A REPORT ON STABILITY OF ISOPERIMETRIC INEQUALITIES VIA OPTIMAL TRANSPORTATION

KENNETH DEMASON

Contents

1. Historical Overview

In this report we discuss the results of [FMP10], which led to quantitative stability of the anisotropic isoperimetric inequality.

1.1. **The Isoperimetric Inequality.** To start, let's introduce the isoperimetric inequality and some known results. Given an open subset $E \subset \mathbb{R}^n$, with say smooth or polygonal boundary, and $|E| < \infty$ the perimeter of E is defined by

$$P(E) = \int_{\partial E} 1 \, d\mathcal{H}^{n-1}$$

where $\mathcal{H}^{n-1}$ is the $(n-1)$-dimensional Hausdorff measure on $\mathbb{R}^n$. Another advantageous definition of the perimeter is

$$P(E) = \lim_{\epsilon \to 0^+} \frac{|E + \epsilon B| - |E|}{\epsilon}$$

where B is the unit ball in $\mathbb{R}^n$. It is well known that the following isoperimetric inequality holds:

$$P(E) \geq n|B|^{1/n}|E|^{(n-1)/n}$$

with equality if and only if $E = B$ up to action by $\mathrm{Iso}(\mathbb{R}^n)$ on B and up to a set of Lebesgue measure zero.

To prove this, one may appeal to either the so-called Knothe map (a monotone rearrangement of sorts) or the Brenier map between the measures $\mu = \chi_E/|E|dx$ and $\nu = \chi_B/|B|dx$. Both of these maps $T : \mathbb{R}^n \to \mathbb{R}^n$ satisfy three critical properties

i) $|T| \leq 1$,
ii) $\operatorname{div} T \geq n(\det \nabla T)^{1/n}$, and

iii) $\det \nabla T = |B|/|E|$.

The proof is essentially due to Cauchy-Schwarz, the divergence theorem, and the arithmetic-geometric inequality:

$$P(E) = \int_{\partial E} 1 \, d\mathcal{H}^{n-1} \geq \int_{\partial E} \langle T, \nu_E \rangle \, d\mathcal{H}^{n-1} = \int_E \operatorname{div} T \, dx$$
$$\geq n \int_E (\det \nabla T)^{1/n} \, dx = n \int_E \left(\frac{|B|}{|E|} \right)^{1/n} dx = n|B|^{1/n}|E|^{(n-1)/n}.$$

The equality case is handled differently depending on which map is used. The Knothe map, being defined in terms of a specific orthonormal basis, is insufficient to handle equality alone – infinitely many such maps must be used to essentially show that the supporting hyperplanes to E coincide almost everywhere to those of B. On the other hand, the Brenier map, being the gradient of a convex function, must be the identity map if equality holds.[1] For more details, see here or [FMP10] Section 1. Through this proof we already see one of the advantages of an optimal transport approach: it immediately provides the existence of a map which has the exact properties necessary to prove the inequality and, moreover, is rigid enough to easily prove the equality case.

1.2. **Stability of the Isoperimetric Inequality.** A natural question to ask is "If E almost saturates the isoperimetric inequality, then is E almost a ball?" This is the nature of stability problems. First, one must determine in which sense closeness (of sets) means. An initial result by Fuglede in [Fug89] established if E is a convex subset of $\mathbb{R}^n$ having the same measure as the unit ball then

$$P(E) \geq n|B|^{1/n}|E|^{(n-1)/n} \left(\min \{ d_H(E, x+B) \mid x \in \mathbb{R}^n \}^{1/\alpha(n)} + 1 \right)$$

where $d_H(\cdot,\cdot)$ is the Hausdorff distance and the exponent $1/\alpha(n)$ is sharp for $n \neq 3$. Here, $\alpha(n) = 2$ for $n = 3$ and $\alpha(n) = 1/2(n+1)$ for $n \geq 4$. In this case, closeness of E to the optimizer B is determined by $\min \{ d_H(E, x+B) \mid x \in \mathbb{R}^n \}$. Is this the correct notion in general though? Let's consider E which is B but with a long, thin tentacle protruding orthogonally a height h and radius ϵ from the surface. Clearly E is not convex. Its perimeter and volume can be approximated by

$$P(E) \approx P(B) + P(B_\epsilon^{n-1})h \approx P(B)$$
$$|E| \approx |B| + |B_\epsilon^{n-1}|h \approx |B|$$

where B_ϵ^{n-1} denotes the ball of dimension $n-1$ and radius ϵ, and ϵ is very small. On the other hand,

$$\min \{ d_H(E, x+B) \mid x \in \mathbb{R}^n \} = d_H(E, B) = \max \left\{ \sup_{y \in Y} d(X, y), \sup_{x \in X} d(x, Y) \right\} \geq h$$

The right hand side of the quantitative inequality above is then approximately bounded below by $n|B|(h^{1/\alpha(n)}+1)$ while the left hand side is approximately $P(B)$. By choosing ϵ and h appropriately, the inequality will eventually fail. So we need a different measurement of closeness.

In [HHW91], Hall, Hayman, and Weitsman considered the *Fraenkel asymmetry of* E defined by

$$\lambda(E) := \min \left\{ \frac{|E \Delta (x + rB)|}{r^n} \;\middle|\; x \in \mathbb{R}^n, \quad r^n|B| = |E| \right\}.$$

Denote also the *isoperimetric deficit of* E by

$$\delta(E) = \frac{P(E)}{n|B|^{1/n}|E|^{(n-1)/n}} - 1.$$

This measures the extent to which the isoperimetric inequality is saturated: $\delta(E) = 0$ if and only if equality holds, and small $\delta(E)$ means E almost saturates the inequality.

In this paper, the authors proved that if E is a smooth open set with $\delta(E)$ small then there exists a straight line such that, if E^* is the Steiner symmetrization of E with respect to this line, then

$$\lambda(E) \leq C(n)\sqrt{\lambda(E^*)}.$$

[1] After assuming E and B at least have the same barycenter and measure.

