THE ANISOTROPIC ISOPERIMETRIC INEQUALITY

ENRIQUE RIVERA

Abstract. We summarize an application of optimal transport to the isoperimetric inequality. In particular, we follow Figali et al to prove the anisotropic inequality, which uses a different notion of perimeter, and to show a sharp decay rate.

1. Introduction

1.1. **Overview.** Optimal mass transport seeks to find the cost minimizing way to transport one distribution of mass to another. This theory has played an important role in the proof of sharp versions several functional and geometric inequalities, including isoperimetric, Sobolev, log Sobolev, and Talagrand inequalities. The classical isoperimetric inequality relates the area and perimeter of a set $S \subset \mathbb{R}^n$, it roughly tells us that the area of a set of fixed perimeter is maximized when S is a ball. We will instead consider a different notion of perimeter, where different directions have different weights; this will be formalized in section 1.3. This will lead to our discussion of the proof of the anisotropic isoperimetric inequality, described in [FMP10].

1.2. **Optimal Transport.** We start with the Monge formulation of optimal transport. The setup is as follows: We are given two probability measures α and β on spaces $\mathcal{X}$ and $\mathcal{Y}$, and a cost function $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$. The goal is to find a map $T : \mathcal{X} \to \mathcal{Y}$ which minimizes

$$\inf_T \left\{ \int_{\mathcal{X}} c(x, T(x)) d\alpha(x) : T_{\#}\alpha = \beta \right\}.$$

Here T is called a Monge map or transport plan, since it tells us where in $\mathcal{Y}$ we transport each point $x \in \mathcal{X}$. In general, this formulation comes with several problems. Firstly, the set of maps T which satisfy the push-forward constraint is nonconvex, which leads to various difficulties in analyzing the problem. More importantly, such a map T may not even exits. Therefore, it becomes useful to study a relaxation of this problem.

The Kantorovich relaxation of optimal transport resolves these by looking for optimal couplings π of the measures instead of optimal maps, that is, we allow for mass-splitting. It is written as

$$\inf_\pi \{ \int_{\mathcal{X}\times\mathcal{Y}} c(x, y) d\pi(x, y) : \pi(\cdot, \mathcal{Y}) = \alpha, \pi(\mathcal{X}, \cdot) = \beta \}.$$

An important theorem in this area is the Monge-Kantorovich equivalence theorem, sometimes called Brenier's theorem or Brenier-McCann, which states that these formulations are equivalent under suitable conditions. We state it below and defer the proof to section 2.

Theorem 1.1. *Taking $\mathcal{X} = \mathcal{Y} = \mathbb{R}^n$ and $c(x, y) = \|x - y\|^2$, if α has density with respect to the Lebesgue measure, then the optimal coupling is unique and is induced by a Monge map*

Date: May 13,2024.

T. —Furthermore, this map is uniquely defined as $T(x) = \nabla\varphi(x)$, where φ is the unique, up to additive constant, convex function such that $(\nabla\varphi)_{\#}\alpha = \beta$.

1.3. **Perimeter.** This is the central concept of study. It turns out to be difficult to define in general, but it can be done. Here is a definition that we will use throughout the rest of the article.

Definition 1.2. *(Total variation measure) If μ is an $\mathbb{R}^n$ valued Borel measure on $\mathbb{R}^n$, we define its total variation measure $|\mu|$ to be the nonnegative measure satisfying*

$$|\mu|(E) = \sup\left\{\sum_{h\in\mathbb{N}} |\mu(E_h)| : \bigsqcup_{h\in\mathbb{N}} E_h \subseteq E\right\},$$

where E is any Borel set.

We start by describing the typical Euclidean notion of perimeter in dimensions $n \geq 2$. For any Borel set $E \subset \mathbb{R}^n$, we say that the perimeter of E is the total variation of the distributional derivative of its characteristic function $|D\chi_E|(\mathbb{R}^n)$, provided that the derivative is a measure. If the boundary of E is smooth, this definition agrees with the $(n-1)$-dimensional Hausdorff measure on $\mathbb{R}^n$ (see [EG15] for a definition) of the boundary:

$$P(E) = \mathcal{H}^{n-1}(\partial E). \tag{1}$$

For non-smooth boundaries, a similar result holds, where instead of the Hausdorff measure of the boundary, we take the Hausdorff measure of the reduced boundary (defined in section 4). This has several nice properties which we would expect from any reasonable notion of perimeter.

