The Pólya–Szegő inequality for the Steiner Symmetrization in k dimensions

Robert Wegner

Geboren am 19. Juli 1998 in Ludwigshafen am Rhein

January 11, 2022

Bachelorarbeit Mathematik

Betreuer: Prof. Dr. Herbert Koch

Zweitgutachter: Dr. Franz Gmeineder

Mathematisches Institut

Mathematisch-Naturwissenschaftliche Fakultät der

Rheinischen Friedrich-Wilhelms-Universität Bonn

Contents

0 Zusammenfassung

Das Ziel dieser Arbeit ist es eine Version Pólya–Szegő Ungleichung für eine generalisierte Version der Steiner Symmetrisierung zu Beweisen. Hier wird nicht nur entlang 1-dimensionaler Linien, sondern k-dimensionaler Hyperebenen symmetrisiert. Dabei folgen wir dem Buch "Rearrangements and Convexity of Level Sets in PDE" ([4]) von Berhnard Kawohl, in dem solche Ungleichungen für verschiedene Symmetrisierungen und insbesondere die standard Steiner Symmetrisierung bewiesen werden. Dies geschieht im 3. Kapitel.
Um daraus die Ungleichung für die generalisierte Version abzuleiten, zeigen wir vorher im 2. Kapitel dass die Steiner Symmetrisierungen in k Dimensionen durch die in einer Dimension approximiert werden kann. Dies folgt daraus, dass jede Menge durch 1-dimensionale Symmetrisierung entlang gewisser Richtungen beliebig nah an einen Ball mit gleichem Volumen angenähert werden kann. Um diese Aussage zu beweisen benutzen wir eine Methode von Herbert Federer aus "Geometric Measure Theory" [3]. Er beweist eine Annäherung an den Ball in der Hausdorff-Metrik. Wir interessieren uns jedoch für eine Annäherung im L^1 Sinn, also konkret im Abstand der Indikatorfunktionen der Mengen. Wir stellen daher, für Mengen mit einer gewissen Regularität die wir "symmetrische Quader Eigenschaft" nennen, eine Verbindung zwischen diesen beiden Abstandsbegriffen her.
Die Beweise werden grundsätzlich detailliert präsentiert, und wir versuchen die Nutzung unbewiesener "ad-hoc" Aussagen auf Situationen zu beschränken in denen sie nicht das Verständnis behindert.

1 Introduction

The Pólya–Szegő inequality is, in broad terms, the following inequality:

$$\|\nabla f^*\| \leq \|\nabla f\|.$$

Here $*$ is an operation that trades information about the function to make it "nicer", for example by introducing a new symmetry. We call this operation a rearrangement or symmetrization. The intuition behind the inequality being true, whenever it is the case, lies in the fact that the rearrangement concentrates the function and in this process smoothes out hills and valleys, reducing the total amount of gradient. The norm here is merely supposed to represent a way of measuring the total amount of gradient. Everything that was left vague is purposefully missing to emphasize that this inequality can appear in many different forms. We will, however, not deal with this generality and instead aim to prove the following theorem.

Theorem 3.5. *Let $1 \leq k \leq d$, $1 \leq p < \infty$, and let Ω be open. Let $F : \mathbb{R}^d \longrightarrow [0, \infty)$ be a convex, radially symmetric, and radially nondecreasing function which fulfills the requirement that the functional $J_F : W^{1,p}(\mathbb{R}^d) \longrightarrow [0, \infty)$, defined by the formula*

$$J_F(f) = \int_{\overline{\Omega}} F(\nabla f) dm^d,$$

is well-defined and continuous. Finally, let $f \in W^{1,p}_{0,\geq}(\overline{\Omega})$. Then the symmetrized function exists and $S^k f \in W^{1,p}_{0,\geq}(S^k\overline{\Omega})$. Furthermore, we have the inequality

$$J_F(S^k f) \leq J_F(f).$$

Here the symmetrization S^k is a generalized version of the well-known Steiner and Schwarz symmetrizations. For better understanding, let us contrast this rather general version of the theorem with the special case $\Omega = \mathbb{R}^d$, $k = d$, $p = 2$, and $F(x) = |x|^2$. Then the statement is that

$$\|\nabla S^d f\|_{L^2}^2 = \|\nabla f\|_{L^2}^2 \quad \forall f \in W^{1,2}(\mathbb{R}^d).$$

The rearrangement S^d is the Schwarz symmetrization and will be defined in the first section. One could say that this version of the Pólya–Szegő inequality is the most minimalistic and most useful one, in the sense that the statement and the requirements are weak, but there are many applications.