Later, Hall proved in [Hal92] that if F is axially symmetric then

$$\lambda(F) \leq C(n)\sqrt{\delta(F)}.$$

But, E^* is axially symmetric. So by applying this with $F = E^*$ one obtains

$$\lambda(E) \leq C(n)\sqrt{\lambda(E^*)} \leq C(n)\delta(E^*)^{1/4}.$$

Now since the Steiner symmetrization decreases perimeter while keeping the measure fixed, we have $\delta(E^*) \leq \delta(E)$. In total:

$$\lambda(E) \leq C(n)\delta(E)^{1/4}.$$

Although the two previous inequalities are sharp, combining the two need not be. Indeed, Hall conjectured that the power $1/4$ is not optimal. Rather, he thought it should be $1/2$.

In [FMP08] this was proved in the affirmative using a series of symmetrization techniques. A few years later, in [CL12] Cicalese and Leonardi reproved this stability theorem via a new technique they called the *selection principle*. Concurrently, the authors of [FMP10] proved a sharp quantitative estimate for the anisotropic perimeter, thus also reproving this past result.

2. Introduction

2.1. **The Anisotropic Perimeter.** We'd like to extend the previous results to a more general setting. One obvious generalization is to change the perimeter functional in a suitable way, for example by weighting points on ∂E differently. To formalize this, let K be a bounded, open, convex subset of $\mathbb{R}^n$ containing the origin. Then we can define a weight function on directions by

$$\|\nu\|_* := \sup_{x\in K} \langle x, \nu\rangle$$

where $\nu \in S^{n-1}$. The set K is called the *Wulff shape* and is uniquely characterized by this relationship. Once again consider an open subset E of $\mathbb{R}^n$ with smooth or polyhedral boundary. Letting ν_E denote its outer unit normal, we define the *anisotropic perimeter* by

$$P_K(E) := \int_{\partial E} \|\nu_E\|_* \, d\mathcal{H}^{n-1},$$

where we have weighted each point $x \in \partial E$ by $\|\nu_E(x)\|_*$. In the isotropic case $K = B$ we obtain $\|\nu\|_* \equiv 1$, and so $P_B(E) = P(E)$.

This generalization provides an interesting problem to study. On the one hand, the anisotropic isoperimetric inequality

$$P_K(E) \geq n|K|^{1/n}|E|^{(n-1)/n}$$

with equality if and only if $E = K$ after affine modifications, is readily proven by the exact same techniques as before *verbatim*! Simply replace B with K everywhere. With this in mind, one would hope to establish a similar quantitative stability result. Define the deficit and asymmetry index as

$$\delta(E) := \frac{P_K(E)}{n|K|^{1/n}|E|^{(n-1)/n}} - 1$$
$$A(E) := \inf\left\{ \frac{|E\Delta(x_0 + rK)|}{|E|} \;\middle|\; x_0 \in \mathbb{R}^n,\ r^n|K| = |E| \right\}.$$

In [FMP10], this is proven

Theorem 1.1. *Let E be a set of finite perimeter with $|E| < \infty$, then there exists $C(n)$ such that*

$$A(E) \leq C(n)\sqrt{\delta(E)}.$$

The exponent $1/2$ is sharp.

On the other hand, this theorem has no chance of being proven via previous techniques. Indeed, the results in [HHW91], [Hal92], and [FMP08] all relied on a symmetrization technique, but the optimizer K of the anisotropic isoperimetric inequality generically has no symmetry. So, the problem itself inherently has no symmetry properties to make use of. Furthermore, the selection principle developed in [CL12] does not apply. At its core, the selection principle relies on regularity properties of almost minimizers. This typically leads to strengthened inequalities – indeed, recently in [FJ14] Fusco and Julin prove a strong form of the quantitative isoperimetric inequality, where a term involving the oscillation on the boundary is added, via the selection principle. Critically, their proof relies on a result in [Fug89] that used explicit knowledge of spherical harmonics and a spectral gap in the Laplace operator. In the anisotropic setting, establishing this spectral gap is much more challenging.[2] The only hope is to adapt carefully the typical proofs of the anisotropic isoperimetric inequality (especially the one exploiting Brenier's theorem) to extract quantitative information.

2.2. **Sets of Finite Perimeter.** As seen in the statement of Theorem 1.1, the paper [FMP10] works in the framework of sets of finite perimeter. For those unfamiliar, a brief overview is given in Section 2.1 of the paper. We will assume the reader has working knowledge of this material. Some standard results, such as the the De Giorgi rectifiability theorem and approximation by smooth, open, bounded sets, are given. There are also some highly technical results, such as the generalized divergence theorem (2.18).

There are two things I would like to point out here for the uninformed reader. First, sets of finite perimeter are a technical tool based on the theory of BV functions. In some sense, a set of finite perimeter admits the weakest regularity on its boundary to allow for a divergence theorem to hold. While the proofs often involve technical measure-theoretic statements, most can be heuristically proven by drawing pictures in the smooth case. For any unproven results about sets of finite perimeter in this survey, I encourage the reader to take a minute to draw the corresponding picture and convince themselves. The second remark is of a possible sign confusion. Given a set of finite perimeter E the measure-theoretic *outer* unit normal, when it exists, is defined by

$$\nu_E(x) = \lim_{r\to 0^+} \frac{-D\chi_E(B_r(x))}{|D\chi_E|(B_r(x))}.$$

Why do we need this negative sign? Recall that the gradient of a function points in the direction of greatest ascent. However, moving from the interior of E to the complement (i.e., in the direction of the outer unit normal) constitutes a *decrease* in χ_E. So, we need to take the opposite sign.

Finally, we point out a remark the authors make. During the proof of Theorem 1.1 we will need to pass to a subset $G \subseteq E$ close to E in the sense that $\delta(G)$, $|E \setminus G|$, and $P_K(G)$ are well controlled (and small). Typically we can assume that E is open, bounded, and has smooth boundary by the usual approximation of sets of finite perimeter. In this case, all of the technical results of sets of finite perimeter can be avoided – by appealing to regularity theory for the Monge-Ampere equation developed by Caffarelli and Urbas, see e.g. [Caf92], one can show the Brenier map $T : E \to K$ is smooth. However, because we work with G directly, and G may not be open, we cannot avoid the technical difficulties.

