partial\mathcal{S}(q)} \bar\alpha\,\partial_n\Phi(q,\cdot)\,d\sigma = 0. \tag{5-13}$$

We recall that the operator $\mathcal{A}$ was introduced in Definition 8. The result above will be proved in Section 7. Note that when $\mathcal{S}(q)$ is a homogeneous disk, an adapted version of Proposition 19 still holds; see Proposition 26 in Section 7. The condition (5-13) will be useful to cancel out the last term of (2-20).

We define

$$g_\varepsilon(t,x) := \beta_\varepsilon(t)\bar{g}(q_0, -2v_0)(x) + \beta_\varepsilon(T-t)\bar{g}(q_1,-2v_1)(x), \tag{5-14}$$

where v_0 and v_1 were defined in (5-3) for (q_1, q_1') in $\bar{B}((q_0,q_0'), r_2)$, and $\bar g$ is given by Proposition 19. The goal is to prove that for ε and $|\gamma|$ small enough, this control drives the system (2-20) with $\alpha = \mathcal{A}[q, g_\varepsilon]$ from (q_0, q_0') to (q_1, q_1'), approximately.

(1) We first observe that

$$F_{1,a}(q)[\mathcal{A}[q,g_\varepsilon]] = \beta_\varepsilon^2(t)F_{1,a}(q)[\mathcal{A}[q,\bar g(q_0,-2v_0)]] + \beta_\varepsilon^2(T-t)F_{1,a}(q)[\mathcal{A}[q,\bar g(q_1,-2v_1)]], \tag{5-15}$$

and is therefore a good candidate to approximate the right-hand side of (5-4) if q is near q_0 for t near 0 and if q is near q_1 for t near T. One then may indeed expect that

$$F_{1,a}(q)[\mathcal{A}[q, \bar{g}(q_0, -2v_0)]] \quad \text{and} \quad F_{1,a}(q)[\mathcal{A}[q, \bar{g}(q_1, -2v_1)]]$$

are close to

$$F_{1,a}(q_0)[\mathcal{A}[q_0, \bar{g}(q_0, -2v_0)]] \quad \text{and} \quad F_{1,a}(q_1)[\mathcal{A}[q_1, \bar{g}(q_1, -2v_1)]],$$

respectively, on the respective supports of $\beta_\varepsilon(\,\cdot\,)$ and $\beta_\varepsilon(T - \,\cdot\,)$. Moreover, according to Proposition 19 these last two terms are equal to v_0 and v_1 (see (5-8) and (5-12)).

(2) Next we will rigorously prove in Proposition 21 below that the conclusion of Lemma 17 for the toy system also holds when one substitutes the term $F_{1,a}(q)[\mathcal{A}[q, g_\varepsilon]]$ in (5-15). This corresponds also to (2-20) with $\gamma = 0$ and the term $F_{1,b}$ and F_2 put to zero.

(3) Finally it will appear that in an appropriate regime, in particular for small ε and $|\gamma|$, the second-to-last term of (2-20) is dominant with respect to the other terms of the right-hand side (here the condition (5-13) above will be essential in order to deal with the last term of (2-20)).

Let us state a proposition summarizing the claims above. According to the Cauchy–Lipschitz theorem there exists a controlled solution $q_{\varepsilon,\gamma}$ associated with the control g_ε introduced in (5-14), starting with the initial condition $q_{\varepsilon,\gamma}(0) = q_0$ and $q'_{\varepsilon,\gamma}(0) = q'_0$, with circulation γ, and lying in $B(q_0, r_1)$ up to some positive time $T_{\varepsilon,\gamma}$. More explicitly $q_{\varepsilon,\gamma}$ satisfies on $[0, T_{\varepsilon,\gamma}]$

$$\begin{aligned}(\mathcal{M}_g + \mathcal{M}_a(q_{\varepsilon,\gamma}))q''_{\varepsilon,\gamma} + \langle \Gamma(q_{\varepsilon,\gamma}), q'_{\varepsilon,\gamma}, q'_{\varepsilon,\gamma}\rangle \\ = \gamma^2 E(q_{\varepsilon,\gamma}) + \gamma q'_{\varepsilon,\gamma} \times B(q_{\varepsilon,\gamma}) + F_1(q_{\varepsilon,\gamma}, q'_{\varepsilon,\gamma}, \gamma)[\mathcal{A}[q_{\varepsilon,\gamma}, g_\varepsilon]] + F_2(q_{\varepsilon,\gamma})[\partial_t \mathcal{A}[q_{\varepsilon,\gamma}, g_\varepsilon]].\end{aligned} \tag{5-16}$$

Observe that due to the choice of the control g_ε in (5-14) the function $q_{\varepsilon,\gamma}$ also depends on (q_1, q'_1) through v_0 and v_1; see their definition in (5-3).

We have the following approximation result.

Proposition 20. *For ε and $|\gamma|$ small enough, $T_{\varepsilon,\gamma} \geqslant T$ and, as ε and $|\gamma|$ converge to 0^+, we have $\|q_{\varepsilon,\gamma} - \bar{q}\|_{T,\varepsilon} \to 0$ and $(q_{\varepsilon,\gamma}, q'_{\varepsilon,\gamma})(T) \to (q_1, q'_1)$, uniformly for (q_1, q'_1) in $\bar{B}((q_0, q'_0), r_2)$.*

This result will be proved in Section 6. Once Proposition 20 is proved, Theorem 14 follows rapidly. Indeed, let us set $\tilde{r} = r_2$; according to Proposition 20, for $\eta > 0$, there exists $\varepsilon = \varepsilon(\eta) > 0$ and $r' = r'(\eta)$ in $(0, \tilde{r})$ such that for any $\gamma \in \mathbb{R}$ with $|\gamma| \le r'$ and for any q'_0 in $\bar{B}(0, \tilde{r})$ the mapping $\mathcal{T}$ defined on $\bar{B}((q_0, q'_0), \tilde{r})$ by setting $\mathcal{T}(q_1, q'_1) = q_{\varepsilon,\gamma}$ has the desired properties. In particular the continuity of $\mathcal{T}$ follows from the regularity of c_0 in Lemma 16 and of the solution of ODEs on their initial data. This ends the proof of Theorem 14.

5D. ***About Remark 2.*** Now that we presented the scheme of proof of Theorem 1 let us explain how to obtain the improvement mentioned in Remark 2. It is actually a direct consequence of the explicit formula for $g_\varepsilon(t, x)$ given in (5-14) and of a change of variable in time. Due to the expression of β_ε given at the beginning of Section 5B one obtains that the total flux through Σ^-, that is, $\int_0^T \int_{\Sigma^-} g_\varepsilon \, d\sigma \, dt$, is of order $\sqrt{\varepsilon}$. Hence one can reduce ε again in order to satisfy the requirement of Remark 2.

On the other hand observe that the time-rescaling argument used in the proof of Theorem 1 from Theorem 13, see (3-2), leaves the total flux through Σ^- invariant, while the number N of steps involved in the end of the same proof does not depend on ε.

6. Closeness of the controlled system to the geodesic: proof of Proposition 20

In this section, we prove Proposition 20.

6A. ***Proof of Proposition 20.*** The proof of Proposition 20 is split in several parts. To compare $q_{\varepsilon,\gamma}$ and $\bar{q}$, we are going to consider an "intermediate trajectory" $\tilde{q}_\varepsilon$ which imitates the trajectory q_ε of the toy model of Lemma 17 by using the part $F_{1,a}$ of the force term. More precisely we define $\tilde{q}_\varepsilon$ by

$$(\mathcal{M}_g + \mathcal{M}_a(\tilde{q}_\varepsilon))\tilde{q}_\varepsilon'' + \langle \Gamma(\tilde{q}_\varepsilon), \tilde{q}_\varepsilon', \tilde{q}_\varepsilon' \rangle = F_{1,a}(\tilde{q}_\varepsilon)[\mathcal{A}[\tilde{q}_\varepsilon, g_\varepsilon]], \quad \text{with } \tilde{q}_\varepsilon(0) = q_0,\ \tilde{q}_\varepsilon'(0) = q_0', \tag{6-1}$$

where g_ε was defined in (5-14) and where the operator $\mathcal{A}$ was introduced in Definition 8. Note that due to the definition of g_ε, the function $\tilde{q}_\varepsilon$ also depends on q_1, q_1'. The statement below is equivalent to Lemma 17 for $\tilde{q}_\varepsilon$, comparing $\tilde{q}_\varepsilon$ to the "target geodesic" $\bar{q}$.

Proposition 21. *There exists $\varepsilon_1 > 0$ such that for any $\varepsilon \in (0, \varepsilon_1]$, for any (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$, the solution $\tilde{q}_\varepsilon$ given by (6-1) lies in the ball $B(q_0, r_1)$ at least up to T. Moreover $\|\tilde{q}_\varepsilon - \bar{q}\|_{T,\varepsilon}$ converges to 0 and $(\tilde{q}_\varepsilon, \tilde{q}_\varepsilon')(T)$ converges to (q_1, q_1') when ε converges to 0^+, uniformly for (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$ for both convergences.*

We recall that the norm $\|\cdot\|_{T,\varepsilon}$ was defined in (5-2). The proof of Proposition 21 can be found in Section 6B.

The following result allows us to deduce the closeness of the trajectories $q_{\varepsilon,0}$, given by (5-16) with $\gamma = 0$, and $\tilde{q}_\varepsilon$ given by (6-1). Let us recall that by the definition of $T_{\varepsilon,\gamma}$ that comes along (5-16), $q_{\varepsilon,0}$ lies in $B(q_0, r_1)$ up to the time $T_{\varepsilon,0}$, which depends on q_1, q_1'.

Proposition 22. *There exists ε_2 in $(0, \varepsilon_1]$ such that for any $\varepsilon \in (0, \varepsilon_2]$, one has $T_{\varepsilon,0} \geq T$. Moreover $\|\tilde{q}_\varepsilon - q_{\varepsilon,0}\|_{C^1([0,T])} \to 0$ when $\varepsilon \to 0^+$, uniformly for (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$.*

The proof of Proposition 22 can be found in Section 6C.

Finally, we have the following estimation of the deviation due to the circulation γ, which will be proved in Section 6D.

Proposition 23. *There exists ε_3 in $(0, \varepsilon_2]$ such that for all $\varepsilon \in (0, \varepsilon_3]$, there exists $\gamma_0 > 0$ such that for any $\gamma \in [-\gamma_0, \gamma_0]$, we have $T_{\varepsilon,\gamma} \geq T$ and $\|q_{\varepsilon,\gamma} - q_{\varepsilon,0}\|_{C^1[0,T]}$ converges to 0 when $\gamma \to 0$, uniformly for (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$.*

Propositions 21, 22 and 23 give us directly the result of Proposition 20.

6B. ***Proof of Proposition 21.*** We proceed as in the proof of Lemma 17 with a few extra complications related to the fact that the right-hand side of (6-1) is more involved than the one of (5-4) and to the fact that we need to obtain uniform convergences with respect to (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$.

As in the proof of Lemma 17 we introduce, for ε in (0, 1), the time

$$T_\varepsilon = \sup\{\hat{T} > 0 : \tilde{q}_\varepsilon(t) \in B(q_0, r_1) \text{ for } t \in (0, \hat{T})\}$$

and we first prove that there exists $\widetilde{T} > 0$ such that for any ε in (0, 1) we have $T_\varepsilon \geq \widetilde{T}$ thanks to an energy estimate. In order to deal with the term coming from (5-15) in the right-hand side of the energy estimate, recalling Remark 4 and the definition of $F_{1,a}$ in (5-8), we observe that for any $R > 0$ there exists $C > 0$ such that for any $q, \tilde{q}$ in $\mathcal{Q}_\delta$, for any v in $B(0, R)$

$$|F_{1,a}(q)[\mathcal{A}[q, \bar{g}(\tilde{q}, v)]]| \leq C. \tag{6-2}$$

This allows us to deduce from the expressions of v_0 and v_1 in (5-3) that there exists $\widetilde{T} > 0$ and $C > 0$ such that for any (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$, for any ε in (0, 1) we have $T_\varepsilon \geq \widetilde{T}$ and $\|\tilde{q}_\varepsilon'\|_{C([0,T_\varepsilon])} \leq C$. We deduce that for $\varepsilon_1 := \widetilde{T}/2$ and ε in $(0, \varepsilon_1)$ we have $T_\varepsilon \geq 2\varepsilon$ and that $\sup_{t \in [0,2\varepsilon]} |\tilde{q}_\varepsilon - q_0|$ converges to 0 as ε goes to 0 uniformly in (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$.

Now let us prove that $\tilde{q}_\varepsilon'(2\varepsilon)$ converges to $c_0(q_1)$ as ε goes to 0 uniformly in (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$. We integrate (6-1) on $[0, 2\varepsilon]$. Thus

$$\begin{aligned}(\mathcal{M}_g + \mathcal{M}_a(\tilde{q}_\varepsilon(2\varepsilon)))\tilde{q}_\varepsilon'(2\varepsilon) = (\mathcal{M}_g + \mathcal{M}_a(q_0))q_0' - \int_0^{2\varepsilon} (D\mathcal{M}_a(\tilde{q}_\varepsilon) \cdot \tilde{q}_\varepsilon') \cdot \tilde{q}_\varepsilon' \, dt \\ - \int_0^{2\varepsilon} \langle \Gamma(\tilde{q}_\varepsilon), \tilde{q}_\varepsilon', \tilde{q}_\varepsilon' \rangle \, dt + \int_0^{2\varepsilon} F_{1,a}(\tilde{q}_\varepsilon)[\mathcal{A}[\tilde{q}_\varepsilon, g_\varepsilon]] \, dt.\end{aligned} \tag{6-3}$$

Then we pass to the limit as ε goes to 0^+ in the last equality. Here we use two extra arguments with respect to the corresponding argument in the proof of Lemma 17. On the one hand we see that the convergences of $\mathcal{M}_a(\tilde{q}_\varepsilon(2\varepsilon))$ to $\mathcal{M}_a(q_0)$ and of the two first terms of the last line to 0, already obtained in the proof of Lemma 17, hold uniformly with respect to (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$, as a consequence of the uniform estimates of $\tilde{q}_\varepsilon - q_0$ and $\tilde{q}_\varepsilon'$ obtained above. On the other hand the term $F_{1,a}$ enjoys the following regularity property with respect to q: we have that $q \mapsto F_{1,a}(q)[\mathcal{A}[q, \bar{g}(q_0, v)]]$ is Lipschitz with respect to q in $\mathcal{Q}_\delta$ uniformly for v in bounded sets of $\mathbb{R}^3$. Therefore using that $\sup_{t \in [0,2\varepsilon]} |\tilde{q}_\varepsilon - q_0|$ converges to 0 as ε goes to 0 uniformly in (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$, the expressions of v_0 and v_1 in (5-3) and that $F_{1,a}(q_0)[\mathcal{A}[q_0, \bar{g}(q_0, -2v_0)]] = v_0$, according to Proposition 19 we deduce that

$$\sup_{t \in [0,2\varepsilon]} |F_{1,a}(\tilde{q}_\varepsilon)[\mathcal{A}[\tilde{q}_\varepsilon, \bar{g}(q_0, -2v_0)]] - v_0|$$

converges to 0 as ε goes to 0 uniformly in (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$. Since for t in $[0, 2\varepsilon]$, (5-15) applied to $q = \tilde{q}_\varepsilon$ is simplified into

$$F_{1,a}(\tilde{q}_\varepsilon)[\mathcal{A}[\tilde{q}_\varepsilon, g_\varepsilon]] = \beta_\varepsilon^2(t) F_{1,a}(\tilde{q}_\varepsilon)[\mathcal{A}[\tilde{q}_\varepsilon, \bar{g}(q_0, -2v_0)]],$$

and $\int_0^{2\varepsilon} \beta_\varepsilon^2(t) \, dt = 1$, we get that the last term in (6-3) converges to v_0 when ε goes to 0. Moreover, due to the choice of v_0 the first and last terms of the right-hand side of (6-3) can be combined at the limit to get $(\mathcal{M}_g + \mathcal{M}_a(q_0))c_0(q_1)$.

Therefore, inverting the matrix in the right-hand side of (6-3) and passing to the limit, we see that $\tilde{q}_\varepsilon'(2\varepsilon)$ converges to $c_0(q_1)$ as ε goes to 0 uniformly in (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$.

When t is in $[2\varepsilon, T - 2\varepsilon]$, (6-1) reduces to a geodesic equation so that the same arguments as in the proof of Lemma 17 apply.

Finally for the last step, for t in $[T - 2\varepsilon, T]$ we proceed in the same way as in the first step. This ends the proof of Proposition 21.

6C. *Proof of Proposition 22.* We begin with the following lemma, which provides a uniform boundedness for the trajectories $q_{\varepsilon,0}$ satisfying (5-16) with $\gamma = 0$, that is,

$$(\mathcal{M}_g + \mathcal{M}_a(q_{\varepsilon,0}))q''_{\varepsilon,0} + \langle \Gamma(q_{\varepsilon,0}), q'_{\varepsilon,0}, q'_{\varepsilon,0}\rangle \\ = F_{1,a}(q_{\varepsilon,0})[\mathcal{A}[q_{\varepsilon,0}, g_\varepsilon]] + F_{1,b}(q_{\varepsilon,0}, q'_{\varepsilon,0})[\mathcal{A}[q_{\varepsilon,0}, g_\varepsilon]] + F_2(q_{\varepsilon,0})[\partial_t \mathcal{A}[q_{\varepsilon,0}, g_\varepsilon]]. \quad (6\text{-}4)$$

We recall that g_ε is given by (5-14) with v_0 and v_1 given by (5-3). The terms $F_{1,a}$ and $F_{1,b}$ were defined in (5-8), (5-9), and F_2 in (2-18). Also we recall that by the definition of $T_{\varepsilon,0}$ (see the definition of $T_{\varepsilon,\gamma}$ in the end of Section 5C), during the time interval $[0, T_{\varepsilon,0}]$, the trajectory $q_{\varepsilon,0}$ remains in $B(q_0, r_1)$.

Lemma 24. *There exists* $\varepsilon_a > 0$ *such that*

$$\sup_{\substack{(q_1,q_1')\in \bar{B}((q_0,q_0'),r_2) \\ \varepsilon\in(0,\varepsilon_a]}} \|q'_{\varepsilon,0}\|_{C([0,T_{\varepsilon,0}])} < +\infty.$$

Proof. First we see that the mappings

$$q \mapsto F_{1,a}(q)[\mathcal{A}[q, \bar{g}(q_0, v)]] \quad \text{and} \quad q \mapsto F_{1,b}(q, \cdot\,)[\mathcal{A}[q, \bar{g}(q_0, v)]]$$

are bounded for q in $\mathcal{Q}_\delta$, uniformly for v in bounded sets of $\mathbb{R}^3$. Let us now focus on the F_2 term. For t in $[0, 2\varepsilon]$, we have $g_\varepsilon(t) = \beta_\varepsilon(t)\bar{g}(q_0, -2v_0)$ so that, by the chain rule, for t in $[0, \min(2\varepsilon, T_{\varepsilon,0})]$,

$$\partial_t \mathcal{A}[q_{\varepsilon,0}, g_\varepsilon] = \beta_\varepsilon D_q \mathcal{A}[q_{\varepsilon,0}, \bar{g}(q_0, -2v_0)] \cdot q'_{\varepsilon,0} + \beta'_\varepsilon \mathcal{A}[q_{\varepsilon,0}, \bar{g}(q_0, -2v_0)].$$

Concerning F_2 we have, using the property (5-13),

$$F_2(q_{\varepsilon,0})[\partial_t \mathcal{A}[q_{\varepsilon,0}, g_\varepsilon]] \\ = \beta_\varepsilon \int_{\partial \mathcal{S}(q_{\varepsilon,0})} (D_q \mathcal{A}[q_{\varepsilon,0}, \bar{g}(q_0, -2v_0)] \cdot q'_{\varepsilon,0})\, \partial_n \Phi(q_{\varepsilon,0}, \cdot\,)\, d\sigma \\ + \beta'_\varepsilon \biggl(\int_{\partial \mathcal{S}(q_{\varepsilon,0})} \mathcal{A}[q_{\varepsilon,0}, \bar{g}(q_0, -2v_0)]\, \partial_n \Phi(q_{\varepsilon,0}, \cdot\,)\, d\sigma - \int_{\partial \mathcal{S}(q_0)} \mathcal{A}[q_0, \bar{g}(q_0, -2v_0)]\, \partial_n \Phi(q_0, \cdot\,)\, d\sigma \biggr).$$

Using that the mapping $q \mapsto \int_{\partial\mathcal{S}(q)} \nabla_q \mathcal{A}[q, \bar{g}(q_0, v)] \otimes \partial_n \Phi(q, \cdot\,)\, d\sigma$ is bounded for q over $\mathcal{Q}_\delta$ and that the mapping $q \mapsto \int_{\partial\mathcal{S}(q)} \mathcal{A}[q, \bar{g}(q_0, v)]\, \partial_n \Phi(q, \cdot\,)\, d\sigma$ is Lipschitz with respect to q in $\mathcal{Q}_\delta$, both uniformly for v in bounded sets of $\mathbb{R}^3$, we see that this involves (recalling the expression of β_ε given at the beginning of Section 5B)

$$|F_2(q_{\varepsilon,0})[\partial_t \mathcal{A}[q_{\varepsilon,0}, g_\varepsilon]]| \lesssim C\biggl(\frac{1}{\varepsilon^{1/2}} |q'_{\varepsilon,0}| + \frac{1}{\varepsilon^{3/2}} |q_{\varepsilon,0} - q_0| \biggr), \quad (6\text{-}5)$$

uniformly for (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$. Then, multiplying (6-4) by $q'_{\varepsilon,0}$ and using once more the identity (2-15), we obtain, for any ε in $(0, 1)$, for t in $[0, \min(2\varepsilon, T_{\varepsilon,0})]$

$$\begin{aligned}(\mathcal{M}_g+\mathcal{M}_a(q_{\varepsilon,0}(t)))q'_{\varepsilon,0}(t)\cdot q'_{\varepsilon,0}(t) \\ = (\mathcal{M}_g+\mathcal{M}_a(q_0))q_0'\cdot q_0' \\ +2\int_0^t \big(F_{1,a}(q_{\varepsilon,0})[\mathcal{A}[q_{\varepsilon,0}, g_\varepsilon]]+F_{1,b}(q_{\varepsilon,0}, q'_{\varepsilon,0})[\mathcal{A}[q_{\varepsilon,0}, g_\varepsilon]]+F_2(q_{\varepsilon,0})[\partial_t\mathcal{A}[q_{\varepsilon,0}, g_\varepsilon]]\big)\cdot q'_{\varepsilon,0}. \quad (6\text{-}6)\end{aligned}$$

Then, using (5-5), the boundedness of the mappings

$$q \mapsto F_{1,a}(q)[\mathcal{A}[q, \bar{g}(q_0, v)]] \quad \text{and} \quad q \mapsto F_{1,b}(q, \cdot)[\mathcal{A}[q, \bar{g}(q_0, v)]]$$

already mentioned above, the definition of β_ε and the bound (6-5), we get

$$|q'_{\varepsilon,0}(t)|^2 \le C\bigg(1+\frac{1}{\varepsilon^{1/2}}\int_0^t |q'_{\varepsilon,0}(s)|^2\,ds + \frac{1}{\varepsilon^{3/2}}\int_0^t |q'_{\varepsilon,0}(s)||q_{\varepsilon,0}(s)-q_0|\,ds\bigg).$$

Then using the mean value theorem and that $t \le 2\varepsilon$, we have

$$|q'_{\varepsilon,0}(t)|^2 \le C\big(1+\varepsilon^{1/2}\sup_{[0,\min(2\varepsilon,T_{\varepsilon,0})]}|q'_{\varepsilon,0}|^2\big),$$

so that for ε small enough, and for t in $[0, \min(2\varepsilon, T_{\varepsilon,0})]$, $|q'_{\varepsilon,0}(t)| \le C$, uniformly for (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$. As a consequence of the usual blow-up criterion for ODEs, we have $T_{\varepsilon,0} \ge 2\varepsilon$.

During the next phase, i.e., for t in $[2\varepsilon, T-2\varepsilon]$, the control is inactive so that (6-4) is a geodesic equation. Then by a simple energy estimate we get again that $|q'_{\varepsilon,0}(t)| \le C$ on $[0, \min(T-2\varepsilon, T_{\varepsilon,0})]$.

Finally if $T_{\varepsilon,0} \ge T-2\varepsilon$, then we deal with the last phase as in the first phase. This concludes the proof of Lemma 24. □

We then conclude the proof of Proposition 22 by a classical comparison argument using Gronwall's lemma and the Lipschitz regularity with respect to q of the various mappings involved ($\mathcal{M}_a$, Γ, $F_{1,a}$, $F_{1,b}$ and F_2). This allows us to prove that there exists ε_2 in $(0, \varepsilon_1]$ such that for any $\varepsilon \in (0, \varepsilon_2]$ we have $T_{\varepsilon,0} \ge T$ and $\|\tilde{q}_\varepsilon - q_{\varepsilon,0}\|_{C^1([0,T])} \to 0$ when $\varepsilon \to 0^+$, uniformly for (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$. This ends the proof of Proposition 22.

6D. ***Proof of Proposition 23.*** First we may extend Lemma 24 to the solutions $q_{\varepsilon,\gamma}$ to (5-16) in the following manner.

Lemma 25. *There exists ε_b in $(0, \varepsilon_2)$ such that $\|q'_{\varepsilon,\gamma}\|_{C([0,T_{\varepsilon,\gamma}])}$ is bounded uniformly in $\varepsilon \in (0, \varepsilon_b]$ for any $\gamma \in [-1, 1]$ and for $(q_1, q_1') \in \bar{B}((q_0, q_0'), r_2)$.*

It is indeed a matter of adding the "electric field" E in (6-6), and noting that E is bounded on Q_δ; the "magnetic field" B gives no contribution to the energy.

We now finish the proof of Proposition 23. Using a comparison argument we obtain that there exists ε_3 in $(0, \varepsilon_b]$ such that for all $\varepsilon \in (0, \varepsilon_3]$, there exists $\gamma_0 > 0$ such that for any $\gamma \in [-\gamma_0, \gamma_0]$, we have $T_{\varepsilon,\gamma} \ge T$ and $\|q_{\varepsilon,\gamma} - q_{\varepsilon,0}\|_{C^1[0,T]}$ converges to 0 when $\gamma \to 0$, uniformly for (q_1, q_1') in $\bar{B}((q_0, q_0'), r_2)$. This concludes the proof of Proposition 23.

7. Design of the control according to the solid position: proof of Proposition 19

This section is devoted to the proof of Proposition 19.

7A. ***The case of a homogeneous disk.*** Before proving Proposition 19 we establish the following similar result concerning the simpler case where the solid is a homogeneous disk. In that case, the statement merely considers q of the form $q = (h, 0)$. Thus in order to simplify the writing, we introduce

$$\mathcal{Q}^h_\delta := \{h \in \mathbb{R}^2 : (h, 0) \in \mathcal{Q}_\delta\}.$$

Also throughout this section when we will write q, it will be understood that q is associated with h by $q = (h, 0)$.

Proposition 26. *Let $\delta > 0$. Then there exists a continuous mapping $\bar{g} : \mathcal{Q}^h_\delta \times \mathbb{R}^2 \to \mathcal{C}$ such that the function $\bar{\alpha} := \mathcal{A}[q, \bar{g}(q, v)]$ in $C^\infty(\overline{\mathcal{F}(q)}; \mathbb{R})$ satisfies*

$$\Delta\bar{\alpha}(q, x) = 0 \quad \text{in } \mathcal{F}(q) \qquad \text{and} \qquad \partial_n\bar{\alpha}(q, x) = 0 \quad \text{on } \partial\mathcal{F}(q) \setminus \Sigma, \tag{7-1}$$

$$\int_{\partial\mathcal{S}(q)} |\nabla\bar{\alpha}(q, x)|^2 n \, d\sigma = v, \tag{7-2}$$

$$\int_{\partial\mathcal{S}(q)} \bar{\alpha}(q, x) n \, d\sigma = 0. \tag{7-3}$$

In order to prove Proposition 26, the mapping $\bar{g}$ will be constructed using a combination of some elementary functions which we introduce in several lemmas.

To begin with, we will make use of the elementary geometrical property that $\{n(q_0, x) : x \in \partial\mathcal{S}(q_0)\}$ is the unit circle $\mathbb{S}^1$ and of the following lemma.

Lemma 27. *There exist three vectors $e_1, e_2, e_3 \in \{n(q_0, x) : x \in \partial\mathcal{S}(q_0)\}$ and positive C^∞ maps $(\mu_i)_{1\leqslant i\leqslant 3} : \mathbb{R}^2 \to \mathbb{R}_+$ such that for any $v \in \mathbb{R}^2$*

$$\sum_{i=1}^{3} \mu_i(v) e_i = v. \tag{7-4}$$

Proof. One may consider for instance $e_1 := (1, 0)$, $e_2 := (0, 1)$, $e_3 := (-1, -1)$, and

$$\mu_1(v) = v_1 + \sqrt{1 + |v_1|^2 + |v_2|^2}, \quad \mu_2(v) = v_2 + \sqrt{1 + |v_1|^2 + |v_2|^2}, \quad \mu_3(v) = \sqrt{1 + |v_1|^2 + |v_2|^2}. \ \square$$

In the next lemma, we introduce some functions that are defined in a neighborhood of $\partial\mathcal{S}(q_0)$ (for some $q_0 = (h_0, 0)$ fixed), satisfying some counterparts of the properties (7-1) and (7-2).

Lemma 28. *There exist families of functions $(\tilde{\alpha}^{i,j}_\varepsilon)_{\varepsilon\in(0,1)}$, $i, j \in \{1, 2, 3\}$, such that for any $i, j \in \{1, 2, 3\}$, for any $\varepsilon \in (0, 1)$ the function $\tilde{\alpha}^{i,j}_\varepsilon$ is defined and harmonic in a closed neighborhood $\mathcal{V}^{i,j}_\varepsilon$ of $\partial\mathcal{S}(q_0)$ and satisfies $\partial_n\tilde{\alpha}^{i,j}_\varepsilon = 0$ on $\partial\mathcal{S}(q_0)$, and moreover one has for any i, j, k, l in $\{1, 2, 3\}$,*

$$\int_{\partial\mathcal{S}(q_0)} \nabla\tilde{\alpha}^{i,j}_\varepsilon \cdot \nabla\tilde{\alpha}^{k,l}_\varepsilon n \, d\sigma \to \delta_{(i,j),(k,l)} e_i \quad \text{as } \varepsilon \to 0^+.$$

Proof. Without loss of generality, we may suppose that $\mathcal{S}(q_0)$ is the unit disk. Consider the parametrization $\{c(s) = (\cos(s), \sin(s)) : s \in [0, 2\pi]\}$ of $\partial\mathcal{S}(q_0)$ and the corresponding s_i such that $n(q_0, c(s_i)) = e_i$, $i \in \{1, 2, 3\}$.

We consider families of smooth functions $\beta_\varepsilon^{i,j} : [0, 2\pi] \to \mathbb{R}$, $i, j \in \{1, 2, 3\}$, $\varepsilon \in (0, 1)$, such that $\operatorname{supp} \beta_\varepsilon^{i,j} \cap \operatorname{supp} \beta_\varepsilon^{k,l} = \varnothing$ whenever $(i, j) \neq (k, l)$, $\operatorname{diam}(\operatorname{supp} \beta_\varepsilon^{i,j}) \to 0$ as $\varepsilon \to 0^+$,

$$\int_0^{2\pi} \beta_\varepsilon^{i,j}(s)\, d\sigma = 0 \qquad \text{and} \qquad \left| \int_0^{2\pi} |\beta_\varepsilon^{i,j}(s)|^2 n(q_0, c(s))\, ds - e_i \right| \to 0 \quad \text{as } \varepsilon \to 0^+.$$

Then we define $\tilde{\alpha}_\varepsilon^{i,j}$ in polar coordinates as the truncated Laurent series

$$\tilde{\alpha}_\varepsilon^{i,j}(r, \vartheta) := \frac{1}{2} \sum_{0<k\leq K} \frac{1}{k}\left(r^k + \frac{1}{r^k}\right)(-\hat{b}_{k,\varepsilon}^{i,j} \cos(k\vartheta) + \hat{a}_{k,\varepsilon}^{i,j} \sin(k\vartheta)),$$

where $\hat{a}_{k,\varepsilon}^{i,j}$ and $\hat{b}_{k,\varepsilon}^{i,j}$ denote the k-th Fourier coefficients of the function $\beta_\varepsilon^{i,j}$. It is elementary to check that the function $\tilde{\alpha}_\varepsilon^{i,j}$ satisfies the required properties for an appropriate choice of K. □

Now, for any $h \in \mathcal{Q}_\delta^h$, we may define

$$\mathcal{V}_\varepsilon^{i,j}(q) := \mathcal{V}_\varepsilon^{i,j} - h_0 + h,$$

which is a neighborhood of $\partial\mathcal{S}(q)$, and

$$\tilde{\alpha}_\varepsilon^{i,j}(q, x) := \tilde{\alpha}_\varepsilon^{i,j}(x + h_0 - h)$$

for each $x \in \mathcal{V}_\varepsilon^{i,j}(q)$. We have for i, j, k, l in $\{1, 2, 3\}$,

$$\int_{\partial\mathcal{S}(q)} \nabla\tilde{\alpha}_\varepsilon^{i,j}(q, x) \cdot \nabla\tilde{\alpha}_\varepsilon^{k,l}(q, x) n(q, x)\, d\sigma = \int_{\partial\mathcal{S}(q_0)} \nabla\tilde{\alpha}_\varepsilon^{i,j}(x) \cdot \nabla\tilde{\alpha}_\varepsilon^{k,l}(x) n(q_0, x)\, d\sigma.$$

Proceeding as in [Glass 2001] (see also [Glass 2012, p. 147–149]) and relying in particular Runge's theorem, we have the following result which asserts the existence of harmonic approximate extensions on the whole fluid domain.

Lemma 29. *There exists a family of functions* $(\alpha_\eta^{i,j})_{\eta\in(0,1)}$, $i, j \in \{1, 2, 3\}$, *harmonic in* $\mathcal{F}(q)$, *satisfying* $\partial_n \alpha_\eta^{i,j}(q, x) = 0$ *on* $\partial\mathcal{F}(q) \setminus \Sigma$, *with for any* k *in* $\mathbb{N}$,

$$\|\alpha_\eta^{i,j}(q, \cdot) - \tilde{\alpha}_\varepsilon^{i,j}(q, \cdot)\|_{C^k(\mathcal{V}_\varepsilon^{i,j}(q)\cap\overline{\mathcal{F}(q)})} \to 0 \quad \textit{when } \eta \to 0^+. \tag{7-5}$$

We now check that the above construction can be made continuous in q.

Lemma 30. *For any* $\nu > 0$, *there exist continuous mappings* $h \in \mathcal{Q}_\delta^h \mapsto \bar{\alpha}^{i,j}(q, \cdot) \in C^\infty(\overline{\mathcal{F}(q)})$, *where* $q = (h, 0)$, $i, j \in \{1, 2, 3\}$, *such that for any* $h \in \mathcal{Q}_\delta^h$ *we have* $\Delta_x \bar{\alpha}^{i,j}(q, x) = 0$ *in* $\mathcal{F}(q)$, $\partial_n \bar{\alpha}^{i,j}(q, x) = 0$ *on* $\partial\mathcal{F}(q) \setminus \Sigma$ *and*

$$\left| \int_{\partial\mathcal{S}(q)} \nabla\bar{\alpha}^{i,j}(q, \cdot) \cdot \nabla\bar{\alpha}^{k,l}(q, \cdot) n\, d\sigma - \delta_{(i,j),(k,l)} e_i \right| \leq \nu. \tag{7-6}$$

Proof. Let us assume that the functions $\alpha_\eta^{i,j}$ were previously defined not only for $h \in \mathcal{Q}_\delta^h$ but for $h \in \overline{\mathcal{Q}}_\delta^h$; this is possible by using a smaller δ. Hence we may for each $h \in \overline{\mathcal{Q}}_\delta^h$ find functions $\alpha_\eta^{i,j}$ (for some $\eta > 0$) satisfying the properties above, and in particular such that (7-6) is valid.

Next we observe that for any $h \in \overline{\mathcal{Q}}_\delta^h$, setting $q = (h, 0)$, the unique solution $\hat{\alpha}_\eta^{i,j}(\tilde{q}, q, \cdot)$ (up to an additive constant) to the Neumann problem

$$\begin{aligned} \Delta_x \hat{\alpha}_\eta^{i,j}(\tilde{q}, q, x) &= 0 && \text{in } \mathcal{F}(\tilde{q}), \\ \partial_n \hat{\alpha}_\eta^{i,j}(\tilde{q}, q, x) &= 0 && \text{on } \partial\mathcal{F}(\tilde{q}) \setminus \Sigma, \\ \partial_n \hat{\alpha}_\eta^{i,j}(\tilde{q}, q, x) &= \partial_n \alpha_\eta^{i,j}(q, x) && \text{on } \Sigma, \end{aligned}$$

is continuous with respect to $\tilde{q} \in \mathcal{Q}_\delta$. It follows that when a family of functions $\alpha_\eta^{i,j}$ satisfies (7-6) at some point $h \in \overline{\mathcal{Q}}_\delta^h$, it satisfies (7-6) (with perhaps 2ν in the right-hand side) in some neighborhood of h. Since $\overline{\mathcal{Q}}_\delta^h$ is compact and can be covered with such neighborhoods, one can extract a finite subcover and use a partition of unity (according to the variable q) adapted to this subcover to conclude: one gets an estimate like (7-6) with $C\nu$ on the right-hand side (for some constant C). It is then just a matter of considering ν/C rather than ν at the beginning. □

Finally our basic building blocks to prove Proposition 26 are given in the following lemma, where we can add the constraint (7-3).

Lemma 31. *For any $\nu > 0$, there exist continuous mappings $q = (h, 0) \in \mathcal{Q}_\delta \mapsto \bar{\alpha}^i(q, \cdot) \in C^\infty(\overline{\mathcal{F}(q)})$, $i \in \{1, 2, 3\}$, such that for any $q = (h, 0) \in \mathcal{Q}_\delta$ we have $\Delta_x \bar{\alpha}^i(q, x) = 0$ in $\mathcal{F}(q)$, $\partial_n \bar{\alpha}^i(q, x) = 0$ on $\partial\mathcal{F}(q) \setminus \Sigma$ and*

$$\left| \int_{\partial\mathcal{S}(q)} \nabla\bar{\alpha}^i(q, \cdot) \cdot \nabla\bar{\alpha}^j(q, \cdot) n \, d\sigma - \delta_{i,j} e_i \right| \le \nu, \tag{7-7}$$

$$\int_{\partial\mathcal{S}(q)} \bar{\alpha}^i(q, \cdot) n \, d\sigma = 0. \tag{7-8}$$

Proof. Consider the functions $\bar{\alpha}^{i,j}$ given by Lemma 30. For any $q = (h, 0) \in \mathcal{Q}_\delta$, for any $i \in \{1, 2, 3\}$ the three vectors $\int_{\partial\mathcal{S}(q)} \bar{\alpha}^{i,j}(q, \cdot) n \, d\sigma$, where $j \in \{1, 2, 3\}$, are linearly dependent in $\mathbb{R}^2$; therefore there exists $\lambda^{i,j}(q) \in \mathbb{R}$ such that

$$\sum_{j=1}^{3} \lambda^{i,j}(q) \int_{\partial\mathcal{S}(q)} \bar{\alpha}^{i,j}(q, \cdot) n \, d\sigma = 0 \quad \text{and} \quad \sum_{j=1}^{3} |\lambda^{i,j}(q)|^2 = 1. \tag{7-9}$$

Then one defines $\bar{\alpha}^i(q, \cdot) := \sum_{j=1}^3 \lambda^{i,j}(q) \bar{\alpha}^{i,j}(q, \cdot)$, and one checks that it satisfies (7-7) with some $C\nu$ in the right-hand side. Again changing ν in ν/C allows us to conclude. □

We are now in position to prove Proposition 26.

Proof of Proposition 26. Let $\delta > 0$. Let $\nu > 0$. We define the mapping $\mathcal{S}$ which with $(h, v) \in \mathcal{Q}_\delta^h \times \mathbb{R}^2$ associates the function

$$\tilde{\alpha}(q, \cdot) := \sum_{i=1}^{3} \sqrt{\mu^i(v)} \bar{\alpha}^i(q, \cdot),$$

in $C^\infty(\overline{\mathcal{F}(q)})$, where the functions μ^i were introduced in Lemma 27 and the functions $\bar{\alpha}^i$ were introduced in Lemma 31. Next we define $\mathcal{T} : \mathcal{Q}^h_\delta \times \mathbb{R}^2 \to \mathcal{Q}^h_\delta \times \mathbb{R}^2$ by

$$(h, v) \mapsto (\mathcal{T}_1, \mathcal{T}_2)(h, v) := \left(h, \int_{\partial\mathcal{S}(q)} |\nabla\tilde{\alpha}(q, \cdot)|^2 n \, d\sigma\right), \quad \text{where } \tilde{\alpha} = \mathcal{S}(h, v).$$

Using (7-4) and (7-7), one checks that $\mathcal{T}$ is smooth and that

$$\frac{\partial\mathcal{T}_2}{\partial v} = \mathrm{Id} + \mathcal{O}(\nu).$$

Hence taking ν sufficiently small, we see that $\partial\mathcal{T}_2/\partial v$ is invertible; hence $\partial\mathcal{T}/\partial(h, v)$ is invertible. Consequently one can use the inverse function theorem on $\mathcal{T}$: for each $h_0 \in \overline{\mathcal{Q}}^h_\delta$ it realizes a local diffeomorphism at$(h_0, 0)$, and hence on $\overline{\mathcal{Q}}^h_\delta \times B(0, r)$ for $r > 0$ small enough. This gives the result of Proposition 26 for v small: given $(h, v) \in \overline{\mathcal{Q}}^h_\delta \times B(0, r)$, we let $(h, \tilde{v}) := \mathcal{T}^{-1}(h, v)$. Then the functions $\bar{\alpha} := \sum_{i=1}^3 \sqrt{\mu^i(\tilde{v})}\bar{\alpha}^i(q, \cdot)$ and $\bar{g} := \mathbb{1}_\Sigma \, \partial_n\bar{\alpha}$ satisfy the requirements. The general case follows by the linearity of (7-1) and (7-3) and by the homogeneity of (7-2). This ends the proof of Proposition 26. □

7B. ***The case when $\mathcal{S}_0$ is not a disk.*** We now get back to the proof of Proposition 19. We will denote by $\mathrm{coni}(A)$ the conical hull of A, namely

$$\mathrm{coni}(A) := \left\{\sum_{i=1}^k \lambda_i a_i : k \in \mathbb{N}^*,\ \lambda_i \geq 0,\ a_i \in A\right\}.$$

The first step is the following elementary geometric lemma.

Lemma 32. *Let $\mathcal{S}_0 \subset \Omega$ bounded, closed, simply connected with smooth boundary, which is not a disk. Then* $\mathrm{coni}\{(n(x), (x - h_0)^\perp \cdot n(x)) : x \in \partial\mathcal{S}_0\} = \mathbb{R}^3$.

Proof. Suppose the contrary. Then there exists a plane separating (in the large sense) the origin in $\mathbb{R}^3$ from the set $\mathrm{coni}(\{(n(x), (x - h_0)^\perp \cdot n(x)) : x \in \partial\mathcal{S}_0\})$. We claim that a normal vector to this plane can be put in the form $(a, b, 1)$, with $a, b \in \mathbb{R}$. Indeed, otherwise it would need to be of the form $(a, b, 0)$, and the separation inequality would give $(a, b) \cdot n(x) \geq 0$ for all $x \in \partial\mathcal{S}_0$. However, since $\partial\mathcal{S}_0$ is a smooth, closed curve, the set $\{n(x) : x \in \partial\mathcal{S}_0\}$ is the unit circle of $\mathbb{R}^2$; therefore we have a contradiction.

Now we deduce that we have the separation property

$$(a, b) \cdot n(x) + (x - h_0)^\perp \cdot n(x) \geq 0 \quad \text{for all } x \in \partial\mathcal{S}_0.$$

Setting $w = (a, b) - h_0^\perp$, this translates into $(w + x^\perp) \cdot n(x) \geq 0$. But using Green's formula, we get

$$0 \leq \int_{\partial\mathcal{S}_0} (w + x^\perp) \cdot n(x) \, d\sigma = \int_{\mathcal{S}_0} \mathrm{div}(w + x^\perp) \, dx = 0,$$

and consequently, we deduce that $(w+x^\perp)\cdot n(x) = 0$ for all x in $\partial\mathcal{S}_0$. This is equivalent to $(x - w^\perp)\cdot\tau(x) = 0$ for all x in $\partial\mathcal{S}_0$. Parametrizing the translated curve $\partial\mathcal{S}_0 - w^\perp$ by $\{c(s) : s \in [0, 1]\}$, it follows that $c(s) \cdot \dot{c}(s) = 0$, for all s in $[0, 1]$, and therefore $|c(s)|^2$ is constant. This means that $\partial\mathcal{S}_0 - w^\perp$ is a circle, so $\mathcal{S}_0$ is a disk, which is a contradiction. □

Fix $q_0 \in Q_\delta$. Recalling the definitions of the Kirchhoff potentials in (2-2) and (2-3), we infer from the previous lemma that

$$\operatorname{coni}\{\partial_n \Phi(q_0, x) : x \in \partial\mathcal{S}_0\} = \mathbb{R}^3.$$

In place of Lemma 27, we have the following lemma which is a straightforward consequence of Lemma 32 and of a repeated application of Carathéodory's theorem on the convex hull.

Lemma 33. *There are some $(x_i)_{i\in\{1,\dots,16\}}$ in $\partial\mathcal{S}_0$ and positive continuous mappings $\mu_i : \mathbb{R}^3 \to \mathbb{R}$, $1 \leqslant i \leqslant 16$, $v \mapsto \mu_i(v)$, such that*

$$\sum_{i=1}^{16} \mu_i(v)\partial_n \Phi(q_0, x_i) = v.$$

We are now in position to establish Proposition 19. We deduce from Lemma 33 that for any $q := (h, \vartheta) \in \overline{\mathcal{Q}}_\delta$, for any v in $\mathbb{R}^3$

$$\sum_{i=1}^{16} \mu_i(v)\, \partial_n \Phi(q, x_i(q)) = \mathcal{R}(\vartheta)v,$$

where $x_i(q) := R(\vartheta)(x_i - h_0) + h$ and $\mathcal{R}(\vartheta)$ denotes the 3×3 rotation matrix defined by

$$\mathcal{R}(\vartheta) := \begin{pmatrix} R(\vartheta) & 0 \\ 0 & 1 \end{pmatrix}.$$

Due to the Riemann mapping theorem, there exists a biholomorphic mapping $\Psi : \bar{\mathbb{C}} \setminus B(0, 1) \to \bar{\mathbb{C}} \setminus \mathcal{S}(q)$ with $\partial\mathcal{S}(q) = \Psi(\partial B(0, 1))$, where $\bar{\mathbb{C}}$ denotes the Riemann sphere. We consider the parametrizations $\{c(s) = (\cos(s), \sin(s)) : s \in [0, 2\pi]\}$ of $\partial B(0, 1)$ and $\{\Psi(c(s)) : s \in [0, 2\pi]\}$ of $\partial\mathcal{S}(q)$, and the corresponding s_i such that $x_i(q) = \Psi(c(s_i))$ for $i \in \{1, \dots, 16\}$.

Then, for any smooth function $\alpha : \partial\mathcal{S}(q) \to \mathbb{R}$, due to the Cauchy–Riemann relations, we have

$$\partial_n \alpha(\Psi(x)) = \frac{1}{\sqrt{|\det(D\Psi(x))|}} \partial_{n_B}(\alpha \circ \Psi)(x),$$

$$\int_{\partial\mathcal{S}(q)} |\nabla\alpha(x)|^2\, \partial_n \Phi(q, x)\, d\sigma = \int_{\partial B(0,1)} |\nabla\alpha(\Psi(x))|^2\, \partial_{n_B} \Phi(q, \Psi(x)) \frac{1}{\sqrt{|\det(D\Psi(x))|}}\, d\sigma$$

for any $x \in \partial B(0, 1)$, where n and n_B respectively denote the normal vectors on $\partial\mathcal{S}(q)$ and $\partial B(0, 1)$. Note that, since Ψ is invertible, we have $|\det(D\Psi(x))| > 0$ for any $x \in \partial B(0, 1)$.

For each $\varepsilon > 0$, $i \in \{1, \dots, 16\}$, $j \in \{1, 2, 3, 4\}$ (here the index j belongs to $\{1, 2, 3, 4\}$ rather than $\{1, 2, 3\}$ in order to adapt the linear dependence argument of Lemma 31 to the case of the three linear constraints (5-13)), we consider families of smooth functions $\beta_\varepsilon^{i,j} : [0, 2\pi] \to \mathbb{R}$ satisfying $\operatorname{supp} \beta_\varepsilon^{i,j} \cap \operatorname{supp} \beta_\varepsilon^{k,l} = \varnothing$ for $(i, j) \neq (k, l)$, $\operatorname{diam}(\operatorname{supp} \beta_\varepsilon^{i,j}) \to 0$ as $\varepsilon \to 0^+$,

$$\int_0^{2\pi} \beta_\varepsilon^{i,j}(s)\, ds = 0,$$

and

$$\left| \int_0^{2\pi} |\beta_\varepsilon^{i,j}(s)|^2\, \partial_n \Phi(q, c(s)) \frac{1}{\sqrt{|\det(D\Psi(c(s)))|}}\, ds - \tilde{e}_i \right| \to 0 \quad \text{as } \varepsilon \to 0^+,$$

where

$$\tilde{e}_i := \frac{1}{\sqrt{|\det(D\Psi(c(s_i)))|}} \partial_n \Phi(q, x_i(q)).$$

Then one may proceed essentially as in the proof of Proposition 26. The details are therefore left to the reader.

Acknowledgements

We would like to thank Jimmy Lamboley and Alexandre Munnier for helpful conversations on shape differentiation. The authors also thank the Agence Nationale de la Recherche, Project DYFICOLTI, grant ANR-13-BS01-0003-01 and Project IFSMACS, grant ANR-15-CE40-0010 for their financial support. Sueur was also partially supported by the Agence Nationale de la Recherche, Project SINGFLOWS grant ANR-18-CE40-0027-01, Project BORDS, grant ANR-16-CE40-0027-01, the Conseil Régional d'Aquitaine, grant 2015.1047.CP, The Simone and Cino Del Duca Foundation, and the H2020-MSCA-ITN-2017 program, Project ConFlex, grant ETN-765579. Furthermore, Kolumbán would also like to thank the Fondation Sciences Mathématiques de Paris for their support in the form of the PGSM Phd Fellowship.

References

[Aronsson 1973] G. Aronsson, "Global controllability and bang-bang steering of certain nonlinear systems", *SIAM J. Control* **11**:4 (1973), 607–619. MR Zbl

[Boulakia and Guerrero 2013] M. Boulakia and S. Guerrero, "Local null controllability of a fluid-solid interaction problem in dimension 3", *J. Eur. Math. Soc.* **15**:3 (2013), 825–856. MR Zbl

[Boulakia and Osses 2008] M. Boulakia and A. Osses, "Local null controllability of a two-dimensional fluid-structure interaction problem", *ESAIM Control Optim. Calc. Var.* **14**:1 (2008), 1–42. MR Zbl

[Bressan 1996] A. Bressan, "Impulsive control systems", pp. 1–22 in *Nonsmooth analysis and geometric methods in deterministic optimal control* (Minneapolis, MN, 1993), edited by B. S. Mordukhovich and H. J. Sussmann, IMA Vol. Math. Appl. **78**, Springer, 1996. MR Zbl

[Brunovský and Lobry 1975] P. Brunovský and C. Lobry, "Controlabilité bang bang, controlabilité différentiable, et perturbation des systèmes non linéaires", *Ann. Mat. Pura Appl.* (4) **105** (1975), 93–119. MR Zbl

[Chambrion and Munnier 2012] T. Chambrion and A. Munnier, "Generic controllability of 3D swimmers in a perfect fluid", *SIAM J. Control Optim.* **50**:5 (2012), 2814–2835. MR Zbl

[Coron 1993] J.-M. Coron, "Contrôlabilité exacte frontière de l'équation d'Euler des fluides parfaits incompressibles bidimensionnels", *C. R. Acad. Sci. Paris Sér. I Math.* **317**:3 (1993), 271–276. MR Zbl

[Coron 1996] J.-M. Coron, "On the controllability of the 2-D incompressible Navier–Stokes equations with the Navier slip boundary conditions", *ESAIM Control Optim. Calc. Var.* **1** (1996), 35–75. MR Zbl

[Coron et al. 2017] J.-M. Coron, F. Marbach, and F. Sueur, "On the controllability of the Navier–Stokes equation in spite of boundary layers", pp. 162–180 in *Mathematical analysis of viscous incompressible fluid* (Kyoto, 2016), edited by Y. Maekawa, RIMS Kôkyûroku **2058**, Kyoto Univ. Inst. Math. Anal., 2017.

[Coron et al. 2020] J.-M. Coron, F. Marbach, and F. Sueur, "Small-time global exact controllability of the Navier–Stokes equation with Navier slip-with-friction boundary conditions", *J. Eur. Math. Soc.* **22**:5 (2020), 1625–1673.

[Gaines 1969] R. Gaines, "Continuous dependence for two-point boundary value problems", *Pacific J. Math.* **28**:2 (1969), 327–336. MR Zbl

[Glass 2000] O. Glass, "Exact boundary controllability of 3-D Euler equation", *ESAIM Control Optim. Calc. Var.* **5** (2000), 1–44. MR Zbl

[Glass 2001] O. Glass, "An addendum to a J. M. Coron theorem concerning the controllability of the Euler system for 2D incompressible inviscid fluids", *J. Math. Pures Appl.* (9) **80**:8 (2001), 845–877. MR Zbl

[Glass 2012] O. Glass, "Some questions of control in fluid mechanics", pp. 131–206 in *Control of partial differential equations* (Cetraro, Italy, 2010), edited by P. Cannarsa and J.-M. Coron, Lecture Notes in Math. **2048**, Springer, 2012. MR

[Glass and Horsin 2010] O. Glass and T. Horsin, "Approximate Lagrangian controllability for the 2-D Euler equation: application to the control of the shape of vortex patches", *J. Math. Pures Appl.* (9) **93**:1 (2010), 61–90. MR Zbl

[Glass and Horsin 2012] O. Glass and T. Horsin, "Prescribing the motion of a set of particles in a three-dimensional perfect fluid", *SIAM J. Control Optim.* **50**:5 (2012), 2726–2742. MR Zbl

[Glass and Horsin 2016] O. Glass and T. Horsin, "Lagrangian controllability at low Reynolds number", *ESAIM Control Optim. Calc. Var.* **22**:4 (2016), 1040–1053. MR Zbl

[Glass and Rosier 2013] O. Glass and L. Rosier, "On the control of the motion of a boat", *Math. Models Methods Appl. Sci.* **23**:4 (2013), 617–670. MR Zbl

[Glass and Sueur 2012] O. Glass and F. Sueur, "The movement of a solid in an incompressible perfect fluid as a geodesic flow", *Proc. Amer. Math. Soc.* **140**:6 (2012), 2155–2168. MR Zbl

[Glass and Sueur 2015] O. Glass and F. Sueur, "Uniqueness results for weak solutions of two-dimensional fluid-solid systems", *Arch. Ration. Mech. Anal.* **218**:2 (2015), 907–944. MR Zbl

[Glass et al. 2012] O. Glass, F. Sueur, and T. Takahashi, "Smoothness of the motion of a rigid body immersed in an incompressible perfect fluid", *Ann. Sci. Éc. Norm. Supér.* (4) **45**:1 (2012), 1–51. MR Zbl

[Glass et al. 2014] O. Glass, C. Lacave, and F. Sueur, "On the motion of a small body immersed in a two-dimensional incompressible perfect fluid", *Bull. Soc. Math. France* **142**:3 (2014), 489–536. MR Zbl

[Glass et al. 2018] O. Glass, A. Munnier, and F. Sueur, "Point vortex dynamics as zero-radius limit of the motion of a rigid body in an irrotational fluid", *Invent. Math.* **214**:1 (2018), 171–287. MR Zbl

[Grasse 1981] K. A. Grasse, "Perturbations of nonlinear controllable systems", *SIAM J. Control Optim.* **19**:2 (1981), 203–220. MR Zbl

[Grasse 1982] K. A. Grasse, "Nonlinear perturbations of control-semilinear control systems", *SIAM J. Control Optim.* **20**:3 (1982), 311–327. MR Zbl

[Henrot and Pierre 2005] A. Henrot and M. Pierre, *Variation et optimisation de formes: une analyse géométrique*, Math. Appl. **48**, Springer, 2005. MR Zbl

[Houot et al. 2010] J. G. Houot, J. San Martin, and M. Tucsnak, "Existence of solutions for the equations modeling the motion of rigid bodies in an ideal fluid", *J. Funct. Anal.* **259**:11 (2010), 2856–2885. MR Zbl

[Imanuvilov and Takahashi 2007] O. Imanuvilov and T. Takahashi, "Exact controllability of a fluid-rigid body system", *J. Math. Pures Appl.* (9) **87**:4 (2007), 408–437. MR Zbl

[Kolumbán 2020] J. Kolumbán, "Control at a distance of the motion of a rigid body immersed in a two-dimensional viscous incompressible fluid", *J. Differential Equations* **269**:1 (2020), 764–831. MR Zbl

[Lohéac and Munnier 2014] J. Lohéac and A. Munnier, "Controllability of 3D low Reynolds number swimmers", *ESAIM Control Optim. Calc. Var.* **20**:1 (2014), 236–268. MR Zbl

[Marsden and Ratiu 1994] J. E. Marsden and T. S. Ratiu, *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems*, Texts in Appl. Math. **17**, Springer, 1994. MR Zbl

[Munnier 2009] A. Munnier, "Locomotion of deformable bodies in an ideal fluid: Newtonian versus Lagrangian formalisms", *J. Nonlinear Sci.* **19**:6 (2009), 665–715. MR Zbl

[Ortega et al. 2005] J. H. Ortega, L. Rosier, and T. Takahashi, "Classical solutions for the equations modelling the motion of a ball in a bidimensional incompressible perfect fluid", *M2AN Math. Model. Numer. Anal.* **39**:1 (2005), 79–108. MR Zbl

[Ortega et al. 2007] J. Ortega, L. Rosier, and T. Takahashi, "On the motion of a rigid body immersed in a bidimensional incompressible perfect fluid", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **24**:1 (2007), 139–165. MR Zbl

[Rosier and Rosier 2009] C. Rosier and L. Rosier, "Smooth solutions for the motion of a ball in an incompressible perfect fluid", *J. Funct. Anal.* **256**:5 (2009), 1618–1641. MR Zbl

[Yudovich 1962] V. I. Yudovich, "The flow of a perfect, incompressible liquid through a given region", *Dokl. Akad. Nauk SSSR* **146** (1962), 561–564. In Russian; translated in *Soviet Physics Dokl.* **7** (1962), 789–791. MR Zbl

Received 13 Jul 2017. Revised 23 Oct 2018. Accepted 18 Apr 2019.

OLIVIER GLASS: glass@ceremade.dauphine.fr
CEREMADE, UMR CNRS 7534, Université de Paris-Dauphine, PSL Research University, Paris, France

JÓZSEF J. KOLUMBÁN: jozsef.kolumban@math.uni-leipzig.de
Institut für Mathematik, Universität Leipzig, Leipzig, Germany

FRANCK SUEUR: franck.sueur@math.u-bordeaux.fr
Institut Universitaire de France, Bordeaux INP, Institut de Mathématiques de Bordeaux, UMR 5251, Université de Bordeaux, Talence, France

ANALYSIS AND PDE
Vol. 13, No. 3, 2020
dx.doi.org/10.2140/apde.2020.13.685

DISTANCE GRAPHS AND SETS OF POSITIVE UPPER DENSITY IN $\mathbb{R}^d$

NEIL LYALL AND ÁKOS MAGYAR

We present a refinement and sharp extension of a result of Bourgain on finding configurations of $k+1$ points in general position in measurable subset of $\mathbb{R}^d$ of positive upper density whenever $d \geq k+1$ to all proper k-degenerate distance graphs.

1. Introduction

1.1. *Background.* A result of Furstenberg, Katznelson, and Weiss [Furstenberg et al. 1990] states that if $A \subseteq \mathbb{R}^2$ has positive upper Banach density, then its distance set $\{|x-x'| : x, x' \in A\}$ contains all sufficiently large numbers. Recall that the *upper Banach density* of a measurable set $A \subseteq \mathbb{R}^d$ is defined by

$$\delta^*(A) = \lim_{N\to\infty} \sup_{t\in\mathbb{R}^d} \frac{|A \cap (t+Q_N)|}{|Q_N|}, \tag{1}$$

where $|\cdot|$ denotes Lebesgue measure on $\mathbb{R}^d$ and Q_N denotes the cube $[-N/2, N/2]^d$.

Note that the distance set of any set of positive Lebesgue measure in $\mathbb{R}^d$ automatically contains all sufficiently small numbers (by for example the Lebesgue density theorem) and that it is easy to construct a set of positive upper density which does not contain a fixed distance by placing small balls centered on an appropriate square lattice.

This result was later reproved using Fourier analytic techniques by Bourgain [1986]. In fact he established the following more general result for all finite point configurations $V = \{v_0, v_1, \ldots, v_k\}$ with the property that $\{v_1 - v_0, \ldots, v_k - v_0\}$ forms a linearly independent collection of vectors in $\mathbb{R}^d$, namely for all nondegenerate simplices. In the sequel we shall refer to such point configurations as being in *general position*.

Theorem 1 [Bourgain 1986]. *Let $\Delta_k \subseteq \mathbb{R}^d$ be a fixed collection of $k+1$ points in general position.*

If $A \subseteq \mathbb{R}^d$ has positive upper Banach density and $d \geq k+1$, then there exists a threshold $\lambda_0 = \lambda_0(A, \Delta_k)$ such that A contains an isometric copy of $\lambda \cdot \Delta_k$ for all $\lambda \geq \lambda_0$.

Recall that a point configuration Δ_k' is said to be an isometric copy of $\lambda \cdot \Delta_k$ if there exists a bijection $\phi : \Delta_k \to \Delta_k'$ such that $|\phi(v) - \phi(w)| = \lambda |v-w|$ for all $v, w \in \Delta_k$.

Bourgain [1986] further demonstrated that no result along the lines of Theorem 1 can hold for configurations that contain any three points in arithmetic progression on a line, specifically showing that

Lyall was partially supported by grant NSF-DMS 1702411 and the Simons Foundation Collaboration Grant for Mathematicians 245792. Magyar was partially supported by grant NSF-DMS 1600840 and ERC-AdG 321104.
MSC2010: 11B30.
Keywords: distance graphs, uniformity norms, geometric Ramsey theory.

for any $d \geq 1$ there are sets of positive upper Banach density in $\mathbb{R}^d$ which do not contain an isometric copy of configurations of the form $\{0, y, 2y\}$ with $|y| = \lambda$ for all sufficiently large λ. However, Ziegler [2006] established the remarkable result that if $A \subseteq \mathbb{R}^d$ with $d \geq 2$ has positive upper density and $V = \{0, v_1, \ldots, v_n\} \subseteq \mathbb{R}^d$, where n can notably be taken arbitrarily large with respect to d, then there does exists a threshold $\lambda_0 = \lambda_0(A, V)$ such that A_ε contains an isometric copy of $\lambda \cdot V$ for all $\lambda \geq \lambda_0$ and any $\varepsilon > 0$, where A_ε denotes the ε-neighborhood of A.

Together these results may be viewed as initial results in *geometric Ramsey theory* where, roughly speaking, one shows that "large" but otherwise arbitrary sets necessarily contain certain geometric configurations. Recently there has been a number of results in this direction in various contexts; see [Bulinski 2018; Bennett et al. 2016; Iosevich and Parshall 2018]. The objective of this article is to present a common extension in the setting of measurable subsets of Euclidean spaces of positive upper Banach density, while simultaneously presenting a new approach to (and refinement of) Theorem 1 based on a simple notion of uniform distribution attached to an appropriate scale. For another instance of this new approach see [Lyall and Magyar 2018] where configurations of points that form the vertices of a rigid geometric square, and more generally the direct product of any two finite point configurations in general position, are addressed.

1.2. ***Distance graphs and main result.*** A *distance graph* $\Gamma = \Gamma(V, E)$ is a connected finite graph with vertex set V contained in $\mathbb{R}^d$ for some $d \geq 1$. We say that Γ is *k-degenerate* if each of its subgraphs contains a vertex with degree at most k; that is, some vertex in the subgraph touches k or fewer of the subgraphs edges. It is thus straightforward to verify, by induction, that if a given graph is k-degenerate, then there exists an ordering of its vertex set $V = \{v_0, v_1, \ldots, v_n\}$ in such a way that $|V_j| \leq k$ for all $1 \leq j \leq n$, where

$$V_j := \{v_i : (v_i, v_j) \in E \text{ with } 0 \leq i < j\} \tag{2}$$

denotes the set of predecessors of the vertex v_j. In this article we shall always assume that the vertices of any given k-degenerate graph have been ordered as such. The *degeneracy* of a graph is defined to be the smallest k for which it is k-degenerate. Finally, we shall refer to a distance graph as *proper* if for every $1 \leq j \leq n$, the set of vertices $v_j \cup V_j$, namely v_j together with its predecessors, are in general position.

Given a distance graph $\Gamma = \Gamma(V, E)$ and $\lambda > 0$ we will say that $\Gamma' = \Gamma'(V', E')$ is *isometric* to $\lambda \cdot \Gamma$ if there exists a bijection $\phi : V \to V'$ such that $(v, w) \in E$ if and only if $(\phi(v), \phi(w)) \in E'$ and $|\phi(v) - \phi(w)| = \lambda |v - w|$, and say that Γ' is a *δ-close isometric* copy of $\lambda \cdot \Gamma$ if one has the additional "angular closeness" property that

$$\frac{(\phi(v) - \phi(w)) \cdot (v - w)}{|\phi(v) - \phi(w)|\,|v - w|} > 1 - \delta \tag{3}$$

for all $(v, w) \in E$. Note that if $\delta > 2$ then a *δ-close isometric* copy is merely an isometric copy. Finally, we say that $A \subseteq \mathbb{R}^d$ contains a distance graph $\Gamma = \Gamma(V, E)$ if $V \subseteq A$.

The main result of this article is the following:

Theorem 2. *Let $\Gamma = \Gamma(V, E)$ be a proper k-degenerate distance graph and $\delta > 0$:*

(i) *If $A \subseteq \mathbb{R}^d$ has positive upper Banach density and $d \geq k+1$, then there exists $\lambda_0 = \lambda_0(A, \Gamma, \delta)$, which tends to infinity as $\delta \to 0^+$, such that A contains a δ-close isometric copy of $\lambda \cdot \Gamma$ for all $\lambda \geq \lambda_0$.*

(ii) *If* $A \subseteq [0, 1]^d$ *with* $|A| > 0$ *and* $d \geq k+1$, *then* A *will contain a* δ*-close isometric copy of* $\lambda \cdot \Gamma$ *for all* λ *in some interval of length at least* $\exp(-C_{\Gamma,\delta}|A|^{-C|V|})$, *with* $C_{\Gamma,\delta}$ *tending to infinity as* $\delta \to 0^+$.

Intuitively one should visualize a distance graph with edges made of rigid rods which can freely turn around the vertices. One should further visualize an isometric copy of a distance graph in a set $A \subseteq \mathbb{R}^d$ as a folding of the graph so that all of its vertices are supported on A, and a δ-close isometric copy of a distance graph in a set $A \subseteq \mathbb{R}^d$, with $\delta > 0$ and small, as a suitably small perturbation of the graph so that all of its vertices are supported on A.

Part (i) of Theorem 2 already constitutes a refinement of Theorem 1 when Γ is simply taken to be a complete distance graph on $(k+1)$ vertices in general position and $\delta > 0$ is taken sufficiently small. In this special case it establishes that positive upper density subsets of $\mathbb{R}^d$ not only contain an isometric copy of all sufficiently large dilates of a given nondegenerate simplex, as already guaranteed by Theorem 1, but that these copies can in fact be found as sufficiently large dilates of a "small rotation" of the original simplex. We further note that in both parts of Theorem 2 the dimension d is restricted only by the "level of degeneracy" of the given distance graph and not on the number of its vertices which could in fact be arbitrarily large. It is important to further observe that the length of the interval of dilations guaranteed by Part (ii) of Theorem 2 depends only on the measure of A and not on the set A itself.

Allowing the edges to rotate around the vertices is essential in our arguments. For example, the authors are unaware of any proof that there are k-equally spaced points along a line in a subset of positive density of $\mathbb{R}^2$, with *arbitrary* large gaps that does not invoking Szemerédi's theorem [1975], and that such a result is in fact not possible for *all* sufficiently large gaps[1]. The reason being that the linear relations between the points of the pattern are no longer there when we allow for rotations of the edges around the vertices. A crucial observation of this note is that in this case the frequency of isometric copies in a given set is controlled by a simple norm, which may be we viewed as a Euclidean analogue of the so-called U^1-seminorm [Tao and Vu 2006, Chapter 11], utilized in additive combinatorics. In the context of finite field geometries, a geometric analogue of the Gowers U^2-uniformity norm was developed in [Lyall et al. 2018] and used to prove that sets of positive density contain isometric copies of all circular quadrilaterals. We hope to address such problems for subsets of positive upper density of Euclidean spaces in the future.

As mentioned above, various special cases of our main result have been established, albeit in different contexts. Indeed, in [Bulinski 2018] the embedding of large copies of trees (1-degenerate distance graphs) was shown for dense subsets of the integer lattice. In [Bennett et al. 2016] it was shown that measurable subsets $A \subseteq [0, 1]^d$ of Hausdorff dimension larger than $(d+1)/2$ contain an isometric copy of $\lambda \cdot \Gamma$ for all λ in some interval, in the special case when Γ is a finite path. Very recently, parallel to our work, embedding of bounded degree distance graphs was addressed for subsets of vector spaces over finite fields [Iosevich and Parshall 2018].

Examples of distance graphs. (a) A nonempty connected graph is 1-degenerate if and only if it is a tree (contains no cycles). Any tree with vertices in $\mathbb{R}^d$ with $d \geq 1$ is isometric to a proper 1-degenerate distance graph in $\mathbb{R}^2$.

[1] For $k = 3$ a result of this type was obtained in [Cook et al. 2017], with the Euclidean distance replaced by the ℓ^p-distance, for all $p \neq 2$.

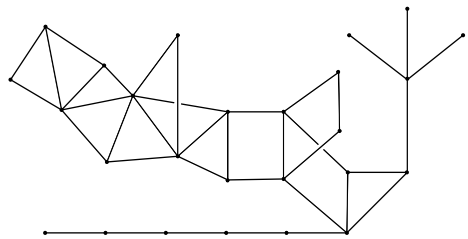

Figure 1. An example of a proper 2-degenerate distance graph in $\mathbb{R}^3$

(b) Cycles with vertices in $\mathbb{R}^d$ with $d \geq 1$ form 2-degenerate distance graphs, but these are not necessarily isometric to a proper 2-degenerate distance graph in $\mathbb{R}^d$ for any $d \geq 1$. Indeed, if $V = \{0, 1, 2\} \subseteq \mathbb{R}$ and $E = \{(0, 1), (1, 2), (0, 2)\}$, then this defines just such a distance graph.

(c) If $V = \{(i, j) : 0 \leq i, j \leq n\} \subseteq \mathbb{R}^2$ and $E = \{((i, j), (i' j')) : |i - i'| + |j - j'| = 1\}$, then this "2-dimensional grid" forms a proper 2-degenerate distance graph in $\mathbb{R}^2$.

In general, one can construct a proper 2-degenerate distance graph in $\mathbb{R}^3$ as follows: Start with any proper cycle with vertices in $\mathbb{R}^3$, such as a proper triangle (three vertices in general position) or four vertices forming a "nonrigid" square (no diagonal edges), and at every step attach an edge (or vertex) of another proper cycle (or tree) to any of the edges (or vertices) of the graph constructed at the previous step. See Figure 1.

(d) A complete graph with vertices $\{v_0, \dots, v_k\} \subseteq \mathbb{R}^k$ forms a proper k-degenerate distance graph if and only if $\{v_0, \dots, v_k\}$ are in general position. Another example of a proper k-degenerate distance graph in $\mathbb{R}^k$ is the "k-dimensional grid" with vertices $V = \{(i_1, \dots, i_k) : 0 \leq i_1, \dots, i_k \leq n\} \subseteq \mathbb{R}^k$ and edges $E = \{((i_1, \dots, i_k), (i'_1, \dots, i'_k)) : |i_1 - i'_1| + \cdots + |i_k - i'_k| = 1\}$.

More generally, one can construct a proper k-degenerate distance graph in $\mathbb{R}^{k+1}$ as follows: Start with any known proper k-degenerate distance graph with vertices in $\mathbb{R}^{k+1}$ and at every step attach another proper ℓ-degenerate distance graph with $\ell \leq k$ to any of the faces, edges, or vertices of the graph constructed at the previous step.

Remark on the sharpness of the dimension condition in Theorem 2. Let $e_1, \dots, e_k$ be the standard basis vectors of $\mathbb{R}^k$ and Δ_+ and Δ_- denote the complete graphs with vertices $\{0, e_1, e_2, \dots, e_k\}$ and $\{0, -e_1, e_2, \dots, e_k\}$ respectively. It is clear that $\Gamma = \Delta_+ \cup \Delta_-$ then defines a proper k-degenerate distance graph with the property that any isometric copy of $\lambda \cdot \Gamma$ in $\mathbb{R}^k$ must contain three collinear points, i.e., a copy of $\{-\lambda e_1, 0, \lambda e_1\}$ obtained by a translation and a rotation. As mentioned above, it was shown in [Bourgain 1986] that there are sets of positive upper Banach density in $\mathbb{R}^k$, for any k, which do not contain such configurations for all large λ. This example shows the sharpness of the dimension condition $d \geq k + 1$ in Theorem 2.

1.3. *Outline of the paper.* In Section 2 we introduce a norm which measures the uniformity of distribution with respect to a scale L. We prove that this norm controls the frequency with which isometric copies of a given distance graph occur in a subset of the unit cube. This is analogous to the so-called von Neumann type inequalities in additive combinatorics; see for example [Tao and Vu 2006, Chapter 11].

In Section 3 we observe that sets of positive density are uniformly distributed with respect to sufficiently large scales which immediately implies Part (i) of Theorem 2. The proof of Part (ii) is also provided in Section 3 and based on a decomposition of a set into uniformly distributed parts.

Section 4 contains an alternative approach inspired by the argument in [Bourgain 1986]. We include this in order to highlight the simplicity and directness of our approach to Part (i) of Theorem 2, but also with the hope that this will serve to clarify Bourgain's approach and emphasize that our approach to Part (ii) of Theorem 2 is in essence a physical space reinterpretation of Bourgain's original.

2. A counting function and generalized von Neumann inequality

We now fix $\delta > 0$ and let $\Gamma = \Gamma(V, E)$ denote a fixed proper k-degenerate distance graph with vertex set $V = \{v_0, v_1, \dots, v_n\}$ with $v_0 = 0$ in $\mathbb{R}^d$ with $d \geq k+1$.

As our arguments are analytic, we need to define a measure on, or at least on a local piece of, the configuration space of all isometric copies of Γ. For each $(v_i, v_j) \in E$ let $t_{ij} = |v_i - v_j|^2$. The configuration space of all isometric copies of Γ, with the vertex v_0 remaining fixed at 0, namely

$$S_\Gamma := \{(x_0, x_1, \dots, x_n) \in \mathbb{R}^{d(n+1)} : x_0 = 0 \text{ and } |x_i - x_j|^2 = t_{ij} \text{ for all } i, j \text{ for which } (v_i, v_j) \in E\} \quad (4)$$

is clearly a real subvariety. We now proceed to give an equivalent description of the S_Γ.

For each point $(x_0, x_1, \dots, x_n) \in \mathbb{R}^{d(n+1)}$ and $1 \leq j \leq n$ we let

$$X_j := \{x_i : i \text{ has the property that } v_i \in V_j\},$$

with $V_j = \{v_i : (v_i, v_j) \in E \text{ with } 0 \leq i < j\}$ as in (2) above. Moreover, for each $(x_0, x_1, \dots, x_n) \in \mathbb{R}^{d(n+1)}$ and $1 \leq j \leq n$, we define the sets

$$S_j = S_j(X_j) := \{x \in \mathbb{R}^d : |x - x_i|^2 = t_{ij} \text{ for all } x_i \in X_j\}. \quad (5)$$

Note that for any finite set X_j, $S_j(X_j)$ is the intersection of $|X_j|$ spheres and hence is itself either a sphere (of some dimension) or empty. It is thus easy to see that

$$(0, x_1, \dots, x_n) \in S_\Gamma \iff x_j \in S_j(X_j) \quad \text{for all } 1 \leq j \leq n \quad (6)$$

since both sides are defined by the same set of equations given in (4) above.

Since Γ is proper there exists a point $(x_0, x_1, \dots, x_n) \in S_\Gamma$ with the property that the sets $\overline{X}_j := \{x_j\} \cup X_j$ are in general position for all $1 \leq j \leq n$; for example one could (and we will) take $x_j = v_j$ for all $0 \leq j \leq n$. We shall refer to such points in S_Γ as *proper*. The following example illustrates that there may exist points of S_Γ that are not proper. Let $V = \{v_0, v_1, v_2, v_3\} \subseteq \mathbb{R}^3$ and $E = \{\{(v_0, v_1), (v_0, v_2), (v_1, v_3), (v_2, v_3)\}$, where

$$v_0 = (0, 0, 0), \quad v_1 = (1, 0, 0), \quad v_2 = (0, 2, 0), \quad v_3 = (1, 2, 0).$$

Then $\Gamma = \Gamma(V, E)$ is a proper 2-degenerate distance graph in $\mathbb{R}^3$. If we let $x_0 = v_0,\ x_1 = v_1,\ x_2 = (2, 0, 0)$, and $x_3 = (3, 0, 0)$, then it is easy to see that (x_0, x_1, x_2, x_3) is a point on S_Γ that is not proper since $\overline{X}_3 = \{x_1, x_2, x_3\}$ is not in general position. It is however clear in this example, and in fact also true in general (although we do not need this fact), that "generic points" of S_Γ will always be proper.

The basic properties of the spheres $S_j(X_j)$ and their geometric relationship with the sets X_j, specifically in the nontrivial case when $|X_j| \geq 2$, are collected in Lemma 3 below.

Lemma 3. *If $(x_0, x_1, \ldots, x_n)$ is a proper point of S_Γ, then for all $1 \leq j \leq n$ for which we have $|X_j| \geq 2$:*

(i) *The affine subspaces spanned by $S_j(X_j)$ and X_j respectively are orthogonal and hence $S_j(X_j)$ is a sphere of dimension $d - |X_j|$, which is at least 1 since $|X_j| \leq k$ and $d \geq k + 1$.*

(ii) *The radius r_j of the sphere $S_j(X_j)$ is positive, equal to the distance from any point $x \in S_j$ to the affine subspace spanned by X_j, and in fact depends continuously on the points in X_j.*

Proof. We first recall the elementary fact that the intersection of any two spheres in $\mathbb{R}^3$ is either empty or a circle lying in a plane perpendicular to the line joining the centers of the spheres. Since any four points in $\mathbb{R}^d$ with $d \geq 3$ span a three-dimensional affine subspace, it follows immediately from the elementary fact above that if we let $x \in S_j(X_j)$ and x_{i_1} and x_{i_2} be any two points in X_j, then the vectors $x - x_j$ and $x_{i_2} - x_{i_1}$ are orthogonal, and hence the affine subspaces spanned by $S_j(X_j)$ and X_j respectively are orthogonal.

Let c_j denote the projection of x_j onto the affine subspace spanned by X_j. We claim that c_j is the center of the sphere $S_j(X_j)$. Indeed, since c_j has the property that the vectors $x - c_j$ and $x_i - c_j$ are orthogonal for all $x \in S_j(X_j)$ and all $x_i \in X_j$, it follows that for any fixed $x_i \in X_j$ we have $|x_j - x_i|^2 = |x - x_i|^2$ for all $x \in S_j(X_j)$, and hence by Pythagoras that $|x - c_j| = |x_j - c_j|$ for all $x \in S_j(X_j)$.

The discussion above implies that the radius r_j of $S_j(X_j)$ is positive and equal to the distance from any point $x \in S_j$, so in particular x_j, to the affine subspace spanned by X_j. Specifically, if $X_j = \{x_{i_1}, \ldots, x_{i_\ell}\}$, then the fact that $\overline{X}_j = \{x_{i_1}, \ldots, x_{i_\ell}, x_j\}$ is in general position ensures that the volume of the ℓ-dimensional fundamental parallelotope determined by the vectors $\{x_j - x_{i_1}, \ldots, x_j - x_{i_\ell}\}$ is nonzero. It is a basic fact, see for example either Section 8.72 in [Shilov 1971] or Theorem 7 in Chapter X of [Birkhoff and Mac Lane 1941], that the volume of this parallelotope is equal to the square root of the so-called Gram determinant, namely the determinant of the (Gram) inner product matrix

$$\det\{(x_j - x_{i_{m_1}}) \cdot (x_j - x_{i_{m_2}})\}_{1 \leq m_1, m_2 \leq \ell}.$$

It thus follows that

$$r_j = \sqrt{\frac{\det\{(x_j - x_{i_{m_1}}) \cdot (x_j - x_{i_{m_2}})\}_{1 \leq m_1, m_2 \leq \ell}}{\det\{(x_{i_\ell} - x_{i_{m_1}}) \cdot (x_{i_\ell} - x_{i_{m_2}})\}_{1 \leq m_1, m_2 \leq \ell - 1}}}, \tag{7}$$

as r_j is the height of our parallelotope if we take its base to be the $(\ell-1)$-dimensional parallelotope determined by the vectors $\{x_{i_\ell} - x_{i_1}, \ldots, x_{i_\ell} - x_{i_{\ell-1}}\}$; see for example Section 8.72 in [Shilov 1971].

Since one can easily see, by expanding $|(x_j - x_{i_{m_2}}) - (x_j - x_{i_{m_1}})|^2$, that

$$\det\{(x_j - x_{i_{m_1}}) \cdot (x_j - x_{i_{m_2}})\}_{1 \leq m_1, m_2 \leq \ell} = \tfrac{1}{2}(t_{i_{m_1} j} + t_{i_{m_2} j} - |x_{i_{m_1}} - x_{i_{m_2}}|^2),$$

it follows that r_j, in addition to being positive, in fact depends continuously on the points in X_j. □

Finally we introduce compactly supported functions $\eta_j \in C^\infty(\mathbb{R}^d)$ with $0 \le \eta_j \le 1$ and $\eta_j(v_j) = 1$ for each $0 \le j \le n$. We further assume that the supports of each η_j have been chosen small enough to ensure that every $(x_0, x_1, \dots, x_n) \in S_\Gamma$ with $x_j \in \operatorname{supp} \eta_j$ corresponds to a δ-close isometric copy of Γ with vertex v_0 remaining fixed at 0.

An important consequence of Part (ii) of Lemma 3 above is that we may also assume that the supports of each η_j have been chosen small enough to ensure that there exists a constant $r_\Gamma > 0$ such that for each $(x_0, x_1, \dots, x_n) \in S_\Gamma$, with $x_j \in \operatorname{supp} \eta_j$, the corresponding spheres $S_j(X_j)$ will all have radius $r_j \ge r_\Gamma$.

Definition 4 (localized counting function). For any $0 < \lambda \ll 1$ and functions

$$f_0, f_1, \dots, f_n : [0,1]^d \to \mathbb{R},$$

with $d \ge k+1$, we define

$$T_{\Gamma,\delta}(f_0, f_1, \dots, f_n)(\lambda) = \iint \cdots \int f_0(x) f_1(x - \lambda x_1) \cdots f_n(x - \lambda x_n)\, d\mu_n(x_n) \cdots d\mu_1(x_1)\, dx, \tag{8}$$

where $d\mu_j(x_j) = \eta_j(x_j)\, d\sigma_j(x_j)$ and σ_j denotes the normalized surface measure on S_j.

Note that if $A \subseteq [0,1]^d$ and $T_{\Gamma,\delta}(1_A, 1_A, \dots, 1_A)(\lambda) > 0$, then A must contain a point configuration $\Gamma' = \{x, x + \lambda x_1, \dots, x + \lambda x_n\}$ with each $x_j \in S_j(X_j)$, and hence a δ-close isometric copy of $\lambda \cdot \Gamma$.

The key to showing that $T_{\Gamma,\delta}(1_A, 1_A, \dots, 1_A)(\lambda)$ is positive for certain sets A is to estimate (8) in terms of a suitable uniformity norm localized to a scale L (related to λ).

Definition 5 ($U^1(L)$-norm). For $0 < L \ll 1$ and functions $f : [0,1]^d \to \mathbb{R}$ we define

$$\|f\|_{U^1(L)} = \|f * \varphi_L\|_2,$$

where $\varphi_L(x) = L^{-d} \varphi(L^{-1} x)$, with $\varphi = 1_{[-1/2,1/2]^d}$.

Note that if $A \subseteq [0,1]^d$ with $\alpha = |A| > 0$ and we define $f_A := 1_A - \alpha 1_{[0,1]^d}$, then

$$\|f_A\|_{U^1(L)}^2 = \int_{\mathbb{R}^d} \left| \frac{|A \cap (t + Q_L)|}{|Q_L|} - |A| \right|^2 dt, \tag{9}$$

where $Q_L = [-L/2, L/2]^d$.

Evidently the $U^1(L)$-norm is measuring the mean-square uniform distribution of A on scale L. The engine that drives our approach to Theorem 2 is the following:

Proposition 6 (generalized von Neumann). *Let* $0 < \varepsilon, \lambda \ll 1$. *For any* $L \le \varepsilon^6 \lambda$, $0 \le m \le n$, *and functions*

$$f_0, f_1, \dots, f_m : [0,1]^d \to [-1,1],$$

we have

$$|T_{\Gamma,\delta}(f_0, f_1, \dots, f_m, 1, \dots, 1)(\lambda)| \le \|f_m\|_{U^1(L)} + O_\Gamma(\varepsilon).$$

Here 1 stands for the indicator function of the unit cube $[0,1]^d$ and $O_\Gamma(\varepsilon)$ means a quantity bounded by $C_\Gamma \varepsilon$ with C_Γ a constant depending only on Γ. We will also use the notation $f \ll_{\Gamma,\delta} g$ to indicate that $|f| \le c_{\Gamma,\delta}\, g$ with a constant $c_{\Gamma,\delta} > 0$, depending on only Γ and δ, that is *sufficiently small* for our purposes.

The above proposition immediately implies the following result for uniformly distributed sets from which we will deduce both parts of Theorem 2 in Section 3 below.

Corollary 7. *Let $\delta > 0$ and Γ be a proper k-degenerate distance graph on $n+1$ vertices in $\mathbb{R}^d$ with $d \geq k+1$.*

Let $\alpha \in (0, 1)$ and $0 < \lambda \leq \varepsilon \ll_{\Gamma,\delta} \alpha^{n+1}$. If $A \subseteq [0, 1]^d$ with $|A| = \alpha$ satisfies $\|f_A\|_{U^1(\varepsilon^6\lambda)} \ll \varepsilon$, then

$$T_{\Gamma,\delta}(1_A, 1_A, \dots, 1_A)(\lambda) \geq \tfrac{1}{2}c_0\alpha^{n+1},$$

where

$$c_0 = c_0(\Gamma, \delta) = \iint \cdots \int d\mu_n(x_n) \cdots d\mu_1(x_1)\, dx.$$

Note that in light of the assumptions that we have placed on the measures μ_j (via the functions η_j), the quantity $c_0(\Gamma, \delta)$ above clearly tends to zero as $\delta \to 0^+$.

Proof of Corollary 7. The result follows immediately from Proposition 6 since

$$T_{\Gamma,\delta}(1_A, \dots, 1_A)(\lambda) = c_0\alpha^{n+1} + \sum_{m=0}^{n} \alpha^{n-m}\, T_{\Gamma,\delta}(\underbrace{1_A, \dots, 1_A}_{m \text{ copies}}, f_A, 1, \dots, 1)(\lambda),$$

where $f_A = 1_A - \alpha 1_{[0,1]^d}$. □

We conclude this section with the proof of Proposition 6.

Proof of Proposition 6. Fix $0 \leq m \leq n$. We have

$$|T_{\Gamma,\delta}(f_0, f_1, \dots, f_m, 1, \dots, 1)(\lambda)| \leq \int \cdots \int \left(\int \left| \iint f_m(x - \lambda x_m) c_{m+1}(x_1, \dots, x_m)\, d\mu_m(x_m) \right| dx \right) d\mu_{m-1}(x_{m-1}) \cdots d\mu_1(x_1),$$

where

$$c_{m+1}(\Gamma, \delta; x_1, \dots, x_m) = \int \cdots \int d\mu_n(x_n) \cdots d\mu_{m+1}(x_{m+1}) \tag{10}$$

if $0 \leq m \leq n-1$ and $c_{n+1} = 1$. It follows from an application of Cauchy–Schwarz and Plancherel that

$$|T_{\Gamma,\delta}(f_0, f_1, \dots, f_m, 1, \dots, 1)(\lambda)|^2 \leq \int |\hat{f}_m(\xi)|^2 I_m(\lambda\xi)\, d\xi, \tag{11}$$

where

$$I_m(\xi) = \int \cdots \int \left|\widehat{c_{m+1}\mu_m}(\xi)\right|^2 d\mu_{m-1}(x_{m-1}) \cdots d\mu_1(x_1), \tag{12}$$

with

$$\widehat{c_{m+1}\mu_m}(\xi) = \int c_{m+1}(x_1, \dots, x_m)\eta_m(x_m)\, e^{-2\pi i x_m \cdot \xi}\, d\sigma_m(x_m)$$

if $2 \leq m \leq n$ and $I_1 = |\widehat{c_2\mu_1}|^2$. In light of the trivial uniform bound $0 \leq I_m(\xi) \leq 1$ and the fact that

$$\|f_m\|^2_{U^1(L)} = \int |\hat{f}_m(\xi)|^2 |\hat{\varphi}(L\xi)|^2\, d\xi,$$

it suffices to establish that

$$I_m(\lambda\xi)(1 - \hat{\varphi}(L\xi)^2) = O_\Gamma(\varepsilon^2). \tag{13}$$

Since $0 \le \hat{\varphi}(\xi)^2 \le 1$ for all $\xi \in \mathbb{R}^d$ and $\hat{\varphi}(0) = 1$ it follows that $0 \le 1 - \hat{\varphi}(L\xi)^2 \le \min\{1, 4\pi L|\xi|\}$. The uniform bound (13) thus reduces to establishing the decay estimate

$$I_m(\xi) \le \min\{1, C_\Gamma |\xi|^{-1/2}\} \tag{14}$$

since this would in turn imply that

$$I_m(\lambda\xi)(1 - \hat{\varphi}(L\xi)^2) \le C_\Gamma \min\{(\lambda|\xi|)^{-1/2}, \varepsilon^6\lambda|\xi|\} \le C_\Gamma \varepsilon^2$$

whenever $L \le \varepsilon^6\lambda$.

To establish (14) we will use the fact that in addition to being trivially bounded by 1, the Fourier transform of $c_{m+1}\mu_m$ also decays for large ξ in certain directions, specifically

$$|\widehat{c_{m+1}\mu_m}(\xi)| \le \min\{1, (r_\Gamma \cdot (\operatorname{dist}(\xi, \operatorname{span} X_m))^{-1/2}\} \tag{15}$$

uniformly over all $x_1, \dots, x_{m-1}$ with $x_j \in \operatorname{supp} \eta_j$. This estimate is an easy consequence of the well-known asymptotic behavior of the Fourier transform of the measure on the unit sphere $S^{d-|X_m|} \subseteq \mathbb{R}^{d-|X_m|+1}$ induced by Lebesgue measure; see for example [Stein 1993].

Using the fact that the measure $d\sigma_{m-1}(x_{m-1}) \cdots d\sigma_1(x_1)$ is invariant under the rotation

$$(x_1, \dots, x_m) \to (Ux_1, \dots, Ux_m),$$

for any $U \in \mathrm{SO}(d)$, together with (15) and the fact that $0 \le \eta_j \le 1$ for $1 \le j \le m$, then gives

$$\begin{aligned} I_m(\xi) &\le C \int \cdots \int (1 + r_\Gamma \cdot \operatorname{dist}(\xi, \operatorname{span} X_m))^{-1}\, d\sigma_{m-1}(x_{m-1}) \cdots d\sigma_1(x_1) \\ &= C \int \cdots \iint_{\mathrm{SO}(d)} (1 + r_\Gamma \cdot \operatorname{dist}(\xi, \operatorname{span} UX_m))^{-1}\, d\mu(U)\, d\sigma_{m-1}(x_{m-1}) \cdots d\sigma_1(x_1) \\ &= C \int \cdots \iint_{S^{d-1}} (1 + r_\Gamma|\xi| \cdot \operatorname{dist}(y, \operatorname{span} X_m))^{-1}\, d\sigma(y)\, d\sigma_{m-1}(x_{m-1}) \cdots d\sigma_1(x_1), \end{aligned} \tag{16}$$

where σ denotes normalized measure on the unit sphere S^{d-1} in $\mathbb{R}^d$ induced by Lebesgue measure. Estimate (14) then follows from the easy observation that the inner integral above satisfies the uniform estimate

$$\int_{S^{d-1}} (1 + r_\Gamma|\xi| \cdot \operatorname{dist}(y, \operatorname{span} X_m))^{-1}\, d\sigma(y) = O((1 + r_\Gamma|\xi|)^{-1/2}). \qquad \square$$

3. Proof of Theorem 2

We will deduce Theorem 2 from Corollary 7 by localizing to cubes on which our set is suitably uniformly distributed. In the case of Part (i) this is achieved as a direct consequence of the definition of upper Banach density, while for Part (ii) this is achieved via an energy increment argument.

3.1. ***Direct proof of Part (i) of Theorem 2.*** Let $\varepsilon > 0$ and $A \subseteq \mathbb{R}^d$ with $\delta^*(A) > 0$.

The following two facts follow immediately from the definition of upper Banach density, see (1):

(i) There exist $M_0 = M_0(A, \varepsilon)$ such that for all $M \geq M_0$ and all $t \in \mathbb{R}^d$

$$\frac{|A \cap (t + Q_M)|}{|Q_M|} \leq \left(1 + \tfrac{1}{3}\varepsilon^4\right) \delta^*(A).$$

(ii) There exist arbitrarily large $N \in \mathbb{R}$ such that

$$\frac{|A \cap (t_0 + Q_N)|}{|Q_N|} \geq \left(1 - \tfrac{1}{3}\varepsilon^4\right) \delta^*(A)$$

for some $t_0 \in \mathbb{R}^d$.

Combining (i) and (ii) above we see that for any $\lambda \geq \lambda_0 := \varepsilon^{-6} M_0$, there exist $N \geq \varepsilon^{-6}\lambda$ and $t_0 \in \mathbb{R}^d$ such that

$$\frac{|A \cap (t + Q_{\varepsilon^6\lambda})|}{|Q_{\varepsilon^6\lambda}|} \leq (1 + \varepsilon^4) \frac{|A \cap (t_0 + Q_N)|}{|Q_N|}$$

for all $t \in \mathbb{R}^d$. Consequently, Theorem 2 reduces, via a rescaling of $A \cap (t_0 + Q_N)$ to a subset of $[0, 1]^d$, to establishing that if Γ is a proper k-degenerate distance graph, $0 < \lambda \leq \varepsilon \ll 1$, and $A \subseteq [0, 1]^d$ is measurable with $|A| > 0$ and the property that

$$\frac{|A \cap (t + Q_{\varepsilon^6\lambda})|}{|Q_{\varepsilon^6\lambda}|} \leq (1 + \varepsilon^4)\,|A|$$

for all $t \in \mathbb{R}^d$, then A contains an isometric copy of $\lambda \cdot \Gamma$.

Now since $A \cap (t + Q_{\varepsilon^6\lambda})$ is only supported in $[-\varepsilon^6\lambda, 1 + \varepsilon^6\lambda]^d$ and

$$|A| = \int_{\mathbb{R}^d} \frac{|A \cap (t + Q_{\varepsilon^6\lambda})|}{|Q_{\varepsilon^6\lambda}|}\, dt$$

it easily follows that

$$\left|\left\{ t \in \mathbb{R}^d : 0 < \frac{|A \cap (t + Q_{\varepsilon^6\lambda})|}{|Q_{\varepsilon^6\lambda}|} \leq (1 - \varepsilon^2)\,|A| \right\}\right| = O(\varepsilon^2)$$

and hence that

$$\| f_A \|^2_{U^1(\varepsilon^6\lambda)} = \int_{\mathbb{R}^d} \left| \frac{|A \cap (t + Q_{\varepsilon^6\lambda})|}{|Q_{\varepsilon^6\lambda}|} - |A| \right|^2 dt = O(\varepsilon^2).$$

The result thus follows from Corollary 7 above provided $\varepsilon \ll_{\Gamma,\delta} \delta^*(A)^{n+1}$. □

3.2. *Proof of Part (ii) of Theorem 2.*

Lemma 8 (localization principle). *Let $A \subseteq [0, 1]^d$ with $d \geq k + 1$ and $|A| = \alpha > 0$.*

Let $\varepsilon > 0$ and $\varepsilon^7 \gg L_1 \gg L_2 \gg \cdots$ be any decreasing sequence with $L_1^{-1} \in \mathbb{N}$ and $L_{j+1} \leq c\varepsilon^7 L_j$ with $L_{j+1} \mid L_j$ for all $j \geq 1$. If we let $\mathcal{G}_j$ denote the partition of $[0, 1]^d$ into cubes of side length L_j, then there exists $1 \leq j \leq C\varepsilon^{-2}$ such that for all but at most εL_j^{-d} of the cubes Q in $\mathcal{G}_j$ the set A will be uniformly distributed on the smaller scale L_{j+1} inside Q in the sense that

$$\frac{1}{|Q|} \int_Q \left| \frac{|A \cap Q \cap (t + Q_{L_{j+1}})|}{|Q_{L_{j+1}}|} - \frac{|A \cap Q|}{|Q|} \right|^2 dt \leq \varepsilon. \tag{17}$$

Before proving Lemma 8 we first show that it, together with Corollary 7 (after rescaling), is sufficient to establish Part (ii) of Theorem 2. Let $\varepsilon \ll_{\Gamma,\delta} \alpha^{n+1}$ and $\{Q_i\}$ denote the cubes of side length L_j in the partition $\mathcal{G}_j$ of $[0,1]^d$ that we obtain from Lemma 8. If we then let $A_i = A \cap Q_i$ and set $\alpha_i = |A\cap Q_i|/|Q_i|$, it follows from Corollary 7 (after rescaling) and Hölder's inequality that for any $\lambda \in (\varepsilon^{-6}L_{j+1}, \varepsilon L_j)$ we have

$$T_{\Gamma,\delta}(1_A,\dots,1_A)(\lambda) \geq \sum_{i=1}^{L_j^{-d}} T_{\Gamma,\delta}(1_{A_i},\dots,1_{A_i})(\lambda) \geq \tfrac{1}{4}c_0 L_j^d \sum_{i=1}^{L_j^{-d}} \alpha_i^{n+1} \geq \tfrac{1}{4}c_0\Bigl(L_j^d \sum_{i=1}^{L_j^{-d}} \alpha_i\Bigr)^{n+1} = \tfrac{1}{4}c_0|A|^{n+1}. \quad (18)$$

Proof of Lemma 8. Let $\{Q_i\}$ denote the cubes of side length L_j in the partition $\mathcal{G}_j$ of $[0,1]^d$ and

$$g_j = 1_A - \mathbb{E}(1_A \mid \mathcal{G}_j),$$

where

$$\mathbb{E}(1_A \mid \mathcal{G}_j)(x) = \frac{|A\cap Q_i|}{|Q_i|}$$

for each $x \in Q_i$. If $\|g_j\|_{U^1(L_{j+1})} \geq \varepsilon$, then by definition

$$\int \Bigl| \frac{1}{|Q_{L_{j+1}}|} \int_{x+Q_{L_{j+1}}} g_j(y)\,dy \Bigr|^2 dx \geq c\varepsilon^2.$$

It follows that there must exist a $x_0 \in [0,1]^d$ for which the shifted grid $x_0 + \mathcal{G}_{j+1}$ satisfies

$$\int |\mathbb{E}(g_j \mid x_0 + \mathcal{G}_{j+1})|^2\,dx \geq c\varepsilon^2,$$

from which one can easily conclude that the (unshifted) refined grid $\mathcal{G}_{j+2}$ satisfies

$$\int |\mathbb{E}(g_j \mid \mathcal{G}_{j+2})|^2\,dx \geq c\varepsilon^2 \quad (19)$$

provided $L_{j+2} \ll \varepsilon^2 L_{j+1}$. By orthogonality, it follows immediately from (19) and the definition of g_j that

$$\|\mathbb{E}(1_A \mid \mathcal{G}_{j+2})\|_2^2 \geq \|\mathbb{E}(1_A \mid \mathcal{G}_j)\|_2^2 + c\varepsilon^2 \quad (20)$$

and hence that there must exist $1 \leq j \leq C\varepsilon^{-2}$ such that $\|g_j\|_{U^1(L_{j+1})} \leq \varepsilon$. From this it follows that

$$\sum_{i=1}^{L_j^{-d}} \int \Bigl| \frac{1}{|Q_{L_{j+1}}|} \int_{x+Q_{L_{j+1}}} (1_{A_i} - \alpha_i 1_{Q_i})(y)\,dy \Bigr|^2 dx \leq C\varepsilon^2$$

provided $L_{j+1} \ll \varepsilon^2 L_j$. □

4. A second proof of Theorem 2

Let $\delta > 0$ and Γ be a proper k-degenerate distance graph in $[0,1]^d$ with $d \geq k+1$. We shall make use of the same notation as in Section 2, specifically for the counting function $T_{\Gamma,\delta}$ as defined in (8), and make the same assumptions on the measures μ_j (via the functions η_j).

4.1. ***Reducing Theorem 2 to a dichotomy between randomness and structure.*** As we shall see, Theorem 2 is an immediate consequence of the following proposition which reveals that if $A \subseteq [0,1]^d$ has positive measure but does not contain an isometric copy of $\lambda \cdot \Gamma$ for all λ in a given interval, then this "nonrandom" behavior is detected by the Fourier transform of the characteristic function of A and results in "structural information", specifically a concentration of its L^2-mass on appropriate annuli.

Proposition 9 (dichotomy). *Let $\delta > 0$ and Γ be a proper k-degenerate distance graph in $[0,1]^d$ with $d \geq k+1$.*

If $A \subseteq [0,1]^d$ with $|A| > 0$, $0 < a \leq b \ll \varepsilon^4$ with $0 < \varepsilon \ll_{\Gamma,\delta} |A|^{n+1}$, and A does not *contain a δ-close isometric copy of $\lambda \cdot \Gamma$ for some λ in $[a,b]$, then*

$$\int_{\varepsilon^2/b \leq |\xi| \leq 1/\varepsilon^2 a} |\hat{1}_A(\xi)|^2 \, d\xi \gg c_1^2 |A|^{2n+2}, \tag{21}$$

with the implied constant above independent of a, b, and ε, and

$$c_1 = c_1(\Gamma, \delta) = \int \cdots \int d\mu_n(x_n) \cdots d\mu_1(x_1).$$

Proof that Proposition 9 implies Theorem 2. We shall first establish Part (ii) of Theorem 2, so we start by letting $A \subseteq [0,1]^d$ with $|A| > 0$. For any fixed $0 < \varepsilon \ll_{\Gamma,\delta} |A|^{n+1}$, let $\{\mathcal{I}_j\}_{j=1}^{J(\varepsilon)}$ denote a sequence of intervals with $\mathcal{I}_j := [a_j, b_j]$ satisfying

$$b_{j+1} \ll \varepsilon^4 a_j \tag{22}$$

and $b_1 \ll \varepsilon^4$ with the property that for each $1 \leq j \leq J(\varepsilon)$ there exists a $\lambda \in \mathcal{I}_j$ such that

$$x + \lambda \cdot U(\Delta) \nsubseteq A \tag{23}$$

for all $x \in A$ and $U \in \mathrm{SO}(d)$. Proposition 9, together with (22), would then imply

$$J(\varepsilon)\varepsilon^2 \leq \sum_{j=1}^{J(\varepsilon)} \int_{\varepsilon^2/b_j \leq |\xi| \leq 1/\varepsilon^2 a_j} |\hat{1}_A(\xi)|^2 \, d\xi \leq \int |\hat{1}_A(\xi)|^2 \, d\xi, \tag{24}$$

a contradiction if $J(\varepsilon) \gg \varepsilon^{-2}$ since by Plancherel we know that $\int |\hat{1}_A(\xi)|^2 \, d\xi = |A| \leq 1$.

To establish Part (i) of Theorem 2 with this approach we will argue indirectly and thus suppose that $A \subseteq \mathbb{R}^d$ is a set with $\delta^*(A) > 0$ for which the conclusion of Part (i) of Theorem 2 fails to hold, namely that there exist arbitrarily large $\lambda \in \mathbb{R}$ for which A does not contain an isometric copy of $\lambda \cdot \Gamma$.

We now let $0 < \alpha < \delta^*(A)$, $0 < \varepsilon \ll_{\Gamma,\delta} \alpha^{n+1}$, and fix $J \gg \varepsilon^{-2}$ as above. By our indirect assumption we can choose a sequence $\{\lambda_j\}_{j=1}^J$ with the property that $\lambda_{j+1} \ll \varepsilon^4 \lambda_j$ for all $1 \leq j \leq J-1$ and A does not contain an isometric copy of $\lambda_j \cdot \Gamma$ for each $1 \leq j \leq J$. It follows from the definition of upper Banach density that there exist $N \in \mathbb{R}$ with $N \gg \lambda_1$ and $t_0 \in \mathbb{R}^d$ for which

$$\frac{|A \cap (t_0 + Q_N)|}{|Q_N|} \geq \alpha.$$

Rescaling $A \cap (t_0 + Q_N)$ to a subset of $[0,1]^d$ and arguing as in the proof of Part (ii) above but this time with $b_j = \lambda_j / N$ again leads to a contradiction. □

4.2. ***Proof of Proposition 9.*** Let $f = 1_A$. We will utilize the existence of a suitably *smoothed* version of f with the certain properties, specifically:

Lemma 10. *For any $\varepsilon > 0$ there exists a function $g : \mathbb{R}^d \to (0, 1]$, an appropriate smoothing of f, such that*

$$|g(x - \lambda z) - g(x)| \ll \varepsilon \tag{25}$$

uniformly in $x \in [0, 1]^d$ and $|z| \le 1$. Moreover, if $\varepsilon \ll |A|^{n+1}$, then

$$\int f(x) g(x)^n \, dx \gg |A|^{n+1}. \tag{26}$$

The proof of Lemma 10 is straightforward and presented in Section 4.3 below. Assuming for now the existence of a function g with property (25) it follows that

$$T_{\Gamma,\delta}(f, f, \dots, f)(\lambda) = c_1 \int f(x) g(x)^n \, dx + \sum_{m=1}^{n} T_{\Gamma,\delta}(f g^{n-m}, f, \dots, f, f - g, \underbrace{1, \dots, 1}_{n-m \text{ copies}})(\lambda) + O(n\varepsilon). \tag{27}$$

If A does *not* contain a δ-close isometric copy of $\lambda \cdot \Gamma$ for some λ in $[a, b]$, then it clearly follows that

$$T_{\Gamma,\delta}(f, f, \dots, f)(\lambda) = 0.$$

In light of (26) and (27) it follows that if $\varepsilon \ll c_1 |A|^{n+1}/n$ then there must exist $1 \le m \le n$ such that

$$\int \cdots \int \left(\int \left| \int [f - g](x - \lambda x_m) c_{m+1}(x_1, \dots, x_m) \, d\mu_m(x_m) \right| dx \right) d\mu_{m-1}(x_{m-1}) \cdots d\mu_1(x_1) \gg c_1 |A|^{n+1}, \tag{28}$$

with c_{m+1} defined as before in (10) above. It then follows from an application of Cauchy–Schwarz and Plancherel that

$$\int |\hat{f}(\xi) - \hat{g}(\xi)|^2 I_m(\lambda \xi) \, d\xi \gg c_1^2 |A|^{2n+2}, \tag{29}$$

with I_m again defined as before in (12) above. The fact that g will be taken to be a sufficient smoothing of f ensures that its Fourier transform satisfies

$$|\hat{f}(\xi) - \hat{g}(\xi)| \le \varepsilon |\hat{f}(\xi)| \tag{30}$$

provided $|\xi| \le \varepsilon^2 b^{-1}$; see Section 4.3 below. This, together with the fact that $I_m(\xi)$ is bounded by 1 uniformly in ξ, and Plancherel, ensures that (29) implies

$$\int_{\varepsilon^2/b \le |\xi|} |\hat{f}(\xi)|^2 I_m(\lambda \xi) \, d\xi \gg c_1^2 |A|^{2n+2} \tag{31}$$

provided $\varepsilon \ll_{\Gamma,\delta} |A|^{n+1}$. Estimate (21), and hence Proposition 9, then follows easily from estimate (31) and our previously established estimates for I_m, namely (14).

4.3. *A smooth cutoff function and proof of Lemma 10.*

4.3.1. *A smooth cutoff function.* Let $\psi : \mathbb{R}^d \to (0, \infty)$ be a Schwartz function that satisfies

$$1 = \hat{\psi}(0) \geq \hat{\psi}(\xi) \geq 0 \qquad \text{and} \qquad \hat{\psi}(\xi) = 0 \quad \text{for } |\xi| > 1.$$

As usual, for any given $t > 0$, we define

$$\psi_t(x) = t^{-d}\psi(t^{-1}x). \tag{32}$$

First we record the trivial observation that

$$\int \psi_t(x)\,dx = \int \psi(x)\,dx = \hat{\psi}(0) = 1, \tag{33}$$

as well as the simple, but important, observation that ψ may be chosen so that

$$|1 - \hat{\psi}_t(\xi)| = |1 - \hat{\psi}(t\xi)| \ll \min\{1, t|\xi|\}. \tag{34}$$

Finally we record a formulation, appropriate to our needs, of the fact that for any given small parameter ε, our cutoff function $\psi_t(x)$ will be essentially supported where $|x| \leq \varepsilon^{-1}t$ and is approximately constant on smaller scales. More precisely:

Lemma 11. *Let* $\varepsilon > 0$ *and* $t > 0$; *then*

$$\int_{|y|\geq\varepsilon^{-1}t} \psi_t(y)\,dy \ll \varepsilon, \tag{35}$$

$$\int \bigl|\psi_t(y - \lambda z) - \psi_t(y)\bigr|\,dy \ll \varepsilon \tag{36}$$

uniformly for $|z| \leq 1$, *provided* $t \gg \varepsilon^{-1}\lambda$.

Proof of Lemma 11. Estimate (35) is easily verified using the fact that ψ is a Schwartz function on $\mathbb{R}^d$ as

$$\int_{|y|\geq\varepsilon^{-1}t} \psi_t(y)\,dy = \int_{|y|\geq\varepsilon^{-1}} \psi(y)\,dy \ll \int_{|y|\geq\varepsilon^{-1}} (1+|y|)^{-d-1}\,dy \ll \varepsilon.$$

To verify estimate (36) we make use of the fact that both ψ and its derivative are rapidly decreasing, specifically

$$\begin{aligned}\int \bigl|\psi_t(y - \lambda z) - \psi_t(x)\bigr|\,dy &\leq \int \bigl|\psi(y - \lambda z/t) - \psi(y)\bigr|\,dy \\ &\ll \frac{\lambda}{t}\int (1+|y|)^{-d-1}\,dy \ll \frac{\lambda}{t}.\end{aligned}$$ □

4.3.2. *Proof of Lemma 10.* Let $g = f * \psi_{\varepsilon^{-1}b}$.

We first note that estimates (30) and (25) follow immediately from (34) and (36) respectively. In order to establish the remaining "main term" estimate (26), we need only establish that if $\varepsilon \ll |A|^{n+1}$, then

$$\int f(x)g(x)\,dx \geq (1 - C\varepsilon)\,|A|^2 \tag{37}$$

for some constant $C > 0$, since by Hölder we would then obtain

$$(1-C\varepsilon)^n |A|^{2n} \le \left(\int f(x)g(x)\,dx\right)^n \le |A|^{n-1} \int f(x)g(x)^n\,dx,$$

from which (26) clearly follows for sufficiently small $\varepsilon > 0$.

To establish (37) we first note that Parseval, the fact that $0 \le \hat{\psi} \le 1$, and a final application of Cauchy–Schwarz give

$$\begin{aligned}\int f(x)g(x)\,dx &= \int |\hat{f}(\xi)|^2 \hat{\psi}(\varepsilon^{-1}b\xi)\,d\xi \\ &\ge \int |\hat{f}(\xi)|^2 |\hat{\psi}(\varepsilon^{-1}b\xi)|^2\,d\xi = \int g(x)^2\,dx \ge \left(\int_{[0,1]^d} g(x)\,dx\right)^2. \qquad (38)\end{aligned}$$

Establishing (37) therefore reduces to showing that if $\varepsilon \ll |A|^3$, then

$$\int_{[0,1]^d} g(x)\,dx \ge (1-C\varepsilon)|A| \qquad (39)$$

for some constant $C > 0$. To establish (39) we use (33) and write

$$\begin{aligned}|A| &= \int_{\mathbb{R}^d} g(x)\,dx \\ &= \int_{[0,1]^d} g(x)\,dx + \int_{\{x\in\mathbb{R}^d:\mathrm{dist}(x,[0,1]^d)\ge\varepsilon^{-2}b\}} g(x)\,dx + \int_{\{x\in\mathbb{R}^d:0<\mathrm{dist}(x,[0,1]^d)<\varepsilon^{-2}b\}} g(x)\,dx. \qquad (40)\end{aligned}$$

The fact that $b \le \varepsilon^4$ ensures that

$$|\{x \in \mathbb{R}^d : 0 < \mathrm{dist}(x,[0,1]^d) < \varepsilon^{-2}b\}| \ll \varepsilon^2 \qquad (41)$$

and hence, since $\varepsilon \ll |A|$ and $0 \le g \le 1$, that

$$\int_{\{x\in\mathbb{R}^d:0<\mathrm{dist}(x,[0,1]^d)<\varepsilon^{-2}b\}} g(x)\,dx \ll \varepsilon^2 \le \varepsilon|A|,$$

while (35) ensures that

$$\int_{\{x\in\mathbb{R}^d:\mathrm{dist}(x,[0,1]^d)\ge\varepsilon^{-2}b\}} g(x)\,dx \le |A| \int_{|y|\gg\varepsilon^{-2}b} \psi_{\varepsilon^{-1}b}(y)\,dy \ll \varepsilon|A|, \qquad (42)$$

which completes the proof. □

References

[Bennett et al. 2016] M. Bennett, A. Iosevich, and K. Taylor, "Finite chains inside thin subsets of $\mathbb{R}^d$", *Anal. PDE* **9**:3 (2016), 597–614. MR Zbl

[Birkhoff and Mac Lane 1941] G. Birkhoff and S. Mac Lane, *A survey of modern algebra*, Macmillan, New York, 1941. 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

[Bulinski 2018] K. Bulinski, "Spherical recurrence and locally isometric embeddings of trees into positive density subsets of $\mathbb{Z}^d$", *Math. Proc. Cambridge Philos. Soc.* **165**:2 (2018), 267–278. MR Zbl

[Cook et al. 2017] B. Cook, Á. Magyar, and M. Pramanik, "A Roth-type theorem for dense subsets of $\mathbb{R}^d$", *Bull. Lond. Math. Soc.* **49**:4 (2017), 676–689. MR Zbl

[Furstenberg et al. 1990] H. Furstenberg, 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

[Iosevich and Parshall 2018] A. Iosevich and H. Parshall, "Embedding distance graphs in finite field vector spaces", preprint, 2018. arXiv

[Lyall and Magyar 2018] N. Lyall and Á. Magyar, "Product of simplices and sets of positive upper density in $\mathbb{R}^d$", *Math. Proc. Cambridge Philos. Soc.* **165**:1 (2018), 25–51. MR Zbl

[Lyall et al. 2018] N. Lyall, Á. Magyar, and H. Parshall, "Spherical configurations over finite fields", preprint, 2018, available at https://tinyurl.com/lyallpdf. To appear in *Amer. J. Math.* **142**: 2 (2020).

[Shilov 1971] G. E. Shilov, *Linear algebra*, Prentice-Hall, Englewood Cliffs, NJ, 1971. MR Zbl

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Math. Series **43**, Princeton Univ. Press, 1993. 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

[Tao and Vu 2006] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Adv. Math. **105**, Cambridge Univ. Press, 2006. MR Zbl

[Ziegler 2006] T. Ziegler, "Nilfactors of $\mathbb{R}^m$-actions and configurations in sets of positive upper density in $\mathbb{R}^m$", *J. Anal. Math.* **99** (2006), 249–266. MR Zbl

Received 23 Mar 2018. Revised 11 Jan 2019. Accepted 13 Mar 2019.

NEIL LYALL: lyall@math.uga.edu
Department of Mathematics, The University of Georgia, Athens, GA, United States

ÁKOS MAGYAR: magyar@math.uga.edu
Department of Mathematics, The University of Georgia, Athens, GA, United States

ANALYSIS AND PDE
Vol. 13, No. 3, 2020
dx.doi.org/10.2140/apde.2020.13.701

ISOLATED SINGULARITIES FOR SEMILINEAR ELLIPTIC SYSTEMS WITH POWER-LAW NONLINEARITY

MARIUS GHERGU, SUNGHAN KIM AND HENRIK SHAHGHOLIAN

We study the system $-\Delta \boldsymbol{u} = |\boldsymbol{u}|^{\alpha-1}\boldsymbol{u}$ with $1 < \alpha \le \frac{n+2}{n-2}$, where $\boldsymbol{u} = (u_1, \dots, u_m)$, $m \ge 1$, is a C^2 nonnegative function that develops an isolated singularity in a domain of $\mathbb{R}^n$, $n \ge 3$. Due to the multiplicity of the components of $\boldsymbol{u}$, we observe a new Pohozaev invariant different than the usual one in the scalar case. Aligned with the classical theory of the scalar equation, we classify the solutions on the whole space as well as the punctured space, and analyze the exact asymptotic behavior of local solutions around the isolated singularity. On a technical level, we adopt the method of moving spheres and the balanced-energy-type monotonicity functionals.

1. Introduction

1A. ***Background.*** This paper concerns the analysis of singular solutions to semilinear elliptic systems with power-law nonlinearity of type

$$-\Delta \boldsymbol{u} = |\boldsymbol{u}|^{\alpha-1}\boldsymbol{u}, \tag{1-1}$$

where $1 < \alpha \le \frac{n+2}{n-2}$, and $\boldsymbol{u} = (u_1, \dots, u_m)$, $m \ge 1$, is a C^2 vector-valued function defined on a domain in $\mathbb{R}^n$, $n \ge 3$. Our primary interest is in the case when each component of $\boldsymbol{u}$ is nonnegative and the domain is of the form $B_R \setminus \{0\}$, with B_R being the ball of radius R centered at the origin. It is by now well known that in cylindrical coordinates $t = -\log|x| \in \mathbb{R}$ and $\theta = x/|x| \in \mathbb{S}^{n-1}$, the transformation

$$\boldsymbol{u}(x) = |x|^{-\frac{2}{\alpha-1}} \boldsymbol{v}\left(-\log|x|, \frac{x}{|x|}\right) \tag{1-2}$$

MSC2010: primary 35J61; secondary 35J75, 35B40, 35C20.
Keywords: elliptic system, isolated singularity, asymptotic behavior, Pohozaev invariant.

yields the system

$$\partial_{tt}\boldsymbol{v} + \mu\partial_t\boldsymbol{v} + \Delta_\theta\boldsymbol{v} - \lambda\boldsymbol{v} + |\boldsymbol{v}|^{\alpha-1}\boldsymbol{v} = 0 \tag{1-3}$$

in $(-\log R, \infty) \times \mathbb{S}^{n-1}$, and vice versa, where Δ_θ is the Laplace–Beltrami operator on $\mathbb{S}^{n-1}$ and λ and μ are the constants fixed throughout this paper by

$$\lambda = \frac{2}{\alpha-1}\Big(n-2-\frac{2}{\alpha-1}\Big), \quad \mu = \frac{4}{\alpha-1} - n + 2. \tag{1-4}$$

The scalar case of this system was introduced in [Lane 1870] and later studied in [Emden 1907] to describe distribution of mass densities in spherical polytropic star in hydrostatic equilibrium. Since its birth, this equation has been used in many applications such as astrophysics, kinetic theory, and quantum mechanics; see [Goenner and Havas 2000]. The Lane–Emden equation has thus been subject to intensive studies in the last few decades and nowadays there is a vast amount of literature treating many aspects of the solutions to this equation and its diverse varieties.

One of the central questions[1] and a technically difficult problem for differential equations and systems is the study of the singular solutions, that is, solutions that develop singularities. In the scalar case, the classical and subsequent works have considered the asymptotic behavior of the solutions close to isolated singularities, with an accurate description of the asymptotic behavior of solutions around such singular points; see, e.g., [Aviles 1983; 1987; Bidaut-Véron and Véron 1991; Chen and Li 1991; Caffarelli et al. 1989; Gidas and Spruck 1981a; 1981b; Korevaar et al. 1999; Véron 1981; 1996].

The system (1-1) can be considered as a generalization of the Lane–Emden equation, and can also be viewed as a strongly coupled system of nonlinear Schrödinger equations (or more precisely the limiting system of the associated blowup solutions). In the latter point of view, there has been some development regarding classification of the global solutions, and compactness of the blowup sequence; see for instance [Chen and Lin 2015; Druet et al. 2010]. In the former point of view, there are many other types of generalizations, among which the Lane–Emden–Fowler systems have received considerable attention. Among possible references, we refer to [Bidaut-Véron and Raoux 1996; Bidaut-Véron and Grillot 1999; Bidaut-Véron and Giacomini 2010; Busca and Manásevich 2002; de Figueiredo and Felmer 1994; Poláčik et al. 2007; Serrin and Zou 1996] for the classification of global solutions, nonexistence theory of singular, positive solutions and local estimates of solutions to the Lane–Emden–Fowler systems. We refer to [Reichel and Zou 2000; Zou 2006] for more general cooperative elliptic systems. One may also consult to [de Figueiredo 2008] for a general theory regarding semilinear elliptic systems. To the best of the authors' knowledge, this is the first paper that conducts a thorough analysis on the qualitative behavior of the system (1-1), particularly regarding the classification of the solutions on the punctured space $\mathbb{R}^n \setminus \{0\}$ with respect to the balanced-energy-type functionals $\big($subcritical case $1 < \alpha < \frac{n+2}{n-2}\big)$ and the Pohozaev identities $\big($critical case $\alpha = \frac{n+2}{n-2}\big)$, as well as the asymptotic behavior of local solutions around the isolated singularities.

[1] To the best of our knowledge there are three central questions in this area. The other two questions refer to the structure of singular sets, see [Pacard 1993], and nonexistence theory, see [Grigor'yan and Sun 2014; Souplet 2009].

The key difference between the system (1-1) and its scalar version is, of course, the multiplicity of the components. The major observation in this paper is that the system (1-1) turns out to be very sensitive to the setting of multiple components in the case of the upper critical exponent $\bigl($that is, $\alpha = \frac{n+2}{n-2}\bigr)$ and lower critical exponent $\bigl($that is, $\alpha = \frac{n}{n-2}\bigr)$. Specifically, in the upper critical case $\alpha = \frac{n+2}{n-2}$, we discover a new Pohozaev invariant different than the usual one. The lower critical case is rather technical and we shall present the discussion on this issue in Section 7D.

Let us briefly illustrate how the new Pohozaev invariant comes into play in the analysis of the system (1-1) in the upper critical case. For the sake of clarity, let us assume that the solution $\boldsymbol{u}$ is rotationally symmetric, so that the cylindrical transformation $\boldsymbol{v}$ is a function of t only. After some manipulation, one can obtain the usual Pohozaev identity,

$$\left|\frac{d\boldsymbol{v}}{dt}\right|^2 = \frac{(n-2)^2}{4}|\boldsymbol{v}|^2 - \frac{n-2}{n}|\boldsymbol{v}|^{\frac{2n}{n-2}} + \kappa \tag{1-5}$$

for the system (1-3), with a constant κ, also known as the usual Pohozaev invariant. Due to the presence of the multiple components, we have

$$\left|\frac{d\boldsymbol{v}}{dt}\right|^2 - \left(\frac{d|\boldsymbol{v}|}{dt}\right)^2 = \frac{1}{|\boldsymbol{v}|^2}\sum_{1\le i<j\le m}\left(v_i\frac{dv_j}{dt} - v_j\frac{dv_i}{dt}\right)^2 \ge 0, \tag{1-6}$$

and the equality on the rightmost side does not hold in general. This shows that κ alone is not enough to analyze the behavior of $|\boldsymbol{v}|$, due to the discrepancy (1-6) between $|d\boldsymbol{v}/dt|$ and $|d|\boldsymbol{v}|/dt|$. In this paper, we find that there is another constant κ_* such that

$$\left(\frac{d|\boldsymbol{v}|}{dt}\right)^2 = \frac{(n-2)^2}{4}|\boldsymbol{v}|^2 - \frac{n-2}{n}|\boldsymbol{v}|^{\frac{2n}{n-2}} + \kappa + \frac{\kappa_*}{|\boldsymbol{v}|^2}, \tag{1-7}$$

and we shall call this constant the new Pohozaev invariant.[2]

Thanks to an anonymous referee, we also observe a more precise characterization of the new invariant. Multiplying by v_i and $-v_j$ in the j-th and respectively in the i-th component of the system (1-3) $\bigl($with $\alpha = \frac{n+2}{n-2}\bigr)$, and then adding the resulting equations together side by side, we deduce that

$$\frac{d}{dt}\left(v_i\frac{dv_j}{dt} - v_j\frac{dv_i}{dt}\right) = 0, \quad 1 \le i, j \le m.$$

Thus for each $1 \le i, j \le m$ there exists a constant k_{ij} such that we have

$$v_i\frac{dv_j}{dt} - v_j\frac{dv_i}{dt} = k_{ij}. \tag{1-8}$$

[2] After this paper was accepted, we discovered a very recent work [Caju et al. 2019], from which we can actually prove that the new Pohozaev invariant in our paper is always zero for nonnegative solutions to the system (1-1). Having said that, some arguments here can be made more direct, without invoking the new Pohozaev invariant. Even so, we believe that our method gives some valuable insight, in particular, on the no-sign solution, where the new Pohozaev invariant becomes nontrivial for systems, while the method in [Caju et al. 2019] only works for nonnegative solutions.

Inserting (1-8) into (1-6) and comparing it with (1-7), we find that

$$\kappa_* = -\sum_{1\le i<j\le m} k_{ij}^2. \tag{1-9}$$

Without the radial symmetry, we obtain a more general formula (2-17) for the new Pohozaev invariant.

We point out that the analysis of the behavior of solutions to system (1-1) involves both κ and κ_*. This is a significant difference from the case of scalar equations, where κ fully determines the behavior of the solution around the isolated singularity, and especially $\kappa = 0$ is a sufficient and necessary condition to have removable singularity.

On the technical level, the system (1-1) exhibits some subtleties compared to the scalar case. One of the main tools we employ in the study of (1-1) is the method of moving spheres, which has been considered in [Jin et al. 2008; Li and Zhang 2003] and then continuously developed especially in the frame of the fractional Laplace operator; see, e.g., [Jin et al. 2014; Caffarelli et al. 2014]. The use of such a method in the case of systems requires particular attention, since the procedure can be continued in some components but should stop in others.

Another technical tool is the balanced-energy-type monotonicity functional (see, e.g., (2-1) below), which yields the Pohozaev identity in the upper critical case $\alpha = \frac{n+2}{n-2}$, combined with the blowup analysis. This energy functional has been a classical tool for the study of scalar case; see, e.g., [Bidaut-Véron and Véron 1991; Aviles 1987; Korevaar et al. 1999] and many others. We believe that the argument presented in this paper regarding the energy functional is more effective, due to an easy observation on the scaling relation (2-3) that is standard in the framework of free boundary problems.

1B. *Main results.* The main results are as follows. First we classify the solutions on the entire space, via the method of moving spheres.

Theorem 1.1. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $\mathbb{R}^n$ with $1 < \alpha \le \frac{n+2}{n-2}$:*

(i) *If $1 < \alpha < \frac{n+2}{n-2}$, then $\boldsymbol{u}$ is trivial.*

(ii) *If $\alpha = \frac{n+2}{n-2}$, then $\boldsymbol{u}$ is of the form*

$$\boldsymbol{u}(x) = \left[(n(n-2))^{\frac{n-2}{4}} \left(\frac{r}{r^2 + |x-z|^2} \right)^{\frac{n-2}{2}} \right] \boldsymbol{e} \tag{1-10}$$

for some $z \in \mathbb{R}^n$, $r \ge 0$, and a unit nonnegative vector $\boldsymbol{e} \in \mathbb{R}^m$.

Remark 1.2. Theorem 1.1(ii) was proved by O. Druet, E. Hebey and J. Vétois [Druet et al. 2010, Proposition 1.1] via the method of moving spheres. Here we include the result and the proof for the reader's convenience.

Next we classify the solutions in the punctured space, through the limiting energy levels or the Pohozaev invariants of the associated energy functional and the blowup analysis, which is standard in the framework of free boundary problems. For the upper critical case $\alpha = \frac{n+2}{n-2}$, we introduce a new Pohozaev invariant, which will play the central role.

