A Short Course on Rearrangement Inequalities

Almut Burchard

June 2009

These notes grew out of introductory courses for graduate students that I gave at the First IMDEA Winter School in Madrid in January 2009, and at the Università di Napoli "Federico II" in April 2009. The manuscript has been been slightly expanded to five sections, each providing material for one or two hours of lecture.

The first section gives an overview of the classical rearrangement inequalities and their applications. The relationship between a function and its level sets is explored through the layer-cake decomposition and the co-area formula. The next couple of sections investigate sequences of rearrangements, such as iterated Steiner symmetrizations and polarizations. Both compactness arguments and the "competing symmetries" technique are described. The fourth section contains the proof of the three fundamental inequalities due to Riesz, Pólya-Szegő and Talenti. Though only the simplest cases are stated, the proofs will more generally. As an afterthought, the fifth section gives a brief introduction to optimal transportation techniques for proving geometric inequalities.

The text is accompanied by exercises, which could serve as starting points for discussions. References at the end of each section are not encyclopedic, but point to books, monographs, surveys, and original papers that the reader might find immediately related to each chapter.

Acknowledgments. I would like to thank Gerhard Huisken (MPI for Gravitational Physics in Golm) and Nicola Fusco (Università di Napoli "Federico II') for their hospitality and generous financial support, and to Adèle Ferone for hosting me at the Università Seconda di Napoli at Caserta and taking care of me in Napoli. Many thanks also to José Arrieta, Aldo Pratelli and Julio Rossi for inviting me to IMDEA.

Contents

1 The symmetric decreasing rearrangement

Rearrangements manipulate the shape of a geometric object while preserving its size. They are used in the Calculus of Variations to find extremals of geometric functionals. Here, we will study the **symmetric decreasing rearrangement**, which replaces a given nonnegative function f by a radial function f^*.

1.1 Definition and basic properties

Let A be a measurable set of finite volume in $\mathbb{R}^n$. Its **symmetric rearrangement** A^* is the open centered ball whose volume agrees with A,

$$A^* = \left\{x \in \mathbb{R}^n \mid \omega_n |x|^n < \mathrm{Vol}(A)\right\}.$$

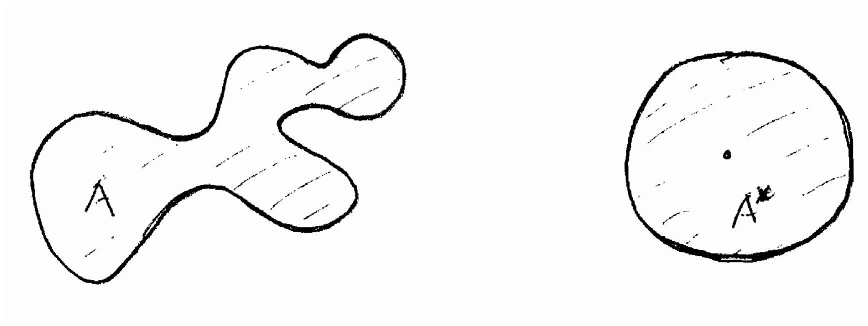


Figure 1.1: A^* is the centered ball of the same volume as A.

Let f be a nonnegative measurable function that **vanishes at infinity**, in the sense that all its positive level sets have finite measure,

$$\mathrm{Vol}\big(\{x \mid f(x) > t\}\big) < \infty, \quad (\text{for all } t > 0).$$

We define define the **symmetric decreasing rearrangement** f^* of f by symmetrizing its the level sets,

$$f^*(x) = \int_0^\infty \mathcal{X}_{\{f(x)>t\}^*}\, dt.$$

Then f^* is lower semicontinuous (since its level sets are open), and is uniquely determined by the **distribution function**

$$\mu_f(t) = \mathrm{Vol}\big(\{x \mid f(x) > t\}\big). \tag{1.1}$$

By construction, f^* is **equimeasurable** with f, i.e., corresponding level sets of the two functions have the same volume,

$$\mu_f(t) = \mu_{f^*}(t), \quad (\text{all } t > 0). \tag{1.2}$$

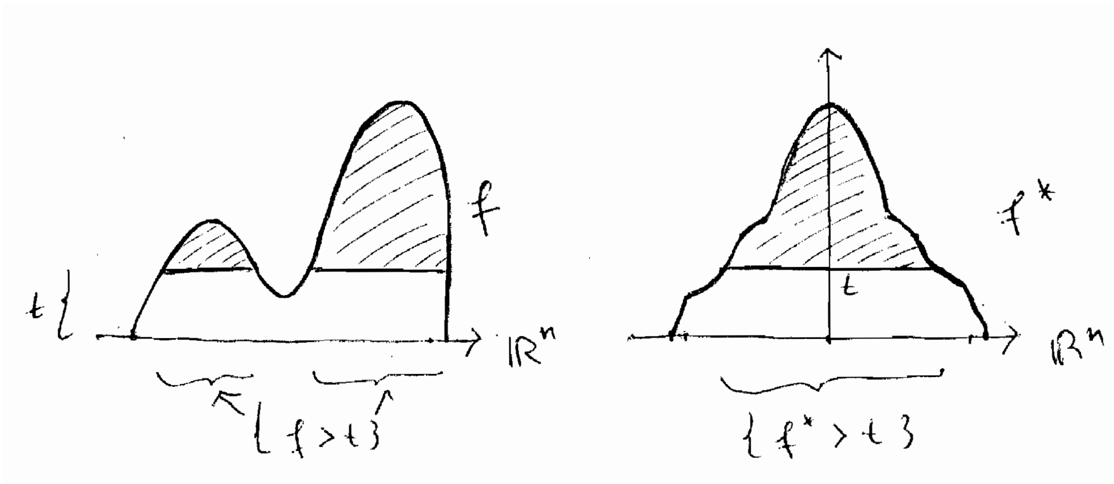


Figure 1.2: f^* is radially decreasing and equimeasurable with f.

Exercise 1.1 *Convince yourself that the definitions of A^* and f^* are consistent,*

$$\mathcal{X}_{A^*} = \big(\mathcal{X}_A\big)^*, \quad \big\{x \mid f(x) > t\big\}^* = \big\{x \mid f^*(x) > t\big\}.$$

Exercise 1.2 *For $a, b > 0$, find the symmetric decreasing rearrangement of the function*

$$f(x) = \begin{cases} 1 - ax\,, & 0 \le x \le a^{-1}\,, \\ 1 + bx\,, & -b^{-1} \le x \le 0\,, \\ 0\,, & \textit{otherwise}\,. \end{cases}$$

Question 1.3 *If f is smooth, does it follow that f^* is differentiable?*

1.2 Functions and their level sets

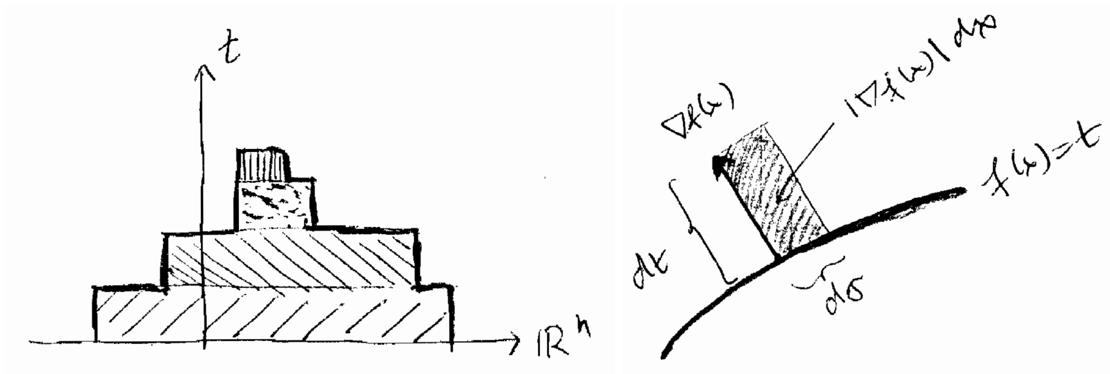
Our definition of f^* used a special case of the **layer-cake decomposition**, which expresses a nonnegative function f in terms of its level sets as

$$f(x) = \int_0^\infty \mathcal{X}_{\{f(x)>t\}}\, dt\,. \tag{1.3}$$

Note that the characteristic function $\mathcal{X}_{\{f(x)>t\}}$ is jointly measurable in x and t, provided that f is measurable. This allows to reduce statements about functions to statements about their level sets:

Lemma 1.4 (Rearrangement preserves L^p-norms) *For every nonnegative function f in $L^p(\mathbb{R}^n)$,*

$$||f||_p = ||f^*||_p \quad 1 \le p \le \infty\,, \tag{1.4}$$



(a) Layer-cake decomposition. (b) Co-area formula.

Figure 1.3: A nonnegative function and its level sets.

PROOF. We use the layer-cake decomposition and Fubini's theorem to write

$$\begin{aligned}\int_{\mathbb{R}^n} |f(x)|^p\,dx &= \int_{\mathbb{R}^n}\int_0^\infty \mathcal{X}_{\{f(x)^p>t\}}\,dtdx\\ &= \int_0^\infty \mathrm{Vol}\big(\{f(x)^p>t\}\big)\,dt\\ &= \int_0^\infty \mathrm{Vol}\big(\big(\{f(x)>s\}\big)\,ps^{p-1}\,ds\\ &= \int_0^\infty \mu_f(s)\,ps^{p-1}\,ds\,.\end{aligned}$$

The claim follows since f^* is equimeasurable with f. □

Exercise 1.5 *Prove that symmetric decreasing rearrangement is* **order-preserving**

$$f(x)\le g(x)\,\textit{for all}\ x\in\mathbb{R}^n \implies f^*(x)\le g^*(x)\,\textit{for all}\ x\in\mathbb{R}^n\,. \tag{1.5}$$

Lemma 1.6 (Hardy-Littlewood inequality) *If f and g are nonnegative measurable functions that vanish at infinity, then*

$$\int fg \le \int f^*g^*\ , \tag{1.6}$$

in the sense that the left hand side is finite whenever the right hand side is finite.

PROOF. Consider first the case where $f=\mathcal{X}_A$ and $g=\mathcal{X}_B$ are characteristic functions of measurable sets A and B of finite volume. The rearrangements A^* and B^* are centered balls, and their intersection $A^*\cap B^*$ is the smaller of the two balls. Thus,

$$\mathrm{Vol}(A^*\cap B^*) = \min\{\mathrm{Vol}(A),\mathrm{Vol}(B)\} \ge \mathrm{Vol}(A\cap B)\,,$$

which proves the inequality in this case. In general, the layer-cake decomposition and Fubini's theorem allow to rewrite the left hand side of Eq. (1.6) as

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)g(x)\,dx &= \int_{\mathbb{R}^n}\int_0^\infty\int_0^\infty \mathcal{X}_{\{f(x)>s\}}\mathcal{X}_{\{g(x)>t\}}\,dsdtdx\\ &= \int_0^\infty\int_0^\infty \mathrm{Vol}\big(\{f>s\}\cap\{g>t\}\big)\,dsdt\,,\end{aligned}$$

and correspondingly for the right hand side. Since we have already shown that the integrand increases under rearrangement, the claim follows. □

Exercise 1.7 *Prove that the symmetric decreasing rearrangement* **decreases L^p-distances**

$$||f-g||_p \geq ||f^*-g^*||_p \quad 1\leq p\leq\infty\,. \tag{1.7}$$

Hint: Write

$$|f(x)-g(x)|^p = p\int_0^\infty [f(x)-t]_+^{p-1}\mathcal{X}_{\{g(x)\leq t\}} + [g(x)-t]_+^{p-1}\mathcal{X}_{\{f(x)\leq t\}}\,dt\,,$$

then replace $\mathcal{X}_{\{f(x)\leq t\}}$ by $1-\mathcal{X}_{\{f(x)>t\}}$ and correspondingly for g.

It is clear from the proofs that Eq. (1.6) and Eq. (1.7) hold with equality, if f and g have the same family of level sets, i.e., if and only if $\big(f(x)-g(x)\big)\big(f(y)-g(y)\big)\geq 0$ for almost all x,y.

We next introduce a useful tool that replaces the layer-cake decomposition for integrals that involve the gradient of a function $f\in W^{1,p}$. The **co-area formula** says that

$$\int g(x)|\nabla f(x)|\,dx = \int_0^\infty\int_{f^{-1}(t)} g(x)\,d\sigma(x)dt$$

for every measurable function g such that the left hand side is well-defined. Some care is needed when evaluating the right hand side: We assign to $f(x)$ the value of the Lebesgue density limit of f (whenever defined), and the integration $d\sigma$ is with respect to $(n-1)$-dimensional Hausdorff measure. An immediate consequence of the co-area formula is the identity

$$||\nabla f||_1 = \int_0^\infty \mathrm{Per}(\{f>t\})\,dt\,.$$

If f is smooth, then the co-area formula defines a local change of variables in any region where ∇f does not vanish. But note that the co-area formula gives no information on the set of critical points of f. For instance,

$$\int_{t_1}^{t_2}\int_{f^{-1}(t)} |\nabla f|^{-1}\,d\sigma dt = \mathrm{Vol}\big(\{x \mid t_1<f(x)\leq t_2\,, |\nabla f(x)|\neq 0\}\big) \tag{1.8}$$

for every interval $(t_1,t_2]$.

Exercise 1.8 *Prove Eq. (1.8), using*

$$|\nabla f(x)|^{-1} = \lim_{\varepsilon \to 0_+} \big(\varepsilon + |\nabla f(x)|\big)^{-1}$$

and the co-area formula. Why can you exchange the limit with the integrals?

1.3 Classical rearrangement inequalities

By construction, the symmetric decreasing rearrangement concentrates the mass of functions near the origin. A subtle expression of this concentration is **Riesz' inequality**

$$\int_{\mathbb{R}^n} f(x)\, g{*}h(x)\, dx \le \int_{\mathbb{R}^n} f^*(x)\, g^*{*}h^*(x)\, dx\,. \tag{1.9}$$

This implies that $g^* * h^*$ dominates $g * h$ in the sense that

$$\int_B (g*h)^*(x)\, dx = \sup_{C:\mathrm{Vol}(C)=\mathrm{Vol}(B)} \int_C g*h(x)\, dx \le \int_B g^* * h^*(x)\, dx$$

for every centered ball B.

Many applications of Riesz' inequality require only the (much simpler) special case where $h(x) = H(|x|)$ is a known strictly symmetrically decreasing function,

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(y)H(|x-y|)\, dxdy \le \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f^*(x)g^*(y)H(|x-y|)\, dxdy\,, \tag{1.10}$$

such as the Coulomb kernel $|x-y|^{-1}$ on $\mathbb{R}^3$ or the heat kernel $(4\pi t)^{-n/2}e^{-\frac{|x-y|^2}{4t}}$. Note that we can recover the Hardy-Littlewood inequality from Eq. (1.10) by letting $t \to 0$ in the heat kernel.

The **Pólya-Szegő** inequality

$$||\nabla f||_p \ge ||\nabla f^*||_p\,, \quad 1 \le p \le \infty\,, \tag{1.11}$$

implies in particular that the kinetic energy in quantum mechanics will decrease under symmetric deceasing rearrangement of the wave function. The case $p = 1$ contains as a limit the **isoperimetric inequality**

$$\mathrm{Per}(A) \ge \mathrm{Per}(A^*)\,, \tag{1.12}$$

which means that balls minimize surface area among all bodies of given volume.