2.3. **Notation.** A brief summary of the notation used in the paper is given below:

- $\|\cdot\|$: The semi-norm defined by $\|x\| = \inf\{\lambda > 0 \mid x/\lambda \in K\}$.
- $|\cdot|$: The norm on $n \times n$ matrices given by $|A| = \sqrt{\mathrm{trace}(A^T A)}$.
- E: A set of finite perimeter, typically with $|E| < \infty$.
- $\mathcal{F}E$: The reduced boundary of E, i.e. the points where a measure-theoretic outer unit normal exists.
- ν_E: The measure-theoretic outer unit normal to E.
- $E^{(\lambda)}$: The set of points having Lebesgue density λ with respect to E.
- $\partial^* E$: The essential boundary of E, i.e. $\mathbb{R}^n \setminus (E^{(0)} \cup E^{(1)})$.
- T: An element of $BV(\mathbb{R}^n; \mathbb{R}^n)$, i.e an L^1 vector field on $\mathbb{R}^n$ with $|DT|(\mathbb{R}^n) < \infty$.

[2] I give here my thanks to Robin Neumayer for pointing this out to me, I originally thought a selection principle argument fails for a different reason. In fact, she was able to establish the spectral gap using results in [FMP10], see [Neu16]

- DT: The $\mathbb{R}^{n\times n}$-valued distributional derivative of T.
- D_sT: The singular part of the distributional derivative of T, i.e. the measure satisfying $DT = \nabla T\ dx + D_sT$.
- $|DT|$: The total variation of DT induced by the metric $|\cdot|$.
- $\|\mu\|_*$: The anisotropic total variation of an $\mathbb{R}^n$-valued Borel measure μ.
- $\operatorname{Div} T$: The distributional divergence of T, which satisfies $\operatorname{Div} T = \operatorname{div} T\ dx + (\operatorname{Div} T)_s$ and $(\operatorname{Div} T)_s = \operatorname{trace}(D_sT)$.
- $\operatorname{tr}_E(T)$: The trace of T relative to E satisfying for almost every $x \in \mathcal{F}E$

$$\lim_{r\to 0}\frac{1}{r^n}\int_{B_r(x)\cap\{y\ |\ \langle y,\nu_E(x)\rangle<0\}} |T(y)-\operatorname{tr}_E(T)(x)|\ dy = 0.$$

All theorems, lemmas, equations, etc. are labeled as in the paper.

3. Initial Estimates from the Anisotropic Isoperimetric Inequality

The authors first revisit the proof of the anisotropic isoperimetric inequality

Theorem 2.3. *Whenever $E \subset \mathbb{R}^n$ is a set of finite perimeter with $|E| < \infty$ we have*

$$P_K(E) \geq n|K|^{1/n}|E|^{(n-1)/n}.$$

to establish the estimate

Corollary 2.4. *Let E be a set of finite perimeter with $|E| = |K|$, and let T be the Brenier map of E into K. If $\delta(E) \leq 1$, then*

$$n|K|\delta(E) \geq \int_{\mathcal{F}E}(1-\|\operatorname{tr}_E(T)\|)\|\nu_E\|_*\ d\mathcal{H}^{n-1}, \tag{2.29}$$

$$9n^2|K|\sqrt{\delta(E)} \geq \int_E |\nabla T(x)-\operatorname{Id}|\ dx + |D_sT|(E^{(1)}) = |DS|(E^{(1)}), \tag{2.30}$$

where $S(x) = T(x) - x$.

The first inequality is immediate. In proving Theorem 2.3, the final chain of inequalities applies the fact that $\|\operatorname{tr}_E(T)\| \leq 1$ to conclude

$$n|K|^{1/n}|E|^{(n-1)/n} \leq \int_{\mathcal{F}E}\|\operatorname{tr}_E(T)\|\|\nu_E\|_*\ d\mathcal{H}^{n-1} \leq \int_{\mathcal{F}E}\|\nu_E\|_*\ d\mathcal{H}^{n-1} = P_K(E),$$

where the first inequality was established previously. Inequality (2.29) is obtained directly from these estimates. The proof of (2.30) is more involved and exploits the following elementary lemma

Lemma 2.5. *Let $0 < \lambda_1 \leq ... \leq \lambda_n$ be positive real numbers, and set*

$$\lambda_A := \frac{1}{n}\sum_{n=1}^{k}\lambda_k, \qquad \lambda_G := \left(\prod_{k=1}^{n}\lambda_k\right)^{1/n}. \tag{2.31}$$

Then

$$7n^2(\lambda_A-\lambda_G) \geq \frac{1}{\lambda_n}\sum_{k=1}^{n}(\lambda_k-\lambda_G)^2. \tag{2.32}$$

which essentially relies on the following inequality

$$\log(s) \leq \log(t) + \frac{s-t}{t} - \frac{(s-t)^2}{2\max\{s,t\}^2}, \quad s,t \in (0,\infty).$$

The proof comes down to applying Lemma 2.5 with the eigenvalues $\{\lambda_k\}_{k=1}^n$ of ∇T. Then, λ_A is giving $\operatorname{div} T(x)/n$ while λ_G is $(\det \nabla T(x))^{1/n}$. Both these quantities appear in the proof of the anisotropic isoperimetric inequality, and by arguing similarly as above for (2.29) one arrives at (2.30).