Theorem 1.3. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $\mathbb{R}^n \setminus \{0\}$ with $1 < \alpha \le \frac{n+2}{n-2}$, and let $\Phi(r, \boldsymbol{u})$ be as in* (2-1) *for all $r > 0$*:

(i) *If $1 < \alpha \le \frac{n}{n-2}$, then $\boldsymbol{u}$ is trivial.*

(ii) *If $\frac{n}{n-2} < \alpha < \frac{n+2}{n-2}$, then $\Phi(r, \boldsymbol{u})$ converges as $r \to 0$ and $r \to \infty$, and*

$$\{\Phi(0+, \boldsymbol{u}), \Phi(+\infty, \boldsymbol{u})\} \subset \left\{ -\frac{\alpha-1}{\alpha+1} \lambda^{\frac{\alpha+1}{\alpha-1}}, 0 \right\} : \tag{1-11}$$

(a) *$\Phi(0+, \boldsymbol{u}) = 0$ if and only if $\boldsymbol{u}$ is trivial.*

(b) *$\Phi(+\infty, \boldsymbol{u}) = -\frac{\alpha-1}{\alpha+1}\lambda^{(\alpha+1)/(\alpha-1)}$ if and only if $\boldsymbol{u}$ is homogeneous of degree $-\frac{2}{\alpha-1}$, and hence of the form*

$$\boldsymbol{u}(x) = \lambda^{\frac{1}{\alpha-1}} |x|^{-\frac{2}{\alpha-1}} \boldsymbol{e}, \tag{1-12}$$

where λ is given by (1-4) *and $\boldsymbol{e} \in \mathbb{R}^m$ is a unit nonnegative vector.*

(iii) *If $\alpha = \frac{n+2}{n-2}$, then $\Phi_*(r, \boldsymbol{u})$ as in* (2-10) *is well-defined for all $r > 0$, and there are constants $\kappa(\boldsymbol{u})$ and $\kappa_*(\boldsymbol{u})$ such that $\kappa(\boldsymbol{u}) = \Phi(r, \boldsymbol{u})$ and $\kappa_*(\boldsymbol{u}) = \Phi_*(r, \boldsymbol{u})$ for all $r > 0$. Moreover,*

$$\kappa(\boldsymbol{u}) \ge -\frac{2}{n}\left(\frac{n-2}{2}\right)^n, \tag{1-13}$$

and

$$-\left(\frac{2}{n}\left(\frac{n-2}{2}\right)^n + \kappa(\boldsymbol{u})\right)\left(\frac{n-2}{2}\right)^{n-2} \le \kappa_*(\boldsymbol{u}) \le 0, \tag{1-14}$$

where the equalities of the lower bounds of both $\kappa(\boldsymbol{u})$ and $\kappa_*(\boldsymbol{u})$ hold only simultaneously:

(a) *$\kappa(\boldsymbol{u}) = \kappa_*(\boldsymbol{u}) = 0$ if and only if $\boldsymbol{u}$ has removable singularity at the origin, hence of the form* (1-10).

(b) *If $\kappa(\boldsymbol{u})^2 + \kappa_*(\boldsymbol{u})^2 > 0$, then $\boldsymbol{u}$ has nonremovable singularity at the origin, and is rotationally symmetric. Moreover, the cylindrical transformation $\boldsymbol{v}$ as in* (1-2) *satisfies* (1-7).

(c) *$\kappa(\boldsymbol{u}) = -\frac{2}{n}\left(\frac{n-2}{2}\right)^n$ and $\kappa_*(\boldsymbol{u}) = 0$ if and only if $\boldsymbol{u}$ is homogeneous of degree $-\frac{n-2}{2}$, and hence is of the form*

$$\boldsymbol{u}(x) = \left[\left(\frac{n-2}{2}\right)^{\frac{n-2}{2}} |x|^{-\frac{n-2}{2}}\right]\boldsymbol{e}, \tag{1-15}$$

where $\boldsymbol{e}$ is a unit nonnegative vector.

The subsequent theorems are concerned with the local solutions in the punctured unit ball. First we deduce the asymptotic radial symmetry by combining the methods of moving spheres and moving planes; a similar argument appears in [Caffarelli et al. 2014, Theorem 1.2]. This result is particularly important to define the second Pohozaev invariant for local solutions.

Theorem 1.4. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1 \setminus \{0\}$ with $1 < \alpha \le \frac{n+2}{n-2}$. Then*

$$\boldsymbol{u}(x) = (1 + O(|x|))\bar{\boldsymbol{u}}(|x|) \quad \text{as } x \to 0, \tag{1-16}$$

where $\bar{\boldsymbol{u}}(r)$ is the average of $\boldsymbol{u}$ over ∂B_r.

Utilizing the classification of solutions in the punctured space and the asymptotic radial symmetry, we obtain the exact asymptotic behavior of local solutions around the singularity.

Theorem 1.5. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1 \setminus \{0\}$ with $1 < \alpha \le \frac{n+2}{n-2}$. Then either $\boldsymbol{u}$ has a removable singularity at the origin, or the following alternatives hold*:

(i) *If $\frac{n}{n-2} < \alpha < \frac{n+2}{n-2}$, then*

$$|\boldsymbol{u}(x)| = (1+o(1))\lambda^{\frac{1}{\alpha-1}}|x|^{-\frac{2}{\alpha-1}} \quad \text{as } x \to 0, \tag{1-17}$$

where λ is given as in (1-4).

(ii) *If $\alpha = \frac{n+2}{n-2}$, then there are $c, C > 0$ such that*

$$c|x|^{-\frac{n-2}{2}} \le |\boldsymbol{u}(x)| \le C|x|^{-\frac{n-2}{2}} \quad \text{as } x \to 0, \tag{1-18}$$

where c depends on $\boldsymbol{u}$, while C is determined by n and m only.

(iii) *If $1 < \alpha < \frac{n}{n-2}$, then there are $c, C > 0$ such that*

$$c|x|^{2-n} \le |\boldsymbol{u}(x)| \le C|x|^{2-n} \quad \text{as } x \to 0, \tag{1-19}$$

where both c and C depend on $\boldsymbol{u}$.

(iv) *If $\alpha = \frac{n}{n-2}$, then*

$$|\boldsymbol{u}(x)| = (1+o(1))\left(\frac{(n-2)^2}{2|x|^2(-\log|x|)}\right)^{\frac{n-2}{2}} \quad \text{as } x \to 0. \tag{1-20}$$

The paper is organized as follows. In the next section, we present the balanced-energy-type monotonicity formula and introduce the second Pohozaev invariants for the upper critical case. In Section 3, we classify the solutions of (1-1) on the whole space, proving Theorem 1.1. In Section 4, we investigate the properties of the solutions on the punctured space, and present the proof of Theorem 1.3. Section 5 is devoted to the a priori estimates for the local solutions, which will play one of the key roles in the subsequent analysis, while we prove the asymptotic radial symmetry, Theorem 1.4, in Section 6. Finally, we derive the exact asymptotic behavior of the local solutions of (1-1) for all $1 < \alpha \le \frac{n+2}{n-2}$ in Section 7. The proofs of parts (i)–(iv) in Theorem 1.5 are presented in the ends of Sections 7A–7D, respectively.

1C. ***Notation and terminology.*** If $|\boldsymbol{u}|$ is bounded in any neighborhood of the origin, we say $|\boldsymbol{u}|$ has a removable singularity. Otherwise, we say that it has a nonremovable singularity.

By $B_r(z) \subset \mathbb{R}^n$ ($n \ge 3$) we denote the ball of radius r centered at z, and $B_r = B_r(0)$. In addition, ω_n is the volume of the unit ball $B_1 \subset \mathbb{R}^n$. Given an open set $\Omega \subset \mathbb{R}^n$, we shall denote by $\partial\Omega$ the topological boundary of Ω. Moreover, when $\partial\Omega$ is C^1, ν denotes the unit normal on $\partial\Omega$ pointing towards the origin. ∇_σ will denote the tangential derivative on $\partial\Omega$.

$\mathbb{S}^{n-1}$ is the unit sphere in $\mathbb{R}^n$, and is also identified with ∂B_1. Note that $n\omega_n$ is the area of $\mathbb{S}^{n-1}$. By ∇_θ and Δ_θ we shall write the derivative and, respectively, the Laplace–Beltrami operator on $\mathbb{S}^{n-1}$.

Any vector in the target space $\mathbb{R}^m$ ($m \ge 1$) is written in bold. Given a vector $\boldsymbol{a} \in \mathbb{R}^m$, we denote by a_i the i-th component of $\boldsymbol{a}$. By $|\boldsymbol{a}|$ we denote its l^2-norm; i.e., $|\boldsymbol{a}| = \left(\sum_{i=1}^m a_i^2\right)^{1/2}$. By $\boldsymbol{a} \ge 0$ (resp., $\boldsymbol{a} \le 0$) or by saying that $\boldsymbol{a}$ is nonnegative (resp., nonpositive) we indicate that $a_i \ge 0$ (resp., $a_i \le 0$) for each $1 \le i \le m$. For two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, we define $\boldsymbol{a} \cdot \boldsymbol{b} = \sum_{i=1}^m a_i b_i$. Also given two vectorial C^1-functions $\boldsymbol{f}$ and $\boldsymbol{g}$, we define $\nabla \boldsymbol{f} : \nabla \boldsymbol{g} = \sum_{i=1}^m (\nabla f_i) \cdot (\nabla g_i)$.

The constants $C, C_0, C_1, C_2, \ldots$ will always be positive, generic, determined by n, m and α only, unless otherwise stated. We shall also call these constants universal. In addition, we shall fix λ, μ, and $\bar\lambda$ throughout the paper as in (1-4) and

$$\bar\lambda = \frac{\alpha-1}{\alpha+1}\lambda^{\frac{\alpha+1}{\alpha-1}}. \tag{1-21}$$

2. Monotonicity formula and Pohozaev invariant

We consider the balanced-energy-type functional

$$\Phi(r,\boldsymbol{u}) = \frac{r^{\mu+1}}{n\omega_n}\int_{\partial B_r}\left(\left|\frac{\partial \boldsymbol{u}}{\partial\nu} - \frac{2}{(\alpha-1)r}\boldsymbol{u}\right|^2 - |\nabla_\sigma \boldsymbol{u}|^2\right)d\sigma + \frac{2r^{\mu+1}}{(\alpha+1)n\omega_n}\int_{\partial B_r}|\boldsymbol{u}|^{\alpha+1}\,d\sigma - \frac{\lambda r^{\mu-1}}{n\omega_n}\int_{\partial B_r}|\boldsymbol{u}|^2\,d\sigma, \tag{2-1}$$

where λ and μ are given as in (1-4). Note that $\lambda \ge 0$ if and only if $\alpha \ge \frac{n}{n-2}$, and $\mu \ge 0$ if and only if $1 < \alpha \le \frac{n+2}{n-2}$.

Let us introduce the scaling function

$$\boldsymbol{u}_r(x) = r^{\frac{2}{\alpha-1}}\boldsymbol{u}(rx). \tag{2-2}$$

Note that the problem (1-1) is preserved under this scaling. That is, if $\boldsymbol{u}$ solves (1-1) in $B_R\setminus\{0\}$ then $\boldsymbol{u}_r$ solves (1-1) in $B_{R/r}\setminus\{0\}$. In terms of $\boldsymbol{u}_r$, one may easily observe that Φ satisfies the scaling relation

$$\Phi(rs,\boldsymbol{u}) = \Phi(s,\boldsymbol{u}_r) \tag{2-3}$$

for any $r, s > 0$.

Recall from (1-2) the cylindrical transformation $\boldsymbol{v}$, in terms of which Φ can be represented as

$$\Phi(r,\boldsymbol{u}) = \Psi(-\log r,\boldsymbol{v}), \tag{2-4}$$

where $\Psi(t,\boldsymbol{v})$ is given by

$$\Psi(t,\boldsymbol{v}) = \frac{1}{n\omega_n}\int_{\mathbb{S}^{n-1}}\left(|\partial_t \boldsymbol{v}|^2 - |\nabla_\theta \boldsymbol{v}|^2 - \lambda|\boldsymbol{v}|^2 + \frac{2}{\alpha+1}|\boldsymbol{v}|^{\alpha+1}\right)d\theta. \tag{2-5}$$

Proposition 2.1. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_R\setminus\{0\}$ with $1<\alpha\le\frac{n+2}{n-2}$, and let $\Phi(r,\boldsymbol{u})$ be as in* (2-1). *One has*

$$\frac{d}{dr}\Phi(r,\boldsymbol{u}) = \frac{2\mu r^{\mu}}{n\omega_n}\int_{\partial B_r}\left|\frac{\partial\boldsymbol{u}}{\partial\nu} - \frac{2}{(\alpha-1)r}\boldsymbol{u}\right|^2 d\sigma, \tag{2-6}$$

where μ is given as in (1-4). *In particular, the following are true*:

(i) *If $1<\alpha<\frac{n+2}{n-2}$, then $\Phi(r,\boldsymbol{u})$ is nondecreasing for $0<r<R$. Moreover, $\Phi(r,\boldsymbol{u})$ is constant for $r_1<r<r_2$ if and only if $\boldsymbol{u}$ is homogeneous of degree $-\frac{2}{\alpha-1}$ in $B_{r_2}\setminus\overline{B}_{r_1}$, i.e.,*

$$\boldsymbol{u}(x) = |x|^{-\frac{2}{\alpha-1}}\boldsymbol{u}\left(\frac{x}{|x|}\right) \quad \text{in } B_{r_2}\setminus\overline{B}_{r_1}. \tag{2-7}$$

(ii) *If $\alpha = \frac{n+2}{n-2}$, then $\Phi(r,\boldsymbol{u})$ is constant for $0<r<R$.*

Proof. The computation is easy if one chooses cylindrical coordinates. Since (2-4) holds with $t = -\log r$,

$$\begin{aligned} r\dot{\Phi}(r,\boldsymbol{u}) = -\Psi'(t,\boldsymbol{v}) &= -\frac{2}{n\omega_n}\int_{\mathbb{S}^{n-1}} ((\partial_{tt}\boldsymbol{v} - \lambda\boldsymbol{v} + |\boldsymbol{v}|^{\alpha-1}\boldsymbol{v})\cdot\partial_t\boldsymbol{v} - \nabla_\theta\boldsymbol{v} : \nabla_\theta\partial_t\boldsymbol{v})\,d\theta \\ &= -\frac{2}{n\omega_n}\int_{\mathbb{S}^{n-1}} (\partial_{tt}\boldsymbol{v} + \Delta_\theta\boldsymbol{v} - \lambda\boldsymbol{v} + |\boldsymbol{v}|^{\alpha-1}\boldsymbol{v})\cdot\partial_t\boldsymbol{v}\,d\theta \\ &= \frac{2\mu}{n\omega_n}\int_{\mathbb{S}^{n-1}} |\partial_t\boldsymbol{v}|^2\,d\theta, \end{aligned}$$

where $\dot{\Phi}$ and Ψ' denote $d\Phi/dr$ and $d\Psi/dt$ respectively, and the right side is evaluated at $t = -\log r$. In addition, when deriving the last equality we used (1-3). Rephrasing the rightmost side in terms of $\boldsymbol{u}$, we arrive at (2-6).

The assertion on the monotonicity of Φ is now clear from (2-6). On the other hand, the assertion on the homogeneity can be shown as follows. We see that if $\alpha \neq \frac{n+2}{n-2}$, then one has $\mu \neq 0$. Hence, the assumption that $\Phi(r,\boldsymbol{u})$ is constant for $r_1 < r < r_2$ along with (2-6) yields that for any $r_1 < r < r_2$

$$\frac{\partial\boldsymbol{u}}{\partial\nu} = \frac{2}{(\alpha-1)r}\boldsymbol{u} \quad \text{on } \partial B_r,$$

where ν is the unit normal pointing towards the origin. Thus, $\boldsymbol{u}$ is homogeneous of degree $-\frac{2}{\alpha-1}$ in $B_{r_2}\setminus\overline{B}_{r_1}$. □

Remark 2.2. As a matter of fact, (2-6) holds for $\alpha > \frac{n+2}{n-2}$, and hence $\Phi(r,\boldsymbol{u})$ is nonincreasing in this case, since $\mu < 0$ for $\alpha > \frac{n+2}{n-2}$.

Remark 2.3. For the case $\alpha = \frac{n+2}{n-2}$, we obtain from Proposition 2.1(ii) a constant $\kappa(\boldsymbol{u})$ such that

$$\kappa(\boldsymbol{u}) = \Phi(r,\boldsymbol{u}) \tag{2-8}$$

for any $0 < r < R$. Since there is a one-to-one correspondence between the nonnegative solutions $\boldsymbol{u}$ of (1-1) and $\boldsymbol{v}$ of (1-3) via the cylindrical transform (1-2), we shall write $\kappa(\boldsymbol{u})$ by $\kappa(\boldsymbol{v})$ as well. In view of (2-4), it is clear that

$$\kappa(\boldsymbol{v}) = \Psi(t,\boldsymbol{v}) \tag{2-9}$$

for any $t > -\log R$. We shall call κ the first Pohozaev invariant.

Let us construct the second Pohozaev invariant in a general setting, that is without rotational symmetry.[3] For $\alpha = \frac{n+2}{n-2}$, let us define, formally for the moment, the quantity

$$\begin{aligned} \Phi_*(r,\boldsymbol{u}) &= \tfrac{1}{4}(r\dot{f}(r,\boldsymbol{u}))^2 - \tfrac{1}{4}((n-2)^2)f(r,\boldsymbol{u})^2 - \kappa(\boldsymbol{u})f(r,\boldsymbol{u}) \\ &\quad + \frac{n-2}{n}f(r,\boldsymbol{u})^{\frac{2n-2}{n-2}} - 2\int_0^r\left(\frac{\rho}{n\omega_n}\int_{\partial B_\rho}|\nabla_\sigma\boldsymbol{u}|^2\,d\sigma\right)\dot{f}(\rho,\boldsymbol{u})\,d\rho \\ &\quad + \frac{2n-2}{n}\int_0^r\left(\frac{\rho}{n\omega_n}\int_{\partial B_\rho}|\boldsymbol{u}|^{\frac{2n}{n-2}}\,d\sigma - f(\rho,\boldsymbol{u})^{\frac{n}{n-2}}\right)\dot{f}(\rho,\boldsymbol{u})\,d\rho, \end{aligned} \tag{2-10}$$

[3]As noted in an earlier footnote, we discovered that the second Pohozaev invariant always becomes trivial for nonnegative singular solutions on the punctured space, $\mathbb{R}^n\setminus\{0\}$, after this paper was accepted. However, this is by no means straightforward for local solutions in a punctured ball, without asymptotic radial symmetry (Theorem 1.4). Moreover, this invariant becomes nontrivial for no-sign solutions. For these reasons, we shall present a general formulation of the second Pohozaev invariant.

where $\dot{f}$ denotes df/dr, and

$$f(r,\boldsymbol{u}) = \frac{1}{n\omega_n r}\int_{\partial B_r} |\boldsymbol{u}|^2\,d\sigma. \tag{2-11}$$

Notice that $\Phi_*(r,\boldsymbol{u})$ is well-defined only if the last two double integrals on the right side are finite. Moreover, once $\Phi_*(r,\boldsymbol{u})$ becomes well-defined, we may also deduce from

$$r\dot{f}(r,\boldsymbol{u}) = -\frac{2}{n\omega_n}\int_{\partial B_r} \boldsymbol{u}\cdot\left(\frac{\partial \boldsymbol{u}}{\partial\nu} - \frac{n-2}{2r}\boldsymbol{u}\right)d\sigma \tag{2-12}$$

a scaling relation of Φ_*,

$$\Phi_*(rs,\boldsymbol{u}) = \Phi_*(s,\boldsymbol{u}_r), \tag{2-13}$$

which holds for any $r, s > 0$. On the other hand, in terms of the cylindrical transformation $\boldsymbol{v}$, one has

$$\Phi_*(r,\boldsymbol{u}) = \Psi_*(-\log r,\boldsymbol{v}), \tag{2-14}$$

where $\Psi_*(t,\boldsymbol{v})$ is given by

$$\begin{aligned}\Psi_*(t,\boldsymbol{v}) &= \tfrac{1}{4}(g'(t,\boldsymbol{v}))^2 - \tfrac{1}{4}((n-2)^2)g(t,\boldsymbol{v})^2 - \kappa(\boldsymbol{v})g(t,\boldsymbol{v})\\ &\quad+\frac{n-2}{n}g(t,\boldsymbol{v})^{\frac{2n-2}{n-2}} + 2\int_t^\infty\left(\frac{1}{n\omega_n}\int_{\mathbb{S}^{n-1}}|\nabla_\theta \boldsymbol{v}|^2\,d\theta\right)g'(\tau,\boldsymbol{v})\,d\tau\\ &\quad-\frac{2n-2}{n}\int_t^\infty\left(\frac{1}{n\omega_n}\int_{\mathbb{S}^{n-1}}|\boldsymbol{v}|^{\frac{2n}{n-2}}\,d\theta - g(\tau,\boldsymbol{v})^{\frac{n}{n-2}}\right)g'(\tau,\boldsymbol{v})\,d\tau,\end{aligned} \tag{2-15}$$

with g' being dg/dt and

$$g(t,\boldsymbol{v}) = \frac{1}{n\omega_n}\int_{\mathbb{S}^{n-1}}|\boldsymbol{v}|^2\,d\theta. \tag{2-16}$$

Proposition 2.4. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_R\setminus\{0\}$ with $\alpha = \frac{n+2}{n-2}$, and let $\Phi_*(r,\boldsymbol{u})$ be as in* (2-10). *Then $\Phi_*(r,\boldsymbol{u})$ is well-defined and is constant for $0 < r < R$.*

We shall postpone the proof to Section 6, since proving the well-definedness of $\Phi_*(r,\boldsymbol{u})$ essentially relies on the asymptotic radial symmetry of local solutions to (1-1) (see Theorem 1.4).

Remark 2.5. Knowing that $\Phi_*(r,\boldsymbol{u})$ is constant, we obtain a constant $\kappa_*(\boldsymbol{u})$ such that

$$\kappa_*(\boldsymbol{u}) = \Phi_*(r,\boldsymbol{u}) \tag{2-17}$$

for any $0 < r < R$. We shall call this constant the second Pohozaev invariant. As with the first Pohozaev invariant, we will also write it by $\kappa_*(\boldsymbol{v})$ whenever $\boldsymbol{v}$ is the cylindrical transformation. Clearly,

$$\kappa_*(\boldsymbol{v}) = \Psi_*(t,\boldsymbol{v}) \tag{2-18}$$

for any $t > -\log R$. In Sections 4 and 7A we will observe that $\kappa_*(\boldsymbol{v}) = 0$ if and only if $\boldsymbol{v}(t,\theta) = (1+o(1))|\boldsymbol{v}(t,\theta)|\boldsymbol{e}$ uniformly for $\theta\in\mathbb{S}^{n-1}$ as $t\to\infty$, with some nonnegative unit vector $\boldsymbol{e}\in\mathbb{R}^m$.

3. Solutions on the whole space

In this section we classify the smooth solutions of (1-1) on the whole space $\mathbb{R}^n$. The analysis is based on the method of moving spheres along with the Kelvin transform, and we follow essentially the argument proposed in [Li and Zhang 2003, Section 2], with only a minor modification. Nevertheless, we shall include the full argument here for the reader's convenience.

Given $z \in \mathbb{R}^n$ and $r > 0$, we shall write $\boldsymbol{u}^*_{z,r}$ for the Kelvin transform of $\boldsymbol{u}$ with respect to the sphere $B_r(z)$; that is,

$$\boldsymbol{u}^*_{z,r}(y) = \left(\frac{r}{|y-z|}\right)^{n-2} \boldsymbol{u}\left(z + \frac{r^2}{|y-z|^2}(y-z)\right). \tag{3-1}$$

Let us remark that if $\boldsymbol{u}$ is a solution of (1-1) in $\mathbb{R}^n$, then

$$-\Delta \boldsymbol{u}^*_{z,r} = \left(\frac{r}{|y-z|}\right)^{(\alpha-1)\mu} |\boldsymbol{u}^*_{z,r}|^{\alpha-1}\boldsymbol{u}^*_{z,r} \quad \text{in } \mathbb{R}^n \setminus \{z\}, \tag{3-2}$$

where μ is given by (1-4). Note that $\mu \geq 0$ if and only if $1 < \alpha \leq \frac{n+2}{n-2}$. The nonnegativity of μ will play a key role when comparing $\boldsymbol{u}$ and $\boldsymbol{u}^*_{z,r}$.

We begin with a basic lemma that holds for any nonnegative, superharmonic function, as a starting point of the method of moving spheres.

Lemma 3.1 [Li and Zhang 2003, Lemma 2.1]. *Let $v \in C^2(\mathbb{R}^n)$ be a superharmonic and nonnegative function on $\mathbb{R}^n$. Then for each $z \in \mathbb{R}^n$, there exists $r_0 > 0$, which may depend on v and z, such that for all $0 < r < r_0$*

$$v^*_{z,r} \leq v \quad \textit{in } \mathbb{R}^n \setminus B_r(z). \tag{3-3}$$

The next lemma is an analogue of [Caffarelli et al. 1989, Lemma 2.4], which claims that either the inequality (3-3) must hold until the solution becomes symmetric (with respect to a sphere) or it must fail on a compact subset of $\mathbb{R}^n$. The proof is given in that of [Li and Zhang 2003, Lemma 2.2], and we shall not repeat it here.

Lemma 3.2. *Let $v \in C^2(\mathbb{R}^n)$, $z \in \mathbb{R}^n$, and $r_0 > 0$ be such that*

$$-\Delta(v - v^*_{z,r_0}) \geq 0 \quad \textit{in } \mathbb{R}^n \setminus \bar{B}_{r_0}(z), \tag{3-4}$$

$$v^*_{z,r_0} < v \quad \textit{in } \mathbb{R}^n \setminus \bar{B}_{r_0}(z). \tag{3-5}$$

Then there is a small $\epsilon > 0$ such that for any $r_0 < r < r_0 + \epsilon$

$$v^*_{z,r} < v \quad \textit{in } \mathbb{R}^n \setminus B_r(z). \tag{3-6}$$

Now let us turn our interest to the nonnegative, smooth global solutions $\boldsymbol{u}$ of (1-1). Given $z \in \mathbb{R}^n$, let us define, for each $1 \leq i \leq m$,

$$r_i(z) = \sup\{r > 0 : (u_i)^*_\rho \leq u_i \text{ in } \mathbb{R}^n \setminus B_\rho(z) \text{ for any } 0 < \rho < r\}. \tag{3-7}$$

Since each component u_i of $\boldsymbol{u}$ is nonnegative and superharmonic, Lemma 3.1 applies to u_i, from which we know that $r_i(z) > 0$ for each $1 \leq i \leq m$. Thus, we have

$$\bar{r}(z) = \inf_{1 \leq i \leq m} r_i(z) > 0. \tag{3-8}$$

Let us remark that we have defined $\bar{r}(z)$ by the infimum, instead of minimum, over a finite set of indices $\{1, 2, \dots, m\}$, since $r_i(z)$ as a supremum could be infinite. Moreover, if $r_i(z) = \infty$ for all $1 \le i \le m$, we shall say that $\bar{r}(z) = \infty$.

The following lemma takes care of the case when $\bar{r}(z)$ is either finite or infinite. The proof is essentially the same as those of [Druet et al. 2010, Lemmas 1.2 and 1.3], which deal with the upper critical case $\alpha = \frac{n+2}{n-2}$ only, whence we shall skip the details.

Lemma 3.3. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $\mathbb{R}^n$ with $1 < \alpha \le \frac{n+2}{n-2}$, $z \in \mathbb{R}^n$ be arbitrary, and $\bar{r}(z)$ be as in* (3-8). *If $\bar{r}(z)$ is finite, then*

$$\boldsymbol{u}^*_{z,\bar{r}(z)} = \boldsymbol{u} \quad \text{in } \mathbb{R}^n \setminus \{z\}. \tag{3-9}$$

If $\bar{r}(z_0) = \infty$ for some $z_0 \in \mathbb{R}^n$, then $\bar{r}(z) = \infty$ for all $z \in \mathbb{R}^n$.

We are now ready to classify the smooth global solutions.

Proof of Theorem 1.1. In view of Lemma 3.3, we observe that $\bar{r}(z)$ defined in (3-8) is either finite or infinite for all $z \in \mathbb{R}^n$. If $\bar{r}(z)$ is finite for all $z \in \mathbb{R}^n$, then we have (3-9) at every point $z \in \mathbb{R}^n$. In this case, we may apply [Li and Zhang 2003, Lemma 11.1]: there are $a_i \ge 0$, $r_i > 0$, and $z_i \in \mathbb{R}^m$ for $1 \le i \le m$ such that

$$u_i(x) = a_i r_i^{-\frac{n-2}{2}} \left(\frac{r_i}{r_i^2 + |x - z_i|^2} \right)^{\frac{n-2}{2}}. \tag{3-10}$$

On the other hand, if $\bar{r}(z)$ is infinite for all $z \in \mathbb{R}^n$, we have (3-7) for all $r > 0$ at any $z \in \mathbb{R}^n$. Due to [Li and Zhang 2003, Lemma 11.2], there are $b_i \ge 0$ for $1 \le i \le m$ such that

$$u_i(x) = b_i. \tag{3-11}$$

Suppose that $\boldsymbol{u}$ satisfies (3-11), that is, $\boldsymbol{u}$ is constant everywhere on $\mathbb{R}^n$. As $\boldsymbol{u}$ is a nonnegative solution of (1-1) in $\mathbb{R}^n$, $\boldsymbol{u}$ must be zero everywhere. Hence, parts (i) and (ii) of Theorem 1.1 are satisfied under this assumption.

Next, let us consider the case that u_i satisfies (3-10) for all $1 \le i \le m$. This part is the same as the proof of [Druet et al. 2010, Proposition 1.1], so we omit the details. □

4. Solutions in punctured space

4A. ***Radial symmetry of singular solutions.*** This section is devoted to the radial symmetry of nonnegative, singular solutions of (1-1). To be more precise, $\boldsymbol{u}$ is a nonnegative solution of (1-1) in the punctured space $\mathbb{R}^n \setminus \{0\}$ that has a nonremovable singularity at the origin, i.e.,

$$\limsup_{x \to 0} |\boldsymbol{u}(x)| = \infty. \tag{4-1}$$

The proof relies again on the method of moving spheres used in the previous section. The proof for the case of a single equation has already been established in [Jin et al. 2008, Proposition 2.1]. Nevertheless, the multiplicity in the components here makes the comparison argument more subtle, as observed in the previous section. Let us also address that the method of moving planes also works, see [Caffarelli et al. 1989, Theorem 8.1], after a suitable modification.

Lemma 4.1. *Let **u** be a nonnegative solution of* (1-1) *in* $\mathbb{R}^n \setminus \{0\}$ *with* $1 < \alpha \le \frac{n+2}{n-2}$. *If **u** satisfies* (4-1), *then **u** is radially symmetric.*

Proof. Let $z \in \mathbb{R}^n \setminus \{0\}$ be arbitrary. Arguing similarly to Lemma 3.1, whose proof can be found in [Li and Zhang 2003, Lemma 2.1], there exists some $0 < r_0 < |z|$ such that for any $0 < r \le r_0$

$$(u_i)^*_{z,r} \le u_i \quad \text{in } \mathbb{R}^n \setminus (B_r(z) \cup \{0\}) \text{ for each } 1 \le i \le m.$$

Hence, one can define, as with (3-7) and (3-8),

$$r_i(z) = \sup\{r > 0 : (u_i)^*_{z,\rho} \le u_i \text{ in } \mathbb{R}^n \setminus (B_\rho(z) \cup \{0\}) \text{ for any } 0 < \rho < r\}$$

and

$$\bar{r}(z) = \inf_{1 \le i \le m} r_i(z).$$

We first claim that

$$0 < \bar{r}(z) \le |z|. \tag{4-2}$$

The positivity of $\bar{r}(z)$ is clear.To prove the second inequality in (4-2), let us first observe that by (4-1), there exist some sequence $x_j \to 0$ and a component u_i such that $u_i(x_j) \to \infty$. If $\bar{r}(z) > |z|$, then by its definition, there should exist $\rho > |z|$ such that

$$(u_i)^*_{z,\rho} \le u_i \quad \text{in } \mathbb{R}^n \setminus B_\rho(z). \tag{4-3}$$

Now let y_j be the reflection of x_j with respect to $\partial B_\rho(z)$; i.e.,

$$y_j = z + \left(\frac{\rho}{|x_j - z|}\right)^2 (x_j - z).$$

Since $x_j \to 0$, we have $y_j \in \mathbb{R}^n \setminus B_\rho(z)$ for all sufficiently large j, and moreover,

$$y_j \to y_0 = \left(1 - \left(\frac{\rho}{|z|}\right)^2\right) z.$$

Thus, if we take ρ close enough to $|z|$, we have $y_0 \ne 0$, whence u_i is smooth at y_0. However, (4-3) implies

$$u_i(y_0) = \lim_{j\to\infty} u_i(y_j) \ge \lim_{j\to\infty} ((u_i)^*_{z,\rho}(y_j)) \ge \left(\frac{|z|}{\rho}\right)^{n-2} \lim_{j\to\infty} u_i(x_j) = \infty,$$

a contradiction.

From (4-2), we can also claim that

$$\bar{r}(z) = |z|.$$

The argument is based on the proof of [Jin et al. 2008, Proposition 2.1] with the corresponding modification shown in Lemma 3.3, which amounts to the number of nontrivial components. The main idea is that if $\bar{r}(z) < |z|$, then (4-1) together with the maximum principle implies

$$u_i > (u_i)^*_{z,\bar{r}(z)} \quad \text{in } \mathbb{R}^n \setminus (\bar{B}_{\bar{r}(z)}(z) \cup \{0\}), \tag{4-4}$$

at least for one $1 \le i \le m$. Then we must have $|\boldsymbol{u}| > |\boldsymbol{u}^*_{z,\bar r(z)}|$ in $\mathbb{R}^n \setminus (\bar B_{\bar r(z)}(z) \cup \{0\})$, and the strong maximum principle yields that the strict inequality in (4-4) must hold for all nontrivial components. Hence, as with Lemma 3.2, we obtain some $\epsilon > 0$ such that (4-4) holds for all $1 \le i \le m$ with $\bar r(z)$ replaced by some $\bar r(z) < r < \bar r(z) + \epsilon$, a contradiction to (4-3). The details are omitted.

To this end, we have proved that for each $z \in \mathbb{R}^n \setminus \{0\}$ and for any $0 < r < |z|$

$$(u_i)^*_{z,r} \le u_i \quad \text{in } \mathbb{R}^n \setminus (B_r(z) \cup \{0\}) \text{ for each } 1 \le i \le m.$$

Thus, one may deduce from [Jin et al. 2008, Lemma 2.1] that u_i is radially symmetric for each $1 \le i \le m$. □

4B. *Limiting energy levels and Pohozaev invariants.* Knowing the radial symmetry of singular solutions, we may classify the nonnegative solutions on the punctured space, using the balanced-energy limit. The idea is to consider both *blowups* and *shrink-downs* of $\boldsymbol{u}$ under the scaling (2-2). Here by saying a blowup or a shrink-down under the scaling $\boldsymbol{u}_r$ we indicate a limit of $\boldsymbol{u}_r$ as $r = r_j \to 0+$, or respectively $r = r_j \to \infty$ in $C^2_{\mathrm{loc}}(\mathbb{R}^n \setminus \{0\}; \mathbb{R}^m)$. The following lemma provides the compactness of the sequence $\boldsymbol{u}_r$ in order to have both the blowups and the shrink-downs.

Lemma 4.2. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $\mathbb{R}^n \setminus \{0\}$ with $1 < \alpha \le \frac{n+2}{n-2}$. If $\boldsymbol{u}$ satisfies* (4-1), *then for each $1 \le i \le m$*

$$u_i(x) \le \left(\frac{\alpha-1}{2n}\right)^{-\frac{1}{\alpha-1}} |x|^{-\frac{2}{\alpha-1}} \quad \text{in } \mathbb{R}^n \setminus \{0\}. \tag{4-5}$$

Proof. Let u_i be a positive component of $\boldsymbol{u}$. Then, since u_i is superharmonic in $\mathbb{R}^n \setminus \{0\}$, it follows from the extended maximum principle [Gilbarg and Serrin 1956, Theorem 1] that

$$\liminf_{x\to 0} u_i(x) > 0. \tag{4-6}$$

Now let $v = u_i^{1-\alpha}$. Then v satisfies, in $\mathbb{R}^n \setminus \{0\}$,

$$\Delta v \ge \frac{\alpha}{\alpha-1}\frac{|\nabla v|^2}{v} + \alpha - 1.$$

Hence, for each $r > 0$, the auxiliary function

$$w(x) = v(x) - \frac{\alpha-1}{2n}|x|^2$$

becomes subharmonic in $B_r \setminus \{0\}$. Then by (4-6), w is bounded around the origin, and thus, it follows from the extended maximum principle [Gilbarg and Serrin 1956, Theorem 1] that

$$0 \le \limsup_{x\to 0} w(x) \le \sup_{\partial B_r} w = \sup_{\partial B_r} v - \frac{\alpha-1}{2n} r^2.$$

In terms of u_i, we obtain

$$\inf_{\partial B_r} u_i \le \left(\frac{\alpha-1}{2n}\right)^{-\frac{1}{\alpha-1}} r^{-\frac{2}{\alpha-1}}.$$

Now the radial symmetry obtained in Lemma 4.1 yields (4-5). □

The next lemma gives the compactness of the sequence $\boldsymbol{u}_r$, and hence the existence of both a blowup and a shrink-down of $\boldsymbol{u}$.

Lemma 4.3. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $\mathbb{R}^n \setminus \{0\}$ with $1 < \alpha \le \frac{n+2}{n-2}$. Then there is some $0 < \gamma < 1$ such that $\boldsymbol{u}_r$ is uniformly bounded in $C^{2,\gamma}(K; \mathbb{R}^m)$ on each compact set $K \subset \mathbb{R}^n \setminus \{0\}$.*

Proof. If $\boldsymbol{u}$ does not satisfy (4-1), then $\boldsymbol{u}$ is bounded around the origin, and the origin becomes a removable singularity. According to Theorem 1.1, if $1 < \alpha < \frac{n+2}{n-2}$, $\boldsymbol{u}$ is trivial, while if $\alpha = \frac{n+2}{n-2}$, $\boldsymbol{u}$ is globally bounded and satisfies $|\boldsymbol{u}(x)| = O(|x|^{2-n})$ as $|x| \to \infty$. Hence, in any case, $\boldsymbol{u}_r$ is bounded uniformly for all $r > 0$ on a fixed compact subset of $\mathbb{R}^n \setminus \{0\}$.

On the other hand, if $\boldsymbol{u}$ satisfies (4-1), Lemma 4.2 implies that $\boldsymbol{u}_r$ is globally bounded in $\mathbb{R}^n \setminus \{0\}$. Thus, regardless of the removability of the singularity at the origin, we know that $\boldsymbol{u}_r$ is uniformly bounded in each compact subset of $\mathbb{R}^n \setminus \{0\}$.

Now since $\boldsymbol{u}_r$ also solves (1-1) in $\mathbb{R}^n \setminus \{0\}$, it follows from the interior regularity theory [Gilbarg and Trudinger 1983, Theorems 6.2 and 6.19] that $\boldsymbol{u}_r$ is uniformly bounded in $C^{2,\gamma}(K; \mathbb{R}^m)$ on each compact set $K \subset \mathbb{R}^n \setminus \{0\}$ for some $0 < \gamma < 1$. □

Let $\Phi(r, \boldsymbol{u})$ be the balanced-energy-type functional defined by (2-1). Recall from Proposition 2.1 that $\Phi(r, \boldsymbol{u})$ is monotone increasing in $r > 0$ for $1 < \alpha < \frac{n+2}{n-2}$, while it is constant for $\alpha = \frac{n+2}{n-2}$.

Lemma 4.4. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $\mathbb{R}^n \setminus \{0\}$ with $1 < \alpha \le \frac{n+2}{n-2}$, and let $\boldsymbol{u}_0$ and $\boldsymbol{u}_\infty$ be respectively a blowup and shrink-down under the scaling $\boldsymbol{u}_r$. Then $\Phi(r, \boldsymbol{u}_0) = \Phi(0+, \boldsymbol{u})$ and $\Phi(r, \boldsymbol{u}_\infty) = \Phi(\infty, \boldsymbol{u})$ for all $r > 0$. In particular, both $\boldsymbol{u}_0$ and $\boldsymbol{u}_\infty$ are homogeneous of degree $-\frac{2}{\alpha-1}$, provided that $1 < \alpha < \frac{n+2}{n-2}$.*

Proof. Since the argument for shrink-downs is the same, we shall only present it for blowups. Let $\boldsymbol{u}_0$ be a blowup with a sequence $r_j \to 0+$. Then due to the scaling relation (2-3), we have, for any $r > 0$,

$$\Phi(r, \boldsymbol{u}_0) = \lim_{j\to\infty} \Phi(r, \boldsymbol{u}_{r_j}) = \lim_{j\to\infty} \Phi(r r_j, \boldsymbol{u}) = \Phi(0+, \boldsymbol{u}),$$

where the existence of $\Phi(0+, \boldsymbol{u})$ follows from the compactness of $\boldsymbol{u}_r$ (Lemma 4.3) and the monotonicity of $\Phi(r, \boldsymbol{u})$ (Proposition 2.1(i)). This proves the first assertion of Lemma 4.4. The second assertion on the homogeneity follows again from Proposition 2.1(i). □

Lemma 4.5. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $\mathbb{R}^n \setminus \{0\}$ with $1 < \alpha \le \frac{n+2}{n-2}$. Suppose further that $\boldsymbol{u}$ is homogeneous of degree $-\frac{2}{\alpha-1}$:*

(i) *If $1 < \alpha \le \frac{n}{n-2}$, then $\boldsymbol{u}$ is trivial.*

(ii) *If $\frac{n}{n-2} < \alpha \le \frac{n+2}{n-2}$, then either $\boldsymbol{u}$ is trivial, or $\boldsymbol{u}$ is of the form* (1-12).

Proof. Since $\boldsymbol{u}$ is homogeneous of degree $-\frac{2}{\alpha-1}$, the cylindrical transform $\boldsymbol{v}$ introduced in (1-2) satisfies

$$\Delta_\theta \boldsymbol{v} - \lambda \boldsymbol{v} + |\boldsymbol{v}|^{\alpha-1}\boldsymbol{v} = 0 \quad \text{on } \mathbb{S}^{n-1}, \tag{4-7}$$

where Δ_θ is the Laplace–Beltrami operator, and λ is given by (1-4).

Case 1: $1 < \alpha \le \frac{n}{n-2}$. In view of (1-4), we have $\lambda \le 0$. As a nonnegative solution of (4-7), we see that each component v_i satisfies $\Delta_\theta v_i \le 0$ on $\mathbb{S}^{n-1}$. This implies that v_i does not attain any strict local minimum on $\mathbb{S}^{n-1}$. As $\mathbb{S}^{n-1}$ is a compact manifold, v_i must be a constant. This argument holds for all $1 \le i \le m$, which makes $\boldsymbol{v}$ a nonnegative, constant vector on $\mathbb{S}^{n-1}$. However, a nonnegative constant solution of (4-7) must be trivial because $\lambda \le 0$. Returning to $\boldsymbol{u}$, it indicates that $\boldsymbol{u}$ is trivial on ∂B_1. As each of its components is nonnegative and superharmonic, $\boldsymbol{u}$ must be trivial in the whole domain, which proves Lemma 4.5(i).

Case 2: $\frac{n}{n-2} < \alpha < \frac{n+2}{n-2}$. Suppose that $\boldsymbol{u}$ is a nontrivial solution in the punctured space. Then by the nonnegativity and the superharmonicity of each component of $\boldsymbol{u}$, we know $|\boldsymbol{u}|$ is positive everywhere. As it is homogeneous of degree $-\frac{2}{\alpha-1}$, $\boldsymbol{u}$ must have a nonremovable singularity at the origin, i.e., (4-1) holds. By Lemma 4.1, $\boldsymbol{u}$ is radially symmetric, whence $\boldsymbol{u}$ is a positive constant vector, $\boldsymbol{a}$, on ∂B_1.

By (4-7) we have $|\boldsymbol{a}| = \lambda^{1/(\alpha-1)}$. By the homogeneity, we see that $\boldsymbol{u}$ is of the form $\lambda^{1/(\alpha-1)}|x|^{-2/(\alpha-1)}\boldsymbol{e}$ with some nonnegative unit vector $\boldsymbol{e} \in \mathbb{R}^m$, proving Lemma 4.5(ii). □

We are in a position to prove Theorem 1.3(i) and (ii).

Proof of Theorem 1.3(i) and (ii). Let $\boldsymbol{u}_0$ and $\boldsymbol{u}_\infty$ be a blowup and, respectively, a shrink-down of $\boldsymbol{u}$. According to Lemma 4.4, both $\boldsymbol{u}_0$ and $\boldsymbol{u}_\infty$ are homogeneous of degree $-\frac{2}{\alpha-1}$. Hence, it follows from Lemma 4.5(i) that if $1 < \alpha \le \frac{n}{n-2}$, both $\boldsymbol{u}_0$ and $\boldsymbol{u}_\infty$ are trivial. This in turn yields by Lemma 4.4 that $\Phi(0+, \boldsymbol{u}) = \Phi(\infty, \boldsymbol{u}) = 0$. Due to the monotonicity of $\Phi(r, \boldsymbol{u})$, we have $\Phi(r, \boldsymbol{u}) = 0$ for all $r > 0$. Thus, by Proposition 2.1(i), $\boldsymbol{u}$ is homogeneous of degree $-\frac{2}{\alpha-1}$. Theorem 1.3(i) is now an immediate consequence of Lemma 4.5(i).

Now let us consider the case $\frac{n}{n-2} < \alpha < \frac{n+2}{n-2}$. By Lemmas 4.4 and 4.5(ii), any blowup $\boldsymbol{u}_0$ is either trivial or of the form (1-12). If $\boldsymbol{u}_0$ is trivial, then clearly $\Phi(r, \boldsymbol{u}_0) = 0$ for all $r > 0$, which along with Lemma 4.4 implies that $\Phi(0+, \boldsymbol{u}) = 0$. On the other hand, if $\boldsymbol{u}_0$ is of the form (1-12), then a simple computation shows that $\Phi(r, \boldsymbol{u}_0) = -\bar{\lambda}$ for all $r > 0$, with $\bar{\lambda}$ given as in (1-21). Thus, again from Lemma 4.4 it follows that $\Phi(0+, \boldsymbol{u}) = -\bar{\lambda}$. The converse statement is obviously true, whence we have proved that $\Phi(0+, \boldsymbol{u}) \in \{-\bar{\lambda}, 0\}$, and $\Phi(0+, \boldsymbol{u}) = 0$ if and only if all the blowups are trivial, while $\Phi(0+, \boldsymbol{u}) = -\bar{\lambda}$ if and only if all the blowups are of the form (1-12).

Further, the same assertion holds for any shrink-down $\boldsymbol{u}_\infty$, proving that $\Phi(\infty, \boldsymbol{u}) \in \{-\bar{\lambda}, 0\}$, and $\Phi(\infty, \boldsymbol{u}) = 0$ if and only if all the shrink-downs are trivial, while $\Phi(\infty, \boldsymbol{u}) = -\bar{\lambda}$ if and only if all the shrink-downs are of the form (1-12).

Now if $\Phi(0+, \boldsymbol{u}) = 0$, then since $\Phi(r, \boldsymbol{u})$ is nondecreasing in r and $\Phi(\infty, \boldsymbol{u}) \in \{-\bar{\lambda}, 0\}$, we must have $\Phi(r, \boldsymbol{u}) = 0$ for all $r > 0$. Hence, by Lemmas 4.4 and 4.5(ii), $\boldsymbol{u}$ is either trivial or of the form (1-12). However, the latter yields that $\Phi(0+, \boldsymbol{u}) = -\bar{\lambda}$, a contradiction. Thus, $\boldsymbol{u}$ must be trivial. Of course, the converse is also true.

Similarly, $\Phi(\infty, \boldsymbol{u}) = -\bar{\lambda}$ implies $\boldsymbol{u}$ is of the form (1-12). This finishes the proof of Theorem 1.3(ii). □

The analysis on the case $\alpha = \frac{n+2}{n-2}$ is more subtle. Our approach relies on the Pohozaev invariants of which the first one $\kappa(\boldsymbol{u})$ was introduced in (2-8). In the following we focus on the second Pohozaev invariant $\kappa_*(\boldsymbol{u})$, which was briefly introduced in Remark 2.5. More importantly, we shall observe that this second invariant appears solely due to the multiplicity of the components of (1-1).

Lemma 4.6. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $\mathbb{R}^n \setminus \{0\}$ with $\alpha = \frac{n+2}{n-2}$. Then $\Phi(r, \boldsymbol{u})$ and $\Phi_*(r, \boldsymbol{u})$ in* (2-1) *and* (2-10) *are well-defined, and there are constants $\kappa(\boldsymbol{u})$ and $\kappa_*(\boldsymbol{u})$ satisfying* (2-8) *and* (2-17) *respectively. Moreover, the inequalities* (1-13) *and* (1-14) *hold and the equalities of the lower bounds only occur simultaneously.*

Proof. The proof can be divided into two cases; first we consider the case where $\boldsymbol{u}$ is not rotationally symmetric, and then we treat the other case. We shall prove the equivalent statements for the cylindrical transformation $\boldsymbol{v}$. Since $\boldsymbol{v}$ will be fixed throughout the proof, we shall omit the dependence of Ψ, Ψ_*, κ and κ_* on $\boldsymbol{v}$ here.

Suppose that $\boldsymbol{u}$ is not rotationally symmetric. Due to Lemma 4.1, $\boldsymbol{u}$ has a removable singularity at the origin. Thus, its cylindrical transformation $\boldsymbol{v}$, given as in (1-2), satisfies

$$|\boldsymbol{v}(t,\theta)| + |\partial_t \boldsymbol{v}(t,\theta)| \le Ce^{-\frac{n-2}{2}t} \quad \text{on } \mathbb{S}^{n-1} \tag{4-8}$$

as $t \to \infty$, with some constant $C > 0$ independent of t. This combined with (2-9) implies

$$\kappa = \lim_{t\to\infty} \Psi(t) = 0. \tag{4-9}$$

On the other hand, the estimate (4-8) also ensures the well-definedness of $\Psi_*(t)$ given by (2-15) for all $t \in \mathbb{R}$. To prove that $\Psi_*(t)$ is constant for any $t \in \mathbb{R}$, we need to compute the derivatives of g, given by (2-16). Utilizing (1-3), (2-9) and (4-9) one can verify that

$$g'' = \frac{2}{n\omega_n} \int_{\mathbb{S}^{n-1}} \left(\frac{(n-2)^2}{2} |\boldsymbol{v}|^2 + 2|\nabla_\theta \boldsymbol{v}|^2 - \frac{2n-2}{n} |\boldsymbol{v}|^{\frac{2n}{n-2}} \right) d\theta,$$

from which it follows that

$$\begin{aligned} \Psi_*'(t) &= g' \left(\frac{g''}{2} - \frac{(n-2)^2}{2} g - \frac{2}{n\omega_n} \int_{\mathbb{S}^{n-1}} \left(|\nabla_\theta \boldsymbol{v}|^2 - \frac{n-1}{n} |\boldsymbol{v}|^{\frac{2n}{n-2}} \right) d\theta \right) \\ &= \frac{g'}{n\omega_n} \int_{\mathbb{S}^{n-1}} \left(|\partial_t \boldsymbol{v}|^2 - \frac{(n-2)^2}{4} |\boldsymbol{v}|^2 - |\nabla_\theta \boldsymbol{v}|^2 + \frac{n-2}{n} |\boldsymbol{v}|^{\frac{2n}{n-2}} \right) d\theta = 0. \end{aligned} \tag{4-10}$$

Thus, $\Psi_*(t)$ is constant for any $t \in \mathbb{R}$, and there must exist a constant $\kappa_*(\boldsymbol{v})$ such that (2-18) holds for all t. Moreover, one can also verify from (4-8) that

$$\kappa_* = \lim_{t\to\infty} \Psi_*(t) = 0.$$

This proves the lemma for the case where $\boldsymbol{u}$ is not rotationally symmetric.

Next we consider the case where $\boldsymbol{u}$ is rotationally symmetric, so that the cylindrical transformation $\boldsymbol{v}$ becomes a function of t only. In this case, we have already observed that (1-7) holds with κ_* given by (1-9). Note that under the rotational symmetry of $\boldsymbol{v}$, g as in (2-16) is identical to $|\boldsymbol{v}|^2$. Hence, one can easily observe from (2-15) and (1-7) that

$$\Psi_*(t) = \frac{(g')^2}{4} - \frac{(n-2)^2}{4} g^2 - \kappa g + \frac{n-2}{n} g^{\frac{2n-2}{n-2}} = \kappa_*, \tag{4-11}$$

as desired.

Let us now prove the bounds in (1-13) and (1-14). Since we have already verified above that $\kappa = \kappa_* = 0$ if $\boldsymbol{v}$ is not rotationally symmetric, it suffices to consider the situation where $\boldsymbol{v}$ is rotationally symmetric. Then one can follow the derivation of (1-9) and verify that $\kappa_* \leq 0$. Hence, we are only left with proving the lower bounds of κ and κ_*.

Set

$$f(s) = \frac{(n-2)^2}{4} s^2 - \frac{n-2}{n} s^{\frac{2n-2}{n-2}} + \kappa s,$$

and let us rephrase the second identity in (4-11) as

$$\frac{(g')^2}{4} = f(g) + \kappa_*. \tag{4-12}$$

Utilizing $\kappa_* \leq 0$ in the identity above, we see that $f(g) \geq 0$. Since either $g(t) = 0$ and $g(t) > 0$ for all t, and $g(t) = 0$ yields $\kappa = 0$, we can focus on the case $g(t) > 0$ for all t. Then $\frac{1}{g} f(g) \geq 0$ as well, from which it follows that

$$\kappa \geq -\frac{(n-2)^2}{4} g + \frac{n-2}{n} g^{\frac{n}{n-2}} \geq -\frac{2}{n}\left(\frac{n-2}{2}\right)^n.$$

This verifies the lower bound (1-13) of κ.

To verify the lower bound (1-14) of κ_*, let us remark that

$$\left(\frac{2}{n}\left(\frac{n-2}{2}\right)^n + \kappa\right)\left(\frac{n-2}{2}\right)^{n-2} = f\left(\left(\frac{n-2}{2}\right)^{n-2}\right).$$

Now suppose towards a contradiction that there is a solution $\boldsymbol{v}$ having $\kappa_* < -f\left(\left(\frac{n-2}{2}\right)^{n-2}\right)$. Then it follows from (4-12) that $\min\{g(t) : t \in \mathbb{R}\} > \left(\frac{n-2}{2}\right)^{n-2}$, or equivalently, $\min\{|\boldsymbol{v}(t)| : t \in \mathbb{R}\} > \left(\frac{n-2}{2}\right)^{(n-2)/2}$. In view of (1-3), this implies

$$v_i'' = \frac{(n-2)^2}{4} v_i - |\boldsymbol{v}|^{\frac{4}{n-2}} v_i \leq -\delta v_i \tag{4-13}$$

for each $1 \leq i \leq m$, where $\delta = \min\{|\boldsymbol{v}(t)| : t \in \mathbb{R}\} - \left(\frac{n-2}{2}\right)^{(n-2)/2} > 0$. Hence, v_i is a concave function. However, (4-5) shows that v_i is uniformly bounded for all t, which indicates that $v_i(t) \to a_i$ and $v_i''(t) \to 0$ as $t \to \infty$ for some $a_i > 0$. However, this is a contradiction against (4-13), which proves the lower bound (1-14) of κ_*.

Finally, let us investigate the scenario when the equalities of the lower bounds in (1-13) and (1-14) hold. Suppose that the equality of the lower bound in (1-14) occurs. That is,

$$\kappa + \left(\frac{n-2}{2}\right)^{2-n} \kappa_* = -\frac{2}{n}\left(\frac{n-2}{2}\right)^n. \tag{4-14}$$

Arguing much as above, one can deduce that $\min\{|\boldsymbol{v}(t)| : t \in \mathbb{R}\} \geq \left(\frac{n-2}{2}\right)^{(n-2)/2}$ and $v_i'' \leq 0$ in $\mathbb{R}$ for each $1 \leq i \leq m$. Again v_i is a concave function that is uniformly bounded in $\mathbb{R}$, so $v_i(t) \to a_i$ for some $a_i \in \mathbb{R}$, and $v_i''(t) \to 0$ as $t \to \infty$. Thus, $|\boldsymbol{v}(t)| \to |\boldsymbol{a}|$ with $\boldsymbol{a} = (a_1, \dots, a_m)$, and it follows from $v_i''(t) \to 0$ and the first equality in (4-13) that $|\boldsymbol{a}| = \left(\frac{n-2}{2}\right)^{(n-2)/2}$. On the other hand, we also have $v_i'(t) \to 0$ as $t \to \infty$, so sending $t \to \infty$ in the second equality of (2-9) yields that

$$\kappa = \lim_{t\to\infty}\left(|\boldsymbol{v}'(t)|^2 - \frac{(n-2)^2}{4}|\boldsymbol{v}(t)|^2 + \frac{n-2}{n}|\boldsymbol{v}(t)|^{\frac{2n}{n-2}}\right) = -\frac{2}{n}\left(\frac{n-2}{2}\right)^n.$$

Thus, (4-14) forces $\kappa_* = 0$, and the final assertion of the lemma is proved. ☐

Let us finish this section by proving Theorem 1.3(iii).

Proof of Theorem 1.3(*iii*). The well-definedness and the bounds of κ and κ_* are proved in Lemma 4.6. The other assertions can be proved as follows.

First consider the assertion (iii)-(a). If $\boldsymbol{u}$ is not radially symmetric, then by Lemma 4.1, $\boldsymbol{u}$ has a removable singularity at the origin, as desired. On the other hand, if $\boldsymbol{u}$ is radially symmetric, one can deduce from (1-7) that the cylindrical transformation $\boldsymbol{v}$, which is now a function of t only, satisfies

$$\left(\frac{d|\boldsymbol{v}|}{dt}\right)^2 = \frac{(n-2)^2}{4}|\boldsymbol{v}|^2 - \frac{n-2}{n}|\boldsymbol{v}|^{\frac{2n}{n-2}}. \tag{4-15}$$

Hence, classical work such as [Fowler 1931; Caffarelli et al. 1989] applies to $|\boldsymbol{v}|$, proving the "only if" part of the assertion (iii)-(a). The "if" part can be verified through a direct computation.

Let us move on to the case $\kappa^2 + \kappa_*^2 > 0$. From the assertion (iii)-(a), we see that $\boldsymbol{u}$ must have a nonremovable singularity at the origin. According to Lemma 4.1, $\boldsymbol{u}$ is radially symmetric, so one can follow the computation in Section 1 and deduce (1-7).

Finally, assume that $\kappa = -\frac{2}{n}\left(\frac{n-2}{2}\right)^n$ and $\kappa_* = 0$. It follows from (1-7) that

$$\left(\frac{d|\boldsymbol{v}|}{dt}\right)^2 - \frac{(n-2)^2}{4}|\boldsymbol{v}|^2 + \frac{n-2}{n}|\boldsymbol{v}|^{\frac{2n}{n-2}} + \frac{2}{n}\left(\frac{n-2}{2}\right)^n = 0,$$

whence $|\boldsymbol{v}|$ has to be constant in $\mathbb{R}$, and the constant has to be $\left(\frac{n-2}{2}\right)^{(n-2)/2}$. In terms of $\boldsymbol{u}$ this implies that $\boldsymbol{u}$ is homogeneous of degree $-\frac{n-2}{2}$ and is of the form (1-15). This constitutes the "only if" part of the assertion (iii)-(c). The "if" part follows easily from a direct computation. ☐

5. A priori estimate and Harnack-type inequality for local solutions

In this section, we prove a priori upper bounds for local solutions of (1-1) in $B_1 \setminus \{0\}$ with $1 < \alpha \leq \frac{n+2}{n-2}$, which further allows us to derive related Harnack inequalities, interior gradient estimates and the compactness of scaling functions. Our analysis is divided into two cases, according to the subcritical range $1 < \alpha < \frac{n+2}{n-2}$ and the critical range $\alpha = \frac{n+2}{n-2}$. The former is based on the nonexistence of the smooth, positive, global solution in Theorem 1.1(i) along with a blowup argument. The latter uses the method of moving spheres presented in the previous section, essentially following [Li and Zhang 2003].

5A. ***A priori bound for*** $\boldsymbol{1 < \alpha < \frac{n+2}{n-2}}$**.** We begin with the upper bound for the subcritical case, which is (much) simpler than the critical case.

Proposition 5.1. *Let* $1 < \alpha < \frac{n+2}{n-2}$ *and suppose that* $\boldsymbol{v} \in C^2(B_1;\mathbb{R}^m) \cap C(\bar{B}_1;\mathbb{R}^m)$ *is a nonnegative solution of*

$$-\Delta \boldsymbol{v} = |\boldsymbol{v}|^{\alpha-1}\boldsymbol{v} \quad \text{in } B_1. \tag{5-1}$$

Then there exists $C > 0$, *depending only on* n, m *and* α, *such that*

$$|\boldsymbol{v}(x)| \le C(1-|x|)^{-\frac{2}{\alpha-1}} \quad \text{in } B_1. \tag{5-2}$$

Proof. Note that $w = v_1 + \cdots + v_m$ satisfies

$$\frac{1}{c}w^\alpha \le -\Delta w \le c w^\alpha$$

for some $c > 1$, depending only on m and α. Thus, we can follow the proof of [Poláčik et al. 2007, Theorem 2.1] and obtain the desired inequality. We omit the details. □

5B. ***A Harnack-type inequality for*** $\boldsymbol{\alpha = \frac{n+2}{n-2}}$**.** Our approach to achieve the Harnack-type inequality for $\alpha = \frac{n+2}{n-2}$ follows the line of the scalar case in [Li and Zhang 2003, Lemma 5.1]. In our system setting, the problem becomes very sensitive to the number of nonzero components, and we modify the proof of [Li and Zhang 2003, Lemma 5.1] in this direction.

Proposition 5.2. *Let* $\boldsymbol{v} \in C^2(B_2;\mathbb{R}^m) \cap C(\bar{B}_2;\mathbb{R}^m)$ *be a nonnegative solution of*

$$-\Delta \boldsymbol{v} = |\boldsymbol{v}|^{\frac{4}{n-2}}\boldsymbol{v} \quad \text{in } B_2. \tag{5-3}$$

Then, there exists $C > 0$ *depending only on* n *and* m, *such that*

$$\big(\min_{i\in I_m} \inf_{\partial B_2} v_i\big)|\boldsymbol{v}(x)| \le C(1-|x|)^{-\frac{n-2}{2}} \quad \text{in } B_1, \tag{5-4}$$

where I_m *is the set of indices* $1 \le i \le m$ *such that* v_i *is nontrivial.*

Proof. If $\boldsymbol{v}$ is trivial, then $I_m = \varnothing$, whence there is nothing to prove. Thus, we shall assume that $\boldsymbol{v}$ is not trivial, so that $I_m \neq \varnothing$. Then for each $i \in I_m$, we know from the superharmonicity and the nonnegativity of v_i that $\inf_{\partial B_2} v_i > 0$, whence $(\min_{i\in I_m} \inf_{\partial B_2} v_i)^{-1}$ is a positive, finite number.

If $|\boldsymbol{v}(x)| \le C_1(1-|x|)^{-(n-2)/2}$ in B_1 for some $C_1 > 0$ depending only on n and m, then the claim (5-4) is true, since the maximum principle and the superharmonicity of each component of $\boldsymbol{v}$ implies that $\inf_{\partial B_2} v_i \le v_i(0)$. Thus, let us assume that for all $j \ge 1$ there are nonnegative solutions $\boldsymbol{v}_j$ of (5-3) and points $x_j \in \bar{B}_1$ such that

$$M_j := \sup_{|x|\le 1} \big((1-|x|)^{\frac{n-2}{2}}|\boldsymbol{v}_j(x)|\big) = (1-|x_j|)^{\frac{n-2}{2}}|\boldsymbol{v}_j(x_j)| \to \infty. \tag{5-5}$$

We know that $x_j \in B_1$ (instead of ∂B_1) since $\boldsymbol{v}_j$ is continuous on $\overline{B}_1$. Moreover, we shall set

$$r_j = \tfrac{1}{2}(1-|x_j|) > 0, \tag{5-6}$$

$$\delta_j = |\boldsymbol{v}_j(x_j)|^{-\frac{\alpha-1}{2}} = 2r_j M_j^{-\frac{2}{n-2}} \to 0, \tag{5-7}$$

$$R_j = \frac{r_j}{\delta_j} = \tfrac{1}{2} M_j^{\frac{2}{n-2}} \to \infty. \tag{5-8}$$

It should be noted that due to (5-5), we have

$$|\boldsymbol{v}_j(x)| \le \left(\frac{1-|x_j|}{1-|x|}\right)^{\frac{2}{\alpha-1}} |\boldsymbol{v}_j(x_j)| \le 2^{\frac{2}{\alpha-1}} |\boldsymbol{v}_j(x_j)| \quad \text{in } B_{r_j}(x_j). \tag{5-9}$$

In addition, inserting (5-6) into (5-5), we obtain

$$|\boldsymbol{v}_j(x_j)| = (2r_j)^{-\frac{2}{\alpha-1}} M_j. \tag{5-10}$$

With (5-9) and (5-10) at hand, one can follow the proof of [Li and Zhang 2003, Lemma 5.1] to deduce that the sequence of the scaled function

$$\boldsymbol{w}_j(x) = \delta_j^{\frac{n-2}{2}} \boldsymbol{v}_j(\delta_j x + x_j) \quad \text{in } B_{R_j}$$

converges to w_0 in $C^2_{\text{loc}}(\mathbb{R}^n;\mathbb{R}^m)$ for certain $\boldsymbol{w}_0 \in C^2(\mathbb{R}^n;\mathbb{R}^m)$, which is a nonnegative solution of

$$-\Delta \boldsymbol{w}_0 = |\boldsymbol{w}_0|^{\frac{4}{n-2}} \boldsymbol{w}_0 \quad \text{in } \mathbb{R}^n \tag{5-11}$$

satisfying

$$|\boldsymbol{w}_0(x)| \le 2^{\frac{2}{\alpha-1}} \quad \text{in } \mathbb{R}^n, \tag{5-12}$$

as well as

$$|\boldsymbol{w}_0(0)| = 1. \tag{5-13}$$

We omit the details here.

With only a minor modification, one may apply Lemma 3.1 to each component $w_{i,j}$ of $\boldsymbol{w}_j$, with $i \in I_m$, and obtain a number $s_{i,j}(z) > 0$, corresponding to each $z \in \mathbb{R}^n$, such that for all $0 < r < s_{i,j}(z)$

$$(w_{i,j})^*_{z,r} \le w_{i,j} \quad \text{in } B_{1/(2\delta_j)}(z) \setminus B_r(z). \tag{5-14}$$

Here we choose j large enough so that $B_{1/(2\delta_j)}(z) \subset B_{1/\delta_j}$, which is possible due to (5-7). One may refer to the proof of [Li and Zhang 2003, Theorem 1.5] for the details.

Let us now replace $s_{i,j}(z)$ by the supremum value of r such that (5-14) holds; that is,

$$s_{i,j}(z) = \sup\{r : (w_{i,j})^*_{z,\rho} \le w_{i,j} \text{ in } B_{1/(2\delta_j)}(z) \setminus B_r(z) \text{ for any } 0 < \rho < r\}. \tag{5-15}$$

Now with $s_{i,j}(z)$ defined as in (5-15), we shall set, analogously to (3-8),

$$\bar{s}_j(z) = \inf_{i \in I_m} s_{i,j}(z). \tag{5-16}$$

Then we have

$$(w_{i,j})^*_{z,\bar{s}_j(z)} \le w_{i,j} \quad \text{in } B_{1/(2\delta_j)}(z) \setminus B_{\bar{s}_j(z)}(z) \text{ for each } i \in I_m, \tag{5-17}$$

and respectively,

$$-\Delta(w_{i,j} - (w_{i,j})^*_{z,\bar{s}_j(z)}) \geq 0 \quad \text{in } B_{1/(2\delta_j)}(z) \setminus \bar{B}_{\bar{s}_j(z)}(z). \tag{5-18}$$

Now let us assume towards a contradiction that

$$\min_{i \in I_m} \inf_{\partial B_2} v_{i,j} \geq j \Big(\sup_{|x| \leq 1} (1 - |x|)^{\frac{n-2}{2}} |\boldsymbol{v}_j(x)| \Big)^{-1} = \frac{j}{M_j}. \tag{5-19}$$

In terms of $w_{i,j}$, one may rewrite (5-19) as

$$\min_{i \in I_m} \inf_{\partial B_{1/\delta_j}} w_{i,j} = \delta_j^{\frac{n-2}{2}} \min_{i \in I_m} \inf_{\partial B_1(x_j)} v_{i,j} \geq \delta_j^{\frac{n-2}{2}} \min_{i \in I_m} \inf_{\partial B_2} v_{i,j} \geq j \delta_j^{n-2}, \tag{5-20}$$

where in the derivation of the first inequality we used the superharmonicity of $v_{i,j}$, the maximum principle and the fact that $B_1(x_j) \subset B_2$, while the second inequality follows from (5-19), (5-7) and the fact that $2r_j = 1 - |x_j| \leq 1$.

In view of (5-20), one may easily deduce that for any $z \in \mathbb{R}^n$

$$\lim_{j \to \infty} \bar{s}_j(z) = \infty. \tag{5-21}$$

Suppose that (5-21) is false, and there exists some $L > 0$, independent of j, such that

$$\bar{s}_j(z) \leq L. \tag{5-22}$$

Then by the definition of the Kelvin transform (see (3-1)), we have, for any $i \in I_m$,

$$\begin{aligned} \sup_{\partial B_{1/(4\delta_j)}(z)} (w_{i,j})^*_{z,\bar{s}_j(z)} &= (4\delta_j \bar{s}_j(z))^{n-2} \sup_{\partial B_{4\delta_j \bar{s}_j^2(z)}} w_{i,j} \\ &\leq (4\delta_j L)^{n-2} \delta_j^{\frac{n-2}{2}} \sup_{B_{4\delta_j^2 L^2}} v_{i,j} \leq (8L)^{n-2} \delta_j^{n-2}, \end{aligned} \tag{5-23}$$

where in deriving the first and the second inequalities we used (5-22) and, respectively, (5-9) with (5-10). According to (5-20) and (5-23), for each $i \in I_m$,

$$\inf_{\partial B_{1/(4\delta_j)}(z)} (w_{i,j} - (w_{i,j})^*_{z,\bar{s}_j(z)}) \geq (j - (8L)^{n-2}) \delta_j^{n-2} > 0 \tag{5-24}$$

for all sufficiently large j, where in the first inequality we used $w_{i,j} \geq \inf_{\partial B_{1/\delta_j}} w_{i,j}$ on $\partial B_{1/(4\delta_j)}(z)$, which follows from the maximum principle, the superharmonicity of $w_{i,j}$ in B_{1/δ_j} and the fact that $B_{1/(4\delta_j)(z)} \subset B_{1/\delta_j}$. With (5-24) at hand, we may apply the maximum principle to (5-18) and observe that for any $i \in I_m$

$$(w_{i,j})^*_{z,\bar{s}_j(z)} < w_{i,j} \quad \text{in } B_{1/(2\delta_j)(z)} \setminus \bar{B}_{\bar{s}_j(z)}(z). \tag{5-25}$$

Now that $w_{i,j}$ satisfies (5-18) and (5-25) for each $i \in I_m$, we can follow a similar argument to that in the proof of [Li and Zhang 2003, Lemma 5.2] and deduce that there exist $\bar{s}_{i,j}(z) > \bar{s}_j(z)$ and $0 < \epsilon_{i,j} < \bar{s}_{i,j}(z) - \bar{s}_j(z)$ such that for any $\bar{s}_j(z) < r < \bar{s}_j(z) + \epsilon_{i,j}$

$$(w_{i,j})^*_{z,\bar{s}_j(z)} < w_{i,j} \quad \text{in } B_{1/(2\delta_j)(z)} \setminus B_r(z) \text{ for each } i \in I_m. \tag{5-26}$$

Clearly, (5-26) violates the definition of $\bar{s}_j(z)$ in (5-16). Hence, the claim (5-21) should be true, under the assumption (5-19).

Knowing that (5-20) is true for all $z \in \mathbb{R}^n$ (under the assumption (5-19)), we have for any $z \in \mathbb{R}^n$ and $r > 0$ that

$$(w_{i,j})^*_{z,r} \le w_{i,j} \quad \text{in } B_{1/(2\delta_j)}(z) \setminus B_r(z) \text{ for any } i \in I_m \tag{5-27}$$

for all sufficiently large j such that $\bar{s}_j(z) > r$. On the other hand, recall from the beginning of this proof that $\boldsymbol{w}_j \to \boldsymbol{w}_0$ in $C^2_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^m)$ with some $\boldsymbol{w}_0 \in C^2(\mathbb{R}^n;\mathbb{R}^m)$ satisfying (5-11), (5-12) and (5-13) with $\alpha = \frac{n+2}{n-2}$. This implies $(\boldsymbol{w}_j)^*_{z,r} \to (\boldsymbol{w}_0)^*_{z,r}$ in $C^2_{\mathrm{loc}}(\mathbb{R}^n \setminus \{z\};\mathbb{R}^m)$ for each $z \in \mathbb{R}^n$ and any $r > 0$. Thus, we may pass to the limit with $j \to \infty$ (possibly along a subsequence) in (5-27) in any compact domain of type $B_R(z) \setminus B_r(z) \subset \mathbb{R}^n \setminus \{z\}$, which gives

$$(w_{i,0})^*_{z,r} \le w_{i,0} \quad \text{in } \mathbb{R}^n \setminus B_r(z) \text{ for any } i \in I_m. \tag{5-28}$$

As $z \in \mathbb{R}^n$ and $r > 0$ in (5-28) are arbitrary, we conclude from [Li and Zhang 2003, Lemma 11.2] that $w_{i,0}$ is constant for each $i \in I_m$. Then as $w_{i,0}$ is a nonnegative (global) solution of (5-11), $w_{i,0}$ must be trivial for each $i \in I_m$. On the other hand, for any $i \notin I_m$, v_i is already trivial and so is the limit $w_{i,0}$. Consequently, $\boldsymbol{w}_0$ is a trivial solution, a contradiction with (5-13). Therefore, the assumption (5-19) must fail, which implies (5-4) with some constant $C > 0$, depending only on n and m. ☐

5C. *Universal upper bounds for $1 < \alpha \le \frac{n+2}{n-2}$.* With Proposition 5.1, we obtain a universal upper estimate for (local) singular solutions in the subcritical case. Let us remark that this bound is not sharp for $1 < \alpha \le \frac{n}{n-2}$, although we obtain a universal constant as well as a universal neighborhood in the estimate. The sharp bounds for those cases will be given separately in Sections 7C and 7D.

Lemma 5.3. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1 \setminus \{0\}$ with $1 < \alpha < \frac{n+2}{n-2}$. Then there exists $C > 0$, depending only on n, m and α, such that*

$$|\boldsymbol{u}(x)| \le C|x|^{-\frac{2}{\alpha-1}} \quad \textit{in } B_{1/2} \setminus \{0\}. \tag{5-29}$$

Proof. Let $x_0 \in B_{1/2} \setminus \{0\}$ and set $r = \frac{1}{2}|x_0|$. Since $\bar{B}_r(x_0) \subset B_1 \setminus \{0\}$, one can define

$$\boldsymbol{v}(x) = r^{\frac{2}{\alpha-1}} \boldsymbol{u}(rx + x_0) \quad \text{in } \bar{B}_1.$$

As $\boldsymbol{u}$ is a nonnegative solution of (1-1) in $B_1 \setminus \{0\}$, we see that $\boldsymbol{v}$ is a nonnegative solution of (5-1). Moreover, $\boldsymbol{v}$ is continuous up to the boundary of B_1. Hence, Proposition 5.1 applies to $\boldsymbol{v}$ and taking $x = 0$ in (5-2) we obtain

$$|\boldsymbol{v}(0)| \le C,$$

which in terms of $\boldsymbol{u}$ can be rephrased as

$$|\boldsymbol{u}(x_0)| \le C r^{-\frac{2}{\alpha-1}}.$$

Since $x_0 \in B_{1/2} \setminus \{0\}$ was arbitrary and $r = \frac{1}{2}|x_0|$, the proof is finished. ☐

Remark 5.4. For $1<\alpha<\frac{n+2}{n-2}$, one may take an alternative approach as follows. Let $w=u_1+u_2+\cdots+u_m$. Then $w\geq 0$ and $\frac{1}{c_1}w\leq|\boldsymbol{u}|\leq c_1 w$ in $B_1\setminus\{0\}$ with $c_1=m^{1/2}$. Hence, w satisfies $\frac{1}{c_2}w^\alpha\leq-\Delta w\leq c_2 w^\alpha$ in $B_1\setminus\{0\}$ with $c_2=m^{(\alpha-1)/2}$. By [Serrin and Zou 2002, Corollary IV] it follows that $w\leq C|x|^{-2/(\alpha-1)}$ in $B_{1/2}\setminus\{0\}$, where C depends only on n, m and α. This together with the inequality $|\boldsymbol{u}|\leq c_1 w$ yields (5-29).

From the Harnack-type inequality in Proposition 5.2, we obtain an upper estimate for the critical case $\alpha=\frac{n+2}{n-2}$.

Lemma 5.5. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1\setminus\{0\}$ with $\alpha=\frac{n+2}{n-2}$. Then there exists $C>0$, depending only on n and m, such that*

$$\Big(\min_{i\in I_m}\inf_{\partial B_{3/4}}u_i\Big)|\boldsymbol{u}(x)|\leq C|x|^{-\frac{n-2}{2}}\quad \text{in } B_{1/2}\setminus\{0\}, \tag{5-30}$$

where I_m consists of all indices $1\leq i\leq m$ such that u_i is nontrivial.

Proof. If $\boldsymbol{u}$ has a removable singularity at the origin, then $-\Delta\boldsymbol{u}=|\boldsymbol{u}|^{4/(n-2)}\boldsymbol{u}$ in B_1 (instead of $B_1\setminus\{0\}$), whence one may apply Proposition 5.2 to $\boldsymbol{u}$ after scaling, and observe that

$$\Big(\min_{i\in I_m}\inf_{\partial B_{3/4}}u_i\Big)|\boldsymbol{u}(x)|\leq C\big(\tfrac{3}{4}-|x|\big)^{-\frac{n-2}{2}}\leq C\big(\tfrac{3}{4}\big)^{-\frac{n-2}{2}}\quad\text{in } B_{1/2}\setminus\{0\},$$

which implies (5-30).

Henceforth, let us assume that $\boldsymbol{u}$ does not have a removable singularity at the origin. Clearly $I_m\neq\varnothing$, and by the superharmonicity and the nonnegativity of u_i with $i\in I_m$, we have $u_i>0$ in $B_1\setminus\{0\}$ for all $i\in I_m$.

Now let $x_0\in B_{1/2}\setminus\{0\}$ and $r=\frac{1}{8}|x_0|$. Since $\overline{B}_{2r}(x_0)\subset B_1\setminus\{0\}$, one can define

$$\boldsymbol{v}(x)=r^{\frac{n-2}{2}}\boldsymbol{u}(rx+x_0)\quad\text{in } \overline{B}_2.$$

Obviously, v_i is nontrivial if and only if $i\in I_m$. On the other hand, as $\boldsymbol{u}$ is a nonnegative solution of (1-1) in $B_1\setminus\{0\}$, $\boldsymbol{v}$ becomes a nonnegative solution of (5-3). Hence, it follows from (5-4) that

$$|\boldsymbol{v}(0)|\leq C\Big(\min_{i\in I_m}\inf_{\partial B_2}v_i\Big)^{-1}=C\Big(\min_{i\in I_m}\inf_{B_{2r}(x_0)}u_i\Big)^{-1}, \tag{5-31}$$

where $C>0$ depends only on n and m.

Now let $J_m\subset I_m$ consist of all components u_i having nonremovable singularity at the origin. Note that J_m may not be equal to I_m. By superharmonicity and positivity, the maximum principle implies that $\liminf_{x\to 0}u_i(x)=\infty$ for each $i\in J_m$. On the other hand, if $i\in I_m\setminus J_m$ (provided that $I_m\setminus J_m\neq\varnothing$), u_i is bounded at the origin, and again by the maximum principle, one has $\liminf_{x\to 0}u_i(x)\geq\inf_{\partial B_{3/4}}u_i$. Hence, one should have $\inf_{\partial B_{2r}(x_0)}u_i\geq\inf_{\partial B_{3/4}}u_i$ for any $i\in I_m$. This along with (5-31) yields

$$|\boldsymbol{u}(x_0)|\leq C\Big(\min_{i\in I_m}\inf_{B_{3/4}}u_i\Big)^{-1}r^{-\frac{n-2}{2}},$$

which proves the lemma. □

Remark 5.6. We shall see in Section 7B that the above estimate can be improved for solutions $\boldsymbol{u}$ with nonremovable singularity at the origin.

Due to Lemmas 5.3 and 5.5, we obtain the standard Harnack inequality and interior gradient estimate.

Lemma 5.7. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in* $B_1 \setminus \{0\}$ *with* $1 < \alpha \le \frac{n+2}{n-2}$. *Then there exists* $C > 0$ *such that for each* $1 \le i \le m$,

$$\sup_{B_r \setminus \bar{B}_{r/2}} u_i \le C \inf_{B_r \setminus \bar{B}_{r/2}} u_i \quad \text{for any } 0 < r < \tfrac{1}{2}, \tag{5-32}$$

and

$$|\nabla u_i(x)| \le C \frac{u_i(x)}{|x|} \quad \text{in } B_{1/2} \setminus \{0\}. \tag{5-33}$$

Moreover, the constant C *in* (5-32) *depends only on* n, m *and* α, *provided that* $1 < \alpha < \frac{n+2}{n-2}$.

Proof. After a scaling argument we may also say that (5-29) and (5-30) hold in $B_{3/4} \setminus \{0\}$, instead of $B_{1/2} \setminus \{0\}$. Consider u_i, $1 \le i \le m$, as a nonnegative solution of $-\Delta u_i = a(x) u_i$ in $B_1 \setminus \{0\}$, where $a(x) = |\boldsymbol{u}|^{\alpha-1}$. Due to (5-29) if $1 < \alpha < \frac{n+2}{n-2}$, and to (5-30) if $\alpha = \frac{n+2}{n-2}$, we know that $0 \le a(x) \le C|x|^{-2}$ in $B_{3/4} \setminus \{0\}$. Thus, (5-32) follows easily from the classical Harnack inequality [Gilbarg and Trudinger 1983, Corollary 9.25]. With (5-32) at hand, one may also prove (5-33) by the classical gradient estimate [Gilbarg and Trudinger 1983, Theorem 3.9]. □

6. Asymptotic radial symmetry of local solutions

This section is devoted to the proof of Theorem 1.4. Let us address that a similar argument was also used in [Caffarelli et al. 2014, Theorem 1.2], which is concerned with fractional Laplacian, scalar equations.

Proof of Theorem 1.4. If the origin is a removable singularity, then the conclusion (1-16) is clear. Hence, we shall assume that the origin is a nonremovable singularity.

Recall from (3-1) that $\boldsymbol{u}^*_{z,r}$ is the Kelvin transform of $\boldsymbol{u}$ with respect to the sphere $\partial B_r(z)$. Since the origin is a nonremovable singularity of $\boldsymbol{u}$, one may prove, with a minor modification of the proof of Lemma 4.1, that there is some small $\epsilon > 0$ such that for any $z \in B_{\epsilon/2} \setminus \{0\}$ and any $0 < r \le |z|$,

$$(u_i)^*_{z,r} \le u_i \quad \text{in } B_1 \setminus (B_r(z) \cup \{0\}) \text{ for each } 1 \le i \le m. \tag{6-1}$$

The key observation here is that (6-1) implies, for any $a > \frac{1}{\epsilon}$ and $e \in \partial B_1$,

$$u_i^*(y) \le u_i^*(y_a) \quad \text{if } y \cdot e > a \text{ and } |y_a| > 1 \text{ for each } 1 \le i \le m, \tag{6-2}$$

where

$$u_i^*(y) = (u_i)^*_{0,1}(y) = |y|^{2-n} u_i(|y|^{-2} y), \qquad y_a = y + 2(a - y \cdot e)e,$$

and $H_a(e)$ is the half-space $\{x : x \cdot e > a\}$. Note that y_a is the reflection point of y with respect to the hyperplane $\partial H_a(e)$. To prove the claim (6-2), let us note first that $y \in B_{1/\epsilon}$ if and only if $y/|y|^2 \in B_\epsilon$. Now we shall choose some $z \in B_{\epsilon/2} \setminus \{0\}$ and some $0 < r < |z|$ such that

$$\frac{y_a}{|y_a|^2} - z = \left(\frac{r}{|y/|y|^2 - z|} \right)^2 \left(\frac{y}{|y|^2} - z \right). \tag{6-3}$$

In other words, $y_a/|y_a|^2$ is the reflection point of $y/|y|^2$ with respect to $\partial B_r(z)$. We shall ask in addition that

$$\frac{|y_a|}{|y|} \le \frac{1}{r}\left|\frac{y}{|y|^2} - z\right|. \tag{6-4}$$

Before we actually find such z and r, let us verify that along with (6-3) and (6-4), (6-1) implies (6-2) as follows.

Given $y \in \mathbb{R}^n$ such that $y \cdot e > a$ and $|y_a| > 1$, and $0 < r < |z| < \frac{\epsilon}{2}$ such that (6-3) and (6-4) hold, let us write by x and $x^*_{z,r}$ the points $y/|y|^2$ and $y_a/|y_a|^2$ respectively. Then since $y \cdot e > a > \frac{1}{\epsilon}$ and $|y_a| > 1$, we have $x \in B_r(z)$, and $x^*_{z,r} \in B_1 \setminus B_r(z)$. Hence, one may proceed, using (6-1), with

$$\begin{aligned} u_i^*(y) &= \frac{1}{|y|^{n-2}}\left(\frac{|x^*_{z,r} - z|}{r}\right)^{n-2} (u_i)^*_{z,r}(x^*_{z,r}) \\ &\le \frac{1}{|y|^{n-2}}\left(\frac{|x^*_{z,r} - z|}{r}\right)^{n-2} u_i(x^*_{z,r}) \le u_i^*(y_a), \end{aligned}$$

proving (6-2), where in deriving the first equality we used (6-3), while the last inequality follows from (6-4). Thus, we only need to prove that there actually exist $0 < r < |z| < \frac{\epsilon}{2}$ satisfying (6-3) and (6-4). However, it only involves an elementary argument to verify (6-3) and (6-4) as well as $0 < r \le |z| < \frac{\epsilon}{2}$, by choosing $r = |z|$ and

$$z = \frac{1}{|y|^2} y + \frac{|y_a|^2}{|y|^2 - |y_a|^2}\left(\frac{1}{|y|^2} y - \frac{1}{|y_a|^2} y_a\right) = \frac{1}{|y|^2 - |y_a|^2}(y - y_a).$$

With the claim (6-2) at hand, one may invoke [Caffarelli et al. 1989, Theorem 6.1 and Corollary 6.2] to finish the proof. That is, from the former one obtains some $C > 0$, independent of ϵ, such that

$$u_i^*(y) \le u_i^*(x) \quad \text{if } |x| > 1 \text{ and } |y| \ge |x| + \frac{C}{\epsilon} \text{ for each } 1 \le i \le m.$$

As u_i^* is a nonnegative superharmonic function, the latter implies

$$u_i^* = \left(1 + O\left(\frac{1}{R}\right)\right)\left(\inf_{\partial B_R} u_i^*\right) \quad \text{uniformly on } \partial B_R \text{ as } R \to \infty,$$

which in terms of u_i implies the asymptotic radial symmetry claimed as in (1-16). □

With the asymptotic radial symmetry as well as the uniform estimate achieved in the previous section, we are ready to prove Proposition 2.4, finally showing the existence of the second Pohozaev invariant; see (2-17).

Proof of Proposition 2.4. Let $\boldsymbol{u}$ be a nonnegative solution of (1-1) in $B_R \setminus \{0\}$ with $\alpha = \frac{n+2}{n-2}$, and let $\Phi_*(r, \boldsymbol{u})$ be as in (2-10). Let us also assume that $\boldsymbol{u}$ is a nontrivial solution. Let us prove the well-definedness of $\Phi_*(r, \boldsymbol{u})$.

In the following, we shall denote by C a positive generic constant independent of r. With $f(r, \boldsymbol{u})$ given as in (2-11), it follows immediately from (5-30) and (5-33) that

$$f(r, \boldsymbol{u}) \le C \quad \text{and} \quad r|\dot{f}(r, \boldsymbol{u})| \le C f(r, \boldsymbol{u}) \quad \text{for any } 0 < r < \tfrac{1}{2}R. \tag{6-5}$$

On the other hand, by the asymptotic radial symmetry (1-16), we have

$$|\Delta(\boldsymbol{u}-\bar{\boldsymbol{u}})| \le C|x||\bar{\boldsymbol{u}}|^{\frac{n+2}{n-2}} \quad \text{in } B_{2r}\setminus \bar{B}_r, \quad \text{as } r\to 0+,$$

where $\bar{\boldsymbol{u}}(r)$ is the average of $\boldsymbol{u}$ over the sphere ∂B_r. Hence, it follows from the interior gradient estimate [Gilbarg and Trudinger 1983, Theorem 3.9] and the Harnack inequality (5-32) that

$$|\nabla(\boldsymbol{u}-\bar{\boldsymbol{u}})| \le C|\boldsymbol{u}| \quad \text{on } \partial B_r,$$

and in particular,

$$|\nabla_\sigma \boldsymbol{u}| \le C|\boldsymbol{u}| \quad \text{on } \partial B_r, \tag{6-6}$$

where $\nabla_\sigma \boldsymbol{u}$ is the tangential derivative of $\boldsymbol{u}$ on ∂B_r.

By means of (6-6) and (6-5), we deduce that

$$\left|\int_0^r \left(\frac{\rho}{n\omega_n}\int_{\partial B_\rho} |\nabla_\sigma \boldsymbol{u}|^2\,d\sigma\right)\dot{f}(\rho,\boldsymbol{u})\,d\rho\right| \le C\int_0^r \rho f(\rho,\boldsymbol{u})^2\,d\rho, \tag{6-7}$$

provided that $r>0$ is sufficiently small. Similarly, one may also prove from (1-16) and (6-5) that

$$\left|\int_0^r \left(\frac{\rho}{n\omega_n}\int_{\partial B_\rho} |\boldsymbol{u}|^{\frac{2n}{n-2}}\,d\rho - f(\rho,\boldsymbol{u})^{\frac{n}{n-2}}\right)\dot{f}(\rho,\boldsymbol{u})\,d\rho\right| \le C\int_0^r \rho f(\rho,\boldsymbol{u})^{\frac{2n-2}{n-2}}\,d\rho. \tag{6-8}$$

By the first inequality in (6-5), we see that the right sides of both (6-7) and (6-8) are of order r^2, proving the well-definedness of $\Phi_*(r,\boldsymbol{u})$.

Proving that $\Phi_*(r,\boldsymbol{u})$ is indeed constant in $0<r<R$ is now easy by considering the cylindrical version $\Psi_*(t,\boldsymbol{v})$ defined as in (2-15). Since the computation is very similar to that of (4-10), we omit the details. □

7. Exact asymptotic behavior of local solutions

With the a priori estimates and the classification of the solutions on the punctured space, we are now ready to investigate exact asymptotic behavior of local solutions near the isolated singularity at the origin. Before we begin our analysis, let us provide the basic integrability of the solution.

Lemma 7.1. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1\setminus\{0\}$ with $\alpha>1$. One has $\boldsymbol{u}\in L^\alpha(B_1;\mathbb{R}^m)$. In particular, if $\alpha\ge \frac{n}{n-2}$, then $\boldsymbol{u}$ is a distribution solution of* (1-1) *in $B_1\setminus\{0\}$ in B_1, that is,*

$$-\int_{B_1} \boldsymbol{u}\cdot\Delta\boldsymbol{v}\,dx = \int_{B_1} |\boldsymbol{u}|^{\alpha-1}\boldsymbol{u}\cdot\boldsymbol{v}\,dx \quad \textit{for any } \boldsymbol{v}\in C_0^\infty(B_1;\mathbb{R}^m).$$

Proof. Recall from the proof of Proposition 5.1 and Remark 5.4 that $w=u_1+\cdots+u_m$ satisfies $\frac{1}{c}w^\alpha \le -\Delta w\le cw^\alpha$, with some $c>1$ depending only on m and α. By [Brézis and Lions 1981], $w\in L^\alpha(B_1)$ which implies that $\boldsymbol{u}\in L^\alpha(B_1;\mathbb{R}^m)$. The second assertion can be proved much as in [Caffarelli et al. 1989], and we omit the details. □

7A. *Case $\frac{n}{n-2} < \alpha < \frac{n+2}{n-2}$.* The upper bound (5-29) and the classification of solutions on the punctured space allow us to capture the exact asymptotic behavior of local solutions to (1-1), by means of the blowup analysis. Let us recall from Section 4 that a blowup $\boldsymbol{u}_0$ is a limit of $\boldsymbol{u}_r$ along a sequence $r = r_j \to 0+$ in $C^2_{\mathrm{loc}}(\mathbb{R}^n \setminus \{0\}; \mathbb{R}^m)$.

Lemma 7.2. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1 \setminus \{0\}$ with $\frac{n}{n-2} < \alpha < \frac{n+2}{n-2}$, and let $\Phi(r, \boldsymbol{u})$ be as in* (2-1). *Then $\Phi(0+, \boldsymbol{u}) \in \{-\bar{\lambda}, 0\}$, where $\bar{\lambda}$ is given by* (1-21). *Moreover, the following are true*:

(i) $\Phi(0+, \boldsymbol{u}) = 0$ *if and only if*

$$|\boldsymbol{u}(x)| = o(|x|^{-\frac{2}{\alpha-1}}) \quad \text{as } x \to 0. \tag{7-1}$$

(ii) $\Phi(0+, \boldsymbol{u}) = -\bar{\lambda}$ *if and only if*

$$|\boldsymbol{u}(x)| = (1 + o(1))\lambda^{\frac{1}{\alpha-1}}|x|^{-\frac{2}{\alpha-1}} \quad \text{as } x \to 0, \tag{7-2}$$

where λ is given by (1-4).

Proof. Due to the estimates (5-29) and (5-33), we know that $\Phi(r, \boldsymbol{u})$ in (2-1) is uniformly bounded for all $0 < r < \frac{1}{2}$. This combined with the monotonicity (Proposition 2.1(i)) implies that $\Phi(0+, \boldsymbol{u})$ exists. Hence, we may argue analogously to the proof of Lemma 4.4 and observe that any blowup $\boldsymbol{u}_0$ of $\boldsymbol{u}$ satisfies $\Phi(r, \boldsymbol{u}_0) = \Phi(0+, \boldsymbol{u})$ for all $r > 0$. As $\boldsymbol{u}_0$ is a nonnegative solution of (1-1) in $\mathbb{R}^n \setminus \{0\}$, it follows from Lemma 4.5(ii) that $\Phi(0+, \boldsymbol{u}) = 0$ if and only if any blowup $\boldsymbol{u}_0$ of $\boldsymbol{u}$ is trivial, while $\Phi(0+, \boldsymbol{u}) = -\bar{\lambda}$ if and only if any blowup of $\boldsymbol{u}_0$ is of the form $\lambda^{1/(\alpha-1)}|x|^{-2/(\alpha-1)}\boldsymbol{e}$ with some nonnegative unit vector $\boldsymbol{e} \in \mathbb{R}^m$. In other words, $\Phi(0+, \boldsymbol{u}) = 0$ if and only if $|\boldsymbol{u}_r| \to 0$ uniformly on ∂B_1, while $\Phi(0+, \boldsymbol{u}) = -\bar{\lambda}$ if and only if $|\boldsymbol{u}_r| \to \lambda^{1/(\alpha-1)}$ uniformly on ∂B_1, where $\boldsymbol{u}_r$ is the scaling function defined by (2-2). □

The next lemma shows that (7-1) is sufficient for the origin to be a removable singularity.

Lemma 7.3. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1 \setminus \{0\}$ with $\frac{n}{n-2} < \alpha < \frac{n+2}{n-2}$. If $\boldsymbol{u}$ satisfies*

$$|\boldsymbol{u}(x)| = o(|x|^{-\frac{2}{\alpha-1}}) \quad \text{as } x \to 0, \tag{7-3}$$

then the origin is a removable singularity.

Proof. Under the assumption (7-3), we claim that

$$|\boldsymbol{u}(x)| \le c|x|^{-\frac{2}{\alpha-1}+\delta} \quad \text{in } B_{r_0} \setminus \{0\} \tag{7-4}$$

for some $\delta > 0$, $r_0 > 0$, and $c > 1$, where c and r_0 may depend on $\boldsymbol{u}$.

Consider the auxiliary function

$$\varphi_\epsilon(x) = (C_0 r_0^{-\delta}|x|^\delta + \epsilon)|x|^{-\frac{2}{\alpha-1}} \quad \text{in } \mathbb{R}^n \setminus \{0\}, \tag{7-5}$$

where $C_0 > 0$ is the (universal) constant from (5-29), $r_0 > 0$ is a small radius to be determined later, and $\epsilon > 0$ is an arbitrary small number. By direct computation, we observe that

$$\Delta\varphi_\epsilon = -(C_0 r_0^{-\delta}(\lambda + \mu\delta - \delta^2)|x|^\delta + \epsilon\lambda)|x|^{\frac{2\alpha}{1-\alpha}} \quad \text{in } \mathbb{R}^n \setminus \{0\},$$

with λ and μ given by (1-4). Note that for $\alpha > \frac{n}{n-2}$, we have $\lambda > 0$. Thus, taking $\delta > 0$ sufficiently small depending only on λ and $|\mu|$, we obtain

$$\Delta\varphi_\epsilon \le -\frac{\lambda}{2|x|^2}\varphi_\epsilon \quad \text{in } \mathbb{R}^n \setminus \{0\}. \tag{7-6}$$

Let us fix $1 \le i \le m$ and consider the i-th component u_i of $\boldsymbol{u}$ as a solution of $\Delta u_i = -a(x)u_i$ in $B_1 \setminus \{0\}$ with $a(x) = |\boldsymbol{u}|^{\alpha-1}$. Due to (7-3), there exists $r_0 > 0$ such that

$$0 \le a(x) \le \frac{\lambda}{2|x|^2}$$

in $B_{r_0} \setminus \{0\}$, and hence, it follows from (7-6) that φ_ϵ is a supersolution of $\Delta u_i = -a(x)u_i$ in $B_{r_0} \setminus \{0\}$. That is,

$$\Delta\varphi_\epsilon \le -a(x)\varphi_\epsilon \quad \text{in } B_{r_0} \setminus \{0\}. \tag{7-7}$$

On the other hand, choosing $C_0 > 0$ to be the constant for which $|\boldsymbol{u}|$ satisfies (5-29), we have $u_i \le C_0 r_0^{-2/(\alpha-1)} \le \varphi_\epsilon$ on ∂B_{r_0}. Utilizing the assumption (7-3) again, one can find a sufficiently small $0 < r < r_0$ such that $u_i \le \epsilon|x|^{-2/(\alpha-1)} \le \varphi_\epsilon$ in $B_r \setminus \{0\}$. Therefore,

$$u_i \le \varphi_\epsilon \quad \text{on } (\partial B_{r_0}) \cup (B_r \setminus \{0\}). \tag{7-8}$$

In view of (7-7) and (7-8), we may apply the maximum principle in $B_{r_0} \setminus B_r$ and obtain $u_i \le \varphi$ in $B_{r_0} \setminus \overline{B}_r$. Combining this inequality with (7-8), we arrive at

$$u_i \le \varphi_\epsilon \quad \text{in } B_{r_0} \setminus \{0\}. \tag{7-9}$$

Since the parameters C_0, r_0, and δ in the definition (7-5) of φ_ϵ are independent of ϵ, we can take $\epsilon \to 0$ in (7-9) and obtain

$$u_i(x) \le C_0 r_0^{-\delta}|x|^{-\frac{2}{\alpha-1}+\delta} \quad \text{in } B_{r_0} \setminus \{0\}.$$

Now that this inequality holds for any $1 \le i \le m$, we arrive at (7-4) with $c = C_0 r_0^{-\delta}\sqrt{m}$.

Since $a(x) = |\boldsymbol{u}|^{\alpha-1}$, we have from (7-4) that $0 \le a(x) \le c|x|^{-2+(\alpha-1)\delta}$ on $B_{r_0} \setminus \{0\}$, which certainly implies $a \in L^{n/(2-\eta)}(B_1)$ for some small $\eta > 0$. According to Lemma 7.1, u_i satisfies $-\Delta u_i = a(x)u_i$ in B_1 in the distributional sense for each $1 \le i \le m$, whence the classical result [Serrin 1964, Theorem 1] yields that u_i has a removable singularity at the origin. □

Remark 7.4. One may have noticed that the proof of Lemma 7.3 works for the upper critical case, $\alpha = \frac{n+2}{n-2}$, without any modification.

We are ready to prove Theorem 1.5(i).

Proof of Theorem 1.5. Suppose that $\boldsymbol{u}$ has a nonremovable singularity at the origin. Then by Lemma 7.3, $\boldsymbol{u}$ does not satisfy (7-3), whence it follows from Lemma 7.2 that $\boldsymbol{u}$ satisfies (7-2), which proves (1-17). □

7B. *Case $\alpha = \frac{n+2}{n-2}$.* The asymptotic behavior for the case $\alpha = \frac{n+2}{n-2}$ becomes more subtle, due to the presence of the second Pohozaev invariant κ_* given by (2-17). The following lemma is the local version of Theorem 1.3(iii). Let us remark that the proof is similar to the classical argument, see the proof of [Caffarelli et al. 1989, Theorem 1.2]; however, the key difference is that we apply the radial symmetry to the second Pohozaev identity (2-17), instead of the first identity (2-8).

Lemma 7.5. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1 \setminus \{0\}$ with $\alpha = \frac{n+2}{n-2}$. Also set $\kappa(\boldsymbol{u})$ and $\kappa_*(\boldsymbol{u})$ as in* (2-8) *and* (2-17) *respectively. Then $\kappa(\boldsymbol{u})$ and $\kappa_*(\boldsymbol{u})$ satisfy* (1-13) *and* (1-14) *respectively. Moreover, the following are true*:

(i) $\kappa(\boldsymbol{u}) = \kappa_*(\boldsymbol{u}) = 0$ *if and only if*

$$|\boldsymbol{u}(x)| = o(|x|^{-\frac{n-2}{2}}) \quad \text{as } x \to 0. \tag{7-10}$$

(ii) $\kappa(\boldsymbol{u})^2 + \kappa_*(\boldsymbol{u})^2 > 0$ *if and only if there are $c, C > 0$ such that*

$$c|x|^{-\frac{n-2}{2}} \le |\boldsymbol{u}(x)| \le C|x|^{-\frac{n-2}{2}} \quad \text{as } x \to 0, \tag{7-11}$$

where c depends on $\boldsymbol{u}$, while C is determined by n and m only.

(iii) $\kappa(\boldsymbol{u}) = -\frac{2}{n}\left(\frac{n-2}{2}\right)^n$ *and $\kappa_*(\boldsymbol{u}) = 0$ if and only if*

$$|\boldsymbol{u}(x)| = (1 + o(1))\left(\frac{n-2}{2}\right)^{\frac{n-2}{2}} |x|^{-\frac{n-2}{2}} \quad \text{as } x \to 0. \tag{7-12}$$

Proof. The existence of $\kappa(\boldsymbol{u})$ and $\kappa_*(\boldsymbol{u})$ are proved in Proposition 2.1(ii) and Proposition 2.4 respectively. Now let $\boldsymbol{u}_0$ be any blowup of $\boldsymbol{u}$, and write $r_j \to 0+$ by the blowup sequence. By the scaling relation (2-3) of $\Phi(r, \boldsymbol{u})$, we see that

$$\kappa(\boldsymbol{u}_0) = \Phi(1, \boldsymbol{u}_0) = \lim_{j\to\infty} \Phi(1, \boldsymbol{u}_{r_j}) = \lim_{j\to\infty} \Phi(r_j, \boldsymbol{u}) = \kappa(\boldsymbol{u}).$$

However, $\boldsymbol{u}_0$ is a nonnegative solution of (1-1) $\left(\text{with } \alpha = \frac{n+2}{n-2}\right)$ in $\mathbb{R}^n \setminus \{0\}$, whence Lemma 4.6 yields $\kappa(\boldsymbol{u}_0)$ satisfies (1-13), and so does $\kappa(\boldsymbol{u})$. Similarly, one may deduce from the scaling relation (2-13) of $\Phi_*(r, \boldsymbol{u})$ that $\kappa_*(\boldsymbol{u}) = \kappa_*(\boldsymbol{u}_0)$, and by Lemma 4.6, $\kappa_*(\boldsymbol{u})$ satisfies (1-14).

Suppose that $\kappa(\boldsymbol{u}) = \kappa_*(\boldsymbol{u}) = 0$, and let $\boldsymbol{v}$ be the cylindrical transformation of $\boldsymbol{u}$ as in (1-2). Rephrasing the estimates (6-7) and (6-8) in terms of $\boldsymbol{v}$, the second Pohozaev identity (2-18) becomes (as $t \to \infty$)

$$(g')^2 = (n-2)^2 g^2 - \frac{4(n-2)}{n} g^{\frac{2n-2}{n-2}} + O\left(\int_t^\infty e^{-2\tau} g(\tau)^2 \, d\tau\right), \tag{7-13}$$

where g is given by (2-16) and $g' = dg/dt$. Since the term $O\left(\int_t^\infty e^{-2\tau} g(\tau)^2 \, d\tau\right)$ decays exponentially, and is comparably smaller than $g(t)$, the behavior of g' is determined by the nonnegative roots of

$$(n-2)^2 g^2 - \frac{4(n-2)}{n} g^{\frac{2n-2}{n-2}} = 0,$$

which are 0 and $\left(\frac{n(n-2)}{4}\right)^{(n-2)/2}$ respectively. In particular, $g(t)$ must be either nonincreasing and converging to 0, or nondecreasing and converging to $\left(\frac{n(n-2)}{4}\right)^{(n-2)/2}$.

If $g(t) \to 0$ as $t \to \infty$, then by the asymptotic radial symmetry we have $|\boldsymbol{v}(t,\cdot)| \to 0$ uniformly on $\mathbb{S}^{n-1}$ as $t \to \infty$. After the inverse cylindrical transform via (1-2), we arrive at (7-10), as desired.

Now let us show that the other alternative, i.e., $g(t) \to \left(\frac{n(n-2)}{4}\right)^{(n-2)/2}$ as $t \to \infty$, cannot occur. Suppose that this is true. Then again from the asymptotic radial symmetry it follows that $|\boldsymbol{u}_r| \to \left(\frac{n(n-2)}{4}\right)^{(n-2)/2}$ uniformly on ∂B_1 as $r \to 0+$. This implies that any blowup $\boldsymbol{u}_0$ of $\boldsymbol{u}$ must be of the form $\left(\frac{n(n-2)}{4}\right)^{(n-2)/2}|x|^{-(n-2)/2}\boldsymbol{e}$ for some nonnegative unit vector $\boldsymbol{e} \in \mathbb{R}^m$. In particular, $\boldsymbol{u}_0$ has a nonremovable singularity at the origin, and hence Theorem 1.3(iii) yields that $\kappa(\boldsymbol{u}_0)$ or $\kappa_*(\boldsymbol{u}_0)$ is nonzero, a contradiction to $\kappa(\boldsymbol{u}) = \kappa(\boldsymbol{u}_0) = 0$ or, respectively, $\kappa_*(\boldsymbol{u}) = \kappa_*(\boldsymbol{u}_0) = 0$. Hence, the assertion (i) is proved.

Now let us consider the case when $\kappa(\boldsymbol{u})^2 + \kappa_*(\boldsymbol{u})^2 > 0$. Let $\boldsymbol{u}_0$ be any blowup of $\boldsymbol{u}$. Then due to the asymptotic radial symmetry of $\boldsymbol{u}$, we know $\boldsymbol{u}_0$ is radially symmetric on the punctured space. Hence, by Lemma 4.2, we have $|\boldsymbol{u}_0| \le C|x|^{-(n-2)/2}$, where $C > 0$ depends only on n and m. Since $\boldsymbol{u}_0$ is an arbitrary blowup of $\boldsymbol{u}$, this proves the upper bound in (7-11).

On the other hand, by Theorem 1.3(iii)-(b), the cylindrical transform $\boldsymbol{v}_0$ of $\boldsymbol{u}_0$ satisfies (1-7). Due to R. H. Fowler [1931], $|\boldsymbol{v}_0|$ has to be bounded uniformly away from zero, with the bound determined solely on the value of n, $\kappa(v_0) = \kappa(u_0) = \kappa(u)$ and $\kappa_*(\boldsymbol{v}_0) = \kappa_*(\boldsymbol{u})$. This proves that $|\boldsymbol{u}_0| \ge c|x|^{-(n-2)/2}$ for some $c > 0$ depending only on n, $\kappa(\boldsymbol{u})$ and $\kappa_*(\boldsymbol{u})$. Since c is independent of the blowup $\boldsymbol{u}_0$, the lower bound in (7-11) is proved. Thus, the assertion (ii) is proved.

The final assertion regarding (7-12) follows immediately from Theorem 1.3(iii)-(c), since the latter implies that the blowup of $\boldsymbol{u}$ is unique and is of the form (1-15), if and only if $\kappa(\boldsymbol{u}) = -\frac{2}{n}\left(\frac{n-2}{2}\right)^n$ and $\kappa_*(\boldsymbol{u}) = 0$. ☐

As with Lemma 7.3, we observe that (7-10) is a sufficient condition to have a removable singularity.

Lemma 7.6. *Let* $\boldsymbol{u}$ *be a nonnegative solution of* (1-1) *in* $B_1 \setminus \{0\}$ *with* $\alpha = \frac{n+2}{n-2}$. *If* $\boldsymbol{u}$ *satisfies*

$$|\boldsymbol{u}(x)| = o(|x|^{-\frac{n-2}{2}}) \quad \text{as } x \to 0,$$

then the origin is a removable singularity.

Proof. As mentioned in Remark 7.4, the same proof of Lemma 7.3 works here as well, whence we leave out the details to the reader. ☐

Proof of Theorem 1.5(ii). Suppose that the origin is a nonremovable singularity, and let us write by κ and κ_* the first and respectively the second Pohozaev invariant. As a contraposition to Lemma 7.6, (7-3) fails. Thus, by Lemma 7.5, one has $\kappa^2 + \kappa_*^2 > 0$. Then the asymptotic bound in (1-18) follows from the second alternative, (7-11), of Lemma 7.5, and the proof is finished. ☐

7C. *Case* $1 < \alpha < \frac{n}{n-2}$. The asymptotic analysis for the case $1 < \alpha < \frac{n}{n-2}$ is very simple. It is noticeable that the monotonicity formula is not required here. We also mention that one can reduce our study to the scalar case by considering $w = u_1 + u_2 + \cdots + u_m \ge 0$, and directly apply the results in [Lions 1980]. Nevertheless, we shall give a more direct proof, for the sake of completeness.

We shall begin with the sharp upper estimate.

Lemma 7.7. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1 \setminus \{0\}$ with $1 < \alpha < \frac{n}{n-2}$. Then there is $C > 0$, depending only $|\boldsymbol{u}|$, such that*

$$|\boldsymbol{u}(x)| \le C|x|^{2-n} \quad \text{as } x \to 0. \tag{7-14}$$

Proof. Lemma 7.1 asserts that $\boldsymbol{u} \in L^\alpha(B_1)$. Since $1 < \alpha < \frac{n}{n-2}$ and $\boldsymbol{u}$ satisfies the Harnack inequality (5-32), it is easy to verify that

$$|\boldsymbol{u}(x)| = o(|x|^{-\frac{2}{\alpha-1}}) \quad \text{as } x \to 0. \tag{7-15}$$

Utilizing (7-15), and noting that $n-2 < \frac{2}{\alpha-1}$, one may argue with a blowup argument to prove that for any $n-2 < q < \frac{2}{\alpha-1}$, there is some $0 < r_q < 1$, depending only on n, m, α, and q, such that

$$|\boldsymbol{u}(x)| < |x|^{-q} \quad \text{in } B_{r_q} \setminus \{0\}. \tag{7-16}$$

Now let r_q be as in (7-16). Due to Lemma 7.1 again, $\Delta\boldsymbol{u} = -|\boldsymbol{u}|^{\alpha-1}\boldsymbol{u} \in L^1(B_1)$, whence one can decompose $\boldsymbol{u}$, in $B_{r_q} \setminus \{0\}$, as

$$\boldsymbol{u}(x) = |x|^{2-n}\boldsymbol{a} - \int_{B_{r_q}} |x-y|^{2-n}\Delta\boldsymbol{u}(y)\,dy + \boldsymbol{h}(x), \tag{7-17}$$

where $\boldsymbol{a}$ is a nonnegative vector in $\mathbb{R}^m$ and $\boldsymbol{h}$ is a nonnegative and harmonic, vectorial function on B_{r_q}. However, owing to the estimate (7-16), it is not hard to see from the equation $\Delta\boldsymbol{u} = -|\boldsymbol{u}|^{\alpha-1}\boldsymbol{u}$ that there is $C_q > 0$, depending only on n, m, α, and q, such that

$$\left|\int_{B_{r_q}} |x-y|^{2-n}\Delta\boldsymbol{u}(y)\,dy\right| \le \int_{B_{r_q}} |x-y|^{2-n}|y|^{-\alpha q}\,dy \le C_q|x|^{2-n}. \tag{7-18}$$

Thus, choosing $n-2 < q < \frac{2}{\alpha-1}$ so as to depend only on n and α, and selecting r_q and C_q in (7-18) correspondingly, we derive the sharp estimate (7-14) from (7-17). □

Next we consider a sufficient condition to have a removable singularity.

Lemma 7.8. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1 \setminus \{0\}$ with $1 < \alpha < \frac{n}{n-2}$. If $\boldsymbol{u}$ satisfies*

$$|\boldsymbol{u}(x)| = o(|x|^{2-n}) \quad \text{as } x \to 0, \tag{7-19}$$

then the origin is a removable singularity.

Proof. Under the assumption (7-19), one has $\boldsymbol{u} \in L^q(B_1;\mathbb{R}^m)$ for any $1 \le q < \frac{n}{n-2}$. Since $1 < \alpha < \frac{n}{n-2}$ and $|\Delta\boldsymbol{u}| \le |\boldsymbol{u}|^\alpha$, we have $-\Delta\boldsymbol{u} \in L^{q/\alpha}(B_1;\mathbb{R}^m)$ for any $\alpha < q < \frac{n}{n-2}$. Thus, the L^p theory [Gilbarg and Trudinger 1983, Theorem 9.9] (applied to each component of $\boldsymbol{u}$) and a bootstrap argument based on the Sobolev inequality yields $\boldsymbol{u} \in W^{2,p}(B_1;\mathbb{R}^m)$ for any $1 < p < \infty$. In particular, it follows from the Sobolev embedding that $\boldsymbol{u} \in C^{1,\gamma}(B_1;\mathbb{R}^m)$ for any $0 < \gamma < 1$, and thus $\boldsymbol{u}$ must have a removable singularity at the origin. □

We are in a position to prove Theorem 1.5(iii).

Proof of Theorem 1.5(*iii*). Suppose that $\boldsymbol{u}$ has a nonremovable singularity at the origin. By Lemma 7.8, we know that $\boldsymbol{u}$ does not satisfy (7-19), or equivalently, there is some $\delta > 0$, a component, say u_1, and a sequence $r_j \to 0+$ such that

$$\sup_{\partial B_{r_j}} u_1 \geq \delta r_j^{2-n}.$$

By the Harnack inequality (5-32), we know that

$$\inf_{\partial B_{r_j}} u_1 \geq c_0 \delta r_j^{2-n},$$

where $c_0 > 0$ depends only on n, m and α. Taking $\delta > 0$ smaller, if necessary, such that $c\delta \leq \inf_{\partial B_{1/2}} u_1$, it follows from the maximum principle that

$$u_1(x) \geq c_0 \delta |x|^{2-n} \quad \text{in } B_{1/2} \setminus \{0\},$$

proving the asymptotic lower bound in (1-19). The asymptotic upper bound in (1-19) is established in Lemma 7.7. Hence, the theorem is proved. □

Remark 7.9. As mentioned in the beginning of this section, the proof of Theorem 1.5(iii) can also be deduced by considering the function $w = u_1 + u_2 + \cdots + u_m \geq 0$. Then w satisfies $C_1 w^\alpha \leq -\Delta w \leq C_2 w^\alpha$ in $B_1 \setminus \{0\}$, where $C_1, C_2 > 0$ depend on n, m and α only, and the claim in Theorem 1.5(iii) follows now from existing results in the literature, such as [Lions 1980, Theorem 2 and Remark 2].

7D. *Case $\alpha = \frac{n}{n-2}$*. The analysis of the lower critical exponent, $\alpha = \frac{n}{n-2}$, exhibits its own subtlety, due to the multiplicity of components in (1-1), as with the upper critical case, $\alpha = \frac{n+2}{n-2}$. To briefly discuss this point, let us first give the asymptotic upper bound.

Lemma 7.10 [Aviles 1987, Lemma 1]. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *with $\alpha = \frac{n}{n-2}$ in $B_1 \setminus \{0\}$. Then for each $1 \leq i \leq m$,*

$$\bar{u}_i(r) \leq \left(\frac{(n-2)^2}{2}\right)^{\frac{n-2}{2}} r^{2-n} (-\log r)^{\frac{2-n}{2}} \quad \textit{as } r \to 0, \tag{7-20}$$

where $\bar{u}_i$ is the average of u_i over the sphere ∂B_r.

Proof. Note that for each $1 \leq i \leq m$, $\bar{u}_i$ satisfies, for $0 < r < 1$,

$$\ddot{\bar{u}}_i + \frac{n-1}{r} \dot{\bar{u}}_i + \bar{u}_i^{\frac{n}{n-2}} = 0,$$

whence the conclusion follows directly from [Aviles 1987, Lemma 1]. □

Let us remark that the constant $\left(\frac{1}{2}(n-2)^2\right)^{(n-2)/2}$ in (7-20) is exact in view of (1-20). Due to the fact that $\boldsymbol{u}$ consists of multiple components, there is not an easy way to prove that $|\bar{\boldsymbol{u}}|$ also satisfies (7-20) with exactly the same constant. This prevents us from applying the argument in [Aviles 1987, Section 2], which deals with the scalar version of (1-1) with $\alpha = \frac{n}{n-2}$. Instead, we mainly follow [Aviles 1987, Section 3], where a sign-changing problem is considered. The idea is to consider several refinements of the usual monotonicity formula $\Psi(t, \boldsymbol{v})$ introduced in (2-5).

Due to the refined upper bound (7-20), we shall consider a new cylindrical transformation $\boldsymbol{\phi}$ defined so as to satisfy

$$\boldsymbol{u}(x) = |x|^{2-n}(-\log|x|)^{\frac{2-n}{2}}\boldsymbol{\phi}\Big(-\log|x|, \frac{x}{|x|}\Big). \tag{7-21}$$

Then the problem (1-1) $\big($with $\alpha = \frac{n}{n-2}\big)$ can be reformulated in terms of $\boldsymbol{\phi}$ as

$$\partial_{tt}\boldsymbol{\phi} + (n-2)\Big(1-\frac{1}{t}\Big)\partial_t\boldsymbol{\phi} + \Delta_\theta\boldsymbol{\phi} = \frac{n-2}{2t}\Big(n-2-\frac{n}{2t}\Big)\boldsymbol{\phi} - \frac{1}{t}|\boldsymbol{\phi}|^{\frac{2}{n-2}}\boldsymbol{\phi}. \tag{7-22}$$

Remark 7.11. Due to the asymptotic radial symmetry (1-16) of $\boldsymbol{u}$, we know $\boldsymbol{\phi}$ satisfies $|\boldsymbol{\phi}-\bar{\boldsymbol{\phi}}| = O(e^{-\gamma t})$ as $t\to\infty$, for some $\gamma > 0$, where $\bar{\boldsymbol{\phi}}(t)$ is the average of $\boldsymbol{\phi}(t,\theta)$ over $\theta\in\mathbb{S}^{n-1}$. In particular, one has (by arguing much as in the derivation of (6-6))

$$|\nabla_\theta\boldsymbol{\phi}(t,\theta)| \le Ce^{-\gamma t} \quad \text{in } (t_0,\infty)\times\mathbb{S}^{n-1} \tag{7-23}$$

for some large t_0 and C independent of t. Moreover, it follows from the sharp estimate (7-20) and the gradient estimate (5-33) that

$$|\boldsymbol{\phi}(t,\theta)| + |\partial_t\boldsymbol{\phi}(t,\theta)| \le C \quad \text{in } (t_0,\infty)\times\mathbb{S}^{n-1}. \tag{7-24}$$

In comparison with (1-3), we obtain the first refinement of the monotonicity formula $\Psi(t,\boldsymbol{v})$, given as

$$E(t,\boldsymbol{\phi}) = \frac{1}{n\omega_n}\int_{\mathbb{S}^{n-1}}\Big(t|\partial_t\boldsymbol{\phi}|^2 - t|\nabla_\theta\boldsymbol{\phi}|^2 + \frac{n-2}{n-1}|\boldsymbol{\phi}|^{\frac{2n-2}{n-2}}\Big)\,d\theta - \frac{n-2}{2n\omega_n}\Big(n-2-\frac{n}{2t}\Big)\int_{\mathbb{S}^{n-1}}|\boldsymbol{\phi}|^2\,d\theta. \tag{7-25}$$

Note that $E(t,\boldsymbol{\phi})$ is well-defined for any t whenever $\boldsymbol{\phi}(t,\cdot)$ is defined on $\mathbb{S}^{n-1}$, due to the smoothness of $\boldsymbol{u}$.

The next lemma is concerned with the monotonicity of $E(t,\boldsymbol{\phi})$.

Lemma 7.12. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1\setminus\{0\}$ with $\alpha = \frac{n}{n-2}$, and $\boldsymbol{\phi}$ be the cylindrical transformation as in* (7-21). *Then*

$$E'(t,\boldsymbol{\phi}) = -\frac{(2n-4)t-2n+3}{n\omega_n}\int_{\mathbb{S}^{n-1}}|\partial_t\boldsymbol{\phi}|^2\,d\theta - \frac{1}{n\omega_n}\int_{\mathbb{S}^{n-1}}\Big(|\nabla_\theta\boldsymbol{\phi}|^2\,d\theta + \frac{n(n-2)}{4t^2}|\boldsymbol{\phi}|^2\Big)\,d\theta. \tag{7-26}$$

In particular, $E(t,\boldsymbol{\phi})$ is nonincreasing for $t > \frac{2n-3}{2n-4}$, and $E(\infty,\boldsymbol{\phi})$ exists.

Proof. The proof of (7-26) follows easily from taking the inner product of (7-22) with $t\partial_t\boldsymbol{\phi}$ and integrating the both sides over $\mathbb{S}^{n-1}$. We omit the details.

With (7-26) at hand, we know that $E(t,\boldsymbol{\phi})$ is nonincreasing for $t > \frac{2n-3}{2n-4}$. Thus, the existence of $E(\infty,\boldsymbol{\phi})$ follows immediately from the fact that $E(t,\boldsymbol{\phi})$ is uniformly bounded from below as $t\to\infty$. However, (7-23) yields

$$\lim_{t\to\infty}\int_{\mathbb{S}^{n-1}} t|\nabla_\theta\boldsymbol{\phi}|^2\,d\theta = 0,$$

which along with (7-24) ensures that

$$\liminf_{t\to\infty} E(t,\boldsymbol{\phi}) > -\infty,$$

as desired. □

In order to have the full strength of the existence of $E(\infty, \boldsymbol{\phi})$, we shall prove the following, which is the system version of [Aviles 1987, Lemma 3.2]. Although the proof is almost identical, we shall present the argument for the sake of completeness.

Lemma 7.13 (essentially due to [Aviles 1987]). *Let $\boldsymbol{\phi}$ be as in Lemma 7.12. Then*

$$\lim_{t\to\infty} \int_{\mathbb{S}^{n-1}} t|\partial_t \boldsymbol{\phi}|^2 \, d\theta = 0. \tag{7-27}$$

Proof. By (7-23) and (7-24), one may integrate the both sides of (7-26) from $t_0 = \frac{2n-3}{2n-4}$ to ∞, and use the existence of $E(\infty, \phi)$ to deduce that

$$\int_{t_0}^{\infty}\int_{\mathbb{S}^{n-1}} \tau|\partial_\tau \boldsymbol{\phi}|^2 \, d\theta \, d\tau < \infty. \tag{7-28}$$

Hence, it is sufficient to prove that $\int_{\mathbb{S}^{n-1}} t|\partial_t \boldsymbol{\phi}|^2 \, d\theta$ is a Cauchy sequence in $t \to \infty$.

In order to do so, we differentiate (7-22) in t and find that $\boldsymbol{\psi} = \partial_t \boldsymbol{\phi}$ solves

$$\begin{aligned} \partial_{tt}\boldsymbol{\psi} + (n-2)\left(1 - \frac{1}{t}\right)\partial_t \boldsymbol{\psi} - \frac{n-2}{2t}\left(n-2-\frac{n+4}{2t}\right)\boldsymbol{\psi} + \Delta_\theta \boldsymbol{\psi} \\ = -\frac{n-2}{2t^2}\left(n-2-\frac{n}{t}\right)\boldsymbol{\phi} + \frac{1}{t}|\boldsymbol{\phi}|^{\frac{2}{n-2}}\left(\frac{1}{t}\boldsymbol{\phi} - \frac{2}{n-2}\frac{\boldsymbol{\phi}\cdot\boldsymbol{\psi}}{|\boldsymbol{\phi}|^2}\boldsymbol{\phi} - \boldsymbol{\psi}\right). \end{aligned} \tag{7-29}$$

Taking the inner product of (7-29) with $t\partial_t\boldsymbol{\psi}$ and integrating over $\mathbb{S}^{n-1}$, one may verify after some computation that the functional

$$\begin{aligned} J(t,\boldsymbol{\psi}) = \frac{1}{n\omega_n}\int_{\mathbb{S}^{n-1}} \left(t|\partial_t\boldsymbol{\psi}|^2 - t|\nabla_\theta\boldsymbol{\psi}|^2 - \frac{n-2}{2}\left(n-2-\frac{n+4}{2t}\right)|\boldsymbol{\psi}|^2\right) d\theta \\ - \frac{1}{n\omega_n}\int_t^\infty\int_{\mathbb{S}^{n-1}} \frac{n-2}{\tau}\left(n-2-\frac{n}{\tau}\right)\boldsymbol{\phi}\cdot\partial_\tau\boldsymbol{\psi}\, d\theta\, d\tau \\ + \frac{1}{n\omega_n}\int_t^\infty\int_{\mathbb{S}^{n-1}} |\boldsymbol{\phi}|^{\frac{2}{n-2}}\left(\frac{1}{\tau}\boldsymbol{\phi} - \frac{2}{n-2}\frac{\boldsymbol{\phi}\cdot\boldsymbol{\psi}}{|\boldsymbol{\phi}|^2}\boldsymbol{\phi} - \boldsymbol{\psi}\right)\cdot\partial_\tau\boldsymbol{\psi}\, d\theta\, d\tau \end{aligned} \tag{7-30}$$

satisfies

$$\begin{aligned} J'(t,\boldsymbol{\psi}) = -\frac{(2n-4)t-2n+3}{n\omega_n}\int_{\mathbb{S}^{n-1}} |\partial_t\boldsymbol{\psi}|^2\, d\theta \\ - \frac{1}{n\omega_n}\int_{\mathbb{S}^{n-1}}\left(|\nabla_\theta\boldsymbol{\psi}|^2 + \frac{(n+4)(n-2)}{t^2}\int_{\mathbb{S}^{n-1}}|\boldsymbol{\psi}|^2\right) d\theta, \end{aligned} \tag{7-31}$$

provided that the last two double integrals in (7-30) are finite, i.e., $J(t,\boldsymbol{\psi})$ is well-defined for all t large.

Assuming for the moment that $J(t,\psi)$ is well-defined for all t large, one may proceed as in the proof of [Aviles 1987, Lemma 3.2]. Note that (7-31) implies the monotonicity of $J(t,\boldsymbol{\psi})$ for $t \geq t_0 = \frac{2n-3}{2n-4}$. Analogous to Remark 7.11, the asymptotic radial symmetry (1-16) implies the exponential decay of $|\nabla_\theta\boldsymbol{\psi}|$ as well as the uniform boundedness of $|\boldsymbol{\psi}|$ and $|\partial_t\boldsymbol{\psi}|$. Hence, one may deduce as in the proof of Lemma 7.12 that $J(t,\boldsymbol{\psi})$ is uniformly bounded from below as $t \to \infty$. As $J(t,\boldsymbol{\psi})$ is nonincreasing in $t \geq t_0$, $J(\infty, \boldsymbol{\psi})$ exists, and thus, integrating (7-30) from t_0 to ∞ yields that

$$\int_{t_0}^{\infty}\int_{\mathbb{S}^{n-1}} \tau|\partial_\tau \boldsymbol{\psi}|^2 \, d\theta \, d\tau < \infty. \tag{7-32}$$

Noting that

$$\left|\frac{d}{dt}\left(t \int_{\mathbb{S}^{n-1}} |\partial_t \boldsymbol{\phi}|^2 \, d\theta\right)\right| \leq \int_{\mathbb{S}^{n-1}} (|\partial_t \boldsymbol{\phi}|^2 + t|\partial_t \boldsymbol{\phi}|^2 + t|\partial_{tt} \boldsymbol{\phi}|^2) \, d\theta,$$

we conclude from (7-28) and (7-32) that $t \int_{\mathbb{S}^{n-1}} |\partial_t \boldsymbol{\phi}|^2 \, d\theta$ is a Cauchy sequence in $t \to \infty$. Thus, (7-27) follows from (7-28).