Let f be a nonnegative smooth function with compact support in $\mathbb{R}^n$, and let u and v be the unique solutions of

$$-\Delta u = f\,, \quad -\Delta v = f^* \tag{1.13}$$

that decay at infinity. **Talenti's inequality** says that

$$u^*(x) \le v(x) \tag{1.14}$$

for all $x \in \mathbb{R}^n$. More general versions of the inequality compare an elliptic problem on a general domain Ω with a radial problem on Ω^*. They are powerful tool for establishing existence and regularity for weak solutions of elliptic equations under minimal regularity assumptions on the coefficients, the data, and the right hand side.

Talenti's inequality is related with a special case of Riesz' inequality as follows. The function u is the electrostatic potential associated with f, given by

$$u(x) = C(n) \int_{\mathbb{R}^n} |x-y|^{-(n-2)} u(y)\, dy\,,$$

and correspondingly for v. Riesz' inequality guarantees that v dominates u in the sense that

$$\int_B u^*(x)\, dx \le \int_B v(x)\, dx \tag{1.15}$$

for every ball B. Talenti's pointwise inequality (1.14) considerably strengthens Eq. (1.15).

Exercise 1.9 *Rewrite the isoperimetric inequality from Eq. (1.12) in the form*

$$\mathrm{Per}(A) \ \ge C(n)\big(\mathrm{Vol}(A)\big)^{\alpha(n)}\,.$$

What is the correct exponent $\alpha(n)$? Determine the sharp constant $C(n)$ in terms of ω_n, the volume of the unit ball.

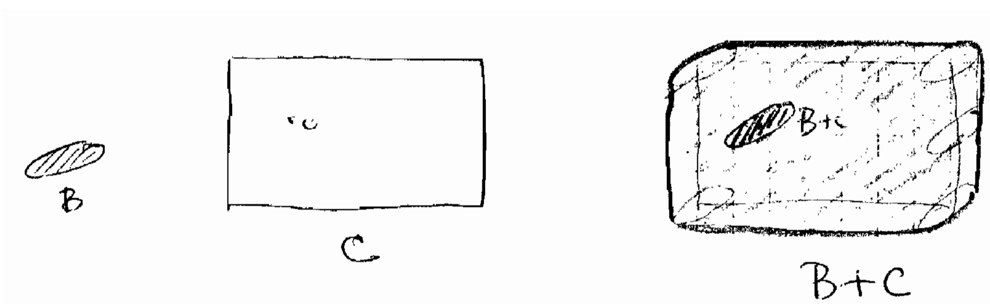


Figure 1.4: The Minkowski sum of two sets

Exercise 1.10 *Define the Minkowski sum of two sets $B, C \subset \mathbb{R}^n$ by $B + C = \{b + c \mid b \in B, c \in C\}$. The* **Brunn-Minkowski inequality** *says that*

$$\mathrm{Vol}(B+C)^{1/n} \ge \mathrm{Vol}(B)^{1/n} + \mathrm{Vol}(C)^{1/n} \tag{1.16}$$

for every pair of nonnegative measurable sets of finite volume. Reformulate this inequality as a geometric relation between the sets B^*, C^*, and $(B+C)^*$.

Exercise 1.11 *Let f be a smooth nonnegative function with compact support in $\mathbb{R}^n$. The function*

$$u(t,x) = \frac{1}{4\pi t} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} f(y)\, dy$$

satisfies the heat equation $\partial_t u = \Delta u$ with initial values $u(0,x) = f(x)$. Combine this with Riesz' inequality to prove the $p = 2$ case of the Pólya-Szegő inequality.

Hint: Differentiate $I(t) = \int u(t,x) f(x)\, dx$ at $t = 0$.

Exercise 1.12 *Let f be a smooth, compactly supported function on $\mathbb{R}^n$, and let u and v be the solutions of Laplace's equations in Eq. (1.13). Show that*

$$||\nabla u||_2 \leq ||\nabla v||_2 \,.$$

Discuss this inequality in light of the Pólya-Szegő inequality.

1.4 Some applications

Theorem 1.13 (The Faber-Krahn inequality) *Let Ω be an open set of finite volume in $\mathbb{R}^n$. Let $\lambda_1(\Omega)$ be the principal eigenvalue of the Dirichlet Laplacian on Ω, i.e., the smallest value of λ for which the problem*

$$\begin{cases} \Delta u = \lambda u \,, & \text{in } \Omega \\ u = 0 \,, & \text{on } \partial\Omega \end{cases}$$

has a non-trivial solution. Then

$$\lambda_1(\Omega) \geq \lambda_1(\Omega^*) \,. \tag{1.17}$$

PROOF. The Rayleigh-Ritz principle characterizes the principal eigenvalue as a minimum,

$$\lambda_1(\Omega) = \inf_{||\phi||_{L^2(\Omega)}=1} \int_\Omega |\nabla\phi|^2 \,.$$

Let ϕ_1 be the normalized minimizing eigenfunction. Since replacing ϕ_1 with $|\phi_1|$ does not change the objective function, we may take ϕ_1 to be nonnegative. It follows from the Pólya-Szegő inequality that

$$\lambda_1(\Omega) = \int_\Omega |\nabla\phi_1|^2 \geq \int_{\Omega^*} |\nabla\phi_1^*|^2 \,.$$

Since $||\phi_1^*||_{L^2(\Omega^*)} = ||\phi_1||_{L^2(\Omega)} = 1$, the Rayleigh-Ritz principle for Ω^* implies that

$$\int_{\Omega^*} |\nabla\phi_1^*|^2 \geq \inf_{||\phi||_{L^2(\Omega)}=1} \int_\Omega |\nabla\phi|^2 = \lambda_1(\Omega^*) \,,$$

and Eq. (1.17) follows. $\square$

The **Sobolev inequality** says that for each $1 \le p < n$, $p^* = \frac{np}{n-p}$ there exists a constant C such that for every function f in $W^{1,p}(\mathbb{R}^n)$,

$$||\nabla f||_p \ge C||f||_{p^*} \,.$$

The sharp constant is given by

$$C_{Sobolev}(n,p) = \inf_{||f||_{p^*}=1} ||\nabla f||_p \,.$$

If we replace f with $|f|$ and then with $|f|^*$ and apply the Pólya-Szegő inequality, we see that it suffices to minimize over symmetric decreasing functions. This transforms the maximization problem to a single-variable problem, which makes it easier to prove that a maximizer exists, and reduces the Euler-Lagrange equation from a partial to an ordinary differential equation. However, this ODE is not easy to analyze, because powers of the radial variable r appear in the coefficients of the differential operators.

Similarly, the optimal constants in the **Hardy-Littlewood-Sobolev inequality**, given by

$$C_{HLS}(p,q,n) = \sup_{||f||_p=||g||_q=1} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{f(x)g(y)}{|x-y|^\lambda}\,dxdy\,, \quad \left(\frac{1}{p}+\frac{1}{q}+\frac{\lambda}{n}=2\right)$$

and in **Young's inequality**, given by

$$C_{Young}(p,q) = \sup_{||f||_p=||g||_q=||h||_r=1} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(g)g(y)h(x-y)\,dxdy\,, \quad \left(\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=2\right)$$

are determined by maximizing over symmetric decreasing functions. Reducing to the radial problem is again a mixed blessing, because the convolution then takes a complicated form.

Exercise 1.14 *The quantum mechanical ground state energy of the hydrogen atom is given by*

$$\inf_{||\psi||_2=1}\left\{\frac{1}{2}\int_{\mathbb{R}^3}|\nabla\psi(x)|^2\,dx - \int_{\mathbb{R}^3}\frac{Const.}{|x|}|\psi(x)|^2\,dx\right\}\,.$$

Here, ψ is a complex-valued function on $\mathbb{R}^3$ that represents the wave function. The minimizing wave function — which is known to be unique up to multiplication by a scalar — is called the **ground state**. *Argue that ψ is symmetric decreasing.*

Exercise 1.15 *Let A be a bounded open set in $\mathbb{R}^n$ with smooth boundary. Its perimeter is just the surface area that we know from multivariable calculus. Sketch three proofs of the isoperimetric inequality, along the following lines: (a) Assume the Pólya-Szegő inequality (1.11). Let ϕ_δ be a* *nonnegative smooth function that increases from* 0 *to* 1 *across a strip of some small width* δ*, and approximate the perimeter by*

$$\mathrm{Per}(A) \approx \int |\nabla \phi_\delta|\, dx\,;$$

(b) assume the Brunn-Minkowski inequality (1.16), let B_δ *be the centered ball of radius* δ*, and approximate the perimeter by*

$$\mathrm{Per}(A) \approx \frac{1}{\delta}\mathrm{Vol}\big((A+B_\delta)\setminus A\big)\,;$$

(c) assume Riesz' inequality (1.9) or Eq. (1.10), and approximate the perimeter by

$$\mathrm{Per}(A) \approx \frac{C(n)}{\delta^{n+1}}\int_A \mathcal{X}_{A^c} * \mathcal{X}_{B_\delta}(x)\, dx\,,$$

where $C(n) = \left(\int_0^\infty \mathcal{X}_{\{x_n<0\}} * \mathcal{X}_B(s\,\mathbf{e}_n)\, ds\right)^{-1} = \frac{n+1}{\omega_{n-1}}$.

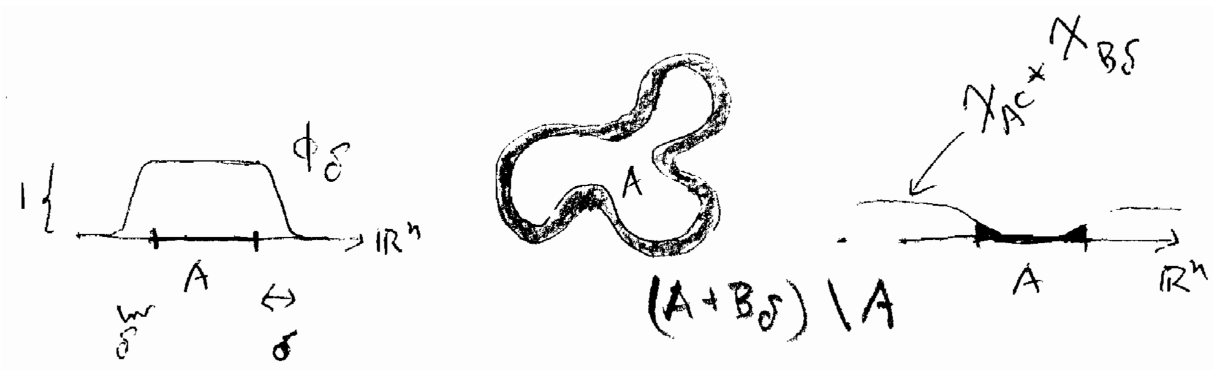


Figure 1.5: Approximations of the perimeter.

1.5 Selected textbooks and monographs

[HLP] G. H. Hardy, J. E. Littlewood, and G. Pólya, “Inequalities”. Cambridge University Press (1952, translated from the 1934 German original).

Chapter 10 contains the first systematic treatment of rearrangement inequalities. The main results discussed there are the Hardy-Littlewood inequality (1.6) and Riesz' inequality (1.9). These inequalities are reduced to sums by discretization and the layer-cake decomposition.

[PS] G. Pólya and G. Szegő, “Isoperimetric inequalities in Mathematical Physics”. Annals of Mathematics Studies, Vol. 27, Princeton University Press (1951).

A natural complement to [HLP] dedicated to the physical and geometrical problems that motivate the study of rearrangements.

[CR] K. M. Chong and N. M. Rice, “Equimeasurable rearrangements of functions”. Queen’s Papers in Pure and Applied Mathematics, Vol. 28, Queen’s University (1971).

A much-cited little book, now sadly out of print.

[B] C. Bandle, Isoperimetric inequalities and applications. Monographs and Studies in Mathematics, Vol. 7 (1980).

Dedicated to geometric boundary value problems and eigenvalue problems. The Pólya-Szegő inequality and Talenti’s comparison principle are derived from the isoperimetric inequality with the co-area formula.

[K] B. Kawohl, “Rearrangements and convexity of level sets in PDE”. Springer Lecture Notes in Mathematics, Vol. 1150 (1985)

A comprehensive account of rearrangement methods in PDE. Many different rearrangements are discussed in addition to the symmetric decreasing rearrangement.

[LL] E. H. Lieb and M. Loss, “Analysis”. AMS Graduate Studies in Mathematics, Vol. 14 (1987, second edition 2001).

Chapter 3 covers much of the same ground as [HLP, Chapter 10], but from a functional analytic point of view. Chapter 4 contains the proofs of Riesz’ inequality and the Hardy-Littlewood-Sobolev inequality to be discussed below.

[AFP] L. Ambrosio, N. Fusco, and D. Pallara, “Functions of bounded variation and free discontinuity problems”. Clarendon Press, Oxford (2000).

Chapters 2 and 3 provide background on sets of finite perimeter, the co-area formula, and the isoperimetric inequality.

[Ke] S. Kesavan, “Symmetrization and applications.” World Scientific Series in Analysis, Vol. 3 (2008).

An exposition at the advanced undergraduate level, in the spirit of [B]. Chapter 5 contains recent applications to geometric eigenvalue problems.

2 Approximation by simpler rearrangements

2.1 Steiner's argument

In two dimensions, the isoperimetric inequality says that the disk minimizes perimeter among all shapes of a given area. Though this was known in antiquity, it took a long time to assemble a proof.

Around 1830, Steiner proposed to prove the inequality for convex sets in two dimensions as follows. Consider a set $C \subset \mathbb{R}^2$ that minimizes perimeter among all convex sets of the same area,

$$\mathrm{Per}(C) = \inf\big\{\mathrm{Per}(A) \mid A \subset \mathbb{R}^2 \text{ convex}\,, Vol(A) = \mathrm{Vol}(C)\big\}\,.$$

Parametrize the upper boundary of C by a function h_+, and the lower boundary by h_-, so that

$$C = \big\{(x,y) \mid a \le x \le b, h_-(x) \le y \le h_+(x)\big\}\,.$$

The area of C can be computed by

$$\mathrm{Vol}(C) = \int_a^b h_+(x) - h_-(x)\,dx\,,$$

and its perimeter by

$$\mathrm{Per}(C) = \big(h_+(a) - h_-(a)\big) + \big(h_+(b) - h_-(b)\big) + \int_a^b \sqrt{1+(h_+')^2} + \sqrt{1+(h_-')^2}\,dx\,.$$

Steiner argues that C must be symmetric under reflection at some horizontal line $y = y_0$. To see this, he compares it with the set

$$\mathcal{S}C = \Big\{(x,y) \mid a \le x \le b, |y| \le \frac{1}{2}\big(h_+(x) - h_-(x)\big)\Big\}\,,$$

which is symmetric about $y = 0$, convex, and has the same area as C.

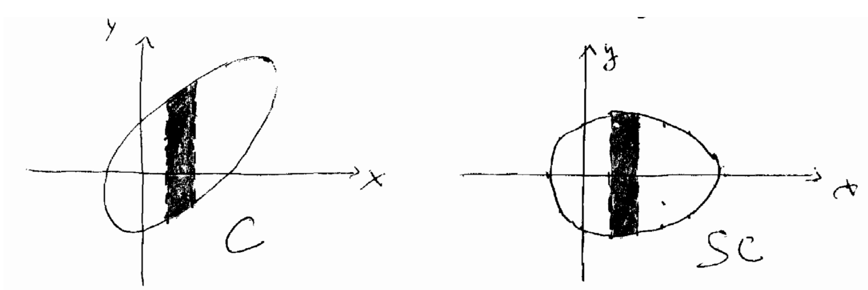


Figure 2.1: The Steiner symmetrization of a convex set C.