4. Stability for the Anisotropic Isoperimetric Inequality

4.1. **Trace inequalities.** Here the authors prove a technical lemma which, morally speaking, bounds the integral of a function along $\mathcal{F}E$ by the integral of its derivative over E. For motivation, consider $f \in W^{1,p}(\Omega)$ where $\Omega \subset \mathbb{R}^n$ is a bounded domain with Lipschitz boundary and $1 \leq p < \infty$. Then the following Poincaré inequality holds: there exists a $C_1 > 0$ such that

$$\int_\Omega \left| f(x) - \frac{1}{|\Omega|} \int_\Omega f(y)\ dy \right|^p\ dx \leq C_1 \int_\Omega |\nabla f(x)|^p\ dx.$$

On the other hand, we also have the existence of a bounded linear operator $\mathrm{Tr} : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ which extends the usual trace for compactly supported smooth functions. Let c denote the average value of f over Ω so that

$$\int_{\partial\Omega} |\operatorname{Tr}(f-c)|^p\ d\mathcal{H}^{n-1} \leq C_2 \int_\Omega |f(x) - c|^p\ dx \leq C \int_\Omega |\nabla f(x)|^p\ dx.$$

What the authors show is that, if E is just a set of finite perimeter then for $p = 1$ a similar trace inequality holds, and moreover the constant C can be deduced somewhat explicitly. It is written in terms of a a parameter $\tau(E)$ defined by

$$\tau(E) = \inf \left\{ \frac{P_K(F)}{\int_{\mathcal{F}F \cap \mathcal{F}E} \|\nu_E\|_*\ d\mathcal{H}^{n-1}} \ \middle|\ F \subset E,\ 0 < |F| \leq \frac{|E|}{2} \right\}.$$

Noting that

$$\mathcal{F}F = [\mathcal{F}F \cap E^{(1)}] \cup [\mathcal{F}F \cap \mathcal{F}E]$$

and $\nu_E = \nu_F$ for $\mathcal{H}^{n-1}$-a.e. point in $\mathcal{F}F \cap \mathcal{F}E$, we can in fact write $\tau(E)$ as

$$\tau(E) = \inf \left\{ 1 + \frac{\int_{\mathcal{F}F \cap E^{(1)}} \|\nu_E\|_*\ d\mathcal{H}^{n-1}}{\int_{\mathcal{F}F \cap \mathcal{F}E} \|\nu_E\|_*\ d\mathcal{H}^{n-1}} \ \middle|\ F \subset E,\ 0 < |F| \leq \frac{|E|}{2} \right\}.$$

Importantly, we see $\tau(E) \geq 1$ and that $\tau(E)$ is measuring the extent to which we can find a small set F (i.e., satisfying $|F| \leq |E|/2$) whose boundary mostly intersects with that of E. For instance, $\tau(E) = 1$ when E is multiply connected and one of its connected regions has small volume (just take the entire region) or when E has outward cusps. One analogy is that E is a container while F is a droplet of liquid on the inner surface, which tends to stick to the surface rather than bend inward.

Before stating the inequality, let us define

$$m_K := \inf\{\|\nu\|_* \mid \nu \in S^{n-1}\}, \qquad M_K := \sup\{\|\nu\|_* \mid \nu \in S^{n-1}\}. \tag{3.1}$$

How do we interpret these? By the Fenchel inequality

$$\langle x, \nu \rangle \leq \|x\| \|\nu\|_*$$

we have that $|x| \leq \|x\| M_K$. On the other hand, since equality is attained in the above at a particular ν_0 we see that

$$|x| \geq \langle x, \nu_0 \rangle = \|x\| \|\nu_0\|_* \geq \|x\| m_K.$$

Together we have

$$\frac{|x|}{M_K} \leq \|x\| \leq \frac{|x|}{m_K}. \tag{3.2}$$

Let us now formally state the trace theorem the authors prove.

Lemma 3.1. *For every function $f \in BV(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ and for every set of finite perimeter E with $|E| < \infty$ we have*

$$\| - Df\|_*(E^{(1)}) \geq \frac{m_K}{M_K}(\tau(E) - 1) \inf_{c \in \mathbb{R}} \int_{\mathcal{F}E} \operatorname{tr}_E(|f - c|) \|\nu_E\|_*\ d\mathcal{H}^{n-1}. \tag{3.3}$$

4.2. **Maximal critical sets.** The trace inequality above is a powerful tool, but it is ineffective due to the parameter $\tau(E)$. A priori there is no good way to estimate this term. The contents of Lemmas 3.2 and 3.3, and Theorem 3.4 are to find a good subset $G \subseteq E$ where $\tau(G)$ can be estimated in terms of known quantities, and such that $|E \setminus G|$ and $\delta(G)$ are well controlled. We first state Lemma 3.2 then discuss its contents.

Lemma 3.2 (Existence of a maximal critical set)**.** *Let E be a set of finite perimeter with $0 < |E| < \infty$, and let $\lambda > 1$. If the family of sets*

$$\Gamma_\lambda = \left\{ F \subseteq E \;\middle|\; 0 < |F| \leq \frac{|E|}{2}, \quad P_K(F) \leq \lambda \int_{\mathcal{F}F \cap \mathcal{F}E} \|\nu_E\|_* \, d\mathcal{H}^{n-1} \right\}$$

is non-empty, then it admits a maximal element with respect to the order relation defined by set inclusion up to sets of measure zero.

What Lemma 3.2 states is that if Γ_λ is nonempty there exists a maximal subset $F_\infty \in \Gamma_\lambda$, and so $\tau(E) \leq \lambda$. The idea now is if we can control $P_K(E \setminus F_\infty)$, then we obtain a *lower bound* of the form $\tau(E \setminus F_\infty) \geq \lambda$. At the same time, we should control $|F_\infty|$ and $\delta(E \setminus F_\infty)$. Controlling these quantities is essential for two reasons. First, it is crucial in showing that $\tau(E \setminus F_\infty) \geq \lambda$. Also, we will apply the trace inequality on the good subset G, and will need a way to connect the deficit and asymmetry index back to E. Controlling the above quantities is critical here too.