- Scaling law: $P(\lambda E) = \lambda^{n-1}P(E)$ for any $\lambda > 0$
- Translation invariance: $P(E + x) = P(E)$ for any $x \in \mathbb{R}^n$
- Invariance under orthogonal transformations: $P(A \cdot E) = P(E)$ for any $A \in O(n)$, where $A \cdot E$ denotes the action of A on E
- Invariance under complements: $P(E) = P(\mathbb{R}^n \setminus E)$

A natural generalization of the Euclidean notion of perimeter is that of anisotropic perimeter. To define this, we start with an open, bounded, and convex set K. We then define a weight function on directions through the Euclidean inner product:

$$\|\nu\|_* = \sup\{x \cdot \nu : x \in K\}.$$

Note that if K is the unit ball, this is simply the Euclidean norm, so this can be viewed as a generalization of it. It comes with a related quantity, which can also be viewed as a generalized Euclidean norm:

$$\|x\| = \inf\{\lambda > 0 : \frac{x}{\lambda} \in K\}.$$

It will always be clear from context whether this notation refers to the Euclidean norm or the above. These are related through a Cauchy-Schwarz type inequality:

$$x \cdot y \leq \|x\|\|y\|_*, \quad \forall x, y \in \mathbb{R}^n.$$

Let E be an open subset of $\mathbb{R}^n$ with smooth boundary. Letting ν_E denote the outer unit normal, the anisotropic perimeter is defined as

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

Comparing this to equation 1, we note that it coincides with our original notion of perimeter when $K = B$ is the unit ball in $\mathbb{R}^n$. For sets of finite perimeter where the boundary is not smooth, we replace the boundary of E with the reduced boundary:

$$P_K(E) = \int_{\mathcal{F}E} \|\nu_E(x)\|_* d\mathcal{H}^{n-1}(x).$$

We will discuss how to make sense of this integral in section 4. We also note that while the anisotropic perimeter satisfies the scaling law and translation invariance, it does not always satisfy invariance under orthogonal transformations or invariance under complements.

1.4. **Isoperimetric Inequalities.** The anisotropic isoperimetric inequality arises in the study of the anisotropic isoperimetric problem, which seeks to minimize

$$\inf\{\frac{P_K(E)}{|E|^{\frac{n-1}{n}}} : 0 < |E| < \infty\},$$

where $|E|$ denotes the Lebesgue measure of set E. The exponent of $\frac{n-1}{n}$ is the only interesting one, otherwise scaling the set could make this arbitrarily small. It turns out that the unique minimizer, up to translations and scalings, is the set K itself. In particular, the anisotropic isoperimetric inequality holds:

Theorem 1.3. *(Anisotropic isoperimetric inequality) For any open, bounded, and convex set K and set of finite perimeter E, the following inequality holds*

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

We may compare this to the standard isoperimetric problem, where the unique minimizer is B, up to translations and scaling, and the inequality holds with $K = B$.

1.5. **Organization.** The remainder of this summary will be devoted to building up the theory of Brenier maps and proving theorem 1.1 in section 2, and using this to prove theorem 1.3 in section 5. To build up to this final proof, we will discuss some properties of convex functions in 3 and introduce concepts of geometric measure theory in 4.

2. Brenier Map

An important concept in the study of optimal transport is duality. It allows us to find the minimum and minimizer by analyzing a different, and often easier, maximization problem. Eventually, it will connect back to the isoperimetric problem, as we will transport the set E to the set K. We begin by setting up necessary definitions.

Definition 2.1. *(Dual problem) The dual of the Kantorovich formulation of optimal transport is as follows: Let α and β be probability measures on $\mathbb{R}^n$. Maximize the value of*

$$\int_{\mathbb{R}^n} \varphi(x)d\alpha(x) + \int_{\mathbb{R}^n} \psi(x)d\beta(x)$$

among all integrable functions such that

$$\varphi(x) + \psi(x) \leq c(x, y), \quad \forall x, y \in \mathbb{R}^n$$

If φ, ψ are maximizers, we call them Kantorovich potentials.