Moreover, by the strict convexity of the function $t \to \sqrt{1+t^2}$, its perimeter satisfies

$$\mathrm{Per}(C) > \mathrm{Per}(\mathcal{S}C)$$

unless $h'_+(x) = -h'_-(x)$. It follows that the optimal set C must be symmetric under reflection at some horizontal line line $y = y_0$. By the same argument, C is symmetric about some vertical line $x = x_0$, and in fact about every line through (x_0, y_0). Thus C is a disk, and we conclude that for any convex set $A \subset \mathbb{R}^2$

$$\mathrm{Per}(A) \geq \mathrm{Per}(A^*) = 2\sqrt{\pi}\big(\mathrm{Vol}(A)\big)^{\frac{1}{2}}\,, \tag{2.1}$$

with equality if and only if A is a disk. $\square$

Exercise 2.1 *Find and discuss all places in Steiner's argument where the convexity of C is used.*

Exercise 2.2 *Let $A \subset \mathbb{R}^2$ be an open set whose boundary is given by a curve of finite length. Prove that the convex hull of A has larger volume, but smaller perimeter than A. Use scaling to extend Eq. (2.1) from convex to general planar sets. What happens to this argument in higher dimensions?*

Question 2.3 *What does Steiner's argument prove, what are its strengths and weaknesses?*

The essence of Steiner's idea is to reduce the two-dimensional isoperimetric inequality to a one-dimensional problem by using a rearrangement along one-dimensional cross sections. His construction can be extended to n dimensions in several ways. If A is a measurable set in $\mathbb{R}^n$, its *Steiner symmetrization* $\mathcal{S}A$ is defined by replacing its intersection with each line $x_1, \dots x_{n-1} = Const.$ by a symmetric interval. Explicitly, if we write $x = (\hat{x}, x_n)$ with $\hat{x} \in \mathbb{R}^{n-1}$, and let $A_{\hat{x}} = \{t \mid \big(\hat{x}, t\big) \in A\}$, then

$$\mathcal{S}A = \big\{(\hat{x}, t) \in \mathbb{R}^n \times \mathbb{R} \mid t \in (A_{\hat{x}})^*\big\}\,.$$

Similarly, we define the **Schwarz symmetrization** $\mathcal{T}A$ of A by replacing its intersection with each hyperplane $x_n = Const.$ by a centered ball. If we write $A^t = \{\hat{x} \mid \big(\hat{x}, t\big) \in A\}$ in $\mathbb{R}^{n-1}$, then

$$\mathcal{T}A = \big\{(\hat{x}, t) \in \mathbb{R}^n \times \mathbb{R} \mid \hat{x} \in (A^t)^*\big\}\,.$$

Steiner and Schwarz symmetrization preserve n-dimensional volume by Fubini's theorem. They also define equimeasurable rearrangements of functions via the layer-cake representation.

2.2 Polarization

Let us carry Steiner's idea one step further and define an even simpler rearrangement that can reduce many inequalities to zero-dimensional, i.e., combinatorial or algebraic problems. Let X_0 be a hyperplane in $\mathbb{R}^n$ that does not contain the origin. Denote by X_+ the resulting open half-space that contains the origin, and by X_- the complementary half-space, and let σ be the reflection that exchanges the two half-spaces. The **two-point rearrangement**, or **polarization** of a function f at X_0 is defined by

$$f^\sigma(x) = \begin{cases} \max\{f(x), f(\sigma x)\} & x \in X_+\,, \\ \min\{f(x), f(\sigma x)\} & x \in X_-\,, \\ f(x) & x \in X_0\,. \end{cases}$$

Exercise 2.4 *If f is a nonnegative measurable function that vanishes at infinity, show that f^σ is equimeasurable with f. You may find it useful to define the polarization of a set A by*

$$\begin{cases} A^\sigma \cap X_+ &= (A \cup \sigma A) \cap X_+\,, \\ A^\sigma \cap X_- &= (A \cap \sigma A) \cap X_-\,, \\ A^\sigma \cap X_0 &= A \cap X_0\,. \end{cases}$$

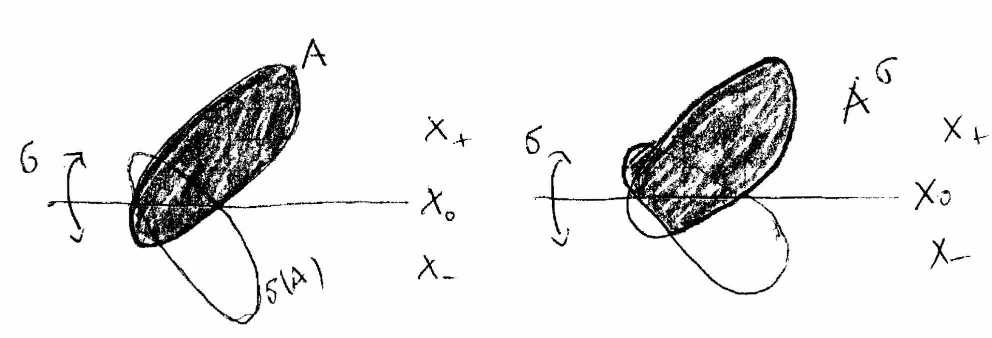


Figure 2.2: Polarization of a set at a hyperplane.

Lemma 2.5 (Polarization improves the modulus of continuity) *Suppose that f is a uniformly continuous function on $\mathbb{R}^n$, i.e., for every $\varepsilon > 0$ there exists a $\delta(\varepsilon) > 0$ such that for all $x, y \in \mathbb{R}^n$,*

$$|x - y| < \delta(\varepsilon) \implies |f(x) - f(y)| < \varepsilon\,.$$

Then, for every reflection σ, the polarization f^σ is uniformly continuous with the same modulus of continuity,

$$|x - y| < \delta(\varepsilon) \implies |f^\sigma(x) - f^\sigma(y)| < \varepsilon\,.$$

PROOF. Let $\varepsilon > 0$ be given, and consider two points x, y with $|x - y| < \delta(\varepsilon)$. If $x, y \in X_+$, then

$$\begin{aligned} |f^\sigma(x) - f^\sigma(y)| &= \big|\max\{f(x), f(\sigma x)\} - \max\{f(y), f(\sigma y)\}\big| \\ &\le \max\big\{|f(x) - f(y)|, |f(\sigma x) - f(\sigma y)|\big\} \\ &< \varepsilon\,, \end{aligned}$$

since $|\sigma x - \sigma y| = |x - y| < \delta(\varepsilon)$. If $x, y \in X_-$, we replace $\max$ by $\min$ in the above argument. If the two points lie in different half-spaces, then

$$\begin{aligned} |f^\sigma(x) - f^\sigma(y)| &\le \max\big\{|f(x) - f(y)|, |f(\sigma x) - f(y)|, |f(x) - f(\sigma y)|, |f(\sigma x) - f(\sigma y)|\big\} \\ &< \varepsilon\,, \end{aligned}$$

since $|\sigma x - y| = |x - \sigma y| \le |x - y| < \varepsilon$. $\square$

Note that the above proof compared only the values of f at four points. A similar reduction occurs for the simple case of Riesz' inequality, which we show next. In the proof, we will use the elementary fact that

$$\big\{(UV{+}uv)W + (Uv{+}uV)w\big\} - \big\{(Uv{+}uV)W + (UV{+}uv)w\big\} = (U{-}u)(V{-}v)(W{-}w)\,. \quad (2.2)$$

Lemma 2.6 (Simple case of Riesz' inequality for polarization) *Let H be a nonincreasing function on the positive half-line with $H(t) \to 0$ as $t \to \infty$. Then, for every pair of nonnegative measurable functions f, g on $\mathbb{R}^n$,*

$$\int\!\!\int f(x)g(y)H(|x-y|)\,dxdy \le \int\!\!\int f^\sigma(x)g^\sigma(y)H(|x-y|)\,dxdy\,. \qquad (2.3)$$

If H is strictly decreasing, then equality holds (with a finite and nonzero value of the functional) if and only if either $f = f^\sigma$ and $g = g^\sigma$, or else $f = f^\sigma \circ \sigma$, $g = g^\sigma \circ \sigma$ almost everywhere.

PROOF. Rewrite the integral on the left hand side as

$$\begin{aligned} I(f,g) \;=\; \int_{X_+}\int_{X_+} & \big\{f(x)g(y) + f(\sigma x)g(\sigma y)\big\}H(|x-y|) + \\ & + \big\{f(x)g(\sigma y) + f(\sigma x)g(y)\big\}H(|x-\sigma y|)\,dxdy\,. \end{aligned}$$

(We have used that $|x-y| = |\sigma x - \sigma y|$ and $|x - \sigma y| = |\sigma x - y|$.)

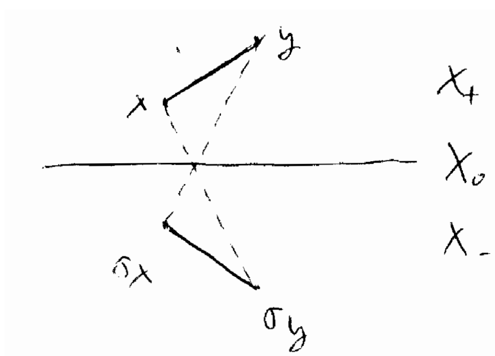


Figure 2.3: $|x-y| = |\sigma x - \sigma y| < |x - \sigma y| = |\sigma x - y|\,.$

We claim that the integrand increases pointwise under polarization. To see this, fix x, y and consider four cases. If $f(x) \ge f(\sigma x)$ and $g(y) \ge g(\sigma y)$, then polarization has no effect on the values of the functions at the points $x, y, \sigma x, \sigma y$. If $f(x) < f(\sigma x)$ and $g(y) < g(\sigma y)$, then polarization switches the values of f, g at x, y with their values at $\sigma x, \sigma y$, and the integrand is again unchanged. If $f(x) \ge f(\sigma x)$ but $g(y) < g(\sigma y)$, then the values are switched for g but not for f, and the integrand increases by the difference

$$\big(f(x) - f(\sigma x)\big)\big(g(\sigma y) - g(y)\big)\big(H(|x-y|) - H(|x-\sigma y|)\big) \ge 0\,,$$

and correspondingly for the remaining case. Set $K(x,y) = H(|x-y|) - H(|x-\sigma y|)$. Since $K(x,y) \geq 0$ for $x, y \in X_+$, we conclude with Eq. (2.2) that

$$\begin{aligned} I(f^\sigma, g^\sigma) - I(f,g) &= \int_{X_+}\int_{X_+} \big[\big(f(x)-f(\sigma x)\big)\big(g(y)-g(\sigma y)\big)\big]_- K(x,y)\,dxdy \\ &\geq 0\,. \end{aligned}$$

If H is strictly decreasing, then $K(x,y) > 0$ for $x, y \in X_+$, proving the equality statement. $\square$

Exercise 2.7 (Hardy-Littlewood inequality for polarization) *Prove that for every pair of non-negative measurable functions f, g*

$$\int fg \leq \int f^\sigma g^\sigma\,,$$

with equality if and only if $\big(f(x)-f(\sigma x)\big)\big(g(x)-g(\sigma x)\big) \geq 0$ almost everywhere on $\mathbb{R}^n$. What is the underlying algebraic identity?

Exercise 2.8 (Polarization decreases L^p-distances) *Let Φ be a smooth function defined on the positive quadrant of $\mathbb{R}^2$ that vanishes if $s = 0$ or $t = 0$. Assume that $\partial_s\partial_t\Phi(s,t) \geq 0$ for all $s, t > 0$. If f and g be nonnegative measurable functions that vanish at infinity, show that*

$$\int \Phi\big(f(x), g(x)\big)\,dx \leq \int \Phi\big(f^\sigma(x), g^\sigma(x)\big)\,dx\,.$$

Conclude that polarization deceases L^p-distances

$$||f-g||_p \geq ||f^\sigma - g^\sigma||_p\,.$$

Exercise 2.9 *Define reflections and polarizations on the unit sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$. Which of the polarization inequalities proved above remain valid?*

2.3 From polarization to symmetric decreasing rearrangement

We would like to use the polarization inequalities from the preceding subsection to prove the corresponding inequalities for the symmetric decreasing rearrangement. Following is the main result from this section.

Theorem 2.10 (Simple case of Riesz' inequality) *Let H be a nonincreasing function on the positive real line with $H(t) \to 0$ as $t \to \infty$. Then, for any pair f, g of nonnegative measurable functions on $\mathbb{R}^n$ that vanish at infinity,*

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(y)H(|x-y|)\,dxdy \le \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f^*(x)g^*(y)H(|x-y|)\,dxdy\,.$$

If H is strictly decreasing, then equality (with a finite and nonzero value of the integral) occurs only if there exists a translation τ such that $f = f^* \circ \tau$ and $g = g^* \circ \tau$ almost everywhere.

We want to make Steiner's argument from Section 2.1 work for polarization. The strategy is to use Lemma 2.5 to show that a maximizer of I **exists** in a suitable class of functions, and then **identify** that f^*, g^* are among the maximizers with the following lemma:

Lemma 2.11 (Characterization of symmetric decreasing functions) *Let f be a nonnegative function on $\mathbb{R}^n$ that vanishes at infinity. Then*

$$f = f^* \iff f = f^\sigma \text{ for all } \sigma\,,$$

and

$$f = f^* \circ \tau \text{ for some translation } \tau \iff \text{for all } \sigma, \text{ either } f = f^\sigma \text{ or } f = f^\sigma \circ \sigma\,.$$

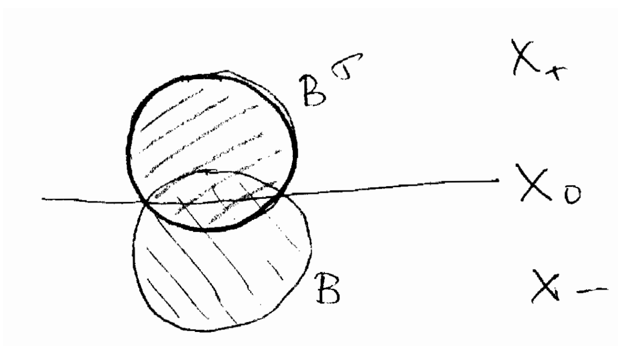


Figure 2.4: Polarization rearranges balls into balls.

PROOF. In both cases, the $\Rightarrow$ implications are straightforward. To prove the reverse implication in the first case, suppose that f is not radially decreasing, and fix two points x_1, x_2 with $|x_1| < |x_2|$ but $f(x_1) < f(x_2)$. Let σ be the reflection that maps x_1 to x_2, and let X be the invariant hyperplane. Then x_1 lies in the half space X_+ that contains the origin, and x_2 lies in the complementary half space X_-. By definition, $f^\sigma(x_1) = f(x_2)$ and $f^\sigma(x_2) = f(x_1)$, showing that $f^\sigma \ne f$.

For the reverse implication in the second case, we assume (by approximation with a **fat layer**

$$f_\varepsilon = \min\{\varepsilon^{-1}, [f-\varepsilon]_+\}\,, \tag{2.4}$$

that f is bounded and integrable. After a suitable translation, its center of mass lies at the origin. Since the center of mass for g^σ lies in X_+, we conclude that $g^\sigma = g$ for all σ. The proof is completed by using the first case. $\square$

Exercise 2.12 *Extend Lemma 2.11 to the unit sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$.*

Lemma 2.11 can be combined with Lemma 2.5 to approximate the symmetric decreasing rearrangement f^* by repeated polarizations of f:

Proposition 2.13 (Approximation of f^* by polarizations) *Assume that f is a nonnegative continuous function with compact support in $\mathbb{R}^n$, and let*

$$\mathrm{Pol}_f = \left\{ f^{\sigma_1,\dots,\sigma_k} \mid k \geq 0, \sigma_1, \dots, \sigma_k \textit{ reflections} \right\}$$

be the set of all functions that can be reached by applying a finite sequence of polarizations to f. There exists a sequence $\{g_k\}_{k\geq 1}$ in Pol_f such that

$$g_k \to f^* \textit{ uniformly}. \tag{2.5}$$

PROOF. Let H be a fixed, strictly decreasing, bounded function on $\mathbb{R}_+$ with $H(t) \to 0$ as $t \to \infty$, and define an auxiliary functional by

$$I(f) = \int f(x) H(|x|)\, dx\, .$$

Since Pol_f is relatively compact in $\mathcal{C}_c$ by Lemma 2.5 and the Arzelà-Ascoli theorem, the functional I assumes its maximum at some function g in the closure of Pol_f. We will show that $g = f^*$.