Sketch proof. The proof proceeds via induction. Let $F_1 \in \Gamma_\lambda$ and, once F_h has been defined for $h \geq 1$, consider

$$\Gamma_\lambda(h) = \{F \in \Gamma_\lambda \mid F_h \subseteq F\}.$$

The set $\Gamma_\lambda(h)$ is clearly non-empty since $F_h \in \Gamma_\lambda(h)$. Then, we can define

$$s_h = \sup_{F \in \Gamma_\lambda(h)} |F| < \infty.$$

Now we inductively choose $F_{h+1} \in \Gamma_\lambda(h)$ such that

$$|F_{h+1}| \geq \frac{|F_h| + s_h}{2} \geq |F_h|;$$

again such a set exists since F_h meets the criterion. In this way the induction always proceeds and, since $F_h \subseteq F_{h+1}$, we may define F_∞ to be the limit of this increasing sequence. One can then show that F_∞ is maximal. □

4.3. **Critical sets in almost optimal sets.** This lemma provides control over $|F|$, $P_K(E \setminus F)$, and in a certain regime $\delta(E \setminus F)$, whenever $F \in \Gamma_\lambda$ with a suitably chosen λ.

Lemma 3.3 (Removal of a critical set)**.** *Let E and F be two sets of finite perimeter, with $F \subseteq E$ such that*

$$0 < |F| \leq \frac{|E|}{2}, \qquad P_K(F) \leq \left(1 + \frac{m_K}{M_K} k(n)\right) \int_{\mathcal{F}F \cap \mathcal{F}E} \|\nu_E\|_* \, d\mathcal{H}^{n-1}.$$

Then

$$|F| \leq \left(\frac{\delta(E)}{k(n)}\right)^{n/(n-1)} |E|, \qquad P_K(E \setminus F) \leq P_K(E),$$

and in particular, provided $\delta(E) \leq k(n)$,

$$\delta(E \setminus F) \leq \frac{3}{k(n)} \delta(E).$$

Here, $k(n)$ is given by

$$k(n) = \frac{2 - 2^{(n-1)/n}}{3}.$$

Sketch proof. As previously seen, we can decompose the reduced boundary $\mathcal{F}F$ into a portion residing in E (that is, $E^{(1)}$) and the portion residing on E (that is, $\mathcal{F}E$). Doing this with F and $G := E \setminus F$, and rewriting $\|\cdot\|_*$ using

$$\|y\|_* \leq \frac{M_K}{m_K} \|-y\|_*$$

for all $y \in \mathbb{R}^n$ yields

$$P_K(E) \geq P_K(G) + (1 - 2k(n))P_K(F),$$

where we have notably used the estimate on $P_K(F)$. By the anisotropic isoperimetric inequality, we obtain

$$P_K(E) \geq n|K|^{1/n} \left(|G|^{(n-1)/n} + (1 - 2k(n))|F|^{(n-1)/n} \right).$$

Define now the strictly concave function $\Psi : [0,1] \to [0, 2^{1/n} - 1]$ by

$$\Psi(s) = s^{(n-1)/n} + (1-s)^{(n-1)/n} - 1$$

The parameter $k(n)$ is chosen so that

$$\Psi(s) \geq 3k(n)s^{(n-1)/n}$$

for $s \in [0, 1/2]$ with equality precisely when $s = 1/2$. Using this in the above, after some rearrangement, we get

$$|F| \leq \left(\frac{\delta(E)}{k(n)} \right)^{n/(n-1)} |E|,$$

and since $k(n) \leq 1/2$ we obtain $P_K(E) \geq P_K(G) = P_K(E \setminus F)$. The bound on $\delta(E \setminus F)$ is obtained by some elementary estimates. □

4.4. **Reduction to a better set.** This theorem makes precise the control on $|E \setminus G|$ and $\delta(G)$ when passing to the good subset $G \subseteq E$, and moreover establishes the bound on $\tau(G)$.

Theorem 3.4. *Let E be a set of finite perimeter, with $0 < |E| < \infty$ and $\delta(E) \leq k(n)^2/8$. Then there exists $G \subseteq E$, having finite perimeter, such that*

$$|E \setminus G| \leq \frac{\delta(E)}{k(n)}|E|, \qquad \delta(G) \leq \frac{3}{k(n)}\delta(E), \tag{3.17}$$

and

$$\tau(G) \geq 1 + \frac{m_K}{M_K}k(n). \tag{3.18}$$

The proof proceeds by applying Lemma 3.2 to the family Γ_λ with $\lambda = 1 + m_K/M_K k(n)$. It is easily verified that the set $G := E \setminus F_\infty$ satisfies (3.17). The difficulty is in showing (3.18). Maximality of F_∞ comes to the rescue here to arrive at a contradiction if $\tau(G) < \lambda$.

4.5. **Proof of Theorem 1.1.** Before finishing we have one small estimate to conclude.

Lemma 3.5. *If E is a set of finite perimeter in $\mathbb{R}^n$ with $|E| < \infty$, then*

$$\int_{\mathcal{F}E} |\|x\| - 1| \|\nu_E(x)\|_* \, d\mathcal{H}^{n-1}(x) \geq \frac{m_K}{M_K}|E \setminus K|. \tag{3.24}$$

Why do we expect such a theorem? Consider the isotropic case, so that $\|\cdot\|_* = 1 = m_K/M_K$. Then, we are saying

$$\int_{\mathcal{F}E} ||x| - 1| \, d\mathcal{H}^{n-1}(x) \geq |E \setminus B|$$

which is believable by writing everything in polar coordinates.

Proof of Theorem 1.1: We finally prove the main theorem. The proof consists of two steps. In the first we obtain a quantitative control over the asymmetry index $A(E)$ in terms of $\sqrt{\delta(E)}$ with a constant dependent on K and n.

In the second step, we exploit a judicious choice of affine transformation L to obtain an estimate like $A(E) \leq C(n)(M_{L(K)}/m_{L(K)})^p \sqrt{\delta(E)}$. By careful choice of set inclusions (for the image $L(K)$) we can actually show that $M_{L(K)}/m_{L(K)} \leq n$, thus proving the theorem.