The proofs in this text will be presented in rather high detail, and the goal is to build everything from the definitions up. We will use some theorems without proof, such as the theorem of Banach-Alaoglu. As many of the proofs are geometric in nature, the reader is urged to try to visualize the presented arguments and results.

This text is based on ideas from [4], a book by Kawohl that discusses the Pólya–Szegő inequality for various rearrangements, in particular the Steiner symmetrization. We emulate some of his proofs and definitions while generalizing in some aspects that were not the goal of Kawohl's work.

1.1 Notation

- We apply functions both with and without parentheses. Generally, we consider a function to contain the information of its domain. This means, for example, that if for some $\Omega \subseteq \mathbb{R}^d$ we have a function $f : \Omega \longrightarrow \mathbb{R}$, we consider f to be different from its the extension by zero to all of $\mathbb{R}^d$. For such a function and some $a \in \mathbb{R}$, we write
$$\{f \geq a\} := \{x \in \Omega : f(x) \geq a\}.$$

- We use the notations $B_r := B_r(0)$ for all $r > 0$ and $B_r^c = \mathbb{R}^d \setminus B_r$. We also write $B_r^k = B_r^{\mathbb{R}^k}$ and define the constant $\omega_d = m^d(B_1^d)$.

- All functions that appear are real valued. As we only symmetrize nonnegative functions, we write $L^p_{\geq}$ and $W^{1,p}_{\geq}$ to denote only the nonnegative functions of the respective spaces.

- We write $\mathcal{L}(\mathbb{R}^d)$ for the Lebesgue measurable subsets of $\mathbb{R}^d$.

- We heavily and suggestively use the notation $(x, y, z) \in \mathbb{R}^k \times \mathbb{R}^{d-k} \times \mathbb{R}$. When doing this, we almost exclusively use x, y and z in the positions given above and denote other points by (x', y', z'), (x'', y'', z''), etc. Here x is the coordinate that we symmetrize in, y represents all the other coordinates, and z, if present, represents the values of functions $\mathbb{R}^d \longrightarrow \mathbb{R}$.

- We use the notation $f_A := \frac{1}{m^d(A)} \int_A f \, dm^d$.

- In an abuse of notation, we write $\|\nabla f\|_{L^p}$ when we really mean $\||\nabla f|\|_{L^p}$.

1.2 Rearranging sets and functions

Although we will concern ourselves with a specific rearrangement, we want to start on broader terms to motivate the idea of rearrangements. One begins by defining some function on the measurable sets $* : \mathcal{L}(\mathbb{R}^d) \longrightarrow \mathcal{L}(\mathbb{R}^d)$. This function should have nice properties, but at the very least it should be

- monotone, so $M \subseteq N \Rightarrow M^* \subseteq N^*$

- equimeasurable, so $m^d(M) = m^d(M^*)$.

As the next step, one defines the rearrangement of functions. For some measurable function $f : \mathbb{R}^d \longrightarrow [0, \infty)$, we define

$$\begin{aligned} f^* : \mathbb{R}^d &\longrightarrow [0, \infty) \cup \{\infty\} \\ x &\longmapsto \sup\{c \geq 0 : x \in \{f \geq c\}^*\}. \end{aligned}$$

There are similar ways to define the arrangement of functions and here we have chosen the definition of Kawohl in [4]. This definition is best understood through the visualization for f^* suggested by the following lemma.

Lemma 1.1. *Let $*$ be a monotone, equimeasurable rearrangement on measurable sets, and let $f : \mathbb{R}^d \longrightarrow [0,\infty)$ be a measurable function. Then*

$$\{f^* \geq c\} = \{f \geq c\}^* \text{ up to a set of measure zero} .$$

Proof. With the definition, we can argue

$$x \in \{f^* \geq c\} \iff f^*(x) \geq c \iff \sup\{b \in \mathbb{R} : x \in \{f \geq b\}^*\} \geq c$$
$$\iff \forall b < c\, \exists b \leq b' \leq c : x \in \{f \geq b'\}^*.$$

Monotonicity implies that $\{f \geq c\}^* \subseteq \{f \geq b'\}^* \subseteq \{f \geq b\}^*$. Applying this yields

$$x \in \{f^* \geq c\} \iff x \in \bigcap_{b<c} \{f \geq b\}^* \supseteq \{f \geq c\}^*.$$