Let $\{g_k\}_{k\geq 1}$ be a sequence in Pol_f that converges uniformly to g. Every polarization g^σ lies also in the closure of Pol_f because g_k^σ converges uniformly to g^σ, and hence $I(g) \geq I(g^\sigma)$ by the maximality of g. On the other hand, $I(g) \leq I(g^\sigma)$ by Exercise 2.7. Thus $I(g) = I(g^\sigma)$. Exercise 2.7 further implies that $\big(g(x) - g(\sigma x)\big)\big(H(|x|) - H(|\sigma x|)\big) \geq 0$, which means that $g(x) = g^\sigma(x)$. Since x and σ were arbitrary, we can use Lemma 2.11 to see that g is radially decreasing. Since g is equimeasurable to f (being a uniform limit of such functions), we conclude that $g = f^*$, proving Eq. (2.5). $\square$

This is easily extended to simultaneous polarizations of m-tuples of functions $(f_1, \dots, f_m)$. We are finally ready to prove the main result of this section:

PROOF OF THEOREM 2.10. Denote the simple Riesz functional by

$$I(f, g) = \int\int f(x) g(y) H(|x-y|)\, dxdy\, .$$

Suppose for the moment that f and g are continuous and compactly supported, and that H is bounded. By Proposition 2.13, there exists a sequence $\{(f_k, g_k)\}_{k\geq 1}$ in the set $\mathrm{Pol}_{f,g}$ that converges uniformly to (f^*, g^*). By Lemma 2.6 and the continuity of I,

$$I(f,g) \leq I(f_k, g_k) \to I(f^*, g^*)\,,$$

proving the inequality for continuous functions of compact support. Since such functions are dense in L^p for $1 \leq p < \infty$, and I is continuous, the inequality holds also there. If f is just measurable, we approximate it with a fat layer as in Eq. (2.4) and correspondingly for g and H. The inequality follows in the the limit $\varepsilon \to 0$ from the monotone convergence theorem.

For the equality statement, assume that H is strictly decreasing, and suppose that $I(f,g) = I(f^*, g^*)$ is finite and non-zero. Then $I(f,g) = I(f^\sigma, g^\sigma)$ for every polarization σ. By Lemma 2.6, either $f^\sigma = f$ or $f^\sigma = f \circ \sigma$, and correspondingly for g. Since this holds for every σ, Lemma 2.11 implies that $f = f^* \circ \tau$ for some translation τ, and similarly for g. Furthermore, the two translations must agree. □

Exercise 2.14 *If f is a uniformly continuous nonnegative function on $\mathbb{R}^n$ that vanishes at infinity, prove that f^* is uniformly continuous with the same modulus of continuity.*

Exercise 2.15 *Let f be a nonnegative measurable function on $\mathbb{R}^n$ that vanishes at infinity, and let Pol_f be the set of functions that can be reached by finitely many polarizations, as in Proposition 2.13. Show that there exists a sequence $\{g_k\}_{k\geq 1}$ in Pol_f such that*

$$g_k \to f^* \text{ in measure}\,.$$

If, moreover, $f \in L^p$, the sequence can be chosen to converge converge strongly in L^p; if $f \in W^{1,p}$, then also

$$\nabla g_k \to \nabla f^* \text{ weakly in } L^p\,.$$

Hint: Consider a fat layer of f, as in Eq. (2.4), and approximate it with a smooth function of compact support. Then apply Proposition 2.13.

The simple case of Riesz' inequality in Eq. (2.3) can be extended to more general integrands $\Phi(t_1, \ldots, t_m)$, with a suitable integral kernel $h(x_1, \ldots, h_m)$. The crucial condition is that Φ should be **supermodular**, in the sense that all its second (mixed) distributional derivatives should be nonnegative. Here is the corresponding version of the Hardy-Littlewood inequality:

Exercise 2.16 (Hardy-Littlewood inequality for general integrands) *Let Φ be a smooth function on the positive cone of $\mathbb{R}^m$ that vanishes on the boundary of the cone. Assume that Φ is supermodular, i.e., its mixed derivatives satisfy $\partial_i \partial_j \Phi(t_1, \ldots, t_m) \geq 0$ for all $i \neq j$. Let $f_1, \ldots, f_m$ be nonnegative measurable functions on $\mathbb{R}^n$ that vanish at infinity.*

$$\int_{\mathbb{R}^n} \Phi\big(f_1(x), \ldots, f_m(x)\big)\, dx \leq \int \Phi\big(f_1^*(x), \ldots, f_m^*(x)\big)\; dx\,.$$

Hint: Use polarization. For $(s_1, \ldots, s_m)$ and $(t_1, \ldots, t_m)$ in the positive cone, show that

$$\begin{aligned}\Phi(s_1, \ldots, s_m) + \Phi(t_1, \ldots, t_m) \;\; &\leq \;\; \Phi\big(\max\{s_1, t_1\}, \ldots, \max\{s_m, t_m\}\big) \\ &\qquad + \Phi\big(\min\{s_1, t_1\}, \ldots, \min\{s_m, t_m\}\big)\,.\end{aligned}$$

2.4 Steiner's argument, revisited

Steiner's argument has two weaknesses: It does not prove that a perimeter-minimizing set exists, and it it relies on a classical definition of perimeter as surface measure that is not appropriate for general measurable sets. Both have been addressed by more resent results in geometric measure theory.

By definition, a measurable set $A \subset \mathbb{R}^n$ has **finite perimeter**, if there exists a vector-valued Radon measure ν of finite total mass such that the integration-by-parts formula

$$\int_A \mathrm{div} F\, dx = -\int F(x) \cdot d\nu(x)$$

holds for every smooth compactly supported vector field F on $\mathbb{R}^n$. It turns out that for a set of finite perimeter, ν is concentrated on the **essential boundary** of A, where both A and A^c have positive density in the sense of Lebesgue. In fact, if σ denotes the $n-1$dimensional Hausdorff measure on the essential boundary of A, then unit outward normal $N(x)$ to A can be defined σ-almost everywhere, and $d\nu(x) = -N(x)d\sigma$. The upshot is that the measure ν, the essential perimeter, and the outward unit normal of A do not depend on sets of measure zero.

In this framework, the existence of a perimeter-minimizing set is guaranteed by **De Giorgi's compactness theorem**: If $\{A_k\}_{k\geq 1}$ is a sequence of measurable sets lying in a common large ball $B \subset \mathbb{R}^n$, and if their perimeters are finite and uniformly bounded, then there exists a subsequence $\{A_{k_j}\}_{i\geq 1}$ that converges with respect to symmetric difference to some limiting set A,

$$\lim_{j\to\infty} \mathrm{Vol}\big(A_{k_j} \setminus A\big) + \mathrm{Vol}\big(A \setminus A_{k_j}\big) = 0\,.$$

A useful fact is that the volume functional is continuous and the perimeter functional is lower semicontinuous,

$$\mathrm{Vol}(A) = \lim_{j\to\infty} \mathrm{Vol}\big(A_{k_j}\big)\,, \quad \mathrm{Per}(A) \leq \liminf_{j\to\infty} \mathrm{Per}\big(A_{k_j}\big)\,.$$

Theorem 2.17 (Isoperimetric inequality) *If $A \subset \mathbb{R}^n$ has finite perimeter, then*

$$\mathrm{Per}(A) \geq \mathrm{Per}(A^*)\,,$$

with equality if and only if A differs from a ball by a set of measure zero.

PROOF. *Compactness.* Let $B \subset \mathbb{R}^n$ be a large closed ball, and let $\{A_k\}_{k\geq 1}$ be a minimizing sequence for the perimeter for a given value of the volume. By De Giorgi's compactness theorem, there exists a subsequence, again denoted by A_k, that converges in symmetric difference to a limiting set A of the same volume. Since the perimeter functional is lower semicontinuous, it follows that A minimizes perimeter for the given volume.

Identification of the minimizer. The next step is to show that the equality $\mathrm{Per}(A) = \mathrm{Per}(\mathcal{S}A)$ implies that almost all vertical cross sections of A are intervals. Since this holds for every direction, A must be convex. Now Steiner's argument implies that A is a ball. $\square$

Exercise 2.18 (Approximation of Steiner symmetrization by polarizations) *Under the assumptions of Proposition 2.13, show that there exists sequences in* Pol_f *that converge uniformly to the Steiner symmetrization* $\mathcal{S}f$*. Conclude that Steiner symmetrization improves the modulus of continuity.*

Exercise 2.19 (Approximation of f^* **by Steiner symmetrizations)** *Assume that* f *is a nonnegative continuous function with compact support, and let* $\mathrm{Steiner}_f$ *be the set of all functions that can be reached by applying a finite sequence of Steiner symmetrizations in different directions. Argue as in Proposition 2.13 that there exists a sequence in* $\mathrm{Steiner}_f$ *that converges uniformly to* f^**.*

2.5 References

[S] J. Steiner , "Einfacher Beweis der isoperimetrischen Hauptsätze". Crelle J. Reine Angew. Math., Vol. 18 (1838); *reprinted in* Gesammelte Werke 2 (Berlin 1882), p. 77-91.

Steiner's "simple proof" of the isoperimetric inequality.

[A] L. V. Ahlfors, "Conformal invariants: Topics in geometric function theory". McGraw-Hill Series in Higher Mathematics (1973)

Riesz' simple inequality (1.10) is proved in one dimension by polarization. Ahlfors credits Littlewood with the idea, but without precise reference.

[BT] A. Baernstein, II and B. A. Taylor, "Spherical rearrangements, subharmonic functions, and $*$-functions in n-space. Duke Math. J., Vol. 66 (1976), p. 245-268.

Polarization is used to prove Riesz' simple inequality on the sphere. Section 2 contains the first proof that the symmetric decreasing rearrangement can be approximated by simpler rearrangements.

[L77] E. H. Lieb, "Existence and uniqueness of the minimizing solution of Chocquard's nonlinear equation". Studies in Applied Mathematics,Vol. 57 (1977), p. 93-105.

Contains a proof of Riesz' simple inequality in $\mathbb{R}^n$ that does not use polarization. If the kernel is strictly decreasing, then optimizers are unique up to translation.

[B] William Beckner, "Sobolev inequalities, the Poison semigroup, and analysis on the sphere $\mathbb{S}^n$". Proc. Nat. Acad. Sci., Vol. 89 (1992), p. 4816-4819.

The equality cases in Eq. (1.10) on $\mathbb{S}^n$ are identified through polarization.

[T] G. Talenti, "The standard isoperimetric theorem". Handbook of Convex Analysis, Vol. A (1993), p. 431-455..

An exposition of the proof of the isoperimetric inequality for sets of finite perimeter.

[BrSo] F. Brock and A. Yu. Solynin, "An approach to symmetrization via polarization". Trans. Amer. Math. Soc. (2. series), Vol. 138 (2000), p. 1759-1796.

A comprehensive study of polarization that builds on Section 2 of [BT].

[CCF] M. Chlebík, A. Cianchi, and N. Fusco, "The perimeter inequality under Steiner symmetrization: Cases of equality". Ann. Math. (2), Vol. 162 (2005), p. 525-555.

[BuH] A. Burchard and H. Hajaiej, "Rearrangement inequalities for functionals with monotone integrands". J. Funct. Anal., Vol. 223 (2006), p. 561-582.

Contains an overview of polarization rtesults from the literature.

3 Competing symmetries

In the last lecture, we discussed approximating the symmetric decreasing rearrangement by sequences of simpler rearrangements. However, we did not construct the approximating sequence explicitly, but relied instead on compactness properties to show that an approximating sequence exists (see Proposition 2.13).

Today we will discuss how symmetrization can be used to construct strongly convergent optimizing sequences in geometric variational problems. Suppose we want to maximize a functional I over a subset X of some function space. The basic idea is to take advantage of two non-commuting operations (the "competing symmetries") that both improve the functional. The first operation, $\mathcal{S}$, is a symmetrization that satisfies $\mathcal{S}^2 = \mathcal{S}$ and strictly improves I,

$$I(f) < I(\mathcal{S}f)\,,$$

unless $f = \mathcal{S}f$. The second operation, $\mathcal{R}$ is a transformation that destroys the symmetry introduced by $\mathcal{S}$ and satisfies

$$I(f) \leq I(\mathcal{R}f)\,.$$

Typically, $\mathcal{R}$ is an isometry of X defined by a Euclidean or conformal transformation that leaves the functional invariant. The goal is to show that the sequence

$$f_k = (\mathcal{S}\mathcal{R})^k f$$

converges to a maximizer.

Below, we will illustrate the technique for two examples, the conformally invariant case of the Hardy-Littlewood-Sobolev inequality for which it as first developed, and the approximation of symmetric decreasing rearrangement by Steiner symmetrizations.

3.1 Conformal invariance and the Hardy-Littlewood-Sobolev inequality

We begin with a brief review of the inequality.

Theorem 3.1 (Hardy-Littlewood 1928, Sobolev 1938.) *Let $p, q > 1$ and $0 < \lambda < n$ be such that $1/p + 1/q + \lambda/n = 2$. There exists a constant $C(n, \lambda, p)$ such that*

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(y)|x-y|^{-\lambda}\,dxdy \leq C(n,\lambda,p)||f||_p||g||_q \tag{3.1}$$

for all $f \in L^p(\mathbb{R}^n), g \in L^q(\mathbb{R}^n)$.

One may ask for the value of the best constant $C(n, \lambda, p)$, and, if it is achieved, for the family of optimizers. In 1983, Lieb showed that for every admissible choice of n, p, q there exist functions f and g with $||f||_p = ||g||_q = 1$ that maximize the right hand side. For $p = q$, he identified all maximizers, and computed the best constant $C(n,p)$. This maximal value is achieved precisely when $f = g$ are scaled and translated multiples of

$$h(x) = \left(\frac{1}{1+|x|^2}\right)^{n/p} .$$

In his proof that maximizers exist, Lieb showed how to modify an arbitrary maximizing sequence so that it converges in L^p. The difficulty in this step is related to the many symmetries of the Hardy-Littlewood-Sobolev functional, particularly the translation and scaling symmetries, which make it easy for a maximizing sequence to converge weakly to zero. On the other hand, he crucially used the symmetries of the equation to identify the optimizers in the case when $p = q$.

For $p = q$, the goal is to maximize the functional

$$I(f) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)f(y)|x-y|^{-\lambda}\, dxdy$$

subject to the constraint $||f||_p = 1$. Let $\mathcal{S}f = f^*$ denote the symmetric decreasing rearrangement of f. By the simple case of Riesz' inequality in Theorem 2.10,

$$I(f) \le I(\mathcal{S}f)\,,$$

with equality if and only if f is a translate of its symmetric decreasing rearrangement.

At the same time, I is conformally invariant, in the following sense. A map γ defined on an open set Ω in $\mathbb{R}^n$ is called **conformal**, if for every pair of smooth curves C_1, C_2 that intersect at some point $x_0 \in \Omega$, the images $\gamma(C_1)$ and $\gamma(C_2)$ intersect at the same angle as the original curves. Equivalently, there exists a scalar function h on Ω such that the derivative of γ satisfies

$$\forall x \in \Omega,\ \forall u, v \in \mathbb{R}^n:\ \langle D\gamma(x)u, D\gamma(x)v\rangle = h(x)^2\langle u, v\rangle\,.$$

The function $h(x)$, which defines the local expansion or contraction of g, is called the **conformal factor**. The Jacobian of γ is given by $|\det D\gamma(x)| = h(x)^n$.