Definition 2.2. *(c-transform) Let $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$. Its c_+ transform $\psi^{c+} : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ is defined as*

$$\psi^{c+} = \inf_{y \in \mathbb{R}^n} c(x, y) - \psi(y).$$

Definition 2.3. *(c-concavity) We say that $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ is c-concave if there exists $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ such that $\varphi = \psi^{c+}$.*

Definition 2.4. *(c-superdifferential) Let $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ be a c-concave function. Define the c-superdifferential as*

$$\partial^{c+}\varphi = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : \varphi(x) + \varphi^{c+}(y) = c(x, y)\}.$$

We now describe a sufficient and necessary condition for c-concavity, which will let us apply results related to the c-superdifferential. From now on, we take the cost function to be squared Euclidean distance.

Proposition 2.5. *Let $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ be a c-concave function. Then φ is c-concave if and only if $\bar{\varphi} : x \mapsto |x|^2 - \varphi(x)$ is convex and lower semicontinuous. Furthermore, $\partial^{c+}\varphi$ is contained in the set of subderivatives of $\bar{\varphi}$.*

An important duality theorem which we will use in our proof of Brenier's theorem is as follows, which we state without proof for the sake of brevity. For a proof, see [AG13].

Theorem 2.6. *(Duality) Let α and β be probability measures on $\mathbb{R}^n$ and let $c : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous function bounded from below. Suppose there exist $a \in L^1(\alpha)$ and $b \in L^1(\beta)$ such that $c(x, y) \leq a(x) + b(y)$. Then the minimum of the Kantorovich problem 1.2 and is equal to the supremum of the dual problem 2.1.*

Furthermore, the supremum of the dual problem is attained, and the maximizing couple (φ, ψ) is of the form (φ, φ^{c+}) for some c-concave function φ.

We now state the reason why c-superdifferentiability will become important. This is a well-known corollary of the so-called fundamental theorem of optimal transport.

Proposition 2.7. *Under the assumptions of the previous theorem, if π is an optimal coupling, then supp$(\pi) \subset \partial^{c+}\varphi$.*

At this point all the necessary theorems and definitions to prove Brenier's theorem, which we will restate here for convenience below for convenience, are set up.

Theorem 2.8. *(Brenier's theorem) Taking $\mathcal{X} = \mathcal{Y} = \mathbb{R}^n$ and $c(x, y) = \|x - y\|^2$, if α has density with respect to the Lebesgue measure, then the optimal coupling is unique and is induced by a Monge map T. —Furthermore, this map is uniquely defined as $T(x) = \nabla\varphi(x)$ where $\varphi : \mathbb{R}^n \to \mathbb{R}$ is some convex function.*

Proof. (Sketch.) For ease of exposition, we assume that $|x|^2$ is integrable with respect to both α and β. These conditions are not truly necessary, but the proof becomes more technical. We also skip the details of uniqueness, since we will not need this for our purposes.
Note that $\|x - y\|^2 \leq (\|x\| + \|y\|)^2 \leq 2\|x\|^2 + 2\|y\|^2$, where the first component is integrable with respect to α, and the second is integrable with respect to β, by our assumption that these were finite. Thus theorem 2.6 applies. That is, the primal problem has a solution, and the dual is maximized at (φ, φ^{c+}) for some c-concave function φ. By proposition 2.7, if π is an optimal coupling, then $\text{supp}(\pi) \subset \partial^{c+}\varphi$. By proposition 2.5, we get that $\bar{\varphi} : x \mapsto |x|^2 - \varphi(x)$

is concave and that $\partial^{c+}\varphi$ is contained in the set of subderivatives of $\bar{\varphi}$. But convex functions are almost everywhere differentiable with respect to the Lebesgue measure, and since α is absolutely continuous, the same holds for this measure. Thus $\nabla\varphi : \mathbb{R}^n \to \mathbb{R}^n$ is defined $\alpha-$a.e. and the map π must be concentrated on its graph. Hence the optimal coupling π is unique and is induced by the gradient of the convex function $\bar{\varphi}$. □

Remark 2.9. *Since monotone maps are of (locally) bounded variation (BV), the Monge map constructed here has (locally) bounded variation. This fact also follows from the fact that distributional second derivative of a convex function is a Radon measure (taking values in $R^{n\times n}$). We discuss this in more detail in the sections which follow. This fact lets us apply several theorems from geometric measure theory (GMT). The map which results from this also has an intricate regularity theory surrounding it. For more details, see for example [AG13], which we have been following in this section.*

Lastly, we include a property that will become useful in section 5.