To finish the proof, we use continuity from above and equimeasurability:

$$m^d \left(\bigcap_{b<c} \{f \geq b\}^* \setminus \{f \geq c\}^* \right) = \lim_{b \nearrow c} m^d \left(\{f \geq b\}^* \setminus \{f \geq c\}^* \right)$$

$$= \lim_{b \nearrow c} m^d \left(\{f \geq b\} \setminus \{f \geq c\} \right) = m^d \left(\bigcap_{b<c} \{f \geq b\} \setminus \{f \geq c\} \right) = m^d(\varnothing) = 0.$$

Since one set contains the other but the measure of their difference is zero, they are equal up to a set of measure zero. □

We can therefore picture the graph of the rearranged function as the resulting shape when each layer of the graph of the original function is rearranged.

1.3 The Schwarz and Steiner symmetrizations

The standard Steiner symmetrization in a direction v rearranges a set by considering the intersection of each line parallel to v with the set, and then replacing it with a closed interval of the same measure, placed symmetrically above and below the hyperplane $v^\perp$. We will use a generalized definition where the intersection of a set with a hyperplane of some dimension k is considered. The intersection is then rearranged to be a closed ball of the same measure, and this process is done for every orthogonally shifted copy of the hyperplane. In the case $k = d$ this yields the Schwarz symmetrization, which merely replaces a set by a closed ball of the same measure.

Definition 1.2 (Schwarz symmetrization)**.** *Let $d \in \mathbb{N}_{>0}$. We define the Schwarz symmetrization as the function*

$$\begin{aligned} \mathrm{SW}_d : \mathcal{L}(\mathbb{R}^d) &\longrightarrow \mathcal{L}(\mathbb{R}^d) \\ \varnothing &\longmapsto \varnothing \\ M &\longmapsto \left\{ x \in \mathbb{R}^d : \omega_d |x|^d \leq m^d(M) \right\} \text{ otherwise.} \end{aligned}$$

Our generalized Steiner symmetrization essentially applies the Schwarz symmetrization to affine k-dimensional subspaces.

Definition 1.3 (Steiner symmetrization)**.** *Let $k, d \in \mathbb{N}$ such that $1 \le k \le d$. We define the Steiner symmetrization in k dimensions as the function*

$$\begin{aligned} S^k : \mathcal{L}(\mathbb{R}^d) &\longrightarrow \mathcal{L}(\mathbb{R}^d) \\ M &\longmapsto \bigcup_{y \in \mathbb{R}^{d-k}} \mathrm{SW}_k(\{x \in \mathbb{R}^k : (x, y) \in M\}) \times \{y\}. \end{aligned}$$

Furthermore, one can symmetrize with respect to a k dimensional subspace V other than $\mathbb{R}^k \times \{0\}^{d-k}$. We define the Steiner symmetrization in k dimensions with respect to V by the formula

$$S_V(M) = \tau S^k \tau^{-1} M,$$

where $\tau : \mathbb{R}^k \times \{0\}^{d-k} \longrightarrow V$ is an isometric isomorphism. Due to the Schwarz symmetrization being invariant under conjugation with isometries, any τ can be chosen here. For a given vector $v \in \mathbb{R}^d$, we denote by S_v the symmetrization with respect to the line spanned by v. When the direction for a Steiner symmetrization in 1 dimension is clear from the context, we do not mention it. When we talk about several Steiner symmetrizations in 1 dimension, we mean that the symmetrizations may be in different directions. We will also use $*$ to denote Steiner symmetrization.

Note that there is some abuse of notation happening here for the case $k = d$. The reader could try to convince himself that the union over $\mathbb{R}^0$ might just, with enough determination, be interpreted in such a way that one ends up with $S^d = \mathrm{SW}_d$. Alternatively, we could just define the case $k = d$ to be exactly that, but we will not make a distinction for this case in proofs. Furthermore, all theorems that only deal with Steiner symmetrization with respect to a single subvector space V will be formulated and proven for the case $V = \mathbb{R}^k \times \{0\}^k$, as simply rotating, using this case, and rotating back solves the other cases.

We will first show the following technical lemma, whose statement should be geometrically intuitive. Afterwards, we will use this lemma in the proof that Steiner symmetrization decreases the diameter of a set.