If γ is a conformal transformation on $\mathbb{R}^n$ with conformal factor h, then

$$\mathcal{U}_\gamma f(x) = f(\gamma(x))h(x)^{n/p}$$

defines a linear isometry on $L^p(\mathbb{R}^n)$. Conformal invariance means that

$$I(f) = I(\mathcal{U}_\gamma f) \tag{3.2}$$

for all conformal transformations γ. This clearly holds for translations, rotations, and dilations For the inversion $\gamma(x) = \frac{x}{|x|^2}$, we compute

$$\left|\frac{x}{|x|^2} - \frac{y}{|y|^2}\right|^2 = \frac{|x-y|^2}{|x|^2|y|^2}\,.$$

Taking $y \to x$ implies that tangent vectors at x are scaled by a conformal factor $|x|^{-2}$. Volumes are scaled by $\det D\gamma(x) = |x|^{-2n}$, and it follows from the Change of Variables formula that invariance holds also here. Every conformal transformation on $\mathbb{R}^n \cup \{\infty\}$ can be written as a composition of translations, rotations, dilations and the inversion.

Exercise 3.2 *We have implicitly used that γ is conformal with conformal factor h, if and only if*

$$\forall x \in \Omega,\ \forall u \in \mathbb{R}^n:\ |D\gamma(x)\, u| = h(x)\, |u|\,.$$

Prove this!

The conformal invariance of the functional I allows to rewrite it as an integral over the unit sphere $\mathbb{S}^n \subset \mathbb{R}^{n+1}$, as follows. The map given by

$$\phi(x) = \frac{1}{|x|^2+1}\left(2x, |x|^2-1\right) \in \mathbb{R}^n \times \mathbb{R}$$

defines a one-to-one correspondence from $\mathbb{R}^n \cup \{\infty\}$ to $\mathbb{S}^n$. The inverse function ϕ^{-1} is called the **stereographic projection** of $\mathbb{S}^n$ to $\mathbb{R}^n \cup \{\infty\}$.

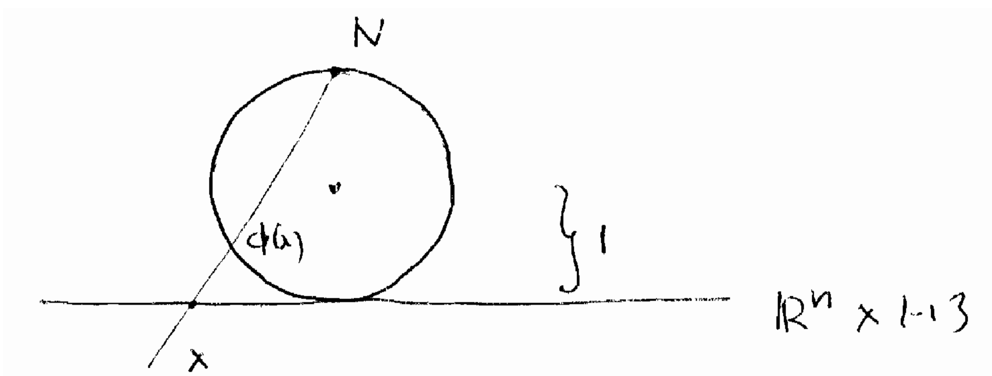


Figure 3.1: The stereographic projection.

We compute for $x, y \in \mathbb{R}^n$

$$|\phi(x)-\phi(y)|^2 = \frac{|x-y|^2}{(1+|x|^2)(1+|y|)^2}\,.$$

Here, the distance on the left hand side is the Euclidean norm on $\mathbb{R}^{n+1}$. This shows that ϕ is conformal, with conformal factor

$$h(x) = \frac{1}{1+|x|^2}\,.$$

Thus, we can define a linear isometry from $L^p(\mathbb{S}^n)$ to $L^p(\mathbb{R}^n)$ by setting

$$\mathcal{U}_\phi F(x) = F(\phi(x))h(x)^{n/p}\,.$$

If $f = \mathcal{U}_\phi F$, then

$$I(f) = \int_{\mathbb{S}^n}\int_{\mathbb{S}^n} F(\xi)F(\eta)|\xi-\eta|^{-\lambda}\, d\xi d\eta\,.$$

The right hand side defines a conformally invariant functional on $L^p(\mathbb{S}^n)$. Abusing notation, we will denote it by $I(F)$. The point is that certain conformal transformations that look complicated on $\mathbb{R}^n$ take a simple form on $\mathbb{S}^n$. Among these are the rotations on $\mathbb{S}^n$ that do not fix the poles. Together with the translations, rotations and dilations on $\mathbb{R}^n$, these generate the entire conformal group.

Exercise 3.3 *Which conformal transformation on $\mathbb{S}^n$ corresponds to the inversion on $\mathbb{R}^n$?*

Proposition 3.4 (Competing symmetries for conformally invariant functionals on $\mathbb{S}^n$) *For $F \in L^p(\mathbb{S}^n)$, define $\mathcal{S}F$ by applying the the symmetric decreasing rearrangement to the corresponding function on $\mathbb{R}^n$,*

$$\mathcal{S}F = \mathcal{U}_\phi{}^{-1}(\mathcal{U}_\phi F)^*\,,$$

and let $\mathcal{R}F$ be given by the rotation

$$\mathcal{R}F(\xi_1,\ldots,\xi_{n+1}) = F(\xi_1,\ldots,\xi_{n-1},-\xi_{n+1},\xi_n)\,.$$

If $F \geq 0$, then $(\mathcal{S}\mathcal{R})^k F$ converges strongly in L^p to the constant function $(n\omega_n)^{-1/p}||F||_{L^p}\cdot 1$.

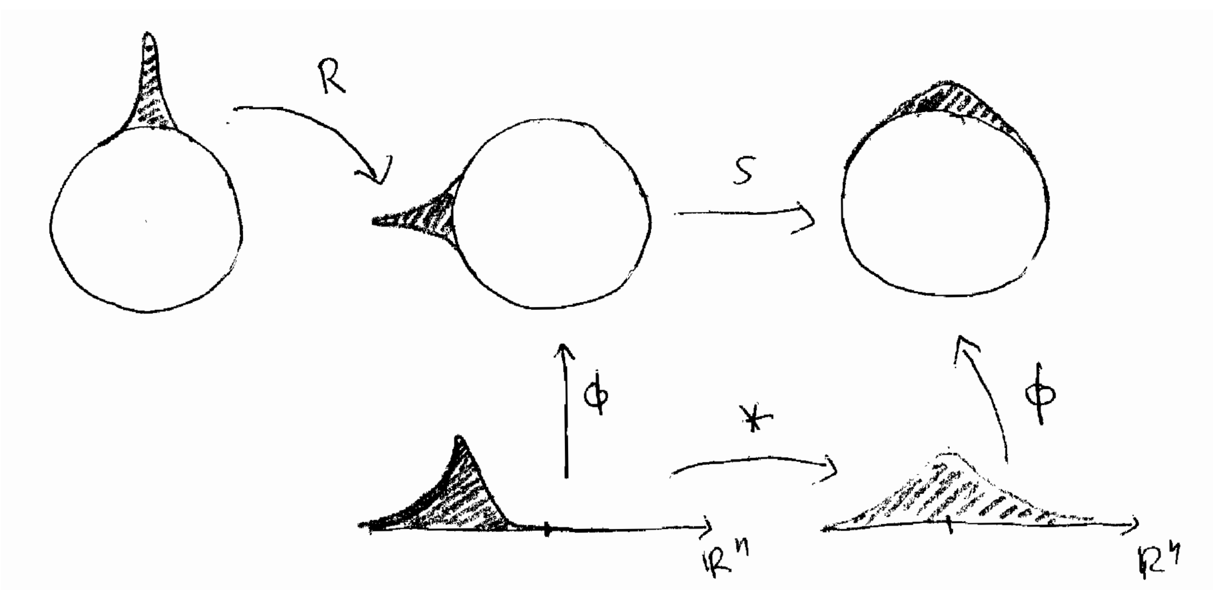


Figure 3.2: The competing symmetries $\mathcal{S}$ and $\mathcal{R}$.

Note that $\mathcal{S}F$ is symmetric decreasing about the south pole, i.e., $\mathcal{S}F(\xi)$ is a nonincreasing function of the last component ξ_{n+1}. By Riesz' inequality and conformal invariance,

$$I(F) \leq I(\mathcal{S}F)\,,\quad I(F) = I(\mathcal{R}F)\,.$$

One complicating factor is that equality in Riesz' inequality occurs whenever f is symmetric decreasing about some point, not necessarily the origin. In the proof, we introduce the auxiliary functional

$$J(F) = \int_{\mathbb{S}^n} F(\xi)\, d\xi\,.$$

Using the isometry of $L^p(\mathbb{S}^n)$ with $L^p(\mathbb{S}^n)$, we can write

$$J(F) = \int_{\mathbb{R}^n} f(x) h(x)^{n/q}\, dx\,, \quad \frac{1}{p} + \frac{1}{q} = 1\,,$$

where $f = \mathcal{U}_\phi F$. By the Hardy-Littlewood inequality and conformal invariance,

$$J(F) \le J(\mathcal{S}F)\,, \quad J(F) = J(\mathcal{R}F)\,.$$

Equality in the Hardy-Littlewood inequality implies that f is already symmetrically decreasing. In particular, $F(\xi)$ can depend only on the last component ξ_{n+1}.

PROOF OF PROPOSITION 3.4. *Compactness.* Assume, by approximation, that F is bounded. Since constant functions are invariant under $\mathcal{S}$ and $\mathcal{R}$ and both operations are order-preserving, the sequence $F_k = (\mathcal{S}\mathcal{R})^k F$ is uniformly bounded. Furthermore, the functions $F_k(\xi)$ are decreasing functions of ξ_{n+1} for $k \ge 1$. We select a subsequence F_{k_j} that converges on all rational points, and thus, (by monotonicity) pointwise almost everywhere to some limiting function G. This is known as *Helly's selection principle*. By dominated convergence, F_{k_j} converges to G strongly in L^p.

Identification of the limit. By construction, G depends only on ξ_{n+1}. On the other hand, $\mathcal{R}G(\xi)$ also depends only on ξ_{n+1}, forcing G to be constant. To see this, we use that the value of the auxiliary functional J increases along the sequence, and so $J(F_k) \to J(G)$ along the entire sequence. It follows that

$$J(G) \le J(\mathcal{S}\mathcal{R}G) = \lim_{k\to\infty} J(\mathcal{S}\mathcal{R}F_k) = \lim_{k\to\infty} J(F_{k+1}) = J(G)\,,$$

which implies that $J(\mathcal{R}G) = J(\mathcal{S}\mathcal{R}G)$, proving the claim. Since $\mathcal{S}$ and $\mathcal{R}$ preserve L^p-norms, we conclude that $G = (n\omega_n)^{-1/p}||F||_p \cdot 1$. We have shown that $F_{k_j} \to G$. Since this limit does not depend on the subsequence, the entire sequence converges and the lemma follows. □

Proposition 3.4 implies that for every nonnegative function $F \in L^p(\mathbb{S}^n)$,

$$I(F) \le I(F_k) \to ||F||^2_{L^p} (n\omega_n)^{-2/p} \int_{\mathbb{S}^n}\int_{\mathbb{S}^n} |\xi - \eta|^{-\lambda}\, d\xi d\eta\,,$$

with equality certainly if F is equivalent to a constant under a conformal change of variables. (Lieb used the sharp rearrangement inequality for I to show that these are all the equality cases.) The sharp Hardy-Littlewood-Sobolev inequality on $\mathbb{R}^n$ follows by transforming back to $L^p(\mathbb{R}^n)$. □

Exercise 3.5 (Conformal invariance of the L^2-norm of the gradient) *Set $r = 2^* = \frac{2n}{n-2}$ for $n \geq 3$. If γ is a conformal transformation on $\mathbb{R}^n$ with conformal factor h, define by*

$$\mathcal{U}_\gamma f(x) = f(\gamma(x))h(x)^{n/r}$$

the corresponding isometry of L^r onto itself. Verify that the functional

$$J(f) = \int_{\mathbb{R}^n} |\nabla f|^2$$

is invariant under $\mathcal{U}_\gamma$.

Exercise 3.6 *Set $r = 2^* = \frac{2n}{n-2}$ for $n \geq 3$. Use competing symmetries to prove that the sharp Sobolev constant in $\mathbb{R}^n$ for $n \geq 3$*

$$C(n,2) = \inf \frac{||\nabla f||_2}{||f||_r}\,, \quad (r = 2^* = \frac{2n}{n-2})$$

is assumed for

$$f(x) = \frac{1}{(1+|x|^2)^{n/r}}\,.$$

3.2 Iterated Steiner and Schwarz symmetrizations

As a second example, we show how to approximate the symmetric decreasing rearrangement by simpler partial symmetrizations. In Section 2.1 we defined the Steiner symmetrization $\mathcal{S}A$ of a measurable set A by symmetrizing its intersection with each line $x_1, \dots x_{n-1} = Const.$, and its *Schwarz symmetrization* $\mathcal{T}A$ by symmetrizing the $n-1$-dimensional cross section $A^t = \{\hat{x} \mid (\hat{x}, t) \in A\}$ in $\mathbb{R}^{n-1}$.

Exercise 3.7 *Give an example of a set $A \subset \mathbb{R}^2$ and a non-trivial rotation $\mathcal{R}$ in the plane such that $\mathcal{TSR}A = A$ but A is not a ball. In particular, the sequence $(\mathcal{TSR})^k A$ does not converge to A^* in symmetric difference.*

Proposition 3.8 (Competing symmetries for Steiner-Schwarz symmetrization) *Let A be a set of finite volume in $\mathbb{R}^n$. Let $\mathcal{TS}A$ be the Steiner-Schwarz symmetrization of A, and let $\mathcal{R}$ be a rotation that acts in the $x_1 - x_n$-coordinate plane as a rotation by an angle θ that is an irrational multiple of π, and fixes the remaining $n-2$ coordinates. Then*

$$(\mathcal{TSR})^k A \to A^* \quad (k \to \infty)$$

with respect to symmetric difference.

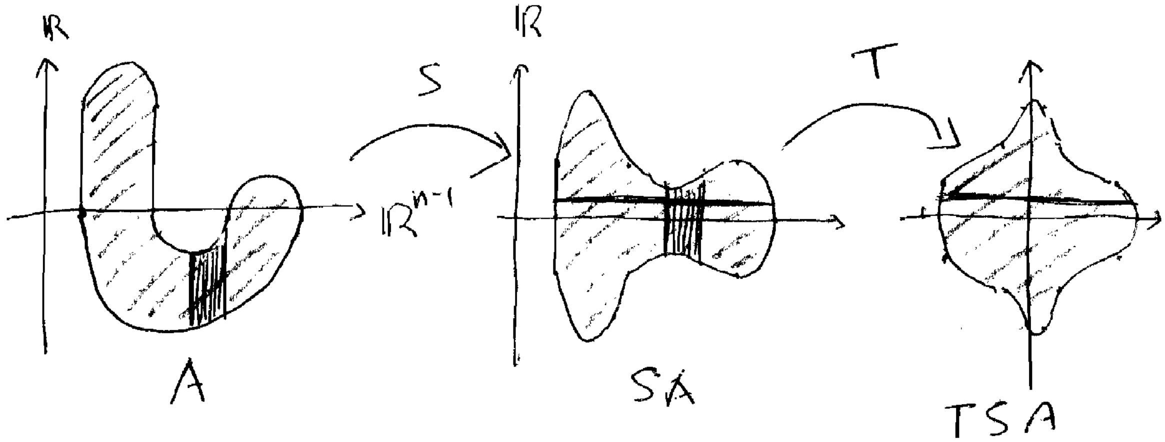


Figure 3.3: Steiner and Schwarz symmetrization.