Lemma 2.10. *Let α and β be probability measures on $\mathbb{R}^n$ with densities with respect to Lebesgue, and let φ be a convex function on $\mathbb{R}^n$ with $\boldsymbol{\nabla}\varphi_{\#}\alpha = \beta$. Then (identifying measures with their densities), we have that $\beta(x) = \frac{\alpha(x)}{\det \boldsymbol{\nabla}^2\phi(x)}$ almost everywhere.*

Proof. This is essentially the Jacobian factor one would expect from a change of variables. See [McC97] for details. □

3. Derivatives of Convex Functions

Brenier's theorem is useful because not only does it give us existence of a map, but some regularity. It turns out that being the gradient of a convex function results in several nice properties, which we will describe in this section, and later use in our proof of the isoperimetric inequality. In the ensuing discussion, we closely follow the main results of [DUD77]. We start with an elementary theorem and follow with an important result.

Theorem 3.1. *(Derivative of a convex function) If $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a convex function, then the derivative $\boldsymbol{\nabla}\varphi$ exists except at most on a countable set.*

Theorem 3.2. *(Second derivative of a convex function) Suppose $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a convex function, with derivative defined everywhere except at a set E. Then the second derivative $\boldsymbol{\nabla}^2\varphi$ exists and is positive semidefinite a.e. with respect to Lebesgue, where the second derivative is defined by taking limits on $\mathbb{R}^n \setminus E$.*

We say that derivatives are taken in the Alexandrov sense if they are defined as in theorem 3.2. It will also be of interest to study the distributional derivatives of convex functions. The following is analogous to the elementary theorem which says that the second derivative of a twice differentiable convex function is positive semi-definite.

Theorem 3.3. *(Distributional second derivative of a convex function) Let $\phi \in \mathcal{D}'(\mathbb{R}^n, \mathbb{R})$ be a distribution. Then $\phi = [\varphi]$ for some convex function φ if and only if $D^2\phi := \{\frac{\partial^2}{\partial x_i \partial x_j}\phi\}$ is a matrix $\{\mu_{ij}\}$ of signed Radon measures on $\mathbb{R}^n$, and is such that for any Borel set B, $\{\mu_{ij}(B)\}$ is a positive semi-definite matrix.*

These two different views are also compatible. Recall that, by the Lebesgue decomposition theorem, any signed measure μ can be decomposed uniquely into $\mu_{ac} + \mu_{sing}$, where μ_{ac} has density with respect to Lebesgue, and μ_{sing} is supported on a set of 0 Lebesgue measure. Note that in particular, both parts are positive semidefinite. We can say the following.

Theorem 3.4. *(Compatibility of notions) Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a convex function. The Hessian matrix of pointwise second derivatives taken in the Alexandrov sense is equal to the density of the absolutely contionuous part of the second distributional derivative (written $\frac{d}{dx}(D^2[\varphi])_{ac}$).*

As a result of this theorem, we can write $T = \boldsymbol{\nabla}\varphi$ and $DT = \boldsymbol{\nabla}T dx + D_sT$, where D_sT is the singular part of the measure. Again, we emphasize that both summands here are positive semidefinite. We fix notation for the rest of the article:

- By a capitalized D we will mean the derivative in the distributional sense
- By a subscript s, we will denote the singular part of a measure in its Lebesgue decomposition

This will become important in section 5, where we will see an interplay of these different types of derivatives.

4. Geometric Measure Theory

In this section, we develop what we saw in 1.3 further. This will let us state and prove our theorems in the most general setting. In particular, the anisotropic isoperimetric inequality will hold for any measureable set (of finite volume). This is in contrast to [AG13], where sufficient smoothness is assumed.

Definition 4.1. *(Finite perimeter) We say a measureable set E has finite perimeter if the distributional gradient of its indicator function $D\chi_E$ is an $\mathbb{R}^n$ valued Borel measure on $\mathbb{R}^n$ with $|D\chi_E|(\mathbb{R}^n) < \infty$.*

Definition 4.2. *(Measure-theoretic outer unit normal) Let E be a set of finite perimeter. The measure-theoretic outer unit normal is*

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

whenever this is a well-defined limit and is a unit vector.