Lemma 1.4. *Let $A, B \subset \mathbb{R}^d$ be measurable. Then*

$$\sup\{|a - b| : a \in \mathrm{SW}_d(A), b \in \mathrm{SW}_d(B)\} \le \sup\{|a - b| : a \in A, b \in B\}.$$

Proof. If both sets have measure zero, then the left hand side of the inequality is zero. If any set has infinite measure, then it must be unbounded and the right hand side is infinite. Now we assume neither of these is the case and let r_A and r_B be the radii of $\mathrm{SW}_d(A)$ and $\mathrm{SW}_d(B)$. We assume, without loss of generality, that $r_A \le r_B$.
It follows from the isodiametric inequality that $\mathrm{diam}(A) \ge 2r_A$ and $\mathrm{diam}(B) \ge 2r_B$ must be true. Let $a \in \mathrm{SW}_d(A)$ and $b \in \mathrm{SW}_d(B)$. Since these two sets are concentric balls, one can see that $|a - b| \le r_A + r_B =: r$.
Now let $(a_n)_{n\in\mathbb{N}}, (a'_n)_{n\in\mathbb{N}} \subset A$ be two sequences such that $|a_n - a'_n| \longrightarrow \mathrm{diam}(A)$. If there is some $n \in \mathbb{N}$ and a $b \in B$ such that $b \notin \overline{B_r(a_n)}$, then $|b - a_n| > r_A + r_B$, and the inequality follows.

Therefore, we assume for every $n \in \mathbb{N}$ that $B \subseteq \overline{B_r(a_n)} \cap \overline{B_r(a_n')}$. However, with some planar geometry we can see that

$$\operatorname{diam}(\overline{B_r(a_n)} \cap \overline{B_r(a_n')}) = \sqrt{r^2 - |a_n - a_n'|^2} \longrightarrow \sqrt{r^2 - \operatorname{diam}(A)^2}$$

$$\leq \sqrt{(r_A + r_B)^2 - r_A^2} \leq \sqrt{r_B^2 + 2r_A r_B} \leq \sqrt{3} r_B \leq \frac{\sqrt{3}}{2} \operatorname{diam}(B),$$

which contradicts the previous. □

We have used the isodiametric inequality in this proof. This might seem unfortunate because the following Lemma will prove that Steiner symmetrization decreases the diameter of a set, and this fact is usually used to prove the isodiametric inequality itself. If we now tried to prove the isodiametric inequality by this method, we would run into a circular argument. This could be resolved by first proving the *lemmas 1.4 and 1,5* for the case $k = 1$, where the isodiametric inequality is trivial, then proving the isodiametric inequality in all dimensions, and then finally proving both lemmas for $k > 1$.

We will now define the Lebesgue metric or Lebesgue distance. For now this only serves as a shorthand notation, but it will be of relevance as measure of distance in the second chapter.

Definition 1.5. *(Lebesgue metric) For two measurable sets of finite measure $A, B \subseteq \mathbb{R}^d$, we define the Lebesgue metric by*

$$d_L(A, B) = m^d(A \setminus B) + m^d(B \setminus A).$$

We could also write $d_L(A, B) = \|\mathbb{1}_A - \mathbb{1}_B\|_{L^1}$. Note that d_L is not a metric when we do not identify sets with a difference that has measure zero.

Now we prove some basic properties of the Steiner symmetrization for our generalized version. Statements (ii) and (iii) are not strictly necessary for the later chapters, but they are classical statements about the standard Steiner symmetrization and therefore worth generalizing. Note that, as mentioned in the introduction, we use the notation $(x, y) \in \mathbb{R}^k \times \mathbb{R}^{d-k}$ when working with the Steiner symmetrization.

Lemma 1.6. *Let $k, d \in \mathbb{N}$ such that $1 \leq k \leq d$, and let $M, N \in \mathcal{L}(\mathbb{R}^d)$ have finite measure. Then the following statements hold.*

(i) *S^k is well-defined, monotone and equimeasurable.*

(ii) $\operatorname{diam}(S^k M) \leq \operatorname{diam}(M)$.

(iii) *If M is convex, then $S^k M$ is convex.*

(iv) *If M is compact, then $S^k M$ is compact.*

(v) $d_L(S^k M, S^k N) \leq d_L(M, N)$.

Proof. We will frequently use the following equivalence for a point $(x, y) \in \mathbb{R}^k \times \mathbb{R}^{d-k}$ to show statements about the Steiner symmetrization:

$$(x, y) \in S^k(M) \iff x \in SW_k(y)$$

$$\iff \{x \in \mathbb{R}^k : (x, y) \in M\} \neq \varnothing \text{ and } \omega_k |x|^k \leq m^k(\{x \in \mathbb{R}^k : (x, y