PROOF. *Compactness.* Assume by approximation that A is bounded, i.e., A lies in a centered ball B. Set $A_k = (\mathcal{TSR})^k A$. By definition of the Steiner and Schwarz symmetrization, there exist nonnegative, nonincreasing functions h_k on the real line such that

$$A_k = \left\{(\hat{x}, x_n) \in \mathbb{R}^n \;\mid\; |\hat{x}| < h_k(|x_n|)\right\}.$$

Since Steiner and Schwarz symmetrization preserve centered balls and reduce symmetric differences, $A_k \subset B$, and thus the functions h_k are uniformly bounded and their support is contained in a common compact interval.

By Helly's selection principle, we can select a subsequence h_{k_j} that converges on all rational points, and thus, (by monotonicity) pointwise almost everywhere to some limiting function. By dominated convergence, the sets A_{k_j} converge with respect to symmetric difference to some limiting set C.

Identification of the limit. Consider the auxiliary functional

$$J(A) = \int_A e^{-(x^2+y^2)}.$$

It has the property that

$$J(A) \leq J(\mathcal{TS}A), \quad J(A) = J(\mathcal{R}A),,$$

with equality in the first inequality only if A agrees with $\mathcal{TS}A$ up to a null set. In particular, A must be symmetric under reflection at the $x_n = 0$ hyperplane and under rotation about the x_n-axis. By monotonicity, the value of J converges along the entire sequence A_k,

$$\lim_{k\to\infty} J(A_k) = J(C).$$

It follows that

$$J(C) = J(\mathcal{TSR}C).$$

By construction C, is symmetric about the x_n-axis, and by the strict rearrangement inequality for J, $\mathcal{R}C$ is also symmetric about the x_n-axis. We conclude that C is symmetric under all rotations at two axes that enclose an angle θ. These rotations generate a dense subgroup of the full rotation group, and thus $C = A^*$. This shows that $A_{k_j} \to A^*$ in symmetric difference. Since the limit does not depend on the subsequence, the entire sequence converges. $\square$

Exercise 3.9 *Let f be a nonnegative continuous function with compact support on $\mathbb{R}^n$. Let $\mathcal{R}$ be the rotation from Proposition 3.8, and set $\mathcal{R}f = f \circ \mathcal{R}^{-1}$. Prove that*

$$(\mathcal{T}\mathcal{S}\mathcal{R})^k f \to f^* \text{ uniformly}.$$

Question 3.10 *How do the sequences obtained by sompeting symmetries differ from the approximations constructed in the previous section, specifically Exercise 2.19?*

These convergence results can be strengthened in different ways. If A is compact, one can show that $A_k \to A^*$ also with respect to the Hausdorff metric. In fact, both the in-radius and the out-radius converge to the radius of the ball A^*. Furthermore, A_k has finite perimeter for $k \geq 2$, and

$$\mathrm{Per}(A_k) \to \mathrm{Per}(A^*).$$

If f is a nonnegative measurable function that vanishes at infinity, it is not hard to show that

$$(\mathcal{T}\mathcal{S}\mathcal{R})^k f \to f^* \text{ in measure}.$$

If $f \in L^p$, the sequence converges strongly in L^p; if $f \in W^{1,p}$, then also

$$\nabla(\mathcal{T}\mathcal{S}\mathcal{R})^k f \to \nabla f^* \text{ weakly in } L^p.$$

Convergence results analogous to Proposition 3.8 can be proved for many other sequences of rearrangements. The following sequence was used by Brascamp, Lieb, and Luttinger (1974):

Question 3.11 *Denote by $\mathcal{S}_i$ Steiner symmetrization with respect to the i-th coordinate axis, and let $\mathcal{R}$ be a rotation in $\mathbb{R}^n$. Under what conditions on $\mathcal{R}$ does the sequence $(\mathcal{S}_n \ldots \mathcal{S}_1 \mathcal{R})^k A$ converge to A^* in symmetric difference for every set $A \subset \mathbb{R}^n$ of finite volume?*

3.3 References

[Lu] L. Lusternik, "Die Brunn-Minkowskische Ungleichung für beliebige meßbare Mengen". Dokl. Acad. Sci. USSR, Vol. 3 (1935), p. 55-58.

Iterated Steiner-Schwarz is used to prove the Brunn-Minkowski inequality. But the proof of the main convergence lemma is incomplete — since convergence is established only along a subsequence, it is not clear why the limit should be a ball.

[So] S. L. Sobolev , "On a theorem of functional analysis". Math. Sb. (N.S.) Vol. 4 (1938), p. 471-497. *Translated in:* AMS Transl. (2) Vol. 34 (1963), p. 39-68.

The Hardy-Littlewood-Sobolev inequality in $\mathbb{R}^n$ is proved with a non-sharp constant. The proof relies on the convergence of the iterated Steiner-Schwarz symmetrization sequence claimed in [Lu].

[BLL] H. J. Brascamp, E. H. Lieb, and J. M. Luttinger, "A general rearrangement inequality for multiple integrals." J. Funct. Anal., Vol. 17 (1974), p. 227-237.

The appendix establishes that sequences of Steiner symmetrizations converge to the symmetric decreasing rearrangement.

[L83] E. H. Lieb, "Sharp constants in the Hardy-Littlewood-Sobolev and related inequalities". Ann. Math. (2), Vol. 118 (1983), p. 349-374.

It is shown that optimizers exist for the Hardy-Littlewood-Sobolev inequality. In the conformally invariant case, the optimizers are identified by a symmetrization argument.

[CL] E. A. Carlen and M. Loss, "Extremals of functionals with competing symmetries". J. Funct. Anal., Vol. 88 (1990), p. 437-456.

The competing symmetries principle is developed abstractly, then applied to the conformally invariant case of the HLS inequality.

[LL] Elliott Lieb and Michael Loss, "Analysis". AMS Graduate Studies in Mathematics, Vol. 14 (1987, second edition 2001).

Chapter 4 explains competing symmetries and the convergence of the iterated Steiner-Schwarz symmetrization sequence claimed from [Lu].

4 Proof of the classical inequalities

4.1 Riesz' inequality

We turn to the general case of Riesz' inequality, where all three functions may vary.

Theorem 4.1 (Riesz' rearrangement inequality) *Let f, g, h be nonnegative measurable functions on $\mathbb{R}^n$ that vanish at infinity. Then*

$$\int\int f(x)g(y)h(x-y)\,dxdy \le \int\int f^*(x)g^*(y)h^*(x-y)\,dxdy\,,$$

in the sense that the left hand side is finite whenever the right hand side is finite.

The theorem was first proved by Riesz (1930) in one dimension. An alternate proof appears in the book of Hardy, Littlewood and Pólya. Sobolev (1937) extended the result to $\mathbb{R}^n$ by an induction over the dimension, similar to our argument above. Brascamp, Lieb, and Luttinger (1974) showed that functionals of the form

$$\int_{(\mathbb{R}^k)^m} \prod_{i=1}^m f_i\left(\sum_{j=1}^n \eta_{ij}x_j\right)\, dx_1, \ldots dx_n$$

can only increase under rearrangement. Here, the η_{ij} form an arbitrary real $n \times m$ matrix.

One interesting property that Riesz' inequality shares with the Brascamp-Lieb-Luttinger inequality is invariance under affine transformations: If L is a linear transformation of determinant ± 1, and $a = b + c \in \mathbb{R}^n$, then replacing $f(x)$ with $f(Lx + a)$, $g(x)$ with $g(Lx + b)$ and $h(x)$ with $h(Lx + c)$ changes neither f^*, g^* and h^* nor the value of the integral.

The characterization of equality cases is complicated, but if two of the three functions are known to have level sets of arbitrary volume, then equality implies that f, g, and h are related to f^*, g^*, and h^* by an affine transformation. If f, g, and h are characteristic functions of measurable sets A, B, and C, then the equality cases depend very much on the relative sizes of the three sets (Burchard 1994).

PROOF OF THEOREM 4.1. By the layer-cake principle, its suffices to show that for every triple of subsets A, B, C of finite volume in $\mathbb{R}^n$,

$$\int_A \mathcal{X}_B * \mathcal{X}_C \le \int_{A^*} \mathcal{X}_{B^*} * \mathcal{X}_{C^*}\,.$$

Denote the left hand side of this inequality by $I(A, B, C)$. We proceed by induction over the dimension.

Base case: $n = 1$. We will use a sliding argument due to Brascamp, Lieb and Luttinger. Consider first the case where A, B, and C are intervals, given by

$$A = a_0 + A^*\,, \quad B = b_0 + B^*\,, \quad C = c_0 + C^*\,.$$

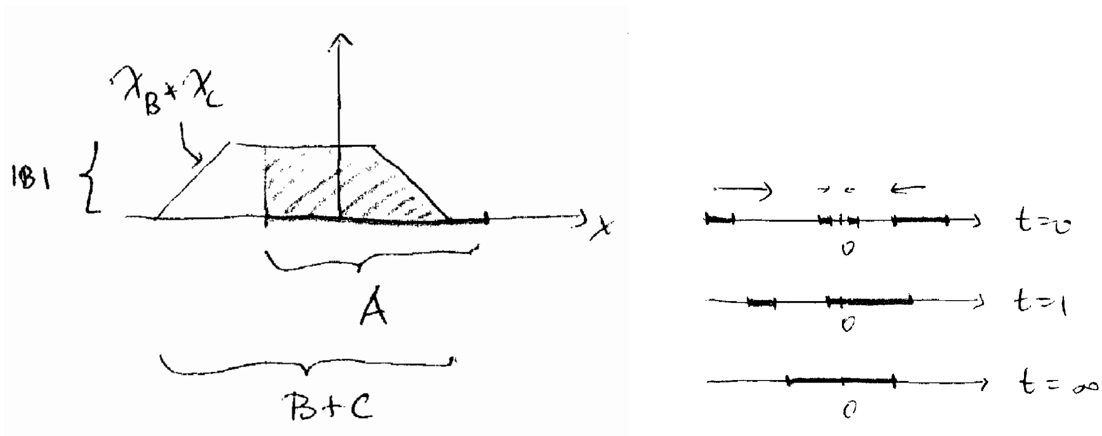
Interpolate between the symmetric rearrangements and the original intervals by setting

$$A(t) = a_0 e^{-t} + A^*\,, \quad B(t) = b_0 e^{-t} + B^*\,, \quad C(t) = c_0 e^{-t} + C^*\,,$$

so that $A(0) = A$ and $A(t) \to A^*$ (in symmetric difference) as $t \to \infty$. Changing variables, we obtain

$$I(A(t), B(t), C(t)) = \int_{(a_0 - b_0 - c_0)e^{-t} + A^*} \mathcal{X}_{B^*} * \mathcal{X}_{C^*}\,.$$

The integrand is an explicitly computable, symmetric decreasing function. It follows that the value $I(A(t), B(t), C(t))$ is nondecreasing in t. This proves the case of single intervals.



(a) The Riesz functional on intervals.

(b) The BLL sliding process.

Figure 4.1: Proof of Riesz' inequality in one dimension.

If A is a finite union of intervals

$$A = \bigcup_{i=1}^{\ell} A_i\,,$$

we first replace any pair of intervals whose closures intersect by their union. The remaining intervals all have positive distance from each other. Set

$$A(t) = \bigcup_{i=1}^{\ell} A_i(t)\,, \quad t \leq t_1$$

where t_1 is the time when the closures of the subintervals $A_i(t)$ first empty intersect. At t_1, fuse the collided intervals these intervals, and continue with the new collection of intervals. Proceed in the same way with the sets B and C. We have already shown that the functional increases with t, and conclude that the claimed inequality holds for finite unions of intervals. For general sets, the inequality follows by approximation, using that I is continuous with respect to symmetric difference, and that Steiner and Schwarz symmetrization can only decrease symmetric difference.

Inductive step: Suppose we have established Riesz' rearrangement inequality in dimensions 1 and $n-1$. Let $A, B, C \subset \mathbb{R}^n$. For $\hat{x} \in \mathbb{R}^{n-1}$, denote by $A_{\hat{x}}$ the intersection of A with the line through $(\hat{x}, 0)$, and correspondingly for B and C. With Fubini's theorem, we can write the functional as

$$I(A,B,C) = \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}^{n-1}} I\big(A_{\hat{x}}, B_{\hat{y}}, C_{\hat{x}-\hat{y}}\big)\, d\hat{x} d\hat{y}\,,$$

and we obtain from the one-dimensional case of Riesz' inequality that I increases under Steiner symmetrization,

$$I(A,B,C) \le I(\mathcal{S}A, \mathcal{S}B, \mathcal{S}C)\,.$$

Similarly, the $n-1$-dimensional case of the inequality implies that I increases under Schwarz symmetrization,

$$I(A,B,C) \le I(\mathcal{T}A, \mathcal{T}B, \mathcal{T}C)\,.$$

I is clearly invariant under simultaneous rotations of A, B, and C. It follows from Proposition 3.8 that

$$I(A,B,C) \le I\big((\mathcal{TSR})^k A, (\mathcal{TSR})^k B, (\mathcal{TSR})^k C\big) \to I(A^*, B^*, C^*)\,. \qquad \square$$

Exercise 4.2 *Let A be a finite union of disjoint intervals, and let $A(t)$ be the set obtained with the Brascamp-Lieb-Luttinger sliding process from the proof of Theorem 4.1. Prove that the center of gravity*

$$a(t) = \mathrm{Vol}(A)^{-1} \int_{A(t)} x\, dx$$

satisfies $\frac{d}{dt} a(t) = -a(t)$.

Exercise 4.3 *Extend the Brascamp-Lieb-Luttinger sliding process to a semigroup of contractions on $L^p(\mathbb{R})$ for $1 \le p < \infty$.*

We next explore the relationship with a fundamental tool of convex analysis, the **Brunn-Minkowski inequality**. The inequality says that for any pair of non-empty measurable sets of finite measure

$$\mathrm{Vol}(B+C)^{1/n} \ge \mathrm{Vol}(B)^{1/n} + \mathrm{Vol}(C)^{1/n}\,,$$

see Exercise 1.6. It is often written in rescaled form as

$$\mathrm{Vol}((1-t)B + tC)^{1/n} \ge (1-t)\mathrm{Vol}(B)^{1/n} + t\mathrm{Vol}(C)^{1/n}\,, \quad 0 \le t \le 1\,.$$

A particular implication is that the cross sectional area of a convex body is a log-concave function of the height. In particular, if the convex body is symmetric under $x \mapsto -x$, then its largest cross section occurs at height 0.

Exercise 4.4 *Show that Riesz' inequality (with $A = B + C$) implies that the Brunn-Minkowski inequality holds for every pair for non-empty open sets of finite volume.*

Hint: Convince yourself that

$$B + C = \left\{ x \in \mathbb{R}^n \mid \mathcal{X}_B * \mathcal{X}_C(x) > 0 \right\},$$

and show that $(B + C)^* \supset B^* + C^*$.

Exercise 4.5 *It is known that equality in the Brunn-Minkowski inequality occurs if and only if B and C are obtained by removing null sets from two scaled and translated versions of the same closed convex set. Use this fact to construct two sets B and C and a polarization σ such that*

$$\mathrm{Vol}(B + C)^{1/n} < \mathrm{Vol}(B^\sigma + C^\sigma)^{1/n}.$$

Exercise 4.6 (Riesz' functional does not satisfy a polarization inequality) *Use Exercises 4.2 and 4.3 to demonstrate by example that the Riesz functional*

$$I(f, g, h) = \int \int f(x) g(y) h(x - y)\, dx dy$$

may decrease as well as increase under polarization.