Definition 4.3. *(Reduced boundary) The reduced boundary of a set of finite perimeter E, denoted $\mathcal{F}E$, is the set of points $x \in \mathbb{R}^n$ for which $\nu_E(x)$ exists (and is a unit vector).*

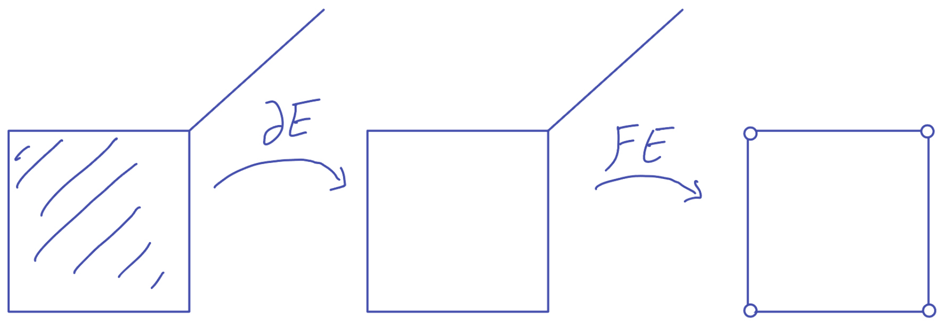
Intuitively, the outer unit normal is the direction in which the negative gradient points, but averaged out to become a more robust value. This is the direction of fastest decrease of the indicator function, so it tells us in which direction to move to exit the set as quickly as possible. Just to check that the objects involved are compatible, note $D\chi_E$ is a vector-valued measure by our finite perimeter assumtion, $|D\chi_E|$ is a scalar-valued measure, and $B_r(x)$ is a set, so we are dividing a vector by a scalar.

Definition 4.4. *(Density) We denote $E^{(\lambda)}$ the set of points $x \in \mathbb{R}^n$ which have density λ with respect to E. That is,*

$$\lim_{r\to 0^+} \frac{|E \cap B_r(x)|}{|B_r(x)|} = \lambda.$$

For example, if $\lambda = 1$, we get the essential interior of the set E. That is, the set of points such that almost every point close to it is in the set. If $\lambda = 0$, we get the essential exterior for similar reasons. It turns out that $\lambda = \frac{1}{2}$ is of particular interest, since $\mathcal{F}E \subseteq E^{(\frac{1}{2})}$ and is equal to it up to $\mathcal{H}^{n-1}$-null sets. In fact, the essential boundary $\partial^* E = \mathbb{R}^n \setminus (E^{(1)} \cup E^{(0)})$ contains both of them and is equal in the same sense.

Example 4.5. *Consider E a unit square with an extra line segment. In this case $E^{(\frac{1}{2})}$ agrees with the reduced boundary, and it is the boundary of unit square without the corners. The measure-theoretic unit normal is undefined on the extra line segment, and it is not a unit vector on the corners. In particular, the line sticking out does not contribute to the perimeter, and we have $P(E) = \mathcal{H}^1(\mathcal{F}E) = 4$.*



Definition 4.6. *(Bounded variation) A function $T : \mathbb{R}^n \to \mathbb{R}^n$ is of bounded variation if its distributional derivative is a finite (vector-valued) Radon measure on $\mathbb{R}^n$. We say that $T \in BV(\mathbb{R}^n; \mathbb{R}^n)$.*

We have already encountered a few examples of these. Sets of finite perimeter are sets with characteristic functions of bounded variation, and the derivatives of convex functions discussed in section 3 also have this property.

Definition 4.7. *(Inner trace) Whenever $T \in BV(\mathbb{R}^n; \mathbb{R}^n)$ and E is of finite perimeter, we have that for $\mathcal{H}^{n-1}$—a.e. $x \in \mathcal{F}E$, there exists a vector $tr_E(T)(x) \in \mathbb{R}^n$ such that*

$$\lim_{r \to 0^+} \frac{2}{|B_r(x)|} \int_{E \cap B_r(x)} |T(y) - \mathrm{tr}_E(T)(x)| dy = 0.$$

We call this vector the inner trace of T on E.

This is a sort of average value of T around x, which would just be $T(x)$ if T were continuous. We could have chosen the normalization to be by $\frac{1}{r^n}$ instead and arrived to the same definition, but writing it like this is more suggestive.

5. Anisotropic Isoperimetric Inequality

The first inequality in this direction is the following.