Step 1: We show

$$A(E) \leq C_0(n, K)\sqrt{\delta(E)}$$

where

$$C_0(n,K) = \frac{181n^3}{(2-2^{(n-1)/n})^{3/2}}\left(\frac{M_K}{m_K}\right)^4.$$

Proof. By considering an approximation scheme with bounded open sets, we may assume that E is bounded (for sets with a cusp tending to infinity, a set may have finite volume but be unbounded). Recall that

$$A(E) := \inf\left\{\frac{|E\Delta(x_0+rK)|}{|E|}\;\middle|\; x_0\in\mathbb{R}^n,\ r^n|K| = |E|\right\}.$$

By monotonicity and properties of the Lebesgue measure,

$$\begin{aligned}|E\Delta(x_0+rK)| &= |E\setminus[x_0+rK]| + |[x_0+rK]\setminus E|\\ &\le |E| + |x_0+rK| = |E| + r^n|K|.\end{aligned}$$

Thus for admissible x_0 and r,

$$A(E) \le \frac{|E|+r^n|K|}{|E|} = 2.$$

If now $\delta(E)\ge k(n)^2/8$ then

$$A(E)\le 2 \le \frac{4\sqrt{2}}{k(n)}\sqrt{\delta(E)} \le C_0(n,K)\sqrt{\delta(E)}$$

so we can assume directly that $\delta(E)\le k(n)^2/8$. This puts us in a regime to apply Theorem 3.4, yielding $G\subseteq E$ such that

$$|E\setminus G| \le \frac{\delta(E)}{k(n)}|E|, \qquad \delta(G)\le\frac{3}{k(n)}\delta(E), \qquad \tau(G)\ge 1+\frac{m_K}{M_K}k(n).$$

We'd like to apply the Brenier theorem to obtain a map from G to K (rather than E to K, since G has a nice trace properties). We'll need to dilate both E and G so that $|G|=|K|$; doing this leaves unchanged all of the above. Now let T be the Brenier map from G to K. Since T takes values in K, it follows that $\|\operatorname{tr}_G(T)\|\le 1$. In a smooth setting, $\operatorname{tr}_G(T)$ yields a map which is the "restriction" of T to ∂G. Since T takes values in K, and $\|x\|$ measures how much you need to dilate $x\in\mathbb{R}^n$ so that $x\in\partial K$, it follows that $\|T(x)\|<1$. Then by continuity we see $\|\operatorname{tr}_G T(x)\|\le 1$.

Define $S(x) := T(x)-x$ and denote by $S^i(x)$ for $1\le i\le n$ the i-th component of $S(x)$. By (2.30) of Corollary 2.4,

$$9n^2|K|\sqrt{\delta(G)} \ge |DS|(G^{(1)}).$$

The right hand side, being the total variation of DS, can hopefully be estimated from below by something like $\|-DS\|_*$, since then we can use the trace inequality. Indeed, the entire point of reducing to the set G is to apply this trace inequality. However, it is valid only for BV functions, and S is a vector field. So, we will apply the trace inequality to the components S^i and add these up. For now, we will try and massage the right-hand side of the above into $\|-DS^i\|(G^{(1)})$ multiplied by some constant. First,

$$|DS(x)| = \sqrt{\operatorname{trace}(DS)^TDS} = \sqrt{\sum_{i=1}^n\sum_{j=1}^n\left|\frac{\partial S^i}{\partial x_j}\right|^2} \ge \sqrt{\sum_{j=1}^n\left|\frac{\partial S^i}{\partial x_j}\right|^2} = |DS^i(x)|$$

when S is differentiable at x. I omit the justification when DS contains a singular piece, and use the previous as justification. Next, since $m_K\le\|\nu\|_*$ for all $\nu\in S^{n-1}$ by definition, applying this with $\nu=-x/|x|$ and applying homogeneity gives $m_K|x|\ge\|-x\|_*$. As $M_K\ge m_k$ (also by definition), we obtain $|x|\ge 1/M_K\|-x\|_*$. With the previous,

$$|DS(x)| \ge \frac{1}{M_K}\|-DS^i(x)\|_*.$$

Of course, the above is also only valid in a pointwise sense. However, when DS is absolutely continuous this is enough to show the same inequality at the level of total variations. Once again I omit the details in the generic case. As mentioned at the end of Section 2.2, we cannot use the regularity theory developed by Caffarelli. It surely works for the interior of G, but G is in general not open.

In any case, by applying the trace inequality (Lemma 3.1) we have

$$9n^2|K|\sqrt{\delta(G)} \geq \frac{1}{M_K}\|-DS^i\|_*(G^{(1)}) \geq \frac{m_K(\tau(G)-1)}{M_K^2}\inf_{c\in\mathbb{R}}\int_{\mathcal{F}G}|\operatorname{tr}_G(S^i-c)|\|\nu_G\|_*\,d\mathcal{H}^{n-1}.$$

By translating as necessary we may assume that $c=0$ and by our estimate of $\tau(G)$ we therefore have

$$\frac{9n^2}{k(n)}\frac{M_K^3}{m_K^2}|K|\sqrt{\delta(G)} \geq \int_{\mathcal{F}G}|\operatorname{tr}_G(S^i)|\|\nu_G\|_*\,d\mathcal{H}^{n-1}.$$

From the other side, let's start estimating $A(G)$. By testing the infimum at $x_0=0$ and $r=1$, and recalling $|G|=|K|$, we see that

$$A(G) \leq \frac{|G\Delta K|}{|G|} \leq \frac{2|G\setminus K|}{|G|} = \frac{2|G\setminus K|}{|K|}.$$

So, to estimate $A(G)$ in terms of $\delta(G)$, we need to bound $\int_{\mathcal{F}G}|\operatorname{tr}_G(S^i)|\|\nu_G\|_*d\mathcal{H}^{n-1}$ below by $|G\setminus K|$. Lemma 3.5 gives a similar result, but we need to connect it to what we have right now. Notice that $\operatorname{tr}_G(\cdot)$ is linear so that $\operatorname{tr}_G(S)=\operatorname{tr}_G(T)-\operatorname{tr}_G(x)=\operatorname{tr}_G(T)-x$, since x is smooth. Recalling that $\|\operatorname{tr}_G(T)\|\leq 1$, we have

$$\begin{aligned}|1-\|x\|| &\leq |1-\|\operatorname{tr}_G(T)\||+|\|\operatorname{tr}_G(T)\|-\|x\||\\ &\leq |1-\|\operatorname{tr}_G(T)\||+\|\operatorname{tr}_G(T)-x\| = (1-\|\operatorname{tr}_G(T)\|)+\|\operatorname{tr}_G(S)\|.\end{aligned}$$