4.2 The Pólya-Szegő inequality

Theorem 4.7 (Pólya-Szegő inequality) *If $f \in W^{1,p}(\mathbb{R}^n)$ for some $1 \leq p \leq \infty$, then*

$$||\nabla f||_p \geq ||\nabla f^*||_p, \quad (1 \leq p \leq \infty).$$

We will give two different proofs of the Pólya-Szegő inequality. The first uses the fact that the Pólya-Szegő functional is invariant under polarization:

Lemma 4.8 (Pólya-Szegő identity for polarization) *If $f \in W^{1,p}(\mathbb{R}^n)$ for some $1 \leq p \leq \infty$, and σ is a reflection, then $|\nabla f^\sigma|$ is equimeasurable with $|\nabla f|$. In particular,*

$$||\nabla f||_p = ||\nabla f^\sigma||_p, \quad (1 \leq p \leq \infty)$$

PROOF. Since the maximum and minimum of $W^{1,p}$-functions is again in $W^{1,p}$, the polarization f^σ is also in $W^{1,p}$. By the usual density arguments, it suffices to compute the gradient for piecewise linear functions away from the singularities. If $x \in X_+$, we obtain for $f(x) \geq f(\sigma x)$ that

$$\nabla f^\sigma(x) = \nabla f(x), \quad \nabla f^\sigma(\sigma x) = \nabla f(\sigma x),$$

and for $f(x) \leq f(\sigma x)$ that

$$\nabla f^\sigma(x) = \sigma \nabla f(\sigma x), \quad \nabla f^\sigma(\sigma x) = \sigma \nabla f(x).$$

Since σ preserves the Euclidean length, this proves the claim. □

PROOF OF THEOREM 4.7 BY POLARIZATION. For $p = \infty$, f is Lipschitz continuous, and $||\nabla f||_\infty$ is its Lipschitz constant. Since rearrangement improves the modulus of continuity, f^* is again Lipschitz with the same constant, proving the claim in this case.

For $1 \leq p < \infty$, we use again the approximation by a fat layer, as in Eq. (2.4). Note that

$$|\nabla f_\varepsilon| = |\nabla f| \mathcal{X}_{\{\varepsilon < f(x) < \varepsilon + \varepsilon^{-1}\}} ,$$

which converges monotonically to $|\nabla f|$ as $\varepsilon \to 0$. So we may assume that f is bounded and integrable. We next approximate f by a sequence $\{g_k\}$ in $\mathcal{S}(f)$ that converges to f strongly in L^p. Since $||\nabla g_k||_p = ||\nabla f||_k$ by Lemma 4.8, the sequence of gradients is uniformly bounded, and we may choose a subsequence, again denoted by g_k, such that

$$\nabla g_k \to \nabla f^* \text{ weakly in } L^p .$$

Since the p-norm is convex, it is weakly lower semicontinuous, and we conclude that

$$||\nabla f||_p = \lim ||\nabla g_k||_p \geq ||\nabla f^*||_p ,$$

as claimed. □

Exercise 4.9 (Concentration inequality for the gradient) *Prove that for every $\varepsilon > 0$*

$$\sup_{A:\mathrm{Vol}(A)=\varepsilon} \int_A |\nabla f| \, dx \geq \sup_{A:\mathrm{Vol}(A)=\varepsilon} \int_A |\nabla f^*| \, dx .$$

Characterizing the equality cases in the Pólya-Szegő inequality turns out to be quite complicated — and the proof we just presented gives no information. More can be inferred from the following proof:

PROOF OF THEOREM 4.7 FROM THE ISOPERIMETRIC INEQUALITY. By the co-area formula,

$$\int |\nabla f(x)|^p \, dx = \int_0^\infty \int_{f^{-1}(t)} |\nabla f|^{p-1} \, d\sigma dt .$$

Consider the inner integral on the right hand side. Jensen's inequality, applied to the convex function $s \mapsto s^{-(p-1)}$ gives

$$\int_{f^{-1}(t)} |\nabla f|^{p-1} \frac{d\sigma}{\mathrm{Per}\big(\{f > t\}\big)} \geq \left(\int_{f^{-1}(t)} |\nabla f|^{-1} \frac{d\sigma}{\mathrm{Per}\big(\{f > t\}\big)} \right)^{-(p-1)} . \tag{4.1}$$

If we replace f with f^*, the perimeter of the level set decreases, but what happens to the remaining integral?

Since the volume of the set of critical points decreases under symmetric decreasing rearrangement, the right hand side of Eq. (1.8) increases. Looking at the left hand sides of Eq. (1.8) for f and f^*, we see that for almost every $t > 0$

$$\int_{f^{-1}(t)} |\nabla f|^{-1}\, d\sigma \leq \int_{(f^*)^{-1}(t)} |\nabla f^*|^{-1}\, d\sigma\,.$$

We combine this with the isoperimetric inequality and Eq. (4.1) to arrive at

$$\begin{aligned}
\int_{f^{-1}(t)} |\nabla f|^{p-1}\, d\sigma &\geq \operatorname{Per}\bigl(\{f > t\}\bigr)^p \cdot \left(\int_{(f)^{-1}(t)} |\nabla f|^{-1}\, d\sigma\right)^{-(p-1)} \\
&\geq \operatorname{Per}\bigl(\{f^* > t\}\bigr)^p \cdot \left(\int_{(f^*)^{-1}(t)} |\nabla f^*|^{-1}\, d\sigma\right)^{-(p-1)} \\
&= \int_{(f^*)^{-1}(t)} |\nabla f^*|^{p-1}\, d\sigma\,.
\end{aligned}$$

In the last step, we have used that Jensen's inequality in Eq. (4.1) holds with equality when $f = f^*$, because $|\nabla f^*|$ is constant on the level surface. The claim follows upon integration over t. $\square$

Suppose that $||\nabla f||_p = ||\nabla f^*||_p$ for some function $f \in W^{1,p}$ and some $p < \infty$. It follows from the equality statement for the isoperimetric inequality that all level sets of f are balls. For $p = 1$, nothing more can be said. For $p > 1$ one might hope that these level sets should also be concentric. but that is false unless the critical points of f with values strictly between 0 and $\sup f$ form a null set.

Exercise 4.10 *Let Ψ be a convex increasing function on the positive half-line with $\Psi(0) = 0$. Use the co-area formula to prove that*

$$\int \Psi\bigl(|\nabla f(x)|\bigr)\, dx \geq \int \Psi\bigl(|\nabla f^*(x)|\bigr)\, dx\,.$$

Hint: The function $s \mapsto s\Psi(s^{-1})$ is convex and decreasing.

4.3 Talenti's inequality

Theorem 4.11 (Talenti's comparison principle for the Laplacian) *Let f be a smooth nonnegative function with compact support on $\mathbb{R}^n$ for some $n > 2$, and let f^* be its symmetric decreasing rearrangement. If u and v vanish at infinity and solve*

$$-\Delta u = f\,, \quad -\Delta v = f^*\,,$$

then $u^*(x) \leq v(x)$ for all $x \in \mathbb{R}^n$.

Note that u and v exist, and are uniquely determined by the equation. Both are nonnegative by the maximum principle. The assumptions on f can be replaced by the condition that f is measurable and decays suitably at infinity.

Exercise 4.12 *Show that necessarily $v = v^*$.*

PROOF OF THEOREM 4.11. The idea is to bound the distribution function of u, given by

$$\mu(t) = \mathrm{Vol}\bigl\{x \mid u(x) > t\}\bigr) ,$$

in terms of f^*. Since μ is non-increasing, it is differentiable at almost every value $t > 0$. By Eq. (1.8), its derivative satisfies for almost every $t > 0$

$$-\mu'(t) \geq \int_{\{u=t\}} |\nabla u|^{-1}\, d\sigma .$$

Writing the integrand on the right hand side as

$$|\nabla u(x)|^{-1} = \sup_{\varepsilon>0} \bigl\{2\varepsilon - \varepsilon^2 |\nabla u(x)|\bigr\} , \tag{4.2}$$

we see that

$$-\mu'(t) \geq \sup_{\varepsilon>0} \left\{ 2\varepsilon \int_{\{u=t\}} 1\, d\sigma - \varepsilon^2 \int_{\{u=t\}} |\nabla u|\, d\sigma \right\} . \tag{4.3}$$

The first term on the right hand side is given by

$$\int_{\{u=t\}} 1\, d\sigma = \mathrm{Per}\bigl(\{u > t\}\bigr) \geq \mathrm{Per}\bigl(\{u^* > t\}\bigr) .$$

For the second term, we use that $\nabla u(x)$ is a negative multiple of the exterior normal $N(x)$ to the level set $\{u = t\}$, and apply Gauss' divergence theorem to see that

$$\int_{\{u=t\}} |\nabla u|\, d\sigma = - \int_{\{u=t\}} \nabla u \cdot N\, d\sigma = \int_{\{u>t\}} -\Delta u(y)\, dy .$$

From the elliptic equation for u and the Hardy-Littlewood inequality, we obtain

$$\int_{\{u=t\}} |\nabla u|\, d\sigma = \int_{\{u>t\}} f(y)\, dy \leq \int_{\{u^*>t\}} f^*(y)\, dy .$$

We insert the inequalities into Eq. (4.3) and minimize over ε to arrive at

$$-\mu'(t) \geq \bigl(\mathrm{Per}(\{u^* > t)\bigr)^2 \left(\int_{\omega_n |y|^n < \mu(t)} f^*(y)\, dy \right)^{-1} .$$

The key observation is that the quantity on the right hand side is determined solely by the distribution functions of u and f. Collecting terms results in

$$-\mu'(t)\int_{\omega_n|y|^n<\mu(t)} f^*(y)\,dx \geq (n\omega_n)^2\left(\frac{\mu(t)}{\omega_n}\right)^{2-\frac{2}{n}} .$$

To express this differential inequality in terms of u^*, we write

$$u^*(x) = \eta(|x|)$$

for some nonincreasing function η, and change variables $t = \eta(r)$. Since $\mu \circ \eta(r) = \omega_n r^n$ by the definition of the distribution function, we compute for the derivatives $\mu'(t)\eta'(r) = n\omega_n r^{n-1}$. In the new variables, we obtain after collecting terms

$$-\eta'(r) \leq \left(n\omega_n r^{n-1}\right)^{-1}\int_{|y|<r} f^*(y)\,dy\,. \tag{4.4}$$

The change of variables is justified so long as $t = \eta(r)$ is a regular value of u^*, i.e., if μ does not jump there. But Eq. (4.4) also holds if μ jumps at $t = \eta(r)$, because $\eta'(r)$ must vanish there while the left hand side is nonnegative. Integrating, and using that η vanishes at infinity, we arrive at

$$\eta(r) \leq \frac{1}{n\omega_n}\int_r^\infty\int_{|y|<s} f^*(y)s^{-n+1}\,dyds\,.$$

If f is already symmetric decreasing, then all the above inequalities hold with equality. We may exchange the integration with Fubini's theorem and compute the integral over s explicitly. This gives

$$u^*(x) \leq v(x) = \frac{1}{n(n-2)\omega_n}\int_{\mathbb{R}^n} f^*(y)\left(\max\{|x|,|y|\}\right)^{-n+2}\,dy\,.$$

The representation for v agrees with the standard representation for v in terms of the fundamental solution of Laplace's equation. $\square$

Exercise 4.13 *Simplify the proof of Talenti's inequality by replacing Eq. (4.2) with Jensen's inequality in Eq. (4.1). When do these inequalities hold with equality?*

Talenti's inequality allows to bound the norm of the solution of an elliptic PDE in any rearrangement-invariant Sobolev space by the norm of the solution of an associated radially symmetric problem. This is particularly useful for defining solutions when the right hand side is only measurable. The inequality extends to elliptic problems of the form

$$-\sum_{-i,j=1}^n \partial_i a_{ij}(x)\partial_j u = f\,,$$

where the matrix $\big(a_{ij}\big)_{i,j=1}^n$ is symmetric and satisfies the ellipticity condition

$$\sum_{i,j=1}^n a_{ij}(x)\,\xi_i\xi_j \geq \sum_{i=1}^n \xi_i^2\,, \quad (\text{for all } x, \xi \in \mathbb{R}^n)\,.$$

It is valid also for Dirichlet problems on bounded domains $A \subset \mathbb{R}^n$; the symmetrized problem then is a Dirichlet problem on A^*.

Exercise 4.14 (Talenti's principle for Schrödinger equations) *Let V be a smooth nonnegative function on $\mathbb{R}^n$ that grows at infinity in the sense that for each $t > 0$, the set*

$$\{x \mid V(x) < t\}$$

has finite measure. Formulate and prove Talenti's comparison principle for the Schrödinger equation with potential V,

$$-\Delta u + V(x)u = f\,.$$

Hint: Define the **symmetric increasing rearrangement** of V by $V_* = -\log\big(e^{-V}\big)^*$, so that

$$\{x \mid V(x) < t\}^* = \{x \mid V_*(x) < t\}$$

for all $t > 0$. The nonnegativity of V guarantees the existence, uniqueness, and positivity of the solution u.

Exercise 4.15 (Talenti's principle for the p-Laplacian) *Fix $1 < p < \infty$, and let f be a smooth nonnegative function with compact support. Use the identity*

$$\frac{1}{a} = \sup_{\varepsilon > 0}\Big\{\varepsilon q - \frac{q}{p}\varepsilon^p a^{\frac{p}{q}}\Big\}\,, \quad \Big(a > 0, \frac{1}{p} + \frac{1}{q} = 1\Big)\,.$$

to extend Talenti's comparison principle to the equation

$$-\mathrm{div}\big(|\nabla u|^{p-2}\nabla u\big) = f\,.$$

Remark: The expression on the left hand side is called the **p-Laplacian** of u. It is quasilinear, but for $p \neq 2$ it is not linear. Weak solutions for this equation can be constructed as minimizers of the convex functional

$$\frac{1}{p}\int_{\mathbb{R}^n} |\nabla u|^p\,dx - \int_{\mathbb{R}^n} f u\,dx$$

in the Sobolev space $W^{1,p}$. Solutions of this equation satisfy a maximum principle; in particular, $u \geq 0$ if $f \geq 0$.

Exercise 4.16 *Use Talenti's principle for the p-Laplacian to show that the solutions of the equations*

$$-\mathrm{div}\big(|\nabla u|^{p-2}\nabla u\big) = f\,, \qquad -\mathrm{div}\big(|\nabla v|^{p-2}\nabla v\big) = f^*$$

satisfy

$$||\nabla u||_p \leq ||\nabla v||_p\,, \quad (1 < p < \infty)\,.$$

4.4 References

[R] F. Riesz, "Sur une inégalité intégrale." J. London Math. Soc., Vol. 5 (1930), p. 162-168.

Proof of Inequality (1.9), where all three functions may vary, in one dimension. (A sketch of this proof is reproduced in [HLP, Chapter 10].) A possible higher-dimensional extension mentioned but not described.

[BLL] H. J. Brascamp, E. H. Lieb, and J. M. Luttinger, "A general rearrangement inequality for multiple integrals." J. Funct. Anal., Vol. 17 (1974), p. 227-237.

The ultimate generalization of Riesz' inequality in $\mathbb{R}^n$, and a new proof (by the sliding argument).

[Bu] A. Burchard, "Cases of equality in the Riesz rearrangement inequality". Ann. Math. (2), Vol. 143 (1996), p. 499-527.

Riesz' proof [R] is modified to identify the equality cases of Eq. (1.9) in one dimension. Higher dimensions are treated inductively, with Steiner and Schwarz symmetrization.