Lemma 5.1. *If H is a positive semidefinite $n \times n$ matrix, then $n(\det(H))^{\frac{1}{n}} \le \mathrm{Tr}(H)$.*

Proof. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of H with multiplicities. Since H is positive semidefinite, all eigenvalues are nonnegative. Thus by the arithmetic mean-geometric mean inequality,

$$n(\det(H))^{\frac{1}{n}} = n(\lambda_1 \ldots \lambda_n)^{\frac{1}{n}} \le \lambda_1 + \ldots + \lambda_n = \mathrm{Tr}(H),$$

as desired. $\square$

The next theorem is known as Lebesgue's differentiation theorem, and the corollary is Lebesgue's density theorem. These can be found in most standard measure theory books, for instance [EG15]. We encourage the reader to go back to the unit square example and check that these hold with $f = \chi_E$.

Theorem 5.2. *(Lebesgue's differentiation theorem) Let $f \in L^1(\mathbb{R}^n)$. Then for a.e. $x \in \mathbb{R}^n$, we have that*

$$\lim_{r \to 0^+} \frac{1}{|B_r(x)|} \int_{B_r(x)} f(y)dy = f(x).$$

Corollary 5.3. *If E is a measureable set such that $|E| < \infty$, then $E^{(1)} = E$ up to null sets.*

Proof. Letting $f = \chi_E$ in the previous theorem, we have that for almost every $x \in \mathbb{R}^n$,

$$\lim_{r \to 0^+} \frac{|E \cap B_r(x)|}{|B_r(x)|} = \chi_E(x).$$

Thus almost all points in E have density 1, so $E \setminus E^{(1)}$ is null, and almost all points in $\mathbb{R}^n \setminus E$ have density 0, so $E^{(1)} \setminus E$ is null. □

The following theorem appears in [FMP10] in a stronger form. It is only needed for a technical point of the main proof, which lets us apply the generalized divergence theorem, which we will state soon.

Proposition 5.4. *(Approximation of finite perimeter sets) For any set of finite perimeter E, there exists a sequence of bounded sets E_h such that*

$$|E_h \Delta E| \to 0 \text{ and } P_K(E_h) \to P_K(E).$$

Remark 5.5. *This proposition is also true if we also require that E_h is open and has smooth or polyhedral boundary. This can simplify our main proof, but does not work for sharper estimates which appear in the paper we are summarizing.*

We have been building up to have the ability to state the following theorem. The idea will be to apply this to a Brenier map on some bounded set, which we will have from theorem 5.4.

Theorem 5.6. *(Generalized divergence theorem) If $T \in BV(\mathbb{R}^n; \mathbb{R}^n)$ and E is a set of finite perimeter, then the divergence theorem holds in the form*

$$\text{Div } T(E^{(1)}) = \int_{\mathcal{F}E} \text{tr}_E(T) \cdot \nu_E d\mathcal{H}^{n-1},$$

where Div *denotes the distributional divergence.*

This is all the theory we need. We can now prove the anisotropic isoperimetric inequality, which we will restate here for convenience. The two big steps are to construct a useful Brenier map, and to analyze it using the divergence theorem. For a discussion of when the inequality is strict, see [FMP10].

Theorem 5.7. *(Anisotropic isoperimetric inequality) For any open, bounded, and convex set K containing the origin and set of finite volume E, the following inequality holds*

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

Proof. If E has infinite perimeter or 0 volume, then there is nothing to prove. So we assume that E is a set of finite perimeter and positive volume. By the scaling properties of anisotropic perimeter and the Lebesgue measure, we have that for any $\lambda > 0$, this inequality is equivalent to

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

Letting $\lambda = (\frac{|K|}{|E|})^{\frac{1}{n}}$, we restrict our attention to sets with $|E| = |K|$. Approximating E as in theorem 5.4, we may further assume that E is bounded. Let $\alpha = \frac{1}{|E|}\chi_E(x)dx$ and $\beta = \frac{1}{|K|}\chi_K(x)dx$. By theorem 1.1, there exists a convex function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $T = \boldsymbol{\nabla}\varphi$ a.e. satisfies $T_{\#}\alpha = \beta$. Modifying T on a set of measure 0, we force $T(E) \subseteq K$ and preserve all other properties. By theorems 3.2 and 3.4, we also have that $\boldsymbol{\nabla} T$ exists and is a positive semidefinite $n \times n$ matrix almost everywhere. By theorem 2.10, it follows that

$$\det \boldsymbol{\nabla} T = 1, \quad \text{for a.e. } x \in E.$$

Then by theorem 5.1 we have that for a.e. $x \in E$,

$$n = n(\det \boldsymbol{\nabla} T(x))^{\frac{1}{n}} \leq \operatorname{Tr} \boldsymbol{\nabla} T(x) = \operatorname{div} T(x)$$