We know how to control the integral of $|1-\|x\||$ over $\mathcal{F}G$, this is the content of Lemma 3.5. How do we control $\|\operatorname{tr}_G(S)\|$ in terms of $|\operatorname{tr}_G(S^i)|$? Simply note that $\sum_{i=1}^n|y_i|\geq|y|\geq m_K\|y\|$ by (3.2), and so

$$\frac{9n^3}{k(n)}\left(\frac{M_K}{m_K}\right)^3|K|\delta(G) \geq \int_{\mathcal{F}G}\|\operatorname{tr}_G(S)\|\|\nu_G\|_*\,d\mathcal{H}^{n-1}.$$

When combined with the above, and applying Lemma 3.5, we get

$$\begin{aligned}\frac{9n^3}{k(n)}\left(\frac{M_K}{m_K}\right)^3|K|\sqrt{\delta(G)}+\int_{\mathcal{F}G}(1-\|\operatorname{tr}_G(T))\|\nu_G\|_*\,d\mathcal{H}^{n-1} &\geq \int_{\mathcal{F}G}|1-\|x\||\|\nu_G\|_*\,d\mathcal{H}^{n-1}\\ &\geq \frac{m_K}{M_K}|G\setminus K| \geq \frac{1}{2}\frac{m_K}{M_K}|K|A(G).\end{aligned}$$

What about the $1-\|\operatorname{tr}_G(T)\|$ term? This is where (2.29) of Corollary 2.4 comes into play. We have that

$$\frac{n^3}{k(n)}\left(\frac{M_K}{m_K}\right)^3|K|\sqrt{\delta(G)} \geq n|K|\delta(G) \geq \int_{\mathcal{F}G}(1-\|\operatorname{tr}_G(T)\|)\|\nu_G\|_*\,d\mathcal{H}^{n-1}$$

where the first inequality makes use of the fact that $\delta(G)\leq 1$, $m_K/M_K\leq 1$, $k(n)\leq 1/2<1$, and $n\geq 2$. The role of this conversion is to place the upper bound in a form which combines nicely with what we previously have. Doing this yields

$$\frac{10n^3}{k(n)}\left(\frac{M_K}{m_K}\right)^3|K|\sqrt{\delta(G)} \geq \frac{1}{2}\frac{m_K}{M_K}|K|A(G)$$

which implies

$$A(G) \leq \frac{20n^3}{k(n)}\left(\frac{M_K}{m_K}\right)^4\sqrt{\delta(G)}.$$

This is of course the type of estimate we want, but it only applies for sets G where the trace inequality holds. However, we constructed G specially so that $\delta(G)$ can be estimated in terms of $\delta(E)$. In particular, we have

$$A(G) \leq \frac{20\sqrt{3}n^3}{k(n)^{3/2}}\left(\frac{M_K}{m_K}\right)^4\sqrt{\delta(E)}.$$

All that is left is to estimate $A(E)$ from above by $A(G)$ (and maybe additional terms involving $\delta(E)$). We have two parameters we can control: x_0 and r. By judicious choices of these, we will arrive at our estimate. First look at the choice of point. Let $\epsilon>0$ and select $x_{G,\epsilon}\in\mathbb{R}^n$ so that

$$|G|A(G)+\epsilon > |G\Delta(x_{G,\epsilon}+K)|.$$

Recall that $d(E,F) = |E\Delta F|$ is a metric on the space of Lebesgue measureable subsets of $\mathbb{R}^n$ (modulo the equivalence relation which identifies sets of the same measure). Then, for $r > 0$

$$\begin{aligned}|E|A(E) &\leq |E\Delta(x_{G,\epsilon} + rK)| \leq |E\Delta G| + |G\Delta(x_{G,\epsilon} + K)| + |(x_{G,\epsilon} + K)\Delta(x_{G,\epsilon} + rK)| \\ &< |E \setminus G| + |G|A(G) + |K\Delta(rK)| + \epsilon\end{aligned}$$

where we have applied the fact that $G \subseteq E$ to conclude $|E\Delta G| = |E \setminus G|$, and also have used the translation invariance of the Lebesgue measure. Dividing by $|E|$, applying the fact that $|G| \leq |E|$, and the arbitrariness of ϵ yields

$$A(E) < \frac{|E \setminus G|}{|E|} + A(G) + \frac{|K\Delta(rK)|}{|E|}.$$

We constructed G so that $|E \setminus G|/|E|$ can be estimated in terms of $\delta(E)$. Applying this yields

$$A(E) < \frac{\delta(E)}{k(n)} + A(G) + \frac{|K\Delta(rK)|}{|E|}.$$

How do we estimate $|K\Delta(rK)|$? Note if $r \geq 1$ then

$$|K\Delta(rK)| = |rK| - |K| = r^n|K| - |G|$$

since we assumed $|K| = |G|$. On the other hand, as $G \subseteq E$ we have $|E \setminus G| = |E| - |G|$. So, selecting $r = r_E := (|E|/|K|)^{1/n}$ works; notably $r_E \geq 1$ as $|K| = |G| \leq |E|$. Thus $|K\Delta(r_E K)| = |E \setminus G| \leq \delta(E)|E|/k(n)$. In total,

$$A(E) < \frac{2\delta(E)}{k(n)} + A(G).$$

As $\delta(E) \leq 1$, by recalling that $k(n) = (2 - 2^{(n-1)/n})/3$

$$\begin{aligned}A(E) \leq \frac{2\delta(E)}{k(n)} + A(G) &\leq \frac{2\sqrt{\delta(E)}}{k(n)} + \frac{20\sqrt{3}n^3}{k(n)^{3/2}}\left(\frac{M_K}{m_K}\right)^4\sqrt{\delta(E)} \\ &\leq \left(\frac{6}{2 - 2^{(n-1)/n}} + \frac{180}{(2 - 2^{(n-1)/n})^{3/2}}\right)n^3\left(\frac{M_K}{m_K}\right)^4\sqrt{\delta(E)} \\ &\leq \frac{181n^3}{(2 - 2^{(n-1)/n})^{3/2}}\left(\frac{M_K}{m_K}\right)^4\sqrt{\delta(E)}\end{aligned}$$

where we have applied the fact that $6/(2-2^{(n-1)/n}) \leq n^3/(2-2^{(n-1)/n})^{3/2}$ for $n \geq 2$. This concludes the first step. □