To this end, we are only left with verifying the well-definedness of $J(t, \boldsymbol{\psi})$ for all $t \geq t_0$ with some t_0 large. As noted above, this boils down to proving that the last two double integrals in (7-30) are finite. Due to the upper estimate (7-20) and (7-28), it suffices to show that

$$\int_{t_0}^{\infty} \frac{1}{t} \int_{\mathbb{S}^{n-1}} (|\boldsymbol{\phi}| + |\boldsymbol{\psi}|)|\partial_t \boldsymbol{\psi}| \, d\theta \, dt < \infty. \tag{7-33}$$

Owing to (7-23) and (7-24), we have, in (7-22) (recall that $\boldsymbol{\psi} = \partial_t \boldsymbol{\phi}$),

$$|\partial_t \boldsymbol{\psi}| = (n-2)|\boldsymbol{\psi}| + O\left(\frac{1}{t}\right), \tag{7-34}$$

so multiplying (7-34) by $\frac{1}{t}|\boldsymbol{\phi}|$ yields

$$\begin{aligned} \int_{t_0}^{\infty} \frac{1}{t} \int_{\mathbb{S}^{n-1}} |\boldsymbol{\phi}||\partial_t \boldsymbol{\psi}| \, d\theta \, dt &\leq (n-2) \int_{t_0}^{\infty} \frac{1}{t} \int_{\mathbb{S}^{n-1}} |\boldsymbol{\phi}||\boldsymbol{\psi}| \, d\theta \, dt + O(1) \\ &\leq \frac{n-2}{2} \int_{t_0}^{\infty}\int_{\mathbb{S}^{n-1}} |\boldsymbol{\psi}|^2 \, d\theta \, dt + O(1) < \infty, \end{aligned} \tag{7-35}$$

where the second inequality follows from

$$|\boldsymbol{\phi}||\boldsymbol{\psi}| \leq \frac{1}{2t}|\boldsymbol{\phi}|^2 + \frac{t}{2}|\boldsymbol{\psi}|^2,$$

while the last inequality is derived from (7-28). On the other hand, multiplying (7-34) by $\frac{1}{t}|\boldsymbol{\psi}|$, we deduce from (7-28) that

$$\int_{t_0}^{\infty} \frac{1}{t} \int_{\mathbb{S}^{n-1}} |\boldsymbol{\psi}||\partial_t \boldsymbol{\psi}| \, d\theta \, dt \leq (n-2) \int_{t_0}^{\infty} \frac{1}{t} \int_{\mathbb{S}^{n-1}} |\boldsymbol{\psi}|^2 \, d\theta \, dt < \infty. \tag{7-36}$$

The claim (7-33) follows readily from (7-35) and (7-36). □

Finally we have the classification of the blowup limit via the limiting energy levels $E(\infty, \boldsymbol{\phi})$.

Lemma 7.14. *Let $\boldsymbol{u}$ be a nonnegative solution of* (1-1) *in $B_1 \setminus \{0\}$ with $\alpha = \frac{n}{n-2}$, and $\boldsymbol{\phi}$ be its cylindrical transform as in* (7-21). *Also let $E(t, \boldsymbol{\phi})$ be as in* (7-25). *Then*

$$E(\infty, \boldsymbol{\phi}) \in \left\{-\frac{1}{n-1}\left(\frac{(n-2)^2}{2}\right)^{n-1}, 0\right\}.$$

Moreover, the following are true:

(i) $E(\infty, \boldsymbol{\phi}) = 0$ *if and only if*

$$|\boldsymbol{u}(x)| = o(|x|^{2-n}(-\log|x|)^{\frac{2-n}{2}}) \quad \text{as } x \to 0. \tag{7-37}$$

(ii) $E(\infty, \boldsymbol{\phi}) = -\frac{1}{n-1}\big(\frac{(n-2)^2}{2}\big)^{n-1}$ *if and only if*

$$|\boldsymbol{u}(x)| = (1+o(1))\left(\frac{(n-2)^2}{2}\right)^{\frac{n-2}{2}} |x|^{2-n}(-\log|x|)^{\frac{2-n}{2}}. \tag{7-38}$$

Proof. Due to Lemma 7.12, (7-23), and (7-27), we have

$$E(\infty, \boldsymbol{\phi}) = \frac{1}{n\omega_n} \lim_{t\to\infty} \int_{\mathbb{S}^{n-1}} \left(\frac{n-2}{n-1}|\boldsymbol{\phi}|^{\frac{2n-2}{n-2}} - \frac{(n-2)^2}{2}|\boldsymbol{\phi}|^2\right) d\theta. \tag{7-39}$$

In fact, (7-23) implies that whenever $\boldsymbol{\phi}(t_j, \theta)$ converges as $t_j \to \infty$, the limit is independent of $\theta \in \mathbb{S}^{n-1}$. Hence, along a convergent sequence $\boldsymbol{\phi}(t_j, \theta) \to \boldsymbol{a}$ (uniformly over $\theta \in \mathbb{S}^{n-1}$), we obtain from (7-39) that

$$E(\infty, \boldsymbol{\phi}) = \frac{n-2}{n-1}|\boldsymbol{a}|^{\frac{2n-2}{n-2}} - \frac{(n-2)^2}{2}|\boldsymbol{a}|^2. \tag{7-40}$$

Since the right-hand side has at most three nonnegative roots, we conclude that the limit value $|\boldsymbol{a}|$ (under the uniform convergence of $|\boldsymbol{\phi}(t, \theta)|$ on $\mathbb{S}^{n-1}$ as $t \to \infty$) is unique.

To compute the limit value $|\boldsymbol{a}|$, let us take the inner product of (7-22) with $\boldsymbol{\phi}$ and integrate the both sides over $(t_0, \infty) \times \mathbb{S}^{n-1}$ (with t_0 large). Then one may easily deduce from (7-23), (7-24), and (7-28) that

$$\left| \int_{t_0}^{\infty} \frac{1}{n\omega_n \tau} \int_{\mathbb{S}^{n-1}} \left(\frac{(n-2)^2}{2} - |\boldsymbol{\phi}|^{\frac{2}{n-2}}\right)|\boldsymbol{\phi}|^2 \, d\theta \, dt \right| < \infty.$$

Now that $|\boldsymbol{\phi}|$ converges to $|\boldsymbol{a}|$ as $t \to \infty$ uniformly on $\mathbb{S}^{n-1}$, we must have either $|\boldsymbol{a}| = 0$ or

$$|\boldsymbol{a}| = \left(\frac{(n-2)^2}{2}\right)^{\frac{n-2}{2}}.$$

Inserting this into (7-40), we deduce that either $E(\infty, \boldsymbol{\phi}) = 0$ if and only if $|\boldsymbol{a}| = 0$, or

$$E(\infty, \boldsymbol{\phi}) = -\frac{1}{n-1}\left(\frac{(n-2)^2}{2}\right)^{n-1}.$$

Obviously, the assertions (7-37) and (7-38) follow immediately via inverse cylindrical transform (7-21). □

We are only left with proving that (7-37) yields the removability of the singularity at the origin.

Lemma 7.15. *Let* $\boldsymbol{u}$ *be a nonnegative solution of* (1-1) *in* $B_1 \setminus \{0\}$ *with* $\alpha = \frac{n}{n-2}$. *Suppose further that* $\boldsymbol{u}$ *satisfies*

$$|\boldsymbol{u}(x)| = o(|x|^{n-2}(-\log|x|)^{\frac{n-2}{2}}) \quad \text{as } x \to 0. \tag{7-41}$$

Then the origin is a removable singularity.

Proof. Under the assumption (7-41), we claim that

$$|\boldsymbol{u}(x)| \le c|x|^{2-n+\delta} \quad \text{in } B_{r_0} \setminus \{0\} \tag{7-42}$$

for some small $\delta > 0$, where $c > 1$ and $r_0 > 0$ may depend on $\boldsymbol{u}$.

Consider the auxiliary function

$$\varphi_\epsilon(x) = (C r_0^{-\delta}|x|^\delta + \epsilon(-\log|x|)^{\frac{2-n}{2}})|x|^{2-n} \quad \text{in } B_{r_0} \setminus \{0\},$$

where $C_0 > 0$ is the (universal) constant chosen from (7-15), $r_0 > 0$ is a small radius to be determined later, and $\epsilon > 0$ is an arbitrary small number. After some computations, one may verify that

$$\Delta\varphi_\epsilon \le \frac{C_1}{|x|^2 \log|x|}\varphi_\epsilon \quad \text{in } B_{r_0} \setminus \{0\},$$

by choosing $\delta, r_0 > 0$ small, $C_1 > 0$ large. Here one may choose δ and C_1 to depend only on n.

Due to the assumption (7-41), we have $a(x) = |\boldsymbol{u}|^{2/(n-2)} = o(-|x|^2 \log|x|)$, whence φ_ϵ becomes a supersolution of $\Delta u_i = -a(x)u_i$ in $B_{r_0} \setminus \{0\}$, by choosing $r_0 > 0$ sufficiently small, where u_i is the i-th component of $\boldsymbol{u}$. The rest of the proof follows the same argument shown in the proof of Lemma 7.3, which eventually leads us to $u_i \le \varphi_\epsilon$ in $B_{r_0} \setminus \{0\}$. Passing to the limit with $\epsilon \to 0$, we get

$$u_i(x) \le C_0 r_0^{-\delta}|x|^{2-n+\delta} \quad \text{in } B_{r_0} \setminus \{0\}.$$

Now that this inequality holds for any $1 \le i \le m$, we arrive at (7-42) with $c = C_0 r_0^{-\delta}\sqrt{m}$.

Thus, it follows from (7-4) that $a(x) = |\boldsymbol{u}|^{2/(n-2)} \in L^{n/(2-\eta)}(B_1)$ for some $\eta > 0$. We know from Lemma 7.1 that u_i is a distribution solution of $-\Delta u_i = a(x)u_i$ in B_1 for each $1 \le i \le m$. Hence, the classical result [Serrin 1964, Theorem 1] implies that u_i has a removable singularity at the origin, and the lemma is proved. □

Theorem 1.5(iv) is now merely a combination of Lemmas 7.14 and 7.15.

Proof of Theorem 1.5(iv). If $\boldsymbol{u}$ has a nonremovable singularity at the origin, then according to Lemma 7.15, $\boldsymbol{u}$ does not satisfy (7-41). By Lemma 7.14, we have (1-20), proving the theorem. □

Acknowledgements

Kim was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (NRF-2014H1A8A1021602). Shahgholian was supported in part by Swedish Research Council.

This work was partly conducted during Ghergu and Kim's visit to Royal Institute of Technology (KTH) in Stockholm. They wish to thank Henrik Shahgholian for the kind invitation and KTH for the hospitality.

The authors would like to thank the anonymous referees for their valuable comments. Especially, we are grateful for one of the referees who pointed out a precise characterization of the new Pohozaev invariant as well as the explicit solution featuring the nontrivial invariant in the two-particle system.

References

[Aviles 1983] P. Aviles, "On isolated singularities in some nonlinear partial differential equations", *Indiana Univ. Math. J.* **32**:5 (1983), 773–791. MR Zbl

[Aviles 1987] P. Aviles, "Local behavior of solutions of some elliptic equations", *Comm. Math. Phys.* **108**:2 (1987), 177–192. MR Zbl

[Bidaut-Véron and Giacomini 2010] M. F. Bidaut-Veron and H. Giacomini, "A new dynamical approach of Emden–Fowler equations and systems", *Adv. Differential Equations* **15**:11-12 (2010), 1033–1082. MR Zbl

[Bidaut-Véron and Grillot 1999] M.-F. Bidaut-Veron and P. Grillot, "Singularities in elliptic systems with absorption terms", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **28**:2 (1999), 229–271. MR Zbl

[Bidaut-Véron and Raoux 1996] M.-F. Bidaut-Véron and T. Raoux, "Asymptotics of solutions of some nonlinear elliptic systems", *Comm. Partial Differential Equations* **21**:7-8 (1996), 1035–1086. MR Zbl

[Bidaut-Véron and Véron 1991] M.-F. Bidaut-Véron and L. Véron, "Nonlinear elliptic equations on compact Riemannian manifolds and asymptotics of Emden equations", *Invent. Math.* **106**:3 (1991), 489–539. MR Zbl

[Brézis and Lions 1981] H. Brézis and P.-L. Lions, "A note on isolated singularities for linear elliptic equations", pp. 263–266 in *Mathematical analysis and applications, A*, edited by L. Nachbin, Adv. in Math. Suppl. Stud. **7a**, Academic Press, New York, 1981. MR Zbl

[Busca and Manásevich 2002] J. Busca and R. Manásevich, "A Liouville-type theorem for Lane–Emden systems", *Indiana Univ. Math. J.* **51**:1 (2002), 37–51. MR Zbl

[Caffarelli et al. 1989] L. A. Caffarelli, B. Gidas, and J. Spruck, "Asymptotic symmetry and local behavior of semilinear elliptic equations with critical Sobolev growth", *Comm. Pure Appl. Math.* **42**:3 (1989), 271–297. MR Zbl

[Caffarelli et al. 2014] L. Caffarelli, T. Jin, Y. Sire, and J. Xiong, "Local analysis of solutions of fractional semi-linear elliptic equations with isolated singularities", *Arch. Ration. Mech. Anal.* **213**:1 (2014), 245–268. MR Zbl

[Caju et al. 2019] R. Caju, J. M. do Ó, and A. S. Santos, "Qualitative properties of positive singular solutions to nonlinear elliptic systems with critical exponent", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **36**:6 (2019), 1575–1601. MR Zbl

[Chen and Li 1991] W. X. Chen and C. Li, "Classification of solutions of some nonlinear elliptic equations", *Duke Math. J.* **63**:3 (1991), 615–622. MR Zbl

[Chen and Lin 2015] Z. Chen and C.-S. Lin, "Removable singularity of positive solutions for a critical elliptic system with isolated singularity", *Math. Ann.* **363**:1-2 (2015), 501–523. MR Zbl

[Druet et al. 2010] O. Druet, E. Hebey, and J. Vétois, "Bounded stability for strongly coupled critical elliptic systems below the geometric threshold of the conformal Laplacian", *J. Funct. Anal.* **258**:3 (2010), 999–1059. MR Zbl

[Emden 1907] R. Emden, *Gaskugeln: Anwendungen der mechanischen Wärmetheorie auf kosmologische und meteorologische Probleme*, Teubner, Leipzig, 1907. Zbl

[de Figueiredo 2008] D. G. de Figueiredo, "Semilinear elliptic systems: existence, multiplicity, symmetry of solutions", pp. 1–48 in *Handbook of differential equations: stationary partial differential equations, V*, edited by M. Chipot, Elsevier, Amsterdam, 2008. MR Zbl

[de Figueiredo and Felmer 1994] D. G. de Figueiredo and P. L. Felmer, "A Liouville-type theorem for elliptic systems", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **21**:3 (1994), 387–397. MR Zbl

[Fowler 1931] R. H. Fowler, "Further studies of Emden's and similar differential equations", *Quart. J. Math. Oxford Ser.* **2**:1 (1931), 259–288. Zbl

[Gidas and Spruck 1981a] B. Gidas and J. Spruck, "Global and local behavior of positive solutions of nonlinear elliptic equations", *Comm. Pure Appl. Math.* **34**:4 (1981), 525–598. MR Zbl

[Gidas and Spruck 1981b] B. Gidas and J. Spruck, "A priori bounds for positive solutions of nonlinear elliptic equations", *Comm. Partial Differential Equations* **6**:8 (1981), 883–901. MR Zbl

[Gilbarg and Serrin 1956] D. Gilbarg and J. Serrin, "On isolated singularities of solutions of second order elliptic differential equations", *J. Anal. Math.* **4** (1956), 309–340. MR Zbl

[Gilbarg and Trudinger 1983] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, 2nd ed., Grundlehren der Math. Wissenschaften **224**, Springer, 1983. MR Zbl

[Goenner and Havas 2000] H. Goenner and P. Havas, "Exact solutions of the generalized Lane–Emden equation", *J. Math. Phys.* **41**:10 (2000), 7029–7042. MR Zbl

[Grigor'yan and Sun 2014] A. Grigor'yan and Y. Sun, "On nonnegative solutions of the inequality $\Delta u + u^\sigma \le 0$ on Riemannian manifolds", *Comm. Pure Appl. Math.* **67**:8 (2014), 1336–1352. MR Zbl

[Jin et al. 2008] Q. Jin, Y. Li, and H. Xu, "Symmetry and asymmetry: the method of moving spheres", *Adv. Differential Equations* **13**:7-8 (2008), 601–640. MR Zbl

[Jin et al. 2014] T. Jin, Y. Li, and J. Xiong, "On a fractional Nirenberg problem, I: Blow up analysis and compactness of solutions", *J. Eur. Math. Soc.* **16**:6 (2014), 1111–1171. MR Zbl

[Korevaar et al. 1999] N. Korevaar, R. Mazzeo, F. Pacard, and R. Schoen, "Refined asymptotics for constant scalar curvature metrics with isolated singularities", *Invent. Math.* **135**:2 (1999), 233–272. MR Zbl

[Lane 1870] J. H. Lane, "On the theoretical temperature of the Sun, under the hypothesis of a gaseous mass maintaining its volume by its internal heat, and depending on the laws of gases known to terrestrial experiment", *Amer. J. Sci. Arts* **50** (1870), 57–74.

[Li and Zhang 2003] Y. Li and L. Zhang, "Liouville-type theorems and Harnack-type inequalities for semilinear elliptic equations", *J. Anal. Math.* **90** (2003), 27–87. MR Zbl

[Lions 1980] P.-L. Lions, "Isolated singularities in semilinear problems", *J. Differential Equations* **38**:3 (1980), 441–450. MR Zbl

[Pacard 1993] F. Pacard, "Partial regularity for weak solutions of a nonlinear elliptic equation", *Manuscripta Math.* **79**:2 (1993), 161–172. MR Zbl

[Poláčik et al. 2007] P. Poláčik, P. Quittner, and P. Souplet, "Singularity and decay estimates in superlinear problems via Liouville-type theorems, I: Elliptic equations and systems", *Duke Math. J.* **139**:3 (2007), 555–579. MR Zbl

[Reichel and Zou 2000] W. Reichel and H. Zou, "Non-existence results for semilinear cooperative elliptic systems via moving spheres", *J. Differential Equations* **161**:1 (2000), 219–243. MR Zbl

[Serrin 1964] J. Serrin, "Local behavior of solutions of quasi-linear equations", *Acta Math.* **111** (1964), 247–302. MR Zbl

[Serrin and Zou 1996] J. Serrin and H. Zou, "Non-existence of positive solutions of Lane–Emden systems", *Differential Integral Equations* **9**:4 (1996), 635–653. MR Zbl

[Serrin and Zou 2002] J. Serrin and H. Zou, "Cauchy–Liouville and universal boundedness theorems for quasilinear elliptic equations and inequalities", *Acta Math.* **189**:1 (2002), 79–142. MR Zbl

[Souplet 2009] P. Souplet, "The proof of the Lane–Emden conjecture in four space dimensions", *Adv. Math.* **221**:5 (2009), 1409–1427. MR Zbl

[Véron 1981] L. Véron, "Comportement asymptotique des solutions d'équations elliptiques semi-linéaires dans $\mathbb{R}^N$", *Ann. Mat. Pura Appl.* (4) **127** (1981), 25–50. MR Zbl

[Véron 1996] L. Véron, *Singularities of solutions of second order quasilinear equations*, Pitman Res. Notes in Math. Series **353**, Longman, Harlow, UK, 1996. MR Zbl

[Zou 2006] H. Zou, "A priori estimates and existence for strongly coupled semilinear cooperative elliptic systems", *Comm. Partial Differential Equations* **31**:4-6 (2006), 735–773. MR Zbl

Received 25 Apr 2018. Revised 22 Jan 2019. Accepted 13 Mar 2019.

MARIUS GHERGU: marius.ghergu@ucd.ie
School of Mathematics and Statistics, University College Dublin, Dublin, Ireland

and

Institute of Mathematics Simion Stoilow of the Romanian Academy, Bucharest, Romania

SUNGHAN KIM: sunghan290@snu.ac.kr
Department of Mathematical Sciences, Seoul National University, Seoul, South Korea

HENRIK SHAHGHOLIAN: henriksh@kth.se
Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 13, No. 3, 2020
dx.doi.org/10.2140/apde.2020.13.741

REGULARITY OF THE FREE BOUNDARY FOR THE VECTORIAL BERNOULLI PROBLEM

DARIO MAZZOLENI, SUSANNA TERRACINI AND BOZHIDAR VELICHKOV

We study the regularity of the free boundary for a vector-valued Bernoulli problem, with no sign assumptions on the boundary data. More precisely, given an open, smooth set of finite measure $D \subset \mathbb{R}^d$, $\Lambda > 0$, and $\phi_i \in H^{1/2}(\partial D)$, we deal with

$$\min\left\{\sum_{i=1}^{k}\int_D |\nabla v_i|^2 + \Lambda\left|\bigcup_{i=1}^{k}\{v_i \neq 0\}\right| : v_i = \phi_i \text{ on } \partial D\right\}.$$

We prove that, for any optimal vector $U = (u_1, \dots, u_k)$, the free boundary $\partial\big(\bigcup_{i=1}^{k}\{u_i \neq 0\}\big) \cap D$ is made of a regular part, which is relatively open and locally the graph of a C^∞ function, a (one-phase) singular part, of Hausdorff dimension at most $d - d^*$, for a $d^* \in \{5, 6, 7\}$, and by a set of branching (two-phase) points, which is relatively closed and of finite $\mathcal{H}^{d-1}$ measure. For this purpose we shall exploit the NTA property of the regular part to reduce ourselves to a scalar one-phase Bernoulli problem.

1. Introduction

Dirichlet problems with free boundary arise in several models describing relevant physical phenomena, such as for example thermal insulation, or flows with two liquids with jets and cavities, and have been the object of an extensive mathematical study in the last decades, starting from the seminal work [Alt and Caffarelli 1981]. The wide literature on these topics has provided many new techniques which have turned out to be extremely useful also in very different fields; see, e.g., [Caffarelli and Salsa 2005]. In two recent papers [Caffarelli et al. 2018; Mazzoleni et al. 2017] the authors dealt with problems connected with vector-valued Bernoulli-type problems, under the assumption that at least one of the components has constant sign, leaving as a main open question the opposite case (we refer to [Spolaor and Velichkov 2019] for the complete analysis in dimension 2). In this paper we give a partial answer to this question, proving some regularity of the free boundary in any dimension and without any assumption on the sign of the components. A key step is to show that in a neighborhood of any flat point, that is, a point of Lebesgue density $\frac{1}{2}$, at least one of the components has in fact a constant sign.

Mazzoleni and Terracini are partially supported by the ERC Advanced Grant 2013 no. 339958 "Complex Patterns for Strongly Interacting Dynamical Systems - COMPAT". Velichkov has been partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) project GeoSpec and the project ANR CoMeDiC.
MSC2010: primary 35R35; secondary 35J60, 49K20.
Keywords: regularity of free boundaries, NTA domains, optimality conditions, viscosity solutions, branching points.

To start with, given a smooth open set $D \subset \mathbb{R}^d$, $\Lambda > 0$, and $\Phi = (\phi_1, \dots, \phi_k) \in H^{1/2}(\partial D; \mathbb{R}^k)$, that is, $\phi_i \in H^{1/2}(\partial D)$ for $i = 1, \dots, k$, we consider the vectorial free boundary problem

$$\min\left\{\int_D |\nabla U|^2\,dx + \Lambda|\Omega_U| : U \in H^1(D; \mathbb{R}^k),\ U = \Phi \text{ on } \partial D\right\}, \tag{1-1}$$

where, for a vector-valued function $U = (u_1, \dots, u_k) : D \to \mathbb{R}^k$, we use the notation

$$|U| := \sqrt{u_1^2 + \cdots + u_k^2}, \quad |\nabla U|^2 := \sum_{i=1}^k |\nabla u_i|^2, \quad \text{and} \quad \Omega_U := \{|U| > 0\} = \bigcup_{i=1}^k \{u_i \neq 0\} \subset D.$$

We will refer to the set $\partial\Omega_U \cap D$ as the *free boundary* given by U. More precisely, for any minimizer U we shall prove that Ω_U is open and that any boundary point has a definite density. We denote by $\Omega_U^{(\gamma)}$ the set of all points having density $\gamma \in [0, 1]$. Hence we can divide the topological boundary into three parts:

$$\operatorname{Reg}(\partial\Omega_U) := \Omega_U^{(1/2)} \cap D, \quad \operatorname{Sing}_2(\partial\Omega_U) := \Omega_U^{(1)} \cap \partial\Omega_U \cap D,$$
$$\operatorname{Sing}_1(\partial\Omega_U) := (\partial\Omega_U \cap D) \setminus (\operatorname{Sing}_2(\partial\Omega_U) \cup \operatorname{Reg}(\partial\Omega_U)).$$

The main result of this paper is the following:

Theorem 1.1. *There exists a solution to problem* (1-1). *Any solution $U \in H^1(D; \mathbb{R}^k)$ is Lipschitz continuous in $D \subset \mathbb{R}^d$ and the set Ω_U has a locally finite perimeter in D. The free boundary $\partial\Omega_U \cap D$ is the union of three disjoint sets*: *a regular part* $\operatorname{Reg}(\partial\Omega_U)$, *a* (*one-phase*) *singular set* $\operatorname{Sing}_1(\partial\Omega_U)$ *and a set of branching points* $\operatorname{Sing}_2(\partial\Omega_U)$. *Moreover*, *we have*:

(1) *The regular part* $\operatorname{Reg}(\partial\Omega_U)$ *is an open subset of $\partial\Omega_U$ and is locally the graph of a C^∞ function.*

(2) *The one-phase singular set* $\operatorname{Sing}_1(\partial\Omega_U)$ *consists only of points in which the Lebesgue density of Ω_U is strictly between $\frac{1}{2}$ and* 1. *Moreover*, *there is a $d^* \in \{5, 6, 7\}$ such that*
   - *if $d < d^*$, then* $\operatorname{Sing}_1(\partial\Omega_U)$ *is empty*;
   - *if $d = d^*$, then the singular set* $\operatorname{Sing}_1(\partial\Omega_U)$ *contains at most a finite number of isolated points*;
   - *if $d > d^*$, then the $(d-d^*)$-dimensional Hausdorff measure of* $\operatorname{Sing}_1(\partial\Omega_U)$ *is locally finite in D.*

(3) *The set of branching points* $\operatorname{Sing}_2(\partial\Omega_U)$ *is a closed set of locally finite $(d-1)$-Hausdorff measure in D and consists only of points in which the Lebesgue density of Ω_U is* 1 *and the blow-up limits are linear functions.*

Our main improvement with respect to the results of the quoted papers [Caffarelli et al. 2018; Kriventsov and Lin 2018; Mazzoleni et al. 2017] is the smoothness of $\operatorname{Reg}(\partial\Omega_U)$ regardless of the sign of the boundary data. The dimension estimates of the singular set $\operatorname{Sing}_1(\partial\Omega_U)$ from claim (2) are already known (see Section 1B below) as they do not depend on the constant-sign assumption on the solution. The analysis of the blow-up limits and the dimension of $\operatorname{Sing}_2(\partial\Omega_U)$ from point (3) are, to our knowledge, new for the vectorial problem, as is the stratification result Theorem 4.3 that we prove in Section 4.

1A. ***Further results on the set of branching points*** $\mathbf{Sing_2(\partial\Omega_U)}$**.** Under the assumption that one of the components of the optimal vector has constant sign [Caffarelli et al. 2018; Mazzoleni et al. 2017], all 1-homogeneous singular solutions are multiples of the same global solution for the one-phase scalar problem. In this case, the singular set of $\partial\Omega_U$ is given precisely by $\mathrm{Sing}_1(\partial\Omega_U)$. Without the constant-sign assumption, the structure of the singular set changes drastically. A set $\mathrm{Sing}_2(\partial\Omega_U)$ of branching points, in which the free boundary may form cusps pointing inwards, might appear. This is natural since the scalar case corresponds to the two-phase Bernoulli problem, for which this is a well-known, though not completely understood, phenomenon. In particular, the dimension of this set of branching points can be as large as the dimension of the regular free boundary.[1] This is somehow natural since for the two-phase case the branching points are contact points of the two level sets $\{u > 0\}$ and $\{u < 0\}$.

The free boundaries around branching points for the vectorial problem have more complex structure. Indeed, even in dimension 2, true cusps may appear on the free boundary, that is, around a branching point x_0, the set $B_r(x_0)\cap\Omega_U$ might stay connected, while the Lebesgue density $|B_r\setminus\Omega_U|/|B_r|$ might decay as r goes to zero (see [Spolaor and Velichkov 2019] for an example of such a free boundary). On the other hand, the nodal set may also degenerate into linear subspace of codimension higher than 1 (see Lemma 2.7 for an example of homogeneous solution with a thin nodal set). In Section 2D we classify the blow-up limits. In Section 4, using a Federer reduction principle, we prove a stratification result, Theorem 4.3, for the branching points, which in particular shows that the only significant (in terms of Hausdorff measure) set of branching points is the one for which the nodal set degenerates into a $(d-1)$-dimensional plane.

1B. ***Remarks on the one-phase singular set*** $\mathbf{Sing_1(\partial\Omega_U)}$**.** The critical dimension d^* is the lowest dimension at which the free boundaries of the one-phase scalar Alt–Caffarelli problem [1981] admit singularities. Caffarelli, Jerison and Kenig proved in [Caffarelli et al. 2004] that $d^* \geq 4$, Jerison and Savin [2015] showed that $d^* \geq 5$, while De Silva and Jerison [2009] gave an example of a singular minimal cone in dimension 7, so $d^* \in \{5, 6, 7\}$. The first claim of Theorem 1.1(2) follows by the fact that at points of the one-phase singular set $\mathrm{Sing}_1(\partial\Omega_U)$ the blow-up limits of the minimizers of (1-1) are multiples of a solution of the one-phase scalar Alt–Caffarelli problem (Section 2D). The second claim of Theorem 1.1(2) was proved in [Mazzoleni et al. 2017, Section 5.5] together with the Hausdorff dimension bound

$$\dim_{\mathcal{H}}(\mathrm{Sing}_1(\partial\Omega_U)) \leq d - d^* \quad \text{for } d > d^*,$$

which follows by a dimension-reduction argument based on the Weiss' monotonicity formula [1999]. The last claim of Theorem 1.1(2) was proved in [Edelen and Engelstein 2019, Theorem 1.15] by a finer argument based on the quantitative dimension reduction of [Naber and Valtorta 2017; 2018]. We notice that [Edelen and Engelstein 2019] contains also a stratification result on $\mathrm{Sing}_1(\partial\Omega_U)$.

1C. ***Connection with shape optimization problems for the eigenvalues of the Dirichlet Laplacian.*** The vectorial Bernoulli problem is strictly related to a whole class of shape optimization problems involving the eigenvalues of the Dirichlet Laplacian. In particular, suppose that $U^* = (u_1^*, \dots, u_k^*)$ is the vector whose

[1] After this paper was accepted, the preprint [De Philippis et al. 2019] appeared, and in it a complete characterization of the two-phase free boundaries is achieved.

components are the Dirichlet eigenfunctions on the set Ω^*, a solution of the shape optimization problem

$$\min\left\{\sum_{j=1}^{k}\lambda_j(\Omega) : \Omega\subset\mathbb{R}^d,\ \Omega \text{ open},\ |\Omega|=1\right\}.$$

It was proved in [Mazzoleni et al. 2017] that U^* is a quasiminimizer of (1-1). Thus, the regularity of the optimal set Ω^* is strongly related to (not to say a consequence of) the regularity of the free boundaries of the solutions of (1-1). A result for more general functionals was proved in [Kriventsov and Lin 2018], still under some structural assumption on the free boundary. We highlight that, having the results of Sections 3A and 3B of the present paper in our hands, a possible way to obtain the $C^{1,\alpha}$ regularity of $\mathrm{Reg}(\partial\Omega_U)$ is to apply [Kriventsov and Lin 2018, Theorem 1.4]. Our approach is different and relies on the application of [Mazzoleni et al. 2017, Theorem 1.4] after having proved that there is locally always a constant-sign component. The result of [Kriventsov and Lin 2018] was then extended by the same authors to general spectral functionals in [Kriventsov and Lin 2019]. The shape optimization problem considered in the latter paper corresponds to (1-1) with sign-changing components. On the other hand the nature of the spectral functionals forces the authors to take a very different road and use an approximation with functionals for which the constant-sign assumption is automatically satisfied. A delicate point concerns the selection of a special representative of the optimal set, which, roughly speaking, corresponds to the largest *quasiopen* set which solves the problem. The problem (1-1) allows a more direct approach and in particular our regularity result holds for the free boundary of *any* optimal vector and allows for a finer classification of the branching points.

1D. ***Plan of the paper and sketch of the proof of Theorem 1.1.*** Since the existence of an optimal vector is nowadays standard, we start Section 2A by proving the Lipschitz continuity of U, which follows by the fact that each component is quasiminimizer for the scalar Alt–Caffarelli functional and so, by [Bucur et al. 2015], is Lipschitz continuous. In Section 2B we prove that the positivity set Ω_U has finite perimeter in D and that the $(d-1)$-Hausdorff measure of $\partial\Omega_U$ is finite. Our argument is different from the classical approach of Alt and Caffarelli and is based on a comparison of the energy of the different level sets of $|U|$.

In Section 2C we summarize the convergence results on the blow-up sequences and Section 2D is dedicated to the classification of the blow-up limits, which are 1-homogeneous global minimizers (that is, globally defined local minimizers) of (1-1) (see Remark 2.9). In Lemma 2.7 we show that a new class of global minimizers appears with respect to the problem considered in [Caffarelli et al. 2018; Mazzoleni et al. 2017]. In Lemma 2.11 we classify the possible blow-up limits according to the Lebesgue density; this is the main result of the section. Finally, in Definition 2.12, we define the sets $\mathrm{Reg}(\partial\Omega_U)$, $\mathrm{Sing}_1(\partial\Omega_U)$, and $\mathrm{Sing}_2(\partial\Omega_U)$.

In Section 3 we prove the smoothness of $\mathrm{Reg}(\partial\Omega_U)$. In Section 3A we prove that on the one-phase free boundary $\mathrm{Reg}(\partial\Omega_U)\cup\mathrm{Sing}_1(\partial\Omega_U)$, U satisfies the extremality condition $|\nabla|U||=\sqrt{\Lambda}$ in a viscosity sense. In Section 3B we prove that $\mathrm{Reg}(\partial\Omega_U)$ is Reifenberg flat and NTA domain.

Section 3C deals with the proof that in a neighborhood of a point $x_0\in\mathrm{Reg}(\partial\Omega_U)$ at least one of the components of U remains strictly positive and (up to a multiplicative constant) controls $|U|$ (see Lemma 3.10). This is the main result of this section and the proof is based on the geometric properties of

NTA domains and on the boundary Harnack principle. In Sections 3D and 3E we prove that $\mathrm{Reg}(\partial\Omega_U)$ is, respectively, $C^{1,\alpha}$ and C^∞. The result of Lemma 3.10 allows us to apply the results from [Mazzoleni et al. 2017]. We give the main steps of the proof for the sake of completeness.

Section 4 is dedicated to the study of the set $\mathrm{Sing}_2(\partial\Omega_U)$ of points $x_0 \in \partial\Omega_U$ in which all the blow-up limits $U_0 \in \mathcal{BU}_U(x_0)$ are linear functions of the form $U_0(x) = Ax$. In Section 4A we prove that the rank of the linear map U_0 depends only on the point x_0 and we define the j-th stratum $\mathcal{S}_j$ as the set of points for which this rank is precisely j. In Section 4B we use a dimension-reduction argument in the spirit of Federer to prove that the Hausdorff dimension of each stratum $\mathcal{S}_j$ is $d - j$. Finally, in Section 4C we give a criterion for the uniqueness of the blow-up limits in terms of the Lebesgue density of Ω_U.

2. Boundary behavior of the solutions

The existence of an optimal vector $U = (u_1, \dots, u_k)$ is standard and follows by the direct method of the calculus of variations; for more details we refer to [Alt and Caffarelli 1981].

2A. *Lipschitz continuity and nondegeneracy.* Any minimizer U has the following properties:

(i) The vector-valued function $U : D \to \mathbb{R}^k$ is locally Lipschitz continuous in D.

(ii) The real-valued function $|U|$ is nondegenerate; i.e., there are constants $c_0 > 0$ and $r_0 > 0$ such that for every $x_0 \in \partial\Omega_U \cap D$ and $r \in (0, r_0]$ we have

$$\fint_{B_r(x_0)} |U|\,dx < c_0 r \quad \Longrightarrow \quad U \equiv 0 \quad \text{in } B_{r/2}(x_0). \tag{2-1}$$

(iii) There are constants ε_0, r_0 such that the *lower density estimate* holds:

$$\varepsilon_0 |B_r| \le |\Omega_U \cap B_r(x_0)| \quad \text{for every } x_0 \in \partial\Omega_U \cap D \text{ and } r \le r_0. \tag{2-2}$$

Remark 2.1. Claim (i) in particular implies that, *for every* minimizer U of (1-1), the set Ω_U is open.

Remark 2.2. It is important to highlight that, unlike the case treated in [Mazzoleni et al. 2017; Caffarelli et al. 2018] where it was assumed at least one component u_i was positive, we cannot hope to have a density estimate *from above* on $\partial\Omega_U \cap D$. Actually, we expect a set of branching points (cusps) will come out. Indeed, the case $k = 1$ corresponds to a scalar *two-phase* problem for which (at least in dimension 2) the set Ω_U is composed of two $C^{1,\alpha}$ sets; see [Spolaor and Velichkov 2019]. At the points of the common boundary of these two sets, the Lebesgue density of Ω_U is 1.

Proof of (i). The Lipschitz continuity of each component u_i, $i = 1, \dots, k$, descends from a quasiminimality property. Indeed, reasoning as in [Mazzoleni et al. 2017, Section 6.2], for every $\tilde u_i : D \to \mathbb{R}$ such that $\tilde u_i - u_i \in H^1_0(D)$ we consider the competitor $\widetilde U := (u_1, \dots, \tilde u_i, \dots, u_k)$. By the optimality of U we have

$$\int_D |\nabla u_i|^2\,dx + \Lambda|\{|U| > 0\}| \le \int_D |\nabla \tilde u_i|^2\,dx + \Lambda|\{|\widetilde U| > 0\}|,$$

which implies that each component u_i is a quasiminimizer of the Dirichlet energy; that is,

$$\int |\nabla u_i|^2\,dx \le \int |\nabla \tilde u_i|^2\,dx + \Lambda|B_r| \quad \text{for every } \tilde u_i \text{ such that } \tilde u_i - u_i \in H^1_0(B_r). \tag{2-3}$$

Applying [Bucur et al. 2015, Theorem 3.3] we get that u_i is Lipschitz continuous in D, and since $i = 1, \dots, k$ is arbitrary, so is U. □

Proof of (ii) *and* (iii). The nondegeneracy of $|U|$ follows by [Mazzoleni et al. 2017, Lemma 2.6], which can be applied since U satisfies the condition (2.9) therein with $K = 0$ and $\varepsilon > 0$. Finally, we notice that the density estimate *from below* (2-2) holds for every Lipschitz function satisfying the nondegeneracy condition (2-1); see for example [Mazzoleni et al. 2017, Lemma 2.11; Alt and Caffarelli 1981]. □

2B. *Finiteness of the perimeter.* For any $U \in H^1(D; \mathbb{R}^k)$ solution of (1-1), the set Ω_U has locally finite perimeter in D and, moreover,

$$\mathcal{H}^{d-1}(\partial\Omega_U \cap K) < \infty \quad \text{for every compact set } K \subset D. \tag{2-4}$$

Remark 2.3. We notice that the condition (2-4) is more general than the finiteness of the perimeter since $\partial\Omega_U$ may contain points x_0 which are in the measure-theoretic interior of Ω_U, that is, $|B_r(x_0) \setminus \Omega_U| = 0$.

In order to prove the claim of this subsection, we will use the following lemma, which holds in general.

Lemma 2.4. *Suppose that $D \subset \mathbb{R}^d$ is an open set and that $\phi : D \to [0, +\infty]$ is a function in $H^1(D)$ for which there exist $\bar\varepsilon > 0$ and $C > 0$ such that*

$$\int_{\{0<\phi\le\varepsilon\}\cap D} |\nabla\phi|^2\,dx + \Lambda|\{0 < \phi \le \varepsilon\} \cap D| \le C\varepsilon \quad \textit{for every } 0 < \varepsilon \le \bar\varepsilon. \tag{2-5}$$

Then $P(\{\phi > 0\}; D) \le C\sqrt{\Lambda}$.

Proof. By the coarea formula, the Cauchy–Schwarz inequality and (2-5), we have that, for every $\varepsilon \le \bar\varepsilon$,

$$\int_0^\varepsilon \mathcal{H}^{d-1}(\{\phi = t\} \cap D)\,dt = \int_{\{0<\phi\le\varepsilon\}} |\nabla\phi|\,dx \le \left(\int_{\{0<\phi\le\varepsilon\}} |\nabla\phi|^2\,dx\right)^{1/2} |\{0 < \phi \le \varepsilon\}|^{1/2} \le \varepsilon C\sqrt{\Lambda}.$$

Taking $\varepsilon = 1/n$, we get that there is $\delta_n \in [0, 1/n]$ such that

$$\mathcal{H}^{d-1}(\partial^*\{\phi > \delta_n\} \cap D) \le n \int_0^{1/n} \mathcal{H}^{d-1}(\{\phi = t\} \cap D)\,dt \le C\sqrt{\Lambda}.$$

Passing to the limit as $n \to \infty$, we obtain $\mathcal{H}^{d-1}(\partial^*\{\phi > 0\} \cap D) \le C\sqrt{\Lambda}$, which concludes the proof of the lemma. □

Lemma 2.5. *Let $D \subset \mathbb{R}^d$ be an open set and $\phi : D \to [0, +\infty)$ be a Lipschitz function such that*:

(a) *ϕ is nondegenerate, that is, there is a constant $c > 0$ such that*

$$\sup_{B_r(x_0)} \phi \ge cr \quad \textit{for every } x_0 \in \partial\Omega_\phi \cap D \textit{ and every } 0 < r < \operatorname{dist}(x_0, \partial D),$$

where we have set $\Omega_\phi = \{\phi > 0\}$.

(b) *There is a constant $C > 0$ such that ϕ satisfies the estimate*

$$|\{0 < \phi \le \varepsilon\} \cap D| \le C\varepsilon \quad \textit{for every } \varepsilon > 0.$$

Then, for every compact set $K \subset \Omega$, we have $\mathcal{H}^{d-1}(K \cap \partial\Omega_\phi) < \infty$.

Proof. Let us first recall that the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}(A)$ of a set $A \subset \mathbb{R}^d$ is given by $\mathcal{H}^{d-1}(A) = \lim_{\delta\to 0} \mathcal{H}^{d-1}_\delta(A)$, where, for every $\delta > 0$, we set

$$\mathcal{H}^{d-1}_\delta(A) = \omega_{d-1} \inf\left\{\sum_{j=1}^{\infty} r_j^{d-1} : \text{for every } B_{r_j}(x_j) \text{ such that } \bigcup_{j=1}^{\infty} B_{r_j}(x_j) \supset A \text{ and } r_j \le \delta\right\}.$$

Let $\delta > 0$ be fixed. By the Vitali covering lemma, there is a family of balls $\{B_\delta(x_j)\}_{j=1}^N$ which is a covering of $K \cap \partial\Omega_\phi$, $x_j \in \partial\Omega_\phi$ for every $j = 1, \dots, N$, and the balls $B_{\delta/5}(x_j)$ are disjoint. The nondegeneracy of ϕ implies that in every ball $B_{\delta/10}(x_j)$ there is a point y_j such that $\phi(y_j) \ge c\delta/10$. The Lipschitz continuity of u implies that $B_{c\delta/10L}(y_j) \subset \Omega_\phi$, where $L = \max\{1, \|\nabla\phi\|_{L^\infty}\}$. On the other hand, since $\phi(x_j) = 0$, we have

$$\phi < L\left(\frac{c\delta}{10L} + \frac{c\delta}{10}\right) = (L+1)\frac{c\delta}{10} \quad \text{on } B_{c\delta/10L}(y_j).$$

This implies that the balls $B_{c\delta/10L}(y_j)$, $j = 1, \dots, N$, are pairwise disjoint and contained in the set

$$\left\{0 < \phi < (L+1)\frac{c\delta}{10}\right\}.$$

Now, the estimate from point (b) implies

$$C(L+1)\frac{c\delta}{10} \ge \sum_{j=1}^{N} |B_{c\delta/10L}(y_j)| \ge N\omega_d \frac{c^d\delta^d}{L^d 10^d},$$

which implies

$$N\, d\omega_d \delta^{d-1} \le dC \frac{10^{d-1}}{c^{d-1}} L^d(L+1).$$

Since, the right-hand side does not depend on δ, we get

$$\mathcal{H}^{d-1}(K \cap \partial\Omega_\phi) \le dC \frac{10^{d-1}}{c^{d-1}} L^d(L+1). \qquad \square$$

Proof of the claim (2-4) *of Section 2B.* We aim to prove an estimate of the form (2-5) for $\phi = |U|$ by constructing a suitable competitor, which is inspired by the approach of Dorin Bucur [2012]. Since we aim to prove a local result, we take $x_0 \in \partial\Omega \cap D$ and $B_r(x_0) \subset D$; moreover, we can assume without loss of generality that $x_0 = 0$ and $r = 1$. Setting $\rho := |U|$, for every $\varepsilon > 0$, we define

$$\widetilde{U} = (\tilde{u}_1, \dots, \tilde{u}_k) := \frac{(\rho-\varepsilon)_+}{\rho} U, \quad \text{where } \tilde{u}_i = \left(1 - \frac{\varepsilon}{\rho}\right)_+ u_i \text{ for every } i = 1, \dots, k,$$

and, for a smooth function $\phi \in C^\infty(D)$ such that $0 \le \phi \le 1$ in D, $\phi = 1$ in $B_{1/2}$, and $\phi = 0$ on $D \setminus B_1$,

$$V = (v_1, \dots, v_k) := (1-\phi)U + \phi\widetilde{U} = \begin{cases} (1-\phi)U & \text{if } |U| = \rho < \varepsilon, \\ (1-\varepsilon\phi/\rho)U & \text{if } |U| = \rho \ge \varepsilon. \end{cases}$$

Thus, clearly V is an admissible competitor in problem (1-1).

We observe that the following relations, which we will use in the rest of the proof, hold true:

$$|\nabla\rho|\le|\nabla U| \quad\text{and}\quad \sum_{i=1}^k u_i\nabla u_i=\rho\nabla\rho \quad\text{in } D,$$
$$\frac{\varepsilon\phi}{\rho}\le 1 \quad\text{in } \{|U|\ge\varepsilon\}.$$

We can now compute on $\{|U|\ge\varepsilon\}$

$$\begin{aligned}|\nabla V|^2-|\nabla U|^2&=\sum_i\left|\nabla\left(1-\frac{\varepsilon\phi}{\rho}\right)u_i\right|^2-|\nabla u_i|^2\\&=\left(-\frac{2\varepsilon\phi}{\rho}+\frac{\varepsilon^2\phi^2}{\rho^2}\right)|\nabla U|^2+\rho^2\left|\nabla\frac{\varepsilon\phi}{\rho}\right|^2-2(\rho-\varepsilon\phi)\nabla\rho\cdot\nabla\frac{\varepsilon\phi}{\rho}\\&=\varepsilon^2|\nabla\phi|^2-2\varepsilon\nabla\phi\cdot\nabla\rho+(|\nabla\rho|^2-|\nabla U|^2)\left(2\varepsilon\frac{\phi}{\rho}-\varepsilon^2\frac{\phi^2}{\rho^2}\right)\le\varepsilon^2|\nabla\phi|^2-2\varepsilon\nabla\phi\cdot\nabla\rho\le C_1\varepsilon,\end{aligned}$$

where C_1 depends only on $\|\nabla\phi\|_{L^\infty}$ and $\|\nabla U\|_{L^\infty}$. Next, on the set $\{|U|<\varepsilon\}$, we compute

$$\begin{aligned}|\nabla U|^2-|\nabla V|^2&=|\nabla U|^2-|\nabla(1-\phi)U|^2\\&=(2\phi-\phi^2)|\nabla U|^2+2(1-\phi)U\nabla\phi\cdot\nabla U+|U|^2|\nabla\phi|^2\ge|\nabla U|^2\mathbb{1}_{B_{1/2}}-C_2\varepsilon,\end{aligned}$$

where again C_2 depends only on $\|\nabla\phi\|_{L^\infty}$ and $\|\nabla U\|_{L^\infty}$. By testing the optimality of U with V we get

$$\int_{B_1}|\nabla U|^2+\Lambda|\{0<|U|\le\varepsilon\}\cap B_1|\le\int_{B_1}|\nabla V|^2+\Lambda|\{|V|>0\}\cap B_1|,$$

so we deduce

$$\int_{\{0<|U|\le\varepsilon\}}(|\nabla U|^2-|\nabla V|^2)+\Lambda|\{0<|U|\le\varepsilon\}\cap B_{1/2}|\le\int_{\{|U|\ge\varepsilon\}}(|\nabla V|^2-|\nabla U|^2)\le C_1\varepsilon,$$

and finally, since $V=0$ on the set $\{0<|U|\le\varepsilon\}\cap\{\phi=1\}$, we get

$$\int_{\{0<|U|\le\varepsilon\}\cap B_{1/2}}|\nabla U|^2+\Lambda|\{0<|U|\le\varepsilon\}\cap B_{1/2}|\le(C_1+C_2)\varepsilon,$$

and, since $|\nabla\rho|\le|\nabla U|$ we obtain the estimate (2-5) for $\rho=|U|$ in the ball $B_{1/2}$. By Lemmas 2.4 and 2.5, we get that Ω_U has locally finite perimeter in D and that (2-4) holds. □

2C. ***Compactness and convergence of the blow-up sequences.*** Let $U:D\to\mathbb{R}^k$ be a solution of (1-1) or, more generally, a Lipschitz function. For $r\in(0,1)$ and $x\in\mathbb{R}^d$ such that $U(x)=0$, we define

$$U_{r,x}(y):=\frac{1}{r}U(x+ry).$$

When $x=0$ we will use the notation $U_r:=U_{r,0}$.

Suppose now that $(r_n)_{n\ge0}\subset\mathbb{R}^+$ and $(x_n)_{n\ge0}\subset D$ are two sequences such that

$$\lim_{n\to\infty}r_n=0,\qquad\lim_{n\to\infty}x_n=x_0\in D,\qquad B_{r_n}(x_n)\subset D\quad\text{and}\quad x_n\in\partial\{|U|>0\}\quad\text{for every } n\ge0. \tag{2-6}$$

Then the sequence $\{U_{r_n,x_n}\}_{n\in\mathbb{N}}$ is uniformly Lipschitz and locally uniformly bounded in $\mathbb{R}^d$. Thus, up to a subsequence, U_{r_n,x_n} converges, as $n\to\infty$, locally uniformly to a Lipschitz continuous function $U_0:\mathbb{R}^d\to\mathbb{R}^k$. Moreover, if U is a minimizer of (1-1), then for every $R>0$ the following properties hold; see [Mazzoleni et al. 2017, Proposition 4.5]:

(i) U_{r_n,x_n} converges to U_0 strongly in $H^1(B_R;\mathbb{R}^k)$.

(ii) The sequence of characteristic functions $\mathbb{1}_{\Omega_n}$ converges in $L^1(B_R)$ to $\mathbb{1}_{\Omega_0}$, where

$$\Omega_n := \{|U_{r_n}|>0\} \quad \text{and} \quad \Omega_0 := \{|U_0|>0\}.$$

(iii) The sequences of closed sets $\overline{\Omega}_n$ and Ω_n^c Hausdorff converge in B_R respectively to $\overline{\Omega}_0$ and Ω_0^c.

(iv) U_0 is nondegenerate at zero; that is, there is a dimensional constant $c_d>0$ such that

$$\|U_0\|_{L^\infty(B_r)} \ge c_d r \quad \text{for every } r>0.$$

Definition 2.6. Let $U:\mathbb{R}^d\to\mathbb{R}^k$ be a Lipschitz function, and r_n and x_n be two sequences satisfying (2-6). We say that the sequence U_{r_n,x_n} is a *blow-up sequence with variable center* (or a *pseudo-blow-up*). If the sequence x_n is constant, $x_n=x_0$ for every $n\ge 0$, we say that U_{r_n,x_0} is a *blow-up sequence with fixed center*. We denote by $\mathcal{BU}_U(x_0)$ the space of all the limits of blow-up sequences with fixed center x_0.

2D. *Classification of the blow-up limits.* In this section we prove that for any $x_0\in\partial\Omega_U\cap D$ the blow-up limits $U_0\in\mathcal{BU}_U(x_0)$ have one of the following forms:

• *Multiples of a scalar solution of the one-phase problem*, that is, there is a 1-homogeneous nonnegative global minimizer $u:\mathbb{R}^d\to\mathbb{R}^+$ of the one-phase Alt–Caffarelli functional

$$\mathcal{F}(u)=\int|\nabla u|^2\,dx+\Lambda|\{u>0\}|$$

such that

$$U_0(x)=\xi u(x), \quad \text{where } \xi\in\mathbb{R}^k \text{ and } |\xi|=1. \tag{2-7}$$

• *Linear functions*, that is, there is a matrix $A=(a_{ij})_{ij}\in\mathcal{M}_{d\times k}(\mathbb{R})$ such that

$$U_0(x)=Ax. \tag{2-8}$$

It was shown in [Mazzoleni et al. 2017] that every function of the form (2-7) is a global solution of (1-1). In the following lemma we classify the linear solutions.

Lemma 2.7. *Let $U:\mathbb{R}^d\to\mathbb{R}^k$ be a linear function, $U(x)=Ax$ with $A=(a_{ij})_{ij}\in M_{d\times k}(\mathbb{R})$. If*

$$\|A\|^2:=\sum_{i=1}^k\sum_{j=1}^d a_{ij}^2\ge\Lambda,$$

then U is a solution of (1-1) *in the unit ball B_1. Moreover, if* rank $A=1$, *then the condition $\|A\|^2\ge\Lambda$ is also necessary.*

Proof. We first prove that if $\|A\|^2 \geq \Lambda$, then $U(x) = Ax$ is a solution of (1-1). Let $U =: (u_1, \dots, u_k)$ and let the vector-valued function $\widetilde{U} = (\tilde{u}_1, \dots, \tilde{u}_k) : B_1 \to \mathbb{R}^d$ be such that $\widetilde{U} = U$ on ∂B_1. We will show that $\widetilde{U}$ has a higher energy than U. Notice that each component u_j, $j = 1, \dots, k$, can be written as $u_j(x) = \alpha_j v_j(x)$, where $\alpha_j \in \mathbb{R}$ and $v_j(x) = x \cdot \nu_j$ for some $\nu_j \in \partial B_1$. We will also write $\tilde{u}_j(x) = \alpha_j \tilde{v}_j(x)$ and we notice that $\tilde{v}_j = v_j$ on ∂B_1. Now since $(v_j)_+$ and $(v_j)_-$ are solutions of the one-phase scalar Alt–Caffarelli problem we have

$$\begin{aligned}\int_{B_1} |\nabla v_j|^2\,dx + |B_1| &= \int_{B_1} |\nabla (v_j)_+|^2\,dx + |\{v_j > 0\} \cap B_1| + \int_{B_1} |\nabla (v_j)_-|^2\,dx + |\{v_j < 0\} \cap B_1| \\ &\leq \int_{B_1} |\nabla (\tilde{v}_j)_+|^2\,dx + |\{\tilde{v}_j > 0\} \cap B_1| + \int_{B_1} |\nabla (\tilde{v}_j)_-|^2\,dx + |\{\tilde{v}_j < 0\} \cap B_1| \\ &\leq \int_{B_1} |\nabla \tilde{v}_j|^2\,dx + |\Omega_{\tilde{u}} \cap B_1|.\end{aligned}$$

Multiplying by α_j^2, taking the sum over j, and using that $\|A\|^2 = \sum_{j=1}^k \alpha_j^2$, we obtain

$$\begin{aligned}\int_{B_1} |\nabla U|^2\,dx + \|A\|^2 |B_1| &= \sum_{j=1}^k \alpha_j^2 \biggl(\int_{B_1} |\nabla v_j|^2\,dx + |B_1| \biggr) \\ &\leq \sum_{j=1}^k \alpha_j^2 \biggl(\int_{B_1} |\nabla \tilde{v}_j|^2\,dx + |\Omega_{\widetilde{U}} \cap B_1| \biggr) = \int_{B_1} |\nabla \widetilde{U}|^2\,dx + \|A\|^2 |\Omega_{\widetilde{U}} \cap B_1|.\end{aligned}$$

Now since $\Lambda \leq \|A\|^2$, we have

$$\int_{B_1} |\nabla U|^2\,dx + \Lambda |B_1| \leq \int_{B_1} |\nabla \widetilde{U}|^2\,dx + \Lambda |\Omega_{\widetilde{U}} \cap B_1|.$$

We will now prove that if rank $A = 1$ and $U(x) = Ax$ is a solution to (1-1), then $\|A\|^2 \geq \Lambda$. Indeed, let $U = (u_1, \dots, u_k)$. The rank-1 condition implies that there is a vector $\nu \in \partial B_1$ such that $u_j(x) = \alpha_j x \cdot \nu_j$ for some $\nu_j \in \partial B_1$. Let

$$u_j^\varepsilon := u_j \circ \Phi_\varepsilon^{-1}, \quad \text{where } \Phi_\varepsilon(x) = x - \varepsilon \phi(x) \nu,$$

$\phi \in C^\infty(B_1)$ being a nonnegative function with compact support in B_1. Setting $U_\varepsilon = (u_1^\varepsilon, \dots, u_k^\varepsilon)$, the optimality of U gives

$$\int_{B_1} |\nabla U_\varepsilon|^2\,dx + |\{|U_\varepsilon| > 0\} \cap B_1| \geq \int_{B_1} |\nabla U|^2\,dx + |\{|U| > 0\} \cap B_1|.$$

Taking the derivative in $\varepsilon = 0$, see [Alt and Caffarelli 1981, Section 2], we get

$$\sum_{j=1}^k \int_{B_1 \cap \partial\Omega_U} |\nabla u_j|^2 \phi\, d\mathcal{H}^{d-1} - \Lambda \int_{B_1 \cap \partial\Omega_U} \phi\, d\mathcal{H}^{d-1} \geq 0.$$

Since $|\nabla u_j|^2 = \alpha_j^2$, we obtain $\sum_{j=1}^k \alpha_j^2 \geq \Lambda$, which proves the claim. □

The classification of the blow-up limits strongly relies on the monotonicity of the vectorial Weiss' boundary-adjusted energy introduced in [Mazzoleni et al. 2017]

$$W(U,x_0,r):=\frac{1}{r^d}\left(\int_{B_r(x_0)}|\nabla U|^2\,dx+\Lambda|\{|U|>0\}\cap B_r(x_0)|\right)-\frac{1}{r^{d+1}}\int_{\partial B_r(x_0)}|U|^2\,d\mathcal{H}^{d-1},\qquad(2\text{-}9)$$

which turns out to be monotone in r. Precisely, by Proposition 3.1 of that paper we have the following estimate.

Lemma 2.8 (Weiss monotonicity formula). *Let $U=(u_1,\dots,u_k)$ be a minimizer for problem* (1-1) *and $x_0\in\partial\Omega_U\cap D$. Then, the function $r\mapsto W(U,x_0,r)$ is nondecreasing and*

$$\frac{d}{dr}W(U,x_0,r)\geq\frac{1}{r^{d+2}}\sum_{i=1}^{k}\int_{\partial B_r(x_0)}|(x-x_0)\cdot\nabla u_i-u_i|^2\,d\mathcal{H}^{d-1}(x);\qquad(2\text{-}10)$$

in particular, the limit $\lim_{r\to 0^+}W(U,x_0,r)$ *exists and is finite.*

Remark 2.9 (homogeneity and minimality of the blow-up limits). As a consequence of the monotonicity formula, we obtain that if U is a solution of (1-1), $x_0\in\partial\Omega_U\cap D$, and $U_0\in\mathcal{BU}_U(x_0)$, then U_0 is a 1-homogeneous global solution of the vectorial Bernoulli problem. Precisely, the fact that U_0 is a global solution follows by [Mazzoleni et al. 2017, Proposition 4.2], while for the homogeneity of U_0 we use the fact that U_0 is a blow-up limit, $U_0=\lim_{n\to\infty}U_{r_n,x_0}$, and the scaling property of the Weiss energy

$$W(U,x_0,rs)=W(U_{r,x_0},s,0)\quad\text{for every } r,s>0,$$

which gives that the function $s\mapsto W(U_0,s,0)$ is constant. In fact, for every $s>0$, we have

$$W(U_0,s,0)=\lim_{n\to\infty}W(U_{r_n,x_0},s,0)=\lim_{n\to\infty}W(U,r_ns,x_0)=\lim_{r\to 0}W(U,r,x_0).$$

Now, the homogeneity of U_0 follows by (2-10) applied to U_0 and its components.

Remark 2.10 (Lebesgue and energy density). Keeping the notation from Remark 2.9, we notice that the homogeneity of the blow-up limits and the strong convergence of the blow-up sequences give

$$W(U_0,1,0)=\Lambda|\{|U_0|>0\}\cap B_1|=\lim_{r\to 0}W(U,r,x_0)=\Lambda\omega_d\lim_{r\to 0}\frac{|\Omega_U\cap B_r(x_0)|}{|B_r|}$$

for every $U_0\in\mathcal{BU}_U(x_0)$. That is, the energy density $\lim_{r\to 0}W(U,r,x_0)$ coincides, up to a multiplicative constant, with the Lebesgue density, which (as a consequence) exists in every point x_0 of the free boundary. In particular, we get

$$\Omega_U^{(\gamma)}=\left\{x\in\partial\Omega_U:\lim_{r\to 0}\frac{|\Omega_U\cap B_r(x)|}{|B_r|}=\gamma\right\}=\left\{x_0\in\partial\Omega_U:\lim_{r\to 0}W(U,x_0,r)=\Lambda\omega_d\gamma\right\}.$$

Lemma 2.11 (structure of the blow-up limits). *Let U be a solution of* (1-1), $x_0\in\partial\Omega_U\cap D$. *Then, there is a dimensional constant $0<\delta<\frac{1}{2}$ such that precisely one of the following holds*:

(i) *The Lebesgue density of Ω_U at x_0 is $\frac{1}{2}$ and every blow-up $U_0\in\mathcal{BU}_U(x_0)$ is of the form*

$$U_0(x)=\xi(x\cdot\nu)_+,\quad\textit{where } \xi\in\mathbb{R}^k,\ |\xi|=\sqrt{\Lambda},\ \nu\in\mathbb{R}^d,\ |\nu|=1.\qquad(2\text{-}11)$$

(ii) *The Lebesgue density of* Ω_U *at* x_0 *satisfies*

$$\tfrac{1}{2}+\delta \le \lim_{r\to 0}\frac{|\Omega_U\cap B_r(x_0)|}{|B_r|}\le 1-\delta,$$

and every blow-up in $\mathcal{BU}_U(x_0)$ *is a one-phase blow-up of the form* (2-7) *with singularity in zero.*

(iii) *The Lebesgue density of* Ω_U *at* x_0 *is* 1 *and every blow-up in* $\mathcal{BU}_U(x_0)$ *is of the form* (2-8).

Proof. Let $x_0 \in \partial\Omega_U \cap D$.

Step 1: The following claim holds true:

$$\begin{aligned} x_0\in\Omega_U^{(1/2)} \quad &\iff \quad \text{there is } U_0\in\mathcal{BU}_U(x_0) \text{ of the form (2-11)}\\ &\iff \quad \text{every } U_0\in\mathcal{BU}_U(x_0) \text{ is of the form (2-11).}\end{aligned}$$

Indeed, if one blow-up is of the form (2-11), then by Remark 2.10 $x_0\in\Omega^{(1/2)}$. On the other, hand, if $x_0\in\Omega^{(1/2)}$ and $U_0\in\mathcal{BU}_U(x_0)$, then again by Remark 2.10 $|\Omega_{U_0}\cap B_1|=\frac{1}{2}|B_1|$. The homogeneity of U_0 and the fact that $\Delta U_0=0$ on Ω_{U_0} imply that each component of U_0 is an eigenfunction on the sphere corresponding to the eigenvalue $d-1$. By the Faber–Krahn inequality on the sphere we get that, up to a rotation, $\Omega_{U_0}=\{x_d>0\}$ and all the components of U_0 are multiples of x_d^+, that is, $U_0(x)=\xi x_d^+$ for some $\xi\in\mathbb{R}^k$. Let ϕ be a compactly supported function and let $\widetilde{U}_0=\xi(x_d^+ +\phi)$. Testing the optimality of U_0 against $\widetilde{U}_0$, it is immediate to check, see [Mazzoleni et al. 2017], that $|\xi|x_d^+$ is a global minimizer of the one-phase Alt–Caffarelli functional. Thus, an internal perturbation, see [Alt and Caffarelli 1981], gives $|\xi|=\sqrt{\Lambda}$ and concludes Step 1.

Step 2: The following claim holds true:

$$\begin{aligned} x_0\in\Omega_U^{(1)} \quad &\iff \quad \text{there is } U_0\in\mathcal{BU}_U(x_0) \text{ of the form (2-8)}\\ &\iff \quad \text{every } U_0\in\mathcal{BU}_U(x_0) \text{ is of the form (2-8).}\end{aligned}$$

Indeed, if one blow-up $U_0\in\mathcal{BU}_U(x_0)$ is of the form (2-8), then by Remark 2.10 $x_0\in\Omega_U^{(1)}$. On the other hand, if $x_0\in\Omega_U^{(1)}$, then still by Remark 2.10 $|U_0\cap B_1|=|B_1|$ and so, the minimality of U_0 implies that U_0 is harmonic in B_1. Now the homogeneity of U_0 implies that it is a linear function, $U_0(x)=Ax$, for some matrix $A=(a_{ij})_{ij}$.

Step 3: Finally, suppose that $x_0\in(\partial\Omega_U\cap D)\setminus(\Omega_U^{(1/2)}\cup\Omega_U^{(1)})$ and let $x_0\in\Omega_U^{(\gamma)}$ for some $\gamma\in\big(0,\frac{1}{2}\big)\cup\big(\frac{1}{2},1\big)$. Let $U_0=(u_1,\dots,u_k)\in\mathcal{BU}_U(x_0)$. Then each component u_i is 1-homogeneous and the functions u_i^+ and u_i^- are eigenfunctions corresponding to the eigenvalue $d-1$ on the spherical sets $\{u_i>0\}\cap\partial B_1$ and $\{u_i<0\}\cap\partial B_1$. Now since the density γ is at most 1, we get that at least one of the sets is empty. Thus, none of the components u_i change sign and they are all multiples of the first eigenfunction on the set $\Omega_{U_0}\cap\partial B_1$; that is, $U_0=\xi|U_0|$ for some $\xi\in\mathbb{R}^k$. Now, reasoning as in [Mazzoleni et al. 2017, Section 5.2], we get that $|\xi|=\Lambda$ and that $|U_0|$ is a global solution of the one-phase scalar functional $u\mapsto\int|\nabla u|^2\,dx+|\{u>0\}|$. In particular, the density estimate for the one-phase Alt–Caffarelli functional implies that $\gamma<1-\delta$ for some dimensional constant $\delta>0$. Now, the fact that the first eigenvalue on $\Omega_{U_0}\cap\partial B_1$ is $d-1$ implies $\gamma\ge\frac{1}{2}$. As in [Mazzoleni et al. 2017, Section 5.2], the improvement of flatness for the scalar problem now implies $\gamma>\frac{1}{2}+\delta$, which concludes the proof. □

Definition 2.12. Let $x_0 \in \partial\Omega_U$. We say that:

- x_0 is a regular point, $x_0 \in \mathrm{Reg}(\partial\Omega_U)$, if Lemma 2.11(i) holds.
- x_0 is a (one-phase) singular point, $x_0 \in \mathrm{Sing}_1(\partial\Omega_U)$, if Lemma 2.11(ii) holds.
- x_0 is a branching point, $x_0 \in \mathrm{Sing}_2(\partial\Omega_U)$, if Lemma 2.11(iii) holds.

In view of Lemma 2.11 we have

$$\mathrm{Reg}(\partial\Omega_U) = \Omega_U^{(1/2)} \cap D, \quad \mathrm{Sing}_2(\partial\Omega_U) = \Omega_U^{(1)} \cap \partial\Omega_U \cap D,$$
$$\mathrm{Sing}_1(\partial\Omega_U) = (\partial\Omega_U \cap D) \setminus (\mathrm{Sing}_2(\partial\Omega_U) \cup \mathrm{Reg}(\partial\Omega_U)).$$

Lemma 2.13. $\mathrm{Sing}_2(\partial\Omega_U)$ *is a closed set and* $\mathrm{Reg}(\partial\Omega_U)$ *is an open subset of* $\partial\Omega_U$.

Proof. We first notice that the function $W(U, x_0, 0) := \lim_{r\to 0^+} W(U, x_0, r)$ is upper semicontinuous in x_0. This follows by the fact that $(x_0, r) \mapsto W(U, x_0, r)$ is increasing in $r > 0$ and continuous in x_0. Thus, the first part of the claim follows since in the points $x_0 \in \mathrm{Sing}_2(\partial\Omega_U)$ the density $W(U, x_0, 0)$ is maximal. The second part of the claim follows by the lower density gap from Lemma 2.11(ii) and the argument of [Mazzoleni et al. 2017, Proposition 5.6]. □

3. Regularity of the one-phase free boundary

Following the argument from [Mazzoleni et al. 2017], we first deduce the optimality condition on the free boundary in a viscosity sense; then we notice that $\mathrm{Reg}(\partial\Omega_U)$ is open and Reifenberg flat. Next we show that around every point of $\mathrm{Reg}(\partial\Omega_U)$ at least one of the components of the optimal vector U has a constant sign. Thus we fall into the framework of [loc. cit.] and can conclude the proof by using the boundary Harnack principle in NTA domains and the regularity of the one-phase free boundaries for the scalar problem. Finally, thanks to Lemma 3.10, we can apply the arguments of [loc. cit., Section 5] in order to obtain the C^∞ regularity of $\mathrm{Reg}(\partial\Omega_U)$, using the component of locally constant sign provided by Lemma 3.10 instead of u_1 in the boundary Harnack principle [loc. cit., Lemma 5.12]. We recall here the updated statements for the reader's sake.

3A. ***The stationarity condition on the free boundary.*** It is well known, see for example [Alt and Caffarelli 1981], that if u is a local minimizer of the Alt–Caffarelli functional

$$H^1_{\mathrm{loc}}(\mathbb{R}^d) \ni u \mapsto \mathcal{F}(u) = \int |\nabla u|^2\,dx + \Lambda|\{u > 0\}|$$

and the boundary $\partial\{u > 0\}$ is smooth, then $|\nabla u| = \sqrt{\Lambda}$ on $\partial\{u > 0\}$. There are various ways to state this optimality for free boundaries that are not a priori smooth; see for example [Alt and Caffarelli 1981; De Silva 2011]. In the case of vector-valued functionals, we use the notion of viscosity solution from [Mazzoleni et al. 2017].

Definition 3.1. Let $\Omega \subset \mathbb{R}^d$ be an open set. We say that the continuous function $U = (u_1, \dots, u_k) : \overline{\Omega} \to \mathbb{R}^k$ is a viscosity solution of the problem

$$-\Delta U = 0 \quad \text{in } \Omega, \qquad U = 0 \quad \text{on } \partial\Omega \cap D, \qquad |\nabla|U|| = \sqrt{\Lambda} \quad \text{on } \partial\Omega \cap D$$

if for every $i = 1, \dots, k$ the component u_i is a solution of the PDE

$$-\Delta u_i = 0 \quad \text{in } \Omega, \qquad u_i = 0 \quad \text{on } \partial\Omega \cap D,$$

and the boundary condition $|\nabla |U|| = \sqrt{\Lambda}$ on $\partial\Omega \cap D$ holds in viscosity sense, that is:

- For every continuous $\phi : \mathbb{R}^d \to \mathbb{R}$, differentiable in $x_0 \in \partial\Omega \cap D$ and such that "ϕ touches $|U|$ from below in x_0" (that is, $|U| - \phi : \overline{\Omega} \to \mathbb{R}$ has a local minimum equal to zero in x_0), we have $|\nabla\phi|(x_0) \le \sqrt{\Lambda}$.
- For every function $\phi : \mathbb{R}^d \to \mathbb{R}$, differentiable in $x_0 \in \partial\Omega \cap D$ and such that "ϕ touches $|U|$ from above in x_0" (that is $|U| - \phi : \overline{\Omega} \to \mathbb{R}$ has a local maximum equal to zero in x_0), we have $|\nabla\phi|(x_0) \ge \sqrt{\Lambda}$.

Lemma 3.2. *Let U be a minimizer for* (1-1) *and $x_0 \in \mathrm{Reg}(\partial\Omega_U) \cup \mathrm{Sing}_1(\partial\Omega_U)$. Then, there is $r > 0$ such that U is a viscosity solution of*

$$-\Delta U = 0 \quad \text{in } \Omega_U \cap B_r(x_0), \qquad U = 0 \quad \text{on } \partial\Omega_U \cap B_r(x_0), \qquad |\nabla|U|| = \sqrt{\Lambda} \quad \text{on } \partial\Omega_U \cap B_r(x_0). \tag{3-1}$$

Proof. Suppose that ϕ touches $|U|$ from above in $y_0 \in B_r(x_0)$. Then $|\phi(y_0)| \ge \Lambda$ precisely as in [Mazzoleni et al. 2017, Lemma 5.2]. If ϕ touches $|U|$ from below in y_0, then every blow-up $U_0 \in \mathcal{BU}_U(y_0)$ is a 1-homogeneous global minimizer of (1-1) such that Ω_{U_0} contains the half-space $\{x : \nabla\phi(y_0) \cdot x < 0\}$. Now since the Lebesgue density of Ω_{U_0} is strictly smaller than 1, the argument of [loc. cit., Lemma 5.2] gives that all the components of U_0 must be multiples of the same global minimizer of the scalar one-phase Alt–Caffarelli problem. Thus $\Omega_{U_0} = \{x : \nabla\phi(y_0) \cdot x < 0\}$ and the conclusion follows as in [loc. cit., Lemma 5.2]. □

3B. *Reifenberg flat and NTA domains.* In this section we briefly recall the basic geometric properties of the Reifenberg flat and NTA domains. The Reifenberg flatness of $\mathrm{Reg}(\partial\Omega_U)$ follows precisely as in [Mazzoleni et al. 2017]. Then a result of [Kenig and Toro 1997] shows that it is also NTA. In the next section we will use the NTA property to prove regularity. For more details on the properties and the structure of the Reifenberg flat domains we refer to [Kenig and Toro 1997], while NTA domains were studied in [Kenig and Toro 1997; Jerison and Kenig 1982].

Definition 3.3 (Reifenberg flat domains). Let $\Omega \subset \mathbb{R}^d$ be an open set and let $0 < \delta < \frac{1}{2}$, $R > 0$. We say that Ω is a (δ, R)-Reifenberg flat domain if:

(1) For every $x \in \partial\Omega$ and every $0 < r \le R$ there is a hyperplane $H = H_{x,r}$ containing x such that

$$\mathrm{dist}_{\mathcal{H}}(B_r(x) \cap H, B_r(x) \cap \partial\Omega) < r\delta.$$

(2) For every $x \in \partial\Omega$, one of the connected components of the open set $B_R(x) \cap \{x : \mathrm{dist}(x, H_{x,R}) > 2\delta R\}$ is contained in Ω, while the other one is contained in $\mathbb{R}^d \setminus \overline{\Omega}$.

Theorem 3.4 (Reifenberg flat implies NTA, [Kenig and Toro 1997, Theorem 3.1]). *There exists a $\delta_0 > 0$ such that if $\Omega \subset \mathbb{R}^d$ is a (δ, R)-Reifenberg flat domain for $\delta < \delta_0$, then it is NTA*; *that is*, *there exist constants $M > 0$ and $r_0 > 0$* (*called NTA constants*) *such that*:

(1) Ω *satisfies the* ***corkscrew condition***; *that is*, *given* $x \in \partial\Omega$ *and* $r \in (0, r_0)$, *there exists* $x_0 \in \Omega$ *such that*

$$M^{-1}r < \operatorname{dist}(x_0, \partial\Omega) < |x - x_0| < r.$$

(2) $\mathbb{R}^d \setminus \Omega$ *satisfies the corkscrew condition.*

(3) *If* $w \in \partial\Omega$ *and* $w_1, w_2 \in B(w, r_0) \cap \Omega$, *then there is a rectifiable curve* $\gamma : [0, 1] \to \Omega$ *with* $\gamma(0) = w_1$ *and* $\gamma(1) = w_2$ *such that* $\mathcal{H}^1(\gamma([0, 1])) \le M|w_1 - w_2|$ *and*

$$\min\{\mathcal{H}^1(\gamma([0, t])), \mathcal{H}^1(\gamma([t, 1]))\} \le M \operatorname{dist}(\gamma(t), \partial\Omega) \quad \textit{for every } t \in [0, 1].$$

Remark 3.5. We note that an NTA domain $\Omega \subset \mathbb{R}^d$ is obviously connected, while its intersection with a ball is not necessarily so. This is due to the fact that an arc, contained in Ω and connecting two points inside the ball, may go out and then back in. On the other hand the NTA condition implies that the two points can be connected with an arc of length comparable to the length of the radius of the ball. Precisely, there exists a constant $M > 0$ such that the following property holds:

For every $x \in \partial\Omega$ *and every* $r > 0$, *there is exactly*
one connected component of $B_r(x) \cap \Omega$ *that intersects* $B_{r/M}(x) \cap \Omega$.

Lemma 3.6. *Let* U *be a solution of* (1-1) *and* $x_0 \in \operatorname{Reg}(\partial\Omega_U)$. *Then* Ω_U *is Reifenberg flat and NTA in a neighborhood of* x_0.

Proof. The proof follows by the same contradiction argument as in [Mazzoleni et al. 2017, Proposition 5.9]. Indeed, suppose that $\operatorname{Reg}(\partial\Omega_U) \ni x_n \to x_0$ and let $r_n \to 0$ be such that $\partial\Omega_U$ is *not* (δ, r_n) flat in $B_{r_n}(x_n)$. Let $U_n := U_{2r_n, x_n}$. Up to a subsequence U_n converges to $U_0 \in H^1(B_1; \mathbb{R}^k)$, which is a solution of (1-1) in B_1. We will prove that U_0 is of the form (2-7); then the conclusion will follow by the Hausdorff convergence of $\partial\Omega_{U_n}$ to $\partial\Omega_{U_0}$. Now, for fixed $0 < r < 1$ we have $W(U_n, 0, r) = W(U, x_n, rr_n) \to W(U_0, x, r)$ as $n \to \infty$. Let now $\varepsilon > 0$ be fixed. Since $x_0 \in \operatorname{Reg}(\partial\Omega_U)$, there is some $R > 0$ such that $W(U, x_0, R) - \Lambda\omega_d/2 \le \varepsilon/2$. By the continuity of W in x we get that for n large enough, $W(U, x_n, R) - \Lambda\omega_d/2 \le \varepsilon$ and, by the monotonicity of W, we have $W(U, x_n, rr_n) - \Lambda\omega_d/2 \le \varepsilon$. Passing to the limit in n we obtain $W(U_0, x, r) - \Lambda\omega_d/2 \le \varepsilon$. Since ε is arbitrary, we get $W(U_0, x, r) = \Lambda\omega_d/2$. Finally, Lemma 2.8 implies that U_0 is 1-homogeneous and $|B_1 \cap \Omega_{U_0}| = \omega_d/2$. Thus, U_0 is necessarily of the form (2-7), which concludes the proof. □

3C. ***Existence of a constant sign component.*** After showing in the previous section that the regular part of the free boundary is an NTA domain, we aim now to apply a boundary Harnack principle on it. It was proved in [Jerison and Kenig 1982] that in any NTA domain $\Omega \subset \mathbb{R}^d$ the boundary Harnack principle does hold; that is, if u and v are positive harmonic functions in Ω, vanishing on the boundary $\partial\Omega \cap B_r$, then

$$\frac{v}{u} \text{ is Hölder continuous on } \overline{\Omega} \cap B_r.$$

The precise statement of the boundary Harnack property for harmonic functions which we will use in Lemma 3.10 is the following [Jerison and Kenig 1982, Theorems 5.1 and 7.9].

Theorem 3.7 (boundary Harnack Principle for NTA Domains). *Let $\Omega \subset \mathbb{R}^d$ be an NTA domain and $A \subset \mathbb{R}^d$ an open set. For any compact $K \subset A$ there exists a constant $C > 0$ such that for all positive harmonic functions u, v vanishing continuously on $\partial\Omega \cap A$ we have*

$$C^{-1}\frac{v(y)}{u(y)} \le \frac{v(x)}{u(x)} \le C\frac{v(y)}{u(y)} \quad \textit{for all } x, y \in K \cap \overline{\Omega}.$$

Moreover, there exists $\beta > 0$, depending only on the NTA constants, such that the function v/u is Hölder continuous of order β in $K \cap \overline{\Omega}$. In particular, for any $y \in \partial\Omega \cap K$, the limit $\lim_{x\to y,\, x\in\Omega} v(x)/u(x)$ exists.

Remark 3.8 (boundary Harnack principle for sign-changing v). Theorem 3.7 still holds in the case when $u > 0$ on the NTA domain Ω and v is a harmonic function on Ω that may change sign. Indeed, if $v : B_1 \cap \Omega \to \mathbb{R}$ is a harmonic function that changes sign in $B_1 \cap \Omega$ and vanishes on $\partial\Omega \cap B_1$, then we consider the harmonic extensions h_+ and h_- solutions of the positive and negative parts of v:

$$\Delta h_\pm = 0 \quad \text{in } \Omega \cap B_1, \qquad h_\pm = 0 \quad \text{on } \partial\Omega \cap B_1, \qquad h_\pm = v_\pm \quad \text{on } \partial B_1 \cap \Omega.$$

Now, by Remark 3.5, each of the functions $h_\pm$ is strictly positive or vanishes identically in $\Omega \cap B_{1/M}$. Thus, the claim follows by the boundary Harnack principle for positive functions applied to h_+ and u (and h_- and u), the fact that $v = h_+ - h_-$, and a standard covering argument.

Remark 3.9. The constants C and β in the boundary Harnack principle do not change under blow-up. That is, given $x_0 = 0 \in \partial\Omega$, there is $r_0 > 0$ such that for all harmonic functions u, v, solutions of

$$\Delta u = \Delta v = 0 \quad \text{in } \Omega_r \cap B_1, \qquad u = v = 0 \quad \text{on } \partial\Omega_r \cap B_1, \qquad \Omega_r := \frac{1}{r}\Omega, \qquad 0 < r < r_0,$$

we have

$$C^{-1}\frac{v(y)}{u(y)} \le \frac{v(x)}{u(x)} \le C\frac{v(y)}{u(y)} \quad \text{for all } x, y \in B_{1/2} \cap \overline{\Omega}_r. \tag{3-2}$$

Following [Mazzoleni et al. 2017] we aim to apply the boundary Harnack principle to the components of the vector U in order to obtain that, for some $i \in \{1, \dots, k\}$, $|\nabla u_i|$ is Hölder continuous on $\partial\Omega_U$ and to apply the known regularity results for the one-phase Bernoulli problem to deduce that $\partial\Omega_U$ is $C^{1,\alpha}$. In our setting the functions u_i, $i = 1, \dots, k$, may change sign, which is a major obstruction since (3-2) can be applied only in the case when the denominator u is strictly positive. In order to overcome this issue, we first show that, at every point x_0 of the regular free boundary $\mathrm{Reg}(\partial\Omega_U)$, there is a neighborhood of x_0 and a component u_i which has constant sign in it.

Lemma 3.10. *Let $U = (u_1, \dots, u_k)$ be a solution for* (1-1). *For all $x_0 \in \mathrm{Reg}(\partial\Omega_U)$, there is $r > 0$ and $i \in \{1, \dots, k\}$ such that the component u_i has constant sign in $B_r(x_0) \cap \Omega_U$. Moreover, there is a constant $C_{\mathrm{sign}} > 0$ such that $C_{\mathrm{sign}} u_i \ge |U|$ in $B_r(x_0) \cap \Omega_U$.*

Proof. Without loss of generality $x_0 = 0$. Let $U_0 \in \mathcal{BU}_U(x_0)$ and $U_n := U_{r_n}$ be a blow-up sequence converging to U_0. By Lemma 2.11 there is a vector $\xi = (\xi_1, \dots, \xi_k) \in \mathbb{R}^k$ such that $|\xi| = \sqrt{\Lambda}$ and $U_0(x) = \xi x_d^+$ up to a rotation of $\mathbb{R}^d$. Now since $|\xi| = \sqrt{\Lambda}$, there is at least one component ξ_i such that $|\xi_i| \ge \sqrt{\Lambda/k}$. Without loss of generality we can assume that $i = 1$ and $\xi_1 \ge \sqrt{\Lambda/k}$.

Let $\Omega_n = \Omega_{U_n}$ and $U_n = (u_{n1}, \dots, u_{nk}) : \Omega_n \cap B_1 \to \mathbb{R}^k$, let u_{n1}^+ and u_{n1}^- be the positive and the negative parts of u_{n1}, and let $\tilde{u}_n^+$ and $\tilde{u}_n^-$ be the solutions of

$$\Delta \tilde{u}_n^{\pm} = 0 \quad \text{in } \Omega_n \cap B_1, \qquad \tilde{u}_{n1}^{\pm} = 0 \quad \text{on } \partial\Omega_n \cap B_1, \qquad \tilde{u}_n^{\pm} = u_{n1}^{\pm} \quad \text{on } \Omega_n \cap \partial B_1.$$

Now, notice that both u_{n1}^+ and u_{n1}^- are subharmonic on $\Omega_n \cap B_1$. Thus,

$$\tilde{u}_n^+ - \tilde{u}_n^- = u_{n1}^+ - u_{n1}^- = u_{n1}, \quad \tilde{u}_n^+ \geq u_{n1}^+ \quad \text{and} \quad \tilde{u}_n^- \geq u_{n1}^- \quad \text{in } \Omega_n \cap B_1.$$

Let M be the constant from Remark 3.5. By the fact that the blow-up limit U_0 has a positive first component, for a fixed n, in the ball $B_{1/M}$ exactly one of the following situations can happen:

(i) $\tilde{u}_n^+ > 0$ and $\tilde{u}_n^- > 0$ in $\Omega_n \cap B_{1/M}$.

(ii) $\tilde{u}_n^+ > 0$ and $\tilde{u}_n^- = 0$ in $\Omega_n \cap B_{1/M}$.

Moreover, again by Remark 3.5 we obtain that in both cases we have $\Omega_n \cap B_{1/M} = \{\tilde{u}_n^+ > 0\} \cap B_{1/M}$, while if (i) holds, then also $\Omega_n \cap B_{1/M} = \{\tilde{u}_n^- > 0\} \cap B_{1/M}$. Now, notice that in the case (ii) the first part of the claim of the lemma is trivial, so we concentrate our attention on the case (i). Let $x_M := e_d/(2M)$ and $r_M := 1/(4M)$. Recall that U_n converges uniformly to U_0 and $\partial\Omega_n$ converges to $\partial\Omega_{U_0} = \{x_d = 0\}$ in the Hausdorff distance. Then, for every $\varepsilon > 0$, there is $n_0 > 0$ such that for every $n \geq n_0$ we have

$$B_{r_M}(x_M) \subset \Omega_n, \qquad u_{n1}^+(x_M) \geq \sqrt{\frac{\Lambda}{k}}\frac{r_M}{2}, \qquad \text{and} \qquad |u_{n1}^-| \leq \varepsilon \quad \text{in } B_1.$$

Now, by the definition of $\tilde{u}_n^+$ and $\tilde{u}_n^-$ and the maximum principle (applied to $\tilde{u}_n^-$), we have

$$\tilde{u}_n^+(x_M) \geq \sqrt{\frac{\Lambda}{k}}\frac{r_M}{2} \quad \text{and} \quad \tilde{u}_n^-(x_M) \leq \varepsilon.$$

Finally, by (3-2), we obtain

$$\frac{\tilde{u}_n^-(x)}{\tilde{u}_n^+(x)} \leq C\frac{\tilde{u}_n^-(x_M)}{\tilde{u}_n^+(x_M)} \leq \varepsilon C\sqrt{\frac{\Lambda}{k}}\frac{r_M}{2} \quad \text{for every } x \in \Omega_n \cap B_{1/2}.$$

Choosing ε such that the right-hand side is smaller than 1, we get

$$u_{n1}(x) = \tilde{u}_n^+(x) - \tilde{u}_n^-(x) > 0 \quad \text{for every } x \in \Omega_n \cap B_{1/2},$$

which proves the first claim. The second part of the statement follows by the boundary Harnack principle applied to u_{n1} and every component u_{ni} for $i = 2, \dots, k$. □

3D. *The regular part of the free boundary is $C^{1,\alpha}$.* In the following lemma we show that the positive optimal component is locally a solution of a one-phase scalar free boundary problem with Hölder condition on the free boundary. The $C^{1,\alpha}$ regularity of $\mathrm{Reg}(\partial\Omega_U)$ then follows by known results on the regularity of the one-phase free boundaries; see [De Silva 2011, Theorem 1.1].

Lemma 3.11. *Let $U = (u_1, \dots, u_k)$ be a minimizer for* (1-1) *and* $0 \in \mathrm{Reg}(\partial\Omega_U)$ *and let the first component be of constant sign in a neighborhood of* 0; *that is, $u_1 > 0$ in $B_{r_0} \cap \Omega_U$. Then there is a constant $0 < c_0 \leq 1$,*

$0 < r \le r_0$, *and a Hölder continuous function* $g : B_r \cap \partial\Omega_U \to [c_0, 1]$ *such that* u_1 *is a viscosity solution to the problem*

$$-\Delta u_1 = 0 \quad \text{in } \Omega_U \cap B_r, \qquad u_1 = 0 \quad \text{on } \partial\Omega_U \cap B_r, \qquad |\nabla u_1| = g\sqrt{\Lambda} \quad \text{on } \partial\Omega_U \cap B_r.$$

Proof. First notice that, by Lemma 3.6, Ω_U is an NTA domain in a neighborhood of 0 and there exists $\beta > 0$, depending only on the NTA constants, such that for $i = 2, \dots, k$, u_i/u_1 is Hölder continuous of order β on $\overline{\Omega}_U \cap B_r$ for some $r \le r_0$. In particular, for every $x_0 \in \Omega^{(1/2)} \cap B_r$, the limit $g_i(x_0) := \lim_{\Omega \ni x \to x_0} u_i(x)/u_1(x)$, exists and $g_i : B_r \cap \partial\Omega \to \mathbb{R}$ is an β-Hölder continuous function. Then we have

$$u_i = g_i u_1 \quad \text{on } B_r \cap \overline{\Omega} \qquad \text{and} \qquad u_1 = g|U| \quad \text{on } B_r \cap \overline{\Omega}, \qquad \text{where } g := (1 + g_2^2 + \cdots + g_k^2)^{-1/2}.$$

We notice that g is a β-Hölder continuous function on $\overline{\Omega} \cap B_r$ for some $\beta > 0$ and is such that $c_0 \le g \le 1$, where $c_0 = 1/C_{\mathrm{sign}}$ and C_{sign} is the constant from Lemma 3.10. Suppose now that the function $\phi \in C^1(\mathbb{R}^d)$ is touching u_1 from below (see Definition 3.1, note that it is local) in a point $x_0 \in \partial\Omega \cap B_r$. For ρ small enough, there is a constant $C > 0$ such that

$$\frac{1}{g(x)} \ge \frac{1}{g(x_0)} - C|x - x_0|^\gamma \ge 0 \quad \text{for every } x \in \overline{\Omega} \cap B_\rho(x_0),$$

and so, setting $\psi(x) = \phi(x)(1/g(x_0) - C|x - x_0|^\gamma)$, we get that $\psi(x_0) = |U|(x_0)$ and

$$\psi(x) \le u_1(x)\left(\frac{1}{g(x_0)} - C|x - x_0|^\gamma\right) \le |U|(x) \quad \text{for every } x \in \overline{\Omega} \cap B_\rho(x_0);$$

that is, in the ball $B_\rho(x_0)$ we have that ψ touches $|U|$ from below in x_0. On the other hand, ψ is differentiable in x_0 and

$$|\nabla\psi(x_0)| = \frac{1}{g(x_0)}|\nabla\phi(x_0)|.$$

Since U is a viscosity solution of (3-1) we obtain

$$\sqrt{\Lambda} \ge |\nabla\psi(x_0)| = \frac{1}{g(x_0)}|\nabla\phi(x_0)|,$$

which gives the claim, the case when ϕ touches u_1 from below being analogous. □

3E. ***Higher regularity*: *the regular part of the free boundary is*** C^∞**.** Thanks to Lemma 3.10, we can apply the arguments of [Mazzoleni et al. 2017, Section 5] in order to obtain the C^∞ regularity of $\mathrm{Reg}(\partial\Omega_U)$, using the component of locally constant sign provided by Lemma 3.10 instead of u_1 in the boundary Harnack principle [loc. cit., Lemma 5.12]. We recall here the updated statements for the reader's sake.

In order to pass from $C^{1,\alpha}$ to C^∞ we need an improved boundary Harnack principle, as was proved in [De Silva and Savin 2015] for harmonic functions.

Theorem 3.12 (improved boundary Harnack principle). *Let* $U = (u_1, \dots, u_k)$ *be a minimizer for* (1-1), $0 \in \mathrm{Reg}(\partial\Omega_U)$, *and let the first component be of constant sign in a neighborhood of* 0, *that is*, $u_1 > 0$ *in* $B_{r_0} \cap \Omega_U$. *There exists* $R_0 < \frac{1}{2}$ *such that*, *if for* $r < \min\{R_0, r_0\}$ *we have* $\mathrm{Reg}(\partial\Omega_U) \cap B_r$ *is of class* $C^{k,\alpha}$

for $k \geq 1$, *then for all* $i = 2, \ldots, k$ *we have*

$$\frac{u_i}{u_1} \text{ is of class } C^{k,\alpha} \text{ on } \overline{\Omega}_U \cap B_r.$$

In particular, for every $x_0 \in \operatorname{Reg}(\partial\Omega_U) \cap B_r$, *the limit* $g_i(x_0) := \lim_{\Omega_U \ni x \to x_0} u_i(x)/u_1(x)$, *exists and* $g_i : B_r \cap \partial\Omega_U \to \mathbb{R}$ *is a* $C^{k,\alpha}$ *function.*

Proof. In order to show the claim, it is enough to apply [De Silva and Savin 2015, Theorem 2.4] for the case $k = 1$ and Theorem 3.1 of the same work for the case $k \geq 2$. □

At this point we are in position to prove the full regularity of $\operatorname{Reg}(\partial\Omega_U)$.

Lemma 3.13. *Let* $U = (u_1, \ldots, u_k)$ *be a minimizer for* (1-1), $0 \in \operatorname{Reg}(\partial\Omega_U)$, *and let the first component be of constant sign in a neighborhood of* 0, *that is,* $u_1 > 0$ *in* $B_{r_0} \cap \Omega_U$. *Then* $\operatorname{Reg}(\partial\Omega_U)$ *is locally a graph of a* C^∞ *function.*

Proof. The smoothness of the free boundary follows by a bootstrap argument as in [Kinderlehrer and Nirenberg 1977]. Let us assume that $\operatorname{Reg}(\partial\Omega_U)$ is locally $C^{k,\alpha}$ regular for some $k \geq 1$, the case $k = 1$ being true thanks to Section 3D. We will prove that $\operatorname{Reg}(\partial\Omega_U)$ is locally $C^{k+1,\alpha}$. By Lemma 3.11 the first component u_1 is locally a (classical) solution to the problem

$$\Delta u_1 = 0 \quad \text{in } \Omega_U, \qquad u_1 = 0 \quad \text{on } \operatorname{Reg}(\partial\Omega_U), \qquad |\nabla u_1| = g\sqrt{\Lambda} \quad \text{on } \operatorname{Reg}(\partial\Omega_U).$$

Now thanks to Theorem 3.12 and the definition of g we have that g is a $C^{k,\alpha}$ function. Now by [Kinderlehrer and Nirenberg 1977, Theorem 2] we have that $\operatorname{Reg}(\partial\Omega_U)$ is locally a graph of a $C^{k+1,\alpha}$ function, and this concludes the proof. □

4. Structure of the branching free boundary

In this section we study in more detail the set of branching points $\operatorname{Sing}_2(\partial\Omega_U)$. By the results of Section 2D we know that for $x_0 \in \partial\Omega_U$ we have

$$x_0 \in \operatorname{Sing}_2(\partial\Omega_U) \iff x_0 \in \Omega_U^{(1)} \iff \text{Every blow-up } U_0 \in \mathcal{BU}_U(x_0) \text{ is a linear function.}$$

In Section 4A we prove that the rank of U_0 depends only on x_0. Then, in Section 4B we stratify the singular set according to the rank at each point and finally, in the last subsection, we give some measure-theoretical criterion for the uniqueness of the blow-up.

4A. *Definition of the strata and decomposition of* $\operatorname{Sing}_2(\partial\Omega_U)$.

Lemma 4.1. *Let* $U = (u_1, \ldots, u_k)$ *be a solution of* (1-1) *and* $Q \in \mathcal{O}(k)$ *be an orthogonal matrix. Then* $V := QU$ *is also a solution of* (1-1) *corresponding to the boundary datum* $Q\Phi$.

Proof. It is sufficient to notice that for every $U : D \to \mathbb{R}^k$ we have $|QU| = |U|$ and $|\nabla(QU)|^2 = |\nabla U|^2$. □

Lemma 4.2. *Let* $U = (u_1, \ldots, u_k)$ *be a solution of* (1-1) *and* $x_0 \in \Omega_U^{(1)} \cap \partial\Omega_U$. *Then every blow-up* $U_0 \in \mathcal{BU}_U(x_0)$ *is a linear function given by a matrix* $A \in M^{d \times k}(\mathbb{R})$, *whose rank does not depend on* U_0 *but only on* x_0 *and* U.

Proof. Without loss of generality we may assume that $x_0 = 0$. Let $U_0 \in \mathcal{BU}_U(0)$, $U_0(x) = Ax$, be a blow-up such that rank $A = j$ for some $j \in \{1, \dots, k\}$. We claim that all the blow-ups in $\mathcal{BU}_U(x_0)$ are of rank j.

We first prove the claim in the case $j = 1$. Indeed, consider a matrix $Q \in \mathcal{O}(k)$ such that $QAx = (\nu \cdot x, 0, \dots, 0)$ for some $\nu \in \mathbb{R}^d$ and consider the vector-valued function $V = (v_1, \dots, v_k) := QU$, which is also a solution of (1-1) by Lemma 4.1. Now, since each of the components v_i is a harmonic function on the set $\{v_i \neq 0\}$, the Alt–Caffarelli–Friedman monotonicity formula (see [Alt et al. 1984]) gives that the function

$$r \mapsto \Phi(r, v_i) := \left(\frac{1}{r^2}\int_{B_r} \frac{|\nabla v_i^+|^2}{|x|^{d-2}}\,dx\right)\left(\frac{1}{r^2}\int_{B_r} \frac{|\nabla v_i^-|^2}{|x|^{d-2}}\,dx\right) = \int_{B_1} \frac{|\nabla (v_i)_r^+|^2}{|x|^{d-2}}\,dx \int_{B_1} \frac{|\nabla (v_i)_r^-|^2}{|x|^{d-2}}\,dx, \quad (4\text{-}1)$$

is increasing in r, where as usual $(v_i)_r(x) := (1/r)v_i(rx)$. Now, since for $i \in \{2, \dots, k\}$ the i-th component of the blow-up $QA \in \mathcal{BU}_V(0)$ constantly vanishes, we have $\Phi(0, v_i) := \lim_{r\to 0} \Phi(r, v_i) = 0$. In particular, the i-th component of any blow-up $V_0 \in \mathcal{BU}_V(0)$ should vanish and so, the only nonvanishing component of V_0 is the first one (recall that the blow-ups are nontrivial by the nondegeneracy of the solutions of (1-1)). Now since $\mathcal{BU}_V(0) = Q(\mathcal{BU}_U(0))$ we obtain that the rank of any blow-up $\mathcal{BU}_U(0)$ is precisely 1, which proves our claim.

Let us now suppose that $2 \le j \le k$ and that the claim holds for all $i \in \{1, \dots, j-1\}$. We will now prove the claim for j. Reasoning as above, we first find a matrix $Q \in \mathcal{O}(k)$ such that the last $k - j$ components of QA vanish; that is, $(QA)_{j+1} = \cdots = (QA)_k = 0 \in \mathbb{R}^k$. Then, we consider the vector-valued function $V = (v_1, \dots, v_k) := QU$ and notice that, for all $i = 1, \dots, k$, the function $r \mapsto \Phi(r, v_i)$ is increasing in r. As above, the strong H^1 convergence of the blow-up sequences implies that $\Phi(0, v_{j+1}) = \cdots = \Phi(0, v_k) = 0$ and that the components $j+1, \dots, k$ of any blow-up $V_0 \in \mathcal{BU}_V(0)$ do vanish identically. Thus, the rank of V_0 is at most j. On the other hand, since the claim does hold for every $i \in \{1, \dots, j-1\}$, the rank of V_0 is precisely j, which concludes the proof. □

Lemma 4.2 allows us to define, for every $j \in \{1, \dots, d\}$, the stratum

$$\mathcal{S}_j := \{x_0 \in \Omega_U^{(1)} \cap \partial\Omega_U : \text{every blow-up } U_0 \in \mathcal{BU}_U(x_0) \text{ has rank } j\}. \quad (4\text{-}2)$$

Again, by Lemma 4.2, the singular set $\partial\Omega_U \cap \Omega_U^{(1)}$ can be decomposed as a disjoint union

$$\Omega_U^{(1)} \cap \partial\Omega_U = \bigcup_{j=1}^{d} \mathcal{S}_j. \quad (4\text{-}3)$$

4B. *Dimension of the strata.* In this subsection we give an estimate on the Hausdorff dimension, $\dim_{\mathcal{H}}$ of the stratum $\mathcal{S}_j$. The proof is based on a well-known technique in geometric measure theory known as the Federer reduction principle.

Given $A \in \mathbb{R}^d$, $0 \le s < \infty$, and $0 < \delta \le \infty$, we recall the notation

$$\mathcal{H}^s_\delta(A) = \frac{\omega_s}{2^s} \inf\left\{\sum_{i=1}^{\infty} (\operatorname{diam} C_i)^s : A \subset \bigcup_{i=1}^{\infty} C_i,\ \operatorname{diam} C_i < \delta\right\}, \quad \mathcal{H}^s(A) = \sup_{\delta \ge 0} \mathcal{H}^s_\delta(A), \quad (4\text{-}4)$$
$$\dim_{\mathcal{H}} A = \inf\{s \ge 0 : \mathcal{H}^s(A) = 0\}.$$

It is well known that $\mathcal{H}^s(A)=0$ if and only if $\mathcal{H}^s_\infty(A)=0$. The other fact (for a proof we refer to [Giusti 1984, Proposition 11.3]) that we will use is

$$\limsup_{r\to 0}\frac{\mathcal{H}^s_\infty(A\cap B_r(x))}{2^{-s}\omega_s r^s}\geq 1 \quad \text{for } \mathcal{H}^s\text{-almost every } x\in A. \tag{4-5}$$

Theorem 4.3. *Let $U:\mathbb{R}^d\supset D\to\mathbb{R}^k$ be a solution of* (1-1) *and $\mathcal{S}_j$ be as in* (4-2). *If $j=d$, then $\mathcal{S}_j$ is a discrete subset of D. More precisely each point of $\mathcal{S}_d$ is isolated in $\partial\Omega_U$. If $1\leq j<d$, then $\mathcal{S}_j$ is a set of Hausdorff dimension $\dim_{\mathcal{H}}\mathcal{S}_j\leq d-j$.*

Proof. We start with the first claim. Suppose that $x_0\in\mathcal{S}_d$ and there is a sequence $\partial\Omega_U\ni x_n\to x_0$. Taking $r_n=|x_n-x_0|$, $\xi_n:=(x_n-x_0)/r_n$, $\xi_0=\lim_{n\to\infty}\xi_n$, and a blow-up limit $U_0\in\mathcal{BU}_U(x_0)$ of the sequence U_{r_n,x_0} we obtain that $\xi_0\in\partial B_1$, $U_0(\xi_0)=0$ and so $\dim\ker U_0\geq 1$, which is a contradiction with the definition of $\mathcal{S}_d$.

Let now $j<d$. Suppose by contradiction that there is $\varepsilon>0$ and a solution U of (1-1) such that $\mathcal{H}^{d-j+\varepsilon}(\mathcal{S}_j)>0$. Then, by (4-5), we get that there is a point $x_0\in\mathcal{S}_j$ such that

$$\limsup_{r\to 0}\frac{\mathcal{H}^{d-j+\varepsilon}_\infty(\partial\Omega_U\cap B_r(x_0))}{r^{d-j+\varepsilon}}\geq\limsup_{r\to 0}\frac{\mathcal{H}^{d-j+\varepsilon}_\infty(\mathcal{S}_j\cap B_r(x_0))}{r^{d-j+\varepsilon}}\geq 2^{-(d-j+\varepsilon)}\omega_{d-j+\varepsilon}. \tag{4-6}$$

Now let $r_n\to 0$ be a sequence realizing the first limsup above and $U_n=U_{r_n,x_0}$ be a blow-up sequence converging to some $U_0\in\mathcal{BU}_U(x_0)$. In particular, $\partial\Omega_{U_n}$ converges in the Hausdorff distance to $\partial\Omega_{U_0}$. Now, since $\mathcal{H}^s_\infty$ is upper semicontinuous with respect to the Hausdorff convergence of sets, (4-6) gives

$$\mathcal{H}^{d-j+\varepsilon}_\infty(\partial\Omega_{U_0}\cap B_1)\geq\lim_{n\to\infty}\mathcal{H}^{d-j+\varepsilon}_\infty(\partial\Omega_{U_n}\cap B_1)\geq 2^{-(d-j+\varepsilon)}\omega_{d-j+\varepsilon},$$

which is in contradiction with the fact that $\mathcal{H}^{d-j+\varepsilon}(\partial\Omega_{U_0}\cap B_1)=0$. □

Remark 4.4. A more refined argument in the spirit of Naber and Valtorta, essentially based on the Weiss monotonicity formula and the structure of the blow-up limits, can be used to deduce that the set $\mathcal{S}_j$ has finite $(d-j)$-dimensional Hausdorff measure. For more details on this technique in the context of the free-boundary problems considered in this paper we refer the reader to [Edelen and Engelstein 2019].

4C. ***A density criterion for the uniqueness of the blow-up limit.*** The uniqueness of the blow-up limit is a central question in free boundary problems and is strictly related to the C^1-rectifiability of the singular set. It remains a major open question even in the case of the two-phase problem corresponding to the case $k=1$. In this last subsection we give a general criterion for the uniqueness of the blow-up at the singular points, which depends only on the Lebesgue density of the positivity set Ω_U (see Proposition 4.5). Now, even if at this point this criterion by itself is not sufficient for the conclusion, it provides a proof of the fact that the regularity of the singular set only reduces to a control over the measure of the nodal set $B_r\setminus\Omega_U$. We prove the lemma by choosing a power rate of convergence, but the argument can be carried out under more general assumptions. For example, a logarithmic decay of the density still translates into a decay of the Weiss energy. This, again implies a blow-up uniqueness and a logarithmic rate of convergence; see [Engelstein et al. 2018]. In this subsection we use the notation $W(U,r):=W(U,0,r)$

and $W_0(U,r) := W_0(U,0,r)$, where

$$W_0(U,x_0,r) = \frac{1}{r^d}\int_{B_r(x_0)} |\nabla U|^2\,dx - \frac{1}{r^{d+1}}\int_{\partial B_r(x_0)} |U|^2\,d\mathcal{H}^{d-1}.$$

Proposition 4.5. *Suppose that U is a solution of* (1-1) *and $x_0 \in \partial\Omega_U$. If there are constants $C > 0$ and $\alpha > 0$ such that*

$$\frac{|B_r(x_0)\setminus\Omega_U|}{r^d} \le Cr^\alpha \quad \textit{for every } 0 < r < \operatorname{dist}(x_0,\partial D),$$

then there is a unique blow-up $U_0 \in \mathcal{BU}_U(x_0)$ and we have the estimate $\|U_{r,x_0} - U_0\|_{L^2(\partial B_1)} \le Cr^\beta$ for some $\beta = \beta(\alpha,d)$.

Proof. Let $x_0 = 0$ and $r > 0$ be fixed. Let $H : B_r \to \mathbb{R}^k$ be the harmonic extension of U in the ball B_r. A classical estimate for harmonic functions, see [Spolaor and Velichkov 2019, Lemma 2.5], states that there is a dimensional constant $\bar{\varepsilon} > 0$ such that

$$(1+\varepsilon)W_0(H,r) \le W_0(Z,r) \quad \text{for every } \varepsilon \in [0,\bar{\varepsilon}], \tag{4-7}$$

where Z is the 1-homogeneous extension of U in the ball $B_r(x_0)$. On the other hand, $|B_r \setminus \Omega_H| = 0$ and so, the optimality of U gives

$$W_0(U,r) \le W_0(H,r) + r^{-d}\Lambda|B_r\setminus\Omega_U|. \tag{4-8}$$

Finally, we notice that for every function U we have the formula

$$\frac{\partial}{\partial r}W_0(U,r) = \frac{d}{r}(W_0(Z,r) - W_0(U,r)) + \frac{1}{r^{d+2}}\sum_{i=1}^k \int_{\partial B_r} |x\cdot\nabla u_i - u_i|^2\,d\mathcal{H}^{d-1}. \tag{4-9}$$

Now, using (4-9), (4-7) and (4-8), we have

$$\begin{aligned}
\frac{\partial}{\partial r}W_0(U,r) &\ge \frac{d}{r}(W_0(Z,r) - W_0(U,r)) \ge \frac{d}{r}(W_0(Z,r) - W_0(H,r) - r^{-d}\Lambda|B_r\setminus\Omega_U|) \\
&\ge \frac{d}{r}(\varepsilon W_0(H,r) - r^{-d}\Lambda|B_r\setminus\Omega_U|) \ge \frac{d}{r}(\varepsilon W_0(U,r) - (1+\varepsilon)r^{-d}\Lambda|B_r\setminus\Omega_U|) \\
&\ge \frac{d\varepsilon}{r}W_0(U,r) - 2d\Lambda C r^{\alpha-1}.
\end{aligned}$$

In particular, this implies that the function

$$r \mapsto \frac{W_0(U,r)}{r^{\varepsilon d}} + \frac{2d\Lambda C}{\alpha - d\varepsilon} r^{\alpha - d\varepsilon}$$

is increasing in r and so, choosing $\varepsilon = \alpha/(2d)$, we get that there is a constant C_{U,x_0} depending on U and the point $x_0 = 0 \in D$ such that

$$W_0(U,r) \le C_{U,x_0} r^{\alpha/2} \quad \text{and} \quad W(U,r) - \Lambda\omega_d = W_0(U,r) - \Lambda\frac{|B_r\setminus\Omega_U|}{r^d} \le W_0(U,r) \le C_{U,x_0} r^{\alpha/2}.$$

Now, the uniqueness of the blow-up and the convergence rate follow by a standard argument; see [Spolaor and Velichkov 2019]. □

References

[Alt and Caffarelli 1981] H. W. Alt and L. A. Caffarelli, "Existence and regularity for a minimum problem with free boundary", *J. Reine Angew. Math.* **325** (1981), 105–144. MR Zbl

[Alt et al. 1984] H. W. Alt, L. A. Caffarelli, and A. Friedman, "Variational problems with two phases and their free boundaries", *Trans. Amer. Math. Soc.* **282**:2 (1984), 431–461. MR Zbl

[Bucur 2012] D. Bucur, "Minimization of the k-th eigenvalue of the Dirichlet Laplacian", *Arch. Ration. Mech. Anal.* **206**:3 (2012), 1073–1083. MR Zbl

[Bucur et al. 2015] D. Bucur, D. Mazzoleni, A. Pratelli, and B. Velichkov, "Lipschitz regularity of the eigenfunctions on optimal domains", *Arch. Ration. Mech. Anal.* **216**:1 (2015), 117–151. MR Zbl

[Caffarelli and Salsa 2005] L. Caffarelli and S. Salsa, *A geometric approach to free boundary problems*, Graduate Studies in Math. **68**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[Caffarelli et al. 2004] L. A. Caffarelli, D. Jerison, and C. E. Kenig, "Global energy minimizers for free boundary problems and full regularity in three dimensions", pp. 83–97 in *Noncompact problems at the intersection of geometry, analysis, and topology* (New Brunswick, NJ, 2001), edited by A. Bahri et al., Contemp. Math. **350**, Amer. Math. Soc., Providence, RI, 2004. MR Zbl

[Caffarelli et al. 2018] L. A. Caffarelli, H. Shahgholian, and K. Yeressian, "A minimization problem with free boundary related to a cooperative system", *Duke Math. J.* **167**:10 (2018), 1825–1882. MR Zbl

[De Philippis et al. 2019] G. De Philippis, L. Spolaor, and B. Velichkov, "Regularity of the free boundary for the two-phase Bernoulli problem", preprint, 2019. arXiv

[De Silva 2011] D. De Silva, "Free boundary regularity for a problem with right hand side", *Interfaces Free Bound.* **13**:2 (2011), 223–238. MR Zbl

[De Silva and Jerison 2009] D. De Silva and D. Jerison, "A singular energy minimizing free boundary", *J. Reine Angew. Math.* **635** (2009), 1–21. MR Zbl

[De Silva and Savin 2015] D. De Silva and O. Savin, "A note on higher regularity boundary Harnack inequality", *Discrete Contin. Dyn. Syst.* **35**:12 (2015), 6155–6163. MR Zbl

[Edelen and Engelstein 2019] N. Edelen and M. Engelstein, "Quantitative stratification for some free-boundary problems", *Trans. Amer. Math. Soc.* **371**:3 (2019), 2043–2072. MR Zbl

[Engelstein et al. 2018] M. Engelstein, L. Spolaor, and B. Velichkov, "Uniqueness of the blow-up at isolated singularities for the Alt–Caffarelli functional", preprint, 2018. arXiv

[Giusti 1984] E. Giusti, *Minimal surfaces and functions of bounded variation*, Monographs in Math. **80**, Birkhäuser, Basel, 1984. MR Zbl

[Jerison and Kenig 1982] D. S. Jerison and C. E. Kenig, "Boundary behavior of harmonic functions in nontangentially accessible domains", *Adv. Math.* **46**:1 (1982), 80–147. MR Zbl

[Jerison and Savin 2015] D. Jerison and O. Savin, "Some remarks on stability of cones for the one-phase free boundary problem", *Geom. Funct. Anal.* **25**:4 (2015), 1240–1257. MR Zbl

[Kenig and Toro 1997] C. E. Kenig and T. Toro, "Harmonic measure on locally flat domains", *Duke Math. J.* **87**:3 (1997), 509–551. 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

[Kriventsov and Lin 2018] D. Kriventsov and F. Lin, "Regularity for shape optimizers: the nondegenerate case", *Comm. Pure Appl. Math.* **71**:8 (2018), 1535–1596. MR Zbl

[Kriventsov and Lin 2019] D. Kriventsov and F. Lin, "Regularity for shape optimizers: the degenerate case", *Comm. Pure Appl. Math.* **72**:8 (2019), 1678–1721. Zbl

[Mazzoleni et al. 2017] D. Mazzoleni, S. Terracini, and B. Velichkov, "Regularity of the optimal sets for some spectral functionals", *Geom. Funct. Anal.* **27**:2 (2017), 373–426. MR Zbl

[Naber and Valtorta 2017] A. Naber and D. Valtorta, "Rectifiable-Reifenberg and the regularity of stationary and minimizing harmonic maps", *Ann. of Math.* (2) **185**:1 (2017), 131–227. MR Zbl

[Naber and Valtorta 2018] A. Naber and D. Valtorta, "Stratification for the singular set of approximate harmonic maps", *Math. Z.* **290**:3-4 (2018), 1415–1455. MR Zbl

[Spolaor and Velichkov 2019] L. Spolaor and B. Velichkov, "An epiperimetric inequality for the regularity of some free boundary problems: the 2-dimensional case", *Comm. Pure Appl. Math.* **72**:2 (2019), 375–421. MR Zbl

[Weiss 1999] G. S. Weiss, "Partial regularity for a minimum problem with free boundary", *J. Geom. Anal.* **9**:2 (1999), 317–326. MR Zbl

Received 27 Apr 2018. Revised 6 Feb 2019. Accepted 3 Apr 2019.

DARIO MAZZOLENI: dariocesare.mazzoleni@unicatt.it
Dipartimento di Matematica e Fisica "N. Tartaglia", Università Cattolica, Brescia, Italy

SUSANNA TERRACINI: susanna.terracini@unito.it
Dipartimento di Matematica "Giuseppe Peano", Università di Torino, Torino, Italy

BOZHIDAR VELICHKOV: bozhidar.velichkov@unina.it
Dipartimento di Matematica e Applicazioni "Renato Caccioppoli", Università degli Studi di Napoli Federico II, Napoli, Italy

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 13, No. 3, 2020
dx.doi.org/10.2140/apde.2020.13.765

ON THE DISCRETE FUGLEDE AND POMPEIU PROBLEMS

GERGELY KISS, ROMANOS DIOGENES MALIKIOSIS, GÁBOR SOMLAI AND MÁTÉ VIZER

We investigate the discrete Fuglede conjecture and the Pompeiu problem on finite abelian groups and develop a strong connection between the two problems. We give a geometric condition under which a multiset of a finite abelian group has the discrete Pompeiu property. Using this description and the revealed connection we prove that Fuglede's conjecture holds for $\mathbb{Z}_{p^n q^2}$, where p and q are different primes. In particular, we show that every spectral subset of $\mathbb{Z}_{p^n q^2}$ tiles the group. Further, using our combinatorial methods we give a simple proof for the statement that Fuglede's conjecture holds for $\mathbb{Z}_p^2$.

1. Introduction

In this article we deal with the discrete version of Fuglede's conjecture and the Pompeiu problem; both originated in analysis. We build a relationship between them that helps us to provide new results for Fuglede's conjecture in the discrete setting.

The following question was asked by Pompeiu [1929]. Take a continuous function f on the plane whose integral is zero on every unit disc. Does it follow that f is constantly zero? The answer to this question is no, but it initiated several different types of investigations in various settings, and in some cases the answer is affirmative for an analogous question. We give an implicit characterization of the non-Pompeiu sets for finite abelian groups.

Fuglede [1974] conjectured that a bounded domain $S \subset \mathbb{R}^d$ tiles the d-dimensional Euclidean space if and only if the set of $L^2(S)$ functions admits an orthogonal basis of exponential functions. This conjecture was disproved by Tao [2004].

A discrete version of Fuglede's conjecture might be formulated in the following way. A subset S of a finite abelian group G tiles G if and only if the character table of G has a submatrix, whose rows are indexed by the elements of S, which is a complex Hadamard matrix. This version of Fuglede's conjecture is not only interesting on its own but also plays a crucial role in the above-mentioned counterexample of Tao. Actually his counterexample (in $\mathbb{R}^5$) is based on a counterexample for elementary abelian p-groups of finite rank.

Fuglede's conjecture is especially interesting for finite cyclic groups, since, e.g., every tiling of $\mathbb{Z}$ is periodic, so it goes back to a tiling of a finite cyclic group. However, not much is known for cyclic groups. A recent paper of the second author and Kolountzakis [Malikiosis and Kolountzakis 2017] shows that Fuglede's conjecture holds for any cyclic group of order $p^n q$, where p and q are different primes.

MSC2010: primary 39B32, 43A46; secondary 13F20, 20K01.
Keywords: Pompeiu problem, spectral set conjecture, tiling, vanishing sums of roots of unity.

Our main contribution towards Fuglede's conjecture for cyclic groups is to connect this problem with the Pompeiu problem, introduce more combinatorial ideas and verify it for yet unknown cases: cyclic groups of order $p^n q^2$, $n \geq 1$ (see Theorem 2.5).

Further using our techniques we give a neat and combinatorial proof for the previously known fact, proved by Iosevich, Mayeli and Pakianathan [Iosevich et al. 2017], that Fuglede's conjecture holds for $\mathbb{Z}_p^2$ (see also Theorem A.1).

Structure of the paper. Section 2 is devoted to a detailed introduction to Fuglede's conjecture and the Pompeiu problem, introducing also the discrete versions of them. Further we establish a connection between the two problems. In Section 3 we give some sort of solution for the Pompeiu problem for abelian groups that we apply later in Section 6. Sections 4 and 5 are preparations for the proof of Theorem 2.5. In Section 4 we reduce the cases to a special one partly based on our results concerning the Pompeiu problem. In Section 5 we prove some technical lemmas, which we use later. Section 6 is devoted to the proof of Theorem 2.5. Finally, in the Appendix we give an alternative proof of Theorem A.1.

2. Fuglede and Pompeiu problems

Fuglede's spectral set conjecture. The original conjecture of Fuglede [1974] was formulated as follows. Let Ω be a measurable subset of $\mathbb{R}^n$ of positive Lebesgue measure. A set $\Omega \subseteq \mathbb{R}^n$ is called *spectral* if there is a set $\Lambda \subset \mathbb{R}^n$ such that $\{e^{i\lambda \cdot x} : \lambda \in \Lambda,\ x \in \Omega\}$ is an orthogonal basis of $L^2(\Omega)$. Then $\Lambda \subseteq \mathbb{R}^n$ is called the *spectrum* of Ω.

We say that S is a *tile* of $\mathbb{R}^n$ if there is a set $T \subset \mathbb{R}^n$ such that almost every point of $\mathbb{R}^n$ can be uniquely written as $s + t$, where $s \in S$ and $t \in T$. In this case, we say that T is the *tiling complement of S*.

Fuglede's spectral set conjecture [1974] (which we just call Fuglede's conjecture) states the following:

Conjecture 1. Ω is spectral if and only if Ω is a tile.

The conjecture was proved by Fuglede [1974] in the special case when the tiling complement or the spectrum is a lattice in $\mathbb{R}^n$. Also it has been verified by Fuglede that the L^2-space over a triangle or a disc does not admit an orthogonal basis of exponentials. (The proof for the disc was corrected by Iosevich, Katz and Pedersen [Iosevich et al. 1999].) The conjecture was further verified in some other cases; see, e.g., [Iosevich et al. 2003; Łaba 2001].

Tao [2004] disproved the Spectral $\Longrightarrow$ Tiling direction of the conjecture by constructing a spectral set in $\mathbb{R}^5$ that does not tile the 5-dimensional space. As an extension of Tao's work, Matolcsi [2005] proved that (the same direction of) the conjecture fails in dimension 4 as well. Further, Kolountzakis and Matolcsi [2006a; 2006b] and Farkas, Matolcsi and Móra [Farkas et al. 2006] provided counterexamples in dimension 3 for each direction of the conjecture.

Discrete abelian groups. Fuglede's conjecture can be naturally stated for other groups, for example $\mathbb{Z}$. These cases are not only interesting on their own, but they also have connection with the original case, since, e.g., in his disproof of the 5-dimensional case, Tao constructed a spectral set in $\mathbb{Z}_3^5$ (containing six elements, hence not a tile, as the cardinality of any tile of a finite abelian group divides the order of the group), then he lifted this counterexample to $\mathbb{R}^5$. A similar strategy was carried out by Kolountzakis and Matolcsi [2006b] in the disproof of the other direction of the original conjecture. We also mention some examples where Fuglede's conjecture holds. These include finite cyclic p-groups [Łaba 2002], $\mathbb{Z}_p \times \mathbb{Z}_p$ [Iosevich et al. 2017], and $\mathbb{Q}_p$ [Fan et al. 2019], the field of p-adic numbers.

Borrowing the notation from [Dutkay and Lai 2014; Malikiosis and Kolountzakis 2017], we write $\boldsymbol{S}-\boldsymbol{T}(G)$ (resp. $\boldsymbol{T}-\boldsymbol{S}(G)$), if the Spectral $\Longrightarrow$ Tiling (resp. Tiling $\Longrightarrow$ Spectral) direction of Fuglede's conjecture holds in G for every bounded subset. The above-mentioned connection between the conjecture on $\mathbb{R}$, on $\mathbb{Z}$ and on finite cyclic groups is summarized below [Dutkay and Lai 2014], where $\boldsymbol{T}-\boldsymbol{S}(\mathbb{Z}_{\mathbb{N}})$ means that $\boldsymbol{T}-\boldsymbol{S}(\mathbb{Z}_n)$ holds for every $n \in \mathbb{N}$:

$$\begin{array}{ccccc} \boldsymbol{T}-\boldsymbol{S}(\mathbb{R}) & \Longleftrightarrow & \boldsymbol{T}-\boldsymbol{S}(\mathbb{Z}) & \Longleftrightarrow & \boldsymbol{T}-\boldsymbol{S}(\mathbb{Z}_{\mathbb{N}}), \\ \boldsymbol{S}-\boldsymbol{T}(\mathbb{R}) & \Longrightarrow & \boldsymbol{S}-\boldsymbol{T}(\mathbb{Z}) & \Longrightarrow & \boldsymbol{S}-\boldsymbol{T}(\mathbb{Z}_{\mathbb{N}}). \end{array}$$

According to this, a counterexample to the Spectral $\Longrightarrow$ Tiling direction in a finite cyclic group can be lifted to a counterexample in $\mathbb{R}$; on the other hand, if the same direction of the conjecture were true for every cyclic group or even in $\mathbb{Z}$, this would hold no meaning for the original conjecture in $\mathbb{R}$.

Concerning tiles in discrete groups it was proved in [Newman 1977] that if S is a finite set, which tiles $\mathbb{Z}$ with tiling complement T, then T is periodic; i.e., $T+N=T$ for some $N \in \mathbb{Z}$. This shows that every tiling of the integers reduces to a tiling of a cyclic group $\mathbb{Z}_N$ for some $N \in \mathbb{N}$.

We also mention a related result of [Rédei 1965]. We say that $A_1+\cdots+A_k$ is a *factorization* of the abelian group G if every element of G can uniquely be written as the sum of one element from each A_i.

Theorem 2.1 [Rédei 1965]. *Let $G=A_1+A_2+\cdots+A_n$ be a factorization of an abelian group G, where each A_i contains 0 and is of prime cardinality. Then at least one of the sets A_i is a subgroup of G.*

Cyclic groups. Surprisingly, despite their previously described role in the discrete version of Fuglede's conjecture, not much is known for cyclic groups. A recent result of [Malikiosis and Kolountzakis 2017] proved Conjecture 2 (see later) for $\mathbb{Z}_{p^n q}$. They also wrote that most likely, their result might be extended to cyclic groups of order having two different prime divisors but they haven't succeeded yet.

As we will mainly deal with cyclic groups, let us state the conjecture again in this setting. First let us define spectral sets and tiles in cyclic groups also.

Definition 2.2. For a set $S \subset \mathbb{Z}_N$, we say that S is *spectral* if $L^2(S)$ has an orthogonal basis of exponentials (indexed by Λ). This is equivalent to the following two conditions holding:

(1) There is $\Lambda \subset \mathbb{Z}_N$ such that any $f : S \to \mathbb{C}$ can be written as the $\mathbb{C}$-linear combination of exponentials of the form
$$\xi_N^{\lambda \cdot x}, \quad \lambda \in \Lambda,$$
where the product $\lambda \cdot x$ is taken modulo N and $\xi_N = e^{2\pi i/N}$.

(2) For any two different $\lambda, \lambda' \in \Lambda$ we have
$$\sum_{x \in S} \xi_N^{(\lambda-\lambda')\cdot x} = 0$$
(i.e., the representations $\chi_\lambda(x)=\xi_N^{\lambda \cdot x}$ and $\chi_{\lambda'}(x)=\xi_N^{\lambda' \cdot x}$ are orthogonal).

We denote $\{\chi_\lambda : \lambda \in \Lambda\}$ by χ_Λ.

Remark 2.3. We note that if S is a spectral set, then $|S| = |\Lambda|$ follows from Definition 2.2. Condition (2) further implies

$$\Lambda - \Lambda \subseteq \{0\} \cup \{x \in \mathbb{Z}_N : \hat{1}_S(x) = 0\}, \tag{1}$$

where 1_S is the characteristic function of S, and $\hat{f}(x) = \sum_{y\in\mathbb{Z}_N} f(y)\xi_N^{-x\cdot y}$ is the discrete Fourier transform of $f : G \mapsto \mathbb{C}$, as usual.

Definition 2.4. Let G be a discrete abelian group. We say that $S \subset G$ *tiles* G if there exists $T \subset G$ such that $S + T = G$, where $S + T$ is the set of elements of G of the form $s + t$, $s \in S$, $t \in T$, counted with multiplicity, so we have each $g \in G$ exactly once. In this case we say that T is a *tiling complement* of S in G.

For cyclic groups Fuglede's conjecture can be stated as follows.

Conjecture 2. For any N and $S \subset \mathbb{Z}_N$ we have that S is spectral if and only if S tiles $\mathbb{Z}_N$.

There has been some recent progress on this conjecture over the last few years. The known results for the Tiling $\Longrightarrow$ Spectral direction follow from [Coven and Meyerowitz 1999; Łaba 2002]. Coven and Meyerowitz proved that if a finite subset A of the integers satisfies two conditions (T1) and (T2) (to be defined below) then it tiles $\mathbb{Z}$ by translations. The inverse holds when the cardinality of A is divisible by at most two primes (corollary to [Coven and Meyerowitz 1999, Theorem B2]). Łaba then connected these properties with Fuglede's conjecture on $\mathbb{Z}$, proving that if A satisfies (T1) and (T2), then it has a spectrum [Łaba 2002, Theorem 1.5(i)]; therefore, if A tiles $\mathbb{Z}$ and its cardinality is divisible by at most two distinct primes, then it must be spectral [Łaba 2002, Corollary 1.6(i)]. The passage to cyclic groups of order $p^n q^m$ is easily done through [Coven and Meyerowitz 1999, Lemma 2.3], which implies that if A tiles $\mathbb{Z}$ and $|A| = p^a q^b$, then there is a (possibly different) tiling of $\mathbb{Z}$ by translates of A with period $N = p^n q^m$ for some $m, n \in \mathbb{N}$. In other words, A could be considered as a subset of $\mathbb{Z}_N$, and the Tiling $\Longrightarrow$ Spectral direction in $\mathbb{Z}_N$ follows verbatim using the above results. For a proof containing all the above arguments strictly in the setting of cyclic groups, we refer the reader to [Malikiosis and Kolountzakis 2017, Section 3].