Also, since the eigenvalues of $D_s T$ are nonnegative, we have that

$$\operatorname{Div} T - \operatorname{div} T dx = (\operatorname{Div} T)_s = \operatorname{Tr}(D_s T)$$

is a positive Radon measure on $\mathbb{R}^n$. By the above discussion, corollary 5.3, and by the generalized divergence theorem 5.6 (using that E is bounded),

$$n|K|^{\frac{1}{n}}|E|^{\frac{n-1}{n}} = n|E| = \int_E n dx$$
$$\leq \int_E \operatorname{div} T(x) dx = \int_{E^{(1)}} \operatorname{div} T(x) dx$$
$$\leq \operatorname{Div} T(E^{(1)}) = \int_{\mathcal{F}E} \operatorname{tr}_E(T) \cdot \nu_E d\mathcal{H}^{n-1}.$$

By 4.7, since T takes values in K on E, we obtain $\|\operatorname{tr}_E(x)\| \leq 1$ for $\mathcal{H}^{n-1}$-a.e. $x \in \mathcal{F}E$. Then by our Cauchy-Schwarz 1.3, we continue our chain of inequalities. We have

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

as desired. □

The bounds contained in this proof can actually be used to prove a sharper inequality. We call $\delta(E) = P_K(E) - n|K|^{\frac{1}{n}}|E|^{\frac{n-1}{n}}$ the isoperimetric deficit of E. If this is deficit small enough, one can argue that there exists some maximal critical subset of E, with volume controlled by $\delta(E)$, which we can remove in order to get a set which is much nicer to deal with. In particular, trace and Sobolev inequalities apply. After several estimates, the following improvement is found:

Definition 5.8. *(Asymmetry index) The asymmetry index of E is defined as*

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

Theorem 5.9. *(Sharp anisotropic isoperimetric inequality) Under the same assumptions as the previous theorem,*

$$P_K(E) \geq n|K|^{\frac{1}{n}}|E|^{\frac{n-1}{n}} \left(1 + \frac{A(E)}{C(n)}\right)^2$$

for some constant $C(n)$ depending on the dimension.

This is sharp in the sense that the exponent of 2 is the smallest for which such an inequality holds. The value of $C(n)$ can be chosen to be

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

For the detailed proof of this theorem, refer to section 3 of [FMP10].

References

[AG13] L. Ambrosio and N. Gigli, "A user's guide to optimal transport," in *Modelling and Optimisation of Flows on Networks: Cetraro, Italy 2009, Editors: Benedetto Piccoli, Michel Rascle*, L. Ambrosio, A. Bressan, D. Helbing, A. Klar, and E. Zuazua, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 1–155, isbn: 978-3-642-32160-3. doi: `10.1007/978-3-642-32160-3_1`. [Online]. Available: `https://doi.org/10.1007/978-3-642-32160-3_1`.

[DUD77] R. M. DUDLEY, "On second derivatives of convex functions," *Mathematica Scandinavica*, vol. 41, no. 1, pp. 159–174, 1977, Publisher: Mathematica Scandinavica, issn: 00255521, 19031807. [Online]. Available: `http://www.jstor.org/stable/24490947` (visited on 04/16/2024).

[EG15] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions* (Textbooks in mathematics), Revised ed. Boca Raton, Fla.: CRC Press, 2015, 299 pp., isbn: 978-1-4822-4239-3 978-1-4822-4238-6.

[FMP10] A. Figalli, F. Maggi, and A. Pratelli, "A mass transportation approach to quantitative isoperimetric inequalities," *Invent. math.*, vol. 182, no. 1, pp. 167–211, Oct. 2010, issn: 0020-9910, 1432-1297. doi: `10.1007/s00222-010-0261-z`. [Online]. Available: `http://link.springer.com/10.1007/s00222-010-0261-z` (visited on 04/16/2024).

[McC97] R. J. McCann, "A convexity principle for interacting gases," *Advances in Mathematics*, vol. 128, no. 1, pp. 153–179, Jun. 1997, issn: 00018708. doi: `10.1006/aima.1997.1634`. [Online]. Available: `https://linkinghub.elsevier.com/retrieve/pii/S0001870897916340` (visited on 04/16/2024).