Step 2: We show

$$A(E) \leq C_0(n)\sqrt{\delta(E)}$$

where

$$C_0(n) = \frac{181n^7}{(2 - 2^{(n-1)/n})^{3/2}}.$$

Proof. By once again approximating E, we may assume E has smooth boundary. By John's Lemma [Joh48] [Theorem III] there exists an affine map $L_0 : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$B_1 \subset L_0(K) \subset B_n$$

and $\det(L_0) > 0$ (of course, by this we mean the determinant of the linear part of the affine map L_0). By scaling, we can find an $r > 0$ and affine map L such that

$$B_r \subset L(K) \subset B_{rn}$$

with $\det(L) = 1$. In general, we can define the relative asymmetry by

$$A(E,F) = \inf\left\{\frac{|E\Delta(x_0 + rF)}{|E|} \;\middle|\; x_0 \in \mathbb{R}^n,\ r^n|F| = |E|\right\}.$$

Of course, $A(E,K)$ is simply $A(E)$. Since L has determinant 1, it preserves volumes and thus $A(L(E), L(K)) = A(E)$. Note that $L(K)$ is still open, convex, and contains the origin: L is invertible,

so is an open map, it is affine and so maps affine hyperplanes to affine hyperplanes, and $0 \in B_r \subset L(K)$. By applying Step 1 to the pair $L(E)$ and $L(K)$ we obtain

$$A(E) = A(L(E), L(K)) \leq \frac{181n^3}{(2-2^{(n-1)/n})^{3/2}} \left(\frac{M_{L(K)}}{m_{L(K)}}\right)^4 \sqrt{\delta(L(E))}.$$

Let's look at the $\delta(L(E))$ term first. Since E has smooth boundary, we can use the limit formulation of the anisotropic perimeter:

$$P_K(E) = \lim_{\epsilon \to 0^+} \frac{|E + \epsilon K| - |E|}{\epsilon}.$$

But now because L is a volume preserving affine map (particularly, $L(E+F) = L(E) + L(F)$),

$$\begin{aligned} P_K(E) &= \lim_{\epsilon \to 0^+} \frac{|E + \epsilon K| - |E|}{\epsilon} = \lim_{\epsilon \to 0^+} \frac{|L(E + \epsilon K)| - |L(E)|}{\epsilon} \\ &= \lim_{\epsilon \to 0^+} \frac{|L(E) + \epsilon L(K)| - |L(E)|}{\epsilon} = P_{L(K)}(L(E)). \end{aligned}$$

It immediately follows that $\delta(L(E)) = \delta(E)$. By now we have essentially the same conclusion as in Step 1, but with some amount of freedom – notably, we have transformed the ratio M_K/m_K to $M_{L(K)}/m_{L(K)}$. And, as $B_r \subset L(K) \subset B_{nr}$ it follows that

$$\begin{aligned} M_{L(K)} &= \max_{\nu \in S^{n-1}} \sup_{x \in L(K)} \langle x, \nu \rangle \leq \max_{\nu \in S^{n-1}} \sup_{x \in B_{nr}} \langle x, \nu \rangle = \max_{\nu \in S^{n-1}} [nr|\nu|^2] = nr; \\ m_{L(K)} &= \min_{\nu \in S^{n-1}} \sup_{x \in L(K)} \langle x, \nu \rangle \geq \min_{\nu \in S^{n-1}} \sup_{x \in B_r} \langle x, \nu \rangle = \min_{\nu \in S^{n-1}} [r|\nu|^2] = r. \end{aligned}$$

Consequently, $M_{L(K)}/m_{L(K)} \leq n$, and we have

$$A(E) \leq \frac{181n^7}{(2-2^{(n-1)/n})^{3/2}} \sqrt{\delta(E)}$$

as desired. □

References

[Caf92] L. A. Caffarelli. The regularity of mappings with a convex potential. J. Amer. Math. Soc., 5(1):99–104, 1992.

[CL12] M. Cicalese and G. P. Leonardi. A selection principle for the sharp quantitative isoperimetric inequality. Arch. Ration. Mech. Anal., 206(2):617–643, 2012.

[FJ14] N. Fusco and V. Julin. A strong form of the quantitative isoperimetric inequality. Calc. Var. Partial Differential Equations, 50(3-4):925–937, 2014.

[FMP08] A. Fusco, F. Maggi, and A. Pratelli. The sharp quantitative isoperimetric inequality. Ann. of Math., 168(2):941–980, 2008.

[FMP10] A. Figalli, F. Maggi, and A. Pratelli. A mass transportation approach to quantitative isoperimetric inequalities. Inv. Math., 182(1):167–211, 2010.

[Fug89] B. Fuglede. Stability in the isoperimetric problem for convex or nearly spherical domains in $\mathbb{R}^n$. Trans. Amer. Math. Soc., 314:619–638, 1989.

[Hal92] R. R. Hall. A quantitative isoperimetric inequality in n-dimensional space. J. Reine Angew. Math., 428:161–176, 1992.

[HHW91] R. R. Hall, W. K. Hayman, and A. W. Weitsman. On asymmetry and capacity. J. d'Analyse Math, 56:87–123, 1991.

[Joh48] F. John. A refined Brunn-Minkowski inequality for convex sets. In Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948, pages 187–204. Interscience Publishers, Inc., New York, N. Y., 1948.

[Neu16] R. Neumayer. A strong form of the quantitative Wulff inequality. SIAM J. Math. Anal., 48(3):1727–1772, 2016.