Concerning the case for N square-free, the Tiling $\Longrightarrow$ Spectral direction in $\mathbb{Z}_N$ follows easily from the fact that any tile of $\mathbb{Z}_N$ is a set of coset representatives of a subgroup of $\mathbb{Z}_N$, again from combined arguments from [Coven and Meyerowitz 1999; Łaba 2002]. This fact was posed as a problem in Tao's blog,[1] which was subsequently solved by Łaba and Meyerowitz (see the comments of that post). Their arguments were based on [Coven and Meyerowitz 1999, Lemma 2.3], which implies that a tile A in $\mathbb{Z}_N$ with N square-free, accepts the subgroup $M\mathbb{Z}_N$ as a tiling complement, where $M = |A|$. This is one instance where the properties (T1) and (T2) hold trivially; thus A is also spectral due to [Łaba 2002, Theorem 1.5(i)]. For a self-contained proof of the Tiling $\Longrightarrow$ Spectral direction in the setting of cyclic groups of square-free order, which is along the same lines, we refer the reader to [Shi 2019].

[1] https://terrytao.wordpress.com/2011/11/19/some-notes-on-the-coven-meyerowitz-conjecture/

The reverse direction, Spectral $\Longrightarrow$ Tiling, is considerably harder, and the best results to this date are the proofs for $N = p^n q$ [Malikiosis and Kolountzakis 2017] and $N = pqr$ [Shi 2019], where p, q, r are distinct primes. The main tool that is introduced in the Spectral $\Longrightarrow$ Tiling direction is the structure of the vanishing sums of roots of unity [Lam and Leung 2000].

In this paper we verify Conjecture 2 for cyclic groups of order $p^n q^2$ by proving the following.

Theorem 2.5. *Let p and q be two different primes. Then we have $\boldsymbol{S} - \boldsymbol{T}$ $(\mathbb{Z}_{p^n q^2})$ for every $n \geq 1$.*

As stated above, $\boldsymbol{T} - \boldsymbol{S}$ $(\mathbb{Z}_{p^n q^m})$ has already been proven [Coven and Meyerowitz 1999; Łaba 2002]. Combining this result and Theorem 2.5 we obtain:

Theorem 2.6. *Let p and q be two different primes. Then Fuglede's conjecture holds for $\mathbb{Z}_{p^n q^2}$, $n \geq 1$.*

Furthermore, using our method, we give in the Appendix a simple proof of the theorem of Iosevich, Mayeli and Pakianathan [Iosevich et al. 2017], stating that Fuglede's conjecture holds for $\mathbb{Z}_p^2$.

Pompeiu problem. The problem goes back to the seminal paper [Pompeiu 1929], where he asked the following question of integral geometry:

Question 1. Let K be a compact set of positive Lebesgue measure. Is it true that if $f : \mathbb{R}^2 \to \mathbb{C}$ is a continuous function that satisfies

$$\int_{\sigma(K)} f(x, y)\, d\lambda_x\, d\lambda_y = 0 \tag{2}$$

for every rigid motion σ (here λ denotes the Lebesgue measure), then f is identically zero (i.e, $f \equiv 0$)?

If K is the closed disc of radius $r > 0$, then the answer is negative. It was shown in [Chakalov 1944], see also [Garofalo 1989], that (2) holds if $f(x, y) = \sin(a(x + iy))$, where $a > 0$, and $J_1(ra) = 0$, where J_λ denotes the Bessel function of order λ. On the other hand, for every nonempty polygon (moreover, for any convex domain with at least one corner) the answer for Question 1 is affirmative by a result of Brown, Schreiber, and Taylor [Brown et al. 1973]. Recently, Ramm [2017] showed that there exists a function $f \not\equiv 0$ that satisfies the 3-dimensional analogue of (2) for a bounded domain $K \subseteq \mathbb{R}^3$ with $\mathcal{C}^1$-smooth boundary if and only if K is a closed ball. Extensive literature is concerned with the Pompeiu problem. For the history of the problem see [Ramm 1997] and the bibliographical survey [Zalcman 1992].

Here we investigate the discrete version of the *Pompeiu problem* on finite abelian groups. We note that it was studied on infinite abelian groups in [Kiss et al. 2018; Puls 2013; Zeilberger 1978].

The discrete version of Pompeiu problem for an abelian group G. In the sequel we denote the binary operation acting on an abelian group G by $+$ (as the usual addition).

Definition 2.7. Let G be an abelian group:

- Let S be a nonempty finite subset of G. We say that S *has the discrete Pompeiu property* (or S is *Pompeiu*) if whenever $f : G \to \mathbb{C}$ satisfies

$$\sum_{s \in S} f(s + x) = 0 \quad \text{for every } x \in G, \tag{3}$$

  then $f \equiv 0$.

We say that S is a *non-Pompeiu set with respect to* f if $f \not\equiv 0$ and satisfies (3).

One can define the discrete Pompeiu property for multisets similarly.

- We call $w : G \to \mathbb{Q}$ a *weight function*[2] defined on G. We say that w *is a Pompeiu weight function* if for any $f : G \to \mathbb{C}$
$$\sum_{g \in G} w(g) f(g+x) = 0 \quad \text{for every } x \in G \tag{4}$$
implies that $f \equiv 0$.

  We say that w is a *non-Pompeiu weight function with respect to* f if $f \not\equiv 0$ and satisfies (4).

Note that S is a Pompeiu set if and only if its characteristic function is a Pompeiu weight function.

Remark 2.8. We can extend the previous definition for arbitrary finite group $(G, \cdot)$ and weight function w as follows.

Let $w : G \to \mathbb{Q}$. We denote by $\mathrm{Cay}(G, w)$ the *Cayley graph* of G with respect to w. The vertex set of $\mathrm{Cay}(G, w)$ is G and g is connected to h by an edge with weight $w(g^{-1}h)$ for every $g, h \in G$. We denote by A_w the adjacency matrix of $\mathrm{Cay}(G, w)$. Using the adjacency matrix A_w of $\mathrm{Cay}(G, w)$ we may also say w is a Pompeiu weight function if and only if $A_w f = 0$ implies $f \equiv 0$. The equation $A_w f = 0$ implies that if $f \not\equiv 0$, then f is an eigenvector of A_w with eigenvalue 0. So w is a Pompeiu weight function if and only if 0 is not an eigenvalue of A_w. In the finite case this is equivalent to A_w being invertible.

We note that if G is a cyclic group, then A_w is a circulant matrix.

The set of irreducible representations of a finite abelian group G will be denoted by $\widetilde{G}$. Every irreducible representation of an abelian group is 1-dimensional (a character). Thus $\widetilde{G}$ is a group which is isomorphic to G. Note that $\widetilde{G}$ is usually called the *dual group* of G.

It is well known [Steinberg 2012] that the set of irreducible representations forms an orthogonal basis of $L^2(G)$ with respect to the natural scalar product
$$[\psi, \chi] := \sum_{g \in G} \psi(g)\overline{\chi(g)}$$
for $\psi, \chi \in \widetilde{G}$. Thus every function $f : G \to \mathbb{C}$ can be uniquely written as
$$f(x) = \sum_{\chi \in \widetilde{G}} c_\chi \chi(x) \quad \text{for all } x \in G, \tag{5}$$
for some $c_\chi \in \mathbb{C}$.

The following proposition can be deduced from [Székelyhidi 2001]. In order to make our paper self-contained, we provide the proof.

Proposition 2.9. *If w is a non-Pompeiu weight function with respect to a function f, then w is a non-Pompeiu weight function with respect to all irreducible representations χ which have nonzero coefficient c_χ in* (5).

[2] Every weight function is a rational constant multiple of a weight function with integer coefficients. The Pompeiu property is invariant under scalar multiplication, and thus we may restrict our attention to those weight functions which take their values in $\mathbb{Z}$.

Proof. Let w be non-Pompeiu with respect to a function f; then $\sum_{s\in G} w(s) f(s+x) = 0$ for every $x \in G$. Using (5) we get

$$0 = \sum_{s\in S} w(s) \sum_{\chi\in\widetilde{G}} c_\chi \chi(s+x) = \sum_{\chi\in\widetilde{G}} c_\chi \sum_{s\in S} w(s)\chi(s+x) = \sum_{\chi\in\widetilde{G}} \Big(c_\chi \sum_{s\in S} w(s)\chi(s)\Big)\chi(x),$$

since χ is a character. This statement holds for every $x \in G$ so we can formulate it as

$$\sum_{\chi\in\widetilde{G}} \Big(c_\chi \sum_{s\in S} w(s)\chi(s)\Big)\chi = 0.$$

Since the irreducible representations are linearly independent over $\mathbb{C}$, the previous equation holds if and only if $\sum_{s\in S} w(s)\chi(s) = 0$ for all χ such that $c_\chi \neq 0$. Multiplying with $\chi(x)$ we obtain $\sum_{s\in S} w(s)\chi(x+s) = 0$. Since this holds for every $x \in G$, this means that w is non-Pompeiu with respect to such χ. □

We note that a stronger result was proved by Babai [1979], who determined the spectrum of Cayley graphs of abelian groups. The set of the eigenvalues of Cay(G, S) is $\{\sum_{s\in S} \chi(s) : \chi \in \widetilde{G}\}$.

Corollary 2.10. *If S is a non-Pompeiu set in a finite abelian group, then S is non-Pompeiu with respect to some irreducible representation of G.*

Remark 2.11. Since the characters (irreducible representations) play the role of exponential functions over the abelian group G, it seems reasonable that the function $\sin(ax)$ can provide an example on the disk for the original Pompieu problem. On the other hand, it is surprising that exponential solutions were not found in the literature.

Connection of the problems.

Proposition 2.12. *Let G be a finite abelian group. If $S \subset G$ is a spectral set with $|S| \geq 2$, then S is a non-Pompeiu set.*

Proof. The spectral property of S requires a set of irreducible representations, of the same cardinality of S, whose restrictions to S are pairwise orthogonal. Assume χ and ψ are different irreducible representations of G, whose restrictions to S are orthogonal. Since $[\chi_{|S}, \psi_{|S}] = [(\chi\bar{\psi})_{|S}, \mathrm{id}_{|S}]$ we obtain a representation $\rho = \chi\bar{\psi}$ such that $\sum_{s\in S} \rho(s) = 0$, which leads us back to the Pompeiu problem. Thus we get that S is a non-Pompeiu set with respect to the irreducible representation ρ. □

3. Pompeiu problem for cyclic groups

In this section we consider the non-Pompeiu sets for abelian groups.

Every representation of a finite abelian group is linear, so it factorizes through a faithful representation of a cyclic group since the finite subgroups of $\mathbb{C} \setminus \{0\}$ are cyclic. This shows that some sort of description for non-Pompeiu sets of finite abelian groups is given by understanding the non-Pompeiu weight functions of cyclic groups with respect to faithful representations.

Let $(\mathbb{Z}_N, +)$ be the cyclic group of order N. Note that for all $k \mid N$ there is a unique normal subgroup $\mathbb{Z}_k \leq \mathbb{Z}_N$ of order k. The group generated this way contains exactly the elements of $\mathbb{Z}_N$ divisible by N/k so this subgroup of $(\mathbb{Z}_N, +)$ will also be denoted by $H_{N/k}$.

We use the following isomorphism between $\mathbb{Z}_N$ and $\widetilde{\mathbb{Z}}_N$: Fix a primitive N-th root of unity α and a generator g of $\mathbb{Z}_N$. Then for any $j \in \mathbb{Z}_N$ the function $\psi_j(g^i) = \alpha^{ji}$ gives a homomorphism from $\mathbb{Z}_N$ to $\mathbb{C}^*$; hence it is an irreducible representation. Now $j \to \psi_j$ gives the isomorphism from $\mathbb{Z}_N$ to $\widetilde{\mathbb{Z}}_N$; throughout the text, we will use the isomorphism that arises from $\alpha = \xi_N$. From now on the subgroup of $\widetilde{\mathbb{Z}}_N$ isomorphic to $H \leq \mathbb{Z}_N$ will be denoted by $\widetilde{H}$.

Hereinafter we use the notion of mask polynomial.

Definition 3.1. Let G be a cyclic group and $w : G \to \mathbb{Q}$ be a weight function. We call

$$m_w(x) = \sum_{h \in G} w(h) x^h$$

the *mask polynomial* of w, where $w(h)$ denotes the weight of $h \in G$. This might be considered as an element of $\mathbb{Q}[x]/(x^n - 1)$. For a (multi-)set S of G we define the *mask polynomial* of S by

$$S(x) = \sum_{s \in S} c_s x^s,$$

where c_s denotes the cardinality of $s \in S$.

Let $\Phi_k(x)$ denote the k-th cyclotomic polynomial, which is of degree $\varphi(k)$, where φ denotes the Euler totient function. Note that for fixed N and prime $p \mid N$ the mask polynomial of $\mathbb{Z}_p \leq \mathbb{Z}_N$ is $\Phi_p(x^{N/p})$. The following is one of the key preliminary observations. Basically, this can be considered as a statement on roots of unity. There is a vast literature on vanishing sums of roots of unity. This particular statement gives a generalization of Theorem 3.3 of [Lam and Leung 2000]. Similar results might appear in other papers.

Proposition 3.2. *Let G be a cyclic group of order N and let α be a primitive N-th root of unity. We denote by P_N the set of prime divisors of N. Further let w be a weighted function. Then w is non-Pompeiu with respect to the faithful representation ψ_α if and only if*

$$w = \sum_{g \in G} \sum_{p \in P_N} w_{p,g} 1_{\mathbb{Z}_p + g}$$

for some $w_{p,g} \in \mathbb{Q}$, where $1_{\mathbb{Z}_p+g}$ denotes the characteristic function of the coset $\mathbb{Z}_p + g$.

Proof. The fact that w is a non-Pompeiu weight function with respect to the faithful representation ψ_α means that α is the root of the mask polynomial m_w of w, since $m_w(\alpha) = \sum_{i=0}^{N-1} w(i)\alpha^i = 0$. On the other hand, for a given $N \in \mathbb{N}$, for every $p \in P_N$ we have that α is the root of the mask polynomial of $\mathbb{Z}_p \leq \mathbb{Z}_N$, that is, $\Phi_p(x^{N/p})$. Indeed, α is a primitive N-th root of unity so $\alpha^{N/p} \neq 1$. Clearly, $\alpha^{N/p}\Phi_p(\alpha^{N/p}) = \Phi_p(\alpha^{N/p})$, so it implies $\Phi_p(\alpha^{N/p}) = 0$.

Then α is also a root of the polynomial $m_w(x) + \sum_{p \in P_N} a_p(x)\Phi_p(x^{N/p})$, where $a_p(x) \in \mathbb{Q}[x]$. By using Euclidean division there are polynomial $q(x), r(x) \in \mathbb{Q}[x]$ such that

$$m_w(x) = q(x)\Phi_N(x) + r(x),$$

with either $r(x)$ the constant zero function or $\deg(r(x)) < \varphi(N)$.

The common roots of the polynomials $\Phi_p(x^{N/p})$ $(p \in P_N)$ are exactly the primitive N-th roots of unity. The multiplicity of these roots in all of these polynomials is 1. These polynomials are all in $\mathbb{Q}[x]$ so the greatest common divisor in the ring $\mathbb{Q}[x]$ of the polynomials $\Phi_p(x^{N/p})$, $p \in P_N$, is $\Phi_N(x)$. Thus

$$\Phi_N(x) = \sum_{p \in P_N} a_p(x)\Phi_p(x^{N/p})$$

for some $a_p(x) \in \mathbb{Q}[x]$. Substituting this into the previous equation we obtain that

$$m_w(x) - \sum_{p \in P_N} q(x)a_p(x)\Phi_p(x^{N/p})$$

is of degree less than $\varphi(N)$ or is the constant zero function. Since $\Phi_N(x)$ is the minimal polynomial of α over $\mathbb{Q}$, we have $m_w(x) - \sum_{p \in P_N} q(x)a_p(x)\Phi_p(x^{N/p}) = 0$. Thus

$$m_w(x) = \sum_{p \in P_N} q(x)a_p(x)\Phi_p(x^{N/p}).$$

It is clear that $x^k\Phi_p(x^{N/p})$ is the mask polynomial of a coset of $\mathbb{Z}_p$ for every $0 \le k < N$. Hence we have

$$w(x) = \sum_{g \in G}\sum_{p \in P_N} w_{p,g} 1_{\mathbb{Z}_p+g}(x)$$

for some $w_{p,g} \in \mathbb{Q}$.

The other direction follows from the fact that $\Phi_p(\alpha^{N/p}) = 0$ for every $p \in P_N$. □

We note that using Proposition 3.2 one can simply construct the asymmetric minimal sums of roots of unity appearing in [Lam and Leung 2000].

In terms of mask polynomials the previous proposition can be stated as follows.

Corollary 3.3. *Let $S(x) \in \mathbb{Z}_{\ge 0}[x]$ with $S(\xi_N) = 0$, where $N = p_1^{m_1} \cdots p_n^{m_n}$ and $p_1, \dots, p_n$ are primes. Then,*

$$S(x) \equiv P_1(x)\Phi_{p_1}(x^{N/p_1}) + \cdots + P_n(x)\Phi_{p_n}(x^{N/p_n}) \bmod (x^N - 1)$$

for some $P_1(x), \dots, P_n(x) \in \mathbb{Q}[x]$.

The following is an easy consequence of Proposition 3.2.

Corollary 3.4. *Let G be a cyclic group of order N and Ψ be a faithful representation of G. Assume w is a non-Pompeiu weight function with respect to Ψ. Then the restriction of w to each $\mathbb{Z}_{\mathsf{Rad}(N)}$-coset is the weighted sum of characteristic functions of $\mathbb{Z}_{p_i}$-cosets, where $\mathsf{Rad}(N)$ denotes the square-free radical of N.*

We will consider

$$\mathbb{Z}_{\prod_{i=1}^d p_i} \cong \prod_{i=1}^d \mathbb{Z}_{p_i}$$

as a grid in $\mathbb{R}^d$, whose points have integer coordinates. More precisely for $\mathbb{Z}_{\prod_{i=1}^d p_i}$ we assign

$$\mathcal{G} = \{x \in \mathbb{Z}^d : 0 \le x_i \le p_i - 1 \text{ for } 1 \le i \le d\},$$

where x_i denotes the i-th coordinate of x. The cosets of $\mathbb{Z}_{p_i}$ coincide with collections of parallel line segments (containing p_i grid points of $\mathcal{G}$). A d-dimensional grid-cuboid will be a collection of 2^d grid points that resembles a d-dimensional cuboid in $\mathbb{R}^d$. Let $P \subset \mathcal{G}$ be a d-dimensional grid-cuboid and fix a point $y \in P$. For a point of $z \in P$ let $\pi(z)$ denote the Hamming distance between z and y. Note that w can also be considered as a function from $\mathcal{G}$ to $\mathbb{Q}$.

The following statement makes the Pompeiu property for weight functions easily recognizable.

Proposition 3.5. *Let w be a non-Pompeiu weight function on the set $\mathbb{Z}_{\prod_{i=1}^d p_i}$, where p_i are mutually different primes. If w is the weighted sum of characteristic functions of $\mathbb{Z}_{p_i}$-cosets, then for every d-dimensional grid-cuboid P we have*

$$\sum_{c\in P}(-1)^{\pi(c)}w(c) = 0. \tag{6}$$

Proof. It is easy to see that each coset of $\mathbb{Z}_{p_i}$ for any $p_i \mid n$ contains either 2 or 0 elements of the cuboid P. Substituting the characteristic function of any coset of $\mathbb{Z}_{p_i}$ as a weight function into the left-hand side of (6), it is clearly reduced to a sum of at most two elements with different signs; thus (6) holds. □

Remark 3.6. The converse of the previous statement also holds. We leave it to the reader to work out the details of the proof.

Now we describe a few special cases which will be later used for the proof of Theorem 2.5. In the proof of the next proposition we use the following definition.

Definition 3.7. Let $S \subseteq \mathbb{Z}_N$. For every $j \in \mathbb{Z}$ and $d \mid N$, we define the subsets

$$S_{j \bmod d} = \{s \in S : s \equiv j \mod d\}.$$

Proposition 3.8. (a) *Every non-Pompeiu set in $\mathbb{Z}_{pq}$ with respect to a faithful representation is either the union of cosets of $\mathbb{Z}_p$ or those of $\mathbb{Z}_q$.*

(b) *Let $N = p^m q^n$ and let S be a non-Pompeiu multiset in $\mathbb{Z}_N$ with respect to a faithful representation. Then there are polynomials $P(x), Q(x) \in \mathbb{Z}_{\geq 0}[x]$ such that*

$$S(x) \equiv P(x)\Phi_p(x^{N/p}) + Q(x)\Phi_q(x^{N/q}) \mod (x^N - 1).$$

Proof. (a) Let S be a non-Pompeiu set in $\mathbb{Z}_{pq}$ with respect to a faithful representation and let w be the characteristic function of S. Using Proposition 3.2 we can write $w = \sum_{i=0}^{q-1} a_i 1_{\mathbb{Z}_p+i} + \sum_{j=0}^{p-1} b_j 1_{\mathbb{Z}_q+j}$, where $a_i, b_j \in \mathbb{Q}$. Then the range of w is $\text{Ran}(w) = \{a_i + b_j : 0 \leq i \leq p-1,\ 0 \leq j \leq q-1\}$. We have $\text{Ran}(w) = \{0, 1\}$. Thus there are at most two different a_i and two different b_j.

One can treat the case when a_i and b_j are constants as functions of i or j, respectively. Thus we may assume that $a_k < a_l$ for some $0 \leq k, l \leq p-1$. Then clearly $a_k + b_j = 0$ and $a_l + b_j = 1$ for all b_j; in particular all b_j are the same. Therefore, we may write

$$w = b + \sum_{i=0}^{p-1} a_i 1_{\mathbb{Z}_p+i} = \sum_{i=0}^{p-1}(b + a_i)1_{\mathbb{Z}_p+i},$$

finishing the proof of the statement.

(b) By Corollary 3.3, it is clear that

$$S(x) \equiv P(x)\Phi_p(x^{N/p}) + Q(x)\Phi_q(x^{N/q}) \bmod (x^N - 1)$$

for some $P(x), Q(x) \in \mathbb{Q}[x]$. Now we show that P and Q can be chosen such that $P(X), Q(x) \in \mathbb{Z}_{\geq 0}[x]$.

The subgroups $\mathbb{Z}_p$ and $\mathbb{Z}_q$ generate $\mathbb{Z}_{pq}$. Thus S can be written as the disjoint union

$$S = \bigcup_{k \in C} S_{k \bmod N/(pq)}$$

for $k = 0, \ldots, N/(pq) - 1$, where k runs through a set of representatives C of the cosets of $\mathbb{Z}_{pq}$. Thus we are given

$$S_{k \bmod N/(pq)} = \sum_{a \in A} c_a(\mathbb{Z}_p + a) + \sum_{b \in B} d_b(\mathbb{Z}_q + b),$$

where $c_a + d_b \in \mathbb{Z}_{\geq 0}$ and A and B are sets of coset representatives of $\mathbb{Z}_p$ and $\mathbb{Z}_q$, respectively, in $\mathbb{Z}_{pq} + k$. We want to modify the coefficients c_a and d_b such that they produce the same multiset and all of them remain nonnegative.

Let $e = c_a + d_b$ be one of the minimal weights of the multiset S. Then the values $d'_x = (c_a + d_x) - e$ are nonnegative for every $x \in B$, and let $c'_y = c_y + d_b$, which are nonnegative since these values are given by the multiset S only.

Now $c'_y + d'_x = ((c_a + d_x) - e) + c_y + d_b = c_y + d_x$ for every $x \in B$ and $y \in A$, finishing the proof of the lemma. □

4. Reduction (of Fuglede's conjecture)

Before we proceed to the proof of Theorem 2.5 we make a few general observations.

Lemma 4.1. *Let G be a finite abelian group. Assume that $S \subset G$ is a spectral set having Λ as a spectrum*:

(a) *$S + t$ is spectral with the same spectrum Λ for every $t \in G$.*

(b) *$\Lambda + \omega$ is a spectrum for S for every $\omega \in G$.*

(c) *S is a spectrum for Λ.*

Proof. (a) If $\sum_{s \in S} \chi_\delta(s) = 0$ for some $\delta \in \Lambda - \Lambda$, then since χ_δ is a homomorphism, we have

$$\sum_{u \in (S+t)} \chi_\delta(u) = \chi_\delta(t) \sum_{s \in S} \chi_\delta(s) = 0.$$

Thus the orthogonality of the representations corresponding to the spectrum is preserved under translation.

(b) Similarly, the orthogonality of the representations corresponding to $\Lambda + \omega$ follows from the fact that $\Lambda - \Lambda = (\Lambda + \omega) - (\Lambda + \omega)$.

(c) This follows from the fact that a finite abelian group is canonically isomorphic to its double dual. □

Corollary 4.2. *It is enough to prove Theorem 2.5 for spectral sets S with $0 \in S$ and with spectrum Λ that contains 0.*

From now on we assume $0 \in S$ and $0 \in \Lambda$.

Lemma 4.3. *Let G be a finite abelian group and let S be spectral in G and such that it does not generate G. Assume that for every proper subgroup H of G we have $\boldsymbol{S}-\boldsymbol{T}(H)$. Then S tiles G.*

Proof. Let S be a spectral set with orthogonal basis $\{\chi_\lambda : \lambda \in \Lambda\} = \chi_\Lambda \subset \widetilde{G}$ and let $\langle S \rangle = H < G$. Since every χ_λ is 1-dimensional, we have $\{\chi_{\lambda|H} : \lambda \in \Lambda\} \subseteq \widetilde{H}$ and clearly these are still orthogonal on S, since $S \subset H$. Then using that $\boldsymbol{S}-\boldsymbol{T}(\mathrm{H})$ holds, there is a set $T \subset H$ with $S+T = H$. Now let U be a complete set of coset representatives of G/H. Then we have $S+(T+U) = G$. □

Now we prove a similar lemma reducing the possible structure of Λ.

Lemma 4.4. *Let G by a cyclic group of order N and let us suppose that $\boldsymbol{S}-\boldsymbol{T}(G/H)$ holds on every proper factor G/H. Let S be a spectral set of G and Λ be the corresponding spectrum. Assume that the intersection of the kernels of the elements of χ_Λ contains $H_{N/\ell} \neq 1$ for some $1 < \ell \mid N$. Then S tiles G.*

Proof. By our assumptions the elements of χ_Λ can be considered as irreducible representations of $G/H_{N/\ell}$ since their kernel is contained in $H_{N/\ell}$.

Let S_ℓ denote the multiset obtained as the image of S by the canonical projection π_ℓ of G to $G/H_{N/\ell} \cong H_\ell$. We claim that the multiset S_ℓ is a set in H_ℓ. Indeed there cannot be two elements of S in the same coset of $H_{N/\ell}$ since otherwise each element of χ_Λ would have the same value on them, contradicting the fact that these representations form a basis of the set of complex-valued functions on S. Thus S_ℓ is a set. Now it is easy to derive that $\Lambda/H_{N/\ell}$ is a spectrum with respect to S_ℓ in $G/H_{N/\ell}$ since $\chi_\lambda(\pi_\ell(s)) = \chi_\lambda(s)$ for every $s \in S$ and $\lambda \in \Lambda$.

We know $\boldsymbol{S}-\boldsymbol{T}(G/H_{N/\ell})$ holds. As S_ℓ is a spectral set in $G/H_{N/\ell}$ there is $T_\ell \subset G/H_{N/\ell}$ with $S_\ell + T_\ell = G/H_{N/\ell}$. Then if T is the preimage of T_l under the canonical projection from G to $G/H_{N/\ell}$, then we have $S+T = G$. □

Observation 4.5. Let us recall that $S(x)$ is the mask polynomial of the spectral set S. Note that for $\chi \in \widetilde{G}$ of order k, we have $\sum_{s\in S} \chi(s) = 0$ is equivalent to the fact that a primitive k-th root of unity ξ_k is a root of $S(x)$. Since $\Phi_k(x)$ is irreducible over $\mathbb{Q}$ we have $\Phi_k(x) \mid S(x)$; hence every primitive k-th root of unity is the root of $S(x)$ and $\sum_{s\in S} \chi'(s) = 0$ for every $\chi' \in \widetilde{G}$ of the same order. If $\Lambda \subseteq G$ is a spectrum of S, the above can be summarized to

$$S(\xi_{\mathrm{ord}(\lambda-\lambda')}) = 0 \tag{7}$$

for every $\lambda \neq \lambda'$ in a spectrum Λ, using (1).

The question whether our techniques can be generalized naturally arises. We point out here that in the next proposition we heavily use the assumption that the order of cyclic groups is divisible by at most two different primes.

Proposition 4.6. *Let G be a cyclic group of order $p^k q^\ell$ and let $|S| \geq 2$ be a spectral set. Assume further that Λ is a spectrum for S such that the elements of χ_Λ do not have a nontrivial common kernel. Then for every faithful representation ψ of G we have $\sum_{s\in S} \psi(s) = 0$.*

Proof. Note that by Observation 4.5, it is enough to prove the statement for one faithful representation.

Since the elements of χ_Λ do not have a common kernel we have a $\lambda_1 \in \Lambda$ with $p \nmid \lambda_1$. If $q \nmid \lambda_1$, then we are done so we assume $q \mid \lambda_1$. Similarly, we might assume that there exists $\lambda_2 \in \Lambda$ with $q \nmid \lambda_2$ and $p \mid \lambda_2$. In this case $\chi_{\lambda_1-\lambda_2}$ generates $\widetilde{G}$ so we have $\sum_{s\in S} \chi_{\lambda_1-\lambda_2}(s) = 0$. □

This has the following interpretation in terms of mask polynomials.

Corollary 4.7. *Let* (S, Λ) *be a spectral pair in* $\mathbb{Z}_N$*, where* $N = p^k q^\ell$*, such that* $0 \in S,\ 0 \in \Lambda$*, and each of* S, Λ *generates* $\mathbb{Z}_N$*. Then*

$$S(\xi_N) = \Lambda(\xi_N) = 0.$$

Proposition 4.8. *Let S be a spectral set in $\mathbb{Z}_N$ and let p be a prime divisor of N. Assume that for every proper factor group $\mathbb{Z}_N/H$ of $\mathbb{Z}_N$ we have $\boldsymbol{S}-\boldsymbol{T}(\mathbb{Z}_N/H)$. Assume further that S is the disjoint union of cosets of $\mathbb{Z}_p$. Then S tiles $\mathbb{Z}_N$.*

Proof. By our assumptions $|S| = pr = |\Lambda|$ for some $r \in \mathbb{N}$ and Λ is a spectrum for S. Thus at least one of the cosets of H_p contains at least r elements of Λ. By Lemma 4.1(b)enumi we may assume that $|H_p \cap \Lambda| \geq r$. The elements $\chi_\Lambda \subseteq \widetilde{H}_p$ are representations having a common kernel $\mathbb{Z}_p = H_{N/p}$. By our assumption S is the disjoint union of $\mathbb{Z}_p$-cosets, so it can be written as $\mathbb{Z}_p + B$ for some $B \subseteq \mathbb{Z}_N/\mathbb{Z}_p$. The representations in $\widetilde{H}_p \cap \chi_\Lambda$ are constant on every coset of $\mathbb{Z}_p$. Hence for every $\chi_1 \neq \chi_2 \in \widetilde{H}_p \cap \chi_\Lambda$ we have

$$\begin{aligned}0 = \sum_{s\in S} \chi_1(s)\bar{\chi}_2(s) &= \sum_{s\in\mathbb{Z}_p+B} \chi_1(s)\bar{\chi}_2(s) = \sum_{t\in B}\sum_{x\in\mathbb{Z}_p} \chi_1(t+x)\bar{\chi}_2(t+x)\\ &= \sum_{t\in B}\sum_{x\in\mathbb{Z}_p} \chi_1(t)\chi_1(x)\bar{\chi}_2(t)\bar{\chi}_2(x) = \sum_{t\in B} p\chi_1(t)\bar{\chi}_2(t) = p\sum_{t\in B}\chi_1(t)\bar{\chi}_2(t),\end{aligned}$$

since the kernel of χ_1 and χ_2 contains $\mathbb{Z}_p$. Thus we obtain a set of $r = |B|$ representations of $\mathbb{Z}_N/\mathbb{Z}_p$, which are mutually orthogonal, hence forming a basis of $L^2(B)$. Thus B is a spectral set in $\mathbb{Z}_N/\mathbb{Z}_p$ and using our assumption we obtain that there exists T with $B + T = \mathbb{Z}_N/\mathbb{Z}_p$. So finally we get $S + T = (\mathbb{Z}_p + B) + T = \mathbb{Z}_p + (B + T) = \mathbb{Z}_p + \mathbb{Z}_N/\mathbb{Z}_p = \mathbb{Z}_N$. □

Before we start to detail the proof of Theorem 2.5 we summarize what we have already proved in the previous sections about the structure of a spectral set S in $\mathbb{Z}_{p^nq^2}$. Note that we may assume by induction on n that $\boldsymbol{S}-\boldsymbol{T}(H)$ holds for every proper subgroup or factor H of $\mathbb{Z}_{p^nq^2}$. Indeed, Fuglede's conjecture holds for $\mathbb{Z}_{pq^2}$ and for $\mathbb{Z}_{p^nq}$ by [Malikiosis and Kolountzakis 2017], which corresponds to the base case of our induction.

If $|S| = 1$, then S is clearly a spectral set and also a tile. By Lemma 4.4 we might assume that the elements of χ_Λ do not have a common kernel so by Proposition 4.6 we might assume that $|S| \geq 2$ is a non-Pompeiu set with respect to a faithful representation of $\mathbb{Z}_{p^nq^2}$. Hence by Proposition 3.2 we have

$$S = \sum_{g\in A} u_g(\mathbb{Z}_p + g) + \sum_{h\in B} v_h(\mathbb{Z}_q + h),$$

where $u_g, v_h \in \mathbb{Q}$ and A and B are sets of coset representatives of $\mathbb{Z}_p$ and $\mathbb{Z}_q$, respectively. Thus S is the weighted sum of cosets of $\mathbb{Z}_p$ and $\mathbb{Z}_q$. Until now we have only seen that the weights are rational numbers. Now we prove that all weights are 0 or 1.

The subgroups $\mathbb{Z}_p$ and $\mathbb{Z}_q$ generate $\mathbb{Z}_{pq}$, so we write S as the disjoint union

$$S = \bigcup_{k \in C} S_{k \bmod N/(pq)},$$

where k runs through a set of representatives C of the cosets of $\mathbb{Z}_{pq}$ for $k = 0, \dots, N/(pq) - 1$. Now

$$S_{k \bmod N/(pq)} = \sum_{\substack{g \in A \\ g+\mathbb{Z}_p \subset k+\mathbb{Z}_{pq}}} u_g(\mathbb{Z}_p + g) + \sum_{\substack{h \in B \\ h+\mathbb{Z}_q \subset k+\mathbb{Z}_{pq}}} v_h(\mathbb{Z}_q + h) \tag{8}$$

for every $k \in C$, so $S_{k \bmod N/(pq)}$ inherits its weights from S. Now it follows from Proposition 3.8 that in (8) $u_g = 0$ for every $g \in A,\ g + \mathbb{Z}_p \subset k + \mathbb{Z}_{pq}$ or $v_h = 0$ for every $h \in B,\ h + \mathbb{Z}_q \subset k + \mathbb{Z}_{pq}$. Since $S_{k \bmod N/(pq)}$ is a set, the remaining coefficients are 0 or 1. Then $S_{k \bmod N/(pq)}$ is the disjoint nontrivial union of $\mathbb{Z}_p$-cosets or $\mathbb{Z}_q$-cosets. Only one type appears for every fixed $k = 0, \dots, N/(pq) - 1$ except in the obvious case as follows:

It can happen that S contains a whole $\mathbb{Z}_{pq}$-coset, in which case it can be considered as the union of only $\mathbb{Z}_p$-cosets and only $\mathbb{Z}_q$-cosets as well. Thus S is the disjoint union of $\mathbb{Z}_p$-cosets and $\mathbb{Z}_q$-cosets.

Beside the case when S contains both $\mathbb{Z}_p$-cosets and $\mathbb{Z}_q$-cosets, by Proposition 4.8 we are done. Thus, we may assume S contains both $\mathbb{Z}_p$-cosets and $\mathbb{Z}_q$-cosets; we shall call such sets *nontrivial unions* of $\mathbb{Z}_p$- and $\mathbb{Z}_q$-cosets, to emphasize that they cannot be expressed as unions consisting solely of $\mathbb{Z}_p$-cosets, or $\mathbb{Z}_q$-cosets.

The above also follows from Corollary 4.7 and the structure of vanishing sums of roots of unity of order N, where N has at most two distinct prime factors [Lam and Leung 2000]. We added also a condition that shows when such a vanishing sum corresponds to a nontrivial union of $\mathbb{Z}_p$- and $\mathbb{Z}_q$-cosets, which is a consequence of Corollary 3.3 and Proposition 3.8(b)enumi or alternatively of Proposition 2.6 in [Malikiosis and Kolountzakis 2017].

Theorem 4.9. *Let $F(x) \in \mathbb{Z}_{\geq 0}[x]$ and $N = p^m q^n$, where p, q are different primes. Then, $F(\xi_N) = 0$ if and only if*

$$F(x) \equiv P(x)\Phi_p(x^{N/p}) + Q(x)\Phi_q(x^{N/q}) \bmod (x^N - 1)$$

for some $P(x), Q(x) \in \mathbb{Z}_{\geq 0}[x]$. If $F(\xi_N^{p^k}) \neq 0$ for some $1 \leq k \leq m$, then we cannot have $P(x) \equiv 0 \bmod (x^N - 1)$, and if $F(\xi_N^{q^\ell}) \neq 0$ for some $1 \leq \ell \leq n$, then we cannot have $Q(x) \equiv 0 \bmod (x^N - 1)$.

We will repeatedly use the above in Section 6 in order to obtain information about the structure of S and Λ from the vanishing of their mask polynomials on various N-th roots of unity. Regarding the case when S is a union of $\mathbb{Z}_p$-cosets (or $\mathbb{Z}_q$-cosets), there is a characterization in terms of the mask polynomial. This follows from a special case of Ma's lemma [1985], see also [Schmidt 2002, Lemma 1.5.1] or [Pott 1995, Corollary 1.2.14], adapted to the cyclic case, using the polynomial notation.

Lemma 4.10. *Suppose that $S(x) \in \mathbb{Z}[x]$, and let $\mathbb{Z}_N$ be a cyclic group such that $p^m \mid N$, but $p^{m+1} \nmid N$. If $S(\xi_d) = 0$ for every $p^m \mid d \mid N$, then*

$$S(x) \equiv P(x)\Phi_p(x^{N/p}) \bmod (x^N - 1).$$

If the coefficients of S are nonnegative, then P can be taken with nonnegative coefficients as well. In particular, if $S \subseteq \mathbb{Z}_N$ satisfies $S(\xi_d) = 0$ for every $p^m \mid d \mid N$, then S is a union of $\mathbb{Z}_p$-cosets.

We summarize the reductions made so far in the following list.

Reduction 1. We may assume that a spectral set $S \subset \mathbb{Z}_{p^n q^2}$, along with a spectrum Λ, has the following structure:

(a) $0 \in S,\ 0 \in \Lambda$ and each of S and Λ generates $\mathbb{Z}_{p^n q^2}$.

(b) Both S and Λ can be written as the disjoint nontrivial union of $\mathbb{Z}_p$-cosets and $\mathbb{Z}_q$-cosets and this holds for $S \cap (\mathbb{Z}_{pq} + g)$ and $\Lambda \cap (\mathbb{Z}_{pq} + h)$ for every $g, h \in \mathbb{Z}_{p^n q^2}$ as well.

(c) There is a $\mathbb{Z}_{pq}$-coset which intersects S and its complement. Further the intersection is the union of $\mathbb{Z}_p$-cosets. The same holds for another $\mathbb{Z}_{pq}$-coset with $\mathbb{Z}_q$-cosets as well.

(d) Fuglede's conjecture holds for all $\mathbb{Z}_M$, with $M \mid p^n q^2,\ M < p^n q^2$ (induction assumption).

Proof. (a) This follows from Lemmas 4.1 and 4.3.

(b) This is an immediate consequence of part (a), Propositions 3.2 and 3.8 and Corollary 4.7.

(c) This follows from Proposition 4.8.

(d) It was proved in [Malikiosis and Kolountzakis 2017] that Fuglede's conjecture holds for $N = p^n q$, and also for $N = pq^2$, so the given statement certainly holds for $p^2 q^2$, which is the base case for the inductive argument. □

Now we turn to the main tool already used in [Malikiosis and Kolountzakis 2017] to prove that a spectral set tiles $\mathbb{Z}_{p^n q^2}$. Clearly, sets coincide with mask polynomials having only coefficients 0 and 1. The following theorem was proved in [Coven and Meyerowitz 1999]. Let H_S be the set of prime powers r^a dividing N such that $\Phi_{r^a}(x) \mid S(x)$.

Theorem 4.11. *If $S \subset \mathbb{Z}_N$ satisfies the following two conditions* (T1) *and* (T2), *then S tiles $\mathbb{Z}_N$.*

(T1) $S(1) = \prod_{d \in H_S} \Phi_d(1)$.

(T2) *For pairwise relative prime elements s_i of H_S, we have $\Phi_{\prod s_i} \mid S(x)$.*

Note that $\Phi_{p^a}(1) = p$ for a prime p and $\Phi_k(1) = 1$ if k has at least two different prime divisors.

5. Preliminary lemmas

We introduce extra notation for divisibility. Fix $N \in \mathbb{N}$. For a natural number k we write $\ell \parallel_N k$ if ℓ is the largest divisor of N, which divides k. In our case N will be $p^n q^2$ so we simply write $\ell \parallel k$.

We review first (1) and (7) for a spectral pair (S, Λ) in $\mathbb{Z}_N$. First, we define as usual

$$\mathbb{Z}_N^\star = \{g \in \mathbb{Z}_N : \gcd(g, N) = 1\},$$

the group of *reduced residues* mod N. It is precisely the subset of elements of N of order exactly N. Similarly, the subset of $\mathbb{Z}_N$ of elements of order N/d, where $d \mid N$, is

$$d\mathbb{Z}_N^\star = \{g \in \mathbb{Z}_N : \gcd(g, N) = d\}.$$

The zero set

$$Z(S) = \{d \in \mathbb{Z}_N : S(\xi_N^d) = 0\}$$

is then a union of subsets of the form $d\mathbb{Z}_N^\star$ for some $d \mid N$, and (1) and (7) can be rewritten as

$$\Lambda - \Lambda \subseteq \{0\} \cup \bigcup_{d \mid N,\, S(\xi_N^d)=0} d\mathbb{Z}_N^\star. \tag{9}$$

Of course, by Lemma 4.1(c)enumi, the roles of S and Λ can be reversed.

Definition 5.1. Let $S \subseteq \mathbb{Z}_N$. Recall that for every $j \in \mathbb{Z}$ and $d \mid N$, we define the subsets

$$S_{j \bmod d} = \{s \in S : s \equiv j \mod d\}.$$

We say that S is *equidistributed* mod d if

$$|S_{j \bmod d}| = \frac{1}{d}|S|$$

for every j. Equivalently, every $\mathbb{Z}_{N/d}$-coset of $\mathbb{Z}_N$ contains the same number of elements of S.

Lemma 5.2. (a) *Assume $\Phi_p(x) \mid S(x)$. Then every $\mathbb{Z}_{N/p}$-coset of $\mathbb{Z}_N$ contains the same number of elements of S.*

(b) *Assume $\Phi_k(x) \mid S(x)$ for every $1 < k \mid d$. Then every $\mathbb{Z}_{N/d}$-coset of $\mathbb{Z}_N$ contains the same number of elements of S.*

Proof. (a) $\Phi_p(x) \mid S(x)$ is equivalent to the fact that S is a non-Pompeiu set with respect to an irreducible representation of order p, whose kernel is $\mathbb{Z}_{N/p}$. It is easy to see that a non-Pompeiu multiset on $\mathbb{Z}_p$ has to be constant, and we obtain the result.

(b) Consider the formula

$$S(x) \equiv \sum_{j=0}^{d-1} |S_{j \bmod d}| x^j \mod (x^d - 1), \tag{10}$$

which holds for every $S \subseteq \mathbb{Z}_N$. It holds that $S(\xi_k) = 0$ for every $1 < k \mid d$ if and only if

$$1 + x + \cdots + x^{d-1} = \prod_{1<k \mid d} \Phi_k(x) \mid S(x),$$

or equivalently $S(x) = (1 + x + \cdots + x^{d-1})G(x)$, where $G(x) \in \mathbb{Z}[x]$. The latter implies

$$S(x) \equiv (1 + x + \cdots + x^{d-1})G(1) \mod (x^d - 1),$$

so by (10) we get $|S_{j \bmod d}| = G(1)$ for all j. Conversely, if $|S_{j \bmod d}| = c$ for all j, then

$$S(x) \equiv c(1 + x + \cdots + x^{d-1}) \mod (x^d - 1),$$

due to (10), which easily gives $S(\xi_k) = 0$ for every $1 < k \mid d$, as desired. □

Let (S, Λ) be a spectral pair in $\mathbb{Z}_N$ satisfying the conditions of Reduction 1, where $N = p^n q^2$. An immediate consequence of Reduction 1(c) is that $S - S$ contains the difference set of both a $\mathbb{Z}_p$-coset and a $\mathbb{Z}_q$-coset; thus

$$\frac{N}{p}\mathbb{Z}_N \cup \frac{N}{q}\mathbb{Z}_N \subseteq S - S,$$

whence

$$\Lambda(\xi_p) = \Lambda(\xi_q) = 0, \tag{11}$$

by (7), and we obtain in particular,

$$|\Lambda_{i \bmod p}| = \frac{1}{p}|\Lambda|, \quad |\Lambda_{j \bmod q}| = \frac{1}{q}|\Lambda| \tag{12}$$

for all i, j, by Lemma 5.2. This shows that pq divides $|S| = |\Lambda|$.

6. Proof of Theorem 2.5

A significant special case will be shown first.

Lemma 6.1. *Let $S \subseteq \mathbb{Z}_N$ be spectral. If $q^2 \mid |S|$, then S tiles $\mathbb{Z}_N$.*

Proof. Let $H_S(p) = \{p^m : S(\xi_{p^m}) = 0,\ 1 \le m \le n\}$, and similarly define $H_\Lambda(p)$ for a spectrum $\Lambda \subseteq \mathbb{Z}_N$. Suppose that

$$H_\Lambda(p) = \{p^{m_1}, \dots, p^{m_k}\},$$

where $1 \le m_1 < m_2 < \cdots < m_k \le n$. For every j,

$$S_{j \bmod q^2} - S_{j \bmod q^2} \subseteq (S - S) \cap q^2\mathbb{Z}_N \subseteq \{0\} \cup \bigcup_{i=0}^{k} \frac{N}{p^{m_i}}\mathbb{Z}_N^\star, \tag{13}$$

by (9). Consider the p-adic expansion of every $s \in S$ taken mod p^n, as follows:

$$s \equiv s_0 + s_1 p + \cdots + s_{n-1}p^{n-1} \bmod p^n, \quad 0 \le s_i \le p-1, \quad 0 \le i \le n-1.$$

Due to (13), the elements of each $S_{j \bmod q^2}$ cannot have the same p-adic digits corresponding to p^{n-m_i}, $1 \le i \le k$, yielding

$$|S_{j \bmod q^2}| \le p^k, \quad 0 \le j < q^2;$$

thus, $|S| \le p^k q^2$. On the other hand, we have

$$\prod_{i=1}^{k} \Phi_{p^{m_i}}(x) \mid \Lambda(x),$$

and putting $x = 1$ we obtain $p^k \mid |\Lambda|$; we then get by hypothesis $p^k q^2 \mid |S|$, whence $|S| = p^k q^2$, and

$$|S_{j \bmod q^2}| = p^k, \quad 0 \le j < q^2.$$

Since S is equidistributed mod q^2, we must also have $S(\xi_q) = S(\xi_{q^2}) = 0$ by Lemma 5.2. We note that each element of $S_{j \bmod q^2}$ is unique mod p^n, so the reduction mod p^n map

$$\pi : \mathbb{Z}_N \mapsto \mathbb{Z}_{p^n}$$

is injective on each $S_{j \bmod p^n}$; fix some j, and let $\pi(S_{j \bmod p^n}) = S'$. Since $q^2 \mid s - s'$ for every $s, s' \in S_{j \bmod q^2}$, we conclude that the order of $s - s'$ in $\mathbb{Z}_N$ is the same as the order of $\pi(s - s')$ in $\mathbb{Z}_{p^n}$, which gives

$$S' - S' \subseteq \{0\} \cup \bigcup_{i=0}^{k} p^{n-m_i} \mathbb{Z}^{\star}_{p^n}.$$

Consider now the set $\Lambda' \subseteq \mathbb{Z}_{p^n}$ whose mask polynomial is given by

$$\Lambda'(x) \equiv \prod_{i=1}^{k} \Phi_{p^{m_i}}(x) \bmod (x^{p^n} - 1).$$

We have $|S'| = |\Lambda'| = p^k$ and

$$S' - S' \subseteq \{0\} \cup \{d \in \mathbb{Z}_{p^n} : \Lambda'(\xi^d_{p^n}) = 0\};$$

therefore, (S', Λ') is a spectral pair in $\mathbb{Z}_{p^n}$ by (9). Since

$$\Phi_{p^{m_i}}(x) = 1 + x^{p^{m_i-1}} + x^{2p^{m_i-1}} + \cdots + x^{(p-1)p^{m_i-1}},$$

we obtain

$$(\Lambda' - \Lambda') \cap p^{n-m_i+1} \mathbb{Z}^{\star}_{p^n} \neq \varnothing, \quad 1 \leq i \leq k;$$

therefore,

$$\bigcup_{i=0}^{k} p^{n-m_i+1} \mathbb{Z}^{\star}_{p^n} \subseteq \{d \in \mathbb{Z}_{p^n} : S'(\xi^d_{p^n}) = 0\},$$

by (9), or equivalently

$$\prod_{i=0}^{k} \Phi_{p^{n-m_i+1}}(x) \mid S_{j \bmod q^2}(x),$$

since

$$S_{j \bmod q^2}(x) \equiv S'(x) \bmod (x^{p^n} - 1).$$

Moreover, by $S(x) = \sum_{j=0}^{q^2-1} S_{j \bmod q^2}(x)$ and $|S| = p^k q^2$, we conclude that

$$H_S = \{p^{n-m_k+1}, \ldots, p^{n-m_1+1}, q, q^2\};$$

hence S satisfies (T1).

Consider next the polynomial $F(X)$ satisfying

$$S_{j \bmod q^2}(x) \equiv x^j F(x^{q^2}) \bmod (x^N - 1)$$

for a fixed j. Since $\Phi_{p^{n-m_i+1}}(x) \mid F(x^{q^2})$ for all $1 \leq i \leq k$ and q^2 is prime to p^{n-m_i+1}, we also get that $\Phi_{p^{n-m_i+1}}(x) \mid F(x)$. Therefore, for $\ell = 1$ or 2 we get

$$S_{j \bmod q^2}(\xi_{p^{n-m_i+1}q^\ell}) = \xi^j_{p^{n-m_i+1}q^\ell} F(\xi^{q^2}_{p^{n-m_i+1}q^\ell}) = \xi^j_{p^{n-m_i+1}q^\ell} F(\xi^{q^{2-\ell}}_{p^{n-m_i+1}}) = 0$$

for all j, which shows that S satisfies (T2). □

We distinguish now the following cases:

Case 1: $S(\xi_N^q) = S(\xi_N^{q^2}) = 0$. Since $S(\xi_N) = 0$ by Corollary 4.7, S is a union of $\mathbb{Z}_p$-cosets by Lemma 4.10 and S tiles due to Reduction 1(c).

Case 2: $S(\xi_N^q) S(\xi_N^{q^2}) \neq 0$. Consider the difference sets $\Lambda_{j \bmod q} - \Lambda_{j \bmod q}$. They are always subsets of $(\Lambda - \Lambda) \cap q\mathbb{Z}_N$, but since they avoid $q\mathbb{Z}_N^\star \cup q^2\mathbb{Z}_N^\star$ in this case by (9), we get

$$\Lambda_{j \bmod q} - \Lambda_{j \bmod q} \subseteq pq\mathbb{Z}_N$$

for all j. This shows that every element of $\Lambda_{j \bmod q}$ has the same remainder mod p, or equivalently, for every j there is an $i = i(j)$ such that

$$\Lambda_{j \bmod q} \subseteq \Lambda_{i(j) \bmod p}.$$

This, in particular, shows that $p < q$, and that every $\Lambda_{i \bmod p}$ is the disjoint union of sets of the form $\Lambda_{j \bmod q}$, namely

$$\Lambda_{i \bmod p} = \bigcup_{i(j)=i} \Lambda_{j \bmod q}.$$

Suppose that the number of sets appearing in the union are ℓ. Then, the above equation along with (12) implies $1/p = \ell/q$, which leads to a contradiction (no such spectrum can exist).

Case 3: $S(\xi_N^q) = 0 \neq S(\xi_N^{q^2})$. We apply Theorem 4.9 to $S(x) \bmod (x^{N/q} - 1)$. We obtain

$$S(x) \equiv P(x)\Phi_p(x^{N/(pq)}) + Q(x)\Phi_q(x^{N/q^2}) \bmod (x^{N/q} - 1),$$

since $S(\xi_{N/q}) = 0$, where $P(x)$ and $Q(x)$ have nonnegative coefficients. Furthermore, since $S(\xi_N^{q^2}) \neq 0$, we cannot have $Q \equiv 0$. Due to the nonnegativity of P and Q, we obtain the existence of $s, s' \in S$ such that

$$s - s' \equiv \frac{N}{q^2} \bmod \frac{N}{q};$$

hence $p^n \mid s - s'$ but $q \nmid s - s'$, yielding $s - s' \in p^n\mathbb{Z}_N^\star$ and

$$\Lambda(\xi_{q^2}) = 0,$$

which further gives $q^2 \mid |\Lambda|$, so by Lemma 6.1, S tiles $\mathbb{Z}_N$.

Case 4: $S(\xi_N^q) \neq 0 = S(\xi_N^{q^2})$. We will prove the following:

Claim 1.
$$(S - S) \cap \frac{N}{pq^2}\mathbb{Z}_N^\star \neq \varnothing.$$

Proof of Claim. By Theorem 4.9, the multiset[3] q^2S is a union of $\mathbb{Z}_p$-cosets, or equivalently

$$|S_{i \bmod p^n}| = |S_{i+kp^{n-1} \bmod p^n}| \tag{14}$$

[3]Here, we consider the elements $q^2 s \bmod N$, $s \in S$, counting multiplicities. For example, if $N = 4$ and $S = \{0, 2\}$, then $2S$ is the multiset whose only element is 0, appearing with multiplicity 2.

for every i, k. We partition the above sets mod $p^n q$:

$$S_{i \bmod p^n} = \bigcup_{\ell=0}^{q-1} S_{i+\ell p^n \bmod p^n q} \quad \text{and} \quad S_{i+kp^{n-1}q \bmod p^n} = \bigcup_{\ell=0}^{q-1} S_{i+kp^{n-1}q+\ell p^n \bmod p^n q}.$$

If for every i there existed some ℓ such that

$$S_{i+kp^{n-1}q \bmod p^n} = S_{i+kp^{n-1}q+\ell p^n \bmod p^n q}$$

for every k, then qS would also be a union of $\mathbb{Z}_p$-cosets. Indeed, as for every i there is at most one value of $0 \le \ell \le q-1$ such that $S_{i+\ell p^n \bmod p^n q} \neq \varnothing$, and by the above condition the cardinalities of $S_{i+kp^{n-1}q+\ell p^n \bmod p^n q}$ are the same for $0 \le k \le p-1$. Therefore, $S(\xi_{p^n q}) = 0$ by Theorem 4.9 (or equivalently by Proposition 3.2), contradicting the hypothesis. Thus, there exists i such that there are nonempty $S_{i+\ell p^n \bmod p^n q}$ and $S_{i+\ell' p^n \bmod p^n q}$, with $0 \le \ell < \ell' \le q-1$. Clearly, $S_{i+\ell p^n \bmod p^n q} \subseteq S_{i \bmod p^n}$, so $S_{i \bmod p^n}$ is nonempty. Using (14) we have $S_{i+p^{n-1} \bmod p^n}$ is nonempty.

Now let $s \in S_{i+p^{n-1} \bmod p^n}$, $s' \in S_{i+\ell p^n \bmod p^n q}$ and $s'' \in S_{i+\ell' p^n \bmod p^n q}$, so that $p^{n-1} \parallel s - s'$ and $p^{n-1} \parallel s - s''$. Since $s'' - s' \equiv (\ell' - \ell) p^n \bmod p^n q$, we get $q \nmid s'' - s'$, so either $q \nmid s - s'$ or $q \nmid s - s''$ would hold, yielding $(S - S) \cap p^{n-1} \mathbb{Z}_N^\star \neq \varnothing$, as desired. □

This implies

$$\Lambda(\xi_{pq^2}) = 0, \tag{15}$$

by (7). If $\Lambda(\xi_{q^2}) = 0$ then we would have $q^2 \mid |S|$ and S would tile $\mathbb{Z}_N$ by virtue of Lemma 6.1. So, we may assume $\Lambda(\xi_{q^2}) \neq 0$.

By (15) and Theorem 4.9 we get

$$\Lambda(x) \equiv \sum_{j=0}^{pq^2-1} |\Lambda_{j \bmod pq^2}| x^j \equiv P(x)\Phi_p(x^{N/p}) + Q(x)\Phi_q(x^{N/q}) \bmod (x^{pq^2} - 1)$$

for some $P(x), Q(x) \in \mathbb{Z}_{\ge 0}[x]$ and $P(x) \not\equiv 0$ by $\Lambda(\xi_{q^2}) \neq 0$. We note that the function $f(j) = |\Lambda_{j \bmod pq^2}|$ restricted on a $\mathbb{Z}_{pq}$-coset of $\mathbb{Z}_{pq^2}$ is supported either on a $\mathbb{Z}_p$-coset or a $\mathbb{Z}_q$-coset; otherwise, there would exist $\lambda \in \Lambda_{j \bmod pq^2}$ and $\lambda' \in \Lambda_{j' \bmod pq^2}$, where j, j' satisfy

$$j - j' \in q\mathbb{Z}^\star_{pq^2}.$$

This shows that $q \parallel \lambda - \lambda'$ and $p \nmid \lambda - \lambda'$; thus $\lambda - \lambda' \in q\mathbb{Z}_N^\star$ and $S(\xi_N^q) = 0$ by (9), contradicting the hypothesis.

Next, we consider a nonempty subset $\Lambda_{j \bmod pq^2}$; the polynomials with nonnegative coefficients $P(x)\Phi_p(x^{N/p})$ and $Q(x)\Phi_q(x^{N/q})$ contribute to the coefficient of x^j of $\Lambda(x) \bmod (x^{pq^2} - 1)$. If both contributions are positive, then all subsets $\Lambda_{j+kq^2 \bmod pq^2}$ and $\Lambda_{j+\ell pq \bmod pq^2}$ are nonempty for $0 < k < p$ and $0 < \ell < q$. Then, for $\lambda \in \Lambda_{j+q^2 \bmod pq^2}$ and $\lambda' \in \Lambda_{j+pq \bmod pq^2}$, we have $q \parallel \lambda - \lambda'$; hence $\lambda - \lambda' \in q\mathbb{Z}_N^\star$, which contradicts $S(\xi_N^q) \neq 0$, due to (7).

Let $\Gamma(x)$ be $\Lambda(x) \bmod (x^{pq^2} - 1)$. The previous argument shows that the coefficient of x^j of $\Gamma(x)$ is determined completely either from $P(x)\Phi_p(x^{N/p})$ or $Q(x)\Phi_q(x^{N/q})$. Moreover, if $q \parallel j - j'$, then we cannot have that both the coefficients of x^j and $x^{j'}$ in $\Gamma(x)$ are nonzero by the same argument. This means that $f(j) = |\Lambda_{j \bmod pq^2}|$ restricted on a $\mathbb{Z}_{pq}$-coset of $\mathbb{Z}_{pq^2}$ is supported either on a $\mathbb{Z}_p$-coset or a $\mathbb{Z}_q$-coset and constant restricted to this coset.

This shows that for each j such that $\Lambda_{j \bmod pq^2} \neq \varnothing$, either

$$|\Lambda_{j+kq^2 \bmod pq^2}| = \frac{1}{p}|\Lambda_{j \bmod q^2}| = \frac{1}{p}|\Lambda_{j \bmod q}|, \quad 0 \le k < p, \tag{16}$$

or

$$|\Lambda_{j+\ell pq \bmod pq^2}| = |\Lambda_{j+\ell pq \bmod q^2}| = \frac{1}{q}|\Lambda_{j \bmod q}|, \quad 0 \le \ell < q, \tag{17}$$

holds. If (17) holds for some j, then $q^2 \mid |\Lambda|$, so by Lemma 6.1 we get that S tiles $\mathbb{Z}_N$. Therefore, we may assume that (16) holds for all j with $\Lambda_{j \bmod pq^2} \neq \varnothing$. For such j, we have

$$\Lambda_{j \bmod q} - \Lambda_{j \bmod q} = \Lambda_{j \bmod q^2} - \Lambda_{j \bmod q^2} \subseteq (\Lambda - \Lambda) \cap q^2\mathbb{Z}_N;$$

hence $\Lambda - \Lambda$ completely avoids $q\mathbb{Z}_N \setminus q^2\mathbb{Z}_N$. On the other hand, $(S - S) \cap p^n\mathbb{Z}_N^\star = \varnothing$ by (9) and the assumption $\Lambda(\xi_{q^2}) \neq 0$; hence the polynomials

$$\bar{S}(x) \equiv S(x)\Phi_q(x^{p^n}) \bmod (x^N - 1),$$
$$\bar{\Lambda}(x) \equiv \Lambda(x)\Phi_q(x^{N/q}) \bmod (x^N - 1)$$

are mask polynomials of subsets of $\mathbb{Z}_N$, say $\bar{S}$ and $\bar{\Lambda}$, i.e., their coefficients are 0 or 1. We claim that $(\bar{S}, \bar{\Lambda})$ is a spectral pair. They obviously have the same cardinality as $q|S|$, and an element of $\bar{\Lambda} - \bar{\Lambda}$ can be expressed as $\lambda - \lambda' + lN/q$, where $\lambda, \lambda' \in \Lambda$, $|l| < q$.

If $\lambda - \lambda' \in p^k\mathbb{Z}_N^\star$, then $q \nmid \lambda - \lambda' + lN/q$; hence $\lambda - \lambda' + lN/q \in p^k\mathbb{Z}_N^\star$ as well, yielding $S(\xi_N^{\lambda-\lambda'+lN/q}) = 0$, since $N/q = p^n q$.

The remaining case is $\lambda - \lambda' \in p^k q^2\mathbb{Z}_N^\star$, where $0 \le k \le n-1$, as $\Lambda - \Lambda$ avoids $q\mathbb{Z}_N \setminus q^2\mathbb{Z}_N$. In this case, $\lambda - \lambda' + lN/p \in p^k q\mathbb{Z}_N^\star$ if $1 \le l \le q-1$, and $\Phi_q(\xi_{p^k q}^{p^n}) = \Phi_q(\xi_q^{p^{n-k}}) = 0$, so

$$\bar{S}(\xi_N^{\lambda-\lambda'+lN/p}) = 0.$$

If $l = 0$, then clearly $\lambda - \lambda' \in \Lambda - \Lambda$. Considering all of these cases we have $\bar{\Lambda} - \bar{\Lambda} \subseteq \{0\} \cup Z(\bar{S})$, proving that the pair $(\bar{S}, \bar{\Lambda})$ is spectral by virtue of (9). Since $q^2 \mid \bar{S}$ we have $\bar{S}$ tiles $\mathbb{Z}_N$ by Lemma 6.1; thus there is $T \subseteq \mathbb{Z}_N$ such that

$$S(x)\Phi_q(x^{p^n})T(x) \equiv \bar{S}(x)T(x) \equiv 1 + x + \cdots + x^{N-1} \bmod (x^N - 1),$$

so $\Phi_q(x^{p^n})T(x)$ is the mask polynomial of a tiling complement of S using Lemma 1.3 in [Coven and Meyerowitz 1999], completing the proof. □

Appendix

Theorem A.1. *Let S be a subset of $\mathbb{Z}_p^2$. Then S tiles $\mathbb{Z}_p^2$ if and only if S is spectral.*

Iosevich et al. [Iosevich et al. 2017] has already proved this theorem, but we provide an easy combinatorial proof for one of the two directions and a short one for the other direction using Theorem 2.1.

Proposition A.2. *Let S be a spectral set of $\mathbb{Z}_p^2$. Then S tiles $\mathbb{Z}_p^2$.*

Proof. Let S be a spectral set. We may assume $|S| > 1$, since one-element sets clearly tile every group. The corresponding spectrum Λ is also of size at least 2. Then there is a nontrivial irreducible representation ψ of $\mathbb{Z}_p^2$ such that $\sum_{s\in S} \psi(s) = 0$. We may also assume $|S| = |\Lambda| < p^2$.

Representations of $\mathbb{Z}_p^2$ can be parametrized by the elements of $\mathbb{Z}_p^2$. For $u \in \mathbb{Z}_p^2$ let $\chi_u(v) = e^{2\pi i\langle u,v\rangle/p}$, where the scalar product of $\langle u, v\rangle$ is taken modulo p. This can be written as $\sum_{j=0}^{p-1} a_j e^{(2\pi i/p)j}$, where the a_j are integers which are determined in the following way.

From now on we may also think of $\mathbb{Z}_p^2$ as a 2-dimensional vector space over $\mathbb{Z}_p$. Cosets of 1-dimensional subspaces are called lines. Let u' be a nonzero vector orthogonal to u. Then $\langle u, v\rangle$ is constant on every coset of the subgroup generated by u'. Basically, we count the intersection of S with the elements the set of lines parallel with $\langle u'\rangle$. We have $\sum_{j=0}^{p-1} a_j e^{(2\pi i/p)j} = 0$ if and only if a_j is a constant sequence so every element of this class of parallel lines contains the same number of elements of S (i.e., S is equidistributed on these set of parallel lines). As a consequence we get that $p \mid |S|$.

If $|S| = p$, then by the previous argument we have that S intersects each element of a class of parallel lines once. Then S clearly tiles $\mathbb{Z}_p^2$.

Thus we may assume $p+1 < 2p \le |\Lambda| = |S| < p^2$. It is enough to show that such spectral set does not exist. Each class of parallel lines consists of p lines. Thus we have that for every class of parallel lines, at least one line contains at least two elements of Λ. Thus using the argument above we have that every element of every class of parallel lines contains the same amount of elements of S. This means that every line contains k elements of S for some fixed number $p > k \ge 2$.

Let $x \in \mathbb{Z}_p^2 \setminus S$. Take every line containing x. These lines give a disjoint cover of $\mathbb{Z}_p^2 \setminus \{x\}$. Since each of them contains $k \ge 2$ elements we have $|S| = (p+1)k$, which is not divisible by p, a contradiction. □

Proposition A.3. *Let S be a set in $\mathbb{Z}_p^2$, which tiles. Then S is spectral.*

Proof. We may assume that $1 < |S| < p^2$. Then S and its tiling complements are of cardinality p. Using Theorem 2.1 we obtain that either S or T is a subgroup of $\mathbb{Z}_p^2$.

Subgroups are clearly spectral sets. If T is a subgroup, then S is a complete set of coset representatives of T. Let U denote the subgroup of $\mathbb{Z}_p^2$ consisting of vectors orthogonal to T. Then clearly S is equidistributed on the orthogonal lines for $u \in U$ so U is a spectrum for S. □

Acknowledgements

The authors wish to thank the anonymous referees, as well as the editor, Professor Izabella Łaba, for improving the quality of our paper with their suggestions. The research of Kiss was supported by the Hungarian National Foundation for Scientific Research, grant no. K124749, internal research project R-AGR-0500 and the Premium Postdoctoral Fellowship of Hungarian Academy of Science. The research of Somlai was supported by project no. ED 181-2019-0030 (Application-specific highly reliable IT solutions) which has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the Thematic Excellence Programme funding scheme. The research of Vizer was supported by the National Research, Development and

Innovation Office - NKFIH under the grants nos. FK 132060, KH130371 and SNN 129364, by the János Bolyai Research Fellowship and by the New National Excellence Program under the grant number no. ÚNKP-19-4-BME-287.

References

[Babai 1979] L. Babai, "Spectra of Cayley graphs", *J. Combin. Theory Ser. B* **27**:2 (1979), 180–189. MR Zbl

[Brown et al. 1973] L. Brown, B. M. Schreiber, and B. A. Taylor, "Spectral synthesis and the Pompeiu problem", *Ann. Inst. Fourier* (*Grenoble*) **23**:3 (1973), 125–154. MR Zbl

[Chakalov 1944] L. Chakalov, "Sur un problème de D. Pompéiu", *Ann.* [*Godishnik*] *Univ. Sofia. Fac. Phys.-Math. Livre 1.* **40** (1944), 1–14. In Bulgarian. MR Zbl

[Coven and Meyerowitz 1999] E. M. Coven and A. Meyerowitz, "Tiling the integers with translates of one finite set", *J. Algebra* **212**:1 (1999), 161–174. MR Zbl

[Dutkay and Lai 2014] D. E. Dutkay and C.-K. Lai, "Some reductions of the spectral set conjecture to integers", *Math. Proc. Cambridge Philos. Soc.* **156**:1 (2014), 123–135. MR Zbl

[Fan et al. 2019] A. Fan, S. Fan, L. Liao, and R. Shi, "Fuglede's conjecture holds in $\mathbb{Q}_p$", *Math. Ann.* **375**:1-2 (2019), 315–341. MR Zbl

[Farkas et al. 2006] B. Farkas, M. Matolcsi, and P. Móra, "On Fuglede's conjecture and the existence of universal spectra", *J. Fourier Anal. Appl.* **12**:5 (2006), 483–494. MR Zbl

[Fuglede 1974] B. Fuglede, "Commuting self-adjoint partial differential operators and a group theoretic problem", *J. Funct. Anal.* **16**:1 (1974), 101–121. MR Zbl

[Garofalo 1989] N. Garofalo, "A new result on the Pompeiu problem", *Rend. Sem. Mat. Univ. Politec. Torino* **47**:special issue (1989), 25–38. MR Zbl

[Iosevich et al. 1999] A. Iosevich, N. Katz, and S. Pedersen, "Fourier bases and a distance problem of Erdős", *Math. Res. Lett.* **6**:2 (1999), 251–255. MR Zbl

[Iosevich et al. 2003] A. Iosevich, N. Katz, and T. Tao, "The Fuglede spectral conjecture holds for convex planar domains", *Math. Res. Lett.* **10**:5 (2003), 559–569. MR Zbl

[Iosevich et al. 2017] A. Iosevich, A. Mayeli, and J. Pakianathan, "The Fuglede conjecture holds in $\mathbb{Z}_p \times \mathbb{Z}_p$", *Anal. PDE* **10**:4 (2017), 757–764. MR Zbl

[Kiss et al. 2018] G. Kiss, M. Laczkovich, and C. Vincze, "The discrete Pompeiu problem on the plane", *Monatsh. Math.* **186**:2 (2018), 299–314. MR Zbl

[Kolountzakis and Matolcsi 2006a] M. N. Kolountzakis and M. Matolcsi, "Complex Hadamard matrices and the spectral set conjecture", *Collect. Math.* **57**:special issue (2006), 281–291. MR Zbl

[Kolountzakis and Matolcsi 2006b] M. N. Kolountzakis and M. Matolcsi, "Tiles with no spectra", *Forum Math.* **18**:3 (2006), 519–528. MR Zbl

[Łaba 2001] I. Łaba, "Fuglede's conjecture for a union of two intervals", *Proc. Amer. Math. Soc.* **129**:10 (2001), 2965–2972. MR Zbl

[Łaba 2002] I. Łaba, "The spectral set conjecture and multiplicative properties of roots of polynomials", *J. London Math. Soc.* (2) **65**:3 (2002), 661–671. MR Zbl

[Lam and Leung 2000] T. Y. Lam and K. H. Leung, "On vanishing sums of roots of unity", *J. Algebra* **224**:1 (2000), 91–109. MR Zbl

[Ma 1985] S.-L. Ma, *Polynomial addition sets*, Ph.D. thesis, University of Hong Kong, 1985, available at https://tinyurl.com/slmphd.

[Malikiosis and Kolountzakis 2017] R. D. Malikiosis and M. N. Kolountzakis, "Fuglede's conjecture on cyclic groups of order $p^n q$", *Discrete Anal.* **2017** (2017), art. id. 12. MR Zbl

[Matolcsi 2005] M. Matolcsi, "Fuglede's conjecture fails in dimension 4", *Proc. Amer. Math. Soc.* **133**:10 (2005), 3021–3026. MR Zbl

[Newman 1977] D. J. Newman, "Tesselation of integers", *J. Number Theory* **9**:1 (1977), 107–111. MR Zbl

[Pompeiu 1929] D. Pompeiu, "Sur une propriété intégrale des fonctions de deux variables réelles", *Bull. Acad. Bruxelles* **5**:15 (1929), 265–269. Zbl

[Pott 1995] A. Pott, *Finite geometry and character theory*, Lecture Notes in Math. **1601**, Springer, 1995. MR Zbl

[Puls 2013] M. J. Puls, "The Pompeiu problem and discrete groups", *Monatsh. Math.* **172**:3-4 (2013), 415–429. MR Zbl

[Ramm 1997] A. G. Ramm, "The Pompeiu problem", *Appl. Anal.* **64**:1-2 (1997), 19–26. MR Zbl

[Ramm 2017] A. G. Ramm, "Solution to the Pompeiu problem and the related symmetry problem", *Appl. Math. Lett.* **63** (2017), 28–33. MR Zbl

[Rédei 1965] L. Rédei, "Die neue Theorie der endlichen abelschen Gruppen und Verallgemeinerung des Hauptsatzes von Hajós", *Acta Math. Acad. Sci. Hungar.* **16**:3-4 (1965), 329–373. MR Zbl

[Schmidt 2002] B. Schmidt, *Characters and cyclotomic fields in finite geometry*, Lecture Notes in Math. **1797**, Springer, 2002. MR Zbl

[Shi 2019] R. Shi, "Fuglede's conjecture holds on cyclic groups $\mathbb{Z}_{pqr}$", *Discrete Anal.* **2019** (2019), art. id. 14. MR Zbl

[Steinberg 2012] B. Steinberg, *Representation theory of finite groups: an introductory approach*, Springer, 2012. MR Zbl

[Székelyhidi 2001] L. Székelyhidi, "A Wiener Tauberian theorem on discrete abelian torsion groups", *Ann. Acad. Paedagog. Crac. Stud. Math.* **4** (2001), 147–150. Zbl

[Tao 2004] T. Tao, "Fuglede's conjecture is false in 5 and higher dimensions", *Math. Res. Lett.* **11**:2 (2004), 251–258. MR Zbl

[Zalcman 1992] L. Zalcman, "A bibliographic survey of the Pompeiu problem", pp. 185–194 in *Approximation by solutions of partial differential equations* (Hanstholm, Denmark, 1991), edited by B. Fuglede et al., NATO Adv. Sci. Inst. Ser. C **365**, Kluwer, Dordrecht, 1992. MR Zbl

[Zeilberger 1978] D. Zeilberger, "Pompeiu's problem on discrete space", *Proc. Nat. Acad. Sci. U.S.A.* **75**:8 (1978), 3555–3556. MR Zbl

Received 10 Jul 2018. Revised 2 Mar 2019. Accepted 8 Apr 2019.

GERGELY KISS: kigergo57@gmail.com
Faculty of Science, Mathematics Research Unit, University of Luxembourg, Luxembourg, Luxembourg

ROMANOS DIOGENES MALIKIOSIS: rwmanos@gmail.com
Department of Mathematics, Aristotle University of Thessaloniki, Thessaloniki, Greece

GÁBOR SOMLAI: gsomlai@cs.elte.hu
MTA-ELTE Geometric and Algebraic Combinatorics Research Group NKFIH 115799 and, Faculty of Science, Institute of Mathematics, Eötvös Loránd University, Budapest, Hungary

MÁTÉ VIZER: vizermate@gmail.com
Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary

ANALYSIS AND PDE
Vol. 13, No. 3, 2020
dx.doi.org/10.2140/apde.2020.13.789

ENERGY CONSERVATION FOR THE COMPRESSIBLE EULER AND NAVIER–STOKES EQUATIONS WITH VACUUM

IBROKHIMBEK AKRAMOV, TOMASZ DĘBIEC, JACK SKIPPER AND EMIL WIEDEMANN

We consider the compressible isentropic Euler equations on $[0, T] \times \mathbb{T}^d$ with a pressure law $p \in C^{1,\gamma-1}$, where $1 \leq \gamma < 2$. This includes all physically relevant cases, e.g., the monoatomic gas. We investigate under what conditions on its regularity a weak solution conserves the energy. Previous results have crucially assumed that $p \in C^2$ in the range of the density; however, for realistic pressure laws this means that we must exclude the vacuum case. Here we improve these results by giving a number of sufficient conditions for the conservation of energy, even for solutions that may exhibit vacuum: firstly, by assuming the velocity to be a divergence-measure field; secondly, imposing extra integrability on $1/\rho$ near a vacuum; thirdly, assuming ρ to be quasinearly subharmonic near a vacuum; and finally, by assuming that u and ρ are Hölder continuous. We then extend these results to show global energy conservation for the domain $[0, T] \times \Omega$ where Ω is bounded with a C^2 boundary. We show that we can extend these results to the compressible Navier–Stokes equations, even with degenerate viscosity.

1. Introduction

In recent years some substantial effort has been directed towards investigating the relation between energy (or, more generally, entropy) conservation and regularity of weak solutions to a given physical system of equations.

Onsager's conjecture states that a weak solution of the (three-dimensional) incompressible Euler system will conserve energy if it is Hölder regular with exponent greater than $\frac{1}{3}$. Otherwise it is possible for solutions to exist where anomalous dissipation of energy occurs. First results towards energy conservation for weak solutions are due to Eyink [1994] and Constantin, E, and Titi [Constantin et al. 1994]; see also [Duchon and Robert 2000]. The sharpest results in optimal Besov spaces are due to Cheskidov et al. [2008] and Fjordholm and Wiedemann [2018]. Further, Bardos and Titi [2018], Bardos, Titi, and Wiedemann [Bardos et al. 2018], and Drivas and Nguyen [2018] have extended these results to consider solutions on a bounded domain.

Investigating the possibility of analogous statements for other systems has become another lively direction of research. Sufficient regularity conditions for the energy to be conserved were studied for a number of models: inhomogeneous incompressible Euler [Chen and Yu 2019] and Navier–Stokes [Leslie and Shvydkoy 2016], compressible Euler [Feireisl et al. 2017], the full Euler system [Drivas and Eyink 2018], compressible Navier–Stokes [Yu 2017], and Euler–Korteweg [Dębiec et al. 2018]. A general class of first-order conservation laws was considered in [Gwiazda et al. 2018], and in [Bardos et al. 2019] on bounded domains.

MSC2010: primary 35Q31; secondary 35Q30, 35L65, 76N10.
Keywords: compressible Euler equations, compressible Navier–Stokes equations, vacuum, Onsager's conjecture, energy conservation.

Another direction of research was aimed towards the construction of $\left(\frac{1}{3}-\varepsilon\right)$-Hölder continuous solutions to the incompressible Euler system that do *not* conserve energy. With the application, and further refinements, of the method of convex integration this was achieved recently in [Isett 2018; Buckmaster et al. 2019]. Thus the famous conjecture of Lars Onsager for the incompressible Euler equations is fully resolved.

One of the major differences between incompressible and compressible fluid dynamics is the possible formation of *vacuum* in the latter case. This means that the density of the fluid may become zero in some region. More precisely, consider the isentropic compressible Euler system

$$\begin{aligned} \partial_t(\rho u) + \operatorname{div}(\rho u \otimes u) + \nabla p(\rho) &= 0, \\ \partial_t \rho + \operatorname{div}(\rho u) &= 0, \end{aligned} \tag{1-1}$$

where u denotes the velocity and ρ the density of the fluid. We will specify the constitutive pressure law $p = p(\rho)$ later. It is classically known that conservation laws like (1-1) may develop singularities (shocks) in finite time, which prohibits the use of a smooth notion of solution. Rather, one works with solutions in the sense of distributions, which may be very rough. Suppose now the density were initially bounded away from zero, $\rho^0 \geq c > 0$. If the solution were smooth, then from the continuity equation $\partial_t \rho + \operatorname{div}(\rho u) = 0$ it would easily follow (see equation (7) in [DiPerna and Lions 1989]) that ρ remains bounded away from zero for all times. More precisely, this requires u to have bounded divergence. However, there seems to be no way to guarantee that the velocity component of a weak solution of (1-1) has bounded divergence, and thus it cannot be excluded that the solution spontaneously develops vacuum in finite time. In fact, to our knowledge it remains an outstanding open question whether this can actually occur for the compressible Euler or even Navier–Stokes equations.

The formation of vacuum constitutes a degeneracy that, in many situations, vastly complicates the mathematical analysis of compressible models. For instance, the compressible Euler equations cease to be strictly hyperbolic in vacuum regions. In the context of the current contribution, densities close to zero invalidate the methods and results from previous works like [Feireisl et al. 2017; Gwiazda et al. 2018; Bardos et al. 2019]: There, it is a crucial assumption that the nonlinearities depend on the dependent variables in a twice continuously differentiable fashion, in order to treat them like a quadratic expression in the commutator estimates. For the system (1-1), a typical and physically reasonable pressure law would be the polytropic one, i.e., $p(\rho) = \rho^\gamma$ with $\gamma > 1$. The second derivative, however, is of order $\rho^{\gamma-2}$ and thus blows up at zero, at least if $\gamma < 2$. But the regime $1 < \gamma < 2$ is precisely the relevant one $\big($for instance, a monoatomic gas has $\gamma = \frac{5}{3}\big)$.

The starting point of our current work is the result of Feireisl, Gwiazda, Świerczewska-Gwiazda, and Wiedemann [Feireisl et al. 2017] for the compressible Euler system, which we quote below. It gives sufficient conditions, in terms of Besov regularity of a weak solution, for energy conservation, but only as long as vacuum is excluded. In the presence of vacuum, the relevant commutator estimate involving the pressure completely breaks down, and it turns out that substantially new techniques are required to fix this. To our knowledge, the only other result on energy conservation for non-C^2 nonlinearities is the one on active scalar equations [Akramov and Wiedemann 2019], using however different techniques.

In the current article, we give a number of sufficient conditions to ensure energy conservation even after possible formation of vacuum.

First (Section 3), we consider the condition that the velocity be a so-called divergence-measure field; this notion is well known in geometric measure theory and hyperbolic conservation laws, but it may seem a bit unmotivated to consider in the present situation. However, justification comes from the compressible Navier–Stokes system, whose a priori estimates ensure this condition. We extensively discuss the ramifications of our result with respect to the Navier–Stokes equations in Section 3A, where we also compare it to recent work of Cheng Yu [2017].

In Section 4, we identify as a sufficient condition for energy conservation an estimate for the quotient between the density and its mollification; see (4-1). This, in itself, may seem rather artificial, and we go on to identify more natural conditions that will ensure (4-1) holds. Arguably, our strongest result is Corollary 4.4: under the slightly stronger assumption of Hölder (instead of Besov) regularity, but with the expected exponents, we can show energy conservation *no matter how the density behaves near vacuum*. It is surprising that this result is completely agnostic to the way that ρ approaches zero. It crucially relies on a new measure-theoretic observation (Lemma 4.3) that may be of independent interest.

If one does want to assume only Besov regularity, then one needs to make further assumptions on the density near vacuum; we show that energy is conserved provided the density descends into vacuum sufficiently fast (Corollary 4.7) or sufficiently slowly (Corollary 4.11).

Finally, in Section 5 we demonstrate how to extend our results, so far shown only under periodic boundary conditions, to the case of a bounded domain.

1A. *The result of Feireisl et al.* To formulate the local or global energy equality for (1-1) it is useful to define the so-called pressure potential by

$$P(\rho)=\rho\int_1^\rho \frac{p(r)}{r^2}\,\mathrm{d}r.$$

The following theorem was proven in [Feireisl et al. 2017, Theorem 4.1].

Theorem 1.1. *Let ρ, u be a solution of* (1-1) *in the sense of distributions. Assume*

$$u\in B_3^{\alpha,\infty}((0,T)\times\mathbb{T}^d),\quad \rho,\rho u\in B_3^{\beta,\infty}((0,T)\times\mathbb{T}^d),\quad 0\le\underline{\rho}\le\rho\le\bar{\rho}\ \text{ a.e. in } (0,T)\times\mathbb{T}^d$$

for some constants $\underline{\rho}$, $\bar{\rho}$, and $0\le\alpha,\beta\le 1$ such that

$$\beta>\max\bigl\{1-2\alpha,\tfrac12(1-\alpha)\bigr\}.$$

Assume further that $p\in C^2[\underline{\rho},\bar{\rho}]$, and in addition

$$p'(0)=0\quad\text{as soon as } \underline{\rho}=0.$$

Then the energy is locally conserved; *i.e.*,

$$\partial_t\bigl(\tfrac12\rho|u|^2+P(\rho)\bigr)+\operatorname{div}\bigl[\bigl(\tfrac12\rho|u|^2+p(\rho)+P(\rho)\bigr)u\bigr]=0 \tag{1-2}$$

in the sense of distributions on $(0,T)\times\mathbb{T}^d$.

Our aim in the current paper is to improve the above theorem by relaxing the C^2 assumption on the pressure. This will allow one, for instance, to apply the theorem in the physically relevant case of the isentropic pressure law $p(\rho) = \kappa\rho^\gamma$ with the adiabatic coefficient $\gamma \in (1, 2)$, without excluding vacuum.

2. Preliminaries

2A. ***Function spaces.*** For $\Omega := (0, T) \times \mathbb{T}^d$ we recall the Besov space $B_p^{\alpha,\infty}(\Omega)$, which is the space of tempered distributions w for which the norm

$$\|w\|_{B_p^{\alpha,\infty}(\Omega)} := \|w\|_{L^p(\Omega)} + \sup_{\xi\in\Omega} \frac{\|w(\cdot + \xi) - w\|_{L^p(\Omega\cap(\Omega-\xi))}}{|\xi|^\alpha} \tag{2-1}$$

is finite. The above norm provides a control over shifts of the distribution w, making Besov spaces a convenient environment for our analysis, as it relies on convolutions with a mollifying kernel.

Let $\eta \in C_c^\infty(\mathbb{R}^N)$ be a positive, radial function of integral 1 with

$$\eta(x) = \begin{cases} 1 & \text{for } |x| \le \frac{1}{3}, \\ 0 & \text{for } |x| \ge 1, \end{cases}$$

and for $N = 1 + d$ set

$$\eta^\varepsilon(x) = \frac{1}{\varepsilon^N}\eta\left(\frac{x}{\varepsilon}\right).$$

We define the notation $w^\varepsilon := \eta^\varepsilon * w$. For any function w, w^ε is well-defined on $\Omega^\varepsilon = \{x \in \Omega : \mathrm{d}(x, \partial\Omega) > \varepsilon\}$.

It is then easy to check that the definition of the Besov spaces implies

$$\|w^\varepsilon - w\|_{L^p(\Omega^\varepsilon)} \le C\varepsilon^\alpha \|w\|_{B_p^{\alpha,\infty}(\Omega)}$$

and

$$\|\nabla w^\varepsilon\|_{L^p(\Omega^\varepsilon)} \le C\varepsilon^{\alpha-1} \|w\|_{B_p^{\alpha,\infty}(\Omega)}.$$

By $\mathscr{M}(\Omega)$ we denote the space of signed Radon measures equipped with the total variation norm

$$\|\mu\|_{TV} := \int_\Omega \mathrm{d}|\mu|.$$

2B. ***Derivation of the local energy equality.*** The starting point in the proof of Theorem 1.1, as well as all our results, is to mollify the Euler equations, then derive the local energy equality for the regularized quantities, and finally estimate commutator errors generated by nonlinear terms. As this strategy is a common part in the proofs of our theorems, we devote this section to the said derivation, omitting the details of passing to the limit under the assumptions of Theorem 1.1.

We begin by mollifying the momentum equation in time and space to obtain

$$\partial_t(\rho u)^\varepsilon + \operatorname{div}(\rho u \otimes u)^\varepsilon + \nabla p^\varepsilon(\rho) = 0, \tag{2-2}$$

or, in terms of commutators

$$\begin{aligned} &\partial_t(\rho^\varepsilon u^\varepsilon) + \operatorname{div}((\rho u)^\varepsilon \otimes u^\varepsilon) + \nabla p(\rho^\varepsilon) \\ &\qquad = \partial_t(\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon) + \operatorname{div}((\rho u)^\varepsilon \otimes u^\varepsilon - (\rho u \otimes u)^\varepsilon) + \nabla(p(\rho^\varepsilon) - p^\varepsilon(\rho)). \end{aligned} \tag{2-3}$$

Making use of the identity

$$\operatorname{div}((\rho u)^\varepsilon \otimes u^\varepsilon) = u^\varepsilon \operatorname{div}(\rho u)^\varepsilon + ((\rho u)^\varepsilon \cdot \nabla) u^\varepsilon,$$

we can see that multiplying (2-3) by u^ε yields

$$\rho^\varepsilon \partial_t\big(\tfrac{1}{2}|u^\varepsilon|^2\big) + ((\rho u)^\varepsilon \cdot \nabla)\tfrac{1}{2}|u^\varepsilon|^2 + \rho^\varepsilon u^\varepsilon \,\nabla(P'(\rho^\varepsilon)) = r_1^\varepsilon + r_2^\varepsilon + r_3^\varepsilon, \tag{2-4}$$

where

$$r_1^\varepsilon = \partial_t(\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon)\cdot u^\varepsilon,$$
$$r_2^\varepsilon = \operatorname{div}((\rho u)^\varepsilon \otimes u^\varepsilon - (\rho u \otimes u)^\varepsilon)\cdot u^\varepsilon,$$
$$r_3^\varepsilon = \nabla(p(\rho^\varepsilon) - p^\varepsilon(\rho))\cdot u^\varepsilon.$$

Using the mollified continuity equation

$$\partial_t \rho^\varepsilon + \operatorname{div}(\rho u)^\varepsilon = 0 \tag{2-5}$$

multiplied by $\frac{1}{2}|u^\varepsilon|^2$, we can rewrite (2-4) as

$$\partial_t\big(\tfrac{1}{2}\rho^\varepsilon|u^\varepsilon|^2\big) + \operatorname{div}\big((\rho u)^\varepsilon\, \tfrac{1}{2}|u^\varepsilon|^2\big) + \rho^\varepsilon u^\varepsilon\,\nabla(P'(\rho^\varepsilon)) = r_1^\varepsilon + r_2^\varepsilon + r_3^\varepsilon. \tag{2-6}$$

On the other hand writing (2-5) in the form

$$\partial_t \rho^\varepsilon + \operatorname{div}(\rho^\varepsilon u^\varepsilon) = \operatorname{div}(\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon)$$

and multiplying by $P'(\rho^\varepsilon)$ we get

$$\partial_t(P(\rho^\varepsilon)) + \operatorname{div}(\rho^\varepsilon u^\varepsilon) P'(\rho^\varepsilon) = \operatorname{div}(\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon)\; P'(\rho^\varepsilon). \tag{2-7}$$

Combining (2-6) and (2-7) we obtain

$$\partial_t\big(\tfrac{1}{2}\rho^\varepsilon|u^\varepsilon|^2 + P(\rho^\varepsilon)\big) + \operatorname{div}\big((\rho u)^\varepsilon\, \tfrac{1}{2}|u^\varepsilon|^2 + \rho^\varepsilon u^\varepsilon P'(\rho^\varepsilon)\big) = r_1^\varepsilon + r_2^\varepsilon + r_3^\varepsilon + s^\varepsilon, \tag{2-8}$$

where we set

$$s^\varepsilon := \operatorname{div}(\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon)\; P'(\rho^\varepsilon).$$

The proof of Theorem 4.1 in [Feireisl et al. 2017] shows that when ρ, u are Besov regular and p is of class C^2, the left-hand side of (2-8) converges to the left-hand side of (1-2) and each term on the right-hand side of (2-8) converges to zero, where each convergence is in the sense of distributions.

3. Energy conservation assuming the divergence of velocity is a bounded measure

Our first result establishes local energy conservation for weak solutions of (1-1) under the additional assumption that the velocity field u is a divergence-measure field.

Remark 3.1. See [Chen and Torres 2005] for details on the role of divergence-measure fields in the theory of hyperbolic conservation laws.

Theorem 3.2. *Let ρ, u be a solution of* (1-1) *in the sense of distributions. Assume*

$$u \in B_3^{\alpha,\infty}((0,T)\times\mathbb{T}^d), \quad \rho, \rho u \in B_3^{\beta,\infty}((0,T)\times\mathbb{T}^d), \quad 0 \le \underline{\rho} \le \rho \le \bar{\rho} \ \text{a.e. in } (0,T)\times\mathbb{T}^d$$

for some constants $\underline{\rho}$, $\bar{\rho}$, and $0 \le \alpha, \beta \le 1$ such that

$$\beta > \max\bigl\{1-2\alpha, \tfrac{1}{2}(1-\alpha)\bigr\}.$$

Assume further that

$$\operatorname{div} u \in \mathscr{M}((0,T)\times\mathbb{T}^d) \quad \text{and} \quad p \in C[\underline{\rho}, \bar{\rho}].$$

Then the energy is locally conserved; *i.e.*,

$$\partial_t\bigl(\tfrac{1}{2}\rho|u|^2 + P(\rho)\bigr) + \operatorname{div}\bigl[\bigl(\tfrac{1}{2}\rho|u|^2 + p(\rho) + P(\rho)\bigr)u\bigr] = 0$$

in the sense of distributions on $(0,T)\times\mathbb{T}^d$.

Proof. Take a sequence $p^\delta \in C^2[\underline{\rho}, \bar{\rho}]$ that converges uniformly to $p \in C[\underline{\rho}, \bar{\rho}]$; that is, for each $\delta > 0$,

$$\|p - p^\delta\|_{L^\infty} \le \delta.$$

Then using p^δ in (2-2) we have

$$\partial_t(\rho u)^\varepsilon + \operatorname{div}(\rho u \otimes u)^\varepsilon + \nabla(p^\delta(\rho))^\varepsilon = \nabla[(p^\delta(\rho))^\varepsilon - p^\varepsilon(\rho)]. \tag{3-1}$$

Now the left-hand side of the last equality satisfies all the conditions of Theorem 1.1, so for each fixed $\delta > 0$ we have, in the limit as $\varepsilon \to 0$,

$$\partial_t\bigl(\tfrac{1}{2}\rho|u|^2 + P^\delta(\rho)\bigr) + \operatorname{div}\bigl[\bigl(\tfrac{1}{2}\rho|u|^2 + p^\delta(\rho) + P^\delta(\rho)\bigr)u\bigr], \tag{3-2}$$

where

$$P^\delta(\rho) := \rho \int_1^\rho \frac{p^\delta(r)}{r^2}\,\mathrm{d}r.$$

We will now show that (3-2) converges as $\delta \to 0$ in the sense of distributions on $(0,T)\times\mathbb{T}^d$ to

$$\partial_t\bigl(\tfrac{1}{2}\rho|u|^2 + P(\rho)\bigr) + \operatorname{div}\bigl[\bigl(\tfrac{1}{2}\rho|u|^2 + p(\rho) + P(\rho)\bigr)u\bigr].$$

Let $\varphi \in C_c^\infty((0,T)\times\mathbb{T}^d)$. From the choice of p^δ we have

$$\left|\int_0^T\int_{\mathbb{T}^d} \nabla\varphi \cdot (p^\delta(\rho) - p(\rho))u \,\mathrm{d}x\,\mathrm{d}t\right| \le C\|\varphi\|_{\mathscr{C}^1}\|p^\delta - p\|_{L^\infty}\|u\|_{L^3} \le C(\varphi, u)\delta.$$

For the terms containing $P^\delta(\rho)$ notice that

$$|P^\delta(\rho) - P(\rho)| \le \rho \int_1^\rho \frac{|p^\delta(r) - p(r)|}{r^2}\,\mathrm{d}r \le \|p^\delta - p\|_{L^\infty}\,\rho\left|\int_1^\rho \frac{1}{r^2}\,\mathrm{d}r\right| \le (1+\rho)\|p^\delta - p\|_{L^\infty}.$$

Hence we can estimate

$$\left|\int_0^T\int_{\mathbb{T}^d} \partial_t\varphi\,(P^\delta(\rho) - P(\rho))\,\mathrm{d}x\,\mathrm{d}t\right| \le C\|\varphi\|_{\mathscr{C}^1}(1 + \|\rho\|_{L^1})\delta \le C(\varphi)\delta,$$

and similarly for the divergence term. It follows that both terms of (3-2) containing P^δ converge as $\delta \to 0$ to the corresponding terms for P.

The final step of the proof is to consider the term coming into (2-8) from the right-hand side of (3-1). We need to show that

$$\nabla[(p^\delta(\rho))^\varepsilon - p^\varepsilon(\rho)] \cdot u^\varepsilon$$

converges to zero in the sense of distributions on $(0,T)\times\mathbb{T}^d$ as first ε and then δ tend to zero. Multiplying by $\varphi \in C_c^\infty((0,T)\times\mathbb{T}^d)$, integrating over time and space, and integrating by parts we obtain

$$\begin{aligned}\int_0^T\int_{\mathbb{T}^d} &\nabla[(p^\delta(\rho))^\varepsilon - p^\varepsilon(\rho)]\varphi u^\varepsilon \,\mathrm{d}x\,\mathrm{d}t \\ &= -\int_0^T\int_{\mathbb{T}^d}[(p^\delta(\rho))^\varepsilon - p^\varepsilon(\rho)]\varphi \operatorname{div} u^\varepsilon \,\mathrm{d}x\,\mathrm{d}t - \int_0^T\int_{\mathbb{T}^d}[(p^\delta(\rho))^\varepsilon - p^\varepsilon(\rho)]\nabla\varphi\cdot u^\varepsilon \,\mathrm{d}x\,\mathrm{d}t. \end{aligned}\tag{3-3}$$

For the second term on the right-hand side of the last equality we see that

$$\begin{aligned}\left|\int_0^T\int_{\mathbb{T}^d}[(p^\delta(\rho))^\varepsilon - p^\varepsilon(\rho)]\nabla\varphi\cdot u^\varepsilon \,\mathrm{d}x\,\mathrm{d}t\right| &= \left|\int_0^T\int_{\mathbb{T}^d}[p^\delta(\rho) - p(\rho)]^\varepsilon\nabla\varphi\cdot u^\varepsilon \,\mathrm{d}x\,\mathrm{d}t\right| \\ &\le C\|\varphi\|_{C^1}\|(p^\delta - p)^\varepsilon\|_{L^\infty}\|u\|_{L^3} \\ &\le C\|\varphi\|_{C^1}\|p^\delta - p\|_{L^\infty}\|u\|_{L^3} \le C\delta.\end{aligned}$$

Finally, for the first term on the right-hand side of (3-3) we invoke the assumption that $\operatorname{div} u$ is a bounded Radon measure to see that

$$\begin{aligned}\left|\int_0^T\int_{\mathbb{T}^d}\varphi[(p^\delta(\rho))^\varepsilon - p^\varepsilon(\rho)]\operatorname{div} u^\varepsilon \,\mathrm{d}x\,\mathrm{d}t\right| &= \left|\int_0^T\int_{\mathbb{T}^d}\varphi[p^\delta(\rho) - p(\rho)]^\varepsilon(\operatorname{div} u)^\varepsilon \,\mathrm{d}x\,\mathrm{d}t\right| \\ &\le \|\varphi\|_{C^0}\|(p^\delta - p)^\varepsilon\|_{L^\infty}\|(\operatorname{div} u)^\varepsilon\|_{L^1} \\ &\le \|\varphi\|_{C^0}\|p^\delta - p\|_{L^\infty}\|\operatorname{div} u\|_{TV} \le C\delta\end{aligned}$$

and so we are done. □

3A. ***Application to the compressible Navier–Stokes equations.*** When studying the result of Theorem 3.2 we see that the condition $\operatorname{div} u \in \mathscr{M}((0,T)\times\mathbb{T}^d)$ is quite a strong assumption for solutions to the compressible Euler equations; however, it is given for the compressible Navier–Stokes equations where one obtains a priori from the diffusion term that $u \in L^2(0,T;H^1)$. Therefore a natural question to ask is what happens when we consider the solutions to the compressible Navier–Stokes equations with vacuum, and how these results relate to the current results in [Yu 2017].

The compressible Navier–Stokes equations are given by

$$\begin{gathered}\partial_t(\rho u) + \operatorname{div}(\rho u\otimes u) + \nabla p(\rho) = \operatorname{div}\mathbb{S}(\nabla u), \\ \partial_t\rho + \operatorname{div}(\rho u) = 0, \\ \mathbb{S}(\nabla u) := \mu\big(\nabla u + (\nabla u)^T - \tfrac{2}{3}\operatorname{div} u\mathbb{I}\big) + \nu\operatorname{div} u\mathbb{I},\end{gathered}\tag{3-4}$$

where we have the constants $\mu > 0$ and $\nu \geq 0$. Here we will use the main properties that $\mathbb{S}(\nabla u)$ is symmetric and positive definite. For degenerate viscosity, the momentum equation becomes, instead,

$$\partial_t(\rho u) + \operatorname{div}(\rho u \otimes u) + \nabla p(\rho) = \operatorname{div}(\rho \mathbb{S}(\nabla u)). \tag{3-5}$$

Corollary 3.3. *Let ρ, u be a solution of* (3-4) *or* (3-5) *in the sense of distributions. Assume*

$$u \in B_3^{\alpha,\infty}((0,T)\times\mathbb{T}^d), \quad u \in L^2(0,T;H^1(\mathbb{T}^d)), \quad \rho, \rho u \in B_3^{\beta,\infty}((0,T)\times\mathbb{T}^d),$$
$$0 \leq \underline{\rho} \leq \rho \leq \bar{\rho} \ \text{ a.e. in } (0,T)\times\mathbb{T}^d$$

for some constants $\underline{\rho}$, $\bar{\rho}$, and $0 \leq \alpha, \beta \leq 1$ such that

$$\beta > \max\bigl\{1-2\alpha, \tfrac{1}{2}(1-\alpha)\bigr\}. \tag{3-6}$$

Assume further that $p \in C[\underline{\rho}, \bar{\rho}]$. Then the energy is locally conserved; *i.e.*,

$$\partial_t\bigl(\tfrac{1}{2}\rho|u|^2 + P(\rho)\bigr) + \mathbb{S}(\nabla u) : \nabla u + \operatorname{div}\bigl[\bigl(\tfrac{1}{2}\rho|u|^2 + p(\rho) + P(\rho) + \mathbb{S}(\nabla u)\bigr)u\bigr] = 0 \tag{3-7}$$

for (3-4) *and*

$$\partial_t\bigl(\tfrac{1}{2}\rho|u|^2 + P(\rho)\bigr) + \rho\mathbb{S}(\nabla u) : \nabla u + \operatorname{div}\bigl[\bigl(\tfrac{1}{2}\rho|u|^2 + p(\rho) + P(\rho) + \rho\mathbb{S}(\nabla u)\bigr)u\bigr] = 0 \tag{3-8}$$

for (3-5), *in the sense of distributions on $(0,T)\times\mathbb{T}^d$.*

Remark 3.4. The condition $\operatorname{div} u \in \mathscr{M}$ is trivially satisfied if we assume that $u \in L^2(0,T;H^1)$ and so does not appear in the statement of Corollary 3.3.

Remark 3.5. For $d \leq 3$ we can use Besov embedding theorems, see [Bahouri et al. 2011], to observe that $H^1 \hookrightarrow B_2^{1,\infty} \hookrightarrow B_3^{2/3,\infty}$ and so assuming that $u \in B_3^{\alpha_1,\infty}(0,T;B_3^{\alpha_2,\infty})$ and $\rho, \rho u \in B_3^{\beta_1,\infty}(0,T;B_3^{\beta_2,\infty})$ we have the same assumptions on the pairs (α_1,β_1) and (α_2,β_2) as (3-6) but can assume that $\alpha_2 \geq \frac{2}{3}$ and remove the assumption that $u \in L^2(0,T;H^1)$.

Remark 3.6. We have assumed that the density ρ is bounded above to simplify the proof, though this is not necessary. Indeed, we can assume that for some $C > 0$, $p^\delta(r) = p(r)$ for $r \geq C$ and so still obtain uniform convergence of p^δ to p for unbounded density.

Proof. We only have to consider the extra term $\operatorname{div}\mathbb{S}(\nabla u)$ in the derivation of the local energy equality that we performed previously. We see that

$$-\int_0^T\!\!\int_{\mathbb{T}^d} \operatorname{div}\mathbb{S}(\nabla u^\varepsilon)\cdot u^\varepsilon\varphi\,\mathrm{d}x\,\mathrm{d}t = \int_0^T\!\!\int_{\mathbb{T}^d} \mathbb{S}(\nabla u^\varepsilon) : \nabla u^\varepsilon\varphi\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\!\!\int_{\mathbb{T}^d} (\mathbb{S}(\nabla u^\varepsilon)u^\varepsilon)\cdot\nabla\varphi\,\mathrm{d}x\,\mathrm{d}t$$

and so obtain (3-7). For (3-8) we perform the same calculation as above; however, with an extra ρ in the equation, the diffusion term is no longer linear and thus we pick up an extra commutator estimate

$$r_d^\varepsilon := \int_0^T\!\!\int_{\mathbb{T}^d} \operatorname{div}(\rho^\varepsilon\mathbb{S}(\nabla u^\varepsilon) - (\rho\mathbb{S}(\nabla u))^\varepsilon)\cdot\varphi u^\varepsilon\,\mathrm{d}x\,\mathrm{d}t.$$

We can perform an integration by parts to obtain

$$|r_d^\varepsilon| \le \left| \int_0^T \int_{\mathbb{T}^d} [(\rho^\varepsilon \mathbb{S}(\nabla u^\varepsilon) - (\rho \mathbb{S}(\nabla u))^\varepsilon) u^\varepsilon] \cdot \nabla\varphi \, \mathrm{d}x \, \mathrm{d}t \right| + \left| \int_0^T \int_{\mathbb{T}^d} (\rho^\varepsilon \mathbb{S}(\nabla u^\varepsilon) - (\rho \mathbb{S}(\nabla u))^\varepsilon) : \nabla u^\varepsilon \varphi \, \mathrm{d}x \, \mathrm{d}t \right|. \tag{3-9}$$

Note the pointwise identity where for any two functions f, g we have

$$f^\varepsilon g^\varepsilon - (fg)^\varepsilon = (f^\varepsilon - f)(g^\varepsilon - g) - \int_{-\varepsilon}^{\varepsilon} \int_{\mathbb{T}^d} \eta^\varepsilon(\tau, \xi)(f(t-\tau, x-\xi) - f(t,x))(g(t-\tau, x-\xi) - g(t,x)) \, \mathrm{d}\xi \, \mathrm{d}\tau. \tag{3-10}$$

Applying this allows us to split the two terms on the right-hand side of (3-9) into four more terms which we can estimate. We focus on the first of these terms only, as the other terms produce the same estimates, after applying Fubini's theorem, as seen in [Feireisl et al. 2017]. We see that

$$\begin{aligned} |r_d^\varepsilon| &\le \left| \int_0^T \int_{\mathbb{T}^d} [(\rho^\varepsilon - \rho)(\mathbb{S}(\nabla u^\varepsilon) - \mathbb{S}(\nabla u)) u^\varepsilon] \cdot \nabla\varphi \, \mathrm{d}x \, \mathrm{d}t \right| + \left| \int_0^T \int_{\mathbb{T}^d} (\rho^\varepsilon - \rho)(\mathbb{S}(\nabla u^\varepsilon) - \mathbb{S}(\nabla u)) : \nabla u^\varepsilon \varphi \, \mathrm{d}x \, \mathrm{d}t \right| \\ &\le \|\varphi\|_{C^1} \|\rho\|_{L^\infty} \|u\|_{L^2} \|\mathbb{S}(\nabla u^\varepsilon) - \mathbb{S}(\nabla u)\|_{L^2} + \|\varphi\|_{C^0} \|\rho\|_{L^\infty} \|\nabla u\|_{L^2} \|\mathbb{S}(\nabla u^\varepsilon) - \mathbb{S}(\nabla u)\|_{L^2}. \end{aligned}$$

Using the a priori estimate that $u \in L^2(0, T; H^1)$ we see that $\|\mathbb{S}(\nabla u^\varepsilon) - \mathbb{S}(\nabla u)\|_{L^2} \to 0$ as $\varepsilon \to 0$ and thus $r_d^\varepsilon \to 0$ as $\varepsilon \to 0$. □

The work of Cheng Yu [2017] also studies energy conservation for the compressible Navier–Stokes systems where a vacuum could occur. The result in [Yu 2017] treats the case where $p(\rho) = \rho^\gamma$ for $\gamma > 1$ and thus where $p \in C^{1,\gamma-1}$, with strong assumptions of spacial regularity where

$$\sqrt{\rho}\nabla u \in L^2(0, T; L^2(\Omega)) \quad \text{and} \quad \frac{\nabla\rho}{\sqrt{\rho}} \in L^\infty(0, T; L^2(\Omega)),$$

among other assumptions. However, [Yu 2017] only assumes integrability in time. The condition $\nabla\rho/\sqrt{\rho} \in L^\infty(0, T; L^2(\Omega))$ restricts the allowable vacuum cases and will only allow vacuum on measure-zero sets with a nice approach to this set. The result presented here complements the result in [Yu 2017] as we show that by assuming some differential regularity in time for both ρ and u then we can weaken the spacial regularity assumptions and only need continuity of the pressure p. Specifically, we can have vacuum on measurable subsets of the domain where the approach to this set can be quite generic.

4. Energy conservation assuming Hölder continuity of the pressure

For the next result we fix $1 < \gamma < 2$ and we will assume that the pressure p is of class $C^{1,(\gamma-1)}$, thus relaxing the regularity assumption of Theorem 1.1. The expense of this relaxation is that we require $\alpha + \gamma\beta > 1$ where before we only needed $\alpha + 2\beta > 1$.

Theorem 4.1. *Let ρ, u be a solution of* (1-1) *in the sense of distributions. Assume*

$$u \in B_p^{\alpha,\infty}((0,T) \times \mathbb{T}^d), \quad \rho, \rho u \in B_q^{\beta,\infty}((0,T) \times \mathbb{T}^d), \quad 0 \le \underline{\rho} \le \rho \le \bar{\rho} \ \text{a.e. in } (0,T) \times \mathbb{T}^d$$

for some constants $\underline{\rho}$, $\bar{\rho}$ *and* $0 \le \alpha, \beta \le 1$ *such that*

$$\frac{1}{p}+\frac{2}{q}\le 1, \quad \frac{2}{p}+\frac{1}{q}\le 1, \quad p,q\ge 2, \quad \alpha+\gamma\beta>1, \quad \text{and} \quad 2\alpha+\beta>1.$$

Define $\mathscr{B}_{\varepsilon^\beta} := \{x : 0 < \rho^\varepsilon(x) < \varepsilon^\beta \text{ and } \rho \neq 0\}$ *and assume that*

$$\left\| \frac{\rho^\varepsilon-\rho}{\rho^\varepsilon} \right\|_{L^q(\mathscr{B}_{\varepsilon^\beta})} \le C(\rho), \tag{4-1}$$

where C *does not depend on* ε. *Assume further that* $p \in C^{1,(\gamma-1)}([\underline{\rho}, \bar{\rho}])$, *and, in addition*

$$p'(0)=0 \quad \text{as soon as } \underline{\rho}=0.$$

Then the energy is locally conserved; *i.e.*, (1-2) *holds in the sense of distributions on* $(0,T)\times\mathbb{T}^d$.

A large part of the proof of this theorem is identical to the proof of Theorem 1.1. In particular we regularize the balance equations to derive an energy balance for the smooth functions ρ^ε and u^ε. Then we need to show that the corresponding commutator errors vanish in the limit $\varepsilon \to 0$. This is done in the same way as in [Feireisl et al. 2017], the only difference being in the terms involving the pressure. In particular, we will have to estimate an appropriate norm of the difference $p(\rho)^\varepsilon - p(\rho^\varepsilon)$. This will be done by means of the following lemma, which is an adaptation to our present case of the argument in [Feireisl et al. 2017, p. 10]; see also [Gwiazda et al. 2018, Lemma 3.1].

Lemma 4.2. *Let* $\gamma \in (1,2)$ *and* $p \in C^{1,\gamma-1}([a,b])$. *If* $\rho \in B^{\beta,\infty}_{\gamma q}(\Omega;[a,b])$, *then*

$$\|p^\varepsilon(\rho)-p(\rho^\varepsilon)\|_{L^q} \le C\varepsilon^{\gamma\beta}\|\rho\|^\gamma_{B^{\beta,\infty}_{\gamma q}}.$$

Proof. First we note that by the fundamental theorem of calculus

$$\begin{aligned} p(s)-p(s_0) = \int_{s_0}^{s} p'(t)\,\mathrm{d}t &= \int_{s_0}^{s} p'(s_0)\,\mathrm{d}t + \int_{s_0}^{s} p'(t)-p'(s_0)\,\mathrm{d}t \\ &= p'(s_0)(s-s_0) + \int_{s_0}^{s} p'(t)-p'(s_0)\,\mathrm{d}t. \end{aligned}$$

Since $p' \in C^{0,\gamma-1}$, we have

$$\left| \int_{s_0}^{s} p'(t)-p'(s_0)\,\mathrm{d}t \right| \le \int_{s_0}^{s} |p'(t)-p'(s_0)|\,\mathrm{d}t \le C \int_{s_0}^{s} \mathrm{d}t \sup_{t\in[s_0,s]} |t-s_0|^{\gamma-1} \le C|s-s_0|^\gamma.$$

Thus,

$$|p(s)-p(s_0)-p'(s_0)(s-s_0)| \le C|s-s_0|^\gamma.$$

As the constant C is independent of s, s_0 we see that

$$|p(\rho^\varepsilon)-p(\rho)-p'(\rho)(\rho^\varepsilon-\rho)| \le C|\rho-\rho^\varepsilon|^\gamma, \tag{4-2}$$

and similarly,

$$|p(\rho(y))-p(\rho(x))-p'(\rho(x))(\rho(y)-\rho(x))| \le C|\rho(x)-\rho(y)|^\gamma. \tag{4-3}$$

Applying convolution against the function η^ε with respect to y in (4-3) and using Jensen's inequality we obtain

$$|p^\varepsilon(\rho)-p(\rho)-p'(\rho)(\rho^\varepsilon-\rho)| \le C|\rho-\rho(\,\cdot\,)|^\gamma *_y \eta^\varepsilon. \tag{4-4}$$

Combining (4-2) and (4-4) we get

$$|p^\varepsilon(\rho)-p(\rho^\varepsilon)|\le C|\rho-\rho^\varepsilon|^\gamma+C|\rho-\rho(\,\cdot\,)|^\gamma *_y \eta^\varepsilon. \tag{4-5}$$

Taking the L^q norm of both sides of (4-5) for the first term on the right-hand side we see that

$$C\||\rho-\rho^\varepsilon|^\gamma\|_{L^q}=C\|\rho-\rho^\varepsilon\|^\gamma_{L^{\gamma q}}.$$

Finally, for the L^q norm of (4-5) for the second term on the right-hand side by Jensen's inequality and Fubini's theorem we have

$$\begin{aligned}
C\||\rho-\rho(\,\cdot\,)|^\gamma *_y \eta^\varepsilon\|_{L^q} &\le C\biggl(\iint |\rho(x)-\rho(x-y)|^{\gamma q}\,\mathrm{d}x\,\eta_\varepsilon(y)\,\mathrm{d}y\biggr)^{1/q}\\
&= C\biggl(\int \|\rho(\,\cdot\,)-\rho(\,\cdot\,-y)\|^{\gamma q}_{L^{\gamma q}}\eta_\varepsilon(y)\,\mathrm{d}y\biggr)^{1/q}\\
&\le C\sup_y|\eta_\varepsilon(y)|^{1/q}\biggl(\int_{\operatorname{supp}\eta_\varepsilon}\|\rho(\,\cdot\,)-\rho(\,\cdot\,-y)\|^{\gamma q}_{L^{\gamma q}}\,\mathrm{d}y\biggr)^{1/q}\\
&\le C\sup_{y\in\operatorname{supp}\eta_\varepsilon}\|\rho(\,\cdot\,)-\rho(\,\cdot\,-y)\|^{\gamma}_{L^{\gamma q}}.
\end{aligned}$$

Finally, we use the definition of the Besov norm and (2-1) to write

$$\begin{aligned}
\|p^\varepsilon(\rho)-p(\rho^\varepsilon)\|_{L^q} &\le C\bigl(\|\rho^\varepsilon-\rho\|^\gamma_{L^{\gamma q}}+\sup_{s\in\operatorname{supp}\eta^\varepsilon}\|\rho(\,\cdot\,)-\rho(\,\cdot\,-s)\|^\gamma_{L^{\gamma q}}\bigr)\\
&\le C\varepsilon^{\gamma\beta}\|\rho\|^\gamma_{B^{\beta,\infty}_{\gamma q}}+\sup_{s\in\operatorname{supp}\eta^\varepsilon}|s|^{\gamma\beta}\|\rho\|^\gamma_{B^{\beta,\infty}_{\gamma q}}\le C\varepsilon^{\gamma\beta}\|\rho\|^\gamma_{B^{\beta,\infty}_{\gamma q}}.
\end{aligned}$$ □

Proof of Theorem 4.1. As remarked above the only novelty needed to establish the desired result is to estimate commutator errors due to nonlinearity of the pressure. Precisely, we need to show that the local versions of r_3^ε and s^ε, which we will denote by R^ε and S^ε, of (2-8) converge to zero as $\varepsilon\to 0$. For a test function $\varphi\in C_c^\infty((0,T)\times\mathbb{T}^d)$ we define

$$\begin{aligned}
R^\varepsilon &:= \int_0^T\!\!\int_{\mathbb{T}^d}\nabla(p(\rho^\varepsilon)-p(\rho)^\varepsilon)\cdot\varphi u^\varepsilon\,\mathrm{d}x\,\mathrm{d}t, \qquad (4\text{-}6)\\
S^\varepsilon &:= \int_0^T\!\!\int_{\mathbb{T}^d}\varphi\operatorname{div}[\rho^\varepsilon u^\varepsilon-(\rho u)^\varepsilon]\,P'(\rho^\varepsilon)\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}$$

Integrating (4-6) by parts and using Lemma 4.2 we obtain the estimate

$$\begin{aligned}
|R^\varepsilon| &\le \|\varphi\|_{\mathscr{C}^1}\int_0^T\!\!\int_{\mathbb{T}^d}|p(\rho)^\varepsilon-p(\rho)^\varepsilon|(|\nabla u^\varepsilon|+|u^\varepsilon|)\,\mathrm{d}x\,\mathrm{d}t\\
&\le C\|\varphi\|_{\mathscr{C}^1}\|p(\rho^\varepsilon)-p(\rho)^\varepsilon\|_{L^{q/2}}(\|\nabla u^\varepsilon\|_{L^p}+\|u^\varepsilon\|_{L^p})\\
&\le C(\varepsilon^{\gamma\beta+(\alpha-1)}+\varepsilon^{\gamma\beta+\alpha})\|\rho\|^\gamma_{B^{\beta,\infty}_{\gamma q/2}}\|u\|_{B^{\alpha,\infty}_p}\\
&\le C(\varepsilon^{\gamma\beta+(\alpha-1)}+\varepsilon^{\gamma\beta+\alpha})\|\rho\|^\gamma_{B^{\beta,\infty}_{q}}\|u\|_{B^{\alpha,\infty}_p},
\end{aligned}$$

where for the last inequality we used that $\frac{1}{2}\gamma q<q$, so we can embed $B^{\beta,\infty}_q$ into $B^{\beta,\infty}_{\gamma q/2}$.

We now investigate the term S^ε and see that we can integrate by parts to obtain

$$\begin{aligned}|S^\varepsilon| &= \left|\int_0^T\int_{\mathbb{T}^d} \varphi\,\mathrm{div}[\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon]\; P'(\rho^\varepsilon)\,\mathrm{d}x\,\mathrm{d}t\right| \\ &\le \int_0^T\int_{\mathbb{T}^d} |\nabla\varphi\cdot[\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon]\; P'(\rho^\varepsilon)|\,\mathrm{d}x\,\mathrm{d}t + \int_0^T\int_{\mathbb{T}^d} |\varphi[\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon]\cdot\nabla P'(\rho^\varepsilon)|\,\mathrm{d}x\,\mathrm{d}t. \end{aligned} \tag{4-7}$$

We make note of the pointwise identity (3-10) but with f and g replaced by ρ and u respectively, that is,

$$\begin{aligned}\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon &= (\rho^\varepsilon - \rho)(u^\varepsilon - u) \\ &\quad - \int_{-\varepsilon}^{\varepsilon}\int_{\mathbb{T}^d} \eta^\varepsilon(\tau,\xi)(\rho(t-\tau,x-\xi)-\rho(t,x))(u(t-\tau,x-\xi)-u(t,x))\,\mathrm{d}\xi\,\mathrm{d}\tau,\end{aligned}$$

and using (3-10) allows us to split first term on the right-hand side of (4-7) into two terms. Here again we focus on the first of these terms only, as the other one produces the same estimates, after applying Fubini's theorem, as seen in [Feireisl et al. 2017]. We see that

$$\int_0^T\int_{\mathbb{T}^d} |\nabla\varphi\cdot(\rho^\varepsilon-\rho)(u^\varepsilon-u)P'(\rho^\varepsilon)|\,\mathrm{d}x\,\mathrm{d}t \le \|\varphi\|_{C^1}\varepsilon^\beta\|\rho\|_{B_q^{\beta,\infty}}\varepsilon^\alpha\|u\|_{B_p^{\alpha,\infty}}\|P'(\rho^\varepsilon)\|_{L^\infty}.$$

We will now focus on the second term on the right-hand side of (4-7), namely,

$$\int_0^T\int_{\mathbb{T}^d} |\varphi[\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon]\cdot\nabla P'(\rho^\varepsilon)|\,\mathrm{d}x\,\mathrm{d}t,$$

and by letting $y=(t,x)$ we split $(0,T)\times\mathbb{T}^d$ into two disjoint domains $\mathscr{A} := \{y : \rho^\varepsilon(y) = 0\}$ and $\mathscr{A}^c$ and see that trivially on $\mathscr{A}$ we have $\rho(y)=0$ a.e. For the integral over $\mathscr{A}$ we note that $\nabla P'(\rho^\varepsilon)$ is a distribution that may have a singular part but we see that $\varphi[\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon]$ is smooth and equals zero on $\mathscr{A}$ and so any singular part vanishes. Thus we are left with

$$\int_{\mathscr{A}^c} |\varphi[\rho^\varepsilon u^\varepsilon - (\rho u)^\varepsilon]\nabla P'(\rho^\varepsilon)|\,\mathrm{d}x\,\mathrm{d}t,$$

and using again the identity (3-10) we obtain

$$\int_{\mathscr{A}^c} |\varphi[(\rho^\varepsilon-\rho)(u^\varepsilon-u)]\nabla P'(\rho^\varepsilon)|\,\mathrm{d}x\,\mathrm{d}t.$$

For the integral over $\mathscr{A}^c$ we see that

$$\int_{\mathscr{A}^c} |\varphi(\rho^\varepsilon-\rho)(u^\varepsilon-u)\nabla P'(\rho^\varepsilon)|\,\mathrm{d}x\,\mathrm{d}t = \int_{\mathscr{A}^c} |\varphi(\rho^\varepsilon-\rho)(u^\varepsilon-u)P''(\rho^\varepsilon)\cdot\nabla\rho^\varepsilon|\,\mathrm{d}x\,\mathrm{d}t$$

and we observe that by the definition of P we have $\rho^\varepsilon P''(\rho^\varepsilon) = p'(\rho^\varepsilon)$, and by assumption p' is bounded. Therefore we have the bound

$$\int_{\mathscr{A}^c} |\varphi(\rho^\varepsilon-\rho)(u^\varepsilon-u)P''(\rho^\varepsilon)\nabla\rho^\varepsilon|\,\mathrm{d}x\,\mathrm{d}t \le \int_{\mathscr{A}^c}\left|\varphi(\rho^\varepsilon-\rho)(u^\varepsilon-u)p'(\rho^\varepsilon)\frac{\nabla\rho^\varepsilon}{\rho^\varepsilon}\right|\mathrm{d}x\,\mathrm{d}t.$$

We have assumed that $p'(0)=0$ and $p'\in C^{0,\gamma-1}$ and so take any ρ_1, ρ_2 such that $p'(\rho_2)=0$ and we obtain

$$|p'(\rho_1)| = |p'(\rho_1)-p'(\rho_2)| \le C|\rho_1-\rho_2|^{\gamma-1} \le C|\rho_1|^{\gamma-1}$$

using the definition of Hölder continuity. Thus letting $\rho_1=\rho^\varepsilon(x)$ for each x we see that $|p'(\rho^\varepsilon)(x)| \le C|\rho^\varepsilon|^{\gamma-1}(x)$ and so we obtain

$$\int_{\mathscr{A}^c} \left|\varphi(\rho^\varepsilon-\rho)(u^\varepsilon-u)p'(\rho^\varepsilon)\frac{\nabla\rho^\varepsilon}{\rho^\varepsilon}\right| \mathrm{d}x\,\mathrm{d}t \le C\int_{\mathscr{A}^c} \left|\varphi(\rho^\varepsilon-\rho)(u^\varepsilon-u)(\rho^\varepsilon)^{\gamma-1}\frac{\nabla\rho^\varepsilon}{\rho^\varepsilon}\right| \mathrm{d}x\,\mathrm{d}t.$$

We will split the integral over $\mathscr{A}^c$ further into different disjoint domains, $\mathscr{B}_{\varepsilon^\beta} := \{y : 0<\rho^\varepsilon(y)<\varepsilon^\beta\}$ and $\mathscr{C}_{\varepsilon^\beta} := \{y : \rho^\varepsilon(y)\ge\varepsilon^\beta\}$. For the integral over $\mathscr{B}_{\varepsilon^\beta}$ we see that

$$\begin{aligned}
&\left|\int_{\mathscr{B}_{\varepsilon^\beta}} \varphi(\rho^\varepsilon-\rho)(u^\varepsilon-u)(\rho^\varepsilon)^{\gamma-1}\frac{\nabla\rho^\varepsilon}{\rho^\varepsilon}\,\mathrm{d}x\,\mathrm{d}t\right| \\
&\qquad\le C\|\varphi\|_{C^0}\varepsilon^{\beta-1+\alpha}\|\rho\|_{B_q^{\beta,\infty}}\|u\|_{B_p^{\alpha,\infty}}\|(\rho^\varepsilon)^{\gamma-1}\|_{L^\infty(\mathscr{B}_{\varepsilon^\beta})}\left\|\frac{\rho^\varepsilon-\rho}{\rho^\varepsilon}\right\|_{L^q(\mathscr{B}_{\varepsilon^\beta})} \\
&\qquad\le C\|\varphi\|_{C^0}\varepsilon^{\gamma\beta-1+\alpha}\|\rho\|_{B_q^{\beta,\infty}}\|u\|_{B_p^{\alpha,\infty}}\left\|\frac{\rho^\varepsilon-\rho}{\rho^\varepsilon}\right\|_{L^q(\mathscr{B}_{\varepsilon^\beta})},
\end{aligned}$$

where for the last line, as $\rho^\varepsilon(y)\le\varepsilon^\beta$, we have $(\rho^\varepsilon(y))^{\gamma-1}\le\varepsilon^{\beta(\gamma-1)}$ as $\gamma-1>0$. We also have the assumption that $\|(\rho^\varepsilon-\rho)/\rho^\varepsilon\|_{L^q(\mathscr{B}_{\varepsilon^\beta})}\le C$ and so we have the bound $C\varepsilon^{\gamma\beta-1+\alpha}$ as wanted. We are left with the integral over $\mathscr{C}_{\varepsilon^\beta}$ and see that

$$\left|\int_{\mathscr{C}_{\varepsilon^\beta}} \varphi(\rho^\varepsilon-\rho)(u^\varepsilon-u)(\rho^\varepsilon)^{\gamma-1}\frac{\nabla\rho^\varepsilon}{\rho^\varepsilon}\,\mathrm{d}x\,\mathrm{d}t\right| \le C\|\varphi\|_{C^0}\varepsilon^{\beta-1+\alpha}\|\rho\|_{B_q^{\beta,\infty}}\|u\|_{B_p^{\alpha,\infty}}\left\|\frac{\rho^\varepsilon-\rho}{(\rho^\varepsilon)^{2-\gamma}}\right\|_{L^q(\mathscr{C}_{\varepsilon^\beta})}.$$

As $\rho^\varepsilon\ge\varepsilon^\beta$, we have $(\rho^\varepsilon)^{-1}\le\varepsilon^{-\beta}$, and so $(\rho^\varepsilon)^{\gamma-2}\le\varepsilon^{\beta(\gamma-2)}$, and we obtain

$$\left\|\frac{\rho^\varepsilon-\rho}{(\rho^\varepsilon)^{2-\gamma}}\right\|_{L^q(\mathscr{C}_{\varepsilon^\beta})} \le \|\rho^\varepsilon-\rho\|_{L^q(\mathscr{C}_{\varepsilon^\beta})}\varepsilon^{\beta(\gamma-2)} \le C\varepsilon^\beta\|\rho\|_{B_q^{\beta,\infty}}\varepsilon^{\beta(\gamma-2)} \le C\varepsilon^{\beta(\gamma-1)}.$$

We are thus done as we have obtained convergence to zero as long as $\gamma\beta+\alpha>1$.

We have thus shown that, under the assumptions of the theorem, we have $R^\varepsilon, S^\varepsilon\to 0$. The result follows. □

We have written Theorem 4.1 in the most general from but observe that the condition

$$\left\|\frac{\rho^\varepsilon-\rho}{\rho^\varepsilon}\right\|_{L^q(\mathscr{B}_{\varepsilon^\beta})} \le C$$

feels rather artificial and is not in the $p\in C^2$ result from [Feireisl et al. 2017]. We will now focus on finding conditions on ρ for different L^q norms that will control this term.