[BZ] J. Brothers and W. Ziemer, "Minimal rearrangements of Sobolev functions". J. Reine Angew. Math., Vol. 384 (1988), 153-179.

The equality cases in the Pólya-Szegő inequality (1.11).

[T76] Giorgo Talenti, "Elliptic equations and rearrangements." Annali della Scuola Normale di Pisa, Classe di Scienze (4. série), Vol. 3 (1976), p. 697-718.

Talenti's comparison principle (1.14) for elliptic equations.

[ATL] A. Alvino, G. Trombetti and P.-L. Lions, "Comparison results for elliptic and parabolic equations via Schwarz symmetrization". Annales de l' I. H. P., Section C, Vol. 7 (1990), p. 37-65.

Talenti's comparison principle is extended to more general elliptic and parabolic equations.

5 Transportation methods for geometric inequalities

5.1 Monge's problem

Rearrangements can be viewed as special instances of **transportation plans**, which move a given mass distribution to another distribution of the same total mass. To give a proper definition, let μ and ν be measures on $\mathbb{R}^n$, with the same total mass $\mu(\mathbb{R}^n) = \nu(\mathbb{R}^n) < \infty$. We say that a map $T : \mathbb{R}^n \to \mathbb{R}^n$ **pushes** μ **forward to** ν, if for every ν-integrable function b on $\mathbb{R}^n$,

$$\int b(x)\, d\mu(x) = \int b(y)\, d\nu(y)\,.$$

In that case, we write $T\#\mu = \nu$.

Monge's problem consists of minimizing a **cost functional**

$$C(\mu, \nu) = \int c(x, Tx)\, d\mu(x)$$

among all maps that push μ forward to ν. Here, $c(x, y)$ is the infinitesimal cost of transporting a unit of mass from the point x to the point y, typically given by an increasing function of the distance $|x - y|$. In general, a minimizing solution to Monge's problem need not exist. The key to solving the problem is a weak formulation due to Kantorovich that replaces the transport map by a **transportation plan**, which is a measure on $\mathbb{R}^n \times \mathbb{R}^n$ that specifies the source and destination location for each piece of mass.

The case where the infinitesimal cost function is given by $c(x, y) = |x - y|^2$ on $\mathbb{R}^n$ plays a special role. In that case, the problem is known to have a unique solution. If the measures $d\mu(x) = F(x)\, dx$ and $\nu(y) = G(y)\, dy$ are absolutely continuous with respect to Lebesgue measure, then **Brenier's theorem** states that the solution is given by a transport map $T = \nabla\phi(x)$, where ϕ is a convex function. The transport map is determined uniquely μ-almost everywhere by the condition that $T = \nabla\phi(x)$ for a convex function $\phi : \mathbb{R}^n \to \mathbb{R}$.

For the applications we will discuss below, we need only the fact that there exists a convex function ϕ on $\mathbb{R}^n$ such that $\nabla\phi$ pushes $F(x)\, dx$ forward to $G(y)\, dy$. This means that

$$\int b(\nabla\phi(x))F(x)\, dx = \int b(y)G(y)\, dy \tag{5.1}$$

for all ν-integrable functions b. If ϕ is twice continuously differentiable and F and G are strictly positive, the Change of Variables formula implies the **Monge-Ampère equation**

$$F(x) = G(\nabla\phi(x)) \det D^2\phi(x)\,. \tag{5.2}$$

Caffarelli regularity theory provides conditions on F and G that guarantee that ϕ solves the Monge-Ampère equation in the classical sense. McCann proved that even if ϕ is just convex, the Monge-Ampère equation holds μ-almost everywhere, if the second derivative is interpreted pointwise (as a limit of a difference quotient, wherever it exists).

Exercise 5.1 *In one dimension, the Brenier map is given by a non-decreasing function. If $d\mu(x) = F(x)\,dx$ and $d\nu(y) = G(y)\,dy$ are probability measures that are absolutely continuous and have positive density with respect to Lebesgue measure, describe this map, and prove that it is continuous and has a continuous inverse.*

Exercise 5.2 (Interpolation by displacement) *Let $\nabla\phi$ be the Brenier map that pushes a probability measure μ forward to another probability measure ν. Define, for $0 \le t \le 1$ the map $T_t = (1-t)I + t\nabla\phi$, and set $\nu_t = T_t\#\mu$. Prove that*

$$\text{supp}\,(\nu_t) \subset (1-t)\,\text{supp}\,(\mu) + t\,\text{support}\,(\nu)\,.$$

5.2 A transportation proof of the isoperimetric inequality

Consider a bounded set $A \subset \mathbb{R}^n$ with smooth boundary, and denote by A^* be the centered ball of the same volume. We will sketch an argument that Brenier's theorem implies the isoperimetric inequality

$$\text{Per}(A) \ge \text{Per}(A^*)\,.$$

Let $T = \nabla\phi$ be the Brenier map that pushes the uniform distribution on A forward to the uniform distribution on A^*, i.e., for every $C \subset \mathbb{R}^n$,

$$\text{Vol}(C \cap A) = \text{Vol}\big(\nabla\phi(C) \cap A^*\big)\,.$$

The Monge-Ampère equation takes the form

$$\det D^2\phi(x) = 1\,, \quad (x \in A)\,.$$

Denote the radius of A^* by R. Since $\nabla\phi$ maps A to A^*, we must have $|\nabla\phi(x)| \le R$ for all $x \in A$. Thus

$$\text{Per}(A) = \int_{\partial A} 1\,d\sigma \ge \frac{1}{R}\int_{\partial A} \nabla\phi(x)\cdot N(x)\,d\sigma(x) = \frac{1}{R}\int_A \Delta\phi(x)\,dx\,, \tag{5.3}$$

where N is the outward unit normal to ∂A, and we have applied Gauss' divergence theorem in the last step. Since ϕ is convex, its Hessian matrix is positive semidefinite, and its eigenvalues satisfy

$$\frac{1}{n}(\lambda_1 + \cdots + \lambda_n) \ge (\lambda_1 \cdots \lambda_n)^{1/n}$$

by the arithmetic-geometric mean inequality. We conclude from the Monge-Ampère equation that

$$\frac{1}{n}\Delta\phi(x) \ge \big(\det D^2\phi(x)\big)^{1/n} = 1\,.$$

Inserting this into Eq. (5.3) we arrive at

$$\text{Per}(A) \ge \frac{n}{R}\int_A 1\,dx = n\omega_n R^{n-1} = \text{Per}(A^*)\,,$$

proving the claim. $\square$

We have not discussed under what hypotheses on A the change of variables and the divergence theorem are applicable, and whether ϕ is indeed twice differentiable. These questions become relevant for the characterization of equality cases, where A is a priori just a measurable set with finite perimeter and no further regularity properties.

Exercise 5.3 (Brunn-Minkowski from optimal transportation) *Let μ, ν be the uniform measures of density one on two sets $A, B \subset \mathbb{R}^n$ of finite and equal volume, and let $\nu_t = T_t \# \mu$ be the interpolation defined in Exercise 5.2. Prove that ν is absolutely continuous with respect to Lebesgue measure,*

$$d\nu_t(x) = \rho_t(x)\, dx\,,$$

and that its density satisfies $\rho_t(x) \le 1$.

(a) Conclude that

$$\mathrm{Vol}\big((1-t)A + tB\big)^{\frac{1}{n}} \ge (1-t)\mathrm{Vol}(A)^{\frac{1}{n}} + t\mathrm{Vol}(B)^{\frac{1}{n}}\,.$$

Hint: *The function $t \mapsto \big(\det DT_t\big)^{\frac{1}{n}}$ is concave (why?).*

(b) Use a scaling argument to obtain the Brunn-Minkowski inequality for sets A, B of arbitrary finite (not necessarily equal) volume, in its more standard form

$$\mathrm{Vol}(A+B)^{\frac{1}{n}} \ge \mathrm{Vol}(A)^{\frac{1}{n}} + \mathrm{Vol}(B)^{\frac{1}{n}}\,.$$

5.3 The sharp Sobolev inequality

Theorem 5.4 *Fix $1 < p < n$, let $q = \frac{p}{p-1}$ be the dual Hölder index, and set $p^* = \frac{np}{n-p}$. There exists a constant $C(n,p)$ such that*

$$||\nabla f||_p \ge C(n,p)||f||_{p^*}\,, \tag{5.4}$$

provided that $|f|$ vanishes at infinity and $|\nabla f| \in L^p$.

PROOF. We will show that

$$\frac{n(n-p)}{(n-1)p} \cdot \sup_{g \in C_c^\infty(\mathbb{R}^n)} \frac{\int g(y)^{\frac{n-1}{n}p^*}\, dy}{\big(\int g(y)^{p^*}|y|^q\big)^{\frac{1}{q}}\, ||g||_{p^*}} \;=\; \inf_{f \in C_c^\infty(\mathbb{R}^n)} \frac{||\nabla f||_p}{||f||_p^*}\,. \tag{5.5}$$

The extremals on the two sides of the equation are assumed for the function $h(x) = (1+|x|^q)^{-\frac{n}{p^*}}$, i.e.,

$$C(n,p) \;=\; \frac{||\nabla h||_p}{||h||_{p^*}} \;=\; \frac{n(n-p)}{(n-1)p} \cdot \frac{\int h(y)^{\frac{n-1}{n}p^*}\, dy}{\big(\int h(y)^{p^*}|y|^q\big)^{\frac{1}{q}}\, ||h||_{p^*}}\,.$$

To establish Eq. (5.5) and the theorem, it suffices to consider smooth functions f and g with compact support that have been normalized to $||f||_{p^*} = ||g||_{p^*} = 1$. By Brenier's theorem, there exists a convex function ϕ that pushes the probability measure $F(x)dx$ forward to the probability measure $G(x)dx$. We apply the transportation equation in Eq. (5.1) and express the integral in the numerator on the left hand side of Eq. (5.5) as

$$\int G(y)^{1-\frac{1}{n}}\,dy = \int F(x)G(\nabla\phi(x))^{-\frac{1}{n}}\,dx = \int F(x)^{1-\frac{1}{n}}\left(\det D^2\phi(x)\right)^{\frac{1}{n}}\,dx\,. \tag{5.6}$$

In the second step, we have used the Monge-Ampère equation in Eq. (5.2) to write

$$G\big(\nabla\phi(x)\big)^{-\frac{1}{n}} = F(x)^{-\frac{1}{n}}\big(\det D^2\phi(x)\big)^{\frac{1}{n}}\,.$$

Since ϕ is convex, the matrix $D^2\phi(x)$ is positive semidefinite, and the arithmetic-geometric mean inequality for its eigenvalues says that $\big(\det D^2\phi(x)\big)^{\frac{1}{n}} \le \frac{1}{n}\Delta\phi(x)$, which integrates to

$$\int F(x)^{1-\frac{1}{n}}\big(\det D^2\phi(x)\big)^{\frac{1}{n}}\,dx \le \frac{1}{n}\int F(x)^{1-\frac{1}{n}}\Delta\phi(x)\,dx\,. \tag{5.7}$$

We next express F in terms of f, integrate by parts, and apply Hölder's inequality

$$\begin{aligned}\frac{1}{n}\int F(x)^{1-\frac{1}{n}}\Delta\phi(x)\,dx &= -\frac{(n-1)p}{n(n-p)}\int f(x)^{\frac{n(p-1)}{n-p}}\nabla\phi(x)\cdot\nabla f(x)\,dx \\ &\le \frac{(n-1)p}{n(n-p)}\left(\int f(x)^{p^*}|\nabla\phi(x)|^q\,dx\right)^{1/q}||\nabla f||_p\,. \end{aligned} \tag{5.8}$$

To eliminate $\nabla\phi$, we use once more the transportation property in Eq. (5.1)

$$\int F(x)|\nabla\phi(x)|^q\,dx = \int G(y)|y|^q\,dy\,.$$

Finally, we insert Eqs. (5.7) and (5.8) into Eq. (5.6) and collect terms to arrive at

$$\frac{n(n-p)}{(n-1)p}\cdot\int g(y)^{\frac{n-1}{n}p^*}\,dy\cdot\left(\int g(y)^{p^*}|y|^q\,dy\right)^{-\frac{1}{q}} \le ||\nabla f||_p\,.$$

Since f and g are arbitrary except for the normalization in L^{p^*}, this inequality remains valid if the left hand side is maximized over g and the right hand side is minimized over f. We conclude that the left hand side of Eq. (5.5) is a lower bound for the right hand side.

For the complementary bound, it suffices to exhibit a pair of functions f, g with $||f||_{p^*} = ||g|_{p^*} = 1$ such that all steps of the proof hold with equality. Consider first Eq. (5.7). Equality in the arithmetic-geometric mean inequality holds if $D^2\phi(x)$ is a diagonal matrix. Let us pick $\phi(x) = \frac{1}{2}|x|^2$ so that $\nabla\phi(x) = x$. This means that the transport map is the identity and $f = g$. Next, look at Hölder's inequality in Eq. (5.8). Equality occurs, if $\nabla f(x)$ and $\nabla\phi(x)$ always point in opposite directions and $|\nabla f|^p = c\cdot f^{p^*}|\nabla\phi|^q$ for some constant c. A direct computation verifies that $f(x) = h(x)$ and $\nabla\phi(x) = x$ satisfy this condition. □

Exercise 5.5 *Verify that the Sobolev functional on the right hand side of Eq. (5.5) is invariant under translation, rotation, dilation, and multiplication by constants. Conclude that the Sobolev optimizers include at least the $n+2$-parameter family of translates, dilates and multiples of h. If h is replaced by another member of this family, what happens to ϕ? What are the symmetries of the left hand side of Eq. (5.5)?*

Exercise 5.6 (Displacement convexity) *Let ρ be a nonnegative integrable function on $\mathbb{R}^n$, and assume that ϕ is convex. Define a measure $d\mu(x) = \rho(x)\,dx$, and let ρ_t be the density of the push-forward of μ under $T_t = (1-t)I + t\nabla\phi$. Prove that the function $t \mapsto \int_{\mathbb{R}^n} \rho_t^{1-\frac{1}{n}}(x)\,dx$ is convex.*

5.4 References

[Br] Y. Brenier, "Polar factorization and monotone rearrangement of vector-valued functions". Comm. Pure Appl. Math., Vol. 44 (1991), p. 375-417.

Solution of the optimal transportation problem for the quadratic cost function.

[C] L. A. Caffarelli, "The regularity of mappings with a convex potential". J. Amer. Math. Soc., Vol. 5 (1992), p. 99-104.

[McC95] R. J. McCann, "Existence and uniqueness of monotone measure-preserving maps". Duke Math. J., Vol. 80 (1995), p. 309-323.

Contains the proof that the Monge-Ampère equation (5.2) holds pointwise almost everywhere.

[McC97] R. J. McCann, "A convexity principle for interacting gases". Adv. Math., Vol. 128 (1997), p. 153-179.

"Displacement convexity" is introduced here.

[CNV] D. Cordero-Erausquin, B. Nazaret, and C. Villani, "A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities." Advances in Mathematics, Vol. 182 (2004), p. 307-332.

The sharp Sobolev inequality and its equality cases, for $1 \leq p < \infty$ and arbitrary norms in $\mathbb{R}^n$.

[V] C. Villani, "Topics in Optimal Transportation". AMS Graduate Studies in Mathematics, Vol. 58 (2003).

Chapter 6 contains a sketch of the argument in [CNV].

[G] R. J. Gardner, "The Brunn-Minkowski inequality". Bull. Amer. Math. Soc., Vol. 39 (2002), p. 355-405.

A survey of the Brunn-Minkowski inequality, its relation with other geometric inequalities, and its various proofs.