Our first result will show that when we assume that $q=1$ and so u, ρ are Hölder continuous, not just Besov functions, we can control this term directly as expected and do not have to ask for any special extra conditions.

Lemma 4.3. *Let $w \in L^1(\Omega)$ be nonnegative, where $\Omega \subset (0,T)\times\mathbb{T}^d$ satisfies $|\Omega| \neq 0$ and $w^\varepsilon|_\Omega > 0$. Then $\|(w^\varepsilon - w)/w^\varepsilon\|_{L^1(\Omega)} \le C$, where C does not depend on ε but may depend on w and Ω.*

Proof. It suffices to show that $\|w/w^\varepsilon\|_{L^1(\Omega)} \le C$. Indeed, since $|\Omega| \le C$,

$$\left\|\frac{w^\varepsilon - w}{w^\varepsilon}\right\|_{L^1(\Omega)} \le \|1\|_{L^1(\Omega)} + \left\|\frac{w}{w^\varepsilon}\right\|_{L^1(\Omega)} = C + \left\|\frac{w}{w^\varepsilon}\right\|_{L^1(\Omega)}.$$

Fix $\varepsilon > 0$, $N = d+1$, and let $\{Q_j\}_{j=1}^n$ be a partition of $(0,T)\times\mathbb{T}^d$ into disjoint cubes with side length ε/C_N, where C_N is a constant depending only on the dimension, and select the cubes such that $|\Omega \cap Q_j| \neq 0$. Decomposing w as $w = \sum_{j=1}^n w\chi_{Q_j}$, we see that

$$\left\|\frac{w}{w^\varepsilon}\right\|_{L^1(\Omega)} = \left\|\frac{\sum_{j=1}^n w\chi_{Q_j}}{\sum_{k=1}^n (w\chi_{Q_k})^\varepsilon}\right\|_{L^1(\Omega)} \le \sum_{j=1}^n \left\|\frac{w\chi_{Q_j}}{(w\chi_{Q_j})^\varepsilon}\right\|_{L^1(\Omega)}.$$

We now want to bound $(w\chi_{Q_j})^\varepsilon$ from below. Recalling from Section 2 that $\eta = 1$ for $|x| < \frac{1}{3}$, we have, for $x \in Q_j$, that

$$\begin{aligned}(w\chi_{Q_j})^\varepsilon(x) &\ge \frac{1}{\varepsilon^N}\int_{\{|(x-y)/\varepsilon| \le 1/3\}} \eta\left(\frac{x-y}{\varepsilon}\right)(w\chi_{Q_j})(y)\,\mathrm{d}y = \frac{1}{\varepsilon^N}\int_{B_{\varepsilon/3}(x)} (w\chi_{Q_j})(y)\,\mathrm{d}y \\ &= \frac{\omega_N}{|B_\varepsilon|}\int_{B_{\varepsilon/3}(x)} (w\chi_{Q_j})(y)\,\mathrm{d}y \ge \frac{\omega_N}{|B_\varepsilon|}\int_{Q_j} w(y)\,\mathrm{d}y,\end{aligned}$$

where we obtain the last inequality provided C_N is large enough so that $B_{\varepsilon/3}(x) \supset Q_j$ for all $x \in Q_j$. Thus we obtain

$$\left\|\frac{w}{w^\varepsilon}\right\|_{L^1(\Omega)} \le \sum_{j=1}^n \left\|\frac{w\chi_{Q_j}}{(w\chi_{Q_j})^\varepsilon}\right\|_{L^1(\Omega)} \le \sum_{j=1}^n \frac{|B_\varepsilon|}{\omega_N}\frac{\int_{Q_j} w\,\mathrm{d}x}{\int_{Q_j} w\,\mathrm{d}x} \le C\sum_{j=1}^n |Q_j| \le C, \tag{4-8}$$

where we have used a dimensional constant to relate the measure of the balls to the associated cubes. □

As a consequence we obtain the following corollary, where by assuming Hölder continuity of u and ρ we obtain a natural extension of Theorem 1.1 to the case where $p \in C^{1,\gamma-1}$.

Corollary 4.4. *Let ρ, u be a solution of* (1-1) *in the sense of distributions. Assume*

$$u \in C^\alpha((0,T)\times\mathbb{T}^d), \quad \rho, \rho u \in C^\beta((0,T)\times\mathbb{T}^d), \quad 0 \le \underline{\rho} \le \rho \le \bar{\rho} \ \text{ a.e. in } (0,T)\times\mathbb{T}^d$$

for some constants $\underline{\rho}$, $\bar{\rho}$ and $0 \le \alpha, \beta \le 1$ such that

$$\alpha + \gamma\beta > 1 \quad \text{and} \quad 2\alpha + \beta > 1.$$

Assume further that $p \in C^{1,(\gamma-1)}([\underline{\rho}, \bar{\rho}])$, and, in addition

$$p'(0) = 0 \quad \text{as soon as } \underline{\rho} = 0.$$

Then the energy is locally conserved; i.e., (1-2) *holds in the sense of distributions on $(0,T)\times\mathbb{T}^d$.*

Proof. For the integral over $\mathscr{B}_{\varepsilon^\beta}$, in the proof of Theorem 4.1, we see that

$$\left|\int_{\mathscr{B}_{\varepsilon^\beta}} \varphi(\rho^\varepsilon-\rho)(u^\varepsilon-u)(\rho^\varepsilon)^{\gamma-1}\frac{\nabla\rho^\varepsilon}{\rho^\varepsilon}\,\mathrm{d}x\,\mathrm{d}t\right|$$
$$\leq C\|\varphi\|_{C^0}\varepsilon^{\beta-1+\alpha}\|\rho\|_{\mathscr{C}^\beta}\|u\|_{\mathscr{C}^\alpha}\|(\rho^\varepsilon)^{\gamma-1}\|_{L^\infty(\mathscr{B}_{\varepsilon^\beta})}\left\|\frac{\rho^\varepsilon-\rho}{\rho^\varepsilon}\right\|_{L^1(\mathscr{B}_{\varepsilon^\beta})}$$
$$\leq C\|\varphi\|_{C^0}\varepsilon^{\gamma\beta-1+\alpha}\|\rho\|_{\mathscr{C}^\beta}\|u\|_{\mathscr{C}^\alpha}.$$

For the other bounds, as we are on a domain with finite measure, we can bound the Besov norms by the Hölder norms. □

Remark 4.5. Notice that the conditions $u\in C^\alpha((0,T)\times\mathbb{T}^d)$ and $\rho\in C^\beta((0,T)\times\mathbb{T}^d)$ imply that $\rho u\in C^{\min(\alpha,\beta)}((0,T)\times\mathbb{T}^d)$. Therefore, if one has $\alpha\geq\beta$, then the requirement that ρu be in $C^\beta((0,T)\times\mathbb{T}^d)$ can be dropped. See also Remark 3.2(2) in [Feireisl et al. 2017].

When we still want to consider Besov spaces for ρ and u we have to consider extra conditions on ρ in order to control the term $\|(\rho^\varepsilon-\rho)/\rho^\varepsilon\|_{L^q(\mathscr{B}_{\varepsilon^\beta})}$. Our first method will be to ask for an integrability condition on $1/\rho$.

Lemma 4.6. *Assume that* $1/w\in L^p$ *and* $w\in L^q$. *Then*

$$\left\|\frac{w^\varepsilon-w}{w^\varepsilon}\right\|_{L^r}\leq C\quad\textit{for }\frac{1}{p}+\frac{1}{q}\leq\frac{1}{r},$$

and in fact if $r<\infty$,

$$\lim_{\varepsilon\to0}\left\|\frac{w^\varepsilon-w}{w^\varepsilon}\right\|_{L^r}=0\quad\textit{for }\frac{1}{p}+\frac{1}{q}\leq\frac{1}{r}.$$

Proof. Using Hölder's inequality and then Jensen's inequality, as the integral of the mollifier is 1 and $1/x$ is a convex function, we get that $\|1/w^\varepsilon\|\leq\|(1/w)^\varepsilon\|\leq\|1/w\|$ and so

$$\left\|\frac{w^\varepsilon-w}{w^\varepsilon}\right\|_{L^r}\leq\|w^\varepsilon-w\|_{L^q}\left\|\frac{1}{w^\varepsilon}\right\|_{L^p}\leq\|w^\varepsilon-w\|_{L^q}\left\|\frac{1}{w}\right\|_{L^p}\leq C.$$

As long as $q<\infty$ we see that this, in fact, converges to zero. □

We now obtain the following corollary adding this condition into Theorem 4.1. We note that when $p=q=3$, we obtain the best result with the weakest integrability assumption in the Besov norms.

Corollary 4.7. *Let* ρ, u *be a solution of* (1-1) *in the sense of distributions. Assume*

$$u\in B_p^{\alpha,\infty}((0,T)\times\mathbb{T}^d),\quad \rho,\rho u\in B_q^{\beta,\infty}((0,T)\times\mathbb{T}^d),\quad 0\leq\underline{\rho}\leq\rho\leq\bar\rho\ \textit{ a.e. in } (0,T)\times\mathbb{T}^d$$

for some constants $\underline{\rho}$, $\bar\rho$ *and* $0\leq\alpha,\beta\leq1$ *such that*

$$\frac{1}{p}+\frac{2}{q}\leq1,\quad \frac{2}{p}+\frac{1}{q}\leq1,\quad p,q\geq2,\quad \alpha+\gamma\beta>1,\quad\textit{and}\quad 2\alpha+\beta>1.$$

Define $\mathscr{E}:=\{x:\rho\neq0\}$ *and assume that*

$$\frac{1}{\rho}\in L^q(\mathscr{E}).$$

Assume further that $p \in C^{1,(\gamma-1)}([\underline{\rho}, \bar{\rho}])$, *and, in addition*

$$p'(0) = 0 \quad \textit{as soon as } \underline{\rho} = 0.$$

Then the energy is locally conserved; *i.e.*, (1-2) *holds in the sense of distributions on* $(0, T) \times \mathbb{T}^d$.

Proof. For the integral over $\mathscr{B}_{\varepsilon^\beta}$, in the proof of Theorem 4.1, we see that as $\rho \in L^\infty$ and $\varepsilon^\beta \geq \rho^\varepsilon$ then

$$\left| \int_0^T \int_{\mathscr{B}_{\varepsilon^\beta}} \varphi(\rho^\varepsilon - \rho)(u^\varepsilon - u)(\rho^\varepsilon)^{\gamma-1} \frac{\nabla \rho^\varepsilon}{\rho^\varepsilon} \,\mathrm{d}x\,\mathrm{d}t \right|$$
$$\leq C \|\varphi\|_{C^0} \varepsilon^{\beta-1+\alpha} \|\rho\|_{B_q^{\beta,\infty}} \|u\|_{B_p^{\alpha,\infty}} \|(\rho^\varepsilon)^{\gamma-1}\|_{L^\infty(\mathscr{B}_{\varepsilon^\beta})} \left\| \frac{\rho^\varepsilon - \rho}{\rho^\varepsilon} \right\|_{L^q(\mathscr{B}_{\varepsilon^\beta})}$$
$$\leq C \|\varphi\|_{C^0} \varepsilon^{\gamma\beta-1+\alpha} \|\rho\|_{B_q^{\beta,\infty}} \|u\|_{B_p^{\alpha,\infty}} \left\| \frac{1}{\rho} \right\|_{L^q(\mathscr{E})},$$

and so we are done, using Lemma 4.6 for the final step. □

Remark 4.8. Even though we have written $1/\rho \in L^q(\mathscr{E})$, we can fix some $\delta > 0$ and only need this condition on some $\mathscr{B}_\delta$, as for $\varepsilon^1 > \varepsilon^2$ we have $\mathscr{B}_{\varepsilon^2} \subset \mathscr{B}_{\varepsilon^1}$, and so when $\varepsilon^\beta < \delta$ we have $\mathscr{B}_{\varepsilon^\beta} \subset \mathscr{B}_\delta$.

One can see that the condition $1/\rho \in L^q(\mathscr{B}_\delta)$ is quite a strong assumption and requires a quick approach of the function to the null set. Above we used conventional bounds to obtain a general integral result but do not consider the local structure of the function. We notice that a pointwise estimate $\rho \leq C\rho^\varepsilon$ would allow us to control the L^q norm of $(\rho^\varepsilon - \rho)/\rho^\varepsilon$ and, though convexity of ρ would do, we will now show a nice link between this and quasinearly subharmonic functions which are much more general functions than subharmonic, quasisubharmonic and nearly subharmonic functions [Pavlović and Riihentaus 2011]. The main motivation for the study of this notion in this paper is that it happens, as will be shown below, to be equivalent to the L^∞-boundedness of our problem term $(\rho^\varepsilon - \rho)/\rho^\varepsilon$.

Definition 4.9. Let $X \subset \mathbb{R}^d$ be a set and $u : X \to [0, +\infty)$ be Borel measurable. Then u is quasinearly subharmonic on X, that is $u \in \mathrm{QNS}(X)$, if there is a constant $\varepsilon_0 = \varepsilon_0(u)$, $0 < \varepsilon_0 < 1$, such that for each open set $O \subset X$, $O \neq X$, for each $x \in O$ and each r, $0 < r \leq \varepsilon_0 \delta^O(x)$, one has $u \in L^1(B_r(x))$ and

$$u(x) \leq \frac{C}{|B_r(x)|} \int_{B_r(x)} u(y)\,\mathrm{d}y \quad \text{for some constant } C \geq 1, \tag{4-9}$$

where C is independent of r, $|B_r(x)| = \omega_d r^d$ is the volume of the ball and

$$\delta^O(x) = \mathrm{dist}(x, O^c) \quad \text{for the complement } O^c \text{ of } O \text{ in } X.$$

Lemma 4.10. *Let* $u : X \to [0, +\infty)$ *be a Borel measurable function. Then u is quasinearly subharmonic if and only if for every* $O \Subset X$ *there exist* M, ε_0 *such that for any* $0 < \varepsilon < \varepsilon_0$

$$u(x) \leq M u^\varepsilon(x) \quad \textit{for any } x \in O.$$

Proof. Let $u : X \to [0, +\infty)$ be a quasinearly subharmonic function. Then for any $\varepsilon < \mathrm{dist}(O, \partial X)$, u^ε is a well-defined smooth function on O. Suppose that $O \Subset X$ is a precompact set. Then $\delta_0 = \mathrm{dist}(O, \partial X)$

is a positive number and for $\varepsilon < \delta_0$

$$O \subset \{x : \operatorname{dist}(x, \partial X) > \varepsilon\}$$

and u^ε is well-defined on O. We prove that there exist M and ε_0 such that

$$u(x) \le M u^\varepsilon(x) \quad \text{for any } x \in O,\ 0 < \varepsilon < \varepsilon_0.$$

Indeed, we have

$$u^\varepsilon(x) = \frac{1}{\varepsilon^d} \int_X \eta\left(\frac{x-y}{\varepsilon}\right) u(y)\,\mathrm{d}y.$$

Note that $y \in X$ for $x \in O$ and $|x - y| < \varepsilon$. Since $u \ge 0$ and recalling that from the definition of η we know that $\eta = 1$ for $|x| < \frac{1}{3}$, we have

$$\begin{aligned} u^\varepsilon(x) &\ge \frac{1}{\varepsilon^d} \int_{\{|(x-y)/\varepsilon| \le 1/3\}} \eta\left(\frac{x-y}{\varepsilon}\right) u(y)\,\mathrm{d}y \\ &= \frac{1}{\varepsilon^d} \int_{\{|(x-y)/\varepsilon| \le 1/3\}} u(y)\,\mathrm{d}y = \frac{\omega_d}{3^d |B_{\varepsilon/3}(x)|} \int_{B_{\varepsilon/3}(x)} u(y)\,\mathrm{d}y \ge \frac{\omega_d u(x)}{3^d C} \end{aligned}$$

for sufficiently small ε. Therefore, we obtain

$$u(x) \le \frac{3^d C u^\varepsilon(x)}{\omega_d} \quad \text{for sufficiently small } \varepsilon \le \varepsilon_0 \delta^O(x).$$

On the other hand, if $u(x) \le M u^\varepsilon(x)$, then we have

$$u(x) \le \frac{M}{\varepsilon^d} \int_X \eta\left(\frac{x-y}{\varepsilon}\right) u(y)\,\mathrm{d}y = \frac{M\omega_d}{\omega_d \varepsilon^d} \int_{|x-y| \le \varepsilon} \eta\left(\frac{x-y}{\varepsilon}\right) u(y)\,\mathrm{d}y \le \frac{M\omega_d}{|B_\varepsilon(x)|} \int_{B_\varepsilon(x)} u(y)\,\mathrm{d}y. \tag{4-10}$$

Hence we deduce

$$u(x) \le \frac{C}{|B_\varepsilon(x)|} \int_{B_\varepsilon(x)} u(y)\,\mathrm{d}y. \qquad \square$$

From this pointwise control showing that $\rho(x) \le M\rho^\varepsilon(x)$ we obtain another corollary to our main result.

Corollary 4.11. *Let ρ, u be a solution of* (1-1) *in the sense of distributions. Assume*

$$u \in B_p^{\alpha,\infty}((0,T) \times \mathbb{T}^d), \quad \rho, \rho u \in B_q^{\beta,\infty}((0,T) \times \mathbb{T}^d), \quad 0 \le \underline{\rho} \le \rho \le \bar{\rho} \ \text{ a.e. in } (0,T) \times \mathbb{T}^d$$

for some constants $\underline{\rho}$, $\bar{\rho}$ and $0 \le \alpha, \beta \le 1$ such that

$$\frac{1}{p} + \frac{2}{q} \le 1, \quad \frac{2}{p} + \frac{1}{q} \le 1, \quad p, q \ge 2, \quad \alpha + \gamma\beta > 1, \quad \text{and} \quad 2\alpha + \beta > 1.$$

Assume that $\rho \in \mathrm{QNS}(\mathscr{B}_\delta)$ for some $\delta > 0$ and $p \in \mathscr{C}^{1,(\gamma-1)}([\underline{\rho}, \bar{\rho}])$ with

$$p'(0) = 0 \quad \text{as soon as } \underline{\rho} = 0.$$

Then the energy is locally conserved; *i.e.*, (1-2) *holds in the sense of distributions on* $(0,T) \times \mathbb{T}^d$.

Proof. For the integral over $\mathscr{B}_{\varepsilon^\beta}$, in the proof of Theorem 4.1, we see that

$$\left\| \frac{\rho^\varepsilon - \rho}{\rho^\varepsilon} \right\|_{L^q(\mathscr{B}_{\varepsilon^\beta})} \leq \left\| \frac{\rho^\varepsilon + C\rho^\varepsilon}{\rho^\varepsilon} \right\|_{L^q(\mathscr{B}_{\varepsilon^\beta})} \leq C \quad \text{for } \varepsilon^\beta < \delta$$

and so we are done. □

Remark 4.12. (1) The condition $\rho \in \mathrm{QNS}(B_\delta)$ deals with ρ, $\rho_\varepsilon = 0$ without splitting into cases and so using this condition the proof is simplified.

(2) We note that this condition is weaker than local convexity of ρ on $\mathscr{B}_\delta$, which would also give the same result.

(3) In view of Lemma 4.10, it is essentially a matter of taste if one prefers to formulate Corollary 4.11 in terms of quasinearly subharmonicity or directly under the assumption $\rho \leq C\rho^\varepsilon$.

4A. *Counterexample for the L^p case.* We indicate in this subsection why Lemma 4.3 is no longer true when the L^1-norm is replaced with the L^p-norm for $p > 1$. This shows that the Hölder assumption of Corollary 4.4 cannot easily be relaxed.

We can see $\rho^\varepsilon(x)$ is like a weighted average of ρ over the ball $B_\varepsilon(x)$ and so heuristically we can see

$$\frac{\rho - \rho^\varepsilon}{\rho^\varepsilon} \simeq \frac{\rho(x) - (1/|B_\varepsilon|)\int_{B_\varepsilon(x)} \rho(y)\,\mathrm{d}y}{(1/|B_\varepsilon|)\int_{B_\varepsilon(x)} \rho(y)\,\mathrm{d}y}$$

(which is rigorous for $\eta_\varepsilon = (1/|B_\varepsilon|)\chi_{B_\varepsilon(0)}(x)$), and assuming the right-hand side is bounded and rearranging gives the condition (4-9). We see that a condition of the form

$$\left\| \frac{\rho(\,\cdot\,) - (1/|B_\varepsilon|)\int_{B_\varepsilon(\,\cdot\,)} \rho(y)\,\mathrm{d}y}{(1/|B_\varepsilon|)\int_{B_\varepsilon(\,\cdot\,)} \rho(y)\,\mathrm{d}y} \right\|_{L^p} < C,$$

in a sense a "relatively weighted L^p mean oscillation condition", could potentially be the weakest condition to control (4-1).

We notice that for the L^1 norm we obtain perfect cancellation in the fraction when calculating (4-8), as a mollifier acts like a local weighted average. However, when we perform the calculation in (4-8), but in L^p, then instead we obtain

$$\sum_{j=1}^{n} \frac{|B_\varepsilon|}{\omega_N} \frac{\|w\chi_{Q_j}\|_{L^p}}{\int_{Q_j} w\,\mathrm{d}x} = \sum_{j=1}^{n} \frac{|B_\varepsilon|}{\omega_N} \frac{\left(\int_{Q_j} w^p\,\mathrm{d}x\right)^{1/p}}{\int_{Q_j} w\,\mathrm{d}x}$$

and if we assume that $w = 1$ then we get $\sum_{j=1}^{n}(|B_\varepsilon|/\omega_N)|Q_j|^{1/p-1}$. As $1/p - 1 < 0$, for certain functions this term could blow up.

In fact if one chooses a function made of separated spikes where the supports get smaller and smaller then we can show this blow-up. We will formulate a simple counterexample so that it is in one dimension, discontinuous and nonnegative, though more regular counterexamples can be constructed in higher dimensions that are, for instance, even smooth and strictly positive.

Firstly, note that if we show that $\|f/f^\varepsilon\|_{L^p}$ blows up as $\varepsilon \to 0$ then $\|f/f^\varepsilon - f^\varepsilon/f^\varepsilon\|_{L^p}$ will also blow up. We can take $x \in \mathbb{T}$ and define our counterexample

$$f(x) := \sum_{i=1}^{\infty} \chi_{[1/i,\,1/i+1/2^i]}(x).$$

It is easy to see that $f \in B_p^{\alpha,\infty}(\mathbb{T})$ for $p > 1$ and any $0 < \alpha < 1-1/p$ by regularizing and using Lemma 2.49 from [Bahouri et al. 2011]. Thus we have the sum of separated spikes so they are further than $1/i^2$ apart yet have supports of size $1/2^i$. Let $\varepsilon = 1/(2i^2)$ and see that as f is nonnegative we can bound the sum below by just the i-th spike and see that as mollification only acts locally, the value on the denominator is only dependent on the i-th spike; thus we obtain

$$\left\| \frac{f}{f^\varepsilon} \right\|_{L^p(\mathbb{T})} \geq \left\| \frac{1}{f^\varepsilon} \right\|_{L^p(1/i,\,1/i+1/2^i)} = \|((\chi_{[1/i,\,1/i+1/2^i]})^{1/(2i^2)})^{-1}\|_{L^p(1/i,\,1/i+1/2^i)}. \tag{4-11}$$

We can then the bound mollification of $\chi_{[1/i,\,1/i+1/2^i]}$ in a similar method to (4-10) but in one dimension and so we can bound (4-11) below by

$$\left\| \frac{f}{f^\varepsilon} \right\|_{L^p(\mathbb{T})} \geq C \frac{2^i}{2i^2} \|1\|_{L^p(1/i,\,1/i+1/2^i)} = C \frac{2^i}{2i^2} 2^{-i/p} = C \frac{2^{i(1-1/p)}}{2i^2}.$$

As f is the sum of infinitely many spikes there will exist an appropriate spike for any ε_i and thus we can send $i \to \infty$ and, as $1 - 1/p > 0$, we have $C 2^{i(1-1/p)}/(2i^2) \to \infty$, which implies $\|f/f^\varepsilon\|_{L^p(\mathbb{T})} \to \infty$.

5. Energy conservation on domains with boundary

We have derived the local energy conservation equations on $(0, T) \times \mathbb{T}^d$ and so for an $\varphi \in C_c^\infty((0, T) \times \mathbb{T}^d)$ we have

$$\int_0^T \int_{\mathbb{T}^d} \partial_t \varphi \cdot \left(\tfrac{1}{2}\rho|u|^2 + P(\rho)\right) + \nabla\varphi \cdot \left[\left(\tfrac{1}{2}\rho|u|^2 + p(\rho) + P(\rho)\right)u\right] \mathrm{d}t\, \mathrm{d}x = 0. \tag{5-1}$$

The local energy equation is derived by taking momentum balance equations and testing with $(\varphi u^\varepsilon)^\varepsilon$ and using that mollification is symmetric to regularize the equation. For the continuity equation we just use φ^ε to test the equation and again move the mollification onto the equation. Once this is done, all the calculations are done locally on $\mathrm{supp}(\varphi)$.

When studying the isentropic Euler equations on a bounded domain with Lipschitz boundary Ω we have

$$\begin{aligned} \partial_t(\rho u) + \operatorname{div}(\rho u \otimes u) + \nabla p(\rho) &= 0 \quad \text{in } [0, T] \times \Omega, \\ \partial_t \rho + \operatorname{div}(\rho u) &= 0 \quad \text{in } [0, T] \times \Omega, \\ u \cdot n &= 0 \quad \text{on } [0, T] \times \partial\Omega, \end{aligned} \tag{5-2}$$

where n denotes the outward normal vector field for $\partial\Omega$. For any $\varphi \in C_c^\infty((0, T) \times \Omega)$ we can find an $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ we have φ^ε, $(\varphi u^\varepsilon)^\varepsilon \in C_c^\infty((0, T) \times \Omega)$, and so we can apply the same method as above to obtain a local energy equation on $(0, T) \times \Omega$ of the form (5-1). Here we are assuming the same conditions on u, ρ and p as in the previous theorems and in the corollaries in Sections 3 and 4, yet making the appropriate changes so that u and ρ are defined on the domain $(0, T) \times \Omega$ rather than $(0, T) \times \mathbb{T}^d$.

The following theorem and its proof follow ideas from [Bardos et al. 2018]:

Theorem 5.1. *Let ρ, u be a solution of* (5-2) *in the sense of distributions. Assume that ρ, u, and p satisfy the conditions necessary to derive the local energy equality* (5-1). *Assume further that $\rho \in L^\infty((0, T) \times \partial\Omega)$, $\partial\Omega$ is C^2, and $u \cdot n$ is continuous at the boundary. Then we have energy conservation on Ω; that is, for $\Theta(t) \in C_c^\infty(0, T)$*

$$\int_0^T \int_\Omega \partial_t \Theta(t) \cdot \left(\tfrac{1}{2}\rho|u|^2 + P(\rho)\right) \mathrm{d}t\, \mathrm{d}x = 0 \tag{5-3}$$

and further if u, ρ are weakly continuous in time then

$$\int_\Omega \tfrac{1}{2}\rho|u|^2(t_1, x) + P(\rho)(t_1, x)\, \mathrm{d}x = \int_\Omega \tfrac{1}{2}\rho|u|^2(t_2, x) + P(\rho)(t_2, x)\, \mathrm{d}x \tag{5-4}$$

for any $t_1, t_2 \in [0, T]$.

Proof. For any $\varphi \in C_c^\infty((0, T) \times \Omega)$ we can find an $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ we have $\varphi^\varepsilon, (\varphi u^\varepsilon)^\varepsilon \in C_c^\infty((0, T) \times \Omega)$ and so assuming sufficient regularity of ρ, u and p we obtain

$$\int_0^T \int_\Omega \partial_t \varphi \cdot \left(\tfrac{1}{2}\rho|u|^2 + P(\rho)\right) + \nabla\varphi \cdot \left[\left(\tfrac{1}{2}\rho|u|^2 + p(\rho) + P(\rho)\right)u\right] \mathrm{d}t\, \mathrm{d}x = 0. \tag{5-5}$$

Let $\chi : \mathbb{R}^+ \to \mathbb{R}$ be a nonnegative, smooth function such that

$$\chi(s) := \begin{cases} 0 & \text{if } s < 1, \\ 1 & \text{if } s > 2, \end{cases}$$

and define for $x \in \overline{\Omega}$ the function $d_{\partial\Omega}(x)$ as the euclidean distance from x to the closest point on the boundary. We can then define for any $\delta > 0$ the composition $\chi(d_{\partial\Omega}(x)/\delta)$ and see that as $\delta \to 0$ so does $\chi(d_{\partial\Omega}(x)/\delta) \to \mathbb{1}_\Omega$. Further, let $\Theta(t) \in C_c^\infty(0, T)$.

We can for any $\delta > 0$ let $\varphi(x, t) = \chi(d_{\partial\Omega}(x)/\delta)\Theta(t)$ in (5-5) and we obtain

$$\int_\Omega \chi\left(\frac{d_{\partial\Omega}(x)}{\delta}\right) \int_0^T \partial_t \Theta(t) \cdot \left(\tfrac{1}{2}\rho|u|^2 + P(\rho)\right) \mathrm{d}x\, \mathrm{d}t \\ + \int_0^T \Theta(t) \int_\Omega \nabla\chi\left(\frac{d_{\partial\Omega}(x)}{\delta}\right) \cdot \left[\left(\tfrac{1}{2}\rho|u|^2 + p(\rho) + P(\rho)\right)u\right] \mathrm{d}t\, \mathrm{d}x = 0,$$

and by the chain rule we see that

$$\nabla\chi\left(\frac{d_{\partial\Omega}(x)}{\delta}\right) = \frac{1}{\delta}\chi'\left(\frac{d_{\partial\Omega}(x)}{\delta}\right)\nabla d_{\partial\Omega}(x),$$

and so

$$0 = \int_\Omega \chi\left(\frac{d_{\partial\Omega}(x)}{\delta}\right) \int_0^T \partial_t \Theta(t) \cdot \left(\tfrac{1}{2}\rho|u|^2 + P(\rho)\right) \mathrm{d}x\, \mathrm{d}t \\ + \int_0^T \Theta(t) \int_\Omega \frac{1}{\delta}\chi'\left(\frac{d_{\partial\Omega}(x)}{\delta}\right) \nabla d_{\partial\Omega}(x) \cdot \left[\left(\tfrac{1}{2}\rho|u|^2 + p(\rho) + P(\rho)\right)u\right] \mathrm{d}t\, \mathrm{d}x. \tag{5-6}$$

As $\chi(d_{\partial\Omega}(x)/\delta) \to \mathbb{1}_\Omega$ strongly, the first integral on the right-hand side of (5-6) will converge to

$$\int_0^T\int_\Omega \partial_t\Theta(t)\cdot\left(\tfrac{1}{2}\rho|u|^2+P(\rho)\right)\mathrm{d}x\,\mathrm{d}t$$

as we wanted. All that is left is to show that the other term on the right-hand side of (5-6) vanishes in the limit.

As $\partial\Omega$ is C^2 we can use [Gilbarg and Trudinger 1977], specifically Lemma 14.16, to see that there exists an $a>0$ such that $d_{\partial\Omega}(x)\in C^2(\Gamma_a)$, where $\Gamma_a := \{x\in\overline{\Omega} : d_{\partial\Omega}(x) < a\}$. Further, in a similar argument to [Bardos et al. 2014, Section 7], when $x\in\Omega$ is sufficiently close to $\partial\Omega$, there exists a unique point $\hat{x}\in\partial\Omega$ such that $x=\hat{x}+n(\hat{x})\,d_{\partial\Omega}(x)$, where $n(\hat{x})$ is the unit outward normal to the boundary at x. We see that we can bound the modulus for the second term on the right-hand side of (5-6) by

$$\begin{aligned}\left\|\chi'\left(\frac{d_{\partial\Omega}}{\delta}\right)\right\|_{L^\infty}\int_0^T\Theta(t)\frac{1}{\delta}\int_{\Gamma_{2\delta}}|\nabla d_{\partial\Omega}(x)\cdot u|\left|\left(\tfrac{1}{2}\rho|u|^2+p(\rho)+P(\rho)\right)\right|\mathrm{d}t\,\mathrm{d}x&\\ \leq C\int_0^T\Theta(t)\frac{1}{\delta}\int_{\Gamma_{2\delta}}|\nabla d_{\partial\Omega}(x)\cdot u|\,\mathrm{d}t\,\mathrm{d}x&\quad(5\text{-}7)\end{aligned}$$

as we know that $\|\chi'(d_{\partial\Omega}/\delta)\|_{L^\infty}\leq C$ and by our assumptions $\left\|\left(\tfrac{1}{2}\rho|u|^2+p(\rho)+P(\rho)\right)\right\|_{L^\infty}\leq C$ as well. For $2\delta<a$ we know that $d_{\partial\Omega}\in C^2$ and furthermore as $\nabla d_{\partial\Omega}\in C^1$, in the region $\Gamma_{2\delta}$, $|\nabla d_{\partial\Omega}(x)\cdot u|\to C|n(\hat{x})\cdot u(\hat{x})|$ as long as $u(x)\to u(\hat{x})$ as $x\to\hat{x}$, and for this the assumption that $u\cdot n$ is continuous at the boundary will suffice. Thus as $\partial\Omega$ is at least Lipschitz so $|\Gamma_{2\delta}|\leq C\delta|\partial\Omega|$ and so we can apply the Lebesgue differentiation theorem to (5-7) and see that as $\delta\to 0$,

$$C\int_0^T\Theta(t)\frac{1}{\delta}\int_{\Gamma_{2\delta}}|\nabla d_{\partial\Omega}(x)\cdot u|\,\mathrm{d}t\,\mathrm{d}x\to C\int_0^T\Theta(t)\int_{\partial\Omega}|n(\hat{x})\cdot u(\hat{x})|\,\mathrm{d}t\,\mathrm{d}\hat{x}=0$$

as $n(\hat{x})\cdot u(\hat{x})=0$ and so we have shown (5-3).

We now want to show (5-4) with the extra assumptions of weak continuity in time of both u and ρ. To do this we define the sequence of functions $\Theta_\nu : [0,T]\to\mathbb{R}$ which are nonnegative and smooth, where for any point $t_1,t_2\in[0,T]$ with $t_1<t_2$ we have

$$\Theta_\nu(\tau) := \begin{cases}0 & \text{if } \tau<t_1+\nu \text{ or } \tau>t_2-\nu,\\ 1 & \text{if } \tau>t_1+2\nu \text{ or } \tau<t_2-2\nu,\end{cases}$$

and see similarly that as $\nu\to 0$ we have $\Theta_\nu(t)\to\mathbb{1}_{[t_1,t_2]}$. We see that $\Theta_\nu\in C_c^\infty(0,T)$ for every $\nu>0$ and so substituting this function into (5-6) we obtain

$$\int_0^T\int_\Omega \partial_t\Theta_\nu(t)\cdot\left(\tfrac{1}{2}\rho|u|^2+P(\rho)\right)\mathrm{d}t\,\mathrm{d}x=0$$

for every ν. From our choice of Θ_ν we see that

$$\begin{aligned}\int_0^T\int_{\mathbb{T}^d}\partial_t\Theta_\nu(t)\cdot\left(\tfrac{1}{2}\rho|u|^2+P(\rho)\right)\mathrm{d}t\,\mathrm{d}x=\int_{t_1}^{t_1+2\nu}\partial_t\Theta_\nu(t)\cdot\int_\Omega\left(\tfrac{1}{2}\rho|u|^2+P(\rho)\right)\mathrm{d}t\,\mathrm{d}x&\\ +\int_{t_2-2\nu}^{t_2}\partial_t\Theta_\nu(t)\cdot\int_\Omega\left(\tfrac{1}{2}\rho|u|^2+P(\rho)\right)\mathrm{d}t\,\mathrm{d}x.&\end{aligned}$$

We know that $\int_{t_1}^{t_1+2\nu} \partial_t \Theta_\nu(t)\,\mathrm{d}t = 1$ and $\int_{t_2-2\nu}^{t_2} \partial_t \Theta_\nu(t)\,\mathrm{d}t = -1$ by the fundamental theorem of calculus and as $\nu \to 0$ these terms approximate the identity at t_1 and t_2, and thus these terms converge to

$$\int_\Omega \tfrac{1}{2}\rho|u|^2(t_1,x) + P(\rho)(t_1,x)\,\mathrm{d}x \quad \text{and} \quad -\int_\Omega \tfrac{1}{2}\rho|u|^2(t_2,x) + P(\rho)(t_2,x)\,\mathrm{d}x$$

respectively, assuming weak continuity of ρ and u in time. Thus we are done. □

Acknowledgements

We would like to thank the universities of Hannover and Warsaw for creating great research environments. Dębiec acknowledges support from the National Science Centre (Poland), 2012/05/E/ST1/02218. The research funding for Skipper is graciously given by German research foundation (DFG) grant no. WI 4840/1-1.

References

[Akramov and Wiedemann 2019] I. Akramov and E. Wiedemann, "Renormalization of active scalar equations", *Nonlinear Anal.* **179** (2019), 254–269. MR Zbl

[Bahouri et al. 2011] H. Bahouri, J.-Y. Chemin, and R. Danchin, *Fourier analysis and nonlinear partial differential equations*, Grundlehren der Math. Wissenschaften **343**, Springer, 2011. MR Zbl

[Bardos and Titi 2018] C. Bardos and E. S. Titi, "Onsager's conjecture for the incompressible Euler equations in bounded domains", *Arch. Ration. Mech. Anal.* **228**:1 (2018), 197–207. MR Zbl

[Bardos et al. 2014] C. Bardos, L. Székelyhidi, Jr., and E. Wiedemann, "Non-uniqueness for the Euler equations: the effect of the boundary", *Uspekhi Mat. Nauk* **69**:2(416) (2014), 3–22. In Russian; translated in *Russian Math. Surv.* **69**:2 (2014), 189–207. MR Zbl

[Bardos et al. 2018] C. Bardos, E. Titi, and E. Wiedemann, "Onsager's conjecture with physical boundaries and an application to the vanishing viscosity limit", preprint, 2018. arXiv

[Bardos et al. 2019] C. Bardos, P. Gwiazda, A. Świerczewska-Gwiazda, E. S. Titi, and E. Wiedemann, "On the extension of Onsager's conjecture for general conservation laws", *J. Nonlinear Sci.* **29**:2 (2019), 501–510. MR Zbl

[Buckmaster et al. 2019] T. Buckmaster, C. de Lellis, L. Székelyhidi, Jr., and V. Vicol, "Onsager's conjecture for admissible weak solutions", *Comm. Pure Appl. Math.* **72**:2 (2019), 229–274. MR Zbl

[Chen and Torres 2005] G.-Q. Chen and M. Torres, "Divergence-measure fields, sets of finite perimeter, and conservation laws", *Arch. Ration. Mech. Anal.* **175**:2 (2005), 245–267. MR Zbl

[Chen and Yu 2019] R. M. Chen and C. Yu, "Onsager's energy conservation for inhomogeneous Euler equations", *J. Math. Pure Appl.* **131** (2019), 1–16. MR Zbl

[Cheskidov et al. 2008] A. Cheskidov, P. Constantin, S. Friedlander, and R. Shvydkoy, "Energy conservation and Onsager's conjecture for the Euler equations", *Nonlinearity* **21**:6 (2008), 1233–1252. MR Zbl

[Constantin et al. 1994] P. Constantin, W. E, and E. S. Titi, "Onsager's conjecture on the energy conservation for solutions of Euler's equation", *Comm. Math. Phys.* **165**:1 (1994), 207–209. MR Zbl

[DiPerna and Lions 1989] R. J. DiPerna and P.-L. Lions, "Ordinary differential equations, transport theory and Sobolev spaces", *Invent. Math.* **98**:3 (1989), 511–547. MR Zbl

[Dębiec et al. 2018] T. Dębiec, P. Gwiazda, A. Świerczewska-Gwiazda, and A. Tzavaras, "Conservation of energy for the Euler–Korteweg equations", *Calc. Var. Partial Differential Equations* **57**:6 (2018), art. id. 160. MR Zbl

[Drivas and Eyink 2018] T. D. Drivas and G. L. Eyink, "An Onsager singularity theorem for turbulent solutions of compressible Euler equations", *Comm. Math. Phys.* **359**:2 (2018), 733–763. MR Zbl

[Drivas and Nguyen 2018] T. D. Drivas and H. Q. Nguyen, "Onsager's conjecture and anomalous dissipation on domains with boundary", *SIAM J. Math. Anal.* **50**:5 (2018), 4785–4811. MR Zbl

[Duchon and Robert 2000] J. Duchon and R. Robert, "Inertial energy dissipation for weak solutions of incompressible Euler and Navier–Stokes equations", *Nonlinearity* **13**:1 (2000), 249–255. MR Zbl

[Eyink 1994] G. L. Eyink, "Energy dissipation without viscosity in ideal hydrodynamics, I: Fourier analysis and local energy transfer", *Phys. D* **78**:3-4 (1994), 222–240. MR Zbl

[Feireisl et al. 2017] E. Feireisl, P. Gwiazda, A. Świerczewska-Gwiazda, and E. Wiedemann, "Regularity and energy conservation for the compressible Euler equations", *Arch. Ration. Mech. Anal.* **223**:3 (2017), 1375–1395. MR Zbl

[Fjordholm and Wiedemann 2018] U. S. Fjordholm and E. Wiedemann, "Statistical solutions and Onsager's conjecture", *Phys. D* **376/377** (2018), 259–265. MR Zbl

[Gilbarg and Trudinger 1977] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Grundlehren der Math. Wissenschaften **224**, Springer, 1977. MR Zbl

[Gwiazda et al. 2018] P. Gwiazda, M. Michálek, and A. Świerczewska-Gwiazda, "A note on weak solutions of conservation laws and energy/entropy conservation", *Arch. Ration. Mech. Anal.* **229**:3 (2018), 1223–1238. MR Zbl

[Isett 2018] P. Isett, "A proof of Onsager's conjecture", *Ann. of Math.* (2) **188**:3 (2018), 871–963. MR Zbl

[Leslie and Shvydkoy 2016] T. M. Leslie and R. Shvydkoy, "The energy balance relation for weak solutions of the density-dependent Navier–Stokes equations", *J. Differential Equations* **261**:6 (2016), 3719–3733. MR Zbl

[Pavlović and Riihentaus 2011] M. Pavlović and J. Riihentaus, "Quasi-nearly subharmonic functions in locally uniformly homogeneous spaces", *Positivity* **15**:1 (2011), 1–10. MR Zbl

[Yu 2017] C. Yu, "Energy conservation for the weak solutions of the compressible Navier–Stokes equations", *Arch. Ration. Mech. Anal.* **225**:3 (2017), 1073–1087. MR Zbl

Received 16 Aug 2018. Revised 18 Dec 2018. Accepted 25 Mar 2019.

IBROKHIMBEK AKRAMOV: ibrokhimbek.akramov@uni-ulm.de
Institute of Applied Analysis, Ulm University, Ulm, Germany

TOMASZ DĘBIEC: t.debiec@mimuw.edu.pl
Institute of Applied Mathematics and Mechanics, University of Warsaw, Warsaw, Poland

JACK SKIPPER: jack.skipper@uni-ulm.de
Institute of Applied Analysis, Ulm University, Ulm, Germany

EMIL WIEDEMANN: emil.wiedemann@uni-ulm.de
Institute of Applied Analysis, Ulm University, Ulm, Germany

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 13, No. 3, 2020
dx.doi.org/10.2140/apde.2020.13.813

A HIGHER-DIMENSIONAL BOURGAIN–DYATLOV FRACTAL UNCERTAINTY PRINCIPLE

RUI HAN AND WILHELM SCHLAG

We establish a version of the fractal uncertainty principle, obtained by Bourgain and Dyatlov in 2016, in higher dimensions. The Fourier support is limited to sets $Y \subset \mathbb{R}^d$ which can be covered by finitely many products of δ-regular sets in one dimension, but relative to arbitrary axes. Our results remain true if Y is distorted by diffeomorphisms. Our method combines the original approach by Bourgain and Dyatlov, in the more quantitative 2017 rendition by Jin and Zhang, with Cartan set techniques.

1. Introduction

Bourgain and Dyatlov [2018] proved the following result.

Theorem 1.1. *Let $X, Y \subset \mathbb{R}$ and $N \geq 1$ be such that $X \subset [-1, 1]$ is δ-regular with constant C_R on scales N^{-1} to 1 and $Y \subset [-N, N]$ is δ-regular with constant C_R on scales 1 to N. Then there exist constants $\beta > 0$ and C depending on δ, C_R so that*

$$\|f\|_{L^2(X)} \leq CN^{-\beta}\|f\|_{L^2(\mathbb{R})}$$

for all $f \in L^2(\mathbb{R})$ with $\operatorname{supp}(\hat{f}) \subset Y$.

The δ-regularity condition is akin to asking for a Frostman measure at dimension δ; see Definition 6.1 below for the precise statement. Theorem 1.1 is most interesting for δ close to 1. For $\delta < \frac{1}{2}$, Cauchy–Schwarz and measure estimates in phase space suffice. The β was made effective later by Jin and Zhang [2017]. Combining this fractal uncertainty principle with earlier results by Dyatlov and Zahl [2016] led to a breakthrough on the existence for an essential spectral gap for convex cocompact hyperbolic surfaces. This refers to a strip to the left of the $\frac{1}{2}$ line in the complex plane in which the Selberg zeta function has only finitely many zeros. This result can be reformulated in terms of strips below the real axis in which the meromorphic continuation of the resolvent of the Laplacian of the hyperbolic surface exhibits only finitely many resonances. This in turn can be rephrased as a decay rate of the resolvent for large energies within such a strip.

For other applications see [Bourgain and Dyatlov 2017; Dyatlov and Jin 2017; 2018], and for a survey [Dyatlov 2017].

It remained an open problem to establish an analogue of Theorem 1.1 in higher dimensions. This is the main goal of this paper.

MSC2010: 32U15, 42B30.
Keywords: uncertainty principle, fractal sets, subharmonic functions, Cartan estimates, Beurling–Malliavin theorem.

We now present our main results. Let $X \subset [-1,1]^d$ be a δ-regular set in the sense of Bourgain and Dyatlov with $\delta \in (0,d)$ and constant C_R, on scales N^{-1} to 1. In [Bourgain and Dyatlov 2018] this concept is defined only on the line, but the definition, together with its main properties, carries over to higher dimensions. Strictly speaking, we do not need the regularity condition per se, but rather the porosity property of such sets as stated precisely in Definition 5.1 below. Second, let $Y \subset [-N,N]^d$ be of the form

$$Y = \left\{ \sum_{i=1}^{d} \xi_i \vec{e}_i : \xi_i \in Y_i \right\}, \tag{1-1}$$

where $\vec{e}_i$ are unit vectors with $|\det(\vec{e}_1, \dots, \vec{e}_d)| \geq \varepsilon_0$, a positive constant (possibly small), and $Y_i \subset [-2N, 2N]$ is a δ_1-regular set with $\delta_1 \in (0,1)$ and constant C_R, on scales 1 to N.

Theorem 1.2. *Let X, Y be as in the previous paragraph in dimension $d \geq 2$. Then there exists a constant $C = C(d, \varepsilon_0, \delta, \delta_1, C_R) > 0$ such that for*

$$\beta = \exp\left\{ -\exp\left[\left(\frac{(C_R^2/\iota)^{\frac{2d-2\delta+2}{d-\delta}}}{\delta_1(1-\delta_1)} \right)^{\frac{2}{1-\delta_1}} \right] \right\},$$

where $\iota > 0$ is a small constant depending on d and ε_0, and for any $f \in L^2(\mathbb{R}^d)$ with $\operatorname{supp}(\hat{f}) \subset Y$ one has

$$\|f\|_{L^2(X)} \leq C N^{-\beta} \|f\|_{L^2(\mathbb{R}^d)} \tag{1-2}$$

for sufficiently large $N \geq N_0(d, \varepsilon_0, \delta, \delta_1, C_R)$.

As a corollary of our main theorem, we allow Y to be covered by the union of a finite number of Y_j's, each satisfying (1-1) but with a uniform ε_0:

$$Y \subset \bigcup_{j=1}^{m} Y_j, \quad \text{where } Y_j = \left\{ \sum_{i=1}^{d} \xi_{j,i} \vec{e}_{j,i} : \xi_{j,i} \in Y_{j,i} \right\}. \tag{1-3}$$

Furthermore, the number m of covers can grow in N. To be specific, we prove:

Corollary 1.3. *Let X be as above and Y be as in* (1-3). *Suppose m grows with N as follows*:

$$m = \lfloor N^{\gamma} \rfloor,$$

in which $0 \leq \gamma < \beta$. Then for any $f \in L^2(\mathbb{R}^d)$ with $\operatorname{supp}(\hat{f}) \subset Y$, and constants C, β in Theorem 1.2, one has

$$\|f\|_{L^2(X)} \leq C N^{\gamma - \beta} \|f\|_2 \tag{1-4}$$

for sufficiently large $N \geq N_0(d, \varepsilon_0, \delta, \delta_1, C_R)$.

Theorem 1.2 and Corollary 1.3 require that the Fourier support Y may be covered by products of regular sets in one dimension *along lines*; see (1-3). Our third result asserts that one may distort these lines by means of diffeomorphisms which are obtained as follows. Let $\Psi_N : [-N,N]^d \to [-N,N]^d$ be a diffeomorphism such that

$$\|D\Phi_N\|_\infty + \|D\Phi_N^{-1}\|_\infty + N\|D^2\Phi_N\|_\infty \leq C(d, D_0), \tag{1-5}$$

where the supremum norm is taken over the cube $[-N, N]^d$. One example of a diffeomorphism satisfying (1-5) is $\Psi_N(x) = N\Psi_0(x/N)$, where Ψ_0 is a diffeomorphism from $[-1, 1]^d$ to $[-1, 1]^d$ such that

$$\|D\Psi_0\|_\infty + \|D\Psi_0^{-1}\|_\infty + \|D^2\Psi_0\|_\infty \le D_0, \tag{1-6}$$

where the supremum norm is taken over the cube.

Theorem 1.4. *Theorem 1.2 remains correct with $\Phi_N(Y)$ in place of Y. Constants depend on D_0, but not on Ψ_0.*

In the following section we demonstrate the Cartan techniques by reproving a certain step in [Bourgain and Dyatlov 2018] which was proved there by means of harmonic measure of the strip with a real line-segment removed. In Section 3 we go beyond the one-dimensional setting via these Cartan methods. The subsequent sections implement the argument in analogy with [Bourgain and Dyatlov 2018] albeit in dimensions and higher. We haven striven to present the argument in a modular fashion. In particular, the delicate Beurling–Malliavin step appears only in Section 6 in order to prove the existence of *damping functions*. We do not use a higher-dimensional version of the Beurling–Malliavin theorem, which appears to be unknown. Rather, we reduce ourselves in that step to the aforementioned product structure of Y (or covers of finitely many of such products) precisely so as to be able to still use the one-dimensional construction of such damping functions. Moreover, as in [Jin and Zhang 2017] it is important for us to use the weaker form of the Beurling–Malliavin theorem obtained via outer functions; see [Mashregi et al. 2005]. Any other construction of damping functions in Section 6 would lead to different formulations of our main theorems in terms of the conditions on Y without needing to change anything in the other sections. Theorem 1.4 is proved in Section 6D. An FUP for Fourier integral operators is presented in Section 6E.

2. L^2 localization in one dimension

Let us first introduce notation. For $\xi = (\xi_1, \xi_2, \dots, \xi_d) \in \mathbb{R}^d$, let

$$|\xi|_1 := \sum_{j=1}^{d} |\xi_k|, \quad |\xi|_2 := \sum_{j=1}^{d} |\xi_k|^2, \quad \text{and} \quad \langle\xi\rangle := (1 + |\xi|_2^2)^{\frac{1}{2}}.$$

Let $e(\theta) := e^{2\pi i\theta}$. For $x \in \mathbb{R}$, let $\lceil x\rceil := \min\{n \in \mathbb{N} : n \ge x\}$, and $\lfloor x\rfloor := \max\{n \in \mathbb{N} : n \le x\}$.

Throughout, we let $\mathcal{R}(q)$ be the rectangle with vertices $\pm iq$, $1 \pm iq$. We begin with quantitative bounds on the Schwarz–Christoffel map from the disk onto a rectangle. The goal is to control this conformal mapping as the eccentricity of $\mathcal{R}(q)$ tends to 0.

Lemma 2.1. *Let $0 < q \le 1$ and define Φ_q to be the unique conformal map, continuous up to the boundary, which takes the unit disk $\mathbb{D}$ onto the rectangle $\mathcal{R}(q)$ and so that $\Phi_q(-1) = 0$ and $\Phi_q(\pm i) = \pm iq$; see Figure 1. Then $\Phi_q(1) = 1$ and $\Phi_q(e^{\pm i\theta(q)}) = 1 \pm iq$, where*

$$\theta(q) = 8\exp\Big(-\frac{\pi}{2q}\Big)(1 + O(q)), \quad q \to 0.$$

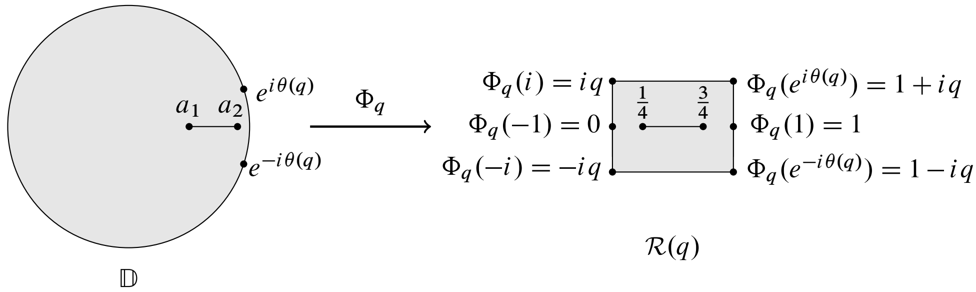

Figure 1. Conformal map Φ_q.

Moreover,

$$\Phi_q([a_1(q), a_2(q)]) = \left[\tfrac{1}{4}, \tfrac{3}{4}\right], \quad a_j(q) = 1 - \delta_j(q),$$

with

$$\delta_1(q) = 4\exp\Big(-\frac{\pi}{8q}\Big)(1 + O(q)), \quad \delta_2(q) = 4\exp\Big(-\frac{3\pi}{8q}\Big)(1 + O(q))$$

as $q \to 0$. *Let* $E \subset [a_1(q), a_2(q)]$ *be a measurable set. Then for sufficiently small* q *one has* $|\Phi_q(E)| \leq 2\delta_2(q)^{-2}|E|$, *where* $|\cdot|$ *denotes Lebesgue measure.*

Proof. Let $0 < k < 1$ and consider the elliptic integral of the first kind

$$\operatorname{arcsn}(z, k) = \int_0^z \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}, \quad \operatorname{Im} z > 0,$$

which maps the upper half-plane onto the rectangle with vertices $\pm L(k)$, $\pm L(k) + iH(k)$; see Figure 2. Here $2L(k)$ and $iH(k)$ are the periods of the elliptic function $\operatorname{sn}(z, k)$ and satisfy, as $k \to 0$,

$$L(k) = \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}} = \tfrac{\pi}{2} + O(k^2),$$

$$H(k) = \int_1^{k^{-1}} \frac{dt}{\sqrt{(t^2-1)(1-k^2t^2)}} = \int_0^\infty \frac{ds}{\sqrt{(1+s^2)(1+k^2s^2)}} = \log 4 - \log k + O(k).$$

The latter expansion is a standard fact; see for example [Abramowitz and Stegun 1966, Section 17.3.26]. Let $q := L(k)/H(k)$ and set

$$F_q(z) = -\frac{i}{H(k)} \operatorname{arcsn}(z, k), \tag{2-1}$$

which maps the upper half-plane onto the rectangle with vertices $\pm iq, 1 \pm iq$. With $k = e^{-(\pi/2)\ell}$,

$$q = \frac{\frac{\pi}{2} + O(k^2)}{\log 4 + \frac{\pi}{2}\ell + O(k)} = \ell^{-1}\Big(1 - \frac{\log 16}{\pi\ell} + O(k)\Big),$$

and thus

$$\ell = q^{-1}\Big(1 - \frac{2\log 4}{\pi} q + O(q^2)\Big), \quad k = 4\exp\Big(-\frac{\pi}{2q}\Big)(1 + O(q)).$$

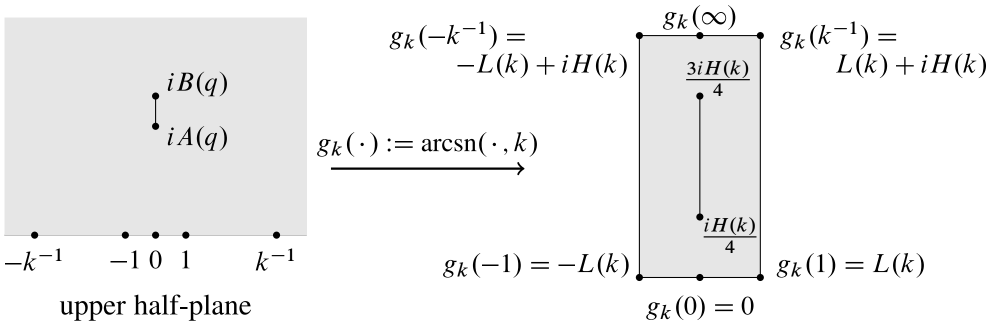

Figure 2. Elliptic integral $\operatorname{arcsn}(z, k)$.

Define $A(q), B(q)$ by $F_q(iA(q)) = \frac{1}{4}$ and $F_q(iB(q)) = \frac{3}{4}$. Thus,

$$\int_0^{A(q)} \frac{ds}{\sqrt{(1+s^2)(1+k^2s^2)}} = \tfrac{1}{4}H(k),$$
$$\int_0^{B(q)} \frac{ds}{\sqrt{(1+s^2)(1+k^2s^2)}} = \tfrac{3}{4}H(k).$$

We make the ansatz $A(q) = ck^{-1/4}(1+\varepsilon(q))$. Then

$$\begin{aligned}\int_0^{A(q)} \frac{ds}{\sqrt{(1+s^2)(1+k^2s^2)}} &= (1+O(k^{\frac{3}{2}}))\int_0^{A(q)} \frac{ds}{\sqrt{1+s^2}}\\ &= \operatorname{arcsinh}(ck^{-\frac{1}{4}}(1+\varepsilon(q)))(1+O(k^{\frac{3}{2}}))\\ &= \log(2ck^{-\frac{1}{4}}(1+\varepsilon(q)))(1+O(k^{\frac{3}{2}}))\\ &= \tfrac{1}{4}(\log 4 - \log k + O(k)).\end{aligned}$$

Hence,

$$\log(2c) - \tfrac{1}{4}\log k + \log(1+\varepsilon(q)) = \tfrac{1}{4}(\log 4 - \log k + O(k)),$$
$$c = \tfrac{1}{2}\sqrt{2}, \quad \varepsilon(q) = O(k),$$
$$A(q) = \tfrac{1}{2}\sqrt{2}\,k^{-\frac{1}{4}}(1+O(k)).$$

Similarly, with $B(q) = \tilde{c}k^{-3/4}(1+\tilde{\varepsilon}(q))$,

$$\log(2\tilde{c}) - \tfrac{3}{4}\log k + \log(1+\tilde{\varepsilon}(q)) = \tfrac{3}{4}(\log 4 - \log k + O(k))(1+O(k^{\frac{1}{2}})),$$
$$\tilde{c} = \sqrt{2}, \quad \tilde{\varepsilon}(q) = O(k^{\frac{1}{2}}\log k),$$

and so

$$B(q) = \sqrt{2}\,k^{-\frac{3}{4}}(1+O(k^{\frac{1}{2}}\log k)).$$

Expressing k in terms of q we obtain

$$A(q) = \tfrac{1}{2}\exp\Big(\frac{\pi}{8q}\Big)(1+O(q)), \quad B(q) = \tfrac{1}{2}\exp\Big(\frac{3\pi}{8q}\Big)(1+O(q)).$$

Next, we conformally map the upper half-plane $\operatorname{Im} z > 0$ onto the unit disk $|w| < 1$ via

$$z = \varphi(w) = i\,\frac{w+1}{1-w}, \quad w = \frac{z-i}{z+i}.$$

One has $\varphi(-1)=0$, $\varphi(\pm i)=\mp 1$, $\varphi(e^{i\theta})=-k^{-1}$ with $\theta=2k+O(k^3)$. Furthermore, $\varphi([a_1(q), a_2(q)]) = i\,[A(q), B(q)]$, where

$$a_1(q) = \frac{A(q)-1}{A(q)+1} = 1 - 2A(q)^{-1} + O(A(q)^{-2}),$$
$$a_2(q) = \frac{B(q)-1}{B(q)+1} = 1 - 2B(q)^{-1} + O(B(q)^{-2}).$$

Setting $a_j(q) = 1 - \delta_j(q)$ we have

$$\delta_1(q) = 4\exp\Bigl(-\frac{\pi}{8q}\Bigr)(1+O(q)), \quad \delta_2(q) = 4\exp\Bigl(-\frac{3\pi}{8q}\Bigr)(1+O(q)),$$

as claimed. The final claim of the lemma follows from

$$|(F_q \circ \varphi)'(w)| \le |F_q'(z)|\,|\varphi'(w)| \le 2(1-|w|)^{-2},$$

where $\varphi(w) = z$, $w \in (0,1)$. We used here that for $z = is$, $s > 0$,

$$|F_q'(z)| = H(k)^{-1}(1+|z|^2)^{-\frac{1}{2}}(1+k^2|z|^2)^{-\frac{1}{2}} \le H(k)^{-1}(1+|z|^2)^{-\frac{1}{2}} \le 1$$

for small q. □

By a subharmonic function v on a domain $\Omega \subset \mathbb{C}$ we mean a function $v : \Omega \to [-\infty, \infty)$, which is upper semicontinuous and satisfies the submean-value property. We recall the basic Riesz representation of a subharmonic function on the disk, albeit with precise quantitative control on the Riesz mass and the harmonic part. In view of Lemma 2.1 we need to consider the case where the lower bound on the subharmonic function is attained arbitrarily close to the boundary of the unit disk.

Lemma 2.2. *Let v be subharmonic on a neighborhood of $\mathbb{D}$, with $v \le M$ on $\mathbb{D}$, and assume $\sup_{\rho\mathbb{D}} v \ge m$ for some $0 < \rho < 1$. Let $\rho < r_1 < r < 1$. Then there exists a nonnegative measure μ on $\mathbb{D}$, called the Riesz measure, with the property that for all $w \in r\mathbb{D}$*

$$v(w) = \int_{r\mathbb{D}} \log|z-w|\,\mu(dz) + h(w), \tag{2-2}$$

with h harmonic on $r\mathbb{D}$. We have the quantitative bounds on the Riesz mass

$$\mu(r\mathbb{D}) \le \frac{M-m}{\log((1+\rho r)/(\rho+r))} \tag{2-3}$$

and on the deviations of the harmonic function

$$\min_{c\in\mathbb{R}} \max_{|w|\le r_1} |h(w)-c| \le \tfrac{1}{2}(M-m)\frac{r+r_1}{r-r_1}\,\frac{\log((1+\rho r)/(1-r^2))}{\log((1+\rho r)/(\rho+r))} =: \varepsilon. \tag{2-4}$$

The constant c which minimizes the left-hand side satisfies

$$c \geq m - \varepsilon - \log(r+\rho)\mu(r\mathbb{D}). \tag{2-5}$$

Proof. We will assume that v is smooth, the general case following by approximation. The Green's function $G : \mathbb{D} \times \mathbb{D} \to \mathbb{R}$ given by

$$G(z,w) := \frac{1}{2\pi} \log\left|\frac{z-w}{1-z\bar{w}}\right|$$

satisfies $\Delta_z G(z,w) = \delta_w$ and $G(z,w) = 0$ when $|z| = 1$.

Let $w \in \mathbb{D}$. By Green's second identity for the domain $\mathbb{D}$, we have

$$v(w) - \int_{\mathbb{D}} G(z,w)\Delta v(z) \ \mathrm{Vol}(dz) = \int_{\partial\mathbb{D}} v(z)\frac{\partial G}{\partial n_z}(z,w)\,\sigma(dz),$$

where Vol is the standard volume measure and σ is the (unnormalized) arc-length measure on the circle $\partial\mathbb{D}$. Since v is smooth and subharmonic, Δv is a nonnegative, continuous function, call it $2\pi\mu$. Therefore

$$v(w) = \int_{\mathbb{D}} 2\pi G(z,w)\,\mu(dz) + h_0(w), \tag{2-6}$$

where

$$h_0(w) := \int_{\partial\mathbb{D}} v(z)\frac{\partial G}{\partial n_z}(z,w)\,\sigma(dz). \tag{2-7}$$

Let $0 < r < 1$. On the disk $r\mathbb{D}$ we have the Riesz representation

$$v(w) = \int_{r\mathbb{D}} \log|z-w|\,\mu(dz) + h(w), \tag{2-8}$$

where

$$h(w) := \int_{\mathbb{D}\setminus r\mathbb{D}} \log\left|\frac{z-w}{1-z\bar{w}}\right|\mu(dz) - \int_{r\mathbb{D}} \log|1-z\bar{w}|\,\mu(dz) + h_0(w) \tag{2-9}$$

is harmonic in $r\mathbb{D}$. Note that $(\partial G/\partial n_z)(z,w)$ is the Poisson kernel, whence

$$h_0(w) = \int_0^1 v(e(\theta))P_{|w|}(\varphi-\theta)\,d\theta, \quad w = |w|e(\varphi). \tag{2-10}$$

We now set out to bound the Riesz measure μ. Without loss of generality, assume $m = v(\rho)$. Then setting $w = \rho$ in (2-6) yields

$$\int_{\mathbb{D}} \log\frac{|1-\rho z|}{|z-\rho|}\,\mu(dz) = h_0(\rho) - v(\rho) \leq M - m, \tag{2-11}$$

in which we used

$$h_0(\rho) \leq M. \tag{2-12}$$

This follows from the maximum principle and the fact that h_0 is the harmonic function on $\mathbb{D}$ with boundary values v by (2-10). By an elementary calculation,

$$\min_{|z|\leq r} \frac{|1-\rho z|}{|z-\rho|} = \frac{1+\rho r}{\rho + r} > 1$$

for all $0 < \rho, r < 1$. Inserting this bound into (2-11) implies

$$\mu(r\,\mathbb{D}) \le \frac{M-m}{\log((1+\rho r)/(\rho+r))}. \tag{2-13}$$

Let $\rho < r_1 < r < 1$. For all $w \in r\mathbb{D}$ we have

$$\begin{aligned} h(w) &= \int_{\mathbb{D}\setminus r\mathbb{D}} 2\pi G(w,z)\,\mu(dz) - \int_{r\mathbb{D}} \log|1-z\bar{w}|\,\mu(dz) + h_0(w) \\ &\le -\log(1-r^2)\mu(r\mathbb{D}) + M =: h^*. \end{aligned} \tag{2-14}$$

By Harnack's inequality on $r_1\mathbb{D}$ we conclude from this that for any $w \in r_1\mathbb{D}$

$$(h^* - h(w)) \le \frac{r+r_1}{r-r_1}(h^* - h(\rho)),$$

whence

$$h(w) \ge \frac{r+r_1}{r-r_1}h(\rho) - \frac{2r_1}{r-r_1}h^*.$$

By (2-8),

$$h(\rho) = v(\rho) - \int_{r\mathbb{D}} \log|z-\rho|\,\mu(dz) \ge m - \log(r+\rho)\mu(r\mathbb{D}) \tag{2-15}$$

and thus

$$h(w) \ge \frac{r+r_1}{r-r_1}(m - \log(r+\rho)\mu(r\mathbb{D})) - \frac{2r_1}{r-r_1}h^* =: h_*.$$

In summary,

$$\begin{aligned} \min_{c\in\mathbb{R}} \max_{|w|\le r_1} |h(w)-c| &\le \frac{1}{2}(h^* - h_*) \\ &= \frac{1}{2}\frac{r+r_1}{r-r_1}(h^* - m + \log(r+\rho)\mu(r\mathbb{D})) \\ &= \frac{1}{2}\frac{r+r_1}{r-r_1}\left(M - m + \log\left(\frac{r+\rho}{1-r^2}\right)\mu(r\mathbb{D})\right). \end{aligned} \tag{2-16}$$

Finally, bounding the μ-mass by (2-13) finally implies

$$\min_{c\in\mathbb{R}} \max_{|w|\le r_1} |h(w)-c| \le \tfrac{1}{2}(M-m)\frac{r+r_1}{r-r_1}\frac{\log((1+\rho r)/(1-r^2))}{\log((1+\rho r)/(\rho+r))} =: \varepsilon,$$

as claimed. Finally, to establish (2-5), we return to (2-15) and note that the left-hand side is at most $c+\varepsilon$ for c the minimizer in the previous line. Then

$$c \ge m - \log(r+\rho)\mu(r\mathbb{D}) - \varepsilon.$$

Note that one may insert (2-13) on the right-hand side to control the mass. ☐

We now apply the Cartan estimate for logarithmic potentials to the Riesz representation (2-2) in order to derive lower bounds on v up to a small measure of exceptions.

Corollary 2.3. *Let v be as in Lemma 2.2 with $\rho = 1-3\delta,\ 0<\delta<\frac{1}{3}$. Then for all $0 < H \le 1$ there exist disks $D(z_j, s_j)$ so that*

$$v(z) \ge m - (M-m)\Big[2\delta^{-3}\log\Big(\frac{2}{\delta}\Big) + \delta^{-2}\log\Big(\frac{2e}{H}\Big)\Big]$$

for all $z \in r_1\mathbb{D} \setminus \bigcup_j D(z_j, s_j)$ with $\sum_j s_j \le 5H$ and $r_1 = 1-2\delta$.

Proof. By Cartan's estimate, for any $H > 0$ there exist disks $D(z_j, s_j)$ such that $\sum_j s_j \le 5H$ and

$$\int_{r\mathbb{D}} \log|w-z|\,\mu(dw) \ge \mu(r\mathbb{D})\log\Big(\frac{H}{e}\Big) \quad \text{for all } z \in r_1\mathbb{D}\setminus\bigcup_j D(z_j, s_j). \tag{2-17}$$

See [Levin 1996, Theorem 3, Section 11.2]. To invoke the measure bound (2-3) we estimate

$$\log\Big(\frac{1+\rho r}{\rho+r}\Big) = \log\Big(\frac{2-4\delta+3\delta^2}{2-4\delta}\Big) = \log\Big(1+\frac{3\delta^2}{2-4\delta}\Big) \ge \log\Big(1+\frac{3\delta^2}{2}\Big) \ge \delta^2$$

since $\delta^2 \le \frac{1}{2}$ and $\log\big(1+\frac{3}{2}x\big) \ge x$ for $0 \le x \le \frac{1}{2}$. Consequently,

$$\mu(r\mathbb{D}) \le \delta^{-2}(M-m).$$

Next,

$$\frac{1+\rho r}{1-r^2} \le \frac{2}{2\delta-\delta^2} \le \delta^{-1}(1+\delta),$$

as well as

$$\frac{r+r_1}{r-r_1} = \frac{2-3\delta}{\delta} \le 2\delta^{-1},$$

whence (2-4) implies

$$\min_{c\in\mathbb{R}}\max_{|w|\le r_1}\ |h(w)-c| \le \varepsilon \le (M-m)\delta^{-3}\log\Big(\frac{2}{\delta}\Big) =: \tilde{\varepsilon}.$$

Finally, by (2-5), one has

$$c \ge m - \varepsilon - \log(r+\rho)\mu(r\mathbb{D}) \ge m - \varepsilon - \log(2)\mu(r\mathbb{D}).$$

In view of (2-2) and the preceding estimates we obtain

$$\begin{aligned} v(z) &\ge c + \mu(r\mathbb{D})\log\Big(\frac{H}{e}\Big) - \varepsilon \ge m - 2\varepsilon + \log\Big(\frac{H}{2e}\Big)\mu(r\mathbb{D}) \\ &\ge m - (M-m)\Big[2\delta^{-3}\log\Big(\frac{2}{\delta}\Big) - \delta^{-2}\log\Big(\frac{H}{2e}\Big)\Big] \end{aligned} \tag{2-18}$$

for all z as in (2-17). □

By means of the conformal transformation Φ_q from Lemma 2.1 we can obtain a version of the Riesz representation theorem on thin rectangles $\mathcal{R}(q)$.

Corollary 2.4. *There exists $q_* \in (0,1]$ with the following property: Let u be subharmonic on $\mathcal{R}(q)$ for some $0<q\le q_*$, continuous up to the boundary. Assume that $u\le M$ on $\mathcal{R}(q)$ and $\max_{x\in[\frac14,\frac34]} u(x)\ge m$. Then*

$$u(x) \ge m-(M-m)\exp\Big(\frac{9\pi}{8q}\Big)\Big[\log(4)+\frac{9\pi}{4q}+\exp\Big(-\frac{3\pi}{8q}\Big)\log\Big(\frac{2e}{H}\Big)\Big] \tag{2-19}$$

for all $x\in\big[\frac14,\frac34\big]\setminus\bigcup_j I_j$, where $\sum_j |I_j|\le 3H\exp\big(\frac{3\pi}{4q}\big)$.

Proof. Let $v=u\circ\Phi_q$, with Φ_q as in Lemma 2.1. Then v satisfies the assumptions of Corollary 2.3 with $\rho\ge 1-\delta_2(q)$, and

$$\begin{aligned}\delta_2(q) &= 4\exp\Big(-\frac{3\pi}{8q}\Big)(1+O(q))\ge 3\delta,\\ \delta &:= \exp\Big(-\frac{3\pi}{8q}\Big)<\frac13,\end{aligned} \tag{2-20}$$

provided q_* is small enough. By Corollary 2.3 we have

$$\begin{aligned} v(z) &\ge m-(M-m)\exp\Big(\frac{9\pi}{8q}\Big)\Big[2\log\Big(\frac{2}{\delta}\Big)+\delta\log\Big(\frac{2e}{H}\Big)\Big]\\ &= m-(M-m)\exp\Big(\frac{9\pi}{8q}\Big)\Big[\log(4)+\frac{9\pi}{4q}+\exp\Big(-\frac{3\pi}{8q}\Big)\log\Big(\frac{2e}{H}\Big)\Big]\end{aligned}$$

for all $z\in r_1\mathbb{D}\setminus\bigcup_j D(z_j,s_j)$, $\sum_j s_j\le 5H$, where $r_1=1-2\delta$. The inverse image of $\big[\frac14,\frac34\big]$ under Φ_q is $[a_1(q),a_2(q)]$. Define $\tilde I_j:=\mathbb{R}\cap D(z_j,s_j)$, $I_j=\Phi_q(\tilde I_j)$, and $E:=\bigcup_j\tilde I_j$ so that $\sum_j|\tilde I_j|\le 10H$. By Lemma 2.1 we have

$$|\Phi_q(E)|\le 20H\delta_2(q)^{-2}<3H\exp\Big(\frac{3\pi}{4q}\Big),$$

as claimed. □

Next, we apply the previous results on subharmonic functions to $\log|F|$, where F is analytic.

Corollary 2.5. *Let F be an analytic function on a neighborhood of $\mathcal{R}(q)$ with $0<q\le q^*$, and F not identically equal to zero. Define*

$$B_1:=\|F\|_{L^2([\frac14,\frac34])},\quad B_2:=\|F\|_{L^2(\partial\mathcal{R}(q))}.$$

Then for some absolute constant C_0, and all $H>0$,

$$\begin{aligned} &B_1^{K+1}\le e^{\frac{C_0K}{q}}B_2^K|F(x)|\\ \textit{holds for any}\quad &K\ge\exp\Big(\frac{9\pi}{8q}\Big)\Big[\log(4)+\frac{9\pi}{4q}+\exp\Big(-\frac{3\pi}{8q}\Big)\log\Big(\frac{2e}{H}\Big)\Big]\end{aligned} \tag{2-21}$$

for all $x\in\big[\frac14,\frac34\big]\setminus\bigcup_j I_j$, where $\sum_j|I_j|\le 3H\exp\big(\frac{3\pi}{4q}\big)$.

Proof. We apply our previous results to $u(z):=\log|F(z)|$, which is subharmonic on a neighborhood of $\mathcal{R}(q)$. However, Corollary 2.4 does not apply directly since we do not have a pointwise upper bound on u. Returning to the subharmonic function $v=u\circ\Phi_q$ on the unit disk $\mathbb{D}$, we note that the pointwise upper bound M on v only entered through the estimate $h_0\le M$; see (2-12), (2-14). The analytic function

$\widetilde{F} = F \circ \Phi_q$ satisfies $\log|\widetilde{F}| = v$. Denoting by

$$P_w(d\theta) = P_{|w|}(d(\theta-\varphi)) = \frac{1-|w|^2}{1-2|w|\cos(2\pi(\theta-\varphi))+|w|^2}$$

the Poisson kernel centered at $w = |w|e(\varphi)$, we estimate h_0 from (2-11) as follows:

$$\begin{aligned} h_0(w) &= \int_0^1 v(e(\theta))\, P_w(d\theta) = \int_0^1 \log|\widetilde{F}(e(\theta))|\, P_w(d\theta) \\ &\le \log\biggl(\int_0^1 |\widetilde{F}(e(\theta))|\, P_w(d\theta)\biggr) \\ &\le \log\biggl(\int_0^1 |\widetilde{F}(e(\theta))|\, d\theta \biggl\|\frac{P_w(d\theta)}{d\theta}\biggr\|_\infty\biggr) \\ &\le \log(B_2) + \log\biggl(\biggl\|\frac{d\theta}{d\sigma}\biggr\|_{L^2(\partial\mathcal{R}(q))}\biggr) + \log\biggl\|\frac{P_w(d\theta)}{d\theta}\biggr\|_\infty, \end{aligned} \tag{2-22}$$

where $d\sigma$ denotes arc-length measure on $\partial\mathcal{R}(q)$, and the correspondence between $\partial\mathbb{D}$ and $\partial\mathcal{R}(q)$ is given by $\xi \mapsto \Phi_q(e(\xi))$. On the one hand,

$$\biggl\|\frac{P_w(d\theta)}{d\theta}\biggr\|_\infty \le 2(1-|w|)^{-1},$$

and on the other hand,

$$\biggl\|\frac{d\theta}{d\sigma}\biggr\|^2_{L^2(\partial\mathcal{R}(q))} = \int_{\partial\mathcal{R}(q)} \biggl|\frac{d\theta}{d\sigma}\biggr|^2 d\sigma = \int_0^1 \biggl|\frac{d\sigma}{d\theta}(\xi)\biggr|^{-1} d\xi. \tag{2-23}$$

Using the notation of Lemma 2.1, the boundary map $\partial\mathbb{D} \to \partial\mathcal{R}(q)$ induced by Φ_q is

$$\begin{gathered} \xi \mapsto \zeta(\xi) := iH(k)^{-1} \operatorname{arcsn}(x(\xi), k), \\ x(\xi) := \varphi(e(\xi)) = -\cot(\pi\xi), \quad x'(\xi) = \pi(1+x(\xi)^2) \end{gathered}$$

where $\varphi(w) = i(w+1)/(1-w)$ takes the disk to the upper half-plane. If $0 < 2\pi\xi < \theta(q)$, then $\zeta(\xi) = 1 + iy(\xi)$, where

$$\frac{dy}{d\xi} = \frac{\pi}{H(k)} \frac{1+x^2}{\sqrt{(x^2-1)(k^2x^2-1)}} \ge \frac{\pi}{kH(k)}, \quad x(\xi) < -k^{-1}.$$

Therefore, this region contributes at most

$$\tfrac{1}{2}kH(k)\theta(q) \lesssim 1 \quad \text{uniformly in } q$$

to the integral in (2-23). Next, if $\theta(q) < 2\pi\xi < \frac{\pi}{2}$, then $\zeta = u + iq$, with

$$\biggl|\frac{du}{d\xi}\biggr| = \frac{\pi}{H(k)} \frac{1+x^2}{\sqrt{(x^2-1)(1-k^2x^2)}} \ge \frac{\pi}{H(k)}, \quad -k^{-1} < x(\xi) < -1,$$

and so this case contributes $\lesssim H(k)$ to (2-23). Finally, the region $\frac{\pi}{2} < 2\pi\xi < 2\pi$ similarly adds $\lesssim H(k)$ to (2-23).

Combining these estimates with (2-22) yields

$$h_0(w) \le \log(B_2) + \log(CH(k)) + \log\left(\frac{2}{\pi(1-r)}\right) \le \log(B_2) + C_0 q^{-1} =: M \tag{2-24}$$

for all $|w| < r = 1-\delta$ with some absolute constant C_0; see (2-20). This bound replaces (2-12) and (2-14) above.

As for the lower bound m on u, one has $m \ge \log(B_1)$ and thus (2-19) holds with

$$M - m \le \log\left(\frac{B_2}{B_1}\right) + C_0 q^{-1}.$$

Finally, (2-21) follows from (2-19) by exponentiating. □

Integrating the previous result over a small set of x yields the following localization estimate for the L^2 norm of F.

Proposition 2.6. *There exists an absolute constant $C_1 > 0$ with the following property: Let F be an analytic function on a neighborhood of $\mathcal{R}(q)$ with $0 < q \le q^*$, and F not identically equal to zero. Define*

$$B_1 := \|F\|_{L^2([\frac{1}{4},\frac{3}{4}])}, \quad B_2 := \|F\|_{L^2(\partial\mathcal{R}(q))}.$$

For any $J \subset \left[\frac{1}{4}, \frac{3}{4}\right]$ some Borel set of positive measure,

$$B_1 \le e^{\frac{C_1}{q}} B_2^{1-\kappa} \|F\|_{L^2(J)}^{\kappa},$$

with $0 < \kappa \le e^{-C_1/q}(\log(1/|J|))^{-1}$.

Proof. We apply Corollary 2.5 with $3H\exp\left(\frac{3\pi}{4q}\right) = |J|/2$. Thus,

$$\begin{gathered} B_1^{K+1}\left(\frac{|J|}{2}\right)^{\frac{1}{2}} \le e^{\frac{C_0 K}{q}} B_2^K \|F\|_{L^2(J)}, \\ K := \exp\left(\frac{9\pi}{8q}\right)\left[\log(4) + \frac{9\pi}{4q} + \exp\left(-\frac{3\pi}{8q}\right)\left(\log\left(\frac{12e}{|J|}\right) + \frac{3\pi}{4q}\right)\right] \end{gathered} \tag{2-25}$$

or

$$B_1 \le e^{\frac{C_0}{q}} \left(\frac{|J|}{2}\right)^{-\frac{\kappa}{2}} B_2^{1-\kappa} \|F\|_{L^2(J)}^{\kappa}, \quad \kappa \le (1+K)^{-1}. \tag{2-26}$$

We write $\kappa \le (1+K)^{-1}$ instead of $\kappa = (1+K)^{-1}$, since we may increase the value of K. One checks that

$$\log\left(\left(\frac{|J|}{2}\right)^{-\frac{\kappa}{2}}\right) \le \frac{\log(2/|J|)}{\exp\left(\frac{9\pi}{8q}\right)\left[\log(4)+\frac{9\pi}{4q}+\exp\left(-\frac{3\pi}{8q}\right)\left(\log(12e/|J|)+\frac{3\pi}{4q}\right)\right]} \le \exp\left(-\frac{3\pi}{4}\right) < 0.1, \tag{2-27}$$

uniformly in $0 < q < 1$ and in $|J|$. Note that

$$K \le \begin{cases} \exp\bigl(\frac{9\pi}{8q}\bigr)\bigl[\log(4) + \frac{9\pi}{4q} + \exp\bigl(-\frac{3\pi}{8q}\bigr)\bigl(\log(12e) + \frac{3\pi}{2q}\bigr)\bigr] & \text{if } \log 2 \le \log(1/|J|) < \frac{3\pi}{4q}, \\ 8\exp\bigl(\frac{9\pi}{8q}\bigr)\bigl[1 + \exp\bigl(-\frac{3\pi}{8q}\bigr)\bigr]\log(1/|J|) & \text{if } \max\bigl(\log 2, \frac{3\pi}{4q}\bigr) \le \log(1/|J|) \end{cases}$$
$$\le e^{\frac{C_2}{q}} \log(1/|J|) - 1$$

for some absolute constant $C_2 > 0$. Taking $C_1 := \max(2C_0, C_2)$ and

$$K_0 := e^{\frac{C_1}{q}} \log\Bigl(\frac{1}{|J|}\Bigr),$$

we conclude from (2-25), (2-26) and (2-27) with the estimate $K \le K_0 - 1$ that

$$B_1 \le e^{\frac{C_0}{q}+0.1} B_2^{1-\kappa} \|F\|_{L^2(J)}^{\kappa} \le e^{\frac{C_1}{q}} B_2^{1-\kappa} \|F\|_{L^2(J)}^{\kappa}, \quad \kappa \le K_0^{-1},$$

as claimed. □

We next apply Proposition 2.6 to a band-limited L^2 function in order to obtain the main result of this section.

Proposition 2.7. *Fix $\lambda \in \bigl(0, \frac{1}{2}\bigr]$ and for each integer n let $I_n \subset [n, n+1]$ be some Borel set with $|I_n| = \lambda$. Let $f \in L^2(\mathbb{R})$ be band-limited; i.e., $\hat{f}$ is of compact support. Then for each $0 < q \le q^*$*

$$\|f\|_{L^2(\mathbb{R})}^2 \le 12 e^{\frac{10C_1}{q}} \Bigl(\sum_n \|f\|_{L^2(I_n)}^2\Bigr)^{\kappa} \|e^{2\pi q|\xi|} \hat{f}(\xi)\|_{L^2(\mathbb{R})}^{2(1-\kappa)}, \tag{2-28}$$

with $0 < \kappa \le e^{-5C_1/q}(-\log\lambda)^{-1}$, and C_1, q^* are as in Proposition 2.6.

Proof. Let F be the entire function with $F = f$ on the real line. Fix $0 \le t \le 1$ and define $\mathcal{R}_{n,t}(q)$ to be the rectangle with vertices $n - 1 - t \pm iq$, $n + 2 + t \pm iq$. We claim that by Proposition 2.6 we have

$$\|f\|_{L^2([n,n+1])} \le e^{\frac{5C_1}{q}} \|F\|_{L^2(\partial\mathcal{R}_{n,t}(q))}^{1-\kappa} \|f\|_{L^2(I_n)}^{\kappa}, \tag{2-29}$$

with $\kappa \le e^{-5C_1/q}\bigl(\log((3+2t)/|I_n|)\bigr)^{-1}$. To see this, we set $n = 0$ without loss of generality, translate $\mathcal{R}_{n,t}(q) \to \mathcal{R}_{n,t}(q) + 1 + t$, and dilate $z \mapsto z/(3+2t)$. After these operations, the transformed interval I_0 lies in

$$\Bigl[\frac{1+t}{3+2t}, \frac{2+t}{3+2t}\Bigr] \subset \Bigl[\frac{1}{4}, \frac{3}{4}\Bigr],$$

and the height q becomes $q/(3+2t) \ge q/5$, whence the claim.

Squaring, summing, and applying Hölder's inequality yields

$$\|f\|_{L^2(\mathbb{R})}^2 \le e^{\frac{10C_1}{q}} \Bigl(\sum_n \|F\|_{L^2(\partial\mathcal{R}_{n,t}(q))}^2\Bigr)^{1-\kappa} \Bigl(\sum_n \|f\|_{L^2(I_n)}^2\Bigr)^{\kappa}.$$

Let $\mathbb{E}$ denote the expected value with respect to $0 \le t \le 1$, uniformly distributed. On the one hand, taking expectations of the previous line yields

$$\|f\|_{L^2(\mathbb{R})}^2 \le e^{\frac{10C_1}{q}} \Bigl(\sum_n \mathbb{E}\|F\|_{L^2(\partial\mathcal{R}_{n,t}(q))}^2\Bigr)^{1-\kappa} \Bigl(\sum_n \|f\|_{L^2(I_n)}^2\Bigr)^{\kappa}. \tag{2-30}$$

On the other hand, since

$$\sup_{0\le t\le 1}\sum_n \mathbb{1}_{[n-1-t,n+2+t)}\le 5, \tag{2-31}$$

we have

$$\sum_n \mathbb{E}\|F\|^2_{L^2(\partial\mathcal{R}_{n,t}(q))} \le 5\|F(\cdot+iq)\|^2_{L^2(\mathbb{R})}+5\|F(\cdot-iq)\|^2_{L^2(\mathbb{R})}+2\sum_n\int_0^1\int_{-q}^{q}|F(n-t+is)|^2\,ds\,dt. \tag{2-32}$$

Since $\|F(\cdot\pm iq)\|_{L^2(\mathbb{R})}=\|e^{\pm 2\pi q\xi}\hat{f}(\xi)\|_{L^2(\mathbb{R})}$ and

$$\begin{aligned}\sum_n\int_0^1\int_{-q}^{q}|F(n-t+is)|^2\,ds\,dt &= \int_{\mathbb{R}}\int_{-q}^{q}|F(x+is)|^2\,ds\,dx\\ &= \int_{-q}^{q}\int_{\mathbb{R}} e^{4\pi s\xi}|\hat{f}(\xi)|^2\,d\xi\,ds \le 2q\|e^{2\pi q|\xi|}\hat{f}(\xi)\|^2_{L^2(\mathbb{R})},\end{aligned}$$

assuming as we may that $q^*\le\frac{1}{2}$, we infer from (2-32) that

$$\sum_n \mathbb{E}\|F\|^2_{L^2(\partial\mathcal{R}_{n,t}(q))}\le 12\|e^{2\pi q|\xi|}\hat{f}(\xi)\|^2_{L^2(\mathbb{R})}.$$

Inserting this into (2-30) concludes the proof. □

3. L^2 localization in higher dimensions

Our goal is to prove a version of Proposition 2.7 for band-limited functions $f\in L^2(\mathbb{R}^d)$, $d\ge 2$. For the sake of simplicity, we first limit ourselves to $d=2$ and begin with a Cartan-type estimate for functions on $\mathbb{D}\times\mathbb{D}$ which are subharmonic relative to each variable.

We begin with the definition of a Cartan-2 set; see [Goldstein and Schlag 2001, Definition 8.1; 2008, Definition 2.12].

Definition 3.1. We say that $\mathcal{B}\subset\mathbb{C}^2$ is a Cartan-2 set with parameter $H>0$ if for all $(z_1,z_2)\in\mathcal{B}$ one has either

- $z_1\in\bigcup_j D(\zeta_j,s_j)$ with $\sum_j s_j\le 5H$,
- or for all other z_1, one has $z_2\in\bigcup_k D(w_k,t_k)$ with $\sum_k t_k\le 5H$ and (w_k,t_k) depend on z_1.

Of particular relevance to us with be the fact that a Cartan-2 set has a real "trace" of small measure.

Lemma 3.1. *Let $\mathcal{B}\subset\prod_{j=1}^2 D(z_{j,0},1)$ be a Cartan-2 set with parameter $H>0$. Then*

$$|\mathcal{B}\cap\mathbb{R}^2|\le 40H.$$

Proof. This follows from Fubini and $|D(\zeta,s)\cap\mathbb{R}|\le 2s$ for all $\zeta\in\mathbb{C}$. □

We can now formulate a Cartan-type bound for plurisubharmonic functions.

Lemma 3.2. *Let* $v : \overline{\mathbb{D}\times\mathbb{D}} \to [-\infty,\infty)$ *be continuous so that* $v = v(z_1,z_2)$ *is separately subharmonic in each variable. Suppose for* $0<\rho<r<1$

$$\max_{|z_1|\le r, |z_2|\le r} \int_{\mathbb{S}^1\times\mathbb{S}^1} v(e(\theta_1),e(\theta_2))\, P_{z_1}(d\theta_1)\, P_{z_2}(d\theta_2) \le M \tag{3-1}$$

and

$$\max_{|z_1|\le \rho, |z_2|\le \rho} v(z_1,z_2) \ge m. \tag{3-2}$$

Let $\rho = r(1-3\delta)$ *with* $0<\delta<\frac{1}{3}$. *Then for any* $0<H\le 1$ *one has*

$$v(z_1,z_2) \ge m-(M-m)(L+1)^2, \quad \text{where } L := 2\delta^{-3}\log\Big(\frac{2}{\delta}\Big)+\delta^{-2}\log\Big(\frac{2e}{H}\Big), \tag{3-3}$$

for all $(z_1,z_2)\in r_1\mathbb{D}\times r_1\mathbb{D}\setminus\mathcal{B}$ *where* $\mathcal{B}$ *is a Cartan-*2 *set with parameter* rH, *and* $r_1 = r(1-2\delta)$.

Proof. The function

$$h(z_1,z_2) := \int_{\mathbb{S}^1\times\mathbb{S}^1} v(e(\theta_1),e(\theta_2))\, P_{z_1}(d\theta_1)\, P_{z_2}(d\theta_2) \tag{3-4}$$

is separately harmonic in each variable, is continuous up to $\partial(\mathbb{D}\times\mathbb{D})$, and satisfies $v\le h$ pointwise. The latter property follows from the pointwise inequalities

$$v(z_1,z_2) \le \int_{\mathbb{S}^1} v(z_1,e(\theta_2))\, P_{z_2}(d\theta_2),$$

which hold due to harmonicity of the right-hand side in z_2, whence

$$v(z_1,z_2) \le \int_{\mathbb{S}^1} v(e(\theta_1),z_2)\, P_{z_1}(d\theta_1) \le \int_{\mathbb{S}^1\times\mathbb{S}^1} v(e(\theta_1),e(\theta_2))\, P_{z_1}(d\theta_1)\, P_{z_2}(d\theta_2) = h(z_1,z_2) \tag{3-5}$$

as claimed. Define

$$\tilde{v}(z_1) := \max_{|z_2|\le\rho} v(z_1,z_2). \tag{3-6}$$

Then $\tilde{v}$ is continuous (by uniform continuity) and subharmonic (as the supremum of a family of subharmonic functions). It satisfies $\tilde{v}(z_1)\le M$ for all $|z_1|\le r$ by (3-1) and (3-5), and $\max_{|z_1|\le\rho}\tilde{v}(z_1)\ge m$. The latter follows from

$$v(z_1,z_2)\le\tilde{v}(z_1) \quad \text{for all } |z_1|\le r,\ |z_2|\le\rho,$$

and (3-2).

We apply Corollary 2.3 to $\tilde{v}$, which requires rescaling from $\mathbb{D}$ to $r\mathbb{D}$. Thus, with $\rho = r(1-3\delta)$, and $r_1 = r(1-2\delta)$,

$$\tilde{v}(z_1) \ge m-(M-m)L =: m^* \tag{3-7}$$

for all $z_1\in r_1\mathbb{D}\setminus\bigcup_j D(\zeta_j,s_j)$ with $\sum_j s_j\le 5rH$. Fix such a *good* z_1. By definition, there exists z_2^* with $|z_2^*|\le\rho$ and $v(z_1,z_2^*)\ge m^*$. On the other hand, $v(z_1,z_2)\le M$ for all $|z_2|\le r$.

Once again, by Corollary 2.3 rescaled from $\mathbb{D}$ to $r\mathbb{D}$, it follows that

$$v(z_1,z_2) \ge m^*-(M-m^*)L \ge m-(M-m)L(2+L) \tag{3-8}$$

for all $z_2\in r_1\mathbb{D}\setminus\bigcup_j D(w_j,t_j)$ with $\sum_j t_j\le 5rH$. These disks depend on z_1. ∎

By means of Lemma 3.2 we establish a two-dimensional analogue of Proposition 2.6.

Proposition 3.3. *Let F be an analytic function of two variables on a neighborhood of $\mathcal{R}(q)\times\mathcal{R}(q)$ with $0<q\le q^*$, and F not identically equal to zero. Define*

$$B_1 := \|F\|_{L^2([\frac14,\frac34]\times[\frac14,\frac34])}, \quad B_2 := \|F\|_{L^2(\partial\mathcal{R}(q)\times\partial\mathcal{R}(q))}.$$

For any $J\subset[\frac14,\frac34]\times[\frac14,\frac34]$ some Borel set of positive measure,

$$B_1 \le e^{\frac{C}{q}} B_2^{1-\kappa}\|F\|_{L^2(J)}^{\kappa},$$

with $0<\kappa\le e^{-C/q}(\log(1/|J|))^{-2}$ with some absolute constant C.

Proof. Set $u(z_1,z_2):=\log|F(z_1,z_2)|$, which is plurisubharmonic on a neighborhood of $\mathcal{R}(q)\times\mathcal{R}(q)$. We pull u back to the polydisk $\mathbb{D}\times\mathbb{D}$, and define

$$v(z_1,z_2)=u(\Phi_q(z_1),\Phi_q(z_2))=\log|\widetilde{F}(z_1,z_2)|, \quad \widetilde{F}(z_1,z_2)=F(\Phi_q(z_1),\Phi_q(z_2)).$$

With h defined as in (3-4), for all $|z_1|,|z_2|\le r$,

$$\begin{aligned} h(z_1,z_2) &= \int_0^1\int_0^1 v(e(\theta_1),e(\theta_2))\,P_{z_1}(d\theta_1)\,P_{z_2}(d\theta_2) \\ &= \int_0^1\int_0^1 \log|\widetilde{F}(e(\theta_1),e(\theta_2))|\,P_{z_1}(d\theta_1)\,P_{z_2}(d\theta_2) \\ &\le \log\Big(\int_0^1\int_0^1 |\widetilde{F}(e(\theta_1),e(\theta_2))|\,P_{z_1}(d\theta_1)\,P_{z_2}(d\theta_2)\Big) \\ &\le \log\Big(\int_0^1\int_0^1 |\widetilde{F}(e(\theta_1),e(\theta_2))|\,d\theta_1\,d\theta_2 \Big\|\frac{P_{z_1}(d\theta)}{d\theta}\Big\|_\infty \Big\|\frac{P_{z_2}(d\theta)}{d\theta}\Big\|_\infty\Big) \\ &\le \log(B_2)+2\log\Big(\Big\|\frac{d\theta}{d\sigma}\Big\|_{L^2(\partial\mathcal{R}(q))}\Big)+2\sup_{|w|\le r}\log\Big\|\frac{P_w(d\theta)}{d\theta}\Big\|_\infty \\ &\le \log(B_2)+\log(Cq^{-1})+2\log\Big(\frac{2}{1-r}\Big), \end{aligned} \tag{3-9}$$

where $d\sigma$ denotes arc-length measure on $\partial\mathcal{R}(q)$; see (2-24). By Lemma 2.1, we can apply Lemma 3.2 to v with $\rho=1-\exp(-A/q)$ with some absolute constant A,

$$m=\log B_1, \quad M=\log(B_2)+3Aq^{-1}, \quad \delta=\exp\Big(-\frac{2A}{q}\Big), \quad r=\rho(1-3\delta)^{-1},$$

and $0<q\le q^*\ll 1$. Thus, for any $H>0$ there exists a Cartan-2 set $\mathcal{B}$ with parameter H such that for

$$r_1=1-\exp\Big(-\frac{A}{q}\Big)<r(1-2\delta),$$

and any $(z_1,z_2)\in r_1\mathbb{D}\times r_1\mathbb{D}\setminus\mathcal{B}$, we have

$$v(z_1,z_2)\ge m-(M-m)(L+1)^2,$$

where

$$L = 2e^{\frac{6A}{q}} \log(2e^{\frac{2A}{q}}) + e^{\frac{4A}{q}} \log\left(\frac{2e}{H}\right) < e^{\frac{8A}{q}} + e^{\frac{4A}{q}} \log\left(\frac{2e}{H}\right) - 1.$$

Returning to the original geometry, and analytic function F, we conclude the following via Lemmas 2.1 and 3.1: with $K := (e^{8A/q} + e^{4A/q} \log(2e/H))^2$,

$$B_1^{K+1} \le e^{\frac{3AK}{q}} |F(x_1, x_2)| B_2^K$$

for all $(x_1, x_2) \in \left[\frac{1}{4}, \frac{3}{4}\right] \times \left[\frac{1}{4}, \frac{3}{4}\right] \setminus \mathcal{E}$, where $\mathcal{E} \subset \mathbb{R}^2$ and $|\mathcal{E}| \le e^{5A/q} H$.

We now pick H so that $e^{5A/q} H = |J|/2$, and integrate over J, and we obtain

$$B_1^{K+1} \left(\frac{|J|}{2}\right)^{\frac{1}{2}} \le e^{\frac{3AK}{q}} B_2^K \|F\|_{L^2(J)}$$

or

$$B_1 \le e^{\frac{3A}{q}} \left(\frac{|J|}{2}\right)^{-\frac{\kappa}{2}} B_2^{1-\kappa} \|F\|_{L^2(J)}^{\kappa}, \quad \kappa \le (1+K)^{-1}. \tag{3-10}$$

We write $\kappa \le (1+K)^{-1}$ instead of $\kappa = (1+K)^{-1}$ since we could increase K. One easily checks that $(|J|/2)^{-\kappa/2} \lesssim 1$, and

$$K \le e^{\frac{C_1}{q}} \left(\log\left(\frac{1}{|J|}\right)\right)^2 - 1,$$

with some absolute constant C_1. Taking $C := \max(4A, C_1)$, and

$$K_0 := e^{\frac{C}{q}} \left(\log\left(\frac{1}{|J|}\right)\right)^2.$$

We conclude from (3-10) with the estimate $K \le K_0 - 1$ that

$$B_1 \le e^{\frac{C}{q}} B_2^{1-\kappa} \|F\|_{L^2(J)}^{\kappa}, \quad \kappa \le K_0^{-1},$$

as claimed. □

In analogy with the one-dimensional case in Proposition 2.7, we can deduce the following L^2 localization result.

Proposition 3.4. *Fix $\lambda \in \left(0, \frac{1}{2}\right]$ and for each integers n_1, n_2 let*

$$I_{n_1,n_2} \subset [n_1, n_1+1] \times [n_2, n_2+1]$$

be some Borel set with $|I_{n_1,n_2}| = \lambda$. Let $f \in L^2(\mathbb{R}^2)$ be band-limited; i.e., $\hat{f}$ is of compact support. Then for each $0 < q \le q^*$

$$\|f\|_{L^2(\mathbb{R}^2)}^2 \le e^{\frac{2C}{q}} \left(\sum_{(n_1,n_2)\in\mathbb{Z}^2} \|f\|_{L^2(I_{n_1,n_2})}^2\right)^{\kappa} \|e^{2\pi q(|\xi_1|+|\xi_2|)} \hat{f}(\xi)\|_{L^2(\mathbb{R}^2)}^{2(1-\kappa)}, \tag{3-11}$$

with $0 < \kappa \le e^{-C/q}(-\log\lambda)^{-2}$, and C some absolute constant.

Proof. Let F be the entire function with $F=f$ on $\mathbb{R}^2$. Fix $0\le t_1,t_2\le 1$ and for $j=1,2$ define $\mathcal{R}_{n,t_j}(q)$ to be the rectangle with vertices $n-1-t_j\pm iq,\ n+2+t_j\pm iq$. We obtain from Proposition 3.3 that for any $n_1,n_2\in\mathbb{Z}$

$$\|f\|_{L^2([n_1,n_1+1]\times[n_2,n_2+1])}\le e^{\frac{5C}{q}}\|F\|^{1-\kappa}_{L^2(\partial\mathcal{R}_{n_1,t_1}(q)\times\partial\mathcal{R}_{n_2,t_2}(q))}\|f\|^{\kappa}_{L^2(I_{n_1,n_2})},$$

with

$$\kappa\le e^{-\frac{5C}{q}}\left(\log\left(\frac{(3+2t_1)(3+2t_2)}{|I_{n_1,n_2}|}\right)\right)^{-2},$$

and C being the absolute constant in Proposition 3.3. Squaring, summing, and applying Hölder's inequality, we have

$$\|f\|^2_{L^2(\mathbb{R}^2)}\le e^{\frac{10C}{q}}\left(\sum_{(n_1,n_2)\in\mathbb{Z}^2}\|F\|^2_{L^2(\partial\mathcal{R}_{n_1,t_1}(q)\times\partial\mathcal{R}_{n_2,t_2}(q))}\right)^{1-\kappa}\left(\sum_{(n_1,n_2)\in\mathbb{Z}^2}\|f\|^2_{L^2(I_{n_1,n_2})}\right)^{\kappa}.$$

Taking expectation of the previous line with respect to $0\le t_1,t_2\le 1$, we obtain

$$\|f\|^2_{L^2(\mathbb{R}^2)}\le e^{\frac{10C}{q}}\left(\sum_{(n_1,n_2)\in\mathbb{Z}^2}\mathbb{E}_{t_1}\mathbb{E}_{t_2}\|F\|^2_{L^2(\partial\mathcal{R}_{n_1,t_1}(q)\times\partial\mathcal{R}_{n_2,t_2}(q))}\right)^{1-\kappa}\left(\sum_{(n_1,n_2)\in\mathbb{Z}^2}\|f\|^2_{L^2(I_{n_1,n_2})}\right)^{\kappa}. \tag{3-12}$$

By decomposing each $\partial\mathcal{R}_{n,t}(q)$ into its four sides, we decompose

$$\sum_{(n_1,n_2)\in\mathbb{Z}^2}\mathbb{E}_{t_1}\mathbb{E}_{t_2}\|F\|^2_{L^2(\partial\mathcal{R}_{n_1,t_1}(q)\times\partial\mathcal{R}_{n_2,t_2}(q))} \tag{3-13}$$

into the following three parts:

Part 1: *vertical and horizontal mixed terms.* This part contains eight terms; each can be bounded in the same way. Taking the left vertical side of $\mathcal{R}_{n_1,t_1}(q)$ and upper horizontal side of $\mathcal{R}_{n_2,t_2}(q)$ for example, we have

$$\begin{aligned}\sum_{(n_1,n_2)\in\mathbb{Z}^2}\mathbb{E}_{t_2}\int_{\mathbb{R}}\mathbb{1}_{[n_2-1-t_2,n_2+2+t_2)}\mathbb{E}_{t_1}\int_{-q}^{q}&|F(n_1-1-t_1+is,x_2+iq)|^2\,ds\,dx_2\\
&\le 5\sum_{n_1\in\mathbb{Z}}\mathbb{E}_{t_1}\int_{\mathbb{R}}\int_{-q}^{q}|F(n_1-1-t_1+is,x_2+iq)|^2\,ds\,dx_2\\
&=5\int_{-q}^{q}\int_{\mathbb{R}^2}|F(x_1+is,x_2+iq)|^2\,dx_1\,dx_2\,ds\\
&\le 5\int_{-q}^{q}\int_{\mathbb{R}^2}e^{4\pi(s\xi_1+q\xi_2)}|\hat f(\xi_1,\xi_2)|^2\,d\xi_1\,d\xi_2\,ds\\
&\le 10q\|e^{2\pi q(|\xi_1|+|\xi_2|)}\hat f(\xi)\|^2_{L^2(\mathbb{R}^2)},\end{aligned}$$

in which we used (2-31) in the first step. Hence, Part 1 contributes in total at most

$$80q\|e^{2\pi q(|\xi_1|+|\xi_2|)}\hat f(\xi)\|^2_{L^2(\mathbb{R}^2)}. \tag{3-14}$$

Part 2: *vertical+vertical sides.* This part contains four terms. Taking the left vertical sides of $\mathcal{R}_{n_1,t_1}(q)$ and $\mathcal{R}_{n_2,t_2}(q)$ for example, we have

$$\begin{aligned}\sum_{(n_1,n_2)\in\mathbb{Z}^2} \mathbb{E}_{t_1}\mathbb{E}_{t_2} \int_{-q}^{q}\int_{-q}^{q} |F(n_1-1-t_1+is_1, n_2-1-t_2+is_2)|^2\, ds_1\, ds_2 \\ = \int_{-q}^{q}\int_{-q}^{q}\int_{\mathbb{R}^2} |F(x_1+is_1, x_2+is_2)|^2\, dx_1\, dx_2\, ds_1\, ds_2 \\ \le 4q^2 \|e^{2\pi q(|\xi_1|+|\xi_2|)}\hat{f}(\xi)\|^2_{L^2(\mathbb{R}^2)}.\end{aligned}$$

Hence, Part 2 contributes in total at most

$$16q^2 \|e^{2\pi q(|\xi_1|+|\xi_2|)}\hat{f}(\xi)\|^2_{L^2(\mathbb{R}^2)}. \tag{3-15}$$

Part 3: *horizontal+horizontal sides.* This part also contains four terms. Taking the upper horizontal sides of $\mathcal{R}_{n_1,t_1}(q)$ and $\mathcal{R}_{n_2,t_2}(q)$ for example, we have

$$\begin{aligned}\sum_{(n_1,n_2)\in\mathbb{Z}^2} \mathbb{E}_{t_1}\mathbb{E}_{t_2} \int_{\mathbb{R}^2} \mathbb{1}_{[n_1-1-t_1,n_1+2+t_1)}\mathbb{1}_{[n_2-1-t_2,n_2+2+t_2)} |F(x_1+iq, x_2+iq)|^2\, dx_1\, dx_2 \\ \le 25 \int_{\mathbb{R}^2} |F(x_1+iq, x_2+iq)|^2\, dx_1\, dx_2 \\ \le 25 \|e^{2\pi q(|\xi_1|+|\xi_2|)}\hat{f}(\xi)\|^2_{L^2(\mathbb{R}^2)},\end{aligned}$$

in which we used (2-31) in the first step. Hence, the contribution of Part 3 is at most

$$100 \|e^{2\pi q(|\xi_1|+|\xi_2|)}\hat{f}(\xi)\|^2_{L^2(\mathbb{R}^2)}. \tag{3-16}$$

Plugging the estimates in (3-14), (3-15) and (3-16) into (3-13), we obtain

$$\begin{aligned}\sum_{(n_1,n_2)\in\mathbb{Z}^2} \mathbb{E}_{t_1}\mathbb{E}_{t_2} \|F\|^2_{L^2(\partial\mathcal{R}_{n_1,t_1}(q)\times\partial\mathcal{R}_{n_2,t_2}(q))} &\le (4q+10)^2 \|e^{2\pi q(|\xi_1|+|\xi_2|)}\hat{f}(\xi)\|^2_{L^2(\mathbb{R}^2)} \\ &\le 144 \|e^{2\pi q(|\xi_1|+|\xi_2|)}\hat{f}(\xi)\|^2_{L^2(\mathbb{R}^2)}\end{aligned} \tag{3-17}$$

for $q \le \frac{1}{2}$. Plugging (3-17) into (3-12) yields

$$\|f\|^2_{L^2(\mathbb{R}^2)} \le 144 e^{\frac{10C}{q}} \Bigg(\sum_{(n_1,n_2)\in\mathbb{Z}^2} \|f\|^2_{L^2(I_{n_1,n_2})}\Bigg)^{\kappa} \|e^{2\pi q(|\xi_1|+|\xi_2|)}\hat{f}(\xi)\|^{2(1-\kappa)}_{L^2(\mathbb{R}^2)},$$

as claimed. ☐

In general dimensions, one can proceed similarly. First, we inductively define Cartan sets in higher dimensions.

Definition 3.2. We say that $\mathcal{B} \subset \mathbb{C}^2$ is a Cartan-d set with parameter $H > 0$ if for all $(z_1, z_2, \dots, z_d) \in \mathcal{B}$ one has either

- $z_1 \in \bigcup_j D(\zeta_j, s_j)$ with $\sum_j s_j \le 5H$ or for all other z_1 one has
- $(z_2, \dots, z_d)$ belongs to a Cartan-$(d-1)$ set with parameter $H > 0$ depending on z_1.

By arguments analogous to those used above for $d=2$, one can exploit these Cartan sets in higher dimensions to obtain the following result. We leave the details to the reader. Throughout, we let $C(d)\ge 1$ be a constant depending only on the dimension d. It is allowed to change its values from line to line.

Proposition 3.5. *Fix $\lambda\in\left(0,\frac{1}{2}\right]$ and for each integer vector $n=(n_1,\dots,n_d)\in\mathbb{Z}^d,\ d\ge 2$, let*

$$I_n\subset\prod_{j=1}^{d}[n_j,n_j+1)$$

be some Borel set with $|I_n|=\lambda$. Let $f\in L^2(\mathbb{R}^d)$ be band-limited; i.e., $\hat{f}$ is of compact support. Then for each $0<q\le q^*=q^*(d)\ll 1$

$$\|f\|^2_{L^2(\mathbb{R}^d)}\le e^{\frac{2C(d)}{q}}\biggl(\sum_{n\in\mathbb{Z}^d}\|f\|^2_{L^2(I_n)}\biggr)^{\kappa}\|e^{2\pi q|\xi|_1}\hat{f}(\xi)\|^{2(1-\kappa)}_{L^2(\mathbb{R}^d)},\tag{3-18}$$

with $0<\kappa\le e^{-C(d)/q}(-\log\lambda)^{-d},\ C(d)\ge 1$ some absolute constant depending on d.

As a precursor to the results of the next section, which involve L^2 functions with Fourier support in thin sets, we now establish an uncertainty principle for $L^2(\mathbb{R}^d)$ functions under a quantitative decay assumption on their Fourier transforms.

Corollary 3.6. *Let $\Theta(\xi)=\Theta(|\xi|_1)=(\log(2+|\xi|_1))^{-\alpha}$, $0<\alpha<1$. Let $\mathcal{S}:=\bigcup_{n\in\mathbb{Z}^d}I_n$ be as in Proposition 3.5. Then*

$$\|f\|_2\le C(d,\alpha,A,\lambda)\|f\|_{L^2(\mathcal{S})}\tag{3-19}$$

for all $f\in L^2(\mathbb{R}^d)$ with $\|e^{\Theta(\xi)|\xi|_1}\hat{f}\|_{L^2(\mathbb{R}^d)}\le A\|f\|_{L^2(\mathbb{R}^d)}$.

Proof. With $0<q$ small to be determined, we fix $R\ge 1$ so that $2\pi q=\Theta(R)$. Split $f=f_1+f_2$, $\hat{f}_1(\xi)=\hat{f}(\xi)\mathbb{1}_{[|\xi|_1\le R]}$. Then by (3-18), and since $2\pi q\le\Theta(\xi)$ for $|\xi|_1\le R$,

$$\|f_1\|_2^2\le e^{\frac{2C(d)}{q}}\|f_1\|^{2\kappa}_{L^2(\mathcal{S})}\|e^{\Theta(\xi)|\xi|_1}\hat{f}_1\|_2^{2(1-\kappa)}\le e^{\frac{2C(d)}{q}}\|f_1\|^{2\kappa}_{L^2(\mathcal{S})}(A\|f\|_2)^{2(1-\kappa)},$$

with

$$\kappa=e^{-\frac{C(d)}{q}}(-\log\lambda)^{-d}=e^{-\frac{2\pi C(d)}{\Theta(R)}}(-\log\lambda)^{-d}.$$

Moreover, since

$$\|f\|_2^2=\|f_1\|_2^2+\|f_2\|_2^2\le e^{\frac{2C(d)}{q}}(\|f\|_{L^2(\mathcal{S})}+\|f_2\|_2)^{2\kappa}(A\|f\|_2)^{2(1-\kappa)}+\|f_2\|_2^2$$

and

$$\|f_2\|_2\le e^{-\Theta(R)R}\|e^{\Theta(\xi)|\xi|_1}\hat{f}\|_2\le Ae^{-\Theta(R)R}\|f\|_2\le\tfrac{1}{2}\|f\|_2,$$

where we chose R large enough depending on $A\ge 1$, it follows that

$$\|f\|_2^2\le 2e^{\frac{2C(d)}{q}}(\|f\|_{L^2(\mathcal{S})}+Ae^{-\Theta(R)R}\|f\|_2)^{2\kappa}(A\|f\|_2)^{2(1-\kappa)},$$

whence

$$\begin{aligned}\|f\|_2&\le 2^{\frac{1}{2\kappa}}A^{\frac{1-\kappa}{\kappa}}e^{\frac{C(d)}{\kappa q}}(\|f\|_{L^2(\mathcal{S})}+Ae^{-\Theta(R)R}\|f\|_2)\\&=2^{\frac{1}{2\kappa}}A^{\frac{1-\kappa}{\kappa}}e^{\frac{C(d)}{\kappa q}}\|f\|_{L^2(\mathcal{S})}+\exp\bigl(-T(R)\bigr)\|f\|_2,\end{aligned}$$

with

$$\begin{aligned} T(R) &= \Theta(R)R - \frac{C(d)}{\kappa q} - \kappa^{-1}\log(\sqrt{2}A) \\ &= \Theta(R)R - \left(\frac{2\pi C(d)}{\Theta(R)} + \log(\sqrt{2}A)\right) e^{\frac{2\pi C(d)}{\Theta(R)}} (-\log\lambda)^d. \end{aligned}$$

In addition to $2A \le e^{\Theta(R)R}$ we require that $T(R) \ge 1$. These conditions hold for sufficiently large R. □

The proof of the corollary gives an explicit and effective dependence of the constant $C(d,\alpha,A,\lambda)$ on A,λ, but we have no need for it. Corollary 3.6 follows (perhaps with a different dependence on the constants) from a quantitative version of the Logvinenko–Sereda theorem; see, e.g., [Kovrijkine 2001; Muscalu and Schlag 2013]. The results in the next section, however, do not.

4. Uncertainty principle with thin Fourier support

We begin with the concept of a damping function.

Definition 4.1. Let Θ be as in Corollary 3.6, with $\alpha \in (0,1)$ fixed. Let $Y \subset \mathbb{R}^d$. We say that Y *admits a damping function* with parameters c_1, c_2, c_3, all falling into the interval $(0,1)$, if there exists a function $\psi \in L^2(\mathbb{R}^d)$ satisfying

- $\operatorname{supp}(\psi) \subset [-c_1, c_1]^d$,
- $\|\widehat{\psi}\|_{L^2([-1,1]^d)} \ge c_2$,
- $|\widehat{\psi}(\xi)| \le \langle\xi\rangle^{-d}$ for all $\xi \in \mathbb{R}^d$,
- $|\widehat{\psi}(\xi)| \le \exp\bigl(-c_3\Theta(|\xi|_1)|\xi|_1\bigr)$ for all $\xi \in Y$.

Lemma 4.1. *Fix $c_1 \in \bigl(0, \tfrac{1}{2}\bigr]$ and for each integer vector $n = (n_1, \dots, n_d) \in \mathbb{Z}^d$, $d \ge 2$, let*

$$I_n \subset \prod_{j=1}^{d} [n_j, n_j + 1)$$

be a square with side length $2c_1$. Define $\mathcal{S} := \bigcup_{n\in\mathbb{Z}^d} I_n$. Suppose $Y \subset \mathbb{R}^d$ is such that $Y + [-2,2]^d$ admits a damping function with parameters c_1, and $c_2, c_3 \in (0,1)$. Then every $f \in L^2(\mathbb{R}^d)$ with $\operatorname{supp}(\hat f) \subset Y$ satisfies

$$\|\hat f\|^2_{L^2([-1,1]^d)} \le C(d) c_2^{-2} \langle R\rangle^{2d} e^{\frac{4\pi C(d)}{c_3\Theta(R)}} \bigl(\|\mathbb{1}_{\mathcal{S}} f\|^{2\kappa}_{H^{-d}} \|f\|^{2(1-\kappa)}_{H^{-d}} + \exp(-2c_3\kappa\Theta(R)R)\,\|f\|^2_{H^{-d}}\bigr) \quad (4\text{-}1)$$

and $\kappa = e^{-2\pi C(d)/(c_3\Theta(R))}(-d\log c_1)^{-d}$, provided $R \ge (2d/c_3)^2$ and $0 < c_3 \le c_3^*(d) := 2\pi q_*$, where q_* is as in Proposition 3.5.

Proof. Let $\eta \in [-2,2]^d$. Set $f_\eta(x) := e^{2\pi i x\cdot\eta} f(x)$, and $g_\eta := f_\eta * \psi$, where ψ is the damping function as in Definition 4.1 associated with $Y + [-2,2]^d$. Split g_η into

$$\begin{aligned} g_\eta &= g_1 + g_2, \\ \operatorname{supp}(\widehat{g_1}) &\subset \{\xi \in \mathbb{R}^d : |\xi|_1 \le R\}, \quad \operatorname{supp}(\widehat{g_2}) \subset \{\xi \in \mathbb{R}^d : |\xi|_1 > R\}, \end{aligned} \quad (4\text{-}2)$$

where $2\pi q = c_3\Theta(R)$. Note that our assumption $c_3 \le 2\pi q_*$ guarantees that $q \le q_*$ holds for any $R \ge 1$. Note also that since $\operatorname{supp}(\psi) \subset [-c_1, c_1]^d$, we have $\mathbb{1}_{\mathcal{S}'} g_\eta = \mathbb{1}_{\mathcal{S}'}(\mathbb{1}_{\mathcal{S}} f_\eta * \psi)$, where $\mathcal{S}' := \bigcup_{n\in\mathbb{Z}^d} I'_n$, with I'_n a square with the same center as I_n, but half the side length. By Proposition 3.5 with $\lambda = c_1^d$ one has

$$\|g_\eta\|_2^2 = \|g_1\|_2^2 + \|g_2\|_2^2 \le e^{\frac{2C(d)}{q}} \big(\|g_\eta\|_{L^2(\mathcal{S}')} + \|g_2\|_2\big)^{2\kappa} \|e^{2\pi q|\xi|_1} \widehat{g_1}\|_2^{2(1-\kappa)} + \|g_2\|_2^2, \tag{4-3}$$

with

$$0 < \kappa \le e^{-\frac{C(d)}{q}} (-d \log c_1)^{-d} = e^{-\frac{2\pi C(d)}{c_3\Theta(R)}} (-d \log c_1)^{-d},$$

$C(d)$ some absolute constant. By construction, $\operatorname{supp}(\widehat{f_\eta}) \subset Y + \eta \subset Y + [-2,2]^d$; hence

$$|\widehat{g_\eta}(\xi)| \le |\widehat{f_\eta}(\xi)| \exp\big(-c_3\Theta(|\xi|_1)|\xi|_1\big) \quad \text{for all } \xi \in \mathbb{R}^d,$$

whence

$$\|e^{2\pi q|\xi|_1}\widehat{g_1}\|_2 = \|e^{c_3\Theta(R)|\xi|_1}\widehat{g_1}\|_2 \le \sup_{|\xi|_1\le R} \langle\xi\rangle^d \|f_\eta\|_{H^{-d}} \le \langle R\rangle^d \, \|f_\eta\|_{H^{-d}},$$

$$\|g_2\|_2 \le \sup_{|\xi|_1\ge R} \exp(-c_3\Theta(|\xi|_1)|\xi|_1)\langle\xi\rangle^d \, \|f_\eta\|_{H^{-d}} \le \exp(-c_3\Theta(R)R)\langle R\rangle^d \, \|f_\eta\|_{H^{-d}},$$

where we used that $|\xi|_2 \le |\xi|_1$, and that $r \mapsto \exp(-c_3\Theta(r)r)\langle r\rangle^d$ is decreasing for large r. To be specific,

$$\exp(-c_3\Theta(r)r)\langle r\rangle^d = \exp(-h(r)), \quad h(r) = c_3(\log(2+r))^{-\alpha}\, r - \frac{d}{2}\log(1+r^2).$$

Differentiating, we obtain

$$\begin{aligned} h'(r) &= c_3(\log(2+r))^{-\alpha}\left[1 - \frac{\alpha r}{2+r}(\log(2+r))^{-1}\right] - \frac{dr}{1+r^2} \\ &\ge \frac{c_3}{2}\,(\log(2+r))^{-\alpha} - dr^{-1} \ge \frac{c_3}{2}\,(\log(2+r))^{-1} - dr^{-1}, \end{aligned}$$

where we used that

$$\frac{\alpha r}{2+r}(\log(2+r))^{-1} \le \frac{1}{2}$$

for all $r \ge 0$. One has $u > \log(2+u^2)$ for $u \ge 2$, say. Hence, if $r \ge (2d/c_3)^2$, then

$$\frac{c_3}{2}\,(\log(2+r))^{-1} - dr^{-1} > 0$$

and thus $h'(r) > 0$. So it suffices to assume that $R \ge (2d/c_3)^2$.

Inserting these bounds into (4-3) yields

$$\begin{aligned} \|g_\eta\|_2^2 \le e^{\frac{2C(d)}{q}} \big(\|\mathbb{1}_{\mathcal{S}} f_\eta\|_{H^{-d}} + \exp(-c_3\Theta(R)R)\langle R\rangle^d \, \|f_\eta\|_{H^{-d}}\big)^{2\kappa} (\langle R\rangle^d \, \|f_\eta\|_{H^{-d}})^{2(1-\kappa)} \\ + \exp(-2c_3\Theta(R)R)\langle R\rangle^{2d} \, \|f_\eta\|_{H^{-d}}^2. \end{aligned}$$

Since $\sup_{\eta\in[-2,2]^d} \|f_\eta\|_{H^{-d}} \le C(d)\|f\|_{H^{-d}}$, we can simplify this further:

$$\|g_\eta\|_2^2 \le C(d)\langle R\rangle^{2d}\, e^{\frac{4\pi C(d)}{c_3\Theta(R)}} \big(\|\mathbb{1}_{\mathcal{S}} f\|_{H^{-d}}^{2\kappa} \|f\|_{H^{-d}}^{2(1-\kappa)} + \exp(-2c_3\kappa\Theta(R)R)\, \|f\|_{H^{-d}}^2\big). \tag{4-4}$$

Finally,

$$\begin{aligned}\|\widehat{f}\|^2_{L^2([-1,1]^d)} &\le c_2^{-2}\int_{[-1,1]^d}|\widehat{f}(\zeta)|^2\,d\zeta\int_{[-1,1]^d}|\widehat{\psi}(\xi)|^2\,d\xi\\ &\le c_2^{-2}\int_{[-1,1]^d}\int_{[-2,2]^d}|\widehat{f}(\xi-\eta)|^2|\widehat{\psi}(\xi)|^2\,d\eta\,d\xi\\ &\le c_2^{-2}\int_{[-2,2]^d}\int_{\mathbb{R}^d}|\widehat{f}(\xi-\eta)|^2|\widehat{\psi}(\xi)|^2\,d\xi\,d\eta = c_2^{-2}\int_{[-2,2]^d}\|g_\eta\|_2^2\,d\eta,\end{aligned}$$

and we are done. □

We now remove the localization in Fourier space on the left-hand side of (4-1) in order to obtain the main result of this section.

Corollary 4.2. *Fix $c_1 \in \left(0, \tfrac{1}{2}\right]$ and for each integer vector $n = (n_1, \dots, n_d) \in \mathbb{Z}^d,\ d \ge 2$, let*

$$I_n \subset \prod_{j=1}^{d}[n_j, n_j + 1)$$

be a square with side length $2c_1$. Define $\mathcal{S} := \bigcup_{n\in\mathbb{Z}^d} I_n$. Suppose $Y \subset [-\alpha_1, \alpha_1]^d \subset \mathbb{R}^d$ with $\alpha_1 \ge 1$ is such that $Y + [-2,2]^d + \eta$ admits a damping function with parameters c_1, and $c_2, c_3 \in (0,1)$ for each $\eta \in [-\alpha_1 - 1, \alpha_1 + 1]^d$. Assume further that $0 < c_3 < c_3^*(d) \ll 1$, with $c_3^*(d)$ as in Lemma 4.1. Then every $f \in L^2(\mathbb{R}^d)$ with $\operatorname{supp}(\hat{f}) \subset Y$ satisfies

$$\|f\|_2 \le C_*\|f\|_{L^2(\mathcal{S})}, \tag{4-5}$$

with constant C_* depending only on d, c_1, c_2, c_3, α explicitly as in (4-15).

Proof. Let $\ell \in (2\mathbb{Z})^d$ be such that $\ell + [-1,1]^d \cap [-\alpha_1, \alpha_1]^d \ne \varnothing$ and define $f_\ell(x) := e^{2\pi i x\cdot\ell} f(x)$ so that $\widehat{f_\ell}(\xi) = \widehat{f}(\xi - \ell)$ and $\operatorname{supp}(\widehat{f_\ell}) \subset Y + \ell$. In order to apply Lemma 4.1, we also need to ensure that $Y + [-2,2]^d + \ell$ admits a damping function. This, however, follows from our assumptions. Hence, for each such ℓ,

$$\|\hat{f}\|^2_{L^2([-1,1]^d+\ell)} \le C(d)c_2^{-2}\,\langle R\rangle^{2d}\,e^{\frac{4\pi C(d)}{c_3\Theta(R)}}\left(\|\mathbb{1}_{\mathcal{S}} f_\ell\|^{2\kappa}_{H^{-d}}\|f_\ell\|^{2(1-\kappa)}_{H^{-d}} + \exp(-2c_3\kappa\Theta(R)R)\,\|f_\ell\|^2_{H^{-d}}\right) \tag{4-6}$$

and $\kappa = e^{-2\pi C(d)/(c_3\Theta(R))}(-d\log c_1)^{-d}$, provided $R \ge (2d/c_3)^2$. Summing over $\ell \in (2\mathbb{Z})^d$, and using Hölder's inequality yields

$$\begin{aligned}\|f\|_2^2 &\le C(d)c_2^{-2}\langle R\rangle^{2d}\,e^{\frac{4\pi C(d)}{c_3\Theta(R)}}\left(\|\mathbb{1}_{\mathcal{S}} f\|_2^{2\kappa}\|f\|_2^{2(1-\kappa)} + \exp(-2c_3\kappa\Theta(R)R)\|f\|_2^2\right)\\ &= C(d)c_2^{-2}\langle R\rangle^{2d}\,e^{\frac{4\pi C(d)}{c_3\Theta(R)}}\|\mathbb{1}_{\mathcal{S}} f\|_2^{2\kappa}\|f\|_2^{2(1-\kappa)} + C(d)c_2^{-2}\langle R\rangle^{2d}\,e^{\frac{4\pi C(d)}{c_3\Theta(R)}}e^{-2c_3\kappa\Theta(R)R}\|f\|_2^2.\end{aligned} \tag{4-7}$$

Suppose further that R satisfies

$$R \geq R_0(d, c_1, c_2, c_3, \alpha) := \max \begin{cases} \text{(i)} & \exp\Big[\Big(\dfrac{16\pi C(d)}{c_3}\Big)^{\frac{1}{1-\alpha}}\Big], \\ \text{(ii)} & \exp(4^{\frac{1}{1-\alpha}}), \\ \text{(iii)} & \Big(\dfrac{(-d\log c_1)^d}{c_3}\Big)^8, \\ \text{(iv)} & \Big(4\log\dfrac{2C(d)}{c_2^2}\Big)^2, \\ \text{(v)} & (8d)^4. \end{cases} \tag{4-8}$$

Note that (i), (ii), (iii) of (4-8) imply

$$e^{-\frac{2\pi C(d)}{c_3\Theta(R)}}(R+2)^{\frac14} \geq 1, \quad \Theta(R)(R+2)^{\frac18} \geq 1, \quad \text{and} \quad \frac{c_3}{(-d\log c_1)^d}(R+2)^{\frac18} \geq 1, \tag{4-9}$$

respectively. Hence multiplying the three inequalities of (4-9) yields

$$c_3\kappa\Theta(R)(R+2) \geq \sqrt{R+2} \quad \text{or} \quad \kappa \geq (c_3\Theta(R)\sqrt{R+2})^{-1}, \tag{4-10}$$

and thus

$$e^{2c_3\kappa\theta(R)R} \geq e^{c_3\kappa\theta(R)(R+2)} \geq e^{\sqrt{R+2}}. \tag{4-11}$$

One also derives from (iv), (v) and (i) that

$$\tfrac14\sqrt{R+2} \geq \log\frac{2C(d)}{c_2^2}, \quad \tfrac12\sqrt{R+2} \geq 2d\log(R+2) \geq \log\langle R\rangle^{2d}, \quad \text{and} \quad \tfrac14\sqrt{R+2} \geq \frac{4\pi C(d)}{c_3\Theta(R)}, \tag{4-12}$$

respectively. Hence by summing up the three inequalities of (4-12), and exponentiating, we obtain

$$e^{\sqrt{R+2}} \geq 2C(d)c_2^{-2}\langle R\rangle^{2d} e^{\frac{4\pi C(d)}{c_3\Theta(R)}}. \tag{4-13}$$

Combining (4-11) with (4-13), we arrive at

$$C(d)c_2^{-2}\langle R\rangle^{2d} e^{\frac{4\pi C(d)}{c_3\Theta(R)}} e^{-2c_3\kappa\Theta(R)R} \leq \tfrac12.$$

Thus (4-7) yields

$$\|f\|_2 \leq \big(2C(d)c_2^{-2}\langle R\rangle^{2d} e^{\frac{4\pi C(d)}{c_3\Theta(R)}}\big)^{\frac{1}{2\kappa}} \|\mathbb{1}_{\mathcal{S}} f\|_2.$$

Combining the estimate of κ in (4-10) with (4-13), we obtain

$$\big(2C(d)c_2^{-2}\langle R\rangle^{2d} e^{\frac{4\pi C(d)}{c_3\Theta(R)}}\big)^{\frac{1}{2\kappa}} \leq e^{\frac{c_3\Theta(R)(R+2)}{2}}.$$

Now we take R_0 as in (4-8) and define R_1 as

$$R_1(d, c_1, c_2, c_3, \alpha) := \max\bigg(\Big(\frac{2d}{c_3}\Big)^2, R_0(d, c_1, c_2, c_3, \alpha)\bigg). \tag{4-14}$$

Then

$$\|f\|_2 \leq C_*(d, c_1, c_2, c_3, \alpha)\|\mathbb{1}_{\mathcal{S}} f\|_2,$$

with

$$C_*(d, c_1, c_2, c_3, \alpha) = e^{\frac{c_3\Theta(R_1)(R_1+2)}{2}}, \tag{4-15}$$

as claimed. □

5. FUP assuming damping functions on Y

In this section we prove, by the same iteration as in [Bourgain and Dyatlov 2018], the fractal uncertainty principle for sets $X \subset [-1,1]^d$ and $Y \subset [-N,N]^d$. On Y we do not impose a geometric condition. Rather, in this section we still restrict ourselves to assuming the existence of damping functions living on Y, as well as on sets derived from Y through translations and dilations; see Definition 4.1. On X we impose a certain tree structure "with gaps"; see [Bourgain and Dyatlov 2018, Lemma 2.10].

Definition 5.1. We say that $X \subset [-1,1]^d \subset \mathbb{R}^d$ *is porous at scale* $L \geq 3$ *with depth* n, where L is an integer, if the following holds: denote by $\mathcal{C}_n$ the cubes obtained from $[-1,1]^d$ by partitioning it into congruent cubes of side length L^{-n}. Thus, $\#\mathcal{C}_n = 2^d L^{nd}$. The condition on X is that for all $Q \in \mathcal{C}_n$ with $Q \cap X \neq \varnothing$, there exists $Q' \in \mathcal{C}_{n+1}$ so that $Q' \subset Q$ and $Q' \cap X = \varnothing$.

It is shown in [Bourgain and Dyatlov 2018] that sets $X \subset \mathbb{R}$ obeying the δ-regularity condition on scales N^{-1} to 1 (see Definition 6.1) satisfy this porosity property at depth n for all $n \geq 0$ with $L^{n+1} \leq N$. We include a d-dimensional analogy in Appendix A; see Lemma A.7. We can now formulate the fractal uncertainty principle, conditionally on the existence of damping functions in Y. As in [Bourgain and Dyatlov 2018] the argument is based on an induction on scales, where at each step a small gain is achieved by means of Corollary 4.2. Recall that $\alpha \in (0,1)$ is the parameter from the damping function.

Theorem 5.1. *Let* $X \subset [-1,1]^d \subset \mathbb{R}^d$ *be porous at scale* $L \geq 3$ *with depth* n *for all* $n \geq 0$ *with* $L^{n+1} \leq N$. *Suppose* $Y \subset [-N,N]^d$ *is such that for all* $n \geq 0$ *with* $L^{n+1} \leq N$ *one has that for all*

$$\eta \in [-NL^{-n}-3, NL^{-n}+3]^d$$

the set

$$L^{-n}Y + [-4,4]^d + \eta \tag{5-1}$$

admits a damping function with parameters $c_1 = (2L)^{-1} \in \left(0, \tfrac{1}{2}\right]$, *and* $c_2, c_3 \in (0,1)$. *Assume* $0 < c_3 < c_3^*(d)$ *as in Corollary 4.2. Then there exists* $\beta = \beta(L, c_2, c_3, d, \alpha) > 0$ *and* $\widetilde{C} = \widetilde{C}(L, c_2, c_3, d, \alpha) > 0$ *so that any* $f \in L^2(\mathbb{R}^d)$ *with* $\operatorname{supp}(\hat{f}) \subset Y$ *satisfies*

$$\|f\|_{L^2(X)} \leq \widetilde{C} N^{-\beta} \|f\|_{L^2(\mathbb{R}^d)} \tag{5-2}$$

for all $N \geq N_0(L, c_2, c_3, d, \alpha)$.

Proof. We pick a nonnegative Schwartz function φ in $\mathbb{R}^d$ with $\operatorname{supp}(\hat{\varphi}) \subset [-1,1]^d$ and $\hat{\varphi}(0) = 1$. With $T \in \mathbb{N}$ to be determined, we set $\psi(x) := L^{Td}\varphi(L^T x)$ so that $\operatorname{supp}(\hat{\psi}) \subset [-L^T, L^T]^d$. Let

$$\mathcal{S}_n := \bigcup_{\substack{Q \in \mathcal{C}_n \\ Q \cap X \neq \varnothing}} Q \quad \text{and} \quad \mathcal{S}_n^* := \mathcal{S}_n + \left[-\frac{L^{-n}}{10}, \frac{L^{-n}}{10}\right]^d, \tag{5-3}$$

and define $\Psi_n := \psi_n * \mathbb{1}_{\mathcal{S}_{n+1}^*}$, where $\psi_k(x) := L^{kd}\psi(L^k x)$. There exists a constant C_φ depending only on φ such that for any $n \geq 0$

$$\Psi_n \geq \left(1 - \frac{C_\varphi}{L^{T-1}}\right)\mathbb{1}_X.$$

Thus, for all $m \geq 1$,

$$\prod_{n=0}^{m-1} \Psi_n \geq \Big(1 - \frac{C_\varphi}{L^{T-1}}\Big)^m \mathbb{1}_X. \tag{5-4}$$

Moreover, if $Q \in \mathcal{C}_{n+1}$ with $n \geq 0$ satisfies $Q \cap X = \varnothing$, denote by Q^* the cube with the same center as Q, but half the side length, i.e., of side length $L^{-(n+1)}/2$. Denote the collection of all such cubes Q^* by U_{n+1}. By the definitions of S^*_{n+1} and Q^*, we clearly have

$$S^*_{n+1} \cap \big(U_{n+1} + \big[-\tfrac{1}{10}L^{-(n+1)}, \tfrac{1}{10}L^{-(n+1)}\big]^d\big) = \varnothing.$$

Then for $x \in U_{n+1}$, and a constant c_φ that depends on φ only, we have

$$\begin{aligned} \Psi_n(x) &= \int_{\mathbb{R}^d} \varphi_{n+T}(x) \mathbb{1}_{S^*_{n+1}}(x-y)\, dy \\ &= \int_{\mathbb{R}^d} \varphi(y) \mathbb{1}_{S^*_{n+1}}(x - L^{-(n+T)} y)\, dy \\ &\leq \int_{\mathbb{R}^d \setminus [-\frac{1}{10}L^{T-1},\, \frac{1}{10}L^{T-1}]^d} \varphi(y)\, dy \leq \frac{c_\varphi}{L^{T-1}}, \end{aligned} \tag{5-5}$$

uniformly in n.

Let $f \in L^2(\mathbb{R}^d)$ with $\operatorname{supp}(\hat{f}) \subset Y$. Then for $m \geq 1$,

$$f_m := \prod_{n=0}^{m-1} \Psi_{nT} \cdot f$$

satisfies

$$\begin{aligned} \operatorname{supp}(\widehat{f_m}) &\subset Y + \sum_{n=0}^{m-1} \operatorname{supp}(\widehat{\psi_{nT}}) \\ &\subset Y + \sum_{n=0}^{m-1} [-L^{(n+1)T}, L^{(n+1)T}]^d = Y + \ell_m[-1,1]^d, \end{aligned} \tag{5-6}$$

where

$$\ell_m := L^T \frac{L^{mT}-1}{L^T - 1}.$$

One has $f_{m+1} = \Psi_{mT} f_m$ for all $m \geq 0$ with $f_0 = f$. We claim that there exists $\gamma_0 = \gamma_0(L, d, c_1, c_2, c_3) \in (0,1)$ with

$$\|f_{m+1}\|_{L^2([-1,1]^d)} \leq (1-\gamma_0)\|f_m\|_{L^2([-1,1]^d)}. \tag{5-7}$$

Define $g_m(x) := f_m(L^{mT} x)$. Then

$$\operatorname{supp}(\widehat{g_m}) \subset L^{-mT} Y + \ell_m L^{-mT}[-1,1]^d \subset L^{-mT} Y + [-2,2]^d, \tag{5-8}$$

where we used

$$\ell_m L^{-mT} \leq \frac{L^T}{L^T - 1} \leq 2.$$

In particular, assuming also that $L^{mT} \le N$,

$$\operatorname{supp}(\widehat{g_m}) \subset [-NL^{-mT}, NL^{-mT}]^d + [-2,2]^d = [-NL^{-mT}-2, NL^{-mT}+2]^d,$$

where $NL^{-mT}+2$ will be our parameter α_1 in Corollary 4.2.

Under this rescaling, the cubes in $\mathcal{C}_{mT}$ turn into unit cubes. Assuming further $L^{mT+1} \le N$, the porosity condition at scale L with depth mT ensures that we always have a "missing cube" of side length L^{-1} inside. In view of our definition of Q^*, we only use the concentric cube of half that side length. In view of the conditions on Y in the theorem we can apply Corollary 4.2 to g_m to obtain the following: with all norms being taken locally on $[-1,1]^d$, and with U_{mT+1} the missing cubes of the next generation as above,

$$\begin{aligned}\|\Psi_{mT} f_m\|_2^2 &\le \|\Psi_{mT}\|_\infty^2 \|f_m\|^2_{L^2([-1,1]^d \setminus U_{mT+1})} + \|\Psi_{mT}\|^2_{L^\infty(U_{mT+1})}\|f_m\|^2_{L^2(U_{mT+1})}\\ &\le \|f_m\|^2_{L^2([-1,1]^d \setminus U_{mT+1})} + \|\Psi_{mT}\|^2_{L^\infty(U_{mT+1})}\|f_m\|^2_{L^2(U_{mT+1})}\\ &= \|f_m\|^2_{L^2([-1,1]^d)} - (1-\|\Psi_{mT}\|^2_{L^\infty(U_{mT+1})})\|f_m\|^2_{L^2(U_{mT+1})}\\ &\le \left(1 - C_*^{-2}\left(1-\frac{c_\varphi^2}{L^{2(T-1)}}\right)\right)\|f_m\|_2^2. \end{aligned} \tag{5-9}$$

To obtain this estimate, we used that

$$\|\Psi_{mT}\|_\infty \le 1, \quad \|\Psi_{mT}\|_{L^\infty(U_{mT+1})} \le \frac{c_\varphi}{L^{T-1}},$$

and

$$\|f_m\|_{L^2(U_{mT+1})} \ge C_*^{-1}\|f_m\|_2^2,$$

with $C_* = C_*(d, L, c_2, c_3, \alpha)$ by Corollary 4.2. Choosing

$$\gamma_0(T) := \frac{1 - c_\varphi^2/L^{2(T-1)}}{2C_*^2}, \tag{5-10}$$

and using $(1-x)^{1/2} \le 1 - x/2$ for $0 \le x \le 1$, we have

$$\left(1 - C_*^{-2}\left(1-\frac{c_\varphi^2}{L^{2(T-1)}}\right)\right)^{\frac{1}{2}} \le 1 - \gamma_0(T).$$

This establishes the claim (5-7).

Applying (5-7) iteratively and using (5-4), we obtain

$$\begin{aligned}\|f\|_{L^2(X)} &\le \left(1 - \frac{C_\varphi}{L^{T-1}}\right)^{-(m+1)} \left\| \prod_{n=0}^{m} \Psi_n f \right\|_{L^2(X)}\\ &\le \left[\left(1-\frac{C_\varphi}{L^{T-1}}\right)^{-1}(1-\gamma_0(T))\right]^{m+1}\|f\|_2 \le \left(1 - \frac{\gamma_0(T)}{2}\right)^{m+1}\|f\|_2. \end{aligned} \tag{5-11}$$

In the last inequality we used

$$1 - \gamma_0(T) \le 1 - \frac{\gamma_0(T)}{2} - \frac{C_\varphi}{L^{T-1}} \le \left(1 - \frac{\gamma_0(T)}{2}\right)\left(1 - \frac{C_\varphi}{L^{T-1}}\right),$$

which requires

$$L^{T-1} - \frac{c_\varphi^2}{L^{T-1}} \geq 4C_\varphi C_*^2 \quad \text{or} \quad T \geq T_0(d, L, c_2, c_3, \alpha) := \left\lceil \frac{\log(2C_\varphi C_*^2 + \sqrt{4C_\varphi^2 C_*^4 + c_\varphi^2})}{\log L} \right\rceil. \tag{5-12}$$

Finally, for any $T \geq T_0$, taking $m \in \mathbb{N}$ be such that $L^{mT+1} \leq N < L^{(m+1)T+1}$, (5-11) yields (5-2) with

$$\beta = -\frac{\log(1-\gamma_0(T)/2)}{T \log L}, \tag{5-13}$$

and

$$\widetilde{C} = \left(1 - \frac{\gamma_0(T)}{2}\right)^{-\frac{1}{T}}, \tag{5-14}$$

as claimed. In the current theorem, we could simply choose $T = T_0$. The flexibility of choosing T will simplify our computations in our proof of Theorem 1.2. □

6. Geometry of Y and damping functions

6A. ***Regular sets.*** We will call a set $I = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d]$ of equal side lengths a *d-dimensional cube* in $\mathbb{R}^d$; we denote its side length by r_I.

Recall the notion of δ-regularity from [Bourgain and Dyatlov 2018, Definition 1.1]; below is a d-dimensional analogy.

Definition 6.1. Suppose $X \subset \mathbb{R}^d$, $X \neq \varnothing$ is closed, and $0 < \delta < d$, $C_R \geq 1$, $0 \leq \alpha_0 \leq \alpha_1 \leq \infty$. Then X is δ-regular on scales α_0 to α_1, with constant C_R, if there exists a Borel measure μ_X with the following properties:

- μ_X is supported on X.
- $\mu_X(I) \leq C_R r_I^\delta$ for each d-dimensional cube I of side length $\alpha_0 \leq r_I \leq \alpha_1$.
- $\mu_X(I) \geq C_R^{-1} r_I^\delta$ for each d-dimensional cube $I \subset \mathbb{R}^d$, centered at a point in X and of side length $\alpha_0 \leq r_I \leq \alpha_1$.

See [Bourgain and Dyatlov 2018, Section 2.2] for the geometry of such sets in $\mathbb{R}$. Loosely speaking, they behave like δ-dimensional fractal sets. The properties of δ-regular sets carry over to higher dimensions. We include some properties in Appendix A.

6B. ***Geometry of Y and damping functions.*** Bourgain and Dyatlov observed that δ-regular sets on $\mathbb{R}$ admit damping functions as in Definition 4.1 above with $\alpha = (1+\delta)/2$. They obtained these functions as a consequence of the Beurling–Malliavin theorem [1962]. However, one does not need the full strength of this theorem. To be more precise, in place of the original Beurling–Malliavin condition $\|(\log \omega)'\|_\infty < \infty$, with ω the weight, a much easier proof is possible (via outer functions) if we assume instead that $\|(H \log \omega)'\|_\infty \ll 1$ where H is the Hilbert transform on $\mathbb{R}$; see [Mashregi et al. 2005, Section 1.14, Theorem 1]. By means of this technique, Jin and Zhang [2017, Lemma 4.1] proved the following quantitative result on damping functions.

Lemma 6.1. *Let $S \geq 1$ be a constant. Let $Y \subset [-SN, SN]$ be δ_1-regular on scales 2 to N, with constant C_R, $0 < \delta_1 < 1$. For any $0 < c_1 < 1$, Y admits a damping function with $\alpha = (1+\delta_1)/2$ and parameters c_1,*

$$c_2 = \iota c_1^6, \quad c_3 = \iota c_1 C_R^{-2} \delta_1 (1-\delta_1), \tag{6-1}$$

where $\iota > 0$ is some small constant that depends on S. Instead of the pointwise global decay of $\langle \xi \rangle^{-1}$ in Definition 4.1, we have

$$|\widehat{\psi}(\xi)| \leq \exp(-c_3 \langle \xi \rangle^{\frac{1}{2}}) \quad \textit{for all } \xi \in \mathbb{R}. \tag{6-2}$$

In this paper we need a slightly different version, where we have pointwise lower bound of $|\widehat{\psi}(\xi)|$ on $\left[-\frac{3}{4}, \frac{3}{4}\right]$. The advantage of a pointwise lower bound over an L^2 bound is that it leads to a lower bound of the product of several $\widehat{\psi}$'s. Let us also note that in Lemma 4.1 of [Jin and Zhang 2017], $S = 1$. But it is clear from their proof that it works for any $S \geq 1$. We will briefly discuss the changes of constants caused by S in Appendix B. We need the extra factor S in our proof of Lemma 6.3.

Lemma 6.2. *Let $S \geq 1$ be a constant. Assume that $Y \subset [-SN, SN]$ is a δ_1-regular set with constant C_R on scales 2 to N and $\delta_1 \in (0, 1)$. Fix $0 < c_1 < 1$; then there exists a function $\psi \in L^2(\mathbb{R})$ such that*

$$\begin{aligned} &\operatorname{supp} \psi \subset \left[-\tfrac{1}{10} c_1, \tfrac{1}{10} c_1\right], \\ &|\widehat{\psi}(\xi)| \leq \exp(-c_3 \langle \xi \rangle^{\frac{1}{2}}) \qquad \textit{for all } \xi \in \mathbb{R}, \\ &|\widehat{\psi}(\xi)| \leq \exp(-c_3 \Theta(|\xi|)|\xi|) \quad \textit{for all } \xi \in Y,\ |\xi| \geq 10, \end{aligned}$$

and

$$|\widehat{\psi}(\xi)| \geq c_2 \quad \textit{for all } \xi \in \left[-\tfrac{3}{4}, \tfrac{3}{4}\right], \tag{6-3}$$

with

$$\alpha = \frac{1+\delta_1}{2}, \quad c_2 = \iota c_1^{10}, \quad c_3 = \iota c_1 C_R^{-2} \delta_1 (1-\delta_1),$$

where $\iota > 0$ is some small constant that depends on S.

We include the proof of Lemma 6.2 in Appendix B.

In higher dimensions, we reduce ourselves to this one-dimensional setting by taking finite unions of products. For simplicity, we restrict ourselves to two dimensions, although the exact analogue can be done in any finite dimension.

Definition 6.2. Pick some $\varepsilon_0 \in (0, 1)$ and let $Y \subset \mathbb{R}^2$ be of the form

$$Y \subset \bigcup_{j=1}^{m} Y_j, \quad \text{where } Y_j = \{\xi_1 \vec{e}_{j,1} + \xi_2 \vec{e}_{j,2} : \xi_i \in Y_{j,i},\ i = 1, 2\}. \tag{6-4}$$

Here $\vec{e}_{j,i} \in \mathbb{S}^1$ with $|\vec{e}_{j,1} \cdot \vec{e}_{j,2}| < 1 - \varepsilon_0$ for all $1 \leq j \leq m$, and $Y_{j,i}$ are δ_1-regular on scales α_0 to α_1 with constant C_R, where $0 < \delta_1 < 1$. In that case Y is called *admissible on scales α_0 to α_1* with parameters δ_1, C_R, ε_0, m. In general dimensions, we require that $\vec{e}_{j,i}$ are unit vectors with $|\det(\vec{e}_{j,1}, \dots, \vec{e}_{j,d})| \geq \varepsilon_0$; see (1-3).

Throughout, we will freeze ε_0 and constants are allowed to depend on it. The admissible sets on scale 2 to N that are contained in $[-N,N]^d$ carry damping functions.

We note that for our proof of Theorem 1.2, we only need $m=1$. We give a construction with arbitrary $m\geq 1$ here, since the construction itself may be of independent interest.

Lemma 6.3. *Let $Y\subset[-N,N]^2$ be admissible on scales 2 to N as in Definition 6.2. Then Y admits a damping function with parameters c_1,*

$$c_2=\iota^{2m+4}c_1^{20m+4}\,m^{-20m}\,C_R^{-8}(\delta_1(1-\delta_1))^4,$$
$$c_3=\iota c_1\,m^{-1}C_R^{-2}\delta_1(1-\delta_1),$$

where $\iota>0$ is a small constant that depends on ε_0.

Remark 6.4. For general dimension d, we can take

$$c_2=\iota^m c_1^{(10m+2)d}\,m^{-10md}\,C_R^{-4d}(\delta_1(1-\delta_1))^{2d},$$
$$c_3=\iota c_1\,m^{-1}C_R^{-2}\delta_1(1-\delta_1),$$

where $\iota>0$ is a small constant that depends on ε_0 and d.

Proof. Let $\psi_{j,i}$ be the damping function associated with $Y_{j,i}\subset[-SN,SN]$, with $S=S(\varepsilon_0)\geq 1$, via Lemma 6.2 with parameters $\tilde{c}_1:=\varepsilon_1 c_1 m^{-1}$, where ε_1 is a small parameter depending on ε_0, and c_2,c_3 are as given by Lemma 6.2, but in terms of $\tilde{c}_1$; i.e.,

$$c_2=\iota\varepsilon_1^{10}\,c_1^{10}m^{-10},$$
$$c_3=c_1\,m^{-1}\iota\varepsilon_1\,C_R^{-2}\delta_1(1-\delta_1),$$

where ι depends ε_0. We will absorb the constant ε_1 into ι. In the following we will also allow ι to change its value from line to line, as long as it only depends on ε_0.

Denote the coordinates associated with the basis $\vec{e}_{j,1}$, $\vec{e}_{j,2}$ by $(\xi_{j,1},\xi_{j,2})$. We set, with $\xi\in\mathbb{R}^2$,

$$\widehat{\psi}(\xi):=\prod_{j=1}^m\widehat{\psi}_j(\xi),\quad \widehat{\psi}_j(\xi):=\widehat{\psi_{j,1}}(\xi_{j,1})\widehat{\psi_{j,2}}(\xi_{j,2}).$$

Then

$$|\widehat{\psi}_j(\xi)|\leq\exp(-c_3\langle\xi_{j,1}\rangle^{\frac{1}{2}})\exp(-c_3\langle\xi_{j,2}\rangle^{\frac{1}{2}})\leq\exp(-c_3\langle\xi\rangle^{\frac{1}{2}}),\tag{6-5}$$

where c_3, more precisely, ι, can change its value in the last line depending on ε_0. Taking products gives

$$|\widehat{\psi}(\xi)|\leq\exp(-mc_3\langle\xi\rangle^{\frac{1}{2}})=\exp(-c_1\nu\langle\xi\rangle^{\frac{1}{2}}),\quad \nu=\iota C_R^{-2}\delta_1(1-\delta_1).\tag{6-6}$$

In particular, $\psi\in L^2(\mathbb{R}^2)$ as well as $\psi_j\in L^2(\mathbb{R}^2)$. Since ψ_j are also compactly supported functions, $\psi_j\in L^1(\mathbb{R}^2)$. Hence in the sense of L^1 functions,

$$\psi=\mathop{\Large *}_{j=1}^m\psi_j,$$

whence

$$\operatorname{supp}(\psi) \subset \sum_{j=1}^{m} \operatorname{supp}(\psi_j) \subset \sum_{j=1}^{m} [-c_1 m^{-1}, c_1 m^{-1}]^2 \subset [-c_1, c_1]^2,$$

where we used that each $\psi_{j,i}$ is a damping function with $\tilde{c}_1 = \varepsilon_1 c_1 m^{-1}$. Next, if $\xi \in Y_j$, then

$$|\widehat{\psi}_j(\xi)| \le \exp(-c_3 \Theta(|\xi_{j,1}|)|\xi_{j,1}|) \exp(-c_3 \Theta(|\xi_{j,2}|)|\xi_{j,2}|) \le \exp(-c_3 \Theta(|\xi|_1)|\xi|_1),$$

where again ι is allowed to change in the second line. Since Y is covered by the union of Y_j, we have

$$|\widehat{\psi}(\xi)| \le \exp(-c_3 \Theta(|\xi|_1)|\xi|_1) \quad \text{for all } \xi \in Y. \tag{6-7}$$

Finally, from (6-3), for each $1 \le j \le m$,

$$|\widehat{\psi}_j(\xi)| \ge c_2^2 \quad \text{for all } \xi_{j,1}, \xi_{j,2} \in \left[-\tfrac{3}{4}, \tfrac{3}{4}\right].$$

Hence,

$$\|\widehat{\psi}\|_{L^2([-1,1]^2)} \ge c_2^{2m} |E|^{\frac{1}{2}},$$

where E is the subset of $[-1,1]^2$ where all conditions $\xi_{j,i} \in \left[-\frac{3}{4}, \frac{3}{4}\right]$, $i = 1, 2$, $1 \le j \le m$, are met. Clearly, $|E|^{1/2}$ is some number depending on ε_0. It follows that

$$\|\widehat{\psi}\|_{L^2([-1,1]^2)} \ge \iota^{2m}\, c_1^{20m} m^{-20m}, \tag{6-8}$$

where ι depends on ε_0.

We required $|\widehat{\psi}(\xi)| \le \langle \xi \rangle^{-2}$ in our definition of damping function; see Definition 4.1. Since for any $0 < \rho < 1$

$$\exp(-\rho \langle \xi \rangle^{\frac{1}{2}}) \le 5\rho^{-4} \langle \xi \rangle^{-2},$$

it follows from (6-6) that $\tilde{\psi} := \frac{1}{5}(c_1 \nu)^4 \psi$ is a damping function in the sense of the definition. Since $\frac{1}{5}(c_1\nu)^4 \le 1$, the decay (6-7) remains intact, as does the support condition. However, (6-8) needs to be modified:

$$\|\widehat{\tilde{\psi}}\|_{L^2([-1,1]^2)} \ge \tfrac{1}{5}(c_1 \nu)^4 \iota^{2m}\, c_1^{20m} m^{-20m} = \tfrac{1}{5} \iota^{2m+4} c_1^{20m+4} m^{-20m} C_R^{-8} (\delta_1(1-\delta_1))^4.$$

Absorbing the $\frac{1}{5}$ into ι, the lemma is proved. □

Finally, we need to check that Y remains admissible if it is transformed by the similarities in (5-1).

Lemma 6.5. *Let $Y \subset [-N, N]^d$ with $N \ge 10$ be admissible on scales 2 to N with parameters δ_1, C_R, ε_0, m. Let $L \ge 4$ be an integer. Then for all integers $n \ge 0$ with $L^{n+1} \le N$ and for all*

$$\eta \in [-NL^{-n} - 3, NL^{-n} + 3]^d,$$

the set

$$L^{-n} Y + [-4, 4]^d + \eta \subset [-(2NL^{-n} + 7), 2NL^{-n} + 7]^d$$

is admissible at scale $S(2NL^{-n} + 7)$ with parameters δ_1, $576 S^2 C_R$, ε_0, m, where $S = S(\varepsilon_0, d) \ge 1$.

Proof. First,

$$L^{-n}Y+[-4,4]^d+\eta\subset[-2NL^{-n}-7,2NL^{-n}+7]^d$$

for all η as above. Second, by (6-4),

$$L^{-n}Y+[-4,4]^d+\eta\subset\bigcup_{j=1}^{m}(L^{-n}Y_j+[-4,4]^d+\eta),$$

where

$$L^{-n}Y_j=\left\{\sum_{k=1}^{d}\xi_k\vec{e}_{j,k}:\xi_k\in L^{-n}Y_{j,k},\ k=1,2,\dots,d\right\},$$

and

$$L^{-n}Y_j+[-4,4]^d+\eta\subset\left\{\sum_{k=1}^{d}\xi_k\vec{e}_{j,k}:\xi_k\in L^{-n}Y_{j,k}+[-4S,4S]+\eta_{j,k},\ k=1,2,\dots,d\right\},$$

where $S=S(\varepsilon_0,d)\geq 1$ and $|\eta_{j,k}|\leq S(NL^{-n}+3)$. By Lemmas 2.1, 2.2, and 2.3 in [Bourgain and Dyatlov 2018], see also Lemmas A.2, A.3, and A.4 with $d=1$, the sets

$$L^{-n}Y_{j,k}+[-4S,4S]+\eta_{j,k}\subset[-S(2NL^{-n}+7),S(2NL^{-n}+7)]$$

are δ_1-regular with constant $576S^2C_R$ on scales 2 to $S(2NL^{-n}+7)$. Indeed, for $n\geq 1$, Lemma A.2 implies that $L^{-n}Y_{j,k}$ is δ_1-regular on scales $2L^{-n}\leq\frac{1}{2}$ to $L^{-n}N$ with constant C_R. Lemma A.4 implies that

$$L^{-n}Y_{j,k}+[-4S,4S]=L^{-n}Y_{j,k}+8S\left[-\tfrac{1}{2},\tfrac{1}{2}\right]$$

is δ_1-regular on scales 1 to $L^{-n}N$ with constant $32SC_R$. Lemma A.3 allows us to increase the upper scale from $L^{-n}N$ to $9SL^{-n}N\geq S(2L^{-n}N+7)$, with changing the constant from $32SC_R$ to $576S^2C_R$. Note that shifting a set does not change its δ_1-regularity; hence $L^{-n}Y_{j,k}+[-4S,4S]+\eta_{j,k}$ is δ_1-regular with constant $576S^2C_R$. The proof for $n=0$ is similar.

The lemma now follows from Definition 6.2. □

6C. *Proof of Theorem 1.2.*

Proof. The proof of Theorem 1.2 is now a corollary to Theorem 5.1 and the considerations in this section, with $m=1$. We will keep track of various constants in order to obtain the effective exponent β.

First, let

$$L:=\lceil(2^{\frac{d}{2}}\sqrt{2d+1}C_R)^{\frac{2}{d-\delta}}\rceil\geq 4$$

be as in (A-3). Lemma A.7 implies that for all $n\geq 0$ with $L^{n+1}\leq N$, X is porous at scale L with depth n. This verifies the porosity condition on X in Theorem 5.1.

Combining Lemma 6.3, more specifically Remark 6.4, with Lemma 6.5, we obtain that for any $n\in\mathbb{N}$ such that $L^{n+1}\leq N$, and for all $\eta\in[-L^{-n}N-3,L^{-n}N+3]^d$, the set

$$L^{-n}Y+[-4,4]^d+\eta$$

admits a damping function with parameters c_1,

$$c_2 = \iota c_1^{12d}(576S^2C_R)^{-4d}(\delta_1(1-\delta_1))^{2d},$$
$$c_3 = \iota c_1(576S^2C_R)^{-2}\delta_1(1-\delta_1),$$

where ι and S are constants depending on ε_0. We absorb the constant S into ι, and allow ι to depend on d as well. Hence we can simply write

$$c_2 = \iota c_1^{12d}C_R^{-4d}(\delta_1(1-\delta_1))^{2d},$$
$$c_3 = \iota c_1 C_R^{-2}\delta_1(1-\delta_1).$$

Note that this verifies the condition on Y in Theorem 5.1.

Before applying Theorem 5.1, let us first determine the constant C_* in Corollary 4.2 with c_1, c_2, c_3 defined above. Recall that

$$C_* = e^{\frac{c_3\Theta(R_1)(R_1+2)}{2}},$$

with $\alpha = (1+\delta_1)/2$ and let

$$R_1 = \max\begin{cases}\exp\Big[\Big(\frac{C_R^2}{\iota c_1\delta_1(1-\delta_1)}\Big)^{\frac{2}{1-\delta_1}}\Big]\\ \exp(4^{\frac{2}{1-\delta_1}})\\ \Big(\frac{C_R^2(-\log c_1)^d}{\iota c_1\delta_1(1-\delta_1)}\Big)^8\\ \Big[4\log\Big(\frac{C_R^{8d}}{\iota c_1^{24d}(\delta_1(1-\delta_1))^{4d}}\Big)\Big]^2\\ (8d)^4\\ \frac{C_R^4}{\iota c_1^2(\delta_1(1-\delta_1))^2}\end{cases} \tag{6-9}$$

be as in (4-14), in which we absorb all the d-dependent constants into ι.

Now we can apply Theorem 5.1 with

$$c_1 = (2L)^{-1} = (2\lceil(2^{\frac{d}{2}}\sqrt{2d+1}C_R)^{\frac{2}{d-\delta}}\rceil)^{-1}.$$

We need to trace out the constant β.

Plugging c_1 into (6-9), and making ι smaller if necessary (depending only on d and ε_0), we have

$$R_1 \le \exp\Bigg[\Bigg(\frac{(C_R^2/\iota)^{\frac{2d-2\delta+2}{d-\delta}}}{\delta_1(1-\delta_1)}\Bigg)^{\frac{2}{1-\delta_1}}\Bigg] =: R_2.$$

This implies

$$C_* = \exp\big(c_1C_R^{-2}\delta_1(1-\delta_1)\Theta(R_1)(R_1+2)\big) \le \exp(R_2).$$

Recall T_0 as in (5-12) and γ_0 as in (5-10). We compute that

$$T_0 = \Bigg\lceil\frac{\log(2C_\varphi C_*^2 + \sqrt{4C_\varphi^2C_*^4 + c_\varphi^2})}{\log L}\Bigg\rceil \le \frac{2\log C_* + \log(5C_\varphi)}{\log L} \le \frac{2R_2 + \log(5C_\varphi)}{\log L} =: T_1, \tag{6-10}$$

and

$$\gamma_0(T_1) = \frac{1 - c_\varphi^2/L^{2(T_1-1)}}{2C_*^2} \ge \frac{1}{4C_*^2} \ge \frac{1}{4}\exp(-2R_2). \tag{6-11}$$

In both inequalities above, we used $C_* \le \exp(R_2)$.

Recall β as in (5-13). Using that $-\log(1-x) \ge x$ for $x < 1$, we have

$$\beta = -\frac{\log(1-\gamma_0(T_1)/2)}{T_1 \log L} \ge \frac{\gamma_0(T_1)}{2T_1 \log L}.$$

Combining this with the estimates of T_1 and $\gamma_0(T_1)$ as in (6-10) and (6-11), we have

$$\beta \ge \exp\biggl\{-\exp\biggl[\biggl(\frac{(C_R^2/\iota)^{\frac{2d-2\delta+2}{d-\delta}}}{\delta_1(1-\delta_1)}\biggr)^{\frac{2}{1-\delta_1}}\biggr]\biggr\},$$

with ι being a small constant depending on ε_0 and d. □

Corollary 1.3 follows from Theorem 1.2 by the triangle inequality.

Remark 6.6. If we try to combine the construction of a damping function for m covers as in Lemma 6.3, with Theorem 5.1, we could allow m to grow in N like $\log\log\log N$. This is worse than the power-law growth obtained via the triangle inequality.

6D. ***Distortion of Y by diffeomorphisms.*** Let $\mathcal{F}_\hbar$ be the unitary semiclassical Fourier transform on $L^2(\mathbb{R}^d)$ defined by

$$\mathcal{F}_\hbar f(\xi) = \hbar^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{2\pi i x\cdot\xi}{\hbar}} f(x)\,dx = \hbar^{-\frac{d}{2}} \hat{f}\Bigl(\frac{\xi}{\hbar}\Bigr).$$

We will use the following proposition which roughly says that the intersection of an admissible set with a cube is still admissible. We only work with admissible sets with $m = 1$ throughout this section.

Proposition 6.7. *Let $Y \subset \mathbb{R}^d$ be an admissible set on scales N^{-1} to 1 with parameters $\delta_1, C_R, \varepsilon_0$. Let $Q \subset \mathbb{R}^d$ be a cube of side length $r_Q \le r_0$. Then*

$$Y \cap Q \subseteq \bigcup_{j=1}^{C(\varepsilon_0,d,r_0)} W_j,$$

where each W_j is contained in a cube of side length $C(\varepsilon_0, d)$, and is admissible on scales N^{-1} to 1 with parameters δ_1, $(4C_R)^{2/(1-\delta_1)}C_R, \varepsilon_0$.

Proof. Let $Y = \{\sum_{k=1}^d \xi_k \vec{e}_k : \xi_k \in Y_k\}$, where $\vec{e}_k \in \mathbb{S}^1$ and $|\det(\vec{e}_1, \dots, \vec{e}_d)| \ge \varepsilon_0$. We cover Q by the smallest parallelepiped $\widetilde{Q}$, whose edges are determined by $\vec{e}_1, \dots, \vec{e}_d$, that contains Q. We can write $\widetilde{Q} = \{\sum_{k=1}^d \xi_k \vec{e}_k : \xi_k \in \widetilde{Q}_k\}$.

By Lemma A.1, there exist disjoint intervals $\mathcal{J}_k$ such that

$$Y_k = \bigcup_{J_{k,\ell} \in \mathcal{J}_k} (Y_k \cap J_{k,\ell}), \qquad \text{with } (4C_R)^{-\frac{2}{1-\delta_1}} \le |J_{k,\ell}| \le 1 \quad \text{for all } J_{k,\ell} \in \mathcal{J}_k,$$

where the $(Y_k \cap J_{k,\ell})$'s are δ_1-regular sets with constant $\widetilde{C}_R = (4C_R)^{2/(1-\delta_1)} C_R$ on scales N^{-1} to 1. For any $\underline{\ell} \in \mathbb{N}^d$, let $Y_{\underline{\ell}} := \{\sum_{k=1}^d \xi_k \vec{e}_k : \xi_k \in Y_k \cap J_{k,\ell_k}\}$. Hence $Y_{\underline{\ell}}$ is admissible on scales N^{-1} to 1 with parameters $\delta_1, \widetilde{C}_R, \varepsilon_0$. Furthermore, $Y_{\underline{\ell}}$ is contained in a cube of side length $C(\varepsilon_0, d)$. Finally note that $\widetilde{Q}_k$ intersects at most finitely many $J_{k,\ell}$'s, and this number depends only on ε_0, d and r_0. □

In this section we prove Theorem 1.4. We need to show that Theorem 1.2 remains valid if an admissible set Y is distorted by a diffeomorphism $\Phi_N(x)$ from the cube $[-N, N]^d \to [-N, N]^d$; see (1-5). The argument is related to Section 4 of [Bourgain and Dyatlov 2018]. Thus, let $Y = \Phi_N(\widetilde{Y})$, where $\widetilde{Y} \subset [-N, N]^d$ is an admissible set with constants C_R, ε_0 on scales 1 to N. Suppose $f \in L^2(\mathbb{R}^d)$ with $\operatorname{supp}(\widehat{f}) \subset Y$ and set $\widehat{g} := \widehat{f} \circ \Phi_N$ so that $\operatorname{supp}(\widehat{g}) \subset \widetilde{Y}$. Furthermore,

$$\begin{aligned} f(x) &= \int_{[-N,N]^d} e^{2\pi i x\cdot\xi}\, \widehat{f}(\xi)\, d\xi = \int_{[-N,N]^d} e^{2\pi i x\cdot\xi}\, \widehat{g}(\Phi_N^{-1}(\xi))\, d\xi \\ &= \int_{[-N,N]^d} e^{2\pi i x\cdot\Phi_N(\eta)}\, \widehat{g}(\eta)\, |\det(D\Phi_N(\eta))|\, d\eta. \end{aligned} \tag{6-12}$$

We claim that for some $\beta > 0$ and $C > 0$ depending on all the same parameters in Theorem 1.2 as well as on D_0

$$\left\| \int_{[-N,N]^d} e^{2\pi i x\cdot\Phi_N(\eta)}\, \widehat{h}(\eta)\, d\eta \right\|_{L^2(X)} \le C N^{-\beta} \|h\|_2 \tag{6-13}$$

for all $h \in L^2$ with $\operatorname{supp}(\widehat{h}) \subset \widetilde{Y}$, in which $\widetilde{Y} \subset [-N, N]^d$ is an admissible set with constants C_R, ε_0 on scales 1 to N. Setting $\widehat{h}(\eta) := \widehat{g}(\eta)\, |\det(D\Phi_N(\eta))|$, we conclude from (6-13) that

$$\|f\|_{L^2(X)} \le C N^{-\beta} \|\widehat{h}\|_2 \le C N^{-\beta} \|\widehat{f}\|_2 = C N^{-\beta} \|f\|_2,$$

with possibly a different constant. So it remains to prove the claim (6-13). We will prove it from another statement, namely

$$\left\| \int_{[-N,N]^d} e^{2\pi i x\cdot\Phi_N(\eta)}\, \mathbb{1}_{\widetilde{Y}}(\eta)\, h(\eta)\, d\eta \right\|_{L^2(X)} \le C N^{-\beta} \|h\|_2 \tag{6-14}$$

for all $h \in L^2$. Notice that by Plancherel we could remove the Fourier transform from h.

To prove (6-14), divide $[-N, N]^d = \bigcup_k Q_k$ into congruent cubes of side length L_N with $\frac{1}{2}\sqrt{N} \le L_N \le \sqrt{N}$. Let $\{\chi_k\}_k$ be a partition of unity adapted to these cubes. With η_k being the center of Q_k,

$$\begin{aligned} \int_{[-N,N]^d} e^{2\pi i x\cdot\Phi_N(\eta)}\, \mathbb{1}_{\widetilde{Y}}(\eta)\, h(\eta)\, d\eta &= \sum_k \int_{\mathbb{R}^d} e^{2\pi i x\cdot\Phi_N(\eta)}\, \chi_k(\eta)\, \mathbb{1}_{\widetilde{Y}}(\eta)\, h(\eta)\, d\eta \\ &= \sum_k \int_{\mathbb{R}^d} e^{2\pi i x\cdot(\Phi_N(\eta_k) + D\Phi_N(\eta_k)(\eta-\eta_k))}\, a_k(x,\eta)\, \mathbb{1}_{\widetilde{Y}}(\eta)\, h(\eta)\, d\eta \\ &=: \sum_k (T_k h)(x), \end{aligned} \tag{6-15}$$

where

$$\begin{aligned} a_k(x,\eta) &:= e^{2\pi i x\cdot R_k(\eta)}\,\chi_k(\eta), \\ R_k(\eta) &:= \int_0^1 (1-t)\langle D^2\Phi_N(\eta_k+t(\eta-\eta_k))(\eta-\eta_k),\eta-\eta_k\rangle\,dt, \end{aligned} \tag{6-16}$$

the latter being the error in the second-order Taylor expansion (we are suppressing the parameter N here). Then

$$\begin{aligned} &\|R_k\|_{L^\infty(\operatorname{supp}\chi_k)} \le C = C(d,D_0), \\ &\|\partial_x^\alpha a_k(x,\eta)\|_{L^\infty([-1,1]^d\times\operatorname{supp}\chi_k)} \le C(d,D_0,\alpha), \quad \operatorname{diam}\operatorname{supp}\chi_k \le C\sqrt{N}, \end{aligned} \tag{6-17}$$

for every multi-index α. By Hörmander's variable-coefficient Plancherel theorem,

$$\max_k \|T_k\|_{2\to 2} \le C(d,D_0). \tag{6-18}$$

This follows by the usual T^*T argument:

$$\begin{aligned} \|T_k h\|_2^2 &= \langle T_k^*T_k h,h\rangle, \\ (T_k^*T_k h)(\eta') &= \int_{\mathbb{R}^d} K_k(\eta',\eta)\,h(\eta)\,d\eta, \\ K_k(\eta',\eta) &= \int_{\mathbb{R}^d} e^{2\pi i x\cdot(\Phi_N(\eta)-\Phi_N(\eta'))}\,\mathbb{1}_{\widetilde{Y}}(\eta)\,\mathbb{1}_{\widetilde{Y}}(\eta')\,\chi_k(\eta)\chi_k(\eta')\,dx. \end{aligned} \tag{6-19}$$

Since $\|\Phi_N(\eta)-\Phi_N(\eta')\| \ge D_0^{-1}\|\eta-\eta'\|$ in the sense of Euclidean lengths, repeated integrations by parts yield the decay

$$|K_k(\eta',\eta)| \le C(d,D_0)\langle\eta-\eta'\rangle^{-d-1},$$

whence (6-18) follows by Schur's test. In particular, $\|\mathbb{1}_X T_k\|_{2\to 2} \le C$ with the same constant as in (6-18).

Next, we would like to show that $\mathbb{1}_X T_k$ and $\mathbb{1}_X T_\ell$ do not interact much for all cubes Q_k, Q_ℓ which are not nearest neighbors. In order to integrate by parts in x, see (6-19), we need to smooth out $\mathbb{1}_X$ at the correct scale. Define

$$X(N^{-\frac{1}{2}}) := X + [-N^{-\frac{1}{2}}, N^{-\frac{1}{2}}]^d.$$

By [Dyatlov and Zahl 2016, Lemma 3.3] there exists a smooth ψ taking values in $[0,1]$ with $\psi = 1$ on X and with $\operatorname{supp}(\psi) \subset X(N^{-1/2})$, as well as so that

$$\|\partial_x^\alpha \psi\|_\infty \le C(\alpha) N^{\frac{|\alpha|}{2}} \tag{6-20}$$

for all multi-indices. Define $S_k := \psi\,T_k$. On the one hand, S_k still obeys (6-18). On the other hand, for any cubes Q_k, Q_ℓ which are not nearest neighbors one has

$$\|S_k^* S_\ell\|_{2\to 2} \le C(d,D_0,p) N^{\frac{p}{2}}\ \operatorname{dist}(Q_k,Q_\ell)^{-p} \tag{6-21}$$

for every positive integer p. This follows from the fact that the kernel of $S_k^*S_\ell$ equals

$$K_{k,\ell}(\eta',\eta) = \int_{\mathbb{R}^d} e^{2\pi i x\cdot(\Phi_N(\eta)-\Phi_N(\eta'))}\,\mathbb{1}_{\widetilde{Y}}(\eta)\,\mathbb{1}_{\widetilde{Y}}(\eta')\,\chi_k(\eta)\chi_\ell(\eta')\,\psi(x)^2\,dx.$$

Using the differential operator

$$\mathcal{L} = \frac{1}{2\pi i}\frac{\Phi_N(\eta)-\Phi_N(\eta')}{\|\Phi_N(\eta)-\Phi_N(\eta')\|^2}\cdot\nabla_x,$$

which obeys

$$\mathcal{L}\, e^{2\pi i x\cdot(\Phi_N(\eta)-\Phi_N(\eta'))} = e^{2\pi i x\cdot(\Phi_N(\eta)-\Phi_N(\eta'))},$$

repeated integration by parts now yields (6-21). Finally, given any k, only a uniformly bounded number of choices of ℓ will satisfy

$$S_k S_\ell^* = \psi\, T_k T_\ell^*\, \psi \neq 0.$$

This is due to the fact that $\chi_k(\eta)\chi_\ell(\eta) = 0$ up to a bounded number of choices of ℓ given k. If we label the cubes by lattice points $\underline{k}\in\mathbb{Z}^d$, then $\eta_{\underline{k}} = L_N\underline{k}$, whence

$$N^{\frac{p}{2}}\,\mathrm{dist}(Q_{\underline{k}},Q_{\underline{\ell}})^{-p} \lesssim N^{\frac{p}{2}}(L_N|\underline{k}-\underline{\ell}|)^{-p} \lesssim |\underline{k}-\underline{\ell}|^{-p},$$

which is summable over $\mathbb{Z}^d$ provided $p > d$. On the other hand, we also have

$$\|S_k^* S_\ell\|_{2\to 2} \le \|S_k\|_{2\to 2}\|S_\ell\|_{2\to 2} \le B^2, \quad B := \sup_j \|S_j\|_{2\to 2}.$$

Combining these two estimates we infer that for any $0<\varepsilon<1$

$$\|S_{\underline{k}}S_{\underline{\ell}}^*\|_{2\to 2} + \|S_{\underline{k}}^* S_{\underline{\ell}}\|_{2\to 2} \le C(d,D_0,\varepsilon)\, B^{2(1-\varepsilon)}\langle \underline{k}-\underline{\ell}\rangle^{-2(d+1)}$$

for all $\underline{k},\underline{\ell}\in\mathbb{Z}^d$. Note that $B\le C(d,D_0)$ by Hörmander's bound (6-18). Hence by Cotlar's lemma,

$$\left\|\int_{[-N,N]^d} e^{2\pi i x\cdot\Phi_N(\eta)}\,\mathbb{1}_{\widetilde{Y}}(\eta)\,h(\eta)\,d\eta\right\|_{L^2(X)} \le C(\varepsilon,d,D_0)\,\max_k \|S_k\|_{2\to 2}^{1-\varepsilon}. \tag{6-22}$$

The claim (6-14) will now follow from (6-22) by applying the fractal uncertainty principle of Theorem 1.2 to each S_k. For this we need to linearize the phase as in (6-15), which in turn makes the localization to scales $\sqrt{N}$ necessary.

To be specific, we reduce (6-14) to the following estimate. Let ψ_0 be compactly supported functions satisfying the bounds

$$\|\partial_x^\alpha \psi_0\|_\infty \le C_s N^s \quad \text{for all } |\alpha| = s \ge 0, \tag{6-23}$$

where $N\ge 1$ is arbitrary and all constant are independent of N. We assume that ψ_0 is supported in a δ-regular set in $[-1,1]^d$ on scales $1/N$ to 1, and with $0<\delta<d$. Let

$$Z = N^{-1}Y_1$$

be a rescaled version of an admissible set Y_1 with constants C_R, δ_1, ε_0 on scales 1 to N. The point is Y_1 is not assumed to be contained in $[-N,N]^d$; hence Theorem 1.2 does not apply directly. Hence we need to use Proposition 6.7 instead, for which we need to make assumptions on supp a. Suppose that the symbol a is smooth and compactly supported with the bounds

$$\|\partial_x^\alpha a(x,\xi)\|_\infty \le C(\alpha) \quad \text{for all } \alpha, \qquad \text{and} \qquad \operatorname{supp} a(x,\cdot)\subset Q, \tag{6-24}$$

where Q is a cube in $\mathbb{R}^d$ that is independent of x, and is of side length $r_Q \le r_0$. Then for some $\beta > 0$ and C as above,

$$\|\psi_0\, A\, \mathbb{1}_Z\, h\|_2 \le C N^{-\beta} \|h\|_2, \tag{6-25}$$

where

$$(Ah)(x) := N^{\frac{d}{2}} \int_{\mathbb{R}^d} e^{2\pi i N x\cdot\xi}\, a(x,\xi) h(\xi)\, d\xi.$$

Indeed,

$$\begin{aligned}
&\left\| \int_{\mathbb{R}^d} e^{2\pi i x\cdot(\Phi_N(\eta_k)+D\Phi_N(\eta_k)(\eta-\eta_k))}\, \psi(x)\, a_k(x,\eta)\, \mathbb{1}_{\widetilde{Y}}(\eta)\, h(\eta)\, d\eta \right\|_2 \\
&\quad\lesssim \left\| \int_{\mathbb{R}^d} e^{2\pi i x\cdot\zeta}\, \psi(x)\, a_k(x, D\Phi_N(\eta_k)^{-1}\zeta+\eta_k)\, \mathbb{1}_{\widetilde{Y}-\eta_k}(D\Phi_N(\eta_k)^{-1}\zeta) h(D\Phi_N(\eta_k)^{-1}\zeta+\eta_k)\, d\zeta \right\|_2 \\
&\quad= N^{\frac{d}{4}} \left\| \int_{\mathbb{R}^d} e^{2\pi i N^{1/2} x\cdot\xi}\, \psi(x)\, \tilde{a}_k(x, N^{\frac12}\xi)\, N^{\frac{d}{4}} \mathbb{1}_{Y_1}(N^{\frac12}\xi)\, \tilde{h}(N^{\frac12}\xi)\, d\xi \right\|_2.
\end{aligned}$$

Here $\tilde{a}$, $\tilde{h}$ signify the functions on the second line but with the linear isomorphism $D\Phi_N(\eta_k)^{-1}$ and the shift η_k included, and $Y_1 = D\Phi_N(\eta_k)(\widetilde{Y}-\eta_k)$ is an admissible set on scales 1 to N with constants that depend on D_0. Note that $\mathbb{1}_{Y_1}(N^{1/2}\xi) = \mathbb{1}_Z(\xi)$, with $Z = N^{-1/2} Y_1$, which is an admissible set on scales $N^{-1/2}$ to 1. By (6-20), $\psi_0(x) := \psi(x)$ satisfies the required bound, and furthermore ψ_0 is supported on $X(N^{-1/2})$, which is a δ-regular set on scales $N^{-1/2}$ to 1; see Lemma A.4. As for the amplitude, ignoring the distinction between $\tilde{a}_k$ and a_k,

$$\begin{aligned}
a_k(x, N^{\frac12}\xi) &:= e^{2\pi i x\cdot R_k(N^{1/2}\xi)}\, \chi_k(N^{\frac12}\xi), \\
R_k(N^{\frac12}\xi) &:= N \int_0^1 (1-t)\langle D^2\Phi_N(\eta_k + t(N^{\frac12}\xi-\eta_k))(\xi-\eta_k'), \xi-\eta_k'\rangle\, dt,
\end{aligned}$$

where $\eta_k' = N^{-1/2}\eta_k$. Setting $a(x,\xi) = a_k(x, N^{1/2}\xi)$, we conclude from (6-17) that a satisfies (6-24) with constant $r_0 = C$, which is an absolute constant. Finally,

$$\|N^{\frac{d}{4}} \tilde{h}(N^{\frac12}\xi)\|_2 \simeq \|h\|_2.$$

Thus, we can apply (6-25) with N replaced by $N^{1/2}$ to obtain a gain of $N^{-\beta/2}$, and we are done.

It remains to prove (6-25). Note that this is equivalent to proving

$$\|\psi_0\, A\, \mathbb{1}_{Z\cap Q}\, h\|_2 \le C N^{-\beta} \|h\|_2. \tag{6-26}$$

By Proposition 6.7, we can cover $Z\cap Q$ by $C(\varepsilon_0, d, r_0)$ many admissible sets W_j with constants δ_1, $\widetilde{C}_R := (4C_R)^{2/(1-\delta_1)} C_R$, $\tilde{\varepsilon}_0 = \tilde{\varepsilon}_0(\varepsilon_0, D_0)$. Hence, via triangle inequality, it suffices to prove (6-26) with $Z\cap Q$ replaced by W_j.

If $a = 1$ on the $\operatorname{supp}(\psi_0)\times W_j$, then this follows immediately from Theorem 1.2 by a rescaling. Indeed, one has by that theorem

$$\begin{aligned}
\left\| N^{\frac{d}{2}} \int_{\mathbb{R}^d} e^{2\pi i N x\cdot\xi}\, \psi_0(x)\, \mathbb{1}_{W_j}(\xi)\, h(\xi)\, d\xi \right\|_2 &= \left\| \int_{\mathbb{R}^d} e^{2\pi i x\cdot\xi}\, \psi_0(x)\, \mathbb{1}_{W_j}\Big(\frac{\xi}{N}\Big)\, N^{-\frac{d}{2}} h\Big(\frac{\xi}{N}\Big)\, d\xi \right\|_2 \\
&\lesssim N^{-\beta} \left\| N^{-\frac{d}{2}} h\Big(\frac{\xi}{N}\Big) \right\|_2 = N^{-\beta} \|h\|_2.
\end{aligned}$$

Let us now consider general a satisfying (6-24). Let $\rho \in (0,1)$ with its value determined later.

Let us note that by the usual A^*A argument, we have Hörmander's bound,

$$\|A\|_{2\to 2} \le C. \tag{6-27}$$

Next we decompose $\psi_0 A \mathbb{1}_{W_j}$ into

$$\psi_0 A \mathbb{1}_{W_j} = \psi_0 \mathcal{F}_\hbar^{-1} A_1 + A_2 \mathcal{F}_\hbar A \mathbb{1}_{W_j},$$
$$A_1 := \mathbb{1}_{\mathbb{R}^d \setminus W_j(N^{-\rho})} \mathcal{F}_\hbar A \mathbb{1}_{W_j}, \quad A_2 := \psi_0 \mathcal{F}_\hbar^{-1} \mathbb{1}_{W_j(N^{-\rho})},$$

where $\hbar = N^{-1}$. Clearly, by (6-27), we have

$$\|\psi_0 A \mathbb{1}_{W_j}\|_{2\to 2} \lesssim \|A_1\|_{2\to 2} + \|A_2\|_{2\to 2}. \tag{6-28}$$

Thus it suffices to bound $\|A_1\|_{2\to 2}$ and $\|A_2\|_{2\to 2}$.

We compute the integral kernel of A_1:

$$K_{A_1}(\xi,\eta) = \mathbb{1}_{\mathbb{R}^d \setminus W_j(N^{-\rho})}(\xi)\, \mathbb{1}_{W_j}(\eta)\, N^d \int_{\mathbb{R}^d} e^{2\pi i N x\cdot(\eta-\xi)} a(x,\eta)\, dx.$$

Note that the Euclidean distance satisfies $\|\eta-\xi\| \ge N^{-\rho}$ on the support of K_{A_1}. Hence by repeated integration by parts in x, we obtain that

$$|K_{A_1}(\xi,\eta)| \le C_{d,\rho} N^{d-\lceil \frac{d+10}{1-\rho} \rceil} \langle \eta-\xi \rangle^{-\lceil \frac{d+10}{1-\rho} \rceil} \le C_{d,\rho} N^{-10}.$$

By Schur's test, we arrive at

$$\|A_1\|_{2\to 2} \le C N^{-10}. \tag{6-29}$$

In view of A_2. Note that

$$W_j(N^{-\rho}) \subset \bigcup_{\substack{\|k\|_\infty \le N^{1-\rho} \\ k \in \mathbb{Z}^d}} (W_j(N^{-1}) + k),$$

and

$$W_j(N^{-1}) \subset \widehat{W}_j := \left\{ \sum_{\ell=1}^d \xi_\ell \vec{e}_\ell : \xi_\ell \in N^{-1} \cdot W_{j,\ell}(2) \right\},$$

which is an admissible set on scales $2N^{-1}$ to 1. Thus by Theorem 1.2 and triangle inequality, we have for $f \in L^2(\mathbb{R}^d)$

$$\begin{aligned} \|A_2 f\| &\le \sum_{\|k\| \le N^{1-\rho}} \|\psi_0 \mathcal{F}_\hbar^{-1} \mathbb{1}_{\widehat{W}_j + k} f\|_2 \\ &\lesssim \sum_{\|k\| \le N^{1-\rho}} \|\mathbb{1}_{\operatorname{supp} \psi_0} \mathcal{F}_\hbar^{-1} \mathbb{1}_{\widehat{W}_j + k} f\|_2 \le C N^{-\beta + d(1-\rho)} \|f\|_2, \end{aligned}$$

where $\hbar = N^{-1}$. Hence for $\rho = 1 - \beta/2d$,

$$\|A_2\|_{2\to 2} \le C N^{-\frac{\beta}{2}}. \tag{6-30}$$

Combining (6-28), (6-29) with (6-30), we obtain (6-25). This concludes the proof of Theorem 1.4.

6E. ***Fourier integral operator.*** In this section, we prove a fractal uncertainty principle for Fourier integral operators on $\mathbb{R}^d$. The proof follows that of the one-dimensional case in [Bourgain and Dyatlov 2018, Section 4]; thus we shall be very brief.

Let

$$(B(\hbar)f)(x) := \hbar^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{2\pi i \Phi(x,y)}{\hbar}} b(x,y) f(y)\, dy, \tag{6-31}$$

where for some open set $U \subset \mathbb{R}^{2d}$

$$\begin{gathered} \Phi \in C^\infty(U;\mathbb{R}), \quad b \in C_0^\infty(U), \quad \det\left(\frac{\partial^2 \Phi}{\partial x_j \partial y_k}\right) \neq 0 \quad \text{on } U, \\ \left(\sup_U \left\| \left(\frac{\partial^2 \Phi}{\partial x_j \partial y_k}\right) \right\|\right) \cdot \left(\sup_U \left\| \left(\frac{\partial^2 \Phi}{\partial x_j \partial y_k}\right)^{-1} \right\|\right) \le C_\Phi \end{gathered} \tag{6-32}$$

for some constant $C_\Phi \ge 1$, in which $\|\cdot\|$ is the matrix norm.

Proposition 6.8. *Let $X, Y \subset [-1,1]^d$. Assume that X is a δ-regular set on scales* 0 *to* 1 *with constant C_R, and Y is an admissible set on scales* 0 *to* 1 *with parameters δ_1, C_R, ε_0. Assume* (6-32) *holds. Then there exist $\beta > 0$, $\rho \in (0,1)$ depending only on δ, δ_1, C_R, ε_0, d, C_Φ, and $C > 0$ depending only on δ, δ_1, C_R, ε_0, d, Φ, b such that for $0 < \hbar < h_0(\Phi) < 1$,*

$$\|\mathbb{1}_{X(\hbar^{\rho/2})} B(\hbar) \mathbb{1}_{Y(\hbar^\rho)}\|_{L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)} \le C\hbar^\beta.$$

Proof. As was pointed out in [Bourgain and Dyatlov 2018], it is enough to prove Proposition 6.8 under the assumption that

$$1 < \left|\det\left(\frac{\partial^2 \Phi}{\partial x_j \partial y_k}\right)\right| < 2 \quad \text{on } U. \tag{6-33}$$

Let $\tilde{h} := \hbar^{1/2}$. Divide $[-2,2]^d = \bigcup_k Q_k$ into congruent cubes of side length L with $\tilde{h}/2 \le L < \tilde{h}$. Let $\{\chi_k\}_k$ be a partition of unity adapted to these cubes. With y_k being the center of Q_k, we have

$$\begin{aligned} &\hbar^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{2\pi i \Phi(x,y)}{\hbar}} b(x,y) \mathbb{1}_{Y(\hbar^\rho)}(y) f(y)\, dy \\ &\qquad = \sum_k \hbar^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{2\pi i \Phi(x,y)}{\hbar}} b(x,y) \chi_k(y) \mathbb{1}_{Y(\hbar^\rho)}(y) f(y)\, dy \\ &\qquad = \sum_k e^{-\frac{2\pi i \Phi(x,y_k)}{\hbar}} \hbar^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{2\pi i \nabla_y \Phi(x,y_k)\cdot(y-y_k)}{\hbar}} \tilde{b}_k(x,y) \mathbb{1}_{Y(\hbar^\rho)}(y) f(y)\, dy \\ &\qquad =: \sum_k (T_k f)(x), \end{aligned}$$

where

$$\begin{aligned} \tilde{b}_k(x,y) &= e^{-\frac{2\pi i \Psi_k(x,y)}{\hbar}} \chi_k(y) b(x,y), \\ \Psi_k(x,y) &= \int_0^1 (1-t)\langle (y-y_k), \mathrm{H}\Phi(x, y_k + t(y-y_k))(y-y_k)\rangle\, dt, \end{aligned} \tag{6-34}$$

in which $\mathrm{H}\Phi(x,\cdot)$ is the Hessian of $\Phi(x,\cdot)$ in the y-variable.

We will prove

$$\|\mathbb{1}_{X(\tilde{h}^\rho)} T_k\|_{L^2\to L^2} \le C\tilde{h}^\beta, \tag{6-35}$$

and the estimate for $\sum_k \mathbb{1}_{X(\tilde{h}^\rho)} T_k$ follows from almost orthogonality and Cotlar's lemma; see the proof of Proposition 4.3 in [Bourgain and Dyatlov 2018].

Let

$$\varphi(x) := \nabla_y \Phi(x, y_k).$$

By (6-33), the Jacobian matrix $\mathrm{J}\varphi$ satisfies $1 < |\det(\mathrm{J}\varphi(x))| < 2$; hence φ admits an inverse function.

We have, by a change variable $x \to \varphi^{-1}(x)$,

$$\begin{aligned}
&\|\mathbb{1}_{X(\tilde{h}^\rho)}(x)(T_k f)(x)\|_{L^2} \\
&\quad = \|\mathbb{1}_{\varphi(X(\tilde{h}^\rho))}(x)|\det(\mathrm{J}\varphi^{-1}(x))|^{\frac{1}{2}}\hbar^{-\frac{d}{2}}\int_{\mathbb{R}^d} e^{-\frac{2\pi i x\cdot y}{\hbar}}\tilde{b}_k(\varphi^{-1}(x), y+y_k)\mathbb{1}_{Y(\hbar^\rho)-y_k}(y) f(y+y_k)\,dy\|_{L^2} \\
&\quad = \|\mathbb{1}_{\varphi(X(\tilde{h}^\rho))}(x)|\det(\mathrm{J}\varphi^{-1}(x))|^{\frac{1}{2}}\int_{\mathbb{R}^d} e^{-\frac{2\pi i x\cdot y}{\tilde{h}}}\tilde{b}_k(\varphi^{-1}(x), \tilde{h}y+y_k)\mathbb{1}_{Y(\hbar^\rho)-y_k}(\tilde{h}y) f(\tilde{h}y+y_k)\,dy\|_{L^2} \\
&\quad \le \|\mathbb{1}_{\varphi(X(\tilde{h}^\rho))}\, A(\hbar)\,\mathbb{1}_{\tilde{h}^{-1}(Y(\tilde{h}^{2\rho})-y_k)}\|_{L^2\to L^2}\cdot\|\tilde{h}^{\frac{d}{2}} f(\tilde{h}y+y_k)\|_{L^2} \\
&\quad = \|\mathbb{1}_{\varphi(X(\tilde{h}^\rho))}\, A(\hbar)\,\mathbb{1}_{Y(\tilde{h}^{2\rho-1})-\tilde{h}^{-1}y_k}\|_{L^2\to L^2}\cdot\|f\|_{L^2},
\end{aligned}$$

where

$$\begin{aligned}
(A(\hbar)f)(x) &= \tilde{h}^{-\frac{d}{2}}\int_{\mathbb{R}^d} e^{-\frac{2\pi i x\cdot y}{\tilde{h}}}\hat{b}_k(x,y) f(y)\,dy, \\
\tilde{b}(x,y) &= |\det(\mathrm{J}\varphi^{-1}(x))|^{\frac{1}{2}}\,\tilde{b}_k(\varphi^{-1}(x), \tilde{h}y+y_k).
\end{aligned} \tag{6-36}$$

Now it suffices to bound

$$\|\mathbb{1}_{\varphi(X(\tilde{h}^\rho))}\, A(\hbar)\,\mathbb{1}_{Y(\tilde{h}^{2\rho-1})-\tilde{h}^{-1}y_k}\|_{L^2\to L^2}. \tag{6-37}$$

Let $\widetilde{X} := \varphi(X)$. By (6-32),

$$(\sup\|\mathrm{J}\varphi\|)\cdot(\sup\|(\mathrm{J}\varphi)^{-1}\|) \le C_\Phi.$$

Note (6-33) implies $C_1 := \sup\|\mathrm{J}\varphi\| \ge 1$ and hence $C_2 := \sup\|(\mathrm{J}\varphi)^{-1}\| \le C_\Phi$. By Lemma A.5, $\widetilde{X}$ is δ-regular with constant $C_R(d\,C_\Phi)^{\delta/2}$ on scales 0 to $d^{-1/2}C_2^{-1}$.

If $d^{-1/2}C_2^{-1} < 1$, Lemma A.3 implies $\widetilde{X}$ is δ-regular with constant

$$2(d^{\frac{1}{2}}C_2)^d C_R(d\,C_\Phi)^{\frac{\delta}{2}} \le 2d^{\frac{d+\delta}{2}}C_R C_\Phi^{d+\frac{\delta}{2}} =: \widetilde{C}_R$$

on scales 0 to 1. If $d^{-1/2}C_2^{-1} \ge 1$, let $\widetilde{C}_R := C_R(d\,C_\Phi)^{\delta/2}$. Hence $\widetilde{X}$ is always δ-regular with constant $\widetilde{C}_R$ on scales 0 to 1.

It is also easy to see that $\varphi(X(\tilde{h}^\rho)) \subseteq \widetilde{X}(C(\Phi)\tilde{h}^\rho)$, where $C(\Phi)$ is a constant depending on Φ. For $0 < \hbar < h_0(\Phi)$, we have $C(\Phi)\tilde{h}^\rho < \tilde{h}^{2\rho-1}$; hence

$$\|\mathbb{1}_{\varphi(X(\tilde{h}^\rho))}A(\hbar)\mathbb{1}_{(Y(\tilde{h}^{2\rho-1})-\tilde{h}^{-1}y_k)}\|_{L^2\to L^2} \le \|\mathbb{1}_{\widetilde{X}(\tilde{h}^{2\rho-1})}A(\hbar)\mathbb{1}_{(Y(\tilde{h}^{2\rho-1})-\tilde{h}^{-1}y_k)}\|_{L^2\to L^2}.$$

Next note that

$$\widetilde{X}(\tilde{h}^{2\rho-1}) \subseteq \bigcup_{\substack{j\in\mathbb{Z}\\ \|j\|\le\tilde{h}^{2\rho-2}}} (\widetilde{X}(\tilde{h}) + \tilde{h}j) =: \bigcup_{\substack{j\in\mathbb{Z}\\ \|j\|\le\tilde{h}^{2\rho-2}}} \widetilde{X}_j,$$
$$Y(\tilde{h}^{2\rho-1}) - \tilde{h}^{-1}y \subseteq \bigcup_{\substack{k\in\mathbb{Z}\\ \|p\|\le\tilde{h}^{2\rho-2}}} (Y(\tilde{h}) - \tilde{h}^{-1}y + \tilde{h}p) =: \bigcup_{\substack{k\in\mathbb{Z}\\ \|p\|\le\tilde{h}^{2\rho-2}}} Y_p. \tag{6-38}$$

Hence, it is eventually reduced to estimating each $\|\mathbb{1}_{\widetilde{X}_j} A(\hbar)\mathbb{1}_{Y_p}\|_{L^2\to L^2}$.

It is easy to check that $\hat{b}_k(x, y)$ satisfy (6-24); hence by (6-25), we have

$$\|\mathbb{1}_{\widetilde{X}_j} A(\hbar)\mathbb{1}_{Y_p}\|_{L^2\to L^2} \le C\tilde{h}^{\beta}$$

for some $\beta > 0$. Choosing $2d(\rho-1) < \beta/2$, we conclude that

$$\|\mathbb{1}_{\widetilde{X}(\tilde{h}^{2\rho-1})} A(\hbar)\mathbb{1}_{(Y(\tilde{h}^{2\rho-1})-\tilde{h}^{-1}y_k)}\|_{L^2\to L^2} \le C\tilde{h}^{\frac{\beta}{2}}$$

by the triangle inequality. □

Appendix A: Regular sets

We show that certain operations preserve the class of δ-regular sets if we allow one to increase the regularity constant and shrink the scales.

The first lemma is from [Bourgain and Dyatlov 2018]. It shows a δ-regular set in $\mathbb{R}^1$, $0 < \delta < 1$, can be split into smaller δ-regular sets.

Lemma A.1. *Let $X \subset \mathbb{R}^1$ be a δ-regular set with constant C_R on scales α_0 to α_1, and assume that $0 < \delta < 1$ and $(4C_R)^{2/(1-\delta)}\alpha_0 \le \rho \le \alpha_1$. Then there exists a collection of disjoint intervals $\mathcal{J}$ such that*

$$X = \bigcup_{J\in\mathcal{J}} (X \cap J), \qquad (4C_R)^{-\frac{2}{1-\delta}}\rho \le |J| \le \rho \quad \textit{for all } J \in \mathcal{J},$$

and each $X \cap J$ is δ-regular with constant $\widetilde{C}_R := (4C_R)^{2/(1-\delta)}C_R$ on scales α_0 to ρ.

The rest of this section concerns δ-regular sets in $\mathbb{R}^d$. We show that certain operations preserve the class of δ-regular sets if we allow one to increase the regularity constant and shrink the scales.

Lemma A.2. *Let X be a δ-regular set with $\delta \in (0, d)$ and constant C_R on scales α_0 to α_1. Fix $\lambda > 0$ and $y \in \mathbb{R}^d$. Then the set $\widetilde{X} := y + \lambda X$ is a δ-regular set with constant C_R on scales $\lambda\alpha_0$ to $\lambda\alpha_1$.*

Proof. Taking the measure

$$\mu_{\widetilde{X}}(A) := \lambda^{\delta}\mu_X(\lambda^{-1}(A - y)),$$

it is easy to verify. □

Lemma A.3. *Let X be a δ-regular set with constant C_R on scales α_0 to α_1. Fix $T > 1$. Then X is δ-regular with constant $\widetilde{C}_R := 2T^dC_R$ on scales α_0 to $T\alpha_1$.*

Proof. Let I be a cube such that $\alpha_0 \le r_I \le T\alpha_1$. For $\alpha_0 \le r_I \le \alpha_1$, the upper bound is immediate. For $\alpha_1 < r_I \le T\alpha_1$, I can be covered by $\lceil T\rceil^d \le 2T^d$ cubes of side length α_1 each; therefore

$$\mu_X(I) \le 2T^d C_R \alpha_1^\delta \le \widetilde{C}_R r_I^\delta.$$

In view of the lower bound estimate, we assume I is centered at a point in X. As before, we may assume $\alpha_1 < r_I \le T\alpha_1$. Let $I' \subset I$ be the cube with the same center and $r_{I'} = \alpha_1$. Then

$$\mu_X(I) \ge \mu_X(I') \ge C_R^{-1}\alpha_1^\delta \ge \widetilde{C}_R^{-1} r_I^\delta,$$

as claimed. □

Lemma A.4. *Let X be a δ-regular set with constant C_R on scales α_0 to α_1. Fix $T \ge 1$:*

(1) *Suppose $\alpha_1 \ge 2\alpha_0$. Then the neighborhood $X + [-T\alpha_0, T\alpha_0]^d$ is δ-regular with constant $\widetilde{C}_R := 4^d T^d C_R$ on scales $2\alpha_0$ to α_1.*

(2) *Suppose that $\alpha_1 \ge T\alpha_0$. Then $X + [-T\alpha_0, T\alpha_0]^d$ is δ-regular with constant $C_R' = 4^d C_R$ on scales $T\alpha_0$ to α_1.*

Proof. Let $\widetilde{X} := X + [-T\alpha_0, T\alpha_0]^d$ and define $\mu_{\widetilde{X}}$ supported on $\widetilde{X}$ by convolution

$$\mu_{\widetilde{X}}(A) := \frac{1}{(T\alpha_0)^d}\int_{[-T\alpha_0, T\alpha_0]^d} \mu_X(A+y)\,dy.$$

Let I be a cube such that $M\alpha_0 \le r_I \le \alpha_1$ with $M \ge 1$. Then

$$\mu_{\widetilde{X}}(I) \le 2^d C_R r_I^\delta,$$

which proves the upper bound estimates for both cases.

Now assume that I is centered at a point $x_1 \in \widetilde{X}$. Take $x_0 \in X$ such that $x_0 \in x_1 + [-T\alpha_0, T\alpha_0]^d$, and let I' be the cube centered at x_0 with side length $r_{I'} = r_I/2$. Then

$$\mu_X(I') \ge C_R^{-1}\left(\frac{r_I}{2}\right)^\delta \ge 2^{-d} C_R^{-1} r_I^\delta.$$

Let $J = x_0 - x_1 + [-\alpha_0/2, \alpha_0/2]^d$; then $J \cap [-T\alpha_0, T\alpha_0]^d$ contains a cube with side length at least $\alpha_0/2$. Clearly, $I' \subset I + y$ for any $y \in J$. Hence

$$\mu_{\widetilde{X}}(I) \ge \frac{1}{(2T)^d}\mu_X(I') \ge \widetilde{C}_R^{-1} r_I^\delta,$$

which proves the lower bound estimate for (1).

Let $J = x_0 - x_1 + [-T\alpha_0/2, T\alpha_0/2]^d$; then $J \cap [-T\alpha_0, T\alpha_0]^d$ contains a cube with side length at least $T\alpha_0/2$. Clearly, $I' \subset I + y$ for any $y \in J$. Hence

$$\mu_{\widetilde{X}}(I) \ge \frac{1}{2^d}\mu_X(I') \ge (C_R')^{-1} r_I^\delta,$$

which proves the lower bound estimate for (2). □

Lemma A.5. *Assume $F:\mathbb{R}^d\to\mathbb{R}^d$ is C^1 diffeomorphism. Let $C_1:=\sup_{x\in\mathbb{R}^d}\|\mathrm{J}F(x)\|$ and $C_2:=\sup_{x\in\mathbb{R}^d}\|\mathrm{J}F^{-1}(x)\|$, where $\mathrm{J}F$ is the Jacobian matrix and $\|\cdot\|$ is the matrix norm. Assume that for some constant $C_F\geq 1$, we have*

$$C_1C_2\leq C_F. \tag{A-1}$$

Let X be a δ-regular set with constant C_R on scales α_0 to $\alpha_1\geq C_F^2\alpha_0$. Then $F(X)$ is a δ-regular set with constant $\widetilde{C}_R:=C_R(d\,C_F)^{\delta/2}$ on scales $d^{1/2}C_1\alpha_0$ to $d^{-1/2}C_2^{-1}\alpha_1$.

Proof. Let $\widetilde{X}:=F(X)$ and define the measure $\mu_{\widetilde{X}}$ supported on $\widetilde{X}$ as

$$\mu_{\widetilde{X}}(A):=C_F^{-\frac{\delta}{2}}C_1^{\delta}\,\mu_X(F^{-1}(A)).$$

Let $\tilde{I}$ be a cube with side length $r_{\tilde{I}}$ with

$$d^{\frac{1}{2}}C_1\alpha_0\leq r_{\tilde{I}}\leq d^{-\frac{1}{2}}C_2^{-1}\alpha_1. \tag{A-2}$$

Clearly, $F^{-1}\tilde{I}$ is contained in a cube of side length r, where $r\leq\sqrt{d}\,C_2r_{\tilde{I}}$. Indeed, let y be the center of $\tilde{I}$. Then for any $x\in\tilde{I}$, we have

$$\|F^{-1}(x)-F^{-1}(y)\|\leq C_2\|x-y\|\leq\frac{\sqrt{d}}{2}C_2r_{\tilde{I}}.$$

Let I be the cube centered at $F^{-1}y$ of side length $\sqrt{d}\,C_2r_{\tilde{I}}\leq\alpha_1$. Then

$$\mu_{\widetilde{X}}(\tilde{I})\leq\mu_X(I)\leq C_F^{-\frac{\delta}{2}}C_1^{\delta}C_R(\sqrt{d}\,C_2r_{\tilde{I}})^{\delta}=C_R(d\,C_F)^{\frac{\delta}{2}}r_{\tilde{I}}^{\frac{\delta}{2}}.$$

If, in addition, $y\in\widetilde{X}$, let $y=F(z)$, where $z\in X$. Then the cube Q centered at z of side length $r=d^{-1/2}C_1^{-1}r_{\tilde{I}}\geq\alpha_0$ is contained in $F^{-1}(\tilde{I})$. Indeed, for any $x\in Q$, we have

$$\|F(x)-F(z)\|\leq\frac{\sqrt{d}}{2}C_1r=\frac{r_{\tilde{I}}}{2}.$$

Hence

$$\mu_{\widetilde{X}}(\tilde{I})=C_F^{-\frac{\delta}{2}}C_1^{\delta}\mu_X(F^{-1}(\tilde{I}))\geq C_F^{-\frac{\delta}{2}}C_1^{\delta}C_R^{-1}(d^{-\frac{1}{2}}C_1^{-1}r_{\tilde{I}})^{\delta}=C_R^{-1}(d\,C_F)^{-\frac{\delta}{2}}r_{\tilde{I}}^{\frac{\delta}{2}}.$$

This proves the claim. □

Lemma A.6. *Let X be a δ-regular set with constant C_R on scales α_0 to α_1, and $0<\delta<d$. Fix an integer*

$$L\geq(2^{\frac{d}{2}}\sqrt{2d+1}C_R)^{\frac{2}{d-\delta}}. \tag{A-3}$$

Assume that I is a cube with $\alpha_0\leq r_I/L\leq r_I\leq\alpha_1$ and $I_1,\dots,I_{L^d}$ is the partition of I into cubes of side length r_I/L. Then there exists ℓ such that $X\cap I_\ell=\varnothing$.

Proof. Using Lemma A.2, it suffices to consider $I=[0,L]^d$, $\alpha_0\leq 1\leq L\leq\alpha_1$. We argue by contradiction. Assume that each I_ℓ intersects X. Then $I'_\ell:=I_\ell+\left[-\frac{1}{2},\frac{1}{2}\right]^d$ contains a unit cube centered at a point in X and thus

$$\mu_X(I'_\ell)\geq C_R^{-1}\quad\text{for all } 1\leq\ell\leq L^d.$$

On the other hand,

$$\bigcup_{\ell=1}^{L^d} I'_\ell = \left[-\tfrac{1}{2}, L+\tfrac{1}{2}\right]^d,$$

and each point in $\left[-\frac{1}{2}, L+\frac{1}{2}\right]^d$ can be covered by at most $2d+1$ of the cubes I'_ℓ. Therefore

$$C_R^{-1}L^d \le \sum_{\ell=1}^{L^d} \mu_X(I'_\ell) \le (2d+1)\mu_X\left(\left[-\tfrac{1}{2}, L+\tfrac{1}{2}\right]^d\right) \le (2d+1)C_R(L+1)^\delta,$$

which contradicts (A-3). □

Recall our definition of $\mathcal{C}_n$ and porosity in Definition 5.1.

Lemma A.7. *Let $X \subset [-1,1]^d$ be a δ-regular set with constant C_R on scales α_0 to α_1. Let L satisfy* (A-3), *and take $n \in \mathbb{Z}$ such that $\alpha_0 \le L^{-n-1} \le L^{-n} \le \alpha_1$. Then X is porous at scale L with depth n.*

Lemma A.8. *Let X be a δ-regular set with constant C_R on scales α_0 to α_1. Let $C \ge 1$ be a constant. Let I be a cube of side length r_I satisfying $\alpha_0 \le r_I \le C\alpha_1$. Let $\rho > 0$ satisfy $\alpha_0 \le \rho \le \min(r_I, \alpha_1)$. Then there exists a nonoverlapping*[1] *collection $\mathcal{J}$ of $N_{\mathcal{J}}$ cubes of side length ρ each such that*

$$X \cap I \subset \bigcup_{J\in\mathcal{J}} J, \quad N_{\mathcal{J}} \le \left(6\left\lceil \frac{3+C}{2}\right\rceil\right)^d C_R^2 \left(\frac{r_I}{\rho}\right)^\delta.$$

We will only use this lemma in dimension 1. Note that in [Bourgain and Dyatlov 2018], this is formulated with $C=1$. We use this form with a constant C in the proof of Lemma 6.2.

Proof. Let $\mathcal{J}$ consist of all cubes of the form $\times_{k=1}^d \rho[j_k, j_k+1]$, $(j_1, j_2, \dots, j_d) \in \mathbb{Z}^d$, which intersect $X \cap I$. Then $X \cap I \subset \bigcup_{J\in\mathcal{J}} J$. Next, we will prove the upper bound on $N_{\mathcal{J}}$.

For each $J \in \mathcal{J}$, let $J' \supset J$ be the cube with the same center and with side length 2ρ. Since J intersects X, J' contains a cube of side length ρ centered at a point in X. Therefore

$$\mu_X(J') \ge C_R^{-1}\rho^\delta.$$

It is also clear that $\bigcup_{J\in\mathcal{J}} J' \subset I\left(\frac{3}{2}\rho\right)$, and each point lies in at most 3^d of the cubes J'.

If $r_I \le \alpha_1$, $I\left(\frac{3}{2}\rho\right)$ can be covered by 4^d cubes of side length r_I. If $\alpha_1 < r_I \le C\alpha_1$, $I\left(\frac{3}{2}\rho\right)$ can be covered by $2^d\lceil(3+C)/2\rceil^d$ cubes of side length α_1. Therefore, we always have

$$N_{\mathcal{J}} \cdot C_R^{-1}\rho^\delta \le \sum_{J\in\mathcal{J}} \mu_X(J') \le 3^d \mu_X\left(\bigcup_{J\in\mathcal{J}} J'\right) \le \left(6\left\lceil\frac{3+C}{2}\right\rceil\right)^d C_R r_I^\delta,$$

and this proves the upper bound on $N_{\mathcal{J}}$. □

Appendix B: Proof of Lemma 6.2

We follow the proofs of Theorem 3.2 and Lemma 4.1 in [Jin and Zhang 2017]. Let us start with introducing some notation.

[1] A collection of cubes is nonoverlapping if the intersection of every two distinct cubes has empty interior.

Hilbert transform. Let $\mathcal{H}_0$ be the standard Hilbert transform defined as convolution with p.v. $\frac{1}{\pi x}$: for $f \in C_0^\infty(\mathbb{R})$ (or more generally, $f \in L^1(\mathbb{R}, \langle x\rangle^{-1}\,dx)$

$$\mathcal{H}_0(f)(x) := \frac{1}{\pi}\lim_{\varepsilon\to 0+}\int_{|x-t|\geq\varepsilon}\frac{f(t)}{x-t}\,dt.$$

Let $\mathcal{H}$ be the modified Hilbert transform with integral kernel that decays like $|x|^{-2}$ as $|x|\to\infty$:

$$\mathcal{H}(f)(x) := \frac{1}{\pi}\lim_{\varepsilon\to 0+}\int_{|x-t|\geq\varepsilon} f(t)\Big(\frac{1}{x-t}+\frac{t}{t^2+1}\Big)\,dt,\quad f \in L^1(\mathbb{R}, \langle x\rangle^{-2}\,dx).$$

The advantage of $\mathcal{H}$ is that it applies to a larger space that contains $L^\infty(\mathbb{R})$ as well as functions that grow like $|x|^{1-\varepsilon}$ as $|x|\to\infty$.

If $f \in L^1(\mathbb{R}, \langle x\rangle^{-1}\,dx)$, then $\mathcal{H}(f)$ differs from $\mathcal{H}_0(f)$ by a constant. Moreover, we have the inversion formula for all $f \in L^1(\mathbb{R}, \langle x\rangle^{-2}\,dx)$ with $\mathcal{H}(f) \in L^1(\mathbb{R}, \langle x\rangle^{-2}\,dx)$:

$$\mathcal{H}(\mathcal{H}(f)) = -f + c(f), \tag{B-1}$$

where $c(f)$ is a real constant depending on f.

We will use the following example later in the proof.

Example B.1 [Jin and Zhang 2017, Example 2.3]. Let $f(x) = \log(x^2+1)$, then we can compute

$$\mathcal{H}(f)'(x) = \mathcal{H}_0(f')(x) = -\frac{2}{x^2+1}. \tag{B-2}$$

Hardy space and outer functions. We recall the definition of Hardy space on the real line

$$H^2 = H^2(\mathbb{R}) = \{f \in L^2(\mathbb{R}) : \operatorname{supp}\hat{f} \subset [0,\infty)\}.$$

If $f \in L^2(\mathbb{R})$, then $f + i\mathcal{H}_0(f) \in H^2(\mathbb{R})$.

The space of modulus of functions in H^2 can be characterized by the logarithmic integral: for $\omega \in L^2$, $\omega \geq 0$, we define

$$\mathcal{L}(\omega) := \int_{\mathbb{R}}\frac{\log\omega(x)}{1+x^2}\,dx.$$

Theorem B.2 [Havin and Jöricke 1994, Section 1.5]. *If $f \in H^2$ and $\mathcal{L}(|f|) = -\infty$, then $f \equiv 0$. On the other hand, if $\omega \in L^2$ and $\mathcal{L}(\omega) > -\infty$, then there exists a function $f \in H^2$ with $|f| = \omega$, unique up to a multiplication by a complex constant with unit modulus.*

If $\mathcal{L}(\omega) > -\infty$. Let $\Omega = -\log\omega$, then $\Omega \in L^1(\mathbb{R}, \langle x\rangle^{-2}\,dx)$. Therefore we can define $\widetilde{\Omega} = \mathcal{H}(\Omega)$ and take

$$f = ae^{-(\Omega + i\widetilde{\Omega})},\quad |a| = 1. \tag{B-3}$$

We call functions of the form (B-3) for general $\Omega \in L^1(\mathbb{R}, \langle x\rangle^{-2}\,dx)$ *outer functions*. The class of outer functions is closed under multiplications. Moreover if two outer functions have the same modulus, then they differ by a complex constant with unit modulus.

The following lemma gives a sufficient condition of a function to be the modulus of the Fourier transform of a function supported in $[0,\sigma]$.

Lemma B.3 [Mashregi et al. 2005, Theorem 1]. *Assume that* $\omega = e^{-\Omega} \in L^2$ *and* $\mathcal{L}(\omega) > -\infty$. *In addition, we assume that* $\omega^2 e^{2\pi i\sigma x}$ *is an outer function. Then there exists* $\psi \in L^2$ *with* $\operatorname{supp}\psi \subset [0,\sigma]$ *and* $|\widehat{\psi}| = \omega$.

An effective multiplier theorem. We prove an effective multiplier theorem. This proof is essentially in [Jin and Zhang 2017, Section 3], the only change we make lies in the definition of $k(x)$ below. Our modified definition makes sure that $k(x)$ is a constant function in a neighborhood of 0, which leads to a pointwise lower bound of $\widehat{\psi}(x)$ on the whole interval $\left[-\frac{3}{4},\frac{3}{4}\right]$.

Theorem B.4. *Assume that* $0<\omega\le 1$ *satisfies* $\mathcal{L}(\omega) > -\infty$, *and*

$$\|\mathcal{H}(\Omega)'\|_{L^\infty} \le \tfrac{\pi}{2}\sigma,$$

where $0<\sigma<\frac{1}{10}$, $\Omega = -\log\omega$. *Then there exists* $\psi \in L^2(\mathbb{R})$ *with*

$$\operatorname{supp}\psi \subset [0,\sigma], \quad |\widehat{\psi}| \le \omega,$$

and

$$|\widehat{\psi}| \ge \frac{\sigma^{10}}{4\times 10^{11}}\omega \quad \text{on } \left[-\tfrac{3}{4},\tfrac{3}{4}\right].$$

Proof. We first set

$$\omega_0(x) = \frac{\omega(x)}{(x^2+T^2)^5}, \quad \Omega_0(x) = -\log(\omega_0(x)),$$

with constant T that will be specified later. We then have

$$\Omega_0 = \Omega + 5\log(x^2+T^2).$$

We compute

$$\mathcal{H}(\log(x^2+T^2))(0) = \lim_{\varepsilon\to 0^+}\int_{|t|\ge\varepsilon} \log(t^2+T^2)\left(\frac{1}{-t} - \frac{t}{t^2+1}\right)dt,$$

in which the integrand is an odd function. Hence the integration is zero. Therefore we have

$$\mathcal{H}(\Omega_0)(0) = \mathcal{H}(\Omega)(0) + 5\mathcal{H}(\log(x^2+T^2)) = \mathcal{H}(\Omega)(0). \tag{B-4}$$

By (B-2), we compute

$$\mathcal{H}(\log(x^2+T^2))' = T^{-1}\mathcal{H}(\log(x^2+1))'\left(\frac{\cdot}{T}\right) = -\frac{2T}{x^2+T^2}.$$

Thus if we choose $T = \frac{20}{\pi\sigma} \ge \frac{200}{\pi} \ge 60$, we have

$$\|\mathcal{H}(\Omega_0)'\|_{L^\infty} \le \|\mathcal{H}(\Omega)'\|_{L^\infty} + 5\|\mathcal{H}(\log(x^2+T^2))'\|_{L^\infty} \le \pi\sigma. \tag{B-5}$$

Let us define

$$s_0(x) = \pi\sigma x + \mathcal{H}(\Omega_0)(x).$$

Hence by (B-4),

$$s_0(0) = \mathcal{H}(\Omega)(0),$$

depending only on ω.

Let $s(x)$ be defined as

$$s(x) = s_0(x) - \pi k(x) - \tfrac{\pi}{2},$$

in which

$$k(x) = \begin{cases} \left\lfloor \tfrac{1}{\pi} s_0(x) \right\rfloor & \text{if } \tfrac{1}{\pi} s_0(0) \in \left[\tfrac{1}{4}, \tfrac{3}{4}\right] \bmod 1, \\ \left\lfloor \tfrac{1}{\pi} s_0(x) - \tfrac{1}{2} \right\rfloor & \text{if } \tfrac{1}{\pi} s_0(0) \in \left[0, \tfrac{1}{4}\right) \bigcup \left(\tfrac{3}{4}, 1\right) \bmod 1. \end{cases} \tag{B-6}$$

Note that our definition of $k(x)$ is different from that in [Jin and Zhang 2017]. We modify the definition in order to make sure $k(x)$ is a constant near $x = 0$. This will be explained and used later in the proof.

By (B-5), $s_0(x)$ is a nondecreasing function and so is k. Note also that by our definition of $s(x)$, we have

$$\|s\|_{L^\infty} \le \pi. \tag{B-7}$$

Let $m = e^{-M}$, where $M = \mathcal{H}(s)$. Next, we will estimate $M(x) = \mathcal{H}(s)(x)$. We split the integral into three parts $M(x) = J_1(x) + J_2(x) + J_3(x)$, where

$$J_1(x) = \frac{1}{\pi} \int_{|x-t|<\frac{1}{2}} \frac{s(t) - s(x)}{x - t}\, dt,$$

$$J_2(x) = \frac{1}{\pi} \int_{|x-t|<\frac{1}{2}} s(t) \frac{t}{t^2 + 1}\, dt,$$

$$J_3(x) = \frac{1}{\pi} \int_{|x-t|\ge\frac{1}{2}} s(t) \left(\frac{1}{x - t} + \frac{t}{t^2 + 1} \right) dt.$$

We estimate J_2 and J_3 in the same way as in [Jin and Zhang 2017]. By (B-7), we have

$$|J_2(x)| \le \tfrac{1}{\pi} \cdot \|s\|_{L^\infty} \cdot \tfrac{1}{2} \le \tfrac{1}{2}. \tag{B-8}$$

Also, we have

$$|J_3(x)| \le \tfrac{1}{\pi} \cdot \|s\|_{L^\infty} \int_{|x-t|\ge\frac{1}{2}} \left| \frac{1}{x - t} + \frac{t}{t^2 + 1} \right| dt \le 6 \log(|x| + 2). \tag{B-9}$$

Finally, we need to bound $|J_1|$. By (B-5), we know $s_0(x) = \pi\sigma x + \mathcal{H}(\Omega_0)(x)$ is nondecreasing with $\|s_0'\|_{L^\infty} \le 2\pi\sigma$. Since we assume $0 < \sigma < \frac{1}{10}$, we have

$$\|\pi^{-1} s_0'\|_{L^\infty} < \tfrac{1}{5}.$$

This leads to the following:

- If $\pi^{-1} s_0(0) \in \left[\tfrac{1}{4}, \tfrac{3}{4}\right] \bmod 1$,

$$\tfrac{1}{\pi} s_0(x) \in (0, 1) \bmod 1 \quad \text{for all } x \in \left[-\tfrac{5}{4}, \tfrac{5}{4}\right].$$

- If $\pi^{-1} s_0(0) \in \left[0, \tfrac{1}{4}\right) \bigcup \left(\tfrac{3}{4}, 1\right) \bmod 1$,

$$\tfrac{1}{\pi} s_0(x) - \tfrac{1}{2} \in (0, 1) \bmod 1 \quad \text{for all } x \in \left[-\tfrac{5}{4}, \tfrac{5}{4}\right].$$

Recalling our definition of $k(x)$ in (B-6), we know in each case $k(x)$ is a constant function on the interval $\left[-\frac{5}{4}, \frac{5}{4}\right]$.

Thus for $x \in \left[-\frac{3}{4}, \frac{3}{4}\right]$, we have

$$|J_1(x)| \le \frac{1}{\pi} \int_{|x-t|<\frac{1}{2}} \left| \frac{s_0(t) - s_0(x)}{x-t} \right| dt \le \frac{1}{\pi} \|s_0'\|_{L^\infty} \le 2\sigma. \tag{B-10}$$

For all x, we only have a lower bound of J_1. Since k is nondecreasing, we have

$$J_1(x) \ge \frac{1}{\pi} \int_{|x-t|<\frac{1}{2}} \frac{s_0(t) - s_0(x)}{x-t}\, dt \ge -2\sigma. \tag{B-11}$$

Now combining (B-8), (B-9) with (B-10), we have the following estimate of M on $\left[-\frac{3}{4}, \frac{3}{4}\right]$:

$$|M(x)| \le 2\sigma + \tfrac{1}{2} + 6 \log \tfrac{11}{4} < 7. \tag{B-12}$$

Using (B-11) instead of (B-10), we obtain that for all x,

$$M(x) \ge -2\sigma - \tfrac{1}{2} - 6 \log\left(|x| + 2\right) > -1 - 6 \log(|x| + 2). \tag{B-13}$$

Next we will apply Lemma B.3 to $\tilde{\omega} = \frac{1}{3} m \omega_0$. We check that $\tilde{\omega}$ satisfies all the assumptions. First, by (B-13), we have

$$0 \le \tilde{\omega} \le \tfrac{1}{3} e(|x| + 2)^6 \omega_0 \le \frac{\omega}{x^2 + T^2}.$$

Hence $0 \le \tilde{\omega} \le \omega$ and $\tilde{\omega} \in L^2$. Moreover

$$\mathcal{L}(\tilde{\omega}) = \mathcal{L}\left(\tfrac{1}{3} m\right) + \mathcal{L}(\omega_0) > -\infty.$$

By the construction $M = \mathcal{H}(s)$ and the inversion formula (B-1), we have

$$\mathcal{H}(-2M - 2\Omega_0) = 2s - 2\mathcal{H}(\Omega_0) - 2c(M) = 2\pi\sigma x - 2\pi k(x) - \pi - 2c(M),$$

where $k(x) \in \mathbb{Z}$ and $c(M)$ is a real constant. Therefore for some constant a with $|a| = 1$, we have

$$\tilde{\omega}^2 e^{2\pi i \sigma x} = \tfrac{1}{9} e^{-2M - 2\Omega_0 + 2\pi i \sigma x} = \tfrac{1}{9} a e^{-2M - 2\Omega_0 + i\mathcal{H}(-2M - 2\Omega_0)},$$

which shows $\tilde{\omega}^2 e^{2\pi i \sigma x}$ is an outer function.

By Lemma B.3, there exists $\psi \in L^2$ with $\operatorname{supp}(\psi) \subset [0, \sigma]$ and $|\widehat{\psi}| \le \tilde{\omega} \le \omega$. Furthermore, on $\left[-\frac{3}{4}, \frac{3}{4}\right]$, by (B-12), and since $T = \frac{20}{\pi\sigma}$, we have

$$|\widehat{\psi}(x)| = \tilde{\omega}(x) \ge \tfrac{1}{3}(1 + T^2)^{-5} e^{-7} \omega(x) \ge \frac{\sigma^{10}}{4 \times 10^{11}} \omega(x),$$

as claimed. □

Multiplier adapted to the regular sets. Now we are in the place to finish the proof of Lemma 6.2.

Proof. The proof is the essentially same as that of Lemma 4.1 of [Jin and Zhang 2017]. We briefly go through the various constants below.

We define $n_1 \in \mathbb{N}$ by $2^{n_1} < S\alpha_1 \le 2^{n_1+1}$. For $1 \le n \le n_1$, let $A_n := [-2^{n+1}, -2^n] \bigcup [2^n, 2^{n+1}]$. Then by Lemma A.8, we have a collection $\mathcal{J}_n$ of N_n intervals of size $\rho_n := n^{-(1+\delta)/2} 2^n$ such that each element is of the form $[j, j+1]$, $j \in \mathbb{Z}$, intersects A_n, and

$$Y \cap A_n \subset \bigcup_{J \in \mathcal{J}_n} J.$$

Moreover, the number N_n satisfies

$$N_n \le 6 \left\lceil \frac{3+S}{2} \right\rceil C_R^2 \left(\frac{2^n}{\rho_n} \right)^{\delta} = 6 \left\lceil \frac{3+S}{2} \right\rceil C_R^2 n^{\frac{\delta(1+\delta)}{2}}. \tag{B-14}$$

Following the proof of [Jin and Zhang 2017], we define a weight function ω such that

$$\begin{aligned} &\omega(\xi) = \exp(-\langle \xi \rangle^{\frac{1}{2}}) \ge 0.3 && \text{for all } \xi \in [-1, 1], \\ &\omega(\xi) \le \exp(-\langle \xi \rangle^{\frac{1}{2}}) && \text{for all } \xi \in \mathbb{R}, \\ &\omega(\xi) \le \exp(-\Theta(|\xi|)|\xi|) && \text{for all } \xi \in Y,\ |\xi| \ge 10, \end{aligned}$$
$$\|\mathcal{H}(\omega)'\|_{L^\infty} \le \frac{\iota^{-1} C_R^2}{\delta_1(1-\delta_1)},$$

where $0 < \iota < 1$ is a constant depending only on S. The dependence comes from the upper bound of N_n in (B-14).

Applying Theorem B.4 to ω^{c_3} with

$$\sigma = \tfrac{1}{5} c_1, \quad c_3 = \tfrac{\pi}{10} \iota c_1 C_R^{-2} \delta_1 (1-\delta_1) < 1.$$

We obtain ψ with

$$\operatorname{supp} \psi \subset \left[0, \tfrac{1}{5} c_1\right],$$
$$\begin{aligned} &|\widehat{\psi}(\xi)| \ge \frac{c_1^{10}}{4 \times 10^{18}} \omega(\xi)^{c_3} \ge \frac{3}{4 \times 10^{19}} c_1^{10} && \text{for all } \xi \in \left[-\tfrac{3}{4}, \tfrac{3}{4}\right], \\ &|\widehat{\psi}(\xi)| \le \exp(-c_3 \langle \xi \rangle^{\frac{1}{2}}) && \text{for all } \xi \in \mathbb{R}, \\ &|\widehat{\psi}(\xi)| \le \exp(-c_3 \Theta(|\xi|)|\xi|) && \text{for all } \xi \in Y,\ |\xi| \ge 10. \end{aligned}$$

Finally, shifting ψ by $\frac{1}{10} c_1$ yields the desired function. □

Acknowledgements

The authors thank Jean Bourgain for his interest in this paper, and thank Semyon Dyatlov for his comments on a preliminary version which improved our presentation. We thank the IAS, Princeton, for its hospitality during the 2017–18 academic year. This material is based upon work supported by the National Science Foundation under grant DMS-1638352. Han was partially supported by NSF grant DMS-1800689 during the preparation of this work. Schlag was partially supported by NSF grant DMS-1500696 during the preparation of this work.

References

[Abramowitz and Stegun 1966] M. Abramowitz and I. A. Stegun (editors), *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, National Bureau of Standards Appl. Math. Series **55**, Dover, New York, 1966.

[Beurling and Malliavin 1962] A. Beurling and P. Malliavin, "On Fourier transforms of measures with compact support", *Acta Math.* **107** (1962), 291–309. MR Zbl

[Bourgain and Dyatlov 2017] J. Bourgain and S. Dyatlov, "Fourier dimension and spectral gaps for hyperbolic surfaces", *Geom. Funct. Anal.* **27**:4 (2017), 744–771. MR Zbl

[Bourgain and Dyatlov 2018] J. Bourgain and S. Dyatlov, "Spectral gaps without the pressure condition", *Ann. of Math.* (2) **187**:3 (2018), 825–867. MR Zbl

[Dyatlov 2017] S. Dyatlov, "Control of eigenfunctions on hyperbolic surfaces: an application of fractal uncertainty principle", art. id. 4 in *Journées équations aux dérivées partielles* (Roscoff, France, 2017), CEDRAM, Grenoble, 2017.

[Dyatlov and Jin 2017] S. Dyatlov and L. Jin, "Resonances for open quantum maps and a fractal uncertainty principle", *Comm. Math. Phys.* **354**:1 (2017), 269–316. MR Zbl

[Dyatlov and Jin 2018] S. Dyatlov and L. Jin, "Semiclassical measures on hyperbolic surfaces have full support", *Acta Math.* **220**:2 (2018), 297–339. MR Zbl

[Dyatlov and Zahl 2016] S. Dyatlov and J. Zahl, "Spectral gaps, additive energy, and a fractal uncertainty principle", *Geom. Funct. Anal.* **26**:4 (2016), 1011–1094. MR Zbl

[Goldstein and Schlag 2001] M. Goldstein and W. Schlag, "Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions", *Ann. of Math.* (2) **154**:1 (2001), 155–203. MR Zbl

[Goldstein and Schlag 2008] M. Goldstein and W. Schlag, "Fine properties of the integrated density of states and a quantitative separation property of the Dirichlet eigenvalues", *Geom. Funct. Anal.* **18**:3 (2008), 755–869. MR Zbl

[Havin and Jöricke 1994] V. Havin and B. Jöricke, *The uncertainty principle in harmonic analysis*, Ergebnisse der Mathematik (3) **28**, Springer, 1994. MR Zbl

[Jin and Zhang 2017] L. Jin and R. Zhang, "Fractal uncertainty principle with explicit exponent", preprint, 2017. arXiv

[Kovrijkine 2001] O. Kovrijkine, "Some results related to the Logvinenko–Sereda theorem", *Proc. Amer. Math. Soc.* **129**:10 (2001), 3037–3047. MR Zbl

[Levin 1996] B. Y. Levin, *Lectures on entire functions*, Translations of Math. Monographs **150**, Amer. Math. Soc., Providence, RI, 1996. MR Zbl

[Mashregi et al. 2005] D. Mashregi, F. L. Nazarov, and V. P. Khavin, "The Beurling–Malliavin multiplier theorem: the seventh proof", *Algebra i Analiz* **17**:5 (2005), 3–68. In Russian; translated in *St. Petersburg Math. J.* **17**:5 (2006), 699–744. MR Zbl

[Muscalu and Schlag 2013] C. Muscalu and W. Schlag, *Classical and multilinear harmonic analysis, I*, Cambridge Studies in Adv. Math. **137**, Cambridge Univ. Press, 2013. MR Zbl

Received 30 Aug 2018. Revised 8 Feb 2019. Accepted 20 Mar 2019.

RUI HAN: rui.han@math.gatech.edu
School of Mathematics, Georgia Institute of Technology, Atlanta, GA, United States

WILHELM SCHLAG: wilhelm.schlag@yale.edu
Department of Mathematics, Yale University, New Haven, CT, United States

ANALYSIS AND PDE
Vol. 13, No. 3, 2020
dx.doi.org/10.2140/apde.2020.13.865

msp

LOCAL MINIMALITY RESULTS FOR THE MUMFORD–SHAH FUNCTIONAL VIA MONOTONICITY

DORIN BUCUR, ILARIA FRAGALÀ AND ALESSANDRO GIACOMINI

Let $\Omega \subseteq \mathbb{R}^2$ be a bounded piecewise $C^{1,1}$ open set with convex corners, and let

$$\mathrm{MS}(u) := \int_\Omega |\nabla u|^2\,dx + \alpha \mathcal{H}^1(J_u) + \beta \int_\Omega |u-g|^2\,dx$$

be the Mumford–Shah functional on the space $\mathrm{SBV}(\Omega)$, where $g \in L^\infty(\Omega)$ and $\alpha, \beta > 0$. We prove that the function $u \in H^1(\Omega)$ such that

$$\begin{cases} -\Delta u + \beta u = \beta g & \text{in } \Omega, \\ \partial u/\partial \nu = 0 & \text{on } \partial\Omega \end{cases}$$

is a local minimizer of MS with respect to the L^1-topology. This is obtained as an application of interior and boundary monotonicity formulas for a weak notion of quasiminimizers of the Mumford–Shah energy. The local minimality result is then extended to more general free discontinuity problems taking into account also boundary conditions.

1. Introduction

The Mumford–Shah functional was introduced in [Mumford and Shah 1985; 1989] in the context of image segmentation and has found important applications in several other fields, including variational theories in fracture mechanics; see [Francfort and Marigo 1998; Bourdin et al. 2008]. It can be considered the typical example of a *free discontinuity functional*, characterized by the coupling of bulk and surface energies.

The *weak* formulation of the functional due to De Giorgi and Ambrosio [1988] takes the form

$$\mathrm{MS}(u) := \int_\Omega |\nabla u|^2\,dx + \alpha \mathcal{H}^1(J_u) + \beta \int_\Omega |u-g|^2\,dx, \tag{1-1}$$

Bucur is member of the Institut Universitaire de France. Fragalà and Giacomini are members of the Gruppo Nazionale per L'Analisi Matematica, la Probabilità e loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).
MSC2010: 35R35, 35A16, 28A75, 35J25, 35Q74, 49J45, 94A08.
Keywords: local minimality, monotonicity formulas, free discontinuity functionals.

where $\Omega \subset \mathbb{R}^2$ is an open bounded set, $\alpha, \beta > 0$ and $g \in L^\infty(\Omega)$. Here $\mathcal{H}^1$ denotes the one-dimensional Hausdorff measure, u belongs to the class of *special functions of bounded variation* $\mathrm{SBV}(\Omega)$ (see Section 2), and J_u is the jump set of u. The last term involving g is usually called the *fidelity term*: in image segmentation, g is the level of gray of the picture, which has to be approximated by choosing conveniently the jump set J_u (the *edges*) and the function u outside it (*regularized image*). When $\beta = 0$, we speak of the homogeneous version of MS.

Within this framework, and in general dimension N, existence of minimizers can be proved easily through the direct method of the calculus of variations, as a direct application of Ambrosio's compactness and lower semicontinuity theorem. Moreover, thanks to the regularity result of De Giorgi, Carriero and Leaci [De Giorgi et al. 1989], minimizers have a topologically closed jump set and are regular outside, yielding an admissible configuration of the original formulation of [Mumford and Shah 1985; 1989], in which the discontinuity set was considered as an independent variable. The regularity of the discontinuity set was then improved by Ambrosio, Fusco and Pallara, who proved that up to an $\mathcal{H}^{N-1}$-negligible set, J_u is a manifold of class $C^{1,\delta}$ for any $\delta < 1$ and of class $C^{1,1}$ if $N = 2$; see also [Bonnet 1996; David 1996].

The issue of detecting minimizers, or more generally *local minimizers* of the Mumford–Shah functional is very delicate. Since MS is only lower semicontinuous (with respect to the natural L^1-topology), necessary conditions for minimality cannot be obtained by "standard" differentiation.

First-order necessary conditions are established for jump sets $\Gamma = J_u$ sufficiently regular by considering *inner variations* (see [Ambrosio et al. 2000, Chapter 7]): they yield that u satisfies an elliptic PDE outside Γ, while its (mean) curvature H_Γ is involved in a transmission condition coupling the values of u and ∇u on both sides of Γ.

For the homogeneous version of MS, a second-order necessary condition for minimality has been proposed by Cagnetti, Mora and Morini [Cagnetti et al. 2008], involving the positive semidefiniteness of a suitable quadratic form defined on $H^1_0(\Gamma)$. Under strict positivity, the authors prove that u is a "local" minimizer among those functions $v \in \mathrm{SBV}(\Omega)$ such that $J_v \subseteq \Phi(\Gamma)$ and $v = u$ on $\partial\Omega$, where Φ is any diffeomorphism of $\mathbb{R}^N$ which is sufficiently C^2-close to the identity and such that $\mathrm{Id} - \Phi$ is compactly supported in Ω. This local minimality was then extended to a full local minimality in the L^1 topology in dimension $N = 2$ by Bonacini and Morini [2015], employing a penalization/regularization technique together with results from the regularity theory of the area and the Mumford–Shah functional.

The aim of this paper is to show that, in dimension $N = 2$ and under mild regularity assumptions on the geometry of Ω, the Mumford–Shah functional (1-1) admits a natural local minimizer with no jumps. More precisely, the main result of the paper is the following.

Theorem 1.1. *Let $\Omega \subseteq \mathbb{R}^2$ be an open, bounded, piecewise $C^{1,1}$-domain with convex corners. Then the function $u \in H^1(\Omega)$ such that*

$$\begin{cases} -\Delta u + \beta u = \beta g & \text{in } \Omega, \\ \partial u / \partial \nu = 0 & \text{on } \partial\Omega \end{cases} \tag{1-2}$$

is a local minimizer of MS *with respect to the L^1-topology.*

In other words, the minimizer of MS within the class of Sobolev functions (which satisfies (1-2)) turns out to be a local minimizer for the natural L^1-topology in the full class of SBV competitors.

A similar result has been obtained by Chambolle, Ponsiglione and the third author [Chambolle et al. 2008] for a generalization of the homogeneous version of the Mumford–Shah functional under boundary conditions, motivated by the study of the issue of *crack initiation* in brittle materials within variational theories of crack propagation; see [Francfort and Marigo 1998; Bourdin et al. 2008]. The technique of [Chambolle et al. 2008] is based essentially on the maximum principle through a truncation argument, and can be used to deal with nonlinear energies for the gradient (with p-growth for example) and more general geometries for Ω, but does not apply to the Mumford–Shah functional since it fails when the fidelity term is present.

Our proof of Theorem 1.1 is based on the use of the monotonicity formula introduced by Luckhaus and the first author [Bucur and Luckhaus 2014] for quasiminimizers of MS, which, in dimension 2, we extend up to boundary points and establish for a quite weak notion of quasiminimizers.

Namely we consider functions $u \in \mathrm{SBV}(\Omega)$ which are *weak local almost-quasiminimizers* of the Mumford–Shah energy *with respect to their own jump sets*, i.e., such that

$$\int_{B_\rho(x)\cap\Omega} |\nabla u|^2\,dx + \mathcal{H}^1(J_u \cap \overline{B}_\rho(x)) \le \int_{B_\rho(x)\cap\Omega} |\nabla v|^2\,dx + \Lambda\mathcal{H}^1(J_v \cap \overline{B}_\rho(x)) + c_\gamma \rho^{1+\gamma} \tag{1-3}$$

for every $v \in \mathrm{SBV}(\Omega)$ with $\{v \neq u\} \subseteq B_\rho(x)$ and $J_v \subseteq J_u$, where $x \in \overline{\Omega}$, $\rho < \rho_0$ are such that

$$\int_{B_\rho(x)\cap\Omega} |\nabla u|^2\,dx + \mathcal{H}^1(J_u \cap \overline{B}_\rho(x)) \le \rho,$$

while $\Lambda \ge 1$ and $\gamma, c_\gamma > 0$. We then prove that if $x \in \Omega$, then the quantity

$$E_x(\rho) := \frac{\mathcal{E}_x(\rho)}{\rho} \wedge 1 + \frac{c_\gamma}{\gamma}\rho^\gamma \tag{1-4}$$

is nondecreasing on $(0, \rho_0 \wedge \mathrm{dist}(x, \partial\Omega))$, where

$$\mathcal{E}_x(\rho) := \int_{\Omega\cap B_\rho(x)} |\nabla u|^2\,dx + \mathcal{H}^1(J_u \cap \overline{B}_\rho(x)),$$

while if $x \in \partial\Omega$, monotonicity holds true for the modification

$$\widetilde{E}_x(\rho) := \frac{\mathcal{E}_x(\rho)}{\rho} \wedge 1 + \frac{c_\gamma}{\gamma}\rho^\gamma + k_\Omega \int_0^\rho \left(\frac{\mathcal{E}_x(r)}{r} \wedge 1\right) dr \tag{1-5}$$

on the interval $(0, \rho_0 \wedge r_\Omega \wedge \mathrm{dist}(x, S \setminus \{x\}))$, where S denotes the set of corners of $\partial\Omega$ and $r_\Omega, k_\Omega > 0$.

The proof of these monotonicities follows that of [Bucur and Luckhaus 2014], which is based on a precise harmonic extension estimate on a ball (see Proposition 3.1). The dimension $N = 2$ gives drastic simplifications in the arguments, and this is the main reason for which we can deal with a weaker class of quasiminimizers (with respect to their own jump set, see also Remark 4.2). The extension of the monotonicity to the boundary requires a generalization of the key harmonic extension estimate to domains of the form $\Omega \cap B_\rho(x)$ (see Proposition 3.7): this motivates the piecewise $C^{1,1}$-regularity assumption for $\partial\Omega$ and the restriction to *convex* angles. We also refer the reader to [Chambolle et al. ≥ 2020] by Chambolle, Séré and Zanini, where a boundary monotonicity formula is proved on flat boundaries, under Neumann conditions.

The use of monotonicity in the proof of Theorem 1.1 is roughly as follows. Assuming by contradiction that there exists $u_n \to u$ strongly in $L^1(\Omega)$ such that $\mathrm{MS}(u_n) < \mathrm{MS}(u)$, it is easily seen (through Ambrosio's theorem) that $\mathcal{H}^1(J_{u_n}) = \varepsilon_n \to 0$. We then consider the *constrained* minimization problem

$$\min_{\substack{w \in \mathrm{SBV}(\Omega) \\ \mathcal{H}^1(J_w) \le \varepsilon_n}} \mathrm{MS}(w)$$

whose minimizers w_n are such that

$$\mathrm{MS}(w_n) \le \mathrm{MS}(u_n) < \mathrm{MS}(u). \tag{1-6}$$

Writing $w_n = u + v_n$, we can show that v_n satisfy the weak minimality property (1-3), so that monotonicity is available, and that

$$\int_\Omega |\nabla v_n|^2 \, dx + \mathcal{H}^1(J_{v_n}) \to 0. \tag{1-7}$$

Since $J_{w_n} = J_{v_n}$, inequality (1-6) can hold only for $J_{v_n} \neq \varnothing$. If $x_n \in J_{v_n}$ is a point of density 1 with $x_n \to x \in \Omega$, monotonicity gives a strictly positive uniform lower bound for $E_{x_n}(\rho)$ defined in (1-4). But (1-7) yields the existence of ρ_n such that $E_{x_n}(\rho_n) \to 0$, a contradiction. If x_n approaches $\partial\Omega$, boundary monotonicity for (1-5) can be employed to get the same conclusion.

In Section 6 we extend the minimality result to a free discontinuity functional of the type

$$F(u) := \int_\Omega A(x)\nabla u \cdot \nabla u \, dx + \int_{J_u} b(x, u^+, u^-, \nu_u) \, d\mathcal{H}^1 + \beta \int_\Omega |u - g|^2 \, dx \tag{1-8}$$

under suitable assumptions on the coefficient A and b (here ν_u denotes the normal vector to J_u, $u^\pm$ are the two traces of u on the jump set, see Section 2) and considering also boundary condition of Dirichlet type on a part Γ_D of the boundary (here $\beta = 0$ is allowed). Again, provided that Ω satisfies some geometric assumptions (see Theorem 6.7), the minimizer of F within the class $H^1(\Omega)$ under boundary conditions is a local minimizer with respect to the L^1 topology in $\mathrm{SBV}(\Omega)$. We thus recover for $\beta = 0$ a particular case of the local minimality result of [Chambolle et al. 2008], and extend it to the case $\beta > 0$, i.e., in the presence of a fidelity term: by employing monotonicity instead of a maximum principle argument, we need to restrict to quadratic gradient energies and require additional geometric assumptions on Γ_D (see also Remark 6.6). Finally, as a numerical consequence of our result, it is worth observing that for any iterative local descent method the regularized image given by (1-2) cannot be used as initial point. In particular, the topological derivative of the Mumford–Shah functional will be nonnegative at any point; hence no jump can be naturally detected in this way.

The paper is organized as follows. In Section 2 we fix the notation and recall some basic facts of the space SBV used in the main proofs. In Section 3 we prove some preliminary results concerning harmonic extensions in corner domains which are pivotal to extend the monotonicity formula up to boundary points in Section 4. Section 5 is devoted to the proof of the main local minimality result, while the extension to the case with boundary conditions with the general form (1-8) is contained in Section 6.

2. Notation and preliminaries

In this section we fix the notation and recall some basic facts concerning the space SBV.

General notation. Throughout the paper $B_\rho(x)$ will denote the open ball of center $x \in \mathbb{R}^2$ and radius $\rho > 0$. We will write B_1 for the ball of center 0 and radius 1, and set $S^1 := \partial B_1$. If $E \subseteq \mathbb{R}^2$ and $x \in \mathbb{R}^2$, $\mathrm{dist}(x, E)$ will stand for the distance between x and E. By $\mathcal{H}^1$ we will denote the one-dimensional Hausdorff measure, which coincides with the usual length measure on sufficiently regular curves.

If $\Omega \subseteq \mathbb{R}^2$ is open, $L^p(\Omega)$ will denote the usual Lebesgue space of measurable functions which are p-summable, while $H^1(\Omega)$ will stand for the Sobolev space of functions in $L^2(\Omega)$ with square integrable gradient.

A function $f : \mathbb{R} \to \mathbb{R}$ is of class $C^{1,1}$ if it belongs to $C^1(\mathbb{R})$ and f' is Lipschitz continuous. We say that f is piecewise $C^{1,1}$ if it is continuous and there exists a (locally finite) subdivision of $\mathbb{R}$ such that on every associated open subinterval f is of class $C^{1,1}$. An open bounded domain $\Omega \subseteq \mathbb{R}^2$ has a piecewise $C^{1,1}$-boundary if for every $x \in \partial\Omega$ there exists a neighborhood U and a piecewise $C^{1,1}$ function f on $\mathbb{R}$ such that, up to a rotation, $\Omega \cap U = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 > f(x_1)\} \cap U$.

Finally, for $a, b \in \mathbb{R}$ we set

$$a \wedge b := \min\{a, b\} \quad \text{and} \quad a \vee b := \max\{a, b\}.$$

Special functions of bounded variation. Let $\Omega \subseteq \mathbb{R}^N$ be an open set. The space $\mathrm{SBV}(\Omega)$ of special functions of bounded variation is given by all functions $u \in L^1(\Omega)$ such that the distributional derivative Du of u can be represented as a vector-valued bounded Radon measure of the form

$$Du = \nabla u \mathcal{L}^N + (u^+ - u^-)\nu_u \mathcal{H}^{N-1} \llcorner J_u.$$

Here $\mathcal{L}^N$ is the Lebesgue measure, $\nabla u \in L^1(\Omega; \mathbb{R}^N)$ is the approximate gradient of u, and J_u is the jump set of u. The set J_u turns out to be countably $\mathcal{H}^{N-1}$-rectifiable; i.e., it is contained up to a set of $\mathcal{H}^{N-1}$-measure zero in the union of C^1-submanifolds of $\mathbb{R}^N$. It is possible to define $\mathcal{H}^{N-1}$-a.e. on J_u an approximate normal denoted by ν_u, as well as traces $u^\pm$. We refer to [Ambrosio et al. 2000] for a detailed account of this topic.

The following result is fundamental when dealing with the analysis of the Mumford–Shah functional.

Theorem 2.1 (Ambrosio's theorem). *Let $\Omega \subseteq \mathbb{R}^N$ be open and bounded, and let $(u_n)_{n\in\mathbb{N}}$ be a sequence in* $\mathrm{SBV}(\Omega)$ *such that*

$$\|\nabla u_n\|_p + \mathcal{H}^{N-1}(J_{u_n}) + \|u_n\|_\infty \le C,$$

where $p > 1$ and $C \ge 0$. Then, there exist a subsequence $(u_{n_k})_{k\in\mathbb{N}}$ and a function $u \in \mathrm{SBV}(\Omega)$ such that

$$\nabla u_{n_k} \rightharpoonup \nabla u \quad \textit{weakly in } L^p(\Omega; \mathbb{R}^N),$$

$$\mathcal{H}^{N-1}(J_u) \le \liminf_{k\to+\infty} \mathcal{H}^{N-1}(J_{u_{n_k}}),$$

$$u_{n_k} \to u \quad \textit{strongly in } L^1(\Omega).$$

3. Harmonic extension results

In this section we collect some harmonic extension estimates on different kinds of domains.

We start by recalling the interior harmonic extension estimate for a disk, which was already exploited in [Bucur and Luckhaus 2014] to obtain the monotonicity formula. As a straightforward extension, we generalize the estimate to the case of circular sectors. Then, through a careful deformation procedure, we handle the case of what we call "corner domains", namely the epigraphs of functions with a corner-type singularity which obey suitable estimates. Finally, via a localization argument, we cover the case of admissible $C^{1,1}$ domains, which will be a crucial tool to extend the monotonicity formula up to the boundary. Below and throughout this section, we denote by ∇_τ the tangential gradient.

Let us start with the simple case of the unit disk B_1.

Proposition 3.1 (harmonic extension estimate in a disk). *Let $w \in H^1(\partial B_1)$, and let $h_w \in H^1(B_1)$ be its harmonic extension. Then*

$$\int_{B_1} |\nabla h_w|^2 \, dx \le \int_{\partial B_1} |\nabla_\tau w|^2 \, d\mathcal{H}^1. \tag{3-1}$$

Proof. Several proofs of this inequality are available in the literature; see for instance [Bucur and Luckhaus 2014] for a proof in N dimensions. For further needs, we give below a short two-dimensional proof employing Fourier expansions. Let us write

$$w(\vartheta) = \sum_n [a_n \cos(n\vartheta) + b_n \sin(n\vartheta)].$$

The harmonic extension on B_1 is given in polar coordinates by

$$h_w(r, \vartheta) = \sum_n r^n [a_n \cos(n\vartheta) + b_n \sin(n\vartheta)].$$

A direct computation shows that

$$\int_{B_1} |\nabla h_w|^2 \, dx = \pi \sum_n n(a_n^2 + b_n^2) \quad \text{and} \quad \int_{\partial B_1} |\nabla_\tau w|^2 \, d\mathcal{H}^1 = \pi \sum_n n^2(a_n^2 + b_n^2),$$

so that the estimate easily follows. □

Let us generalize the estimate (3-1) to the case of circular sectors. Given $\vartheta_0 \in [0, 2\pi]$, we consider the circular unit sector and associated arc given in polar coordinates by

$$S_{\vartheta_0} := \left\{(r, \vartheta) : r \in [0, 1], \vartheta \in \left[-\frac{\vartheta_0}{2}, \frac{\vartheta_0}{2}\right]\right\} \quad \text{and} \quad \Gamma_{\vartheta_0} := \left\{(r, \vartheta) : r = 1, \vartheta \in \left[-\frac{\vartheta_0}{2}, \frac{\vartheta_0}{2}\right]\right\}.$$

The following extension of Proposition 3.1 holds true.

Proposition 3.2 (harmonic extension estimate in a sector). *Let $w \in H^1(\Gamma_{\vartheta_0})$, and let $h_w \in H^1(S_{\vartheta_0})$ denote a harmonic extension of w. Then the inequality*

$$\int_{S_{\vartheta_0}} |\nabla h_w|^2 \, dx \le \int_{\Gamma_{\vartheta_0}} |\nabla_\tau w|^2 \, d\mathcal{H}^1$$

holds true provided that one of the following assumptions is fulfilled:

(a) $\vartheta_0 \in]0, \pi]$ *and* h_w *satisfies Neumann conditions or Dirichlet homogeneous conditions on* $\partial S_{\vartheta_0} \setminus \Gamma_{\vartheta_0}$.

(b) $\vartheta_0 \in]0, \pi/2]$ *and* h_w *satisfies homogeneous Dirichlet conditions for* $\vartheta = -\vartheta_0/2$ *and Neumann conditions for* $\vartheta = \vartheta_0/2$.

Above, a harmonic function $h \in H^1(S_{\vartheta_0})$ is said to satisfy Neumann conditions on a portion of the boundary $\Gamma_N \subset \partial S_{\vartheta_0}$ and Dirichlet conditions g on the complement $\Gamma_D = \partial S_{\vartheta_0} \setminus \Gamma_N$ for some function $g \in H^1(S_{\vartheta_0})$ if h is a minimizer of

$$\min\left\{\int_{S_{\vartheta_0}} |\nabla v|^2\, dx : v \in H^1(S_{\vartheta_0}),\ v = g \text{ on } \Gamma_D\right\}.$$

Proof. Let us consider the case of homogeneous Dirichlet conditions. We may develop w as

$$w = \sum_{j=1}^{\infty} c_j \sin\left[j\frac{\pi}{\vartheta_0}\left(\vartheta + \frac{\vartheta_0}{2}\right)\right]$$

so that the associated harmonic extension takes the form

$$h_w = \sum_{j=1}^{\infty} c_j r^{j\pi/\vartheta_0} \sin\left[j\frac{\pi}{\vartheta_0}\left(\vartheta + \frac{\vartheta_0}{2}\right)\right].$$

A straightforward computation for the energies of the gradients in polar coordinates shows that

$$\frac{\pi}{\vartheta_0}\int_{S_{\vartheta_0}} |\nabla h_w|^2\, dx \le \int_{\Gamma_{\vartheta_0}} |\nabla_\tau w|^2\, d\mathcal{H}^1,$$

so that the result follows.

Let us come to the case of Neumann conditions. The case of the semicircle, i.e., $\vartheta_0 = \pi$, is easily obtained through an even extension of w to ∂B_1 and employing Proposition 3.1. If $\vartheta_0 \in]0, \pi[$, we pass from w to a function $\tilde{w}$ by prolonging the constant values of the extremes. Applying the inequality on the semicircle we may write

$$\int_{S_{\vartheta_0}} |\nabla h_w|^2\, dx \le \int_{S_{\vartheta_0}} |\nabla h_{\tilde{w}}|^2\, dx \le \int_{S_\pi} |\nabla h_{\tilde{w}}|^2\, dx \le \int_{\Gamma_\pi} |\nabla_\tau \tilde{w}|^2\, d\mathcal{H}^1 = \int_{\Gamma_{\vartheta_0}} |\nabla_\tau w|^2\, d\mathcal{H}^1,$$

so that the result follows, and point (a) is proved.

Let us come finally to point (b) which follows by combing the idea used above. If $\vartheta_0 = \pi/2$, we can perform an even extension to a semicircle and employ the harmonic estimate with Dirichlet conditions of point (a), and then restrict to $\Gamma_{\pi/2}$ to get the result. If $\vartheta_0 \in]0, \pi/2[$, we can proceed by extension to $\Gamma_{\pi/2}$ and using the associated inequality. □

We are now going to consider the case of *corner domains*. By corner domain, we mean the epigraph (see Figure 1)

$$\Omega_f := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 > f(x_2)\} \tag{3-2}$$

of a given function f of the form

$$f(x) := \begin{cases} f_1(x) & \text{if } x < 0, \\ f_2(x) & \text{if } x \ge 0, \end{cases} \tag{3-3}$$

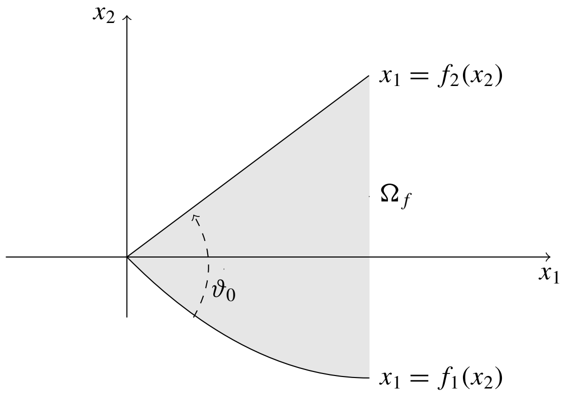

Figure 1. The domain Ω_f.

with $f_1 \in C^1(]-\infty, 0])$ and $f_2 \in C^1([0, +\infty[)$ nonnegative functions satisfying

$$f_1(0) = f_2(0) = 0, \quad f_2'(0) = -f_1'(0). \tag{3-4}$$

We define

$$\vartheta_0(\Omega_f) := \text{angle of } \Omega_f \text{ at the origin}, \tag{3-5}$$

meant as the angle at the origin of the sector $\{x_1 \geq f_1'(0)x_2\} \cap \{x_1 \geq f_2'(0)x_2\}$. In particular, $\vartheta_0(\Omega_f) = \pi$ in the case $\partial\Omega_f$ is smooth ($f_1'(0) = f_2'(0) = 0$). We also set

$$\partial_1\Omega_f := \{(f_1(x), x) : x < 0\} \cap \bar{B}_1(0) \quad \text{and} \quad \partial_2\Omega_f := \{(f_2(x), x) : x \geq 0\} \cap \bar{B}_1(0). \tag{3-6}$$

The following result holds true.

Proposition 3.3 (harmonic extension estimate in corner domains). *For $\rho > 0$, let Ω_{f_ρ} be a family of corner domains as in* (3-2), *where each function f_ρ is of the kind* (3-3)–(3-4). *Assume further that*:

- *The left and right derivatives of f_ρ at the origin are independent of ρ*; *i.e.*,

$$f_{\rho,2}'(0) = -f_{\rho,1}'(0) =: \lambda \geq 0. \tag{3-7}$$

- *$f_{\rho,1}'$, $f_{\rho,2}'$ satisfy the following Lipschitz-type estimates for some constant $c > 0$*:

$$\begin{aligned} &\textit{for all } x \in [-1, 0], \quad |f_{\rho,1}'(x) + \lambda| \leq c\rho|x|, \\ &\textit{for all } x \in [0, 1], \quad\;\; |f_{\rho,2}'(x) - \lambda| \leq c\rho x. \end{aligned} \tag{3-8}$$

Then there exist $\rho_0 > 0$ and $c_0 > 0$ such that, *for every $\rho \in (0, \rho_0]$ and $w \in H^1(\Omega_{f_\rho} \cap \partial B_1(0))$*, *denoting by h_w a harmonic extension of w to $\Omega_{f_\rho} \cap B_1(0)$*, *it holds that*

$$(1 - c_0\rho) \int_{\Omega_{f_\rho} \cap B_1(0)} |\nabla h_w|^2 \, dx \leq \int_{\Omega_{f_\rho} \cap \partial B_1(0)} |\nabla_\tau w|^2 \, d\mathcal{H}^1, \tag{3-9}$$

provided that $\vartheta_0 := \vartheta_0(\Omega_{f_\rho})$ (the inner angle at the origin of Ω_{f_ρ}) and h_w fulfill one of the following assumptions:

(a) $\vartheta_0 \in]0, \pi]$ *and* h_w *satisfies Neumann conditions or Dirichlet homogeneous conditions on* $\partial_1\Omega_{f_\rho} \cup \partial_2\Omega_{f_\rho}$.

(b) $\vartheta_0 \in]0, \pi/2]$ *and* h_w *satisfies homogeneous Dirichlet conditions on* $\partial_1\Omega_{f_\rho}$ *and Neumann conditions on* $\partial_2\Omega_{f_\rho}$.

In order to prove the previous proposition, we need some preliminary work. Let us consider the planar domain given in polar coordinates (r, ϑ) by

$$C_{\vartheta_0,g} := \left\{(r, \vartheta) : r \in [0, 1],\ \vartheta \in \left[0, \frac{\vartheta_0}{2} + g(r)\right]\right\}, \tag{3-10}$$

where

$$\vartheta_0 \in]0, \pi] \quad \text{and} \quad g \in C^1([0, 1]),\ g(0) = 0,\ g(r) \in \left[-\frac{\vartheta_0}{2}, -\frac{\vartheta_0}{2} + \frac{\pi}{2}\right]. \tag{3-11}$$

The domain $C_{\vartheta_0,g}$ can be mapped to the unit sector

$$S^+_{\vartheta_0} := \left\{(r, \vartheta) : r \in [0, 1],\ \vartheta \in \left[0, \frac{\vartheta_0}{2}\right]\right\}$$

by means of the transformation defined by

$$T : C_{\vartheta_0,g} \to S^+_{\vartheta_0}, \quad T(r, \vartheta) := \left(r, \frac{\vartheta_0}{2}\frac{\vartheta}{\vartheta_0/2 + g(r)}\right), \tag{3-12}$$

with inverse

$$T^{-1}(r', \vartheta') := \left(r', \frac{2}{\vartheta_0}\left(\frac{\vartheta_0}{2} + g(r')\right)\vartheta'\right).$$

Notice that T maps $C_{\vartheta_0,g} \cap \partial B_1(0)$ onto $\Gamma^+_{\vartheta_0} := \{(r, \vartheta) : r = 1, \vartheta \in [0, \vartheta_0/2]\}$. Accordingly, with a given function v defined on $C_{\vartheta_0,g}$ (resp. on $C_{\vartheta_0,g} \cap \partial B_1(0)$), we can associate the function $v^\sharp$ defined on $S^+_{\vartheta_0}$ (resp. $\Gamma^+_{\vartheta_0}$) by

$$v^\sharp := v \circ T^{-1}. \tag{3-13}$$

Lemma 3.4. *For* $\rho > 0$, *let* C_{ϑ_0,g_ρ} *be a family of domains as in* (3-10), *with* ϑ_0 (*independent of* ρ) *and* g_ρ *as in* (3-11). *Assume further that* g'_ρ *satisfy the following* L^∞ *estimate for some constant* $c > 0$:

$$|g'_\rho(r)| \le c\rho \quad \text{for every } r \in [0, 1]. \tag{3-14}$$

Let $T : C_{\vartheta_0,g_\rho} \to S^+_{\vartheta_0}$ *be defined as in* (3-12). *There exist* $\rho_0 > 0$ *and* $c_0 > 0$ *such that for every* $\rho \in (0, \rho_0]$ *the following items hold true*:

(a) *If* $v^\sharp \in H^1(S^+_{\vartheta_0})$ *is associated with* $v \in H^1(C_{\vartheta_0,g_\rho})$ *as in* (3-13), *then*

$$\int_{S^+_{\vartheta_0}} |\nabla v^\sharp|^2\, dx' \ge (1 - c_0\rho) \int_{C_{\vartheta_0,g_\rho}} |\nabla v|^2\, dx.$$

(b) *If* $w^\sharp \in H^1(\Gamma^+_{\vartheta_0})$ *is associated with* $w \in H^1(C_{\vartheta_0,g} \cap \partial B_1(0))$ *as in* (3-13), *then*

$$\int_{\Gamma^+_{\vartheta_0}} |\nabla_\tau w^\sharp|^2\, d\mathcal{H}^1 \le (1 + c_0\rho) \int_{C_{\vartheta_0,g_\rho} \cap \partial B_1(0)} |\nabla_\tau w|^2\, d\mathcal{H}^1.$$

Proof. Let us prove item (a). We write for simplicity g in place of g_ρ. By (3-13), we have

$$v(r,\vartheta)=v^\sharp(T(r,\vartheta))=v^\sharp\left(r,\frac{\vartheta_0}{2}\frac{\vartheta}{\vartheta_0/2+g(r)}\right).$$

Computing the partial derivatives of the function v, we obtain

$$\begin{aligned}\partial_r v(r,\vartheta)&=\partial_{r'}v^\sharp(T(r,\vartheta))+\partial_{\vartheta'}v^\sharp(T(r,\vartheta))\frac{\vartheta_0}{2}\frac{-\vartheta g'(r)}{(\vartheta_0/2+g(r))^2},\\ \partial_\vartheta v(r,\vartheta)&=\partial_{\vartheta'}v^\sharp(T(r,\vartheta))\frac{\vartheta_0}{2}\frac{1}{\vartheta_0/2+g(r)},\end{aligned}$$

so that

$$\begin{aligned}\partial_{r'}v^\sharp(T(r,\vartheta))&=\partial_r v(r,\vartheta)+\frac{\vartheta g'(r)}{\vartheta_0/2+g(r)}\partial_\vartheta v(r,\vartheta),\\ \partial_{\vartheta'}v^\sharp(T(r,\vartheta))&=\frac{2}{\vartheta_0}\left(\frac{\vartheta_0}{2}+g(r)\right)\partial_\vartheta v(r,\vartheta).\end{aligned}\tag{3-15}$$

Since the jacobian of T is given by

$$J_T(r,\vartheta)=\frac{\vartheta_0}{2}\frac{1}{\vartheta_0/2+g(r)},$$

the change of variable formula together with (3-15) yields

$$\int_{S^+_{\vartheta_0}}\left[(\partial_{r'}v^\sharp)^2+\left(\frac{\partial_{\vartheta'}v^\sharp}{r'}\right)^2\right]r'\,dr'\,d\vartheta'=\int_{C_{\vartheta_0,g}}\left[(\partial_r v)^2+\left(\frac{\partial_\vartheta v}{r}\right)^2+e(\rho)\right]\frac{\vartheta_0}{2}\frac{1}{\vartheta_0/2+g(r)}r\,dr\,d\vartheta,\tag{3-16}$$

where

$$e(\rho):=2\frac{\vartheta g'(r)}{\vartheta_0/2+g(r)}\partial_r v\partial_\vartheta v+\left(\frac{\vartheta g'(r)}{\vartheta_0/2+g(r)}\right)^2(\partial_\vartheta v)^2+\left[\left(\frac{2}{\vartheta_0}\left(\frac{\vartheta_0}{2}+g(r)\right)\right)^2-1\right]\left(\frac{\partial_\vartheta v}{r}\right)^2.\tag{3-17}$$

It is easy to check that there exist $\rho_0,\tilde c>0$ (depending only on the constant c in (3-14) and on ϑ_0) such that for every $\rho\in(0,\rho_0]$

$$-\tilde c\rho\left[(\partial_r v)^2+\left(\frac{\partial_\vartheta v}{r}\right)^2\right]\le e(\rho)\le\tilde c\rho\left[(\partial_r v)^2+\left(\frac{\partial_\vartheta v}{r}\right)^2\right].\tag{3-18}$$

Indeed, let us show for instance how the first term on the right-hand side of (3-17) can be handled (the other terms being similar). Notice first that the assumption (3-14) gives also $|g(r)|\le c\rho$ for every $r\in[0,1]$. Then, taking ρ sufficiently small we obtain, for a suitable constant $\tilde c>0$,

$$\begin{aligned}\left|2\frac{\vartheta g'(r)}{\vartheta_0/2+g(r)}\partial_r v\partial_\vartheta v\right|&\le\left|\frac{\vartheta g'(r)}{\vartheta_0/2+g(r)}\right|\left[r^2(\partial_r v)^2+\left(\frac{\partial_\vartheta v}{r}\right)^2\right]\\ &\le\frac{(\vartheta_0/2+c\rho)c\rho}{\vartheta_0/2-c\rho}\left[(\partial_r v)^2+\left(\frac{\partial_\vartheta v}{r}\right)^2\right]\le\tilde c\rho\left((\partial_r v)^2+\left(\frac{\partial_\vartheta v}{r}\right)^2\right).\end{aligned}$$

In view of (3-18), coming back to (3-16) we get

$$\int_{S^+_{\vartheta_0}}\left[(\partial_{r'}v^\sharp)^2+\left(\frac{\partial_{\vartheta'}v^\sharp}{r'}\right)^2\right]r'\,dr'\,d\vartheta' \geq (1-\tilde{c}\rho)\int_{C_{\vartheta_0}}\left[(\partial_r v)^2+\left(\frac{\partial_\vartheta v}{r}\right)^2\right]\frac{\vartheta_0}{2}\frac{1}{\vartheta_0/2+g(r)}r\,dr\,d\vartheta$$
$$\geq \frac{1-\tilde{c}\rho}{1-(2/\vartheta_0)c\rho}\int_{C_{\vartheta_0,g}}\left[(\partial_r v)^2+\left(\frac{\partial_\vartheta v}{r}\right)^2\right]r\,dr\,d\vartheta,$$

from which part (a) of the statement follows.

The proof of part (b) is analogous, by using the lower bound inequality in (3-18) in place of the upper bound one. □

We are now in a position to prove Proposition 3.3.

Proof of Proposition 3.3.. Throughout the proof we write for simplicity f in place of f_ρ and g in place of g_ρ. We claim that

$$\Omega_f^+ := \Omega_f \cap B_1(0) \cap \{x_2 > 0\}$$

corresponds to the domain $C_{\vartheta_0,g}$ defined in (3-10), with ϑ_0 and g satisfying (3-11) and (3-14), if we take

$$\vartheta_0 = \vartheta_0(\Omega_f), \quad g(r) := \arccos\left(\frac{f_2(x_2)}{\sqrt{x_2^2+f_2(x_2)^2}}\right) - \frac{\vartheta_0}{2}, \tag{3-19}$$

where $\vartheta_0(\Omega_f)$ is the angle of Ω_f at the origin defined according to (3-5), and $x_2 = x_2(r)$ is obtained by inverting the relation

$$r = \sqrt{x_2^2 + f_2(x_2)^2}. \tag{3-20}$$

To prove the claim notice firstly that $x_2(r)$ is well-defined because (3-20) defines a bijection if ρ is sufficiently small: if $f_2'(0) \neq 0$, this follows immediately from (3-8); if $f_2'(0) = 0$, setting $h(x_2) := x_2^2 + f_2(x_2)^2$ and using again (3-8), we get

$$h'(x_2) = 2x_2 + 2f_2(x_1)f_2'(x_2) \geq 2x_2(1-c_2^2\rho^2) > 0 \quad \text{for } x_2 > 0 \text{ and } \rho \ll 1.$$

Next observe that, since by assumption f_2 is nonnegative, the domain Ω_f^+ can be written in polar coordinates (r, ϑ) as

$$0 \leq r \leq 1, \quad 0 \leq \vartheta \leq \varphi(r) := \arccos\left(\frac{f_2(x_2)}{\sqrt{x_2^2+f_2(x_2)^2}}\right) \tag{3-21}$$

(with $x_2 = x_2(r)$ as above). In view of (3-21), and noticing that the function φ satisfies

$$\varphi(0^+) = \arccos\left(\frac{f_2'(0)}{\sqrt{1+f_2'(0)^2}}\right) = \frac{\vartheta_0(\Omega_f)}{2},$$

we see that Ω_f^+ corresponds to $C_{\vartheta_0,g}$ with ϑ_0 and g as in (3-19). To achieve the proof of the claim, it remains to check that ϑ_0 and g satisfy (3-11) and (3-14). Since ϑ_0 is the angle of Ω_f at the origin, recalling that f_1 and f_2 are nonnegative by assumption, it is clear that $\vartheta_0 \in [0, \pi[$. It is also immediate from the above definition of g that $g \in C^1([0, 1])$ and that $g(0) = \varphi(0) - \vartheta_0/2 = 0$. Moreover, since the function φ in (3-21) takes values into $[0, \pi/2]$, g satisfies the bounds in (3-11). Finally, taking into account (3-8), a straightforward computation shows that there exists $c > 0$ such that (3-14) is fulfilled.

A similar change of variable can be operated on the set $\Omega_f^- := \Omega_f \cap B_1(0) \cap \{x_2 < 0\}$. We conclude that we can map $\Omega_f \cap B_1(0)$ onto the sector S_{ϑ_0}, with an estimate on the L^2-norm of the gradients according to Lemma 3.4.

We are now ready to prove the estimate (3-9). Let us consider firstly the case of Neumann conditions. Given $w \in H^1(\Omega_f \cap \partial B_1(0))$, let $w^\sharp \in H^1(\Gamma_{\vartheta_0})$ be associated with w according to (3-13). If $h_{w^\sharp}$ is the harmonic extension of $w^\sharp$ to S_{ϑ_0} with Neumann conditions on $\partial S_{\vartheta_0} \setminus \Gamma_{\vartheta_0}$, thanks to Proposition 3.2(a) we have

$$\int_{S_{\vartheta_0}} |\nabla h_{w^\sharp}|^2 \, dx \le \int_{\Gamma_{\vartheta_0}} |\nabla_\tau w^\sharp|^2 \, d\mathcal{H}^1. \tag{3-22}$$

Coming back to the original domain, if we consider the function $h \in H^1(\Omega_f \cap B_1(0))$ defined by $h := h_{w^\sharp} \circ T$ (so that $h^\sharp = h_{w^\sharp}$), by using in the order Lemma 3.4(a), (3-22), and Lemma 3.4(b), we infer

$$\begin{aligned}(1 - c_0\rho) \int_{\Omega_f \cap B_1(0)} |\nabla h|^2 \, dx &\le \int_{S_{\vartheta_0}} |\nabla h_{w^\sharp}|^2 \, dx \\ &\le \int_{\Gamma_{\vartheta_0}} |\nabla_\tau w^\sharp|^2 \, d\mathcal{H}^1 \le (1 + c_0\rho) \int_{\Omega_f \cap \partial B_1(0)} |\nabla_\tau w|^2 \, d\mathcal{H}^1.\end{aligned}$$

Since the harmonic extension h_w of w which satisfies Neumann conditions on $\partial_1\Omega_f \cup \partial_2\Omega_f$ satisfies the inequality

$$\int_{\Omega_f \cap B_1(0)} |\nabla h_w|^2 \, dx \le \int_{\Omega_f \cap B_1(0)} |\nabla h|^2 \, dx,$$

the result follows.

The case of Dirichlet homogeneous conditions or mixed Dirichlet/Neumann conditions can be handled in a similar way, by employing the corresponding inequalities on sectors established in Proposition 3.2. □

In order to handle more general geometries and also boundary conditions, the following definition will be useful.

Definition 3.5 (admissible domains). Let $\Omega \subset \mathbb{R}^2$ be an open bounded set with a piecewise $C^{1,1}$-boundary. Let S denote the set of corners of $\partial\Omega$, and let $\gamma_1, \dots, \gamma_k$ be the open arcs (connected components) of $\partial\Omega \setminus S$. We will say that Ω is *admissible* if $\partial\Omega$ can be decomposed as

$$\partial\Omega = \Gamma_D \cup \Gamma_N \cup S, \tag{3-23}$$

where $\Gamma_D \cup \Gamma_N$ is a partition of $\{\gamma_1, \dots, \gamma_k\}$ and satisfies the following conditions:

(i) The angle formed (on the side of Ω) by any pair of consecutive arcs in Γ_D or in Γ_N is less than or equal to π.

(ii) The angle formed (on the side of Ω) by any arc of Γ_D adjacent to an arc of Γ_N is less than or equal to $\pi/2$.

Remark 3.6. In our further considerations Γ_D and Γ_N stand for the arcs with Dirichlet and Neumann boundary conditions respectively For $\Gamma_D = \partial\Omega$ or $\Gamma_N = \partial\Omega$, admissibility according to the above definition reduces to the assumption that Ω has convex corners. In particular, smooth $C^{1,1}$ domains are admissible.

Proposition 3.7 (harmonic extension estimate in admissible domains). *Let $\Omega \subset \mathbb{R}^2$ be an open bounded set with a piecewise $C^{1,1}$-boundary, which is admissible according to Definition 3.5. Let $\partial\Omega$ be decomposed as in* (3-23). *There exist $\rho_\Omega > 0$ and $c_\Omega > 0$ such that, for every $x \in \partial\Omega$ and every $\rho < \rho_\Omega \wedge \operatorname{dist}(x, S\setminus\{x\})$, if $w \in H^1(\Omega \cap \partial B_\rho(x))$ and h_w denotes a harmonic extension of w to $\Omega \cap B_\rho(x)$, it holds that*

$$\frac{(1 - c_\Omega \rho)}{\rho} \int_{\Omega\cap B_\rho(x)} |\nabla h_w|^2\, dx \le \int_{\Omega\cap\partial B_\rho(x)} |\nabla_\tau w|^2\, d\mathcal{H}^1, \tag{3-24}$$

provided one of the following assumptions is fulfilled:

(a) *h_w satisfies Neumann conditions or homogeneous Dirichlet conditions on $\partial\Omega \cap B_\rho(x)$.*

(b) *h_w satisfies homogeneous Dirichlet conditions on $\Gamma_D \cap B_\rho(x)$ and Neumann conditions on $\Gamma_N \cap B_\rho(x)$.*

Proof. Let us first consider the case when h_w satisfies Neumann conditions on $\partial\Omega \cap B_\rho(x)$. Since Ω is compact and with a piecewise $C^{1,1}$-boundary, there exists $\rho_\Omega > 0$ such that, for every $x \in \partial\Omega$ and $\rho < \rho_\Omega$, we have that $\Omega \cap B_\rho(x)$ is given in a suitable coordinate system with center in x by the intersection with $B_\rho(x)$ of the epigraph of a piecewise $C^{1,1}$-function $f_x : \mathbb{R} \to \mathbb{R}$. In view of the compactness of $\partial\Omega$, it is not restrictive to assume

$$\sup_{x\in\partial\Omega} \|f_x\|_{C^{1,1}_{pw}} = C < +\infty. \tag{3-25}$$

Let $x \in \partial\Omega$ be a smooth point, and let

$$\rho < \rho_\Omega \wedge \operatorname{dist}(x, S).$$

Up to a translation and a rotation, we may assume that $x = 0$ and that $\Omega \cap B_\rho(x)$ is given by

$$\Omega_f \cap B_\rho(0),$$

where

$$\Omega_f := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 > f(x_2)\}$$

for a suitable $f \in C^{1,1}(\mathbb{R})$ with $f(0) = f'(0) = 0$ and $\|f\|_{C^{1,1}} \le C$. If we rescale to unit size, i.e., we consider the map

$$x \mapsto x/\rho,$$

then $\Omega_f \cap B_\rho(0)$ is transformed into the set $\Omega_{f_\rho} \cap B_1(0)$, where

$$f_\rho(s) := \frac{1}{\rho} f(\rho s).$$

Since $f_\rho'(s) = f'(\rho s)$ and $f(0) = f'(0) = 0$, in view of (3-25) we deduce that for every $s \in [-1, 1]$

$$|f_\rho'(s)| \le c\rho|s|.$$

Hence, the family of functions f_ρ satisfies conditions (3-7) and (3-8). Therefore, up to reducing ρ_Ω, we can apply Proposition 3.3 (under assumption (a), the case of Neumann conditions) to the function

$$\tilde{w}(y) := w(\rho y) \in H^1(\Omega_{f_\rho} \cap \partial B_1(0)).$$

Rescaling back to size ρ, we obtain easily the result.

If x is a convex corner for $\partial\Omega$, the proof is similar, using again Proposition 3.3 (under assumption (a), the case of Neumann conditions).

The cases when h_w satisfies homogeneous Dirichlet conditions on $\partial\Omega \cap B_\rho(x)$, or homogeneous Dirichlet conditions on $\Gamma_D \cap B_\rho(x)$ and Neumann conditions on $\Gamma_N \cap B_\rho(x)$, can be settled in the analogous way, by using the parts of Proposition 3.3 in which the corresponding boundary conditions are considered. □

Remark 3.8. An inspection in the proof of Proposition 3.7 shows that the constants c_Ω, ρ_Ω remain bounded if Ω is replaced by the domain $L(\Omega)$ where $L : \mathbb{R}^2 \to \mathbb{R}^2$ varies in a family of linear transformations with bounded norm. This observation will be useful in Section 6.

4. The monotonicity formula up to the boundary

In this section we prove a monotonicity formula up to boundary points for a suitable notion of quasiminimizers of the Mumford–Shah functional in dimension 2. We point out that this notion is much weaker than the classical one employed in [Bucur and Luckhaus 2014], since the family of test functions is much smaller. The results of this section are typically two-dimensional and cannot, a priori, be extended in N dimensions.

Let us start with the following definition.

Definition 4.1 (weak minimizers). Let $\Omega \subseteq \mathbb{R}^2$ be an open set, and $u \in \mathrm{SBV}(\Omega)$. We say that u is a weak local almost-quasiminimizer of the Mumford–Shah energy with respect to its own jump set at the point $x \in \overline{\Omega}$ if there exist $\rho_x > 0$, $\gamma > 0$, $c_\gamma > 0$ and $\Lambda \geq 1$ such that for every $\rho < \rho_x$ with

$$\int_{\Omega \cap B_\rho(x)} |\nabla u|^2 \, dx + \mathcal{H}^1(J_u \cap \overline{B}_\rho(x)) \leq \rho \tag{4-1}$$

and for every $v \in \mathrm{SBV}(\Omega)$ with $\{v \neq u\} \subseteq B_\rho(x)$ and $J_v \subseteq J_u$ we have

$$\int_{\Omega \cap B_\rho(x)} |\nabla u|^2 \, dx + \mathcal{H}^1(J_u \cap \overline{B}_\rho(x)) \leq \int_{\Omega \cap B_\rho(x)} |\nabla v|^2 \, dx + \Lambda \mathcal{H}^1(J_v \cap \overline{B}_\rho(x)) + c_\gamma \rho^{1+\gamma}. \tag{4-2}$$

Remark 4.2. Notice that local almost-quasiminimizers of the Mumford–Shah functional considered in [Bucur and Luckhaus 2014] (in particular absolute minimizers) are weak local almost-quasiminimizers with respect to their own jump set: indeed they satisfy the minimality property (4-2) on *every* ball $B_\rho(x)$ (with ρ sufficiently small) not necessarily satisfying the energy bound (4-1), and for every competitor v such that $\{v \neq u\} \subseteq B_\rho(x)$, without the restriction $J_v \subseteq J_u$.

With a given weak local almost-quasiminimizer u according to the previous definition, we associate for every $x \in \bar{\Omega}$ and $\rho > 0$ the quantity

$$\mathcal{E}_x(\rho) := \int_{\Omega \cap B_\rho(x)} |\nabla u|^2 \, dx + \mathcal{H}^1(J_u \cap \bar{B}_\rho(x)). \tag{4-3}$$

Our monotonicity formula at interior points reads as follows.

Theorem 4.3 (interior monotonicity). *Let $\Omega \subset \mathbb{R}^2$ be open, and let $u \in \mathrm{SBV}(\Omega)$ be a weak local almost-quasiminimizer for the Mumford–Shah energy with respect to its own jump set at the point $x \in \Omega$ according to Definition 4.1. Let $\mathcal{E}_x(\rho)$ be associated with u as in* (4-3).

Then the quantity

$$E_x(\rho) := \frac{\mathcal{E}_x(\rho)}{\rho} \wedge 1 + \frac{c_\gamma}{\gamma} \rho^\gamma \tag{4-4}$$

is nondecreasing on $(0, \rho_x \wedge \operatorname{dist}(x, \partial\Omega))$.

This formula appears to have a similar expression to the one in [Bucur and Luckhaus 2014], but applies to a much weaker notion of minimizer. In particular there is no natural upper bound for $\mathcal{E}_x(\rho)/\rho$ as in [Bucur and Luckhaus 2014] (such a bound is usually obtained by using as a test for minimality the function $u 1_{\Omega \setminus B_\rho(x)}$, which is not a priori an admissible competitor for *weak* local almost-quasiminimizers with respect to their own jump set).

Coming to boundary points, we are going to state a monotonicity formula for domains which are admissible according to Definition 3.5. First, let us consider the case of $C^{1,1}$-domains with convex corners (see Remark 3.6), with no imposition of Dirichlet boundary condition. In the case of a flat boundary, such a monotonicity formula has been obtained in [Chambolle et al. $\geq$ 2020], in which case $k_\Omega = 0$.

Theorem 4.4 (boundary monotonicity). *Let $\Omega \subset \mathbb{R}^2$ be an open, bounded, piecewise $C^{1,1}$-domain with convex corners, and let $u \in \mathrm{SBV}(\Omega)$ be a weak local almost-quasiminimizer for the Mumford–Shah energy with respect to its own jump set at the point $x \in \partial\Omega$ according to Definition 4.1. Let $\mathcal{E}_x(\rho)$ be associated with u as in* (4-3).

Then there exist $r_\Omega > 0$ and $k_\Omega \geq 0$ such that the quantity

$$\widetilde{E}_x(\rho) := \frac{\mathcal{E}_x(\rho)}{\rho} \wedge 1 + \frac{c_\gamma}{\gamma} \rho^\gamma + k_\Omega \int_0^\rho \left(\frac{\mathcal{E}_x(r)}{r} \wedge 1 \right) dr \tag{4-5}$$

is nondecreasing on $(0, \rho_x \wedge r_\Omega \wedge \operatorname{dist}(x, S \setminus \{x\}))$, *where S denotes the set of corners of $\partial\Omega$.*

To formulate a monotonicity result which takes into account also boundary conditions, we consider the case when Ω is an admissible domain according to Definition 3.5, and homogeneous Dirichlet boundary conditions are imposed on the (nonempty) portion Γ_D of its boundary. To this aim, let us consider an open bounded domain Ω' such that $\Omega \Subset \Omega' \subset \mathbb{R}^2$ and with (see Figure 2)

$$\partial\Omega \cap \Omega' = \Gamma_D. \tag{4-6}$$

Setting

$$\mathcal{A}_0 := \{u \in \mathrm{SBV}(\Omega') : u = 0 \text{ on } \Omega' \setminus \bar{\Omega}\}. \tag{4-7}$$

we adapt the notion of weak local almost-quasiminimizers as follows.

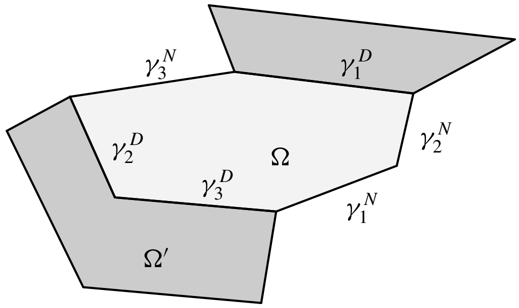

Figure 2. The domain Ω'.

Definition 4.5 (weak minimizers in $\mathcal{A}_0$). We say that $u \in \mathcal{A}_0$ is a weak local almost-quasiminimizer in $\mathcal{A}_0$ of the Mumford–Shah energy with respect to its own jump set at the point $x \in \overline{\Omega}$ if there exist $\rho_x > 0$, $\gamma > 0$, $c_\gamma > 0$ and $\Lambda \geq 1$ such that for every $\rho < \rho_x$ with

$$\int_{\Omega \cap B_\rho(x)} |\nabla u|^2 \, dx + \mathcal{H}^1(J_u \cap \bar{B}_\rho(x)) \leq \rho$$

and for every $v \in \mathcal{A}_0$ with $\{v \neq u\} \subseteq B_\rho(x)$ and $J_v \subseteq J_u$ we have

$$\int_{\Omega \cap B_\rho(x)} |\nabla u|^2 \, dx + \mathcal{H}^1(J_u \cap \bar{B}_\rho(x)) \leq \int_{\Omega \cap B_\rho(x)} |\nabla v|^2 \, dx + \Lambda \mathcal{H}^1(J_v \cap \bar{B}_\rho(x)) + c_\gamma \rho^{1+\gamma}.$$

The following variant of Theorem 4.4 holds true.

Theorem 4.6 (boundary monotonicity in $\mathcal{A}_0$). *Let $\Omega \subset \mathbb{R}^2$ be an admissible domain according to Definition 3.5, and let $u \in \mathcal{A}_0$ be a weak local almost-quasiminimizer in $\mathcal{A}_0$ for the Mumford–Shah energy with respect to its own jump set at the point $x \in \partial\Omega$ according to Definition 4.5. Let $\mathcal{E}_x(\rho)$ be associated with u as in* (4-3).

Then there exist $r_\Omega > 0$ and $k_\Omega > 0$ such that the quantity

$$\widetilde{E}_x(\rho) := \frac{\mathcal{E}_x(\rho)}{\rho} \wedge 1 + \frac{c_\gamma}{\gamma} \rho^\gamma + k_\Omega \int_0^\rho \left(\frac{\mathcal{E}_x(r)}{r} \wedge 1 \right) dr \tag{4-8}$$

is nondecreasing on $(0, \rho_x \wedge r_\Omega \wedge \operatorname{dist}(x, S \setminus \{x\}))$, where S denotes the set of corners of $\partial\Omega$.

In order to establish these results, we need to revisit the proof of the monotonicity formula given in [Bucur and Luckhaus 2014]. The new ingredients here are the jump constraint (for which the two-dimensional setting is crucial) and the fact that the point x can belong to the boundary. We start with the boundary case, which contains the relevant modifications with respect to the case treated in [Bucur and Luckhaus 2014], and then we go back to the (simpler) interior case.

Proof of Theorems 4.4 and 4.6. Let us consider first the case of Theorem 4.4. Following [Bucur and Luckhaus 2014], to prove the result it is enough to show that there exist $r_\Omega > 0$ and $k_\Omega > 0$ such that, for $\rho \in (0, \rho_x \wedge r_\Omega \wedge \operatorname{dist}(x, S \setminus \{x\}))$, it holds that $\widetilde{E}'_x(\rho) \geq 0$ at almost every differentiability point ρ of $\widetilde{E}_x$

such that $\mathcal{E}_x(\rho) < \rho$. Let ρ_Ω and c_Ω be as in Proposition 3.7, and let us choose $k_\Omega = 2c_\Omega$ and $r_\Omega < \rho_\Omega$. We argue by contradiction. Assume that

$$\widetilde{E}'_x(\rho) = \frac{\mathcal{E}'_x(\rho)}{\rho} - \frac{\mathcal{E}_x(\rho)}{\rho^2} + c_\gamma \rho^{\gamma-1} + k_\Omega \frac{\mathcal{E}_x(\rho)}{\rho} < 0,$$

so that

$$\rho\, \mathcal{E}'_x(\rho) + c_\gamma \rho^{\gamma+1} + k_\Omega \rho \mathcal{E}_x(\rho) < \mathcal{E}_x(\rho). \tag{4-9}$$

In particular, we infer $\mathcal{E}'_x(\rho) < 1$ so that from

$$\int_{\Omega\cap\partial B_\rho(x)} |\nabla u|^2\, dx + \mathcal{H}^0(J_u \cap \partial B_\rho(x)) \le \mathcal{E}'_x(\rho) < 1$$

we get $J_u \cap \partial B_\rho(x) = \varnothing$. This means that the restriction of the SBV function u on $\partial B_\rho(x) \cap \Omega$ is a Sobolev function which we denote by w. Notice that (4-9) gives

$$\rho \int_{\Omega\cap\partial B_\rho(x)} |\nabla_\tau w|^2\, d\mathcal{H}^1 + c_\gamma \rho^{\gamma+1} + k_\Omega \rho \mathcal{E}_x(\rho) < \mathcal{E}_x(\rho). \tag{4-10}$$

Let h_w denote a harmonic extension h_w of w to $\Omega \cap B_\rho(x)$ which satisfies Neumann conditions on $\partial\Omega \cap B_\rho(x)$. By Proposition 3.7, we have

$$\frac{(1 - c_\Omega \rho)}{\rho} \int_{\Omega\cap B_\rho(x)} |\nabla h_w|^2\, dx \le \int_{\Omega\cap\partial B_\rho(x)} |\nabla_\tau w|^2\, d\mathcal{H}^1.$$

From (4-10) we deduce that

$$(1 - c_\Omega \rho) \int_{\Omega\cap B_\rho(x)} |\nabla h_w|^2\, dx + c_\gamma \rho^{\gamma+1} + k_\Omega \rho \mathcal{E}_x(\rho) < \mathcal{E}_x(\rho).$$

Up to reducing r_Ω if necessary, we infer in particular

$$\int_{\Omega\cap B_\rho(x)} |\nabla h_w|^2\, dx \le 2\mathcal{E}_x(\rho)$$

so that for every $\rho < r_\Omega \wedge \operatorname{dist}(x, S \setminus \{x\})$

$$\int_{\Omega\cap B_\rho(x)} |\nabla h_w|^2\, dx + c_\gamma \rho^{\gamma+1} + (k_\Omega - 2c_\Omega)\rho \mathcal{E}_x(\rho) < \mathcal{E}_x(\rho). \tag{4-11}$$

We now consider the admissible competitor for u given by

$$v(y) := \begin{cases} h_w(y) & \text{if } y \in \Omega \cap B_\rho(x), \\ u(y) & \text{otherwise.} \end{cases} \tag{4-12}$$

Notice that $J_v \subseteq J_u$, since v has no jumps inside $\Omega \cap B_\rho(x)$ and coincides with u on $\Omega \cap \partial B_\rho(x)$. Then, recalling that we have chosen $k_\Omega = 2c_\Omega$, inequality (4-11) for $\rho < \rho_x \wedge r_\Omega \wedge \operatorname{dist}(x, S \setminus \{x\})$ contradicts the weak local almost-quasiminimality property of u according to Definitions 4.1 (recall that $\mathcal{E}_x(\rho) < \rho$).

Coming to Theorem 4.6, we can follow the previous arguments by considering the harmonic extension h_w which satisfies homogeneous Dirichlet conditions on $\Gamma_D \cap B_\rho(x)$ and Neumann conditions on $\Gamma_N \cap B_\rho(x)$, and using again Proposition 3.7. The contradiction then follows by noting that the competitor v in (4-12) belongs to $\mathcal{A}_0$. □

Remark 4.7. Notice that the uniformity property of Remark 3.8 holds also for the constants r_Ω and k_Ω.

Proof of Theorem 4.3. The proof reduces essentially to the original case of [Bucur and Luckhaus 2014] by noting that, since we work in dimension 2, the key competitors involved in the arguments turn out to have a jump set contained in J_u. More precisely, we can follow the proof of Theorem 4.4, by using the estimate of Proposition 3.1 in place of that of Proposition 3.7: in this way we can choose $k_\Omega = 0$ and obtain the monotonicity for the energy in the simpler form (4-4). □

5. The local minimality result

Let $\Omega \subseteq \mathbb{R}^2$ be a bounded open set, and let

$$\mathrm{MS}(u) := \int_\Omega |\nabla u|^2\,dx + \alpha\mathcal{H}^1(J_u) + \beta\int_\Omega |u-g|^2\,dx$$

denote the Mumford–Shah functional on $\mathrm{SBV}(\Omega)$, where $g \in L^\infty(\Omega)$ and $\alpha, \beta > 0$.

The following result holds true.

Theorem 5.1 (small jump sets are not convenient). *Let $\Omega \subseteq \mathbb{R}^2$ be an open, bounded, piecewise $C^{1,1}$-domain with convex corners and let $g \in L^\infty(\Omega)$. Let $U \in H^1(\Omega)$ be the solution to*

$$\min_{v\in H^1(\Omega)} \int_\Omega |\nabla v|^2\,dx + \beta\int_\Omega |v-g|^2\,dx,$$

i.e., such that

$$\begin{cases} -\Delta U + \beta U = \beta g & \text{in } \Omega, \\ \partial U/\partial \nu = 0 & \text{on } \partial\Omega. \end{cases} \tag{5-1}$$

Then there exists $\varepsilon > 0$ such that for every $u \in \mathrm{SBV}(\Omega)$ with $\mathcal{H}^1(J_u) < \varepsilon$ we have

$$\mathrm{MS}(U) < \mathrm{MS}(u).$$

The main result of the paper, already stated in the Introduction as Theorem 1.1, is a simple consequence of the previous theorem. For convenience, we restate it hereafter:

Theorem 5.2 (local minimality in L^1). *Under the assumptions of Theorem 5.1, the function U is a local minimizer for the Mumford–Shah functional in $\mathrm{SBV}(\Omega)$ with respect to the L^1-topology.*

Proof. Assume by contradiction that there exists $v_n \in \mathrm{SBV}(\Omega)$ such that

$$v_n \to U \quad \text{strongly in } L^1(\Omega),$$

with

$$\mathrm{MS}(v_n) < \mathrm{MS}(U). \tag{5-2}$$

By truncation we may assume that $\|v_n\|_\infty \le \|g\|_\infty$, so that the convergence holds also in $L^2(\Omega)$. By (5-2), we may apply Ambrosio's theorem (see Theorem 2.1) and deduce

$$\int_\Omega |\nabla U|^2\,dx \le \liminf_n \int_\Omega |\nabla v_n|^2\,dx,$$

so that, since the fidelity terms are converging, we infer $\lim_n \mathcal{H}^1(J_{v_n}) = 0$: but then (5-2) is in contradiction with Theorem 5.1. □

Remark 5.3. Under suitable regularity assumptions on g, it has been proved in [Alberti et al. 2003, Sections 5.1 and 5.3] by using the *calibration method* that U is a *global minimizer* for MS if β is sufficiently small or if β is sufficiently large.

The proof of Theorem 5.1 rests on a suitable use of the monotonicity formulas up to the boundary developed in Section 4.

Proof of Theorem 5.1. First of all, by considering the change of variable $x \mapsto \sqrt{\alpha}x$, it is not restrictive to assume $\alpha = 1$.

Let $u_\varepsilon \in \mathrm{SBV}(\Omega)$ be a minimizer of

$$\min_{\substack{u\in \mathrm{SBV}(\Omega)\\ \mathcal{H}^1(J_u)\le\varepsilon}} \mathrm{MS}(u).$$

We shall prove by a contradiction argument that $J_{u_\varepsilon} = \varnothing$ for ε small enough, so that $u_\varepsilon = U$ and the proof follows.

Let ε_n be an infinitesimal sequence, and let us denote by u_n the corresponding functions u_{ε_n}. By truncation we may assume that

$$\|u_n\|_\infty \le \|g\|_\infty \quad \text{and} \quad \|U\|_\infty \le \|g\|_\infty. \tag{5-3}$$

Comparing u_n with the zero function we get

$$\int_\Omega |\nabla u_n|^2\,dx + \mathcal{H}^1(J_{u_n}) + \beta\int_\Omega |u_n - g|^2\,dx \le \beta\int_\Omega g^2\,dx. \tag{5-4}$$

We will concentrate on

$$v_n := u_n - U.$$

We divide the proof into several steps.

Step 1: regularity for U. In view of the bound (5-3) and of the regularity of $\partial\Omega$, we infer from the elliptic problem (5-1) satisfied by U that $U \in H^2(\Omega)$; see [Grisvard 1985, Theorem 3.2.1.3 and Remark 3.2.4.6]. In particular we deduce that $\nabla U \in L^p(\Omega)$ for every $p > 1$. Then we may write for every $x \in \overline{\Omega}$, $\rho > 0$ and $p > 4$

$$\int_{\Omega\cap B_\rho(x)} |\nabla U|^2\,dx \le \left(\int_{\Omega\cap B_\rho(x)} |\nabla U|^p\,dx\right)^{2/p} |\Omega\cap B_\rho(x)|^{1-2/p}.$$

As a consequence, for every $\gamma \in (0, 1)$ we get the estimate

$$\int_{\Omega\cap B_\rho(x)} |\nabla U|^2\,dx \le c_1\rho^{1+\gamma} \tag{5-5}$$

for some $c_1 > 0$ (here, c_1 depends on γ, Ω, β, g, but not on ρ).

Step 2: weak local almost-quasiminimality of v_n with respect to its own jump set. Let us show that there exist $\delta > 0$ and $c_\delta > 0$ such that for every $x \in \overline{\Omega}$, for every $\rho \leq 1$ with

$$\int_{\Omega \cap B_\rho(x)} |\nabla v_n|^2 \, dx + \mathcal{H}^1(J_{v_n} \cap \overline{B}_\rho(x)) \leq \rho, \tag{5-6}$$

and for every $v \in \mathrm{SBV}(\Omega)$ with $\{v \neq v_n\} \subseteq B_\rho(x)$ and $\mathcal{H}^1(J_v) \leq \varepsilon_n$ we have

$$\int_{\Omega \cap B_\rho(x)} |\nabla v_n|^2 \, dx + \mathcal{H}^1(J_{v_n} \cap \overline{B}_\rho(x)) \leq \int_{\Omega \cap B_\rho(x)} |\nabla v|^2 \, dx + \mathcal{H}^1(J_v \cap \overline{B}_\rho(x)) + c_\delta \rho^{1+\delta}. \tag{5-7}$$

In particular we get that v_n is a weak local almost-quasiminimizer of the Mumford–Shah energy with respect to its own jump set at any point $x \in \overline{\Omega}$ according to Definition 4.1.

Recalling (5-3) we have $\|v_n\|_\infty \leq 2\|g\|_\infty$, so that it is not restrictive to assume that also

$$\|v\|_\infty \leq 2\|g\|_\infty. \tag{5-8}$$

Moreover, we may assume

$$\int_{\Omega \cap B_\rho(x)} |\nabla v|^2 \, dx < \int_{\Omega \cap B_\rho(x)} |\nabla v_n|^2 \, dx + \mathcal{H}^1(J_{v_n} \cap \overline{B}_\rho(x)), \tag{5-9}$$

since otherwise (5-7) is immediately satisfied.

Since

$$\{v + U \neq u_n\} \subseteq B_\rho(x)$$

by the minimality of u_n we get

$$\int_{\Omega \cap B_\rho(x)} |\nabla v_n + \nabla U|^2 \, dx + \mathcal{H}^1(J_{v_n} \cap \overline{B}_\rho(x)) + \beta \int_{\Omega \cap B_\rho(x)} |u_n + U - g|^2 \, dx$$
$$\leq \int_{\Omega \cap B_\rho(x)} |\nabla v + \nabla U|^2 \, dx + \mathcal{H}^1(J_v \cap \overline{B}_\rho(x)) + \beta \int_{\Omega \cap B_\rho(x)} |v + U - g|^2 \, dx$$

so that

$$\int_{\Omega \cap B_\rho(x)} |\nabla v_n|^2 \, dx + \mathcal{H}^1(J_{v_n} \cap \overline{B}_\rho(x)) \leq \int_{\Omega \cap B_\rho(x)} |\nabla v|^2 \, dx + \mathcal{H}^1(J_v \cap \overline{B}_\rho(x)) + c(\rho),$$

where

$$c(\rho) := 2 \int_{\Omega \cap B_\rho(x)} \nabla U \cdot (\nabla v - \nabla v_n) \, dx + \beta \int_{\Omega \cap B_\rho(x)} |v + U - g|^2 \, dx - \beta \int_{\Omega \cap B_\rho(x)} |v_n + U - g|^2 \, dx.$$

Recalling (5-3) and (5-8), we have

$$\int_{\Omega \cap B_\rho(x)} |v + U - g|^2 \, dx - \int_{\Omega \cap B_\rho(x)} |v_n + U - g|^2 \, dx \leq c_2 \rho^2 \tag{5-10}$$

for some $c_2 > 0$. Thanks to the estimate (5-5) for ∇U obtained in Step 1, and taking into account (5-9) and (5-6), we have

$$\left| \int_{\Omega \cap B_\rho(x)} \nabla U \cdot (\nabla v - \nabla v_n) \, dx \right| \leq \|\nabla U\|_{L^2(\Omega \cap B_\rho(x);\mathbb{R}^2)} (\|\nabla v\|_{L^2(\Omega \cap B_\rho(x);\mathbb{R}^2)} + \|\nabla v_n\|_{L^2(\Omega \cap B_\rho(x);\mathbb{R}^2)})$$
$$\leq c_3 \rho^{(1+\gamma)/2} \rho^{1/2} = c_3 \rho^{1+\gamma/2} \tag{5-11}$$

for some $c_3 > 0$.

Collecting (5-10) and (5-11), we get for every $\rho \le 1$

$$c(\rho) \le c_\delta \rho^{1+\delta}$$

for $\delta = \gamma/2$ and $c_\delta > 0$, so that inequality (5-7) is proved.

Step 3: vanishing energy. We claim that

$$\int_\Omega |\nabla v_n|^2\,dx + \mathcal{H}^1(J_{v_n}) \to 0. \tag{5-12}$$

Indeed, in view of (5-4) and of (5-3) we may apply Ambrosio's theorem (see Theorem 2.1) to the sequence $(u_n)_{n\in\mathbb{N}}$: there exists $u \in \mathrm{SBV}(\Omega)$ such that up to a subsequence

$$\begin{aligned} u_n &\to u \quad \text{strongly in } L^2(\Omega),\\ \nabla u_n &\rightharpoonup \nabla u \quad \text{weakly in } L^2(\Omega;\mathbb{R}^2) \end{aligned} \tag{5-13}$$

and

$$\mathcal{H}^1(J_u) \le \liminf_n \mathcal{H}^1(J_{u_n}) = 0. \tag{5-14}$$

In particular, $u \in H^1(\Omega)$. Moreover, by the minimality of u_n, we deduce that in the limit

$$\int_\Omega |\nabla u|^2\,dx + \beta\int_\Omega |u-g|^2\,dx \le \int_\Omega |\nabla \varphi|^2\,dx + \beta\int_\Omega |\varphi - g|^2\,dx$$

for every $\varphi \in H^1(\Omega)$, which yields $u = U$. Passing to the limit in the inequality

$$\int_\Omega |\nabla u_n|^2\,dx + \mathcal{H}^1(J_{u_n}) + \beta\int_\Omega |u_n - g|^2\,dx \le \int_\Omega |\nabla U|^2\,dx + \beta\int_\Omega |U-g|^2\,dx,$$

we deduce

$$\limsup_n \int_\Omega |\nabla u_n|^2\,dx \le \int_\Omega |\nabla U|^2\,dx,$$

which together with the weak convergence (5-13) yields

$$\nabla u_n \to \nabla U \quad \text{strongly in } L^2(\Omega;\mathbb{R}^2).$$

Recalling that $v_n = u_n - U$ and that $\mathcal{H}^1(J_{v_n}) = \mathcal{H}^1(J_{u_n}) = \varepsilon_n$, claim (5-12) follows.

Step 4: conclusion. We can now conclude the proof via a contradiction argument. Assume that $\mathcal{H}^1(J_{u_n}) > 0$ for every n. Since $J_{v_n} = J_{u_n}$, this implies that for every n

$$\mathcal{H}^1(J_{v_n}) > 0. \tag{5-15}$$

To derive a contradiction from (5-15) when n is large enough, we consider for every n a point $x_n \in J_{v_n}$ of density 1, i.e., such that

$$\lim_{\rho\to 0^+} \frac{\mathcal{H}^1(J_{v_n} \cap B_\rho(x_n))}{2\rho} = 1. \tag{5-16}$$

Let $y_n \in \partial\Omega$ be a projection of x_n on $\partial\Omega$; since Ω has convex corners, y_n is a smooth point of $\partial\Omega$. Let us set

$$d_n := \operatorname{dist}(x_n, \partial\Omega) = |x_n - y_n| \quad \text{and} \quad d_n' := \operatorname{dist}(y_n, S),$$

where S is the set of corners of Ω.

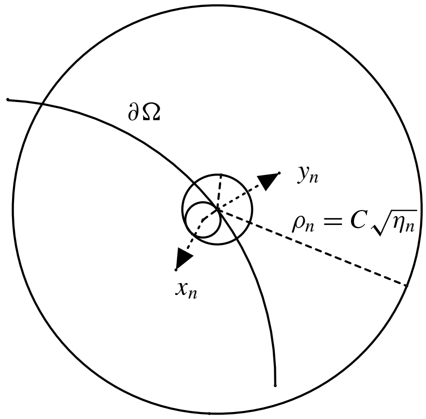

Figure 3. Illustration of Case 2.

For $x \in \overline{\Omega}$, we set

$$\mathcal{E}_x^n(\rho) := \int_{\Omega\cap B_\rho(x)} |\nabla v_n|^2\,dx + \mathcal{H}^1(J_{v_n} \cap \overline{B}_\rho(x)).$$

Notice that, in view of Step 3,

$$\eta_n := \int_\Omega |\nabla v_n|^2\,dx + \mathcal{H}^1(J_{v_n}) \to 0 \tag{5-17}$$

and

$$\mathcal{E}_x^n(\rho) \le \eta_n. \tag{5-18}$$

We distinguish four cases.

Case 1: Assume that

$$\limsup_n \frac{d_n}{\sqrt{\eta_n}} > 0. \tag{5-19}$$

Thanks to Step 2 and the interior monotonicity formula of Theorem 4.3 we infer that the map

$$\rho \mapsto \frac{\mathcal{E}_{x_n}^n(\rho)}{\rho} \wedge 1 + c_\gamma \rho^\gamma \tag{5-20}$$

is nondecreasing on $(0, d_n \wedge 1)$. Thanks to (5-19), possibly passing to a subsequence, for n large we may choose as an admissible radius $\rho_n = C\sqrt{\eta_n}$ for some $C > 0$, and write

$$1 \le \frac{\mathcal{E}_{x_n}^n(\rho_n)}{\rho_n} \wedge 1 + c_\gamma \rho_n^\gamma \le \frac{\mathcal{E}_{x_n}^n(\rho_n)}{\rho_n} + c_\gamma \rho_n^\gamma \le \frac{\eta_n}{C\sqrt{\eta_n}} + c_\gamma (C\sqrt{\eta_n})^\gamma,$$

where the first inequality comes from monotonicity at x_n and (5-16), and the last one by (5-18). In view of (5-17), the above relation gives a contradiction for n large enough.

Case 2: Assume that

$$\lim_n \frac{d_n}{\sqrt{\eta_n}} = 0 \quad \text{and} \quad \limsup_n \frac{d_n'}{\sqrt{\eta_n}} > 0. \tag{5-21}$$

By Step 2 and the boundary monotonicity formula of Theorem 4.4 we infer that the map

$$\rho \mapsto \frac{\mathcal{E}^n_{y_n}(\rho)}{\rho} \wedge 1 + c_\gamma \rho^\gamma + k_\Omega \int_0^\rho \left(\frac{\mathcal{E}^n_{y_n}(r)}{r} \wedge 1 \right) dr \tag{5-22}$$

is nondecreasing on $(0, 1 \wedge r_\Omega \wedge d'_n)$, where $r_\Omega, k_\Omega > 0$. Thanks to (5-21), possibly passing to a subsequence, for n large we may choose as an admissible radius $\rho_n = C\sqrt{\eta_n}$, assume that $\rho_n > 2d_n$, and write

$$\begin{aligned}\frac{\mathcal{E}^n_{y_n}(\rho_n)}{\rho_n} \wedge 1 + c_\gamma \rho_n^\gamma + k_\Omega \int_0^{\rho_n} \left(\frac{\mathcal{E}^n_{y_n}(r)}{r} \wedge 1 \right) dr &\ge \frac{\mathcal{E}^n_{y_n}(2d_n)}{2d_n} \wedge 1 + c_\gamma (2d_n)^\gamma + k_\Omega \int_0^{2d_n} \left(\frac{\mathcal{E}^n_{y_n}(r)}{r} \wedge 1 \right) dr \\ &\ge \frac{\mathcal{E}^n_{x_n}(d_n)}{2d_n} \wedge 1 + c_\gamma d_n^\gamma \ge \frac{1}{2}\left[\frac{\mathcal{E}^n_{x_n}(d_n)}{d_n} \wedge 1 + c_\gamma d_n^\gamma \right] \ge \frac{1}{2},\end{aligned}$$

where we have used in the first inequality monotonicity at y_n, then the inclusion $B_{d_n}(x_n) \subseteq B_{2d_n}(y_n)$ (see Figure 3), and finally monotonicity at x_n combined with (5-16). We infer

$$\frac{1}{2} \le \frac{\mathcal{E}^n_{y_n}(\rho_n)}{\rho_n} \wedge 1 + c_\gamma \rho_n^\gamma + k_\Omega \int_0^{\rho_n} \left(\frac{\mathcal{E}^n_{y_n}(r)}{r} \wedge 1 \right) dr \le \frac{\eta_n}{C\sqrt{\eta_n}} + c_\gamma (C\sqrt{\eta_n})^\gamma + k_\Omega C\sqrt{\eta_n}, \tag{5-23}$$

which is a contradiction for n large in view of (5-17).

<u>Case 3</u>: Assume that

$$\lim_n \frac{d_n}{\sqrt{\eta_n}} = 0, \quad \lim_n \frac{d'_n}{\sqrt{\eta_n}} = 0 \quad \text{and} \quad \lim_n \frac{d_n}{d'_n} = 0. \tag{5-24}$$

Then there exists a vertex $z \in S$ such that, possibly passing to a subsequence, $x_n, y_n \to z$. Set

$$d'_n = \operatorname{dist}(y_n, z).$$

We choose $\rho_n := \sqrt{\eta_n}$. Since for n large we have $\rho_n \ge 2d'_n$, we may write using the monotonicity at z and the inclusion $B_{d'_n}(y_n) \subseteq B_{2d'_n}(z)$ (see Figure 4)

$$\begin{aligned}\frac{\mathcal{E}^n_{z}(\rho_n)}{\rho_n} \wedge 1 + c_\gamma \rho_n^\gamma + k_\Omega \int_0^{\rho_n} \left(\frac{\mathcal{E}^n_{z}(r)}{r} \wedge 1 \right) dr &\ge \frac{\mathcal{E}^n_{z}(2d'_n)}{2d'_n} \wedge 1 + c_\gamma (2d'_n)^\gamma + k_\Omega \int_0^{2d'_n} \left(\frac{\mathcal{E}^n_{z}(r)}{r} \wedge 1 \right) dr \\ &\ge \frac{\mathcal{E}^n_{y_n}(d'_n)}{2d'_n} \wedge 1 + c_\gamma (d'_n)^\gamma.\end{aligned}$$

On the other hand, since for n large we also have $d'_n \ge 2d_n$, by monotonicity at y_n and the inclusion $B_{d_n}(x_n) \subseteq B_{2d_n}(y_n)$ (see again Figure 4), we have

$$\begin{aligned}\frac{\mathcal{E}^n_{y_n}(d'_n)}{2d'_n} \wedge 1 + c_\gamma (d'_n)^\gamma + k_\Omega \int_0^{d'_n} \left(\frac{\mathcal{E}^n_{y_n}(r)}{r} \wedge 1 \right) dr &\ge \frac{1}{2}\left[\frac{\mathcal{E}^n_{y_n}(d'_n)}{d'_n} \wedge 1 + c_\gamma (d'_n)^\gamma + k_\Omega \int_0^{d'_n} \left(\frac{\mathcal{E}^n_{y_n}(r)}{r} \wedge 1 \right) dr \right] \\ &\ge \frac{1}{2}\left[\frac{\mathcal{E}^n_{y_n}(2d_n)}{2d_n} \wedge 1 + c_\gamma (2d_n)^\gamma + k_\Omega \int_0^{2d_n} \left(\frac{\mathcal{E}^n_{y_n}(r)}{r} \wedge 1 \right) dr \right] \\ &\ge \frac{1}{2}\left[\frac{\mathcal{E}^n_{x_n}(d_n)}{2d_n} \wedge 1 + c_\gamma d_n^\gamma \right] \ge \frac{1}{4}\left[\frac{\mathcal{E}^n_{x_n}(d_n)}{d_n} \wedge 1 + c_\gamma d_n^\gamma \right] \ge \frac{1}{4},\end{aligned}$$

the last inequality coming from monotonicity at x_n combined with (5-16).

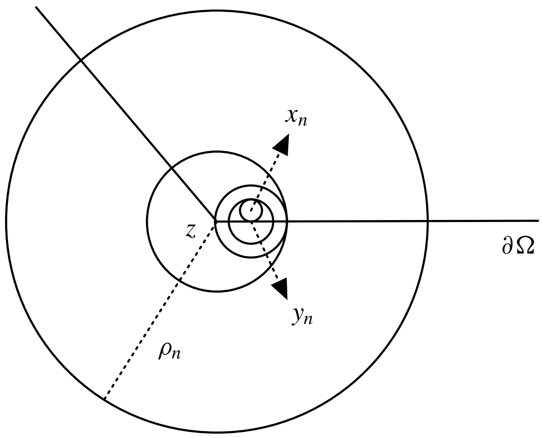

Figure 4. Illustration of Case 3.

Collecting the previous inequalities we obtain

$$\begin{aligned}\frac{1}{4} &\le k_\Omega \int_0^{d'_n} \left(\frac{\mathcal{E}^n_{y_n}(r)}{r} \wedge 1 \right) dr + \frac{\mathcal{E}^n_z(\rho_n)}{\rho_n} \wedge 1 + c_\gamma \rho_n^\gamma + k_\Omega \int_0^{\rho_n} \left(\frac{\mathcal{E}^n_z(r)}{r} \wedge 1 \right) dr \\ &\le k_\Omega d'_n + \frac{\eta_n}{\sqrt{\eta_n}} + c_\gamma (\sqrt{\eta_n})^\gamma + k_\Omega \sqrt{\eta_n},\end{aligned}$$

which yields a contradiction for n large.

Case 4: Assume finally that

$$\lim_n \frac{d_n}{\sqrt{\eta_n}} = 0, \quad \lim_n \frac{d'_n}{\sqrt{\eta_n}} = 0 \quad \text{and} \quad \limsup_n \frac{d_n}{d'_n} > 0. \tag{5-25}$$

Let $z \in S$ be a vertex such that, up to a subsequence if necessary, $x_n, y_n \to z$, so that $d'_n = \operatorname{dist}(y_n, z)$. Set $d''_n := \operatorname{dist}(x_n, z)$. Up to a subsequence we may assume that, for n large, $d'_n \le C d_n$ with $C > 0$, so that from the inequalities

$$d_n = \operatorname{dist}(x_n, \partial\Omega) \le d''_n = \operatorname{dist}(x_n, z) \le \operatorname{dist}(x_n, y_n) + \operatorname{dist}(y_n, z) = d_n + d'_n \le (1 + C) d_n \tag{5-26}$$

we infer that d_n and d''_n are comparable.

Then, if we choose $\rho_n := \sqrt{\eta_n}$, in view of (5-25) and (5-26), for n large we have $\rho_n \ge 2d''_n$. By monotonicity at z and the inclusion $B_{d''_n}(x_n) \subseteq B_{2d''_n}(z)$ (see Figure 5), we obtain

$$\begin{aligned}\frac{\mathcal{E}^n_z(\rho_n)}{\rho_n} \wedge 1 + c_\gamma \rho_n^\gamma + k_\Omega \int_0^{\rho_n} \left(\frac{\mathcal{E}^n_z(r)}{r} \wedge 1 \right) dr &\ge \frac{\mathcal{E}^n_z(2d''_n)}{2d''_n} \wedge 1 + c_\gamma (2d''_n)^\gamma + k_\Omega \int_0^{2d''_n} \left(\frac{\mathcal{E}^n_z(r)}{r} \wedge 1 \right) dr \\ &\ge \frac{\mathcal{E}^n_{x_n}(d''_n)}{2d''_n} \wedge 1 + c_\gamma (d''_n)^\gamma \ge \frac{\mathcal{E}^n_{x_n}(d_n)}{2(1+C)d_n} \wedge 1 + c_\gamma d_n^\gamma \\ &\ge \frac{1}{2(1+C)} \left[\frac{\mathcal{E}^n_{x_n}(d_n)}{d_n} \wedge 1 + c_\gamma d_n^\gamma \right] \ge \frac{1}{2(1+C)},\end{aligned}$$

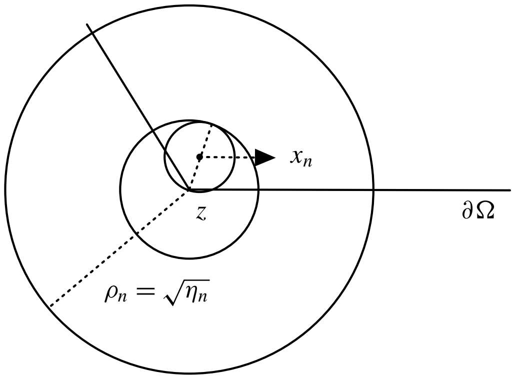

Figure 5. Illustration of Case 4.

where as above the last inequality comes from monotonicity at x_n combined with (5-16). Then,

$$\frac{1}{2(1+C)} \le \frac{\mathcal{E}_z^n(\rho_n)}{\rho_n} \wedge 1 + c_\gamma \rho_n^\gamma + k_\Omega \int_0^{\rho_n} \left(\frac{\mathcal{E}_z^n(r)}{r} \wedge 1 \right) dr \le \frac{\eta_n}{\sqrt{\eta_n}} + c_\gamma (\sqrt{\eta_n})^\gamma + k_\Omega \sqrt{\eta_n},$$

which yields a contradiction for n large. □

Remark 5.4. For later use, let us notice that in order to carry out Step 4 in the previous proof, the monotonicity properties given by Theorems 4.3 and 4.4 can be replaced respectively by *interior and boundary quasimonotonicity properties* of the following type: if u is a weak local almost-quasiminimizer for the Mumford–Shah energy with respect to its own jump set at every point of $\overline{\Omega}$, and E_x and $\widetilde{E}_x$ are defined respectively by (4-4) and (4-5), there exist $c > 1$ and $\tilde{r}_\Omega > 0$ such that

$$E_x(\rho_2) \ge \frac{1}{c} E_x\left(\frac{\rho_1}{c}\right) \quad \text{if } x \in \Omega \text{ and } \rho_1 \le \rho_2 \le \tilde{r}_\Omega \wedge \operatorname{dist}(x, \partial\Omega),$$
$$\widetilde{E}_x(\rho_2) \ge \frac{1}{c} \widetilde{E}_x\left(\frac{\rho_1}{c}\right) \quad \text{if } x \in \partial\Omega \text{ and } \rho_1 \le \rho_2 \le \tilde{r}_\Omega \wedge \operatorname{dist}(x, S \setminus \{x\}).$$

Indeed, considering for example Case 2, we can choose again $\rho_n := C\sqrt{\eta_n}$ for some $C > 0$, and use the *boundary quasimonotonicity* at y_n for the radii $\rho_n \ge 2cd_n$ to write

$$\begin{aligned}\frac{\mathcal{E}_{y_n}^n(\rho_n)}{\rho_n} \wedge 1 + c_\gamma \rho_n^\gamma + k_\Omega \int_0^{\rho_n} \left(\frac{\mathcal{E}_{y_n}^n(r)}{r} \wedge 1 \right) dr &\ge \frac{1}{c}\left[\frac{\mathcal{E}_{y_n}^n(2d_n)}{2d_n} \wedge 1 + c_\gamma (2d_n)^\gamma + k_\Omega \int_0^{2d_n} \left(\frac{\mathcal{E}_{y_n}^n(r)}{r} \wedge 1 \right) dr \right] \\ &\ge \frac{1}{2c}\left[\frac{\mathcal{E}_{x_n}^n(d_n)}{d_n} \wedge 1 + c_\gamma d_n^\gamma \right]. \end{aligned} \tag{5-27}$$

Since x_n is a point of density 1 for J_{v_n}, there exists r_n such that $cr_n \le d_n$ and

$$\frac{\mathcal{H}^1(J_{v_n} \cap \overline{B}_{r_n}(x_n))}{r_n} \ge \frac{3}{2}.$$

Then, by using the *interior quasimonotonicity* at x_n for the radii $cr_n \leq d_n$, we get

$$\frac{1}{2c}\left[\frac{\mathcal{E}^n_{x_n}(d_n)}{d_n} \wedge 1 + c_\gamma d_n^\gamma\right] \geq \frac{1}{2c^2}\left[\frac{\mathcal{E}^n_{x_n}(r_n)}{r_n} \wedge 1 + c_\gamma r_n^\gamma\right] \geq \frac{1}{2c^2}.$$

But this lower bound is incompatible with (5-27) since the first line of (5-27) is going to zero as $n \to \infty$ (see (5-23)). The other cases can be treated similarly.

6. The case of general energies with boundary conditions

In this section we deal with the local minimality of the Sobolev minimizer of a generalization of the Mumford–Shah functional under prescribed boundary conditions.

6A. ***Setting of the problem and the local minimality result.*** Given an open bounded set $\Omega \subset \mathbb{R}^2$, which is admissible according to Definition 3.5, we want to impose Dirichlet boundary conditions on the (nonempty) portion Γ_D of its boundary. To this aim, similarly as in Section 4 we consider an open bounded domain Ω' such that $\Omega \Subset \Omega' \subset \mathbb{R}^2$ and with $\partial\Omega \cap \Omega' = \Gamma_D$ (see Figure 2). Then, given $w \in H^1(\Omega') \cap L^\infty(\Omega')$, we set

$$\mathcal{A}_w := \{u \in \mathrm{SBV}(\Omega') : u = w \text{ on } \Omega' \setminus \overline{\Omega}\}. \tag{6-1}$$

We are interested in the minimization on the class $\mathcal{A}_w$ of the Mumford–Shah-type functional

$$F(u) := \int_{\Omega'} A(x)\nabla u \cdot \nabla u \, dx + \int_{J_u} b(x, u^+, u^-, \nu_u)\, d\mathcal{H}^1 + \beta \int_{\Omega'} |u - g|^2 \, dx. \tag{6-2}$$

Here ν_u denotes a normal along J_u, and $u^\pm$ the associated traces of u (see Section 2), while the assumptions satisfied by the matrix A, the function b, and the datum g are specified below.

Remark 6.1. Notice that working on the set $\mathcal{A}_w$ the boundary condition

$$u = w \quad \text{on } \Gamma_D$$

is taken into account in a relaxed sense: indeed, the parts of Γ_D on which $u \neq w$ are contained in J_u, so that they turn out to be penalized by the functional F. This is usual in variational problems for functions of bounded variation (like for example the graph area problem). Finally, observe that the bulk terms provide a fixed contribution on $\Omega' \setminus \Omega$, since $u = w$ on this set, so that for the minimization it suffices simply to integrate on Ω.

Assumptions on the functional. Concerning the volume terms in (6-2), we require

$$A \in C^{0,\delta}(\overline{\Omega}'; M^{2\times 2}_{\mathrm{sym}}), \qquad c_1^A|\eta|^2 \leq A(x)\eta \cdot \eta \leq c_2^A|\eta|^2 \quad \text{for every } x \in \Omega' \text{ and } \eta \in \mathbb{R}^2, \tag{6-3}$$

and

$$g \in L^\infty(\Omega'), \quad \beta \geq 0 \tag{6-4}$$

for suitable constants $c_1^A, c_2^A > 0$ and $\delta \in (0, 1)$.

Concerning the surface term

$$\Psi(u) := \int_{J_u} b(x, u^+, u^-, \nu_u)\, d\mathcal{H}^1,$$

we require for $b : \Omega' \times \mathbb{R} \times \mathbb{R} \times S^1 \to [0, +\infty[$ to be such that

$$c_1^b \le b(x, s_1, s_2, \nu) \le c_2^b + \Phi(s_1, s_2), \tag{6-5}$$

where $c_1^b, c_2^b > 0$ and $\Phi : \mathbb{R}^2 \to [0, +\infty[$. In addition we ask that

$$u \mapsto \Psi(u) \text{ is l.s.c with respect to the weak convergence in } \mathrm{SBV}(\Omega), \tag{6-6}$$

and the *monotonicity under truncation* satisfies

$$\Psi((u \wedge c_2) \vee c_1) \le \Psi(u) \quad \text{for every } c_1 \le c_2. \tag{6-7}$$

Remark 6.2 (example of admissible surface energies). An admissible surface term could be given for example by

$$b(x, s_1, s_2, \nu) := \varphi(x, \nu) + |s_1 - s_2|,$$

where $\varphi : \Omega' \times \mathbb{R}^2 \to]0, +\infty[$ is such that

$$x \mapsto \varphi(x, \nu) \quad \text{is lower semicontinuous on } \Omega' \text{ for every } \nu \in S^1,$$

$$\nu \mapsto \varphi(x, \nu) \quad \text{is convex, positively one homogeneous on } \mathbb{R}^2 \text{ for every } x \in \Omega',$$

and

$$0 < c_1^b \le \varphi(x, \nu) \le c_2^b.$$

Requirements (6-5) and (6-7) are immediately fulfilled. The lower semicontinuity (6-6) is a consequence of the Reshetnyak theorem (see [Ambrosio et al. 2000, Theorem 2.38]) and of the lower semicontinuity result in SBV [Ambrosio et al. 2000, Theorem 5.22].

A-admissible domains. To control the interaction between Neumann and Dirichlet conditions in connection with monotonicity, we need to introduce the following property concerning the interplay between the geometry of Ω and the bulk energy.

Definition 6.3 (A-admissible domains). Let $\Omega \subset \mathbb{R}^2$ be an open bounded set with a piecewise $C^{1,1}$-boundary, decomposed as in (3-23), and let A satisfy (6-3). We say that Ω is A-admissible if, for every $x \in \overline{\Omega}$, the domain $\Omega_x := A(x)^{-1/2}\Omega$ is admissible according to Definition 3.5.

Remark 6.4. Notice that admissibility according to Definition 3.5 is equivalent to Id-admissibility according to Definition 6.3. In particular, when $\Gamma_D = \partial\Omega$ or $\Gamma_N = \partial\Omega$, A-admissibility reduces to the assumption of convex corners, and smooth $C^{1,1}$ domains are always A-admissible.

The local minimality result under boundary conditions. The analogue of Theorem 5.1 under boundary conditions is the following.

Theorem 6.5 (small jump sets are not convenient under boundary conditions). *Assume the functional F satisfies assumptions* (6-3)–(6-7). *Let $\Omega \subseteq \mathbb{R}^2$ be A-admissible according to Definition 6.3, and let $w \in H^1(\Omega') \cap L^\infty(\Omega')$. Denoting by $U_w \in H^1(\Omega')$ the solution to*

$$\min_{v \in \mathcal{A}_w \cap H^1(\Omega')} F(v),$$

assume that U_w admits at most uniformly weak singularities in $\overline{\Omega}$; i.e., there exist $C > 0$ and $\alpha > 0$ such that for every $x \in \overline{\Omega}$ and $\rho \leq 1$

$$\int_{\Omega \cap B_\rho(x)} |\nabla U_w|^2 \, dx \leq C \rho^{1+\alpha}. \tag{6-8}$$

Then there exists $\varepsilon > 0$ such that for every $u \in \mathcal{A}_w$ with $\mathcal{H}^1(J_u) < \varepsilon$ we have

$$F(U_w) < F(u).$$

Remark 6.6. Assumptions on the kind of *singularities* which the function U_w can exhibit are essential to exclude that small jump sets are not convenient. Indeed, assume for example that

$$\int_{\Omega \cap B_\rho(x)} |\nabla U_w|^2 \, dx \approx c \rho^\gamma, \tag{6-9}$$

with $0 < \gamma < 1$ and $c > 0$ at some point $x \in \overline{\Omega}$: this means that for small ρ, the singularity has an energy content in $\Omega \cap B_\rho(x)$ which is higher than the length of $\partial(\Omega \cap B_\rho(x))$. It turns out that the admissible competitor

$$v := \begin{cases} 0 & \text{in } \Omega \cap B_\rho(x), \\ U_w & \text{otherwise} \end{cases}$$

is such that $F(v) < F(U_w)$ if ρ is sufficiently small, so that the minimality property of Theorem 6.5 cannot hold.

Theorem 6.5 states that the energy content given by (6-8) is not enough to destroy the minimality of U_w, which is preserved even if small jump sets are allowed. In the case $\Gamma_D = \partial\Omega$, inequality (6-8) holds for example if $w \in H^2(\Omega')$ and A is Lipschitz continuous on $\overline{\Omega}'$, since elliptic regularity gives $U_w \in H^2(\Omega)$; see, e.g., [Grisvard 1985, 3.2.1.2].

In the language of fracture mechanics, the situation (6-9) is defined as a *strong singularity*: the "elastic energy" stored in $\Omega \cap B_\rho(x)$ is much higher than the energy required to create a crack along $\partial(\Omega \cap B_\rho(x))$ so that the elastic configuration is not in equilibrium, and the creation of a small crack is energetically convenient. "*Weak singularities*" are on the contrary compatible with local equilibrium.

As mentioned in the Introduction, the minimality property of Theorem 6.5 (without fidelity term but with nonlinear bulk energies and general Lipschitz boundaries) was derived in [Chambolle et al. 2008, Theorem 1] under the additional assumption that the competitors have a closed jump set with a preset number of connected components, while the full SBV case was derived under the stronger assumption $U_w \in C^1(\overline{\Omega})$; see [Chambolle et al. 2008, Theorem 6].

As in Section 5, we can draw the following local minimality result with respect to the L^1 topology under boundary conditions.

Theorem 6.7 (local minimality in L^1 under boundary conditions). *Under the assumptions of Theorem 6.5, the function U_w is a local minimizer for the Mumford–Shah-type energy F on $\mathcal{A}_w$ with respect to the L^1-topology.*

Proof. Assume by contradiction that there exists $v_n \in \mathcal{A}_w$ such that

$$v_n \to U_w \quad \text{strongly in } L^1(\Omega'),$$

with

$$F(v_n) < F(U_w). \tag{6-10}$$

It is not restrictive to assume that $\|v_n\|_\infty \le \|g\|_\infty + \|w\|_\infty$, so that the convergence holds also in $L^2(\Omega')$ and is weak in $\mathrm{SBV}(\Omega')$ (thanks to the coercivity assumptions on A and b). By Ambrosio's theorem we deduce

$$\liminf_n F(v_n) \ge F(U_w),$$

so that from (6-10) we infer $\lim_n \mathcal{H}^1(J_{v_n}) = 0$; but then (6-10) is in contrast with Theorem 6.5. □

6B. ***Proof of Theorem 6.5.*** Let us start by deriving some properties of a solution $v \in \mathcal{A}_0$ to

$$\min_{\substack{h\in\mathcal{A}_0\\ \mathcal{H}^1(J_h)\le\varepsilon}} F(U_w + h). \tag{6-11}$$

They are stated in two separate lemmas below, and concern respectively a uniform almost-quasiminimality property, and a crucial quasimonotonicity property. Some preliminary notation and remarks are in order.

Firstly notice that existence of minimizers to (6-11) is guaranteed by the application of Ambrosio's theorem in view of the assumptions (6-3)–(6-7) on the terms appearing in F. Notice that we may assume (thanks in particular to the truncation assumption (6-7) for the surface energy)

$$\|U_w\|_\infty \le \|w\|_\infty + \|g\|_\infty \quad \text{and} \quad \|v\|_\infty \le 2\|w\|_\infty + \|g\|_\infty. \tag{6-12}$$

For every $\xi \in \overline{\Omega}$, let us consider the matrix

$$L_\xi := A(\xi)^{1/2}$$

and the sets

$$\widetilde{\Omega}_\xi := L_\xi^{-1}\Omega \quad \text{and} \quad \widetilde{\Omega}'_\xi := L_\xi^{-1}\Omega'.$$

To every function $u \in \mathrm{SBV}(\Omega')$ let us associate the function $\tilde{u}_\xi \in \mathrm{SBV}(\widetilde{\Omega}'_\xi)$ given by

$$\tilde{u}_\xi(y) := \sqrt{\frac{2\det L_\xi}{c_1^b\sqrt{c_1^A}}}\, u(L_\xi y), \tag{6-13}$$

where c_1^A and c_1^b are the constants appearing in (6-3) and (6-5). Let us denote by $\widetilde{\mathcal{A}}_0$ the space of functions associated to $\mathcal{A}_0$ under the previous transformation, and let $\tilde{\xi}$ be the point corresponding to ξ.

Lemma 6.8 (uniform almost-quasiminimality). *Under the assumptions of Theorem 6.5, let $v \in \mathcal{A}_0$ be a solution to the minimization problem* (6-11). *There exist $\rho_0 > 0,\ \Lambda \geq 1,\ \gamma > 0,\ c_\gamma > 0$ such that for every $\xi \in \overline{\Omega}$, for every $\rho < \rho_0$ with*

$$\int_{B_\rho(\tilde{\xi})\cap\widetilde{\Omega}_\xi} |\nabla \tilde{v}_\xi|^2\,dx + \mathcal{H}^1(J_{\tilde{v}_\xi} \cap \overline{B}_\rho(\tilde{\xi})) \leq \rho, \tag{6-14}$$

and for every $\tilde{h} \in \widetilde{\mathcal{A}}_0$ with $\{\tilde{h} \neq \tilde{v}_\xi\} \subseteq B_\rho(\tilde{x})$ and $J_{\tilde{h}} \subseteq J_{\tilde{v}_\xi}$ we have

$$\int_{B_\rho(\tilde{\xi})\cap\widetilde{\Omega}_\xi} |\nabla \tilde{v}_\xi|^2\,dx + \mathcal{H}^1(J_{\tilde{v}_\xi} \cap \overline{B}_\rho(\tilde{\xi})) \leq \int_{B_\rho(\tilde{\xi})\cap\widetilde{\Omega}_\xi} |\nabla \tilde{h}|^2\,dx + \Lambda\mathcal{H}^1(J_{\tilde{h}} \cap \overline{B}_\rho(\tilde{\xi})) + c_\gamma \rho^{1+\gamma}.$$

In other words, uniformly in $\xi \in \overline{\Omega}$, the function $\tilde{v}_\xi$ is a weak local almost-quasiminimizer in $\widetilde{\mathcal{A}}_0$ of the Mumford–Shah energy with respect to its own jump set at the point $\tilde{\xi}$.

Proof. Let us divide the proof in two steps.

Step 1: assume that $A(\xi) = \mathrm{Id}$. Let us show that if $h \in \mathcal{A}_0$ with $\{h \neq v\} \subseteq B_\rho(\xi),\ \rho \leq 1$ and $J_h \subseteq J_v$, we have

$$\int_{\Omega\cap B_\rho(\xi)} |\nabla v|^2\,dx + c_1^b\mathcal{H}^1(J_v \cap \overline{B}_\rho(\xi)) \\ \leq \int_{\Omega\cap B_\rho(\xi)} |\nabla h|^2\,dx + C_1\mathcal{H}^1(J_h \cap \overline{B}_\rho(\xi)) + 2c_A(1+c_1^b)\rho^{1+\delta} + C_2\rho^{(1+\alpha/2)\wedge 2}, \tag{6-15}$$

where c_1^b is the constant appearing in the estimates (6-5) for the surface energy b, c_A is the Holder constant of $A \in C^{0,\delta}(\overline{\Omega}', M^{2\times 2}_{\mathrm{sym}})$, α is the constant appearing in (6-8), and $C_1, C_2 > 0$ are suitable constants depending only on the data of the problem.

Indeed, recalling (6-12) and using a truncation argument, it is not restrictive to assume that

$$\|h\|_\infty \leq 2\|w\|_\infty + \|g\|_\infty. \tag{6-16}$$

Moreover, we may assume

$$\int_{\Omega\cap B_\rho(\xi)} |\nabla h|^2\,dx < \int_{\Omega\cap B_\rho(\xi)} |\nabla v|^2\,dx + c_1^b\mathcal{H}^1(J_v \cap \overline{B}_\rho(\xi)) \tag{6-17}$$

since otherwise (6-15) is immediately satisfied.

By the minimality of v (since the inclusion $J_h \subseteq J_v$ ensures that $\mathcal{H}^1(J_h) \leq \varepsilon$), we get

$$\int_{\Omega\cap B_\rho(\xi)} A(x)\nabla v\cdot\nabla v\,dx + c_1^b\mathcal{H}^1(J_v \cap \overline{B}_\rho(\xi)) \\ \leq \int_{\Omega\cap B_\rho(\xi)} A(x)\nabla h\cdot\nabla h\,dx + C_1\mathcal{H}^1(J_h \cap \overline{B}_\rho(\xi)) + c(\rho), \tag{6-18}$$

where

$$c(\rho) := 2\int_{\Omega\cap B_\rho(\xi)} A(x)\nabla U_w\cdot(\nabla h - \nabla v)\,dx + \beta\int_{\Omega\cap B_\rho(\xi)} (h-v)(2U_w + h + v - 2g)\,dx$$

and

$$C_1 := c_2^b + \Phi(2\|w\|_\infty + \|g\|_\infty, -(2\|w\|_\infty + \|g\|_\infty)),$$

where c_2^b and Φ appear in (6-5).

In view of the energy bounds (6-8), (6-14), (6-17) and of the L^∞-bounds (6-12) and (6-16) we deduce for every $\rho \le 1$

$$\begin{aligned} c(\rho) &\le 2c_2^A \int_{\Omega\cap B_\rho(\xi)} |\nabla U_w||\nabla h - \nabla v|\,dx + \widetilde{C}\rho^2 \\ &\le 2c_2^A \|\nabla U_w\|_{L^2(\Omega\cap B_\rho(\xi))}(\|\nabla h\|_{L^2(\Omega\cap B_\rho(\xi))} + \|\nabla v\|_{L^2(\Omega\cap B_\rho(\xi))}) + \widetilde{C}\rho^2 \\ &\le 2c_2^A \sqrt{C}\rho^{(1+\alpha)/2}(\sqrt{(1+c_1^b)\rho} + \sqrt{\rho}) + \widetilde{C}\rho^2 \le C_2\rho^{(1+\alpha/2)\wedge 2} \end{aligned}$$

for a suitable $C_2 > 0$.

Since $A(\xi) = \mathrm{Id}$ and $A \in C^{0,\delta}(\overline{\Omega}'; M^{2\times 2}_{\mathrm{sym}})$, coming back to (6-18) we may write

$$\begin{aligned} &\int_{\Omega\cap B_\rho(\xi)} |\nabla v|^2\,dx - c_A\rho^\delta \int_{\Omega\cap B_\rho(\xi)} |\nabla v|^2\,dx + c_1^b \mathcal{H}^1(J_v \cap \overline{B}_\rho(x)) \\ &\qquad \le \int_{\Omega\cap B_\rho(\xi)} |\nabla h|^2\,dx + c_A\rho^\delta \int_{\Omega\cap B_\rho(\xi)} |\nabla h|^2\,dx + C_1\mathcal{H}^1(J_h \cap \overline{B}_\rho(\xi)) + C_2\rho^{(1+\alpha/2)\wedge 2}, \end{aligned}$$

where $c_A > 0$ is the Hölder constant of A, so that taking into account (6-17) and (6-14) we infer

$$\begin{aligned} &\int_{\Omega\cap B_\rho(\xi)} |\nabla v|^2\,dx + c_1^b \mathcal{H}^1(J_v \cap \overline{B}_\rho(\xi)) \\ &\qquad \le \int_{\Omega\cap B_\rho(\xi)} |\nabla h|^2\,dx + C_1\mathcal{H}^1(J_h \cap \overline{B}_\rho(\xi)) + 2c_A(1+c_1^b)\rho^{1+\delta} + C_2\rho^{(1+\alpha/2)\wedge 2}, \end{aligned}$$

so that (6-15) follows.

Step 2: let us come to the general case. Under the transformation (6-13), a direct computation shows that the functional F is transformed up to a multiplicative constant into the functional on $\mathrm{SBV}(\widetilde{\Omega}'_\xi)$ given by

$$\widetilde{F}(\tilde{u}) := \int_{\widetilde{\Omega}'_\xi} \tilde{A}(y)\nabla\tilde{u}\cdot\nabla\tilde{u}\,dy + \int_{J_{\tilde{u}}} \tilde{b}(y, \tilde{u}^+, \tilde{u}^-, \nu_{\tilde{u}})\,d\mathcal{H}^1(y) + \beta \int_{\widetilde{\Omega}'_\xi} |\tilde{u} - \tilde{g}_\xi|^2\,dy,$$

with

$$\tilde{A}(y) := L_\xi^{-1} A(L_\xi y) L_\xi^{-1}$$

and

$$\tilde{b}(y, s_1, s_2, \nu) := \frac{2}{c_1^b\sqrt{c_1^A}}\, b\biggl(L_\xi y, \sqrt{\frac{2\det L_\xi}{c_1^b}}s_1, \sqrt{\frac{2\det L_\xi}{c_1^b}}s_2, L_\xi\nu\biggr)|L_\xi\nu^\perp|,$$

where $\nu^\perp$ is a unit vector orthogonal to ν (so that $|L_\xi \nu^\perp_{\tilde{u}}|$ turns out to be the one-dimensional jacobian of the transformation $y \mapsto L_\xi y$ involved in the change of variable to pass from an integral on J_u to that on $J_{\tilde{u}}$, see [Ambrosio et al. 2000, Theorem 2.91]). Notice that by construction

$$\tilde{b}(y, s_1, s_2, \nu) \ge 2. \tag{6-19}$$

Clearly the function

$$\tilde{u} := \widetilde{(U_w)}_\xi + \tilde{v}_\xi$$

is a minimizer of the functional $\widetilde{F}$ among the functions in $\mathrm{SBV}(\widetilde{\Omega}'_\xi)$ such that $\tilde{u} = \tilde{w}_\xi$ on $\widetilde{\Omega}'_\xi \setminus \widetilde{\Omega}_\xi$ and $\mathcal{H}^1(L_\xi J_{\tilde{v}}) \le \varepsilon$.

Notice that $\tilde{A}(\tilde{\xi}) = \mathrm{Id}$. Therefore, taking into account (6-19), inequality (6-15) of Step 1 gives that, for every $\tilde{h} \in \widetilde{\mathcal{A}}_0$ with $\{\tilde{h} \neq \tilde{v}_\xi\} \subseteq B_\rho(\tilde{\xi})$ and $J_{\tilde{h}} \subseteq J_{\tilde{v}_\xi}$ (which is an admissible competitor for $\tilde{v}_\xi$ since automatically $\mathcal{H}^1(L_\xi J_{\tilde{h}}) \le \varepsilon$),

$$\int_{\widetilde{\Omega}_\xi \cap B_\rho(\tilde{\xi})} |\nabla \tilde{v}_\xi|^2\,dx + \mathcal{H}^1(J_{\tilde{v}_\xi} \cap \overline{B}_\rho(\tilde{\xi})) \\ \le \int_{\widetilde{\Omega}_\xi \cap B_\rho(\tilde{\xi})} |\nabla \tilde{h}|^2\,dx + \widetilde{C}_1 \mathcal{H}^1(J_{\tilde{h}} \cap \overline{B}_\rho(\tilde{\xi})) + 2c_{\tilde{A}}(1 + c_1^{\tilde{b}})\rho^{1+\delta} + \widetilde{C}_2 \rho^{(1+\alpha/2)\wedge 2} \tag{6-20}$$

if ρ is sufficiently small. The statement follows thanks to the uniform estimates on L_ξ coming from (6-3). □

Lemma 6.9 (quasimonotonicity). *Under the assumptions of Theorem 6.5, let $v \in \mathcal{A}_0$ be a solution to the minimization problem* (6-11).

There exist $c > 1$, $\tilde{r}_\Omega, \tilde{k}_\Omega > 0$ such that, if $E_x(\rho)$, $\widetilde{E}_x(\rho)$ are defined respectively as in (4-4), (4-5), *where γ, c_γ are the constants of Lemma 6.8, and*

$$\mathcal{E}_x(\rho) := \int_{\Omega \cap B_\rho(x)} |\nabla v|^2\,dx + \mathcal{H}^1(J_v \cap \overline{B}_\rho(x)),$$

the following properties hold true:

(a) *For every $x \in \Omega$ and for every $\rho_1 < \rho_2 < \tilde{r}_\Omega \wedge \mathrm{dist}(x, \partial\Omega)$*

$$E_x(\rho_2) \ge \frac{1}{c} E_x\left(\frac{\rho_1}{c}\right). \tag{6-21}$$

(b) *For every $x \in \partial\Omega$ and for every $\rho_1 < \rho_2 < \tilde{r}_\Omega \wedge \mathrm{dist}(x, S \setminus \{x\})$, where S denotes the set of corners of $\partial\Omega$,*

$$\widetilde{E}_x(\rho_2) \ge \frac{1}{c} \widetilde{E}_x\left(\frac{\rho_1}{c}\right).$$

Proof. Let us start with point (a). Thanks to Lemma 6.8, we have that $\tilde{v}_x \in \widetilde{\mathcal{A}}_0$ is a weak local almost-quasiminimizer (in $\widetilde{\mathcal{A}}_0$) for the Mumford–Shah energy with respect to its own jump set at the point $\tilde{x} = L_x^{-1} x$, so that the quantity

$$\rho \mapsto \frac{\int_{\widetilde{\Omega}_x \cap B_\rho(\tilde{x})} |\nabla \tilde{v}_x|^2\,dy + \mathcal{H}^1(J_{\tilde{v}_x} \cap B_\rho(\tilde{x}))}{\rho} \wedge 1 + \frac{c_\gamma}{\gamma}\rho^\gamma \tag{6-22}$$

is nondecreasing on $(0, \rho_0 \wedge r_{\widetilde{\Omega}_x} \wedge \mathrm{dist}(\tilde{x}, \partial\widetilde{\Omega}_x))$ in view of the interior monotonicity formula of Theorem 4.4. Coming back to the domain Ω, thanks to the bounds on A, we find universal constants $c_i > 0$ such that

$$c_3 \mathcal{E}_x(c_4\rho) \le \int_{\widetilde{\Omega}_x \cap B_\rho(\tilde{x})} |\nabla \tilde{v}_x|^2\,dy + \mathcal{H}^1(J_{\tilde{v}_x} \cap \overline{B}_\rho(\tilde{x})) \le c_1 \mathcal{E}_x(c_2\rho).$$

This estimate, together with the monotonicity of (6-22) and Remarks 3.8 and 4.7, yields easily point (a) of the statement for a suitable $\tilde{r}_\Omega > 0$.

The proof of point (b) is similar: we need to invoke, in place of Theorem 4.4, the boundary monotonicity formula of since Ω is A-admissible by hypothesis, together with Remarks 3.8 and 4.7. □

We are now in a position to prove Theorem 6.5.

Proof of Theorem 6.5. It suffices to show that, for ε small enough, a minimizer $v_\varepsilon \in \mathrm{SBV}(\Omega)$ of

$$\min_{\substack{v\in\mathcal{A}_0\\ \mathcal{H}^1(J_v)\le\varepsilon}} F(U_w+v)$$

is such that $J_{v_\varepsilon} = \varnothing$, which gives $v_\varepsilon = 0$, and the proof follows.

We proceed by contradiction, assuming that there exists $\varepsilon_n \to 0$ such that, for the corresponding $v_n := v_{\varepsilon_n}$, it holds that $\mathcal{H}^1(J_{v_n}) > 0$. Let us divide the proof in two steps.

Step 1: vanishing energy. We claim that

$$\int_{\Omega'} |\nabla v_n|^2\,dx + \mathcal{H}^1(J_{v_n}) \to 0. \tag{6-23}$$

Notice indeed that

$$F(U_w+v_n) \le F(U_w)$$

so that we get easily, taking into account (6-3)–(6-7) and (6-12)

$$\int_{\Omega'} |\nabla v_n|^2\,dx + \mathcal{H}^1(J_{v_n}) + \|v_n\|_\infty \le C.$$

By applying Ambrosio's theorem to the sequence $(v_n)_{n\in\mathbb{N}}$, there exists $v \in \mathcal{A}_0$ such that, up to a subsequence,

$$\begin{aligned} v_n &\to v && \text{strongly in } L^2(\Omega'),\\ \nabla v_n &\rightharpoonup \nabla v && \text{weakly in } L^2(\Omega';\mathbb{R}^2) \end{aligned} \tag{6-24}$$

and

$$\mathcal{H}^1(J_v) \le \liminf_n \mathcal{H}^1(J_{v_n}) \le \lim_n \varepsilon_n = 0. \tag{6-25}$$

In particular $v \in H^1(\Omega')$. Moreover, by the minimality of v_n, in the limit we deduce that

$$F(U_w+v) \le F(U_w+\varphi)$$

for every $\varphi \in H^1(\Omega')$ with $\varphi = 0$ on $\Omega' \setminus \overline{\Omega}$, which yields $v = 0$ in view of the definition of U_w. Passing to the limit in the inequality

$$\begin{aligned} \int_{\Omega'} A(x)[\nabla U_w + \nabla v_n][\nabla U_w + \nabla v_n]\,dx + \int_{J_{v_n}} b(x, v_n^+, v_n^-, \nu_{v_n})\,d\mathcal{H}^1 + \beta\int_{\Omega'} |U_w + v_n - g|^2 \\ \le \int_{\Omega'} A(x)\nabla U_w \nabla U_w\,dx + \beta\int_{\Omega'} |U_w - g|^2 \end{aligned}$$

and taking into account (6-24), (6-25) and the coercivity of A and b, we deduce that

$$\nabla v_n \to 0 \quad \text{strongly in } L^2(\Omega'; \mathbb{R}^2),$$

so that the claim follows.

Step 2: conclusion. By Step 1 we have

$$\eta_n := \int_{\Omega'} |\nabla v_n|^2 \, dx + \mathcal{H}^1(J_{v_n}) \to 0.$$

In view of the *quasimonotonicity* properties enjoyed by v by Lemma 6.9, and taking into account Remark 5.4, we can repeat the arguments of Step 4 in the proof of Theorem 5.1 and get a contradiction, so that the conclusion follows. □

Acknowledgments

Bucur was supported by the ANR-15-CE40-0006 COMEDIC project, the "Geometry and Spectral Optimization" research programme LabEx PERSYVAL-Lab GeoSpec ANR-11-LABX-0025-01 and the Institut Universitaire de France.

References

[Alberti et al. 2003] G. Alberti, G. Bouchitté, and G. Dal Maso, "The calibration method for the Mumford–Shah functional and free-discontinuity problems", *Calc. Var. Partial Differential Equations* **16**:3 (2003), 299–333. MR Zbl

[Ambrosio et al. 2000] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Univ. Press, 2000. MR Zbl

[Bonacini and Morini 2015] M. Bonacini and M. Morini, "Stable regular critical points of the Mumford–Shah functional are local minimizers", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **32**:3 (2015), 533–570. MR Zbl

[Bonnet 1996] A. Bonnet, "On the regularity of edges in image segmentation", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **13**:4 (1996), 485–528. MR Zbl

[Bourdin et al. 2008] B. Bourdin, G. A. Francfort, and J.-J. Marigo, "The variational approach to fracture", *J. Elasticity* **91**:1-3 (2008), 5–148. MR Zbl

[Bucur and Luckhaus 2014] D. Bucur and S. Luckhaus, "Monotonicity formula and regularity for general free discontinuity problems", *Arch. Ration. Mech. Anal.* **211**:2 (2014), 489–511. MR Zbl

[Cagnetti et al. 2008] F. Cagnetti, M. G. Mora, and M. Morini, "A second order minimality condition for the Mumford–Shah functional", *Calc. Var. Partial Differential Equations* **33**:1 (2008), 37–74. MR Zbl

[Chambolle et al. 2008] A. Chambolle, A. Giacomini, and M. Ponsiglione, "Crack initiation in brittle materials", *Arch. Ration. Mech. Anal.* **188**:2 (2008), 309–349. MR Zbl

[Chambolle et al. ≥ 2020] A. Chambolle, E. Séré, and C. Zanini, "Progressive water-waves: a global variational approach", in progress.

[David 1996] G. David, "C^1-arcs for minimizers of the Mumford–Shah functional", *SIAM J. Appl. Math.* **56**:3 (1996), 783–888. MR Zbl

[De Giorgi and Ambrosio 1988] E. De Giorgi and L. Ambrosio, "Un nuovo tipo di funzionale del calcolo delle variazioni", *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* (8) **82**:2 (1988), 199–210. MR Zbl

[De Giorgi et al. 1989] E. De Giorgi, M. Carriero, and A. Leaci, "Existence theorem for a minimum problem with free discontinuity set", *Arch. Ration. Mech. Anal.* **108**:3 (1989), 195–218. MR Zbl

[Francfort and Marigo 1998] G. A. Francfort and J.-J. Marigo, "Revisiting brittle fracture as an energy minimization problem", *J. Mech. Phys. Solids* **46**:8 (1998), 1319–1342. MR Zbl

[Grisvard 1985] P. Grisvard, *Elliptic problems in nonsmooth domains*, Monographs and Studies in Math. **24**, Pitman, Boston, 1985. MR Zbl

[Mumford and Shah 1985] D. Mumford and J. Shah, "Boundary detection by minimizing functionals, I", pp. 22–26 in *IEEE Proceedings on Computer Vision and Pattern Recognition* (San Francisco, 1985), IEEE, Washington, DC, 1985.

[Mumford and Shah 1989] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems", *Comm. Pure Appl. Math.* **42**:5 (1989), 577–685. MR Zbl

Received 16 Oct 2018. Revised 9 Feb 2019. Accepted 3 Apr 2019.

DORIN BUCUR: dorin.bucur@univ-savoie.fr
Université Savoie Mont Blanc, CNRS, LAMA, Chambéry, France

ILARIA FRAGALÀ: ilaria.fragala@polimi.it
Dipartimento di Matematica, Politecnico di Milano, Milano, Italy

ALESSANDRO GIACOMINI: alessandro.giacomini@unibs.it
DICATAM, Sezione di Matematica, Università degli studi di Brescia, Brescia, Italy

ANALYSIS AND PDE
Vol. 13, No. 3, 2020
dx.doi.org/10.2140/apde.2020.13.901

THE GRADIENT FLOW OF THE MÖBIUS ENERGY ε-REGULARITY AND CONSEQUENCES

SIMON BLATT

We study the gradient flow of the Möbius energy introduced by O'Hara (*Topology* **30**:2 (1991), 241–247). We will show a fundamental ε-regularity result that allows us to bound the infinity norm of all derivatives for some time if the energy is small on a certain scale. This result enables us to characterize the formation of a singularity in terms of concentrations of energy and allows us to construct a blow-up profile at a possible singularity. This solves one of the open problems listed by Zheng-Xu He (*Comm. Pure Appl. Math.* **53**:4 (2000), 399–431).

Ruling out blow-ups for planar curves, we will prove that the flow transforms every planar curve into a round circle.

1. Introduction

In their seminal paper, Freedman, He, and Wang [Freedman et al. 1994] suggested the study of the negative gradient flow of the Möbius energy introduced by O'Hara [1991]. For a closed curve $\gamma \in C^{0,1}(\mathbb{R}/l\mathbb{Z}, \mathbb{R}^n)$, $l > 0$, this energy is given by

$$E(\gamma) := \iint_{(\mathbb{R}/l\mathbb{Z})^2} \left(\frac{1}{|\gamma(x) - \gamma(y)|^2} - \frac{1}{d_\gamma(x, y)^2} \right) |\gamma'(x)|\,|\gamma'(y)|\, dx\, dy, \tag{1-1}$$

where $d_\gamma(x, y)$ denotes the distance of the two points $\gamma(x)$, $\gamma(y)$ along γ. Among many other things, Freedman, He, and Wang showed that curves of finite energy are tame and that the Möbius energy can be minimized within every prime knot class. Abrams et al. [2003] proved that the circle minimizes the energy among all closed curves. It is an open problem whether these energies can be minimized within composite knot classes or not.

The evolution equation is governed by the law

$$\partial_t \gamma = -\mathcal{H}\gamma, \tag{1-2}$$

where

$$\mathcal{H}\gamma(x) := 2\,\text{p.v.} \int_{-l/2}^{l/2} \left(2 \frac{P^{\perp}_{\gamma'}(\gamma(x+w) - \gamma(x))}{|\gamma(x+w) - \gamma(x)|^2} - \kappa_\gamma(x) \right) \frac{|\gamma'(x+w)|\, dw}{|\gamma(x+w) - \gamma(x)|^2}$$

and

$$P^{\perp}_{\gamma'} w := w - \left\langle w, \frac{\gamma'}{|\gamma'|} \right\rangle \frac{\gamma'}{|\gamma'|}$$

MSC2010: primary 53C44; secondary 35S10.
Keywords: Möbius energy, geometric evolution equations, gradient flow, long-time existence.

denotes the orthogonal projection onto the normal part along the curve γ [Freedman et al. 1994, Lemma 6.1]. Here, p.v.$\int_{-l/2}^{l/2}$ denotes Cauchy's principal value, i.e., is an abbreviation for $\lim_{\varepsilon\downarrow 0}\int_{I_{l,\varepsilon}}$, where $I_{l,\varepsilon}=[-l/2,l/2]\setminus(-\varepsilon,\varepsilon)$.

If γ is parametrized by arc-length, this further reduces to

$$\mathcal{H}\gamma(x) := 2\,\mathrm{p.v.}\int_{-l/2}^{l/2}\biggl(2\frac{P^{\perp}_{\gamma'}(\gamma(x+w)-\gamma(x))}{|\gamma(x+w)-\gamma(x)|^2}-\gamma''(x)\biggr)\frac{dw}{|\gamma(x+w)-\gamma(x)|^2}. \tag{1-3}$$

Zheng-Xu He [2000, Theorem 2.1] observed that (1-2) is a quasilinear equation of third order and stated a short-time existence result for smooth curves using the Nash–Moser implicit function theorem. Using refined estimates, in [Blatt 2012b] we proved short-time existence for embedded $C^{2+\alpha}$-curves by Banach's fixed-point theorem. Furthermore, we have shown, using a Łojasiewicz–Simon gradient estimate, that local minimizers of the energy are attractive in the sense that there is a $C^{2+\alpha}$-neighborhood of initial data for which the flow exists for all time and converges to a local minimizer. Lin and Schwetlick [2010] considered the elastic energy plus some positive multiple of the Möbius energy and the length. They could show long-time existence for the related negative gradient flow and convergence to critical points by essentially treating the flow as a perturbation of the elastic flow investigated in [Dziuk et al. 2002].

In this paper we derive an ε-regularity result for the evolution equation (1-2) that will be essential in the analysis of the long-time behavior of the flow. As for the Willmore flow [Kuwert and Schätzle 2002] or the biharmonic and polyharmonic heat flow in the critical dimension [Lamm 2004; Gastel 2006] a quantum of the energy has to concentrate whenever a singularity forms.

For any measurable subset $A\subset\mathbb{R}^n$ we define the localized energy

$$E_A(\gamma) := \iint_{(\gamma^{-1}(A))^2}\biggl(\frac{1}{|\gamma(x)-\gamma(y)|^2}-\frac{1}{d_\gamma(x,y)^2}\biggr)|\gamma'(x)|\,|\gamma'(y)|\,dx\,dy. \tag{1-4}$$

Theorem 3.1 (ε-regularity). *There are constants $\varepsilon_0>0$ and $C_k<\infty$, $k\in\mathbb{N}$, depending only on n and $E(\gamma_0)$ such that the following holds: Let γ_t, $t\in[0,T)$, be a maximal smooth solution of* (1-2) *and let $t_0\in[0,T)$, $r>0$, be such that*

$$\sup_{x\in\mathbb{R}^n} E_{B_r(x)}(\gamma)\le\varepsilon_0.$$

Then $T>t_0+r^3$ and

$$\|\partial_s^k\gamma_{t_0+r^3}\|_{L^\infty}\le\frac{C_k}{(rt)^{(k-1)/3}}\quad \textit{for all } t\in(t_0,t_0+r^3].$$

Though the structure of this result is similar to many well-known ε-regularity results for critical evolution equations, due to the nonlocality of the equation one has to develop new techniques in order to prove this theorem. These techniques will certainly be applicable to other nonlocal geometric partial differential equations. The main strategy is to consider the evolution of localized energies and derive differential inequalities. Due to the nonlocality of the equation, however, nonlocal terms appear in these inequalities which make it impossible to apply Gronwall's lemma. We will see that instead a "point-picking method" will help us out.

As a first consequence of this result we prove the following concentration compactness alternative for the flow.

Theorem 4.1 (characterization of singularities). *Let $\gamma \in C^\infty([0,T) \times \mathbb{R}/\mathbb{Z}, \mathbb{R}^n)$ be a maximal smooth solution of* (1-2). *There is a constant $\varepsilon_0 > 0$ depending only on n and $E(\gamma_0)$ such that if $T < \infty$ there are times $t_k \uparrow T$, points $x_k \in \mathbb{R}^n$, and radii $r_k \downarrow 0$, with*

$$E_{B_{r_k}(x_k)}(\gamma_{t_k}) \geq \varepsilon_0.$$

If a singularity occurs, then, by choosing the points x_j in the last theorem more carefully, we can furthermore construct a so-called *blow-up profile*. It is simpler to formulate this theorem using the intrinsically defined local energies

$$E^{\text{int}}_{B_r(x_0)}(\gamma) := \int_{d_\gamma(y,x_0)\leq r} \int_{d_\gamma(x,x_0)\leq r} \left(\frac{1}{|\gamma(x)-\gamma(y)|^2} - \frac{1}{d_\gamma(x,y)^2}\right)|\gamma'(x)||\gamma'(y)|\,dx\,dy.$$

Theorem 4.2 (blow-up profile). *There is an $\varepsilon_0 > 0$ such that the following holds: Assume that γ_t is a solution to* (1-2) *that develops a singularity in finite time, i.e., $T < \infty$ and $r_j \to 0$. Then there are points x_j and times $t_j \to T$ such that*

$$E^{\text{int}}_{B_{r_j}(x_j)}(t_j) \geq \varepsilon_0.$$

Let us now choose the points $x_j \in \mathbb{R}$ and times $t_j \in [0,T)$ such that

$$\sup_{\tau\in[0,t_j],\, x\in\Gamma_\tau} E^{\text{int}}_{B_{r_j}(x)}(\gamma_{t_j}) \leq E^{\text{int}}_{B_{r_j}(x)}(\gamma_{t_j}) = \varepsilon_0,$$

and let $\tilde{\gamma}_j$ be reparametrizations by arc-length of the rescaled and translated curves

$$r_j^{-1}(\gamma_{t_j} - x_j)$$

such that $\tilde{\gamma}_j(0) \in B_2(0)$. Then these curves subconverge locally in C^∞ to an embedded closed or open curve $\tilde{\gamma}_\infty : I \to \mathbb{R}^n$, $I = \mathbb{R}/l\mathbb{Z}$ or $I = \mathbb{R}$ resp., parametrized by arc-length. This curve satisfies

$$\text{p.v.}\int_{-l/2}^{l/2} \left(2\frac{P^\perp_\tau(\tilde{\gamma}(y)-\tilde{\gamma}(x))}{|\gamma(y)-\gamma(x)|^2} - \kappa_\gamma(x)\right)\frac{dy}{|\gamma(y)-\gamma(x)|^2} = 0 \quad \textit{for all } x \in I, \tag{1-5}$$

and

$$E^{\text{int}}_{\bar{B}_1(0)}(\tilde{\gamma}_\infty) \geq \varepsilon_0.$$

This solves Problem 2 of the open problems list in [He 2000]. In the last part of this paper, we deduce a geometric interpretation of the Euler–Lagrange equation of the Möbius energy. In the case of codimension 1, he could show that the only closed critical curves of the Möbius energy are the circles. We will see that unfortunately the blow-up profiles are noncompact. Therefore we cannot apply this result of He in this context. Our new interpretation of the Euler–Lagrange equation allows us to show that the only planar solutions to the Euler–Lagrange equation (1-5) are straight lines and circles. Combining this result with a careful analysis of the asymptotic behavior of the flow, we can finally show:

Theorem 4.8 (planar curves). *Let $\gamma_0 \subset \mathbb{R}^2$ be a closed smoothly embedded curve. Then the negative gradient flow of the Möbius energy exists for all times and converges to a round circle as time goes to infinity.*

Though from the topological point of view the case of planar curves is of no interest, the techniques that lead to this last result reduce the study of the flow to the study of compact and noncompact smooth solutions of the Euler–Lagrange equation (1-5) in the very intuitive geometric form (4-3). Surprisingly, in the classification of planar blow-up profiles this equation is only used in one point, which gives hope that this geometric version of the equation might help to classify blow-up profiles in other situations.

2. Preliminaries and notation

As for most of our estimates the precise algebraic form of the terms does not matter, we will use the following notation to describe the essential structure of the terms.

For two Euclidean vectors v, w, we denote by $v * w$ a bilinear operator in v and w into another Euclidean vector space. For a regular curve γ, let $\partial_s = \partial_x/|\gamma'|$ denote the derivative with respect to arc-length. For $\mu, \nu \in \mathbb{N}$, a regular curve $\gamma \in C^\infty(\mathbb{R}/\mathbb{Z}, \mathbb{R}^n)$, and a function $f : \mathbb{R}/\mathbb{Z} \to \mathbb{R}^k$, let $P^\mu_\nu(f)$ be a linear combination of terms of the form $\partial_s^{j_1} f * \cdots * \partial_s^{j,\nu} f$, where $j_1 + \cdots + j_\nu = \mu$. Furthermore, given a second function $g : \mathbb{R}/\mathbb{Z} \to \mathbb{R}^k$, the expression $P^\mu_\nu(g, f)$ denotes a linear combination of terms of the form $\partial_s^{j_1} g * \partial_s^{j_2} f * \partial_s^{j_3} f * \cdots * \partial_s^{j,\nu} f$, where $j_1 + \cdots + j_\nu = \mu$.

2A. ***Decomposition of the gradient and the operator Q.*** We will always assume that our curve is parametrized by arc-length at the fixed time t we currently consider. Whenever we have to estimate $\mathcal{H}$ we will write it as

$$P^\perp_{\gamma'} \widetilde{\mathcal{H}}, \tag{2-1}$$

where

$$\widetilde{\mathcal{H}}\gamma(x) = 2\,\text{p.v.}\int_{-l/2}^{l/2} \left(2\frac{\gamma(u+w)-\gamma(u)-w\gamma'(u)}{|\gamma(u+w)-\gamma(u)|^2} - \gamma''(x) \right) \frac{dw}{|\gamma(u+w)-\gamma(u)|^2},$$

and take the decomposition

$$\widetilde{\mathcal{H}}\gamma = Q\gamma + R_1\gamma + R_2\gamma = Q\gamma + R\gamma, \tag{2-2}$$

where

$$\begin{aligned} Q\gamma(x) &= 2\,\text{p.v.}\int_{-l/2}^{l/2} \left(2\frac{\gamma(x+w)-\gamma(x)-w\gamma'(x)}{w^4} - \frac{\kappa(x)}{|w|^2} \right) dw \\ &= 4\,\text{p.v.}\int_{-l/2}^{l/2} \int_0^1 (1-s) \frac{\kappa(x+sw)-\kappa(x)}{|w|^2}\, ds\, dw = \widetilde{Q}\kappa(x), \\ R_1\gamma(x) &= 4\int_{-l/2}^{l/2} (\gamma(x+w)-\gamma(x)-w\gamma'(x)) \left(\frac{1}{|\gamma(x+w)-\gamma(x)|^4} - \frac{1}{w^4} \right) dw, \\ R_2\gamma(x) &=: 2\int_{-l/2}^{l/2} \kappa(x) \left(\frac{1}{w^2} - \frac{1}{|\gamma(x+w)-\gamma(x)|^2} \right) dw. \end{aligned}$$

He observed that the operator Q can be written as a multiple of the fractional Laplacian $(-\Delta)^{3/2}$ plus an operator of order 2 [He 2000]. Let us state the consequences of his result for the operator $\widetilde{Q}$ of order 1:

Lemma 2.1. *For every smooth function $f \in C^\infty(\mathbb{R}/l\mathbb{Z}, \mathbb{R}^n)$ we have*

$$\widetilde{Q} f = \frac{1}{l} \sum_{k \in \mathbb{Z}} \frac{\lambda_k}{k} \hat{f}(k),$$

where $\hat{f}(k)$ denotes the k-th Fourier coefficient and $\lambda_k = \pi/3 + O(1/k)$. Hence, for $l \geq 1$ we have

$$\left| \tfrac{1}{9}\pi^2 \|f'\|^2_{L^2} - \|\widetilde{Q} f\|^2_{L^2} \right| \leq C \|f\|^2_{L^2}.$$

Let us add another useful identity for the operator Q to the two identities we already have given above. For smooth f, g we observe, using first partial integration and then discrete partial integration,

$$\begin{aligned}
&\int_{\mathbb{R}/l\mathbb{Z}} \text{p.v.}\int_{-l/2}^{l/2} \int_0^1 (1-s) \frac{f''(x+sw) - f''(x)}{|w|^2}\, ds\, dw\, g(x)\, dx \\
&\quad = \int_{\mathbb{R}/l\mathbb{Z}} \text{p.v.}\int_{-l/2}^{l/2} \int_0^1 (1-s) \frac{f'(x+sw) - f'(x)}{|w|^2} g'(x)\, dw\, dx \\
&\quad = \frac{1}{2}\biggl(\int_{\mathbb{R}/l\mathbb{Z}} \text{p.v.}\int_{-l/2}^{l/2} \int_0^1 (1-s) \frac{f'(x+sw) - f'(x)}{|w|^2} g'(x)\, ds\, dw\, dx \\
&\qquad\qquad - \int_{\mathbb{R}/l\mathbb{Z}} \text{p.v.}\int_{-l/2}^{l/2} \int_0^1 (1-s) \frac{f'(x+sw) - f'(x)}{|w|^2} g'(x+sw)\, dw\, dx \biggr) \\
&\quad = \frac{1}{2} \int_{\mathbb{R}/l\mathbb{Z}} \text{p.v.}\int_{-l/2}^{l/2} \int_0^1 (1-s) \frac{(f'(x+sw) - f'(x))(g'(x+sw) - g'(x))}{w^2}\, ds\, dw\, dx.
\end{aligned}$$

Hence, as we do not need the principal value to make sense of the last expression we have

$$\int_{\mathbb{R}/l\mathbb{Z}} \langle Qf, g\rangle\, ds = 2 \int_{\mathbb{R}/l\mathbb{Z}} \int_{-l/2}^{l/2} \int_0^1 (1-s) \frac{(f'(x+sw) - f'(x))(g'(x+sw) - g'(x))}{w^2}\, ds\, dw\, dx. \tag{2-3}$$

2B. *Coercivity of the Möbius energy and bi-Lipschitz estimates.* Of fundamental importance in the following is the deep connection between the Möbius energy and fractional Sobolev spaces observed in [Blatt 2012a], which was sharpened in [Blatt 2018, Theorem 3.2]. We showed there that the Möbius energy of an embedded curve parametrized by arc-length is finite if and only if the curve is of class $w^{3/2,2}$. More precisely, we have

Theorem 2.2 (characterization of finite energy curves). *Let $\gamma \in C^1(\mathbb{R}/l\mathbb{Z}, \mathbb{R}^n)$ be a curve parametrized by arc-length. Then the energy $E(\gamma)$ is finite if and only if $\gamma \in W^{3/2,2}$. Moreover there exists a constant $C < \infty$ not depending on γ such that*

$$\|\gamma'\|_{W^{3/2,2}} \leq C(E(\gamma)). \tag{2-4}$$

So in particular, for a solution of the gradient flow (1-2), the $W^{3/2,2}$-norm of the gradient after reparametrizing the curve by arc-length is uniformly bounded in time. An essential ingredient of the proof of the theorem above and the analysis in this article is the following bi-Lipschitz estimate for curves of finite energy of [O'Hara 1991]. This bi-Lipschitz constant is also well known under the term *Gromov distortion*.

Lemma 2.3 (bi-Lipschitz estimate). *For an injective curve $\gamma \in W^{3/2,2}(\mathbb{R}/l\mathbb{Z}, \mathbb{R}^n)$ we get the following bound of the Gromov distortion*:

$$\beta = \beta(\gamma) = \sup_{x \neq y} \frac{d_\gamma(x,y)}{|\gamma(x)-\gamma(y)|} \leq 18 e^{E(\gamma)/4}.$$

If γ is parametrized by arc-length, we obtain

$$\frac{|w|}{|\gamma(x+w)-\gamma(x)|} \leq 18 e^{E(\gamma)/4} \quad \text{for all } x, w \in \mathbb{R},\ |w| \leq \frac{l}{2}. \tag{2-5}$$

Let us sketch how this bi-Lipschitz estimate was used in [Blatt 2012a] to prove Theorem 2.2. For a curve $\gamma \in W^{3/2,2}(\mathbb{R}/l\mathbb{Z}, \mathbb{R}^n)$ parametrized by arc-length, $x \in \mathbb{R}/l\mathbb{Z}$, and $0 < |w| < l/2$, we deduce using this bi-Lipschitz estimate the following estimate for the integrand of the energy:

$$\begin{aligned}\frac{1}{|\gamma(x+w)|-\gamma(x)|^2} - \frac{1}{w^2} &= \frac{w^2}{|\gamma(x+w)-\gamma(x)|^2}\,\frac{1-|\gamma(x+w)-\gamma(x)|^2/w^2}{|w^2|}\\ &\leq \frac{\beta}{2}\int_0^1\!\!\int_0^1 \frac{|\gamma'(x+s_1w)-\gamma'(x+s_2w)|^2}{w^2}\,ds_1\,ds_2\\ &\leq 2\beta\int_0^1 \frac{|\gamma(x+sw)-\gamma(x)|^2}{w^2}\,ds.\end{aligned} \tag{2-6}$$

One then derives the statement of Theorem 2.2 by basically integrating this inequality over all x and w.

More generally, for $\alpha \geq 0$ and using that the function

$$x \to \frac{1-x^{2+\alpha}}{1-x^2}$$

is locally bounded on $(0, \infty)$ we get

$$\begin{aligned}\frac{|w|^\alpha}{|\gamma(x+w)|-\gamma(x)|^{\alpha+2}} - \frac{1}{w^2} &= \frac{|w|^{2+\alpha}}{|\gamma(x+w)-\gamma(x)|^{2+\alpha}}\,\frac{1-|\gamma(x+w)-\gamma(x)|^{2+\alpha}/|w|^{2+\alpha}}{|w^2|}\\ &\leq C\,\frac{1-|\gamma(x+w)-\gamma(x)|^2/|w|^2}{|w^2|}\\ &\leq C\int_0^1\!\!\int_0^1 \frac{|\gamma'(x+s_1w)-\gamma'(x+s_2w)|^2}{w^2}\,ds_1\,ds_2\\ &\leq C\int_0^1 \frac{|\gamma(x+sw)-\gamma(x)|^2}{w^2}\,ds,\end{aligned} \tag{2-7}$$

where the constant C depends only on an upper bound for β like $E(\gamma)$ and α. Furthermore, we have

$$\frac{1}{|\gamma(x+w)-\gamma(x)|^2} - \frac{1}{w^2} = \frac{w^2/|\gamma(x+w)-\gamma(x)|^2-1}{w^2} \leq \frac{18^2 e^{E(\gamma)/2}-1}{w^2}.$$

So we get the rough estimate

$$\int_{B_r(x)} \int_{l/2 \geq |w| \geq \Lambda r} \left(\frac{1}{|\gamma(x+w) - \gamma(x)|^2} - \frac{1}{w^2} \right) dw\, dx \leq C(\beta) \int_{\Lambda r}^{\infty} \frac{dw}{w^2} \leq C(\beta). \tag{2-8}$$

2C. ***Fractional Sobolev spaces and Besov spaces.*** In our calculation, fractional Sobolev spaces as well as Besov spaces naturally appear. For an introduction to Besov spaces we refer to [Triebel 1983; 1992]. Let $f \in L^1(\mathbb{R}/l\mathbb{Z})$. For $s \in (0, 1)$ and $p, q \in [1, \infty)$ and for open subsets $\Omega \subset \mathbb{R}/l\mathbb{Z}$ we also consider the *Besov-type* seminorm

$$|f|_{B^s_{p,q}(B_R(x))} := \left(\int_{B_R(x)} \frac{\left(\int_{-R/2}^{R/2} |f'(u+w) - f'(u)|^p \, du \right)^{q/p}}{|w|^{1+qs}} \, dw \right)^{1/q}. \tag{2-9}$$

It is shown in the Appendix that

$$|f|_{B^s_{p,q}(B_R(x))} \leq C \|f\|_{B^s_{p,q}(B_{2R}(x))},$$
$$\|f\|_{B^s_{p,q}(B_R(x))} \leq C\left(|f|_{B^s_{p,q}(B_{2R}(x))} + \|f\|_{L^p(B_{2R}(x))}\right).$$

3. An ε-regularity result

In this section we prove the main result of this article, an ε-regularity result for the flow (1-2):

Theorem 3.1 (ε-regularity). *There are constants $\varepsilon > 0$ and $C_k < \infty$, $k \in \mathbb{N}$, depending only on n and $E(\gamma_0)$ such that the following holds: Let γ_t, $t \in [0, T)$ be a maximal smooth solution of* (1-2) *and let $t_0 \in [0, T)$, r be such that*

$$\sup_{x \in \mathbb{R}^n} E_{B_r(x)}(\gamma_{t_0}) \leq \varepsilon. \tag{3-1}$$

Then $T > t_0 + r^3$ and

$$\|\partial_s^k \gamma_{t_0 + r^3}\|_{L^\infty} \leq \frac{C_k}{(rt)^{(k-1)/3}} \quad \textit{for all } t \in (t_0, t_0 + r^3].$$

Remark 3.2. Note that the assumptions in the theorem are highly nonlocal. It is a very interesting and challenging question whether one can prove a local version of this regularity theorem.

Clearly, one only has to prove Theorem 3.1 for the special case $t_0 = 0$ and $r = 1$. Scaling and translation in time then give the full statement.

We will prove Theorem 3.1 in three steps using energy estimates for this special case. First we control the energy within a ball of radius 1 at later times, before we estimate the elastic energy, i.e., the L^2-norm of the curvature. In a last step we will then bound higher-order energies. The general strategy will always be to derive evolution equations for the quantities and use the quasilinear structure together with interpolation estimates in order to derive differential inequalities (see Lemmas 3.10, 3.19, and 3.20).

Due to the nonlocal structure of the inequalities, though we start with local quantities these differential inequalities are also nonlocal, which makes the usual application of Gronwall's lemma impossible. A kind of point-picking method will help us there.

3A. ***Estimates for the energy density.*** Let us fix a radial cutoff function $\phi(x) = \phi(|x|) \in C_c^\infty(\mathbb{R}^n)$ such that

$$\chi_{B_1(0)} \leq \phi \leq \chi_{B_2(0)}.$$

For $x_0 \in \mathbb{R}^n$ we set $\phi_{x_0}(x) := \phi(x - x_0)$ and define the localized energy

$$E^{\phi_{x_0}}(\gamma) := \iint_{(\mathbb{R}/\mathbb{Z})^2} \left(\frac{1}{|\gamma(x)-\gamma(y)|^2} - \frac{1}{d_\gamma(x,y)^2} \right) |\gamma'(x)||\gamma'(y)| \phi_{x_0}(\gamma(x))\, dx\, dy. \tag{3-2}$$

A straightforward calculation leads to the following evolution equation for E^ϕ. We leave the proof to the reader.

Lemma 3.3 (evolution equation for local density). *Let γ_t be parametrized by arc-length and $(d/dt)\gamma_t = V$ be orthogonal to γ_t. Then we have*

$$\begin{aligned}
\frac{d}{dt} E^\phi(\gamma_t) &= 2\,\text{p.v.} \int_{-l/2}^{l/2} \int_{I_{l,\varepsilon}} \left\langle 2 \frac{\gamma(x+w)-\gamma(x)}{|\gamma(x+w)-\gamma(x)|^4} - \frac{\kappa_\gamma(x)}{|\gamma(x+w)-\gamma(x)|^2}, V(x) \right\rangle \phi(\gamma(x))\, dw\, dx \\
&\quad + 2 \int_{\mathbb{R}/l\mathbb{Z}} \int_{-l/2}^{l/2} \frac{\left\langle \gamma(x+w)-\gamma(x) - w\gamma'(x) - \frac{1}{2}|\gamma(x+w)-\gamma(x)|^2 \kappa(x), V(x) \right\rangle}{|\gamma(x+w)-\gamma(x)|^4} \\
&\qquad\qquad \times (\phi(\gamma(x+w)) - \phi(\gamma(x)))\, dw\, dx \\
&\quad + \int_{\mathbb{R}/l\mathbb{Z}} \int_{-l/2}^{l/2} \frac{\langle \kappa(x), V(x) \rangle}{|w|^2} \left(\phi(\gamma(x+w) + \phi(\gamma(x)) - 2 \int_0^1 \phi(\gamma(x+\tau w)) d\tau \right) dw\, dx \\
&\quad + \int_{\mathbb{R}/l\mathbb{Z}} \int_{-l/2}^{l/2} \left(\frac{1}{|\gamma(x+w)-\gamma(x)|^2} - \frac{1}{|w|^2} \right) \langle V(x), \nabla\phi(\gamma(x)) \rangle\, dw\, dx \\
&=: I_1 + I_2 + I_3 + I_4,
\end{aligned}$$

where l is the length of γ_t.

In the rest of this section we estimate these terms for the case

$$V = \mathcal{H}\gamma.$$

To make the calculations and formulas as simple as possible, we always assume that the curve γ_t is parametrized by arc-length at the current time t. We will use the intrinsically defined quantities

$$\begin{aligned}
M_{3/2} = M_{3/2}(t) &= \sup_{x \in \mathbb{R}/l\mathbb{Z}} \int_{-1/2}^{1/2} \int_{-l/2}^{l/2} \frac{|\gamma'(x+w)-\gamma'(x)|^2}{w^2}\, dw\, dx, \\
S_3(x) = S_3(x,t) &= \|\partial_s^3 \gamma_t\|^2_{L^2(B_\Lambda(x))} + \sum_{j=1}^\infty \frac{\|\partial_s^3 \gamma_t\|^2_{L^2(B_{\Lambda+j}(x) \setminus B_{\Lambda+(j-1)}(x))}}{(\Lambda/2+j)^2}, \\
\widetilde{S}_3(x) = \widetilde{S}_3(x,t) &= \|\mathcal{H}\gamma_t(x)\|^2_{L^2(B_\Lambda(0))} + \sum_{j=1}^\infty \frac{\|\mathcal{H}\gamma_t\|^2_{L^2(B_{\Lambda+j}(x))}}{(\Lambda/2+j)^2}
\end{aligned}$$

for $\Lambda = 1000 \cdot 18 e^{E(\gamma_0)/4} < \infty$. Note, that due to Lemma 2.3 the quantity $18e^{E(\gamma_0)/4}$ bounds the Gromov distortion of the curves γ_t for all t. Hence, Λ is large compared with the Gromov distortion of γ.

To prevent complicated terms in the estimates, we will assume throughout this section that

$$M_{3/2} \leq 1.$$

Furthermore, we will assume that $\gamma(0) \in B_2(0)$ to get some preliminary estimates in terms of the intrinsically defined quantities above. In the final differential inequality we will use the extrinsic quantity

$$S_3^{\text{ext}}(x,t) := \|\mathcal{H}\gamma(x)\|^2_{L^2(\gamma^{-1}(B_1(0)))} + \sum_{j=1}^{\infty} \frac{\|\mathcal{H}\gamma\|^2_{L^2(\gamma^{-1}(B_{j+1}(x)\setminus B_j(x))}}{j^2}$$

in place of $S_3(0)$.

Let us start with the following easy, but useful lemma that will help us to control the part of the integrals defining I_i, $i = 1, \ldots, 4$, for the case that $|w|$ is large.

Lemma 3.4. *For all $s \in [0,1]$, $p \in [1,\infty)$, and $x \in B_\beta(0)$ we have*

$$\int_{|w|\geq\Lambda} \frac{|f(x+sw)|^p}{w^2}\,dw \leq C\|f\|^p_{L^p(B_\Lambda(0))} + \sum_{j\in\mathbb{N}} \frac{\|f\|^p_{L^p(B_{\Lambda+j}(0)-B_{\Lambda+j-1}(0))}}{(\Lambda+j)^2}$$

Proof. The statement obviously holds for $s=0$. For $s>0$ we get substituting $\tilde{w} = sw$

$$\begin{aligned}\int_{|w|\geq\Lambda} \frac{|f(x+sw)|^p}{w^2}\,dw &= s\int_{\Lambda/2\geq|\tilde{w}|\geq s\Lambda/2} \frac{|f(x+\tilde{w})|^p}{\tilde{w}^2}\,d\tilde{w} + s\int_{|\tilde{w}|\geq\Lambda/2} \frac{|f(x+\tilde{w})|^p}{\tilde{w}^2}\,d\tilde{w} \\ &\leq C\|f\|^p_{L^p(B_\Lambda)} + s\int_{|\tilde{w}|\geq\Lambda/4} \frac{|f(\tilde{w})|^p}{\tilde{w}^2}\,d\tilde{w} \\ &\leq C\|f\|^p_{L^p(B_\Lambda)} + \sum_{j\in\mathbb{N}} \frac{\|f\|^p_{L^p(B_{\Lambda+j}(0)\setminus B_{\Lambda+j-1}(0))}}{(\Lambda+j)^2}. \qquad \square\end{aligned}$$

We start with estimating the term I_1, which contains the terms of highest order. The guideline for estimating the remainder terms will be throughout this section to distinguish between areas where $|w|$ is small and where $|w|$ is big. Combining this idea with the commutator estimates and interpolation inequalities in the Appendix (see Lemmas A.3 and A.4) we obtain the desired estimates.

Lemma 3.5 (estimate for I_1). *Let γ be parametrized by arc-length, $M_{3/2} \leq 1$, and $\gamma(0) \in B_2(0)$. Then there is a constant $\alpha > 0$*

$$I_1 = -\int_{\mathbb{R}/l\mathbb{Z}} |\mathcal{H}\gamma|^2 \phi(\gamma(x))\,dx = -\int_{\mathbb{R}/\mathbb{Z}} |Q\gamma(x)|^2\phi(\gamma(x))\,dx + R_I,$$

where for all $\varepsilon > 0$

$$R_I \leq (C_\varepsilon M^\alpha_{3/2} + \varepsilon)S_3(0) + C_\varepsilon$$

for some $C_\varepsilon < \infty$.

Proof. We have

$$\mathcal{H}\gamma(x) = P^\perp_{\gamma'(x)}(Q\gamma(x) + R_1\gamma(x) + R_2\gamma(x)),$$

where

$$\begin{aligned} Q\gamma(x) &= 2\lim_{\varepsilon\downarrow 0}\int_{I_{l,\varepsilon}}\left(2\frac{\gamma(x+w)-\gamma(x)-w\gamma'(x)}{w^4}-\frac{\kappa(x)}{|w|^2}\right)dw \\ &= 4\lim_{\varepsilon\downarrow 0}\int_{I_{l,\varepsilon}}\int_0^1(1-s)\frac{\kappa(x+sw)-\kappa(x)}{|w|^2}\,dw=\widetilde{Q}\kappa(x), \\ R_1\gamma(x) &= 4\int_{I_l}(\gamma(x+w)-\gamma(x)-w\gamma'(x))\left(\frac{1}{|\gamma(x+w)-\gamma(x)|^4}-\frac{1}{w^4}\right)dw, \\ R_2\gamma(x) &= 2\int_{I_l}\kappa(x)\left(\frac{1}{w^2}-\frac{1}{|\gamma(x+w)-\gamma(x)|^2}\right)dw. \end{aligned} \tag{3-3}$$

The bi-Lipschitz estimate together with $\gamma(0)\in B_2(0)$ tells us $\phi(\gamma(x))=0$ for all $x\notin B_{2\beta}(0)$. This yields

$$\begin{aligned} -\int_{\mathbb{R}/l\mathbb{Z}}|P_\gamma^\perp(Q\gamma(x))|^2\phi(\gamma(x))\,dx &= -\int_{\mathbb{R}/l\mathbb{Z}}|Q\gamma(x)|^2\phi(\gamma(x))\,dx+\int_{\mathbb{R}/l\mathbb{Z}}|\langle Q\gamma(x),\gamma'\rangle|^2\phi(\gamma(x))\,dx \\ &\le -\int_{\mathbb{R}/l\mathbb{Z}}|Q\gamma(x)|^2\phi(\gamma(x))\,dx+\int_{-2\beta}^{2\beta}|\langle Q\gamma(x),\gamma'\rangle|^2\,dx. \end{aligned} \tag{3-4}$$

Using that $\langle\kappa,\gamma'\rangle=0$ and that $\widetilde{Q}$ is a linear operator, we get

$$\begin{aligned} |\langle Q\gamma(x),\gamma'\rangle| &= |\langle\widetilde{Q}\kappa(x),\gamma'\rangle| = |\langle\widetilde{Q}\kappa(x),\gamma'\rangle-\widetilde{Q}[\langle\kappa,\gamma'\rangle](x)| \\ &= \left|\sum_{i=1}^n(\widetilde{Q}[\kappa_i](x)\gamma_i'(x)-\widetilde{Q}[\kappa_i\gamma_i'](x))\right|. \end{aligned}$$

Hence, applying first the commutator estimate (Lemma A.4) and then the interpolation estimates (Lemma A.3) we obtain

$$\begin{aligned} \|\langle Q\gamma,\gamma'\rangle\|_{L^2(B_{2\beta}(0))} &\le C(\|\kappa\|_{B^{1/2}_{4,2}(B_\Lambda(0))}\|\gamma'\|_{B^{1/2}_{4,2}(B_\Lambda(0))}+\|\kappa\|_{L^2(B_\Lambda(0))}(\|\gamma'\|_{C^{0,1}(B_\Lambda(0))}+1)) \\ &\le C(M_{3/2}^{1/2}S_3^{1/2}(0)+M_{3/2})\le C(M_{3/2}^{1/2}S_3^{1/2}(0)+1). \end{aligned} \tag{3-5}$$

Using Taylor's theorem and (2-6), we get

$$\begin{aligned} |R_1\gamma(x)| &= 4\left|\int_{-l/2}^{l/2}\int_0^1(1-s)\kappa(x+sw)\left(\frac{w^2}{|\gamma(x+w)-\gamma(x)|^2}-\frac{1}{w^2}\right)ds\,dw\right| \\ &\le C\int_{I_l}\iiint_{[0,1]^3}|\kappa(x+s_1w)|\frac{|\gamma'(x+s_2w)-\gamma'(x+s_3w)|^2}{|w|^2}\,ds_1\,ds_2\,ds_3\,dw \\ &\le C\int_{I_l}\iint_{[0,1]^2}|\kappa(x+s_1w)|\frac{|\gamma'(x+s_2w)-\gamma'(x)|^2}{|w|^2}\,ds_1\,ds_2\,dw \\ &= C(R_{11}\gamma(x)+R_{12}\gamma(x)), \end{aligned}$$

where

$$\begin{aligned} R_{11}\gamma(x) &= \int_{|w|\le\Lambda/2}\iint_{[0,1]^2}g_{w,s_1,s_2}(x)\,ds_1\,ds_2\,dw, \\ R_{12}\gamma(x) &= \int_{l/2\ge|w|\ge\Lambda/2}\iint_{[0,1]^2}g_{w,s_1,s_2}(x)\,ds_1\,ds_2\,dw, \end{aligned}$$

and

$$g_{w,s_1,s_2}(x) := |\kappa(x+s_1w)| \frac{|\gamma'(x+s_2w)-\gamma'(x)|^2}{|w|^2}.$$

Since

$$\int_{B_{2\beta}(0)} |g_{w,s_1,s_2}(x)|^2\,dx \le \|\kappa\|^2_{L^4(B_\Lambda(0))} \frac{\|\gamma'(\,\cdot\,+s_2w)-\gamma'\|^4_{L^8(B_{2\beta}(0))}}{|w|^2},$$

we get

$$\begin{aligned}\|R_{11}\gamma(x)\|_{L^2(B_{2\beta}(0))} &\le C\|\kappa\|_{L^4(B_\Lambda(0))} \int_{|w|\le\Lambda/2} \iint_0^1 \frac{\|\gamma'(\,\cdot\,+s_2w)-\gamma'\|^2_{L^8(B_{2\beta}(0))}}{|w|^2}\,ds\,dw \\ &\le C\|\kappa\|_{L^4(B_\Lambda(0))} \|\gamma'\|^2_{B^{1/2}_{8,2}(B_\Lambda(0))}.\end{aligned}$$

Furthermore, since $|\gamma'|\equiv 1$ we get by Cauchy's inequality and Lemma 3.4

$$\begin{aligned}|R_{12}\gamma(x)| &\le 4\int_{l/2\ge|w|\ge\Lambda/2}\int_0^1 \frac{|\kappa(x+sw)|}{|w|^2}\,ds\,dw \\ &\le C\left(\int_0^1\int_{l/2\ge|w|\ge\Lambda/2} \frac{|\kappa(x+sw)|^2}{|w|^2}\,dw+1\right) \\ &\le C\left(\|\kappa\|^2_{L^2(B_\Lambda(0))} + \sum_{j=1}^\infty \frac{\|\kappa\|^2_{L^2(B_{\Lambda+j}-B_{\Lambda+j-1})}}{(\Lambda+j)^2}+1\right).\end{aligned}$$

Thus

$$\|R_{12}\gamma(x)\|_{L^2(B_{2\beta}(0))} \le C\left(\|\kappa\|^2_{L^2(B_\Lambda(0))} + \sum_{j=1}^\infty \frac{\|\kappa\|^2_{L^2(B_{\Lambda+j}-B_{\Lambda+j-1})}}{(\Lambda+j)^2}+1\right).$$

Together with the interpolation inequalities from Lemma A.3 this leads to

$$\begin{aligned}\|R_1\gamma\|^2_{L^2(B_{2\beta}(0))} &\le C\left(\|\kappa\|^2_{L^4(B_\Lambda(0))}\|\gamma'\|^4_{B^{1/2}_{8,2}(B_\Lambda(0))} + \|\kappa\|^2_{L^2(B_\Lambda(0))} + \sum_{j=1}^\infty \frac{\|\kappa\|^2_{L^2(B_{\Lambda+j}-B_{\Lambda+j-1})}}{(\Lambda+j)^2}+1\right) \\ &\le C\left(M^2_{3/2}S_3 + \|\partial_s^3\gamma\|^{4/3}_{L^2(B_\Lambda)} + \sum_{j=1}^\infty \frac{\|\partial_s^3\gamma\|^{4/3}_{L^2(B_{\Lambda+j}-B_{\Lambda+j-1})}}{(\Lambda+j)^2}+1\right) \\ &\le (CM^2_{3/2}+\varepsilon)S_3 + C_\varepsilon,\end{aligned}$$

where we have used the Cauchy inequality in the last step.

In the same way, one deals with the term R_2 to get

$$\|P^\perp_{\gamma'}R\|^2_{L^2(B_{2\beta}(0))} \le \|R\|^2_{B_{2\beta}(0)} \le (CM^3_{3/2}+\varepsilon)S_3 + C_\varepsilon. \tag{3-6}$$

From (3-4), (3-5), and (3-6) the assertion follows. □

Lemma 3.6 (estimate for I_2). *Let $M_{3/2}\le 1$ and $\gamma(0)\in B_2(0)$. For all $\varepsilon>0$*

$$|I_2| \le \varepsilon(S_3+\widetilde{S}_3)+C_\varepsilon$$

for some $C_\varepsilon<\infty$ depending only on ε and $E(\gamma_0)$.

Proof. We take the decomposition

$$\begin{aligned} I_2 &= 2\int_{\mathbb{R}/l\mathbb{Z}}\int_{-l/2}^{l/2} \frac{\left\langle \gamma(x+w)-\gamma(x)-w\gamma'(x)-\frac{1}{2}|\gamma(x+w)-\gamma(x)|^2\kappa(x),\, V(x)\right\rangle}{|\gamma(x+w)-\gamma(x)|^4} \\ &\qquad\qquad \times (\phi(\gamma(x+w))-\phi(\gamma(x)))\,dw\,dx \\ &= 2\int_{\mathbb{R}/l\mathbb{Z}}\int_{-l/2}^{l/2} \frac{\left\langle \gamma(x+w)-\gamma(x)-w\gamma'(x)-\frac{1}{2}|w|^2\kappa(x),\, V(x)\right\rangle}{|\gamma(x+w)-\gamma(x)|^4}(\phi(\gamma(x+w))-\phi(\gamma(x)))\,dw\,dx \\ &\quad -\frac{1}{2}\int_{\mathbb{R}/l\mathbb{Z}}\int_{-l/2}^{l/2} \frac{|\gamma(x+w)-\gamma(x)|^2-|w|^2}{|\gamma(x+w)-\gamma(x)|^4}\langle\kappa(x),\, V(x)\rangle(\phi(\gamma(x+w))-\phi(\gamma(x)))\,dw\,dx \\ &=: I_{21}+I_{22}. \end{aligned}$$

Using the bi-Lipschitz estimate (2-5) and Taylor's approximation up to the first order, we get

$$I_{21} \le C\int_{\mathbb{R}/l\mathbb{Z}}\int_{-l/2}^{l/2}\int_0^1 \frac{|\kappa(x+sw)-\kappa(x)|}{w^2}|V(x)|\,|\phi(\gamma(x+w)-\phi(\gamma(x))|\,ds\,dw\,dx.$$

Observing that

$$\phi(\gamma(x+w))-\phi(\gamma(x))=0$$

if both $|x|, |x+w| \ge 2\beta$, this can be estimated by

$$\begin{aligned} I_{21} &\le C\int_{B_{\Lambda/2}(0)}\int_{-l/2}^{l/2}\int_0^1 \frac{|\kappa(x+sw)-\kappa(x)|}{|w|}|V(x)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,ds\,dw\,dx \\ &\quad + C\int_{\mathbb{R}/l\mathbb{Z}}\int_{x+w\in B_{\Lambda/2}}\int_0^1 \frac{|\kappa(x+sw)-\kappa(x)|}{|w|^2}|V(x)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,ds\,dw\,dx \\ &\le C\int_{B_{\Lambda/2}(0)}\int_{-l/2}^{l/2}\int_0^1 \frac{|\kappa(x+sw)-\kappa(x)|}{|w|}|V(x)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,ds\,dw\,dx \\ &\quad + C\int_{\mathbb{R}/l\mathbb{Z}}\int_{x+w\in B_{\Lambda/2}}\int_0^1 \frac{|\kappa(x+sw)-\kappa(x+w)|}{|w|^2}|V(x)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,ds\,dw\,dx \\ &\quad\quad + C\int_{\mathbb{R}/l\mathbb{Z}}\int_{x+w\in B_{\Lambda/2}}\int_0^1 \frac{|\kappa(x+w)-\kappa(x)|}{|w|^2}|V(x)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,ds\,dw\,dx \\ &\le C\int_{B_{\Lambda/2}(0)}\int_{-l/2}^{l/2}\int_0^1 \frac{|\kappa(x+sw)-\kappa(x)|}{|w|}|V(x)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,ds\,dw\,dx \\ &\quad + C\int_{B_{\Lambda/2}}\int_{-l/2}^{l/2}\int_0^1 \frac{|\kappa(x+sw)-\kappa(x)|}{|w|^2}|V(x+w)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,ds\,dw\,dx \\ &\quad\quad + C\int_{B_{\Lambda/2}}\int_{-l/2}^{l/2} \frac{|\kappa(x+w)-\kappa(x)|}{|w|^2}|V(x+w)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,dw\,dx \\ &\le C\sup_{s_1,s_2\in[0,1]}\int_{B_{\Lambda/2}}\int_{-l/2}^{l/2} \frac{|\kappa(x+s_1w)-\kappa(x)|}{|w|^2}|V(x+s_2w)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,dw\,dx. \end{aligned}$$

To estimate this last supremum we decompose the integral into

$$\int_{B_{\Lambda/2}}\int_{-l/2}^{l/2}\frac{|\kappa(x+s_1w)-\kappa(x)|}{|w|^2}|V(x+s_2w)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,dw\,dx$$
$$\leq \int_{B_{\Lambda/2}}\int_{-\Lambda/2}^{\Lambda/2}\frac{|\kappa(x+s_1w)-\kappa(x)|}{|w|}|V(x+s_2w)|\,dw\,dx + \int_{B_{\Lambda/2}}\int_{|w|\geq\Lambda/2}\frac{|\kappa(x+s_1w)-\kappa(x)|}{|w|^2}|V(x+s_2w)|\,dw\,dx.$$

Then we can estimate the first term by

$$C\|\kappa\|_{B^{1/2}_{2,2}(B_\Lambda)}\|V\|_{L^2(B_{\Lambda/2})}\leq \varepsilon\widetilde{S}_3 + C_\varepsilon\|\kappa\|^2_{B^{1/2}_{2,2}(B_\Lambda)}\leq \varepsilon(\widetilde{S}_3+S_3)+C_\varepsilon,$$

where we used the interpolation estimates in Lemma A.3 and $M_{3/2}\leq 1$. We estimate the second term using Lemma 3.4 and then again the interpolation estimates yield

$$C\int_{x\in B_{\Lambda/2}}\int_{w\geq\Lambda/2}\frac{|\kappa(x+s_1w)|^2+|V(x+s_2w)|^2}{|w|^2}\,ds$$
$$\leq \varepsilon S_3 + C_\varepsilon\left(\|\kappa\|^2_{L^2(B_\Lambda(0))}+\sum_{j=1}^\infty\frac{\|\kappa\|^2_{L^2(B_{\Lambda+j}-B_{\Lambda+j-1})}}{(\Lambda+j)^2}\right)$$
$$\leq \varepsilon S_3+\varepsilon\|\partial_s^3\gamma\|^2_{B_\Lambda(0)}+\varepsilon\sum_{j=1}^\infty\frac{\|\partial_s^3\gamma\|^2_{L^2(B_{\Lambda+j}-B_{\Lambda+j-1})}}{(\Lambda+j)^2}+C_\varepsilon\left(1+\sum_{j=1}^\infty\frac{1}{(\Lambda+j)^2}\right)$$
$$\leq \varepsilon(S_3+\widetilde{S}_3)+C_\varepsilon.$$

Hence,

$$I_{21}\leq \varepsilon(S_3+\widetilde{S}_3)+C_\varepsilon.$$

Similarly, we get

$$I_{22}\leq C\int_{\mathbb{R}/l\mathbb{Z}}\int_{-l/2}^{l/2}\int_{[0,1]^2}\frac{|\gamma'(x+s_1w)-\gamma'(x+s_2w)|^2}{|w|^2}|\kappa(x)|\,|V(x)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,dw\,dx$$
$$\leq C\int_{\mathbb{R}/l\mathbb{Z}}\int_{-l/2}^{l/2}\int_0^1\frac{|\gamma'(x+s_1w)-\gamma'(x)|^2}{|w|^2}|\kappa(x)|\,|V(x)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,dw\,dx$$
$$\leq \sup_{s_1,s_2\in[0,1]} C\int_{B_{\Lambda/2}(0)}\int_{|w|\leq l/2}\int_0^1\frac{|\gamma'(x+s_1w)-\gamma'(x)|^2}{|w|^2}|\kappa(x+s_2w)| \times|V(x+s_2w)|\,|\phi(\gamma(x+w))-\phi(\gamma(x))|\,dw\,dx,$$

which as above can be estimated by

$$\varepsilon(S_3+\widetilde{S}_3)+C_\varepsilon.$$

□

Lemma 3.7 (estimate for I_3). *Let $M_{3/2}\leq 1$ and $\gamma(0)\in B_2(0)$. Given $\varepsilon>0$ we have*

$$|I_3|\leq \varepsilon(S_3+\widetilde{S}_3)+C_\varepsilon$$

for some $C_\varepsilon<\infty$.

Proof. We use that

$$\left|\phi(\gamma(x+w))+\phi(\gamma(x))-2\int_0^1\phi(\gamma(x+\tau w))\,d\tau\right|=C|w|^2$$

and for $x\notin B_{\Lambda/2}(0)$

$$\left|\phi(\gamma(x+w))+\phi(\gamma(x))-2\int_0^1\phi(\gamma(x+\tau w))\,d\tau\right|\le\begin{cases}0 & \text{for } |w|\le|x|-2\beta,\\ 2 & \text{for } |x|-2\beta\le|w|\le|x|+2\beta,\\ 2/|w| & \text{for } |x|+2\beta\le|w|\end{cases}$$

to get

$$\int_{-l/2}^{l/2}\frac{\left|\phi(\gamma(x+w))+\phi(\gamma(x))-2\int_0^1\phi(\gamma(x+\tau w))\,d\tau\right|}{w^2}\,dw\le\frac{C}{x^2}$$

if $|x|\ge\Lambda/2$ and

$$\int_{-l/2}^{l/2}\frac{\left|\phi(\gamma(x+w))+\phi(\gamma(x))-2\int_0^1\phi(\gamma(x+\tau w))\,d\tau\right|}{w^2}\,dw\le C$$

if $|x|\le\Lambda/2$. These estimates then imply

$$I_3\le C\left(\int_{B_{\Lambda/2}(0)}|\kappa(x)||V(x)|\,dx+\int_{\mathbb{R}/l\mathbb{Z}-B_{\Lambda/2}(0)}\frac{|\kappa(x)||V(x)|}{|x|^2}\,dx\right).$$

From here again Hölder's inequality together with Lemma 3.4 and the interpolation inequalities of Lemma A.3 imply the assertion of the lemma as in the proof of Lemma 3.6. □

Lemma 3.8 (estimate for I_4). *Let $M_{3/2}\le 1$ and $\gamma(0)\in B_2(0)$. For all $\varepsilon>0$*

$$|I_4|\le\varepsilon(S_3+\widetilde{S}_3)+C_\varepsilon$$

for some $C_\varepsilon<\infty$.

Proof. To estimate I_4 we use (2-6) to get

$$\begin{aligned}|I_4|&=\left|\int_{\mathbb{R}/l\mathbb{Z}}\int_{-l/2}^{l/2}\left(\frac{1}{|\gamma(x+w)-\gamma(x)|^2}-\frac{1}{|w|^2}\right)\langle V(x),\nabla\phi(\gamma(x))\rangle\,dw\,dx\right|\\&\le C\int_{\mathbb{R}/l\mathbb{Z}}\int_{-l/2}^{l/2}\iint_{[0,1]^2}\frac{|\gamma'(x+s_1w)-\gamma'(x+s_2w)|^2}{|w|^2}|V(x)||\nabla\phi(\gamma(x))|\,ds_1\,ds_2\,dw\,dx\\&\le C\int_{B_{2\beta}(0)}\int_{-l/2}^{l/2}\int_0^1\frac{|\gamma'(x+sw)-\gamma'(x)|^2}{|w|^2}|V(x)|\,ds\,dw\,dx\\&\le C\int_{-l/2}^{l/2}\int_0^1\left(\int_{x\in B_{2\beta}(0)}\frac{|\gamma'(x+s_1w)-\gamma'(x)|^4}{|w|^2}\,dx\right)^{1/2}dw\,ds\|V\|_{L^2(B_2(0))}.\end{aligned}$$

Since

$$\begin{aligned}\int_{-l/2}^{l/2}\int_0^1\Big(\int_{x\in B_{2\beta}(0)}\frac{|\gamma'(x+s_1w)-\gamma'(x)|^4}{|w|^2}\,dx\Big)^{1/2}dw\,ds\\ &\le \int_0^1\int_{-sl/2}^{sl/2}\Big(\int_{x\in B_{2\beta}(0)}\frac{|\gamma'(x+w)-\gamma'(x)|^4}{|w|^2}\,dx\Big)^{1/2}dw\,ds\\ &\le \int_0^1\int_{-\Lambda/2}^{\Lambda/2}\Big(\int_{x\in B_{2}(0)}\frac{|\gamma'(x+w)-\gamma'(x)|^4}{|w|^2}\,dx\Big)^{1/2}dw\,ds+C\\ &\le C(\|\gamma'\|^2_{B^{1/2}_{4,2}(B_\Lambda(0))}+1),\end{aligned}$$

again the interpolation Lemma A.3 gives the assertion. □

The final ingredient shows that the summands in $\widetilde{S}_3$ and S_3 are essentially the same.

Lemma 3.9. *Let $M_{3/2}\le 1$ and $\gamma(0)\in B_2(0)$. For all $\varepsilon>0$ we have*

$$\int_{B_1(0)}|\mathcal{H}|^2\,dx\le C\int_{B_{4\beta}(0)}|\partial_s^3\gamma|^2\,ds+(C_\varepsilon M_{3/2}^\alpha+\varepsilon)S_3+C_\varepsilon,$$

and hence in particular

$$\widetilde{S}_3\le C(S_3+1).$$

Furthermore, we have for all $\varepsilon>0$

$$\int_{B_1(0)}|\partial_s^3\gamma|^2\,dx\le C\int_{B_{4\beta}(0)}|\mathcal{H}\gamma|^2\,ds+(C_\varepsilon M_{3/2}^\alpha+\varepsilon)S_3+C_\varepsilon,$$

and hence in particular

$$S_3\le C\widetilde{S}_3+(C_\varepsilon M_{3/2}^\alpha+\varepsilon)S_3+C_\varepsilon$$

for some $C_\varepsilon<\infty$ depending ε and the bi-Lipschitz constant of γ. If $M_{3/2}$ is small enough, we have

$$S_3\le C(\widetilde{S}_3+1).$$

Proof. Lemma 3.5 tells us that

$$\Big|\int_{\mathbb{R}/\mathbb{Z}}|\mathcal{H}\gamma|^2\phi(\gamma)\,dx-\int_{\mathbb{R}/\mathbb{Z}}|Q\gamma|^2\phi(\gamma)\,dx\Big|\le C(M_{3/2}^\alpha+\varepsilon)S_3(x)+C_\varepsilon,$$

and hence in particular

$$\Big|\int_{B_1(x)}|\mathcal{H}\gamma|^2\,dx-\int_{B_1(0)}|Q\gamma|^2\,dx\Big|\le C(M_{3/2}^\alpha+\varepsilon)S_3(x)+C_\varepsilon. \tag{3-7}$$

Let $\psi\in C^\infty(\mathbb{R})$ be such that $\chi_{B_2(0)}\le\phi\le\chi_{B_4(0)}$. We get

$$\begin{aligned}\|\widetilde{Q}(\kappa)\|_{L^2(B_2(0))}&=\|\psi\widetilde{Q}(\kappa)\|_{L^2(B_2(0))}\\ &\le\|\widetilde{Q}[\psi\kappa]-\psi\widetilde{Q}[\kappa]-\kappa\widetilde{Q}[\psi]\|_{L^2(B_2(0))}+\|\widetilde{Q}[\psi\kappa]\|_{L^2(B_2(0))}+\|\kappa\widetilde{Q}[\psi]\|_{L^2(B_2(0))}.\end{aligned}$$

The commutator estimate (Lemma A.4) and the interpolation estimate (Lemma A.3) tell us that

$$\begin{aligned}\|\widetilde{Q}[\psi\kappa]-\psi\widetilde{Q}[\kappa]-\kappa Q[\psi]\|^2_{L^2(B_2(0))} &\leq C\|\kappa\|^2_{B^{1/2}_{4,2}(B_\Lambda(0))}\|\psi\|^2_{B^{1/2}_{4,2}B_\Lambda(0))}+C\sum_{j=1}^{\infty}\frac{\|\kappa\|^4_{L^4(B_{\Lambda+j}\setminus B_{\Lambda+j-1})}}{(\Lambda+j)^2}\\ &\leq \varepsilon S_3+C_\varepsilon.\end{aligned}$$

As by Lemma 2.1

$$\begin{aligned}\|\widetilde{Q}[\psi\kappa]\|_{L^2(B_2(0))} &\leq C\|\psi\kappa\|_{W^{1,2}(\mathbb{R}/l\mathbb{Z})}\leq C(\|\partial_s\psi\kappa\|_{L^2(\mathbb{R}/l\mathbb{Z})}+\|\psi\kappa\|_{L^2})\\ &\leq C(\|\partial^3\gamma\|_{L^2(B_4(0))}+\|\kappa\|_{L^2(B_2(0))})\end{aligned}$$

and

$$\|\kappa\widetilde{Q}[\psi]\|_{L^2(B_2(0))}\leq C\|\kappa\|_{L^2(B_2(0))},$$

we get using again the interpolation estimates

$$\|Q(\gamma)\|^2_{L^2(B_2(0))}=\|\widetilde{Q}(\kappa)\|^2_{L^2(B_2(0))}\leq C\|\partial_s^k\gamma\|_{L^2(B_4(0)}+\varepsilon S_3+C_\varepsilon. \tag{3-8}$$

The estimates (3-7) and (3-8) imply the first inequality. Summing up yields the second.

On the other hand, for a cutoff function $\psi\in C^\infty(\mathbb{R})$ such that $\chi_{B_{1/2}(0)}\leq\psi\leq\chi_{B_1(0)}$ we have

$$\|Q\gamma\|_{L^2(B_1(0))}\geq\|\psi Q\gamma\|_{L^2(B_1(0))},$$

which implies as above

$$\|Q\gamma\|_{L^2(B_1(0))}\geq\|Q(\psi\kappa)\|_{L^2}-\varepsilon S_3+C_\varepsilon.$$

Using Lemma 2.1 we get

$$\|\nabla(\psi\kappa)\|^2_{L^2}\leq\|Q(\psi\kappa)\|^2_{L^2}+\|\kappa\|^2_{L^2(B_1(0))}+\varepsilon S_3+C_\varepsilon$$

and hence using an interpolation estimate

$$\begin{aligned}\|\nabla\kappa\|^2_{L^2(B_{1/2}(0))} &\leq\|Q(\psi\kappa)\|^2_{L^2(B_1(0))}+C\|\kappa\|^2_{L^2(B_1(0))}+\varepsilon S_3+C_\varepsilon\\ &\leq C\|Q(\psi\kappa)\|^2_{L^2(B_1(0))}+\varepsilon S_3+C_\varepsilon.\end{aligned}$$

Using (3-7) we obtain

$$\|\nabla\kappa\|^2_{L^2(B_{1/2}(0))}\leq\|\mathcal{H}\gamma\|^2_{L^2(B_1(0))}+C(M^\alpha_{3/2}+\varepsilon)S_3(x)+C_\varepsilon,$$

and covering the ball $B_1(0)$ by balls of radius $\frac{1}{2}$ we get

$$\|\nabla\kappa\|^2_{L^2(B_1(0))}\leq\|\mathcal{H}\gamma\|^2_{L^2(B_2(0))}+C(M^\alpha_{3/2}+\varepsilon)S_3(x)+C_\varepsilon.$$

This implies the remaining three inequalities of the lemma. □

Gathering all the estimates above, we can now show:

Lemma 3.10 (differential inequality). *For $1>\varepsilon>0$ there is a constant $C_\varepsilon<\infty$ such that*

$$\frac{d}{dt}E_\phi(\gamma_t)+\int_{\mathbb{R}/l\mathbb{Z}}|\mathcal{H}\gamma_t|^2\phi\leq(C_\varepsilon M^\alpha_{3/2}(t)+\varepsilon)\widetilde{S}^{\text{ext}}_3(0,t)+C_\varepsilon$$

whenever $M_{3/2}$ is sufficiently small.

Proof. If $\gamma_t(\mathbb{R}/\mathbb{Z}) \cap B_2(0) = \varnothing$ we have

$$\frac{d}{dt} E_\phi(\gamma_t) = -\int_{\mathbb{R}/l\mathbb{Z}} |\mathcal{H}\gamma_t|^2 \phi,$$

since both sides of the equation are vanishing. Let us now assume that $\gamma(x_t, t) \in B_2(0)$ for some $x_t \in \mathbb{R}/\mathbb{Z}$. Then the Lemmas 3.5, 3.6, 3.7, 3.8, and 3.9 tell us that

$$\frac{d}{dt} E_\phi(\gamma_t) + \int_{\mathbb{R}/l\mathbb{Z}} |\mathcal{H}\gamma_t|^2 \phi \le (C_\varepsilon M^\alpha_{3/2}(t) + \varepsilon)\widetilde{S}_3(x_t, t) + C_\varepsilon.$$

It is an easy exercise to show using the bi-Lipschitz estimate that

$$\widetilde{S}_3(x_t, t) \le C \widetilde{S}^{\text{ext}}_3(0, t),$$

where the constant C depends on the bi-Lipschitz constant of γ. Hence,

$$\frac{d}{dt} E_\phi(\gamma_t) + \int_{\mathbb{R}/l\mathbb{Z}} |\mathcal{H}\gamma_t|^2 \phi \le (C_\varepsilon M^\alpha_{3/2}(t) + \varepsilon)\widetilde{S}^{\text{ext}}_3(0, t) + C_\varepsilon. \qquad \square$$

Exploiting this result, we get:

Proposition 3.11. *For every* $\delta > 0$ *there are constants* $\varepsilon_0 > 0$ *and* $C < \infty$ *such that* $\sup_{x\in\mathbb{R}^n} E_{B_1(x)}(\gamma_{t_0}) \le \varepsilon_0$ *for some* $t_0 \in [0, T)$ *implies*

$$\int_{t_0}^t \int_{\gamma_t^{-1}(B_1(x))} |\mathcal{H}\gamma_\tau|^2 \, ds \, d\tau \le C \quad \textit{and} \quad E_{B_1(x)}(\gamma_\tau) \le \delta$$

for all $\tau \in [t_0, \min\{T, t_0 + 1\})$ *and* $x \in \mathbb{R}^n$.

Proof. We assume without loss of generality that $t_0 = 0$. Clearly we only have to show the claim under the additional assumption that $\delta > 0$ is small. Furthermore, it is enough to show that

$$\int_{t_0}^t \int_{\gamma_t^{-1}(B_1(x))} |\partial_s^3 \gamma_\tau|^2 \, ds \, d\tau \le C \quad \text{and} \quad E_{B_1(x)}(\gamma_\tau) \le \delta$$

for all $\tau \in [t_0, \min\{T, t_0 + \varepsilon_2\})$ and $x \in \mathbb{R}^n$ for a sufficiently small ε_2. One then obtains the assertion in its original form by applying the preliminary result to the rescaled flow

$$\tilde{\gamma}(x, t) := \frac{1}{\sqrt[3]{\varepsilon_2}} \gamma\left(\frac{x}{\sqrt[3]{\varepsilon_2}}, \frac{t}{\varepsilon_2}\right),$$

which satisfies by a standard covering argument

$$E_{B_1(0)}(\tilde{\gamma}) \le C_n \frac{1}{(\varepsilon_2)^{n/3}} \varepsilon_1.$$

Lemma 3.10 tells us that

$$\frac{d}{dt} E_\phi(\gamma_t) + \int_{\mathbb{R}/l\mathbb{Z}} \int_{B_{1/4}(0)} |\mathcal{H}\gamma|^2 \phi \le (C_\varepsilon M^\alpha_{3/2} + \varepsilon) S^{\text{ext}}_3 + C_\varepsilon. \tag{3-9}$$

Let us assume that t_1 is the first time such that

$$\sup_{x\in\mathbb{R}^n} E_{B_1}(\gamma_{t_1}) \geq \delta.$$

We set

$$IM_3 := \sup_{x\in\mathbb{R}^n} \int_0^{t_1}\int_{\gamma^{-1}(B_{1/4}(x))} |\mathcal{H}\gamma|^2\, ds\, d\tau.$$

After a translation we can assume that

$$IM_3 = \int_0^{t_1}\int_{\gamma^{-1}(B_{1/4}(0))} |\mathcal{H}\gamma|^2\, ds\, d\tau.$$

Due to the definition of $\widetilde{S}_3(x)$ and Lemma 3.9 we know that

$$\int_{t_0}^{t_1} \widetilde{S}_3\, d\tau \leq C\int_0^{t_1} \widetilde{S}_3\, d\tau + C(M_{3/2}^{\alpha}+\varepsilon)\int_0^{t_1} \widetilde{S}_3\, d\tau + Ct_1 \leq CIM_3 + C(t_0-t_1).$$

Integrating (3-9) and using $\chi_{B_1(0)} \leq \phi \leq \chi_{B_2(0)}$ we hence get

$$E_\phi(\gamma_{t_1}) + c_0 IM_3 \leq E_\phi(\gamma_0) + C(\delta^\alpha+\varepsilon)IM_3 + C_\varepsilon t_1 \leq \varepsilon_0 + C(\delta^\alpha+\varepsilon)IM_3 + C_\varepsilon t_1. \tag{3-10}$$

If $C(\delta^\alpha+\varepsilon) \leq c_0/2$, this implies

$$\frac{c_0}{2} IM_3 \leq \varepsilon_0 + Ct.$$

Plugging this back into the inequality (3-10), we get for all $x \in \mathbb{R}^n$

$$E_{B_1(x)}(\gamma_{t_1}) \leq E_{\phi_x}(\gamma_{t_1}) \leq \varepsilon_0 + C(\delta^\alpha+\varepsilon)(\varepsilon_0+Ct) + C_\varepsilon t_1 < \delta$$

if we first choose $\varepsilon_0 > 0$ and then t small enough. □

3B. *Estimating the elastic energy.* In this section we derive estimates from the evolution equations of energies containing higher order terms. The following lemma was proven in [Blatt 2018]:

Lemma 3.12 (evolution of higher-order energies). *Let γ be a family of curves moving with normal speed V. Then*

$$\begin{aligned}\partial_t \int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa|^2 \phi\, ds &= 2\int_{\mathbb{R}/\mathbb{Z}} \langle \partial_s^{k+2} V, \partial_s^k\kappa\rangle \phi\, ds + 2\int \langle P_2^k(V,\kappa)\tau, \partial_s^{k+1}\kappa\rangle\phi\, ds \\ &\quad + 2\int \langle P_3^k(V,\kappa), \partial_s^k\kappa\rangle \phi\, ds - \int |\partial_s^k\kappa|^2 \langle \kappa, V\rangle \phi\, ds + \int_{\mathbb{R}/l\mathbb{Z}} |\partial_s^k\kappa| \nabla_V \phi\, ds. \end{aligned}\tag{3-11}$$

In this section, we will derive estimates for the right-hand side of (3-11) for the case that $V = -\mathcal{H}$. We use both the evolution equations from Lemma 3.12 and these estimates to bound the so-called *elastic energy* of the curve γ, i.e., the L^2-norm of its curvature.

Proposition 3.13 (estimate for the elastic energy). *Let $\gamma : [0,T) \times \mathbb{R}/\mathbb{R} \to \mathbb{R}^n$, $T > 1$, be a smooth solution of* (1-2). *There is an $\varepsilon_0 > 0$ depending only on n such that*

$$\sup_{(x,t)\in\mathbb{R}^n\times(0,1)} E_{B_1(x)}(\gamma(\,\cdot\,,t)) < \varepsilon_0$$

implies

$$\sup_{x\in\mathbb{R}/\mathbb{Z}}\int_{B_1(x)\cap\gamma_1}|\kappa_{\gamma_1}|^2\,ds\le C,\quad \textit{and}\quad \inf_{t\in[0,1]}\int_{B_1(x)}|\partial_s\kappa_t|^2\,ds\le C.$$

3B1. *Preliminary estimates.* To estimate the respective integrals appearing on the right-hand side of (3-11) we have to distinguish as before between $|w|$ big and $|w|$ small. The next lemma helps us to deal with the part where $|w|$ is big:

Lemma 3.14. *Let us assume that $p\in[1,\infty)$, $l_i\in\mathbb{N}$, $l_i\ge 2$, $p_i\in[1,\infty)$ for $i=1,\dots,r$, and let $\in\mathbb{N}$ be chosen such that*

$$l_i\le m,\quad l_i-\frac{1}{p_i}\le m-\frac{1}{2},\quad \textit{and}\quad \sum_{i=1}^{n}\frac{1}{p_i}=\frac{1}{p}.$$

For $\Lambda=1000\cdot 18e^{E(\gamma_0)/4}$ we set

$$g(x):=\int_{l\ge|w|\ge\Lambda}\int_{s\in[0,1]^r}\frac{\prod_{i=1}^r|\partial^{l_i}\gamma(x+s_iw)|}{|w|^2}\,ds\,dw$$

and assume that $M_{3/2}\le 1$. Then there is a constant $\beta_1,\beta_2>0$ such that

$$\|g\|^p_{L^p(B_1(0))}\le C(M_{3/2}^{\beta_1}+M_{3/2}^{\beta_2})\sum_{i=1}^{r}\left(\|\partial^m\gamma\|^{\theta_i}_{L^2(B_\Lambda(0))}+\sum_{j=1}^{\infty}\frac{\|\partial^m\gamma\|^{\theta_i}_{L^2(B_{\Lambda+j}(0)\setminus B_{\Lambda+j-1}(0))}}{(\Lambda+j)^2}\right)+C,$$

where

$$\theta_i=p\frac{l_i-\frac{1}{p_i}}{m-\frac{1}{2}}$$

and $C<\infty$ only depends on n and $E(\gamma_0)$.

Proof. Using Jensen's inequality, we obtain

$$\begin{aligned}\int_{B_1(0)}|g(x)|^p\,dx&=\int_{B_1(0)}\left(\int_{l\ge|w|\ge\Lambda}\int_{s\in[0,1]^r}\frac{\prod_{i=1}^r|\partial^{l_i}\gamma(x+s_iw)|}{|w|^2}\,ds\,dw\right)^p dx\\&\le C\int_{B_1(0)}\int_{l\ge|w|\ge\Lambda}\int_{s\in[0,1]^r}\frac{\prod_{i=1}^r|\partial^{l_i}\gamma(x+s_iw)|^p}{|w|^2}\,ds\,dw\,dx.\end{aligned}$$

As $\sum_{i=0}^r 1/p_i=1/p$, we get by Cauchy's inequality

$$\int_{B_1(0)}|g(x)|^p\,dx\le\frac{C}{\Lambda^{p-1}}\int_{B_1(0)}\int_{l\ge|w|\ge\Lambda}\int_{s\in[0,1]^r}\frac{\sum_{i=1}^r|\partial^{l_i}\gamma(x+s_iw)|^{p_i}}{|w|^2}\,ds\,dw\,dx.\tag{3-12}$$

We can estimate the summands further substituting $\tilde{w}=sw$ by

$$\begin{aligned}&\int_{B_1(0)}\int_{l\ge|w|\ge\Lambda}\int_0^1\frac{|\partial^{l_i}\gamma(x+s_iw)|^{p_i}}{|w|^2}\,ds\,dw\,dx\\&\quad\le\int_{B_1(0)}\int_0^1 s\int_{\Lambda\ge|\tilde{w}|\ge s\Lambda}\frac{|\partial^{l_i}\gamma(x+\tilde{w})|^{p_i}}{|\tilde{w}|^2}\,ds\,d\tilde{w}\,dx+\int_{B_1(0)}\int_0^1\int_{l\ge|\tilde{w}|\ge\Lambda}\frac{|\partial^{l_i}\gamma(x+\tilde{w})|^{p_i}}{|\tilde{w}|^2}\,ds\,d\tilde{w}\,dx\\&\quad\le C\|\partial^{l_i}\gamma\|^{p_i}_{L^{p_i}(B_\Lambda(0))}+\int_{l\ge|y|\ge\Lambda/2}\frac{|\partial^{l_i}\gamma(y)|^{p_i}}{|y|^2}\,dy.\end{aligned}\tag{3-13}$$

Applying a Gagliardo–Nirenberg-type inequality (Lemma A.3), we obtain for

$$\theta_i = p_i \frac{l_i - 1 - \frac{1}{p_i}}{m - \frac{1}{2}}$$

that

$$\|\partial^{l_i}\gamma\|^{p_i}_{L^{p_i}(B_1(0))} \leq C\|\partial^m \gamma\|^{\theta_i}_{L^2(B_1(0))} M_{3/2}^{(p_i-\theta_i)/2} + M_{3/2}^{p_i/2}.$$

Scaling this inequality, we get

$$\|\partial^{l_i}\gamma\|^{p_i}_{L^{p_i}(B_\Lambda(0))} \leq C(\|\partial^m \gamma\|^{\theta_i}_{L^2(B_\Lambda(0))} M_{3/2}^{(p_i-\theta)/2} + M_{3/2}^{p_i/2}).$$

Furthermore,

$$\begin{aligned}\int_{l\geq|y|\geq\Lambda/2} \frac{|\partial^{l_i}\gamma(y)|^{p_i}}{|y|^2}\,dy &\leq \frac{C}{\Lambda^2}\|\partial^{l_i}\gamma\|^{p_i}_{L^{p_i}(B_\Lambda(0))} + C\sum_{j=1}^{\infty} \frac{\|\partial^{l_i}\gamma\|^{p_i}_{L^{p_i}(A_{\Lambda+j}(0))}}{(\Lambda+j)^2} \\ &\leq CM_{3/2}^{(p_i-\theta)/2}\Big(\|\partial^{l_i}\gamma\|^{\theta_i}_{L^2(B_\Lambda(0))} + \sum_{j=1}^{\infty}\|\partial^{l_i}\gamma\|^{\theta}_{L^2(B_\Lambda(0))}\Big) + CM_{3/2}^{p_i/2}.\end{aligned}$$

From (3-12) and (3-13) the assertion follows. □

The second ingredient is the following lemma, which helps to deal with small $|w|$.

Lemma 3.15. *Let*

$$\begin{aligned} g(x) := \int_{|w|\leq\Lambda}\int_{s\in[0,1]^r}\int_{\tau_1,\tau_2\in[0,1]} \prod_{i=1}^{r} |\partial^{l_i}\gamma(x+s_i w)| \\ \times \frac{|\partial^{l_{r+1}}\gamma(x+\tau_1 w) - \partial^{l_{r+1}}\gamma(x+\tau_2 w)||\partial^{l_{r+2}}\gamma(x+\tau_1 w) - \partial^{l_{r+2}}\gamma(x+\tau_2 w)|}{w^2}\,ds\,d\tau_1\,d\tau_2\,dw.\end{aligned}$$

Let $\tilde{l}_i = l_i$ *for* $i = 1, \dots, r$ *and* $\tilde{l}_i = l_i + \frac{1}{2}$ *for* $i = r+1, r+2$. *If* $\tilde{l}_i \leq m$ *and* $\tilde{l}_i - \frac{1}{p_i} < m - \frac{1}{2}$ *for all* $i = 1, \dots, r+2$, *then there is a constant* $\alpha > 0$ *such that*

$$\|g\|_{L^p(B_\Lambda)} \leq (C_\varepsilon M_{3/2}^{\alpha} + \varepsilon)(\|\partial^m\gamma\|^{\theta}_{L^2(B_{2\Lambda}} + C_\varepsilon),$$

where

$$\theta = \frac{\sum_{i=1}^{r+2}(\tilde{l}_i - 1) - \frac{1}{p}}{m - \frac{3}{2}}.$$

Proof. We write

$$g(x) = \int_{I_l}\int_{s\in[0,1]^r}\iint_{\tau_1,\tau_2\in[0,1]} g_{x,\tau_1,\tau_2}(x,w)\,ds\,d\tau_1\,d\tau_2\,dw,$$

where

$$g_{s,\tau_1,\tau_2}(x,w) := \prod_{i=1}^{r}|\partial^{l_i}\gamma(x+s_i w)| \frac{|\partial^{l_{r+1}}\gamma(x+\tau_1 w) - \partial^{l_{r+1}}\gamma(x+\tau_2 w)||\partial^{l_{r+2}}\gamma(x+\tau_1 w) - \partial^{l_{r+2}}\gamma(x+\tau_2 w)|}{w^2}.$$

Using Hölder's inequality, we get for $|w| \leq \Lambda$

$$\|g_{s,\tau_1,\tau_2}(\,\cdot\,,w)\|_{L^p(B_\Lambda(0))} \leq \prod_{i=1}^{r}\|\partial^{l_i}\gamma\|_{L^{p_i}(B_{2\Lambda}(0))} \frac{\prod_{j=1,2}\|\partial^{l_{r+j}}\gamma(\,\cdot\,+(\tau_1-\tau_2)w) - \partial^{l_{r+j}}\gamma\|_{L^{p_{r+j}}(\Lambda)}}{w^2}.$$

Since

$$\begin{aligned}\iint_{\tau_1,\tau_2\in[0,1]}\int_{|w|\le\Lambda}&\frac{\prod_{j=1,2}\|\partial^{l_{r+j}}\gamma(\,\cdot\,+(\tau_1-\tau_2)w)-\partial^{l_{r+j}}\gamma\|^2_{L^{p_{r+j}}(\Lambda)}}{w^2}\,dw\,d\tau_1\,d\tau_2\\ &\le\iint_{\tau_1,\tau_2\in[0,1]}\prod_{j=1,2}\biggl(\frac{\int_{|w|\le\Lambda}\|\partial^{l_{r+j}}\gamma(\,\cdot\,+(\tau_1-\tau_2)w)-\partial^{l_{r+j}}\gamma\|^2_{L^{p_{r+j}}(\Lambda)}}{w^2}\,dw\biggr)^{1/2}d\tau_1\,d\tau_2\\ &\le\iint_{\tau_1,\tau_2\in[0,1]}|\tau_1-\tau_2|^{1/2}\prod_{j=1,2}\biggl(\frac{\int_{|\tilde w|\le\Lambda}\|\partial^{l_{r+j}}\gamma(\,\cdot\,+\tilde w)-\partial^{l_{r+j}}\gamma\|^2_{L^{p_{r+j}}(\Lambda)}}{\tilde w^2}\,d\tilde w\biggr)^{1/2}d\tau_1\,d\tau_2\\ &\le\prod_{j=1,2}\|\partial^{l_{r+j}}\gamma\|_{B(B_{3\Lambda}(p)},\end{aligned}$$

we obtain

$$\|g\|_{L^p(B_\Lambda(0))}\le\prod_{j=1,2}\prod_{i=1}^{r}\|\partial^{l_i}\gamma\|_{L^{p_i}(B_{2\Lambda}(0))}\|\partial^{l_{r+j}}\gamma\|_{B^{p_{r+j},2}_{1/2}(B_{4\Lambda}(0))}.$$

As above, the assertion follows from the Gagliardo–Nirenberg interpolation estimates in Lemma A.3. □

3B2. *Estimating the derivatives of $\mathcal{H}$.* For $k\in\mathbb{N}_0$, $s\in(0,1)$, we define

$$S_{k+s}(x)=\iint_{B_\Lambda(x)}\frac{|\partial^k\gamma(y)-\partial^k\gamma(z)|^2}{|y-z|^{1+2s}}\,dz\,dy+\sum_{j=1}^{\infty}\frac{1}{(\Lambda+j)^2}\iint_{B_{\Lambda+j}(x)\setminus B_{\Lambda+j-1}(x)}\frac{|\partial^k\gamma(y)-\partial^k\gamma(z)|^2}{|y-z|^{1+2s}}\,dz\,dy,$$

$$S_k(x)=\|\partial_s^k\gamma\|^2_{L^2(B_\Lambda(0)}+\sum_{j=1}^{\infty}\frac{\|\partial_s^k\gamma\|^2_{L^2(B_{\Lambda+j}(x)\setminus B_{\Lambda+j-1}(x))}}{(\Lambda+j)^2},$$

and

$$\tilde M_{k+3/2}=\tilde M^{\phi}_{k+3/2}(x)=\int_{\mathbb{R}/l\mathbb{Z}}\int_{-1}^{1}\frac{|\partial^{k+1}\gamma(y+w)-\partial^{k+1}\gamma(y)|^2}{w^2}\phi_x(\gamma(y))\,dw\,dy.$$

As before, we will assume that $\gamma(0)\in B_2(0)$ to get some preliminary estimates in terms of the intrinsically defined quantities above. In the final differential inequality we will use the extrinsic quantity

$$S^{\text{ext}}_{k+s}(x)=\iint_{\gamma^{-1}(B_1(x))}\frac{|\partial^k\gamma(y)-\partial^k\gamma(z)|^2}{|y-z|^{1+2s}}\,dz\,dy+\sum_{j=1}^{\infty}\frac{1}{j^2}\iint_{\gamma^{-1}(B_{j+1}(x)\setminus B_j(x))}\frac{|\partial^k\gamma(y)-\partial^k\gamma(z)|^2}{|y-z|^{1+2s}}\,dz\,dy$$

in place of $S_{k+s}(0)$.

We start with an estimate for $\widetilde{\mathcal{H}}$. Again we use the decomposition

$$\widetilde{\mathcal{H}}\gamma(x)=Q\gamma(x)+R_1\gamma(x)+R_2\gamma(x)=Q\gamma(x)+R\gamma(x),$$

where

$$\begin{aligned}Q\gamma(x)&=2\lim_{\varepsilon\downarrow0}\int_{I_{l,\varepsilon}}\biggl(2\frac{\gamma(x+w)-\gamma(x)-w\gamma'(x)}{w^4}-\frac{\kappa(x)}{|w|^2}\biggr)dw\\ &=4\lim_{\varepsilon\downarrow0}\int_{I_{l,\varepsilon}}\int_0^1(1-s)\frac{\kappa(x+sw)-\kappa(x)}{|w|^2}\,dw=\widetilde{Q}\kappa(x),\end{aligned}$$

$$R_1\gamma(x) = 4\int_{I_l} (\gamma(x+w)-\gamma(x)-w\gamma'(x))\left(\frac{1}{|\gamma(x+w)-\gamma(x)|^4}-\frac{1}{w^4}\right)dw,$$

$$R_2\gamma(x) = 2\int_{I_l} \kappa(x)\left(\frac{1}{w^2}-\frac{1}{|\gamma(x+w)-\gamma(x)|^2}\right)dw,$$

and set $R = R_1 + R_2$.

Lemma 3.16. *Let $M_{3/2} \le 1$ and $\gamma(0) \in B_2(0)$. For all $\tilde{k} \ge k$ there is a constant $\alpha > 0$ that for all $\varepsilon > 0$*

$$\|\partial_s^k P^{\perp}_{\gamma'}(R)\|_{L^2(B_1(0))} + \|\partial_s^k R\|^2_{L^2(B_1(0))} \le (C_\varepsilon M_{3/2}^{\alpha_1} + \varepsilon) S^{\theta}_{\tilde{k}+2}(0) + C_\varepsilon$$

for some constant $C_\varepsilon < \infty$, where $\theta = (2k+3)/(2\tilde{k}+1)$. Hence, for every $\varepsilon > 0$ and $k_1 > k$ there is a $C_\varepsilon < \infty$ such that

$$\|\partial_s^k \widetilde{\mathcal{H}} - \partial_s^k Q\|^2_{L^2(B_1(0))} \le \varepsilon S_{\tilde{k}+3} + C_\varepsilon.$$

Proof. First we will show that the two summands building R can be brought into a common form and can thus be dealt with simultaneously.

To this end we first use Taylor's theorem to rewrite

$$R_1\gamma(x) = 4\int_{I_l}\int_0^1 \kappa(x+sw)\left(\frac{1}{|\gamma(x+w)-\gamma(x)|^4}-\frac{1}{w^4}\right)ds\,dw.$$

For $\beta > 0$ we observe

$$\begin{aligned}\frac{1}{|\gamma(u+w)-\gamma(u)|^\beta}-\frac{1}{|w|^\beta} &= \frac{|w|^\beta}{|\gamma(u+w)-\gamma(u)|^\beta}\cdot\frac{1-|\gamma(u+w)-\gamma(u)|^\beta/|w|^\beta}{|w|^\beta}\\ &= G^{(\beta)}\left(\frac{\gamma(u+w)-\gamma(u)}{w}\right)\frac{2-2|\gamma(u+w)-\gamma(u)|^2/w^2}{|w|^\beta}\\ &= \int_0^1\int_0^1 G^{(\beta)}\left(\frac{\gamma(u+w)-\gamma(u)}{w}\right)\frac{|\gamma'(u+\tau_1 w)-\gamma'(u+\tau_2 w)|^2}{|w|^\beta}\,d\tau_1\,d\tau_2,\end{aligned}$$

where

$$G^{(\beta)}(z) := \frac{1}{2|z|^\beta}\cdot\frac{1-|z|^\beta}{1-|z|^2}$$

is an analytic function away from the origin. Defining

$$g^{(\alpha,\beta)}_{s_1,\tau_1,\tau_2}(u,w) := G^{(\beta)}\left(\frac{\gamma(u+w)-\gamma(u)}{w}\right)\frac{|\gamma'(u+\tau_1 w)-\gamma'(u+\tau_2 w)|^2}{|w|^\alpha}\kappa(u+s_1 w)$$

we thus get

$$R\gamma(x) = 4\int_{w\in I_l}\iint_{[0,1]^2}\int_0^1 g^{4,2}_{s_1,\tau_1,\tau_2}(x,w)\,d\tau_1\,d\tau_2\,ds\,dw - 2\int_{w\in I_l}\iint_{[0,1]^2} g^{2,2}_{0,\tau_1,\tau_2}(x,w)\,d\tau_1\,d\tau_2\,dw. \quad (3\text{-}14)$$

We now give the details of the estimate for the first term. The second term can be estimated analogously.

We differentiate under the integral to get

$$\begin{aligned}\partial^k R_1\gamma(x) &= 4\int_{w\in I_l}\iint_{[0,1]^2}\int_0^1 \partial_x^k g^{4,2}_{s_1,\tau_1,\tau_2}(x,w)\,d\tau_1\,d\tau_2\,ds\,dw\\ &= 4\int_{|w|\geq\Lambda}\iint_{[0,1]^2}\int_0^1 \partial_x^k g^{4,2}_{s_1,\tau_1,\tau_2}(x,w)\,d\tau_1\,d\tau_2\,ds\,dw\\ &\qquad + \int_{|w|\leq\Lambda}\iint_{[0,1]^2}\int_0^1 \partial_x^k g^{4,2}_{s_1,\tau_1,\tau_2}(x,w)\,d\tau_1\,d\tau_2\,ds\,dw.\end{aligned}$$

The product rule and Faà di Bruno's formula tell us that

$$\begin{aligned}&\partial_x^k g^{2,2}_{s_1,\tau_1,\tau_2}(x,w)\\ &= \sum_{l_1+l_2+l_3+l_4=k}\Bigl(\sum_{\pi\in\Pi_{l_1}}(\partial^{|\pi|}G^\beta)\Bigl(\frac{\gamma(x+w)-\gamma(x)}{w^2}\Bigr)\prod_{B\in\pi}\frac{\partial^{|B|}\gamma(x+w)-\partial^{|B|}\gamma(x)}{w}\Bigr)\\ &\qquad\times\frac{(\partial^{l_2+1}\gamma(x+\tau_1 w)-\partial^{l_2+1}\gamma(x+\tau_2 w))(\partial^{l_3+1}\gamma(x+\tau_1 w)-\partial^{l_3+1}\gamma(x+\tau_2 w))}{w^2}\partial^{l_4}\kappa(x+s_1 w),\end{aligned}$$

where Π_{l_1} denotes the set of all partitions of the set $\{1,\dots,l_1\}$. Using the fundamental theorem of calculus this can be brought into the form

$$\begin{aligned}&\partial_x^k g^{2,2}_{s_1,\tau_1,\tau_2}(x,w)\\ &= \sum_{l_1+l_2+l_3+l_4=k}\Bigl(\sum_{\pi\in\Pi_{l_1}}(\partial^{|\pi|}G^\beta)\Bigl(\frac{\gamma(x+w)-\gamma(x)}{w^2}\Bigr)\prod_{B\in\pi}\int_0^1\partial^{|B|+1}\gamma(x+s_B w)\,ds_B\Bigr)\\ &\qquad\times\frac{(\partial^{l_2+1}\gamma(x+\tau_1 w)-\partial^{l_2+1}\gamma(x+\tau_2 w))(\partial^{l_3+1}\gamma(x+\tau_1 w)-\partial^{l_3+1}\gamma(x+\tau_2 w))}{w^2}\partial^{l_4}\kappa(x+s_1 w).\end{aligned}$$

We choose $p_B = p_i = (\#\pi+4)p$ and observe that

$$|B|+1-\tfrac{1}{p_i}\leq |B|+1\leq k+1\leq k+2-\tfrac{1}{2}$$

and

$$l_i+\tfrac{3}{2}-\tfrac{1}{2}\leq l_i+\tfrac{3}{2}\leq k+\tfrac{3}{2}\leq k+2-\tfrac{1}{2}.$$

Hence, we can apply Lemmas 3.14 and 3.15 to get an estimate as claimed for each of the summands with

$$\theta\leq\frac{(l_1+l_2+l_3+l_4+5-3)-\frac{1}{2}}{\tilde{k}-\frac{3}{2}}.$$

Using the identity

$$P^{\perp}_{\gamma'}(R)=R-\langle R,\gamma'\rangle\gamma'$$

and treating the second term in this difference in the same way as above, we get the second estimate in the assertion. □

Lemma 3.17. *We have*

$$\|\partial_s^k\mathcal{H}-\partial_s^k Q\|^2_{L^2(B_{2\beta}(0))}\leq (C_\varepsilon M^\alpha_{3/2}+\varepsilon)(S_{k+3}+C_\varepsilon)$$

for suitable constants $\alpha>0$ *and*

$$\|\partial_s^k\mathcal{H}\|_{L^2(B_{2\beta}(0))}\leq (C_\varepsilon M^\alpha_{3/2}+\varepsilon)S_{k+3}+C_\varepsilon(M^\beta_2+1).$$

Proof. We use

$$\mathcal{H}\gamma(x) = P^{\perp}_{\gamma'}\widetilde{\mathcal{H}}\gamma(x) = \widetilde{\mathcal{H}}\gamma(x) - \langle \widetilde{\mathcal{H}}\gamma(x), \gamma' \rangle \gamma'$$

together with the decomposition

$$\widetilde{\mathcal{H}} = Q + R$$

to write

$$\mathcal{H} = Q - P^{T}_{\gamma'}Q + P^{\perp}_{\gamma'}R,$$

where $P^{T}_{\gamma'}$ denotes the projection onto the tangential part.

Lemma 3.16 tells us that

$$\|\partial_s^k P^{\perp}_{\gamma'}R\|^2_{L^2} \leq C(M^{\alpha}_{3/2} + \varepsilon)S_{k+3} + C_{\varepsilon}.$$

To deal with the term containing Q we use $P^{T}_{\gamma'}Q = \langle Q\gamma, \gamma' \rangle \gamma'$. Leibniz's rule yields

$$\partial_s^k(\langle Q, \gamma' \rangle \gamma') = \langle \partial_s^k Q, \gamma' \rangle \gamma' + I_1,$$

where I_1 is a linear combination of terms

$$Q[\partial_s^{k_1}\gamma]\partial_s^{k_2}\gamma'\partial_s^{k_3}\gamma',$$

with $k_1, k_2, k_3 \in \mathbb{N}_0$, $k_1 + k_2 + k_3 = k$, and $k_2 + k_3 \geq 1$. By Hölder's inequality the L^2-norm over $B_{2\beta}(0)$ of all these terms can be estimated by

$$C\|Q\partial_s^{k_1}\gamma\|_{L^2(B_{2\beta})}\|\partial_s^{k}\gamma'\|_{L^4(B_{2\beta})}\|\partial_s^{k_3}\gamma'\|_{L^4(B_{2\beta})}.$$

As in the proof of Lemma 3.9 we see that

$$\|Q\partial_s^{k_1}\gamma\|_{L^2(B_{2\beta}(0)} \leq C\|\partial_s^{3+k_1}\gamma\|_{L^2(B_{4\beta}(0))} + \varepsilon S_{k_1+3} + C_{\varepsilon}.$$

Hence, the interpolation estimates give

$$I_1 \leq \varepsilon S_{k+7/2} + C_{\varepsilon}.$$

We now pick up the argument from the proof of Lemma 3.5 to estimate the term

$$\langle \partial_s^k Q\gamma, \gamma' \rangle \gamma'.$$

Using the linearity of Q, we can rewrite

$$\langle \partial_s^k Q, \gamma' \rangle = \sum_{i=1}^{n} \langle \widetilde{Q}[\partial_s^k \kappa_i]\gamma_i' - \widetilde{Q}[\partial_s^k \kappa_i \gamma_i'].$$

From Lemma A.4 we then get

$$\begin{aligned}
&\|\langle \partial_s^k Q\gamma, \gamma' \rangle \gamma'\|_{L^2(B^2(0))} \\
&\quad \leq C\left(\|\partial_s^k \kappa\|_{B^{1/2}_{4,2}(B_3(0))}\|\gamma'\|_{B^{1/2}_{4,2}(B_3(0))} + \sum_{j \in \mathbb{N}} \frac{\|\partial_s^k \kappa\|^2_{L^4}}{(\Lambda + j)^2} + 1\right) + C\|\partial_s^k \kappa\|_{L^2(B_3(0))}(\|\gamma'\|_{C^{0,1}(B_3(0))} + 1) \\
&\quad \leq (C_{\varepsilon}M^{1/2}_{3/2} + \varepsilon)S^{1/2}_{k+3} + C_{\varepsilon}. \qquad \square
\end{aligned}$$

3B3. *Estimate of the highest-order term.*

Lemma 3.18. *If $M_{3/2} \le 1$ and $\gamma(0) \in B_2(0)$, we have*

$$-\int_{\mathbb{R}/l\mathbb{Z}} \langle \partial_s^{k+2}\mathcal{H}\gamma, \partial_s^k \kappa\rangle \phi \, ds \le -\tilde{M}_{k+7/2}(0) + C M_{3/2}^{\alpha} S_{k+7/2}(0).$$

Proof. The main strategy is to use partial integration to move $1+\frac{1}{2}$ derivatives from the term $\partial_s^{k+2}\mathcal{H}$ to the term $\partial_s^k\kappa$. But first we want to get rid of the projection onto the normal part contained in the definition of $\mathcal{H}$. We have

$$\mathcal{H}\gamma(x) = P^{\perp}_{\gamma'(x)}(Q\gamma(x) + R\gamma(x)).$$

Let us first deal with the terms containing R. Integration by parts gives

$$\begin{aligned} -\int_{\mathbb{R}/l\mathbb{Z}} &\partial_s^{k+2}(P^{\perp}_{\gamma'(x)}(R\gamma(x)))\partial_s^k\kappa\phi\, ds \\ &= \int_{\mathbb{R}/l\mathbb{Z}} \partial_s^{k+1}(P^{\perp}_{\gamma'(x)}(R\gamma(x)))\partial_s^{k+1}\kappa\phi\, ds + \int_{\mathbb{R}/l\mathbb{Z}} \partial_s^{k+1}(P^{\perp}_{\gamma'(x)}(R\gamma(x)))\partial_s^{k}\kappa\phi'\, ds, \end{aligned}$$

which we can estimate using the product rule, Hölder's inequality, and Lemma 3.16 by

$$(C M_{3/2}^{\alpha} + \varepsilon) S_{k+7/2}^{(2(k+1)+3)/(4(k+2))} \|\partial_s^k\kappa\|_{L^2} \le (C M_{3/2}^{\alpha} + \varepsilon) S_{k+7/2} + C_{\varepsilon}.$$

So we get

$$-\int_{\mathbb{R}/l\mathbb{Z}} \partial_s^{k+2}(P^{\perp}_{\gamma'(x)}(R\gamma(x)))\partial_s^k\kappa\phi\, ds \le (C M_{3/2}^{\alpha} + \varepsilon) S_{k+7/2} + C_{\varepsilon}. \tag{3-15}$$

To estimate

$$\int_{\mathbb{R}/l\mathbb{Z}} \langle \partial_s^{k+2} P^{\perp}_{\gamma'} Q, \partial_s^k\kappa\rangle \phi\, ds$$

we write

$$P^{\perp}_{\gamma'(x)} Q\gamma(x) = Q\gamma(x) - \langle Q\gamma(x), \gamma'(x)\rangle \gamma'(x) = Q\gamma - P^{T}_{\gamma'} Q.$$

Using

$$\langle Q\gamma(x), \gamma'(x)\rangle = 2\int_{-l}^{l}\int_0^1 (1-s)\frac{(\gamma'(x+sw)-\gamma'(x))(\kappa(x+sw)-\kappa(x))}{w^2}\, ds\, dw\, dx$$

we get from Lemmas 3.14 and 3.15

$$\|\partial_s^{k+1} P^{T}_{\gamma'}(Q)\|^2_{L^2(B_{2\beta}(0))} \le C(M_{3/2}^{\alpha} + \varepsilon) S_{k+7/2}^{(2k+3)/(2(2k+4))}.$$

Hence, Cauchy's inequality implies

$$-\int_{\mathbb{R}/l\mathbb{Z}} \partial_s^{k+1} P^{T}_{\gamma'} Q\gamma\, \partial_s^{k+1}\kappa\phi\, ds \le \varepsilon S_{k+7/2} + C_{\varepsilon}. \tag{3-16}$$

The term

$$\int_{\mathbb{R}/l\mathbb{Z}} \langle \partial_s^{k+2} Q, \partial_s^k\kappa\rangle\, ds$$

can be rewritten using (2-3) as

$$\begin{aligned}
&\int_{\mathbb{R}/l\mathbb{Z}} \langle \partial_s^{k+2} Q\gamma, \partial^k \kappa\phi\rangle\, ds \\
&= \int_{\mathbb{R}/l\mathbb{Z}} \langle Q\partial^k \kappa, \partial^k \kappa\rangle \phi\, ds \\
&= 2\int_{\mathbb{R}/l\mathbb{Z}} \int_{-l}^{l} \int_0^1 (1-s) \frac{|\partial_s^{k+1}\kappa(x+sw) - \partial_s^{k+1}\kappa(x)|^2}{w^2} \phi(\gamma(x))\, ds\, dw\, dx \\
&\quad + 2\int_{\mathbb{R}/l\mathbb{Z}} \int_{-l}^{l} \int_0^1 (1-s) \frac{(\partial_s^{k+1}\kappa(x+sw) - \partial_s^{k+1}\kappa(x))(\phi(\gamma(x+sw)) - \phi(\gamma(x)))}{w^2} \partial_x^{k+1}\kappa(x+sw)\, ds\, dw\, dx.
\end{aligned}$$

We observe that

$$\begin{aligned}
&\int_{\mathbb{R}/l\mathbb{Z}} \int_{-l}^{l} \int_0^1 (1-s) \frac{|\partial_s^{k+1}\kappa(x+sw) - \partial_s^{k+1}\kappa(x)|^2}{w^2} \phi(x)\, ds\, dw\, dx \\
&\qquad\geq \int_{B_1(0)} \int_{-l}^{l} \int_0^1 (1-s) \frac{|\partial_s^{k+1}\kappa(x+sw) - \partial_s^{k+1}\kappa(x)|^2}{w^2}\, ds\, dw\, dx \\
&\qquad\geq \int_{B_1(0)} \int_0^1 (1-s)s \int_{-sl}^{sl} \frac{|\partial_s^{k+1}\kappa(x+\tilde{w}) - \partial_s^{k+1}\kappa(x)|^2}{\tilde{w}^2}\, d\tilde{w}\, ds\, dx \\
&\qquad\geq c_0 \int_{B_1(0)} \int_{-l/2}^{l/2} \frac{|\partial_s^{k+1}\kappa(x+\tilde{w}) - \partial_s^{k+1}\kappa(x)|^2}{\tilde{w}^2}\, d\tilde{w}\, dx \geq c_0 \tilde{M}_{k+7/2}(0).
\end{aligned}$$

Furthermore, we take the decomposition

$$\begin{aligned}
&\left| \int_{\mathbb{R}/l\mathbb{Z}} \int_{-l}^{l} \int_0^1 (1-s) \frac{(\partial_s^{k+1}\kappa(x+sw) - \partial_s^{k+1}\kappa(x))(\phi(\gamma(x+sw)) - \phi(\gamma(x)))}{w^2} \partial_x^{k+1}\kappa(x+sw)\, ds\, dw\, dx \right| \\
&\qquad \leq \int_{B_\Lambda(0)} \int_{|w|\geq\Lambda} \int_0^1 (1-s) \frac{|\partial_s^{k+1}\kappa(x+sw) - \partial_s^{k+1}\kappa(x)|}{w^2} |\partial_s^{k+1}\kappa(x+sw)|\, ds\, dw\, dx \\
&\qquad\qquad + \int_{B_\Lambda(0)} \int_{|w|\leq\Lambda} \int_0^1 (1-s) \frac{|\partial_s^{k+1}\kappa(x+sw) - \partial_s^{k+1}\kappa(x)|}{|w|} |\partial_s^{k+1}\kappa(x+sw)|\, ds\, dw\, dx.
\end{aligned}$$

We use Lemmas 3.14 and 3.15 to estimate the first term and the second term by

$$S_{k+7/2}^{\theta} M_{3/2}^{1-\theta} + M_{3/2},$$

where

$$\theta = \frac{2k+3+\frac{1}{2}}{2k+4} < 1.$$

Hence, Cauchy's inequality yields

$$-\int_{\mathbb{R}/l\mathbb{Z}} \langle \partial_s^{k+2} Q\gamma, \partial^k \kappa\phi\rangle\, ds \leq -c_0 M_{k+7/2}(0) + \varepsilon S_{k+7/2} + C_\varepsilon. \tag{3-17}$$

The inequalities (3-15), (3-16), and (3-17) prove the statement of the lemma. ☐

Lemma 3.19 (differential inequality for energies of higher order). *For every $\varepsilon > 0$ there is a constant $C_\varepsilon < \infty$ depending only on ε, n, and k and $c_k > 0$ such that*

$$\partial_t \int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa|^2 \phi\, ds + c_k \tilde{M}_{k+7/2} \leq C(M_{3/2}^{\alpha} + \varepsilon) S_{k+7/2}^{\mathrm{ext}}(0) + C_\varepsilon.$$

Proof. From (3-11) we get

$$\partial_t \int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa|^2 \phi\, ds = 2 \int_{\mathbb{R}/\mathbb{Z}} \langle \partial_s^{k+2} V, \partial_s^k \kappa \rangle \phi\, ds + 2 \int \langle P_2^k(V,\kappa)\tau, \partial_s^{k+1}\kappa \rangle \phi\, ds \\ + 2 \int \langle P_3^k(V,\kappa), \partial_s^k \kappa \rangle \phi\, ds - \int |\partial_s^k \kappa|^2 \langle \kappa, V \rangle \phi\, ds + \int_{\mathbb{R}/l\mathbb{Z}} |\partial_s^k \kappa| \nabla_V \phi\, ds. \quad (3\text{-}18)$$

Lemma 3.18 gives

$$2 \int_{\mathbb{R}/\mathbb{Z}} \langle \partial_s^{k+2} V, \partial_s^k \kappa \rangle \phi\, ds \leq -\tilde{M}_{k+7/2}(0) + (M_{3/2}^{\alpha} + \varepsilon)\varepsilon S_{k+3/2} + C_\varepsilon.$$

Let $k_1 + k_2 = k$. Hölder's inequality, standard interpolation estimates and Lemma 3.17 give

$$\int_{\mathbb{R}/l\mathbb{Z}} \partial_s^{k_1} V * \partial_s^{k_2} \kappa * \tau * \partial_s^{k+1} \kappa \phi\, ds \leq \|\partial_s^{k_1} V\|_{L^2} \|\partial_s^{k_2} \kappa\|_{L^4} \|\partial_s^{k+1}\kappa\|_{L^4} \\ \leq C(M_{3/2}^{\alpha} + \varepsilon) S_{k+7/2} + C_\varepsilon,$$

and hence

$$\int \langle P_2^k(V,\kappa)\tau, \partial_s^{k+1}\kappa \rangle \phi\, ds \leq C(M_{3/2}^{\alpha} + \varepsilon) S_{k+7/2} + C_\varepsilon.$$

Similarly we get the estimate

$$\int \langle P_3^k(V,\kappa), \partial_s^k \kappa \rangle \phi\, ds \leq C(M_{3/2}^{\alpha} + \varepsilon) S_{k+7/2} + C_\varepsilon,$$

and

$$\int_{\mathbb{R}/l\mathbb{Z}} |\partial_s^k \kappa|^2 \nabla_V \phi\, ds \leq \|\partial_s^k \kappa\|_{L^4}^2 \|V\|_{L^2} \leq (C_\varepsilon M_{3/2}^{\alpha} + \varepsilon) S_{k+7/2} + C_\varepsilon.$$

Hence, we have

$$2 \int \langle P_2^k(V,\kappa)\tau, \partial_s^{k+1}\kappa \rangle \phi\, ds + 2 \int \langle P_3^k(V,\kappa), \partial_s^k \kappa \rangle \phi\, ds \\ - \int |\partial_s^k \kappa|^2 \langle \kappa, V \rangle \phi\, ds + \int_{\mathbb{R}/l\mathbb{Z}} |\partial_s^k \kappa| \nabla_V \phi\, ds \leq 5\varepsilon S_{k+7/2} + C_\varepsilon (M_2^{\beta} + 1).$$

Together, these estimates imply

$$\partial_t \int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa|^2 \phi\, ds + \tilde{M}_{k+7/2}(0) \leq C(M_{3/2}^{\alpha} + \varepsilon) S_{k+7/2} + C_\varepsilon.$$

As $S_{k+7/2} \leq C S_{k+7/2}^{\mathrm{ext}}$ we get the assertion. □

3B4. *Proof of Proposition 3.13.* We get from Lemma 3.19

$$\partial_t \int_{\mathbb{R}/\mathbb{Z}} |\kappa|^2 \phi\, ds + c_0 \tilde{M}_{7/2}(0) \le (C_\varepsilon M_{3/2}^\alpha + \varepsilon)(S_{7/2}^{\text{ext}}(x) + C_\varepsilon).$$

Integrating this inequality and using that $M_{3/2} \le \varepsilon_0 < 1$ we get

$$\begin{aligned}\int_{\mathbb{R}/l\mathbb{Z}} |\kappa_{\gamma_1}|^2 \phi\, ds + c_0 \int_\tau^1 \tilde{M}_{7/2}(0,t)\, dt &\le \int_{\mathbb{R}/l\mathbb{Z}} |\kappa_{\gamma_\tau}|^2 \phi\, ds + (C_\varepsilon \varepsilon_0^\alpha + \varepsilon) \int_\tau^1 S_{7/2}^{\text{ext}}(x,t)\, dt + C_\varepsilon(1-\tau) \\ &\le \int_{\mathbb{R}/l\mathbb{Z}} |\kappa_{\gamma_\tau}|^2\, ds + (C_\varepsilon \varepsilon_0^\alpha + \varepsilon) \int_\tau^1 S_{7/2}^{\text{ext}}(x,t)\, dt + C_\varepsilon(1-\tau). \quad (3\text{-}19)\end{aligned}$$

Integrating again over $\tau \in \left[0, \frac{1}{2}\right]$ yields

$$\begin{aligned} c_0 \int_0^{1/2}\int_\tau^1 \tilde{M}_{7/2}(0,t)\, dt &\le \int_0^{1/2}\int_{\mathbb{R}/l\mathbb{Z}} |\kappa_{\gamma_\tau}|^2 \phi(\gamma)\, ds + (C_\varepsilon \varepsilon_0^\alpha + \varepsilon) \int_0^{1/2}\int_\tau^1 S_{7/2}^{\text{ext}}(x,t)\, dt\, d\tau + C_\varepsilon(1-\tau) \\ &\le \int_0^{1/2}\int_{\mathbb{R}/l\mathbb{Z}} |\kappa_{\gamma_\tau}|^2 \phi(\gamma)\, ds + (C_\varepsilon \varepsilon_0^\alpha + \varepsilon) \int_0^{1/2}\int_\tau^1 S_{7/2}^{\text{ext}}(x,t)\, dt + C_\varepsilon.\end{aligned}$$

We can estimate the first term, using interpolation estimates as in Section 3A, by

$$C\left(\int_0^{1/2} \widetilde{S}_3^{\text{ext}}\, dt + 1\right) \le C(IM_3 + 1),$$

which is bounded by Proposition 3.11. Assuming that

$$IM_{7/2} = \sup_{x\in\mathbb{R}^n} \int_0^{1/2}\int_{t_1}^1 \tilde{M}_{7/2}(x,t)\, d\tau\, dt = \int_0^{1/2}\int_\tau^1 \tilde{M}_{7/2}(0,t)\, dt,$$

and using that

$$\sup_{x\in\mathbb{R}/l\mathbb{Z}} \int_0^{1/2}\int_\tau^1 S_{7/2}^{\text{ext}}(x,t)\, d\tau\, dt = \int_0^{1/2}\int_\tau^1 \tilde{M}_{7/2}(0,t)\, d\tau\, dt \le C\, IM_{7/2},$$

we deduce

$$c_0 IM_{7/2} \le C + (C_\varepsilon \varepsilon_0^\alpha + \varepsilon) IM_{7/2} + C_\varepsilon.$$

Choose first $\varepsilon > 0$ and then $\varepsilon_0 > 0$ sufficiently small; then we get

$$IM_{7/2} \le C.$$

Plugging this back into (3-19) we get the assertion

3C. ***Estimates for higher-order energies.*** It is tempting to just iterate the above argument to get control of higher-order energies. Unfortunately, one would have to adapt ε_1 in each of the steps which would not yield to the desired result. Instead we improve the differential estimate from the end of the last subsection assuming that M_2 is finite. By literally the same argument as in the proof of the Lemmas in the last subsection but interpolating in all the arguments between $W^{k+7/2,2}$ and $W^{2,2}$ instead of $W^{k+7/2,2}$ and $W^{3/2,2}$ we get:

Lemma 3.20 (differential inequality for energies of higher order). *For every $\varepsilon > 0$ there is a constant C_ε depending only on ε, n, and k and a constant $c_k > 0$ such that*

$$\partial_t \int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa|^2 \phi \, ds + c_k M_{k+7/2}(0) \leq \varepsilon S^{\text{ext}}_{k+7/2} + C_\varepsilon M_2^\beta .$$

Now we are finally able to conclude the proof of the ε-regularity theorem. We prove inductively the following statement

Proposition 3.21. *There is an $\varepsilon_0 > 0$ and constant $C_k < \infty$ such that*

$$\sup_{x\in\mathbb{R}^n} E_{B_1(x)}(\gamma_0) \leq \varepsilon_0$$

implies

$$\sup_{x\in\mathbb{R}^n} \|\partial_s^k \kappa_{\gamma(t)}\|_{L^2(B_1(x))} \leq \frac{C}{t^{k/3-1/2}}.$$

Proof. We prove by induction on k that

$$\|\partial_s^k \kappa_{\gamma(t)}\|_{L^2(B_1(x))} \leq \frac{C_k}{t^{k/3-1/2}}$$

and

$$\sup_{x\in\mathbb{R}^n} \int_{t/2}^t M_{k+7/2}(x,t)\, dt \leq \frac{C_k}{t^{k/3-3/2}}.$$

Again by scaling properties of the solution it is enough to show these inequalities for $t = 1$. Let us fix $\varepsilon_0 > 0$ such that we can apply Propositions 3.11 and 3.13, i.e., such that the Möbius energies on balls of radius 1 are small and the elastic energy on unit balls is bounded for times larger than $t = \frac{1}{4}$. Hence, the statement is true for $k = 0$.

Let us assume that we have the bound claimed for $k - 1$ and let $t = \frac{1}{2}$. By Lemma 3.20 for every $\varepsilon > 0$ we have

$$\partial_t \int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa|^2 \phi \, ds + c_k \tilde{M}_{7/2+k}(x) \leq \varepsilon S^{\text{ext}}_{k+7/2} + C_\varepsilon. \tag{3-20}$$

Integrating the inequality, we get for all $0 < \tau < 1$ that

$$\int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa_1|^2 \phi \, ds + c_k \int_\tau^1 \tilde{M}_{7/2+k}(x,t)\, dt \leq \int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa_\tau|^2 \phi \, ds + \varepsilon \int_\tau^1 S^{\text{ext}}_{k+7/2}(x,t)\, dt + C_\varepsilon (1-t).$$

We integrate this inequality for $\tau \in \left[\frac{1}{4}, \frac{1}{2}\right]$ to get

$$\frac{1}{4} \int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa_1|^2 \phi \, ds + c_k \int_{1/4}^{1/2} \int_\tau^1 \tilde{M}_{7/2+k}(x,t)\, dt\, d\tau \\ \leq \int_{1/4}^{1/2} \int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa_\tau|^2 \phi \, ds\, dt\, d\tau + \varepsilon \int_{1/4}^{1/2} \int_\tau^1 S^{\text{ext}}_{k+7/2}(x,t)\, dt + C_\varepsilon .$$

Since interpolation estimates yield

$$\int_{1/4}^{1/2} \int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa_\tau|^2 \phi \, ds\, dt\, d\tau \leq C \int_{1/4}^{1/2} (S^{\text{ext}}_{k+3/2} + 1)\, dt \leq C \sup_{x\in\mathbb{R}^n} \int_{1/4}^{1/2} (M_{(k-1)+7/2}(x,t) + 1)\, dt \leq C,$$

by the induction hypotheses, we deduce that

$$\frac{1}{4}\int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa_1|^2 \phi \, ds + c_k \int_{1/4}^{1/2}\int_{\tau}^{1} \tilde{M}_{7/2+k}(x,t)\, dt\, d\tau \le C + \varepsilon \int_{1/4}^{1/2}\int_{\tau}^{1} S^{\text{ext}}_{k+7/2}(x,t)\, dt + C_\varepsilon. \tag{3-21}$$

Let us now assume that the supremum

$$IM_{k+7/2} = \sup_{x\in\mathbb{R}^n} \int_{1/4}^{1/2}\int_{\tau}^{1} \tilde{M}_{k+7/2}(x,t)\, dt$$

is attained in the point $x = 0$. Since

$$\int_{1/4}^{1/2}\int_{\tau}^{1} S^{\text{ext}}_{k+7/2}(x,t)\, dt \le C\, IM_{k+7/2},$$

we deduce from (3-21)

$$IM_{k+7/2} \le \varepsilon IM_{k+7/2} + C_\varepsilon.$$

Choosing $\varepsilon > 0$ small enough and absorbing, we get

$$IM_{k+7/2} \le C.$$

Hence, in particular

$$\int_{1/2}^{1} S^{\text{ext}}_{k+7/2}(x,t)\, dt \le C \quad \text{for all } x \in \mathbb{R}^n.$$

Plugging this back into (3-21), we derive

$$\int_{\mathbb{R}/\mathbb{Z}} |\partial_s^k \kappa(s,1)|^2 \phi_x \, ds \le C \quad \text{for all } x \in \mathbb{R}^n. \qquad \square$$

Proof of Theorem 3.1. Using scaled Sobolev embeddings we get the claimed estimates from Proposition 3.21 as long as the flow exists. So the only thing left is to show that $T > 1$. But this follows by standard methods from the uniform estimates in Proposition 3.21. □

4. Applications

4A. *Blow-up profiles.* Using Theorem 3.1, we get the following classification of finite time blow-up.

Theorem 4.1 (characterization of singularities). *Let $\gamma \in C^\infty([0,T) \times \mathbb{R}/\mathbb{Z}, \mathbb{R}^n)$ be a maximal smooth solution of* (1-2). *There is a constant $\varepsilon_0 > 0$ depending only on n and $E(\gamma_0)$ such that if $T < \infty$ there are times $t_k \uparrow T$, points $x_k \in \mathbb{R}^n$, and radii $r_k \downarrow 0$ with*

$$E_{B_{r_k}(x_k)}(\gamma_{t_k}) \ge \varepsilon_0.$$

Proof. Let us assume that $T < \infty$ and that there is an $r > 0$ such that for all $t \in [0,T)$ and all $x \in \mathbb{R}^n$ we have

$$E_{B_{r_j}(x_j)}(\gamma(t)) \le \varepsilon_0.$$

Then Theorem 3.1 would tell us that $T > t_j + r_j^3 \to T + r^3$. □

Picking the concentration times more carefully, we can construct a blow-up profile at a singularity. As mentioned in the Introduction, we localize the energy intrinsically for this purpose; i.e., we work with $E^{\text{int}}_{B_r(x)}$ instead of $E_{B_r(x)}$. We do this for the simple reason that $E^{\text{int}}_{B_r(x)}$ is continuous in r and x.

In the rest of this article we will express from time to time the integrals occurring as integrals over the image

$$\Gamma_t := \gamma(\mathbb{R}/l\mathbb{Z}, t).$$

Theorem 4.2 (blow-up profiles). *There is an $\varepsilon_0 > 0$ such that the following holds: Assume that γ_t is a solution of* (1-2) *that develops a singularity in finite time, i.e., $T < \infty$ and $r_j \to 0$. Then there are points x_j and times $t_j \to T$ such that*

$$E^{\text{int}}_{B_{r_j}(x_j)}(t_j) \geq \varepsilon_0.$$

Let us now choose the points $x_j \in \mathbb{R}$ and times $t_j \in [0, T)$ such that

$$\sup_{\tau \in [0,t_j], x \in \Gamma_\tau} E^{\text{int}}_{B_{r_j}(x)}(\gamma_{t_j}) \leq E^{\text{int}}_{B_{r_j}(x)}(\gamma_{t_j}) = \varepsilon_0,$$

and let $\tilde{\gamma}_j$ be reparametrizations by arc-length of the rescaled and translated curves

$$r_j^{-1}(\gamma_{t_j} - x_j)$$

such that $\tilde{\gamma}_j(0) \in B_2(0)$. Then these curves subconverge locally in C^∞ to an embedded closed or open curve $\tilde{\gamma}_\infty : I \to \mathbb{R}^n$, $I = \mathbb{R}/l\mathbb{Z}$ or $I = \mathbb{R}$ resp., parametrized by arc-length. This curve satisfies

$$\text{p.v.} \int_{-l/2}^{l/2} \left(2 \frac{P^\perp_\tau(\tilde{\gamma}(y) - \tilde{\gamma}(x))}{|\gamma(y) - \gamma(x)|^2} - \kappa_\gamma(x) \right) \frac{dy}{|\gamma(y) - \gamma(x)|^2} = 0 \quad \textit{for all } x \in I, \tag{4-1}$$

and

$$E^{\text{int}}_{\bar{B}_1(0)}(\tilde{\gamma}_\infty) \geq \varepsilon_0.$$

Proof. The first statement is an immediate consequence of Theorem 4.1 and the bi-Lipschitz estimate (2-5). We consider the rescaled flows

$$\tilde{\gamma}^{(j)}(x, t) := \frac{1}{r_j}(\gamma(x, r_j^3 t + t_j) - x_j)$$

for $t \in (-t_j/r_j^3, 0]$ which still solve (1-2). Under the assumptions of the theorem we get

$$E^{\text{int}}_{B_1(0)}(\tilde{\gamma}_t^{(j)}) \leq \varepsilon_0 \quad \text{for all } t \in \left[-\frac{t_j}{r_j^3}, 0 \right],$$

and thus from the bi-Lipschitz estimate

$$E_{B_{\beta^{-1}}(0)}(\tilde{\gamma}_t^{(j)}) \leq \varepsilon_0 \quad \text{for all } t \in \left[-\frac{t_j}{r_j^3}, 0 \right].$$

Hence we can apply Theorem 3.1 to find

$$\|\partial_s^k \tilde{\gamma}_t\|_{C^k} \leq C_k$$

for all $k \in \mathbb{N}$ and $t \in [-t_j/r_j^3 + 1, 0)$. As $-t_j/r_j^3 \to -\infty$, we can use the Arzelà–Ascoli theorem to get, after going to a subsequence,

$$\tilde{\gamma}_j \to \tilde{\gamma}$$

locally smoothly in time and space. Since all derivatives of γ_∞ are uniformly bounded we furthermore deduce that

$$\mathcal{H}\tilde{\gamma}_\infty(x) = \text{p.v.}\int_I \left(2\frac{P_\tau^\perp(\tilde{\gamma}(y) - \tilde{\gamma}(x))}{|\gamma(y) - \gamma(x)|^2} - \kappa_\gamma(x) \right) \frac{dy}{|\gamma(y) - \gamma(x)|^2}$$

is well-defined. Furthermore, we have

$$\int_{-\delta_0}^0 \int_{\mathbb{R}/l_t\mathbb{Z}} |\mathcal{H}(\gamma_t^{(j)})(x)|^2 \, dx \, dt = E(\gamma_{t_j - r_j^3}) - E(\gamma_{t_j}) \to 0 \tag{4-2}$$

for some subsequence j and hence after going to a subsequence

$$\mathcal{H}\tilde{\gamma}^{(j)}(x) \to 0$$

pointwise almost everywhere. We now show that

$$\mathcal{H}\tilde{\gamma}_j \to \mathcal{H}\tilde{\gamma}_\infty$$

pointwise. For this purpose we again use the decomposition

$$\widetilde{\mathcal{H}\gamma} = Q\gamma + R_1\gamma + R_2\gamma.$$

As

$$\begin{aligned} \frac{|w|^\alpha}{|\gamma(x+w) - \gamma(x)|^{2+\alpha}} - \frac{1}{w^2} &\le C \frac{\int_0^1 \int_0^1 |\gamma'(x + s_1 w) - \gamma'(x + s_2 w)|^2 \, ds_1 \, ds_2}{|w|^2} \\ &\le C \min\left\{ \|\kappa\|_{L^\infty(B_R(x))}, \frac{1}{|w|^2} \right\}, \end{aligned}$$

we get that the integrands of both $R_1(\gamma_j)$ and $R_2(\gamma_j)$ are uniformly bounded. As all the integrands also converge pointwise to the integrands of $R_1(\gamma_\infty)$ and $R_2(\gamma_\infty)$, the dominant convergence theorem yields

$$R(\gamma_j) \to R(\gamma_\infty).$$

For the integrand of Q we use Taylor's approximation up to order 2 to get

$$\frac{\gamma(x+w) - \gamma(x) - w\gamma'(x) - \frac{1}{2}w^2\gamma''(x)}{w^4} = \frac{\int_0^1 (1-s)^2 \gamma'''(x + sw) \, ds}{w}$$

and write

$$Q\gamma = \int_{\mathbb{R}/l\mathbb{Z}} I \, dw,$$

where

$$\begin{cases} I(x, w) := \left(\int_0^1 (1-s)^2 \gamma'''(x + sw) - \gamma'''(x) \, ds \right) / 2w & \text{for } |w| \le 1, \\ I(x, w) := \left(\int_0^1 (1-s) \gamma''(x + sw) - \gamma''(x) \, ds \right) / |w|^2 & \text{else.} \end{cases}$$

The mean value theorem tells us that $|I(x,w)| \le C\|\gamma''''\|_{L^\infty(B_1(x))}$ if $|w| \le 1$, and $I(x,w) \le w^{-2}\|\gamma''\|_{L^\infty}\}$ else. We get using the dominated convergence theorem $Q\gamma_j \to Q\gamma_\infty$. This completes the proof of

$$\mathcal{H}\tilde{\gamma}^{(j)} \to \mathcal{H}\tilde{\gamma}_\infty$$

pointwise.

We get in view of (4-2)

$$\int_{-\delta_0}^{0}\int_\Gamma |\mathcal{H}\tilde{\gamma}_\infty|^2\, d\mathcal{H}^1(x)\, dt \le \lim_{j\to\infty}\int_{-\delta_0}^{0}\int_{\mathbb{R}/l_t\mathbb{Z}} |\mathcal{H}(\gamma_t^{(j)})(x)|^2\, dx\, dt = 0.$$

Since $\tilde{\gamma}_\infty$ is smooth, we obtain $\mathcal{H}\tilde{\gamma}_\infty \equiv 0$. Furthermore, the local smooth convergence together with $E^{\text{int}}_{B_1(0)}(\widetilde{\Gamma}_j) = \varepsilon_0$ implies

$$E^{\text{int}}_{\overline{B}_1(0)}(\widetilde{\Gamma}_\infty) \ge \varepsilon_0. \qquad \square$$

Using the evolutionary attractivity of critical points proven in [Blatt 2012b] we can further show that the blow-up profile cannot be compact.

Proposition 4.3 (blow-ups profiles are never compact). *The blow-up profile constructed in Theorem 4.2 cannot be compact.*

Proof. Let us assume that $\tilde{\gamma}_\infty$ is compact, i.e., that $\tilde{\gamma}_\infty \in C^\infty(\mathbb{R}/l\mathbb{Z}, \mathbb{R}^n)$ for suitable l. Then there would be a subsequence of $\tilde{\gamma}_j$ converging smoothly to the critical point $\tilde{\gamma}_\infty$ of E. Since furthermore $E(\gamma_t) \ge E(\tilde{\gamma}_\infty)$, we get from [Blatt 2012b, Theorem 1.5] for all $t \in [0, T)$, that for j large enough the flow $\tilde{\gamma}_t$ exists for all time and converges to a stationary point of E, which contradicts the assumption $T < \infty$. $\square$

4B. *Planar curves.* For a regular curve γ the *curvature vector* κ is given by

$$\kappa = \frac{\gamma''}{|\gamma'|^2} - \frac{\langle \gamma'', \gamma'\rangle}{|\gamma'|^4}\gamma',$$

which is equal to γ'' if γ is parametrized by arc-length.

Given two points $x, y \in I$ there is either a unique circle or a straight line — which we like to think of as a degenerate circle — going through $\gamma(x)$ and $\gamma(y)$ and being tangent to γ at x. See Figure 1. Note that this is the same circle used to define the integral tangent-point energies. We denote by $\kappa_\Gamma(x, y)$ the curvature vector of this circle in x and set $\kappa_\Gamma(x, y) = 0$ if the tangent on Γ in x is pointing in the direction of y — which is the curvature of the straight line.

Lemma 4.4. *We have*

$$\kappa_\gamma(x, y) = 2\frac{P^\perp_{\gamma'(x)}(\gamma(y) - \gamma(x))}{|\gamma(x) - \gamma(y)|^2}.$$

Proof. If the vectors γ' and $\gamma(x) - \gamma(y)$ are colinear, both sides of the identity obviously vanish. So we can assume that $P^\perp_{\gamma'(x)}(\gamma(y) - \gamma(x)) \neq 0$. The circle going through $\gamma(x)$ that is tangential to γ in the point x with curvature vector $\kappa = aP^\perp(\gamma(y) - \gamma(x))$ is the set of all points $z \in \mathbb{R}^n$ satisfying

$$\left| z - \gamma(x) - \frac{\kappa}{|\kappa|^2}\right|^2 = \frac{1}{|\kappa|^2}.$$

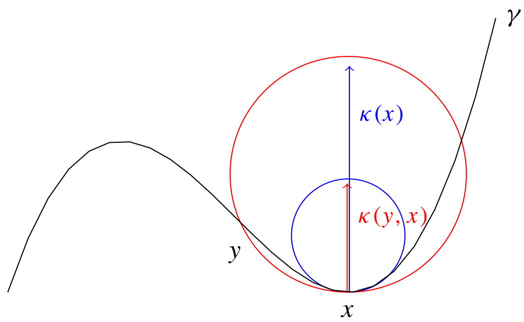

Figure 1. This picture shows the two circles playing a role in the geometric interpretation of the Euler–Lagrange equation of the Möbius energy: the inner circle, with curvature vector $\kappa(x)$, is the osculating circle at x, while the outer circle, with curvature vector $\kappa(x, y)$, is the circle going through x and y and tangent to Γ at x.

This circle contains $\gamma(y)$ if and only if

$$|\gamma(y)-\gamma(x)|^2 = 2\frac{\langle\kappa, \gamma(x)-\gamma(y)\rangle}{\kappa^2} = \frac{2}{a}.$$

Thus, $a = 2/|\gamma(x)-\gamma(y)|^2$ which proves the lemma. □

Using Lemma 4.4 we immediately get the following geometric interpretation of (4-1).

Lemma 4.5 (geometric interpretation of the Euler–Lagrange equation). *The curve γ parametrized by arc-length satisfies $\mathcal{H}\gamma \equiv 0$ if and only if*

$$\lim_{\varepsilon\downarrow 0}\int_{\Gamma\setminus B_\varepsilon(x)} \frac{\kappa_\gamma(x, y)-\kappa_\gamma(x)}{|x-y|^2}\, d\mathcal{H}^1(y) = 0 \tag{4-3}$$

for all $x \in I$.

In codimension 1, (4-3) is equivalent to

$$\lim_{\varepsilon\searrow 0}\int_{I/B_\varepsilon(x)} \frac{\langle\kappa_\gamma(x, y)-\kappa_\gamma(x), n(x)\rangle}{|x-y|^2}\, d\mathcal{H}^1(y) = 0, \tag{4-4}$$

where n is a unit normal along γ We are now looking for a situation that implies that the integrand on the left-hand side of (4-4) has a sign and thus must vanish identically. For $x \in I$ in which the curvature of γ does not vanish, we denote by $\mathrm{OB}(x)$ the open ball whose boundary is the osculating circle along γ in x; i.e.,

$$\mathrm{OB}(x) := B_{1/|\kappa(x)|}\left(\gamma(x) + \frac{\kappa}{|\kappa|^2}\right).$$

Lemma 4.6. *If there is a point $x \in I$ such that*

$$\mathrm{OB}(x) \cap \gamma(I) = \varnothing$$

or

$$\gamma(I) \subset \overline{\mathrm{OB}(x)}$$

then

$$\Gamma = \partial\, \mathrm{OB}(x),$$

i.e., Γ *is a circle.*

Proof. If $\Gamma \cap \mathrm{OB}(x) = \varnothing$, we get

$$\langle \kappa_\Gamma(x, y), n(x)\rangle \le \langle \kappa_\Gamma(x), n(x)\rangle,$$

and if $\Gamma \subset \overline{\mathrm{OB}(x)}$

$$\langle \kappa_\Gamma(x, y), n(x)\rangle \ge \langle \kappa_\Gamma(x), n(x)\rangle.$$

So in both cases

$$\langle \kappa_\Gamma(x, y), n(x)\rangle - \langle \kappa_\Gamma(x), n(x)\rangle$$

has a sign that is independent of $y \in \Gamma$.

Since $\mathcal{H}\gamma \equiv 0$ implies

$$\lim_{\varepsilon \searrow 0} \int_{\Gamma / B_\varepsilon(x)} \frac{\langle (\kappa_\gamma(x, y) - \kappa_\gamma(x)), n(x)\rangle}{|\gamma(y) - \gamma(x)|^2}\, d\mathcal{H}^1(y) = 0$$

and the integrand has a sign, we get

$$\langle (\kappa_\gamma(x, y) - \kappa_\gamma(x)), n(x)\rangle = 0$$

for all $y \in \Gamma$. But this implies

$$\kappa_\gamma(x, y) = \kappa_\gamma(x)$$

for all $y \in \Gamma$, which by the definition of $\kappa_\Gamma(x, y)$ implies

$$y \in \partial\, \mathrm{OB}(x). \qquad \square$$

Theorem 4.7. *Let* $\Gamma : I \to \mathbb{R}^2$ *be a properly embedded smooth curve parametrized by arc-length satisfying*

$$\mathrm{p.v.}\int_I \frac{\kappa_\gamma(x, y) - \kappa_\gamma(x)}{|\gamma(y) - \gamma(x)|^2}\, dy = 0.$$

Then γ *is either a straight line or a circle.*

Proof. Let us assume that γ is not a straight line. We will show that then there is a point $x \in I$ with $\kappa(x) \neq 0$ and

$$\mathrm{OB}(x) \cap \gamma(I) = \varnothing,$$

where $\mathrm{OB}(x)$ is the open ball surrounded by the osculating circle on γ at x; i.e.,

$$\mathrm{OB}(x) := \left\{ y \in \mathbb{R}^2 : \left| y - \left(\gamma(x) + \frac{\kappa(x)}{|\kappa(x)|^2} \right) \right| \le \frac{1}{|\kappa(x)|} \right\}.$$

Then the statement follows from Lemma 4.6. We construct this point as follows: As $\Gamma = \gamma(I)$ is not a straight line, we find a point $x_1 \in \Gamma$ with $\kappa_\Gamma(x_1) \neq 0$. Let n be the continuous unit normal field pointing in the direction of $\kappa_\Gamma(x_1)$ at the point x_1. Then if $\mathrm{OB}(x_1) \cap \Gamma = \varnothing$, we set $x = x_1$. If on the other hand $\mathrm{OB}(x_1) \cap \Gamma \neq \varnothing$, there is a ball $B_1 \subset \mathrm{OB}(x_1)$ touching Γ in x_1 and at least one other point. Let x_1' be one of these touching points nearest to x_1 and let Γ_1 denote the closed curve consisting of the arc of Γ between x_1 and x_1' and the part of the boundary of B_1 that makes this curve C^1 and let Ω_1 be the open set bounded by this curve.

We now start an iterative scheme in order to find the desired point x. So let $x_2 \in \Gamma$ be the point on the part of the curve between x_1 and x_1' which divides this arc into two parts of equal length. Note that $x_2 \notin \bar{B}_1$. We choose

$$r_2 := \sup\left\{r : B_r\left(x_2 + \frac{1}{r}n(x_2)\right) \subset \Omega_1\right\}$$

Then either $B_2 = B_{r_2}(x_2 + (1/r_2)n_2)$ touches Γ in x up to second order and we set $x = x_2$ and have found our point x, or we can choose $x_2' \in \Gamma_1$ to be one of the nearest points on Γ_1 touching $B_2 = B_{r_2}(x_2 + (1/r_2)n_2)$. But then x_2' must belong to the arc of Γ between x_1 and x_1' since else B_2 touches B_1 from within and hence $B_2 \subset B_1$, which is not possible, as $x_2 \in \bar{B}_2$ but $x_2 \notin \bar{B}_1$. Hence,

$$d_\Gamma(x_2, x_2') \leq \tfrac{1}{2} d_\Gamma(x_1, x_1'). \tag{4-5}$$

Then we repeat the construction above, and either get our point x in a finite number of steps, or get a sequence of points x_i, x_i' and balls $B_i \cap \Gamma = \varnothing$ such that B_i touches Γ in x_i, x_i', the intervals x_i, x_i' are nested, and the diameter of the balls B_i is bounded by the diameter of Γ_1 and from below by

$$\|\kappa_\Gamma|_{[x_1, x_1']}\|_{L^\infty}^{-1} > 0.$$

In the latter case, there is a point $x \in \Gamma$ with

$$x = \lim_{i\to\infty} x_i = \lim_{i\to\infty} x_i'$$

and it is well known that

$$r_i \to \frac{1}{|\kappa(x)|}.$$

We get for every $r < 1/|\kappa(x)|$ that

$$B_r\left(x + \frac{1}{r}n(x)\right) \subset B_n$$

for n large enough. Hence,

$$B_r \cap \Gamma = \varnothing \quad \text{for all } r < \frac{1}{\kappa(x)},$$

which implies

$$\mathrm{OB}(x) \cap \Gamma = \varnothing. \qquad \square$$

Using the characterization of the solutions to (1-2) we can now show:

Theorem 4.8 (the evolution of planar curves). *Let $\gamma_0 \subset \mathbb{R}^2$ be a closed smoothly embedded curve. Then the negative gradient flow of the Möbius energy exists for all times and converges to a round circle as time goes to infinity.*

Proof. Let us first prove the long-time existence of the flow. Assume that a singularity occurs after finite time. Then we construct a blow-up profile $\tilde{\gamma}_\infty$ as described in Theorem 3.1. But Theorem 4.7

implies that this blow-up must be a circle or straight line, which is not possible due to Proposition 4.3 and $E^{\text{int}}_{B_1(0)}(\tilde{\gamma}_\infty) \neq 0$.

To prove the statement about the asymptotic behavior of the flow, we let, for $t \in (0, \infty)$ and $\varepsilon_0 > \varepsilon > 0$ small enough, the radius $r_t > 0$ and $x_t \in \gamma_t$ be such that

$$E_{B^{\text{int}}_{r_t}(x_t)}(\gamma_t) = \sup_{x \in \gamma_t} E_{B^{\text{int}}_{r_t}(x)}(\gamma_t) = \varepsilon.$$

Let us assume that

$$M := \liminf_{t \in [0,\infty)} \frac{r_{t+r_t^3/2}}{r_t} < \infty. \tag{4-6}$$

Then we can choose a sequence $t_j \to \infty$ such that

$$r_{t_j + r_{t_j}^3/2} \le 2M r_{t_j}.$$

As in Theorem 4.2, let $\tilde{\gamma}_j$ be reparametrizations of the rescaled curves

$$\frac{1}{r_j}\{\gamma_{t_j + r_{t_j}^3/2} - x_{t_j + r_{t_j}^3/2}\}$$

by arc-length such that $\tilde{\gamma}_j(0) = 0$. Then these curves γ_j subconverge locally smoothly to a curve γ_∞ satisfying $\mathcal{H}\gamma_\infty \equiv 0$, which is not a straight line. Hence, due to Theorem 4.7 γ_∞ is a circle. Since $\tilde{\gamma}_{t_j} \to \gamma_\infty$ smoothly we get that for j large enough, the flow starting with $\tilde{\gamma}_j$ converges smoothly to a circle as time goes to infinity. Hence, the same is true for γ_t.

Let us assume that (4-6) was wrong and let L_t denote the length of the curve γ_t. Then for every $\Lambda > 0$ there is a t_0 such that $r_{t+r_t^3/2} \ge \Lambda r_t$ for all $t \ge t_0$. We iteratively define $t_{j+1} := t_j + r_j^3/2$, where $r_j := r(t_j)$, and get

$$r_j \ge \Lambda^j r_{t_0}. \tag{4-7}$$

Scaling our a priori estimates in Theorem 3.1 we obtain

$$|\mathcal{H}\gamma_{\tilde{t}}| \le \frac{C}{(\tilde{t} - t)^{2/3} r_t^2}$$

for all times $\tilde{t} \in t + (0, r_t^3)$ and hence

$$\left| \frac{d}{dt}\Big|_{t=\tilde{t}} L_t \right| \le 2 \sup \left| \frac{d}{dt}\Big|_{t=\tilde{t}} \gamma \right| \le \frac{C}{(\tilde{t} - t)^{2/3} r_t^2}.$$

Integrating this inequality we obtain

$$L_{t+r_t^3/2} \le C \frac{r_t^3}{r_t^2} + L_t \le C L_t.$$

Hence,

$$L_{t_j+1} \le C L_{t_j}$$

and thus

$$r_j \le L_{t_j} \le C^j L_{t_0},$$

which contradicts (4-7) for $\Lambda > C$ and j large enough. □

Appendix: Besov spaces, commutator estimates and interpolation inequalities

For the convenience of the reader, let us gather some well-known and not so well-known facts about Besov spaces in this section. We will stick to the notation used in [Triebel 1983] and will assume that the reader is familiar with the definition of the Besov spaces $B^s_{p,q}(\mathbb{R}^n)$ on $\mathbb{R}^n$ and the respective spaces $B^s_{p,q}(\Omega)$ on smooth domains $\Omega \subset \mathbb{R}^n$ as defined in Sections 2.3.1 and 3.2.2 of [Triebel 1983].

Essential for our analysis is the following characterization of these spaces using finite differences. For an arbitrary function $f : \mathbb{R}^n \to \mathbb{R}$ these are inductively defined by

$$(\Delta^1_h f)(x) := f(x+h) - f(x), \quad (\Delta^l_h f) = \Delta^1_h \Delta^{l-1}_h f \quad \text{for } l = 2, 3, \dots.$$

Furthermore, for a set $\Omega \subset \mathbb{R}^n$ we set $\Omega_{h,l} = \bigcap_{j=0}^{l} \{x \in \Omega : x + jh \in \Omega\}$.

Lemma A.1 (equivalent norms, see [Triebel 1983, Sections 2.5.12, 3.4.2, and 2.5.10]). *The following estimates hold*:

(1) *For* $0 < p, q \leq \infty$, *we have* $s > \tilde{\sigma}_p := n(1/\min\{p, 1\} - 1)$. *If* $M > s$ *and* M *an integer*, *then*

$$\|f|B^s_{p,q}(\mathbb{R}^n)\|^{(2)}_M := \|f\|_{L^p(\mathbb{R}^n)} + \left(\int_{\mathbb{R}^n} \frac{\|\Delta^M_h f|L^p(\mathbb{R}^n)\|^q}{|h|^{n+sq}}\, dh\right)^{1/q}$$

is an equivalent quasinorm on $B^s_{p,q}(\mathbb{R}^n)$.

(2) *If* $\Omega \subset \mathbb{R}^n$ *is a smooth domain* $1 < p < \infty$,. $s > 0$, *and* k, l *integers with* $0 \leq k < s$ *and* $s < l + k$, *then*

$$\|f|B^s_{p,q}(\Omega)\|^{(2)} := \|f|L^p(\Omega)\| + \sum_{|\alpha| \leq k} \left\| \left(\int_{\mathbb{R}^n} \left(\int_{\Omega_{h,l}} |\Delta^l_h \partial^\alpha f(x)|^q dx\right)^{q/p} \frac{dh}{|h|^{n+sq}}\right)^{1/q} \,\middle|\, L^p(\Omega)\right\|$$

is an equivalent quasinorm on $B^s_{p,q}(\Omega)$.

As an easy consequence, we get

Lemma A.2. *For* $1 < p, q < \infty$ *and* $1 > s > 0$ *we have*

$$\left(\int_{B_1(0)} \frac{\|\Delta^M_h|L^p(B_1(0))\|^q}{|h|^{n+sq}}\, dh\right)^{1/q} \leq C\|f|B^s_{p.q}(B_2(0))\|$$

and

$$\|f|B^s_{p.q}(B_1(0))\| \leq C\left(\int_{B_2(0)} \frac{\|\Delta^M_h|L^p(B_2(0))\|^q}{|h|^{n+sq}}\, dh\right)^{1/q}.$$

Proof. From the definition of the norm, we deduce that there is an extension $\tilde{f}$ of $f|B_2(0)$ such that

$$\|f\|_{B^s_{p,q}(B_1(0))} \leq \|\tilde{f}\|_{B^s_{p,q}(\mathbb{R}^n)} \leq 2\|f\|_{B^s_{p,q}(B_2(0))}.$$

Lemma A.1 gives

$$\begin{aligned}\left(\int_{B_1(0)} \frac{\|\Delta^M_h|L^p(B_1(0))\|^q}{|h|^{n+sq}}\, dh\right)^{1/q} &\leq \|\tilde{f}|L^p(\mathbb{R}^n)\| + \left(\int_{\mathbb{R}^n} \frac{\|\Delta^M_h \tilde{f}|L^p(\mathbb{R}^n)\|^q}{|h|^{n+sq}}\, dh\right)^{1/q}\\ &\leq C\|\tilde{f}|B^s_{p,q}(\mathbb{R}^n)\| \leq C\|f|B^s_{p,q}(B_2(0))\|.\end{aligned}$$

To get the second estimate, we extend $f|B_1(0)$ to a function $\tilde{f}$ such that

$$\|f\|^M_{B^s_{p,q}(B_1(0))} \le \|\tilde{f}\|_{B^s_{p,q}(\mathbb{R}^n)} \le 2\|f\|^M_{B^s_{p,q}(B_2(0))}.$$

and argue as above. □

We will now state the following interpolation inequalities in Besov-space. Since it seems to be hard to find a proof of this result in the literature, we include a proof here for the sake of completeness.

Lemma A.3 (interpolation inequalities). *Let $\Omega \subset \mathbb{R}^n$ be a smooth bounded domain. If $0 \le s_1 < s_2 < s_3$ and $p \in [2, \infty)$ satisfy $s_0 - n/2 < s_1 - n/p < s_w - n/2$ then*

$$\|f|B^{s_1}_{p_1,q}(\Omega)\| \le C\|f|B^{s_0}_{2,2}(\Omega)\|^{1-\theta}\|f|B^{s_2}_{2,2}(\Omega)\|^{\theta}$$

for all $q \in [1, \infty]$ with $C = C(s, p, q, \Omega)$ and

$$\theta = \frac{(s_1 - n/p) - (s_0 - n/2)}{s_2 - s_0}.$$

Proof. By [Lunardi 1995, Proposition 1.3.2] we have to show that the real interpolation space

$$(B^{s_0}_{2,2}(\Omega), B^{s_0}_{2,2}(\Omega))_{\theta,1}$$

is continuously embedded in $B^{s_1}_{p,q}$. But this is indeed the case, as

$$(B^{s_0}_{2,2}(\Omega), B^{s_0}_{2,2}(\Omega))_{\theta,1} = B^{\tilde{s}}_{2,1}(\Omega),$$

with $\tilde{s} = (1-\theta)s_0 + \theta s_2 = s_1 + n/2 - n/p > s_1$ by [Triebel 1992, p. 204] and the Sobolev embedding for Besov spaces [Triebel 1992, p. 196] tells us that $B^{\tilde{s}}_{2,1}(\Omega)$ is continuously embedded into $B^{s_1}_{p_1,1}(\Omega) \subset B^{s_1}_{p_1,q}(\Omega)$ for all $q \in [1, \infty]$. □

One of the most important tools in this article is the following commutator estimate for our operator $\widetilde{Q}$. This is a very special case of well-known fractional Leibniz rule and known commutator estimates of Kato and Ponce, which we still decide to prove here in order to make the article as easily accessible as possible.

Lemma A.4 (commutator estimates). *For $f, g \in C^\infty([-\Lambda, \Lambda])$ we have*

$$\|\widetilde{Q}[fg] - g\widetilde{Q}[f] - f\widetilde{Q}[g]\|_{L^r(B_{1/2}(0))}$$
$$\le C\left(\|f\|_{B^{1/4}_{2p,2}(B_\Lambda(0))}\|g\|_{B^{1/4}_{2p,2}(B_\Lambda(0))} + \sum_{j=1}^{\infty}\frac{\|f\|^2_{L^{2p}(B_{\Lambda+j+1}(0)\setminus B_{\Lambda+j}(0))} + \|g\|^2_{L^{2p}(B_{\Lambda+j+1}(0)\setminus B_{\Lambda+j}(0))}}{(\Lambda+j)^2}\right).$$

Proof. Remember that

$$\frac{\widetilde{Q}f(x)}{4} := \text{ p.v.}\int_{-l/2}^{l/2}\int_0^1 (1-s)\frac{\kappa(x+sw)-\kappa(x)}{|w|^2}\,dw = \widetilde{Q}\kappa(x).$$

Since

$$\begin{aligned}
&\frac{\widetilde{Q}^s[fg]-g\widetilde{Q}^s[f]-f\widetilde{Q}^s[g]}{4}\\
&=\text{p.v.}\int_{-l/2}^{l/2}\int_0^1(1-s)\frac{f(x+sw)g(x+sw)-f(x)g(x)-(f(x+sw)-f(x))g(x)-(g(x+sw)-g(x))f(x)}{|w|^2}\,dw\\
&=\text{p.v.}\int_{-l/2}^{l/2}\int_0^1(1-s)\frac{f(x+sw)g(x+sw)-f(x+sw)g(x)-f(x)g(x+sw)+f(x)g(x)}{|w|^2}\,dw\\
&=\int_{-l/2}^{l/2}\int_0^1(1-s)\frac{(f(x+sw)-f(x))(g(x+sw)-g(x))}{|w|^2}\,dw,
\end{aligned}$$

we get

$$\begin{aligned}
&\|\widetilde{Q}^s[fg]-g\widetilde{Q}^s[f]-f\widetilde{Q}^s[g]\|_{L^p(B_1(0)}\\
&\le C\int_{-l/2}^{l/2}\int_0^1(1-s)\biggl(\int_{B_1(0)}\biggl(\frac{(f(x+sw)-f(x))(g(x+sw)-g(x))}{|w|^2}\biggr)^p dx\biggr)^{1/p} dw\\
&\le C\int_0^1 s(1-s)\biggl(\int_{B_1(0)}\int_{-l/2}^{l/2}\biggl(\frac{(f(x+w)-f(x))(g(x+w)-g(x))}{|w|^2}\biggr)^p dx\biggr)^{1/p} dw\\
&\le C\biggl(\int_{B_1(0)}\int_{-\Lambda/2}^{\Lambda/2}\biggl(\frac{(f(x+w)-f(x))(g(x+w)-g(x))}{|w|^2}\biggr)^p dx\biggr)^{1/p} dw\\
&\qquad\qquad +C\biggl(\int_{B_1(0)}\int_{|w|\ge\Lambda/2}\biggl(\frac{(f(x+w)-f(x))(g(x+w)-g(x))}{|w|^2}\biggr)^p dx\biggr)^{1/p} dw\\
&\le \|f\|_{B^{1/2}_{4,2}(B_\Lambda(0)}\|g\|_{B^{1/2}_{4,2}(B_\Lambda(0)}+C\biggl(\int_{B_1(0)}\int_{|w|\ge\Lambda/2}\biggl(\frac{(f(x+w)-f(x))(g(x+w)-g(x))}{|w|^2}\biggr)^p dx\biggr)^{1/p} dw.
\end{aligned}$$

Factoring and using the Cauchy inequality, we can estimate it by

$$\begin{aligned}
&C\int_{|w|\ge\Lambda}\biggl(\int_{B_1(0)}\frac{|f(x)|^{2p}+|g(x)|^{2p}}{w^2}\,dx\biggr)^{1/p} dw\\
&\qquad +\int_{|w|\ge\Lambda}\biggl(\int_{B_1(0)}\frac{|f(x+w)|^{2p}+|g(x+w)|^{2p}}{w^2}\,dx\biggr)^{1/p} dw\\
&\qquad\qquad \le C\|f\|^2_{L^p(B_\Lambda(0)}\|g\|^2_{L^p(B_\Lambda(0)}+\sum_{j\in\mathbb{N}}\frac{\|f\|^2_{L^{2p}(B_{\Lambda+j}\setminus B_{\Lambda+j-1})}+\|g\|_{L^{2p}(B_{\Lambda+j}\setminus B_{\Lambda+j-1})}}{(\Lambda+j)^2}. \qquad \square
\end{aligned}$$

References

[Abrams et al. 2003] A. Abrams, J. Cantarella, J. H. G. Fu, M. Ghomi, and R. Howard, "Circles minimize most knot energies", *Topology* **42**:2 (2003), 381–394. MR Zbl

[Blatt 2012a] S. Blatt, "Boundedness and regularizing effects of O'Hara's knot energies", *J. Knot Theory Ramifications* **21**:1 (2012), art. id. 1250010. MR Zbl

[Blatt 2012b] S. Blatt, "The gradient flow of the Möbius energy near local minimizers", *Calc. Var. Partial Differential Equations* **43**:3-4 (2012), 403–439. MR Zbl

[Blatt 2018] S. Blatt, "The gradient flow of O'Hara's knot energies", *Math. Ann.* **370**:3-4 (2018), 993–1061. MR Zbl

[Dziuk et al. 2002] G. Dziuk, E. Kuwert, and R. Schätzle, "Evolution of elastic curves in $\mathbb{R}^n$: existence and computation", *SIAM J. Math. Anal.* **33**:5 (2002), 1228–1245. MR Zbl

[Freedman et al. 1994] M. H. Freedman, Z.-X. He, and Z. Wang, "Möbius energy of knots and unknots", *Ann. of Math.* (2) **139**:1 (1994), 1–50. MR Zbl

[Gastel 2006] A. Gastel, "The extrinsic polyharmonic map heat flow in the critical dimension", *Adv. Geom.* **6**:4 (2006), 501–521. MR Zbl

[He 2000] Z.-X. He, "The Euler–Lagrange equation and heat flow for the Möbius energy", *Comm. Pure Appl. Math.* **53**:4 (2000), 399–431. MR Zbl

[Kuwert and Schätzle 2002] E. Kuwert and R. Schätzle, "Gradient flow for the Willmore functional", *Comm. Anal. Geom.* **10**:2 (2002), 307–339. MR Zbl

[Lamm 2004] T. Lamm, "Heat flow for extrinsic biharmonic maps with small initial energy", *Ann. Global Anal. Geom.* **26**:4 (2004), 369–384. MR Zbl

[Lin and Schwetlick 2010] C.-C. Lin and H. R. Schwetlick, "On a flow to untangle elastic knots", *Calc. Var. Partial Differential Equations* **39**:3-4 (2010), 621–647. MR Zbl

[Lunardi 1995] A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems*, Prog. Nonlinear Diff. Equ. Appl. **16**, Birkhäuser, Basel, 1995. MR Zbl

[O'Hara 1991] J. O'Hara, "Energy of a knot", *Topology* **30**:2 (1991), 241–247. MR Zbl

[Triebel 1983] H. Triebel, *Theory of function spaces*, Monographs in Math. **78**, Birkhäuser, Basel, 1983. MR Zbl

[Triebel 1992] H. Triebel, *Theory of function spaces, II*, Monographs in Math. **84**, Birkhäuser, Basel, 1992. MR Zbl

Received 30 Oct 2018. Accepted 7 Mar 2019.

SIMON BLATT: simon.blatt@sbg.ac.at
Paris Lodron Universität Salzburg, Salzburg, Austria

ANALYSIS AND PDE
Vol. 13, No. 3, 2020
dx.doi.org/10.2140/apde.2020.13.943

msp

CORRECTION TO THE ARTICLE THE HEAT KERNEL ON AN ASYMPTOTICALLY CONIC MANIFOLD

DAVID A. SHER

Volume **6**:7 (2013), 1755–1791

We present a minor correction, which leaves the main results unchanged.

In the introduction of [Sher 2013], a theorem of Cheng, Li, and Yau has been misstated. Here is the correct version:

Theorem 1 [Cheng et al. 1981]. *For any $T > 0$, there exist nonzero constants C_1 and C_2 such that the heat kernel on M, denoted by $H^M(t, z, z')$, satisfies, for any $z, z' \in M$ and any $t \in (0, T]$,*

$$H^M(t, z, z') \le \frac{C_1}{t^{n/2}} e^{-|z-z'|^2/(C_2 t)}. \tag{1}$$

Note in particular that the constants C_1 and C_2 may depend on T. The incorrect version of Theorem 1 was used later in [Sher 2013], in Section 2.5, in the proof of Theorem 2. Specifically, Theorem 8 does not give the claimed order-n decay of the heat kernel at the face zf, and the corrected version of Theorem 1 is not sufficient to do so.

However, Theorem 2 is still true. An alternative approach to this proof is already outlined in [Sher 2013], but we also present another version suggested by Pierre Albin (personal communication, 2016). Specifically, since we know the heat kernel is polyhomogeneous, it has an expansion at zf. If that expansion is trivial, the leading order of the heat kernel at zf is ∞ and we are done. Otherwise, the expansion must be of the form

$$F(w, z, z') = w^{s_0} (\log w)^j a_0(z, z') + \text{ (lower order terms)}$$

for some $s_0 \in \mathbb{R}$, $j \in \mathbb{N}_0$, with $a_0(z, z')$ not identically zero. Applying the heat operator to this heat kernel gives zero by definition, and in these coordinates the heat operator is $-w^2\partial_w + \Delta_z$. Since the kernel has a polyhomogeneous conormal expansion we may apply this operator term by term. The leading-order term of the result is

$$w^{s_0} (\log w)^j \Delta_z a_0(z, z'),$$

with all other terms lower order. This term must be zero, so $\Delta_z a_0(z, z')$ must be zero, and therefore $a_0(z, z')$ is harmonic for each z' and nonvanishing for at least some z'. By the maximum principle, $a_0(z, z')$ cannot decay at infinity. So the leading-order term of $F(w, z, z')$ at zf is $w^{s_0}(\log w)^j$ times a

MSC2010: 58J05, 58J35, 58J52.
Keywords: heat kernel, geometric microlocal analysis, asymptotically conic manifolds.

term which *does not vanish* at the left face lf_0. Since $w^{s_0}(\log w)^j$ has index set (s_0, j) at lf_0, the index set of $F(w, z, z')$ at lf_0 must contain a term no better than (s_0, j) — in particular $F(w, z, z')$ cannot decay to any order better than $w^{s_0}(\log w)^j$ at lf_0. However, we already know that the leading order of $F(w, z, z')$ at lf_0 is n. Thus $s_0 \geq n$, with $j = 0$ if $s_0 = n$. This shows that the leading order of the heat kernel at zf must actually be at least n, filling the gap in the proof of Theorem 2.

Acknowledgements

The author would like to thank Pierre Albin for finding this error and suggesting this alternative approach.

References

[Cheng et al. 1981] S. Y. Cheng, P. Li, and S. T. Yau, "On the upper estimate of the heat kernel of a complete Riemannian manifold", *Amer. J. Math.* **103**:5 (1981), 1021–1063. MR Zbl

[Sher 2013] D. A. Sher, "The heat kernel on an asymptotically conic manifold", *Anal. PDE* **6**:7 (2013), 1755–1791. MR Zbl

Received 6 Dec 2018. Revised 10 Feb 2019. Accepted 21 Mar 2019.

DAVID A. SHER: dsher@depaul.edu
Department of Mathematical Sciences, DePaul University, Chicago, IL, 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 13 No. 3 2020

2157-5045(2020)13:3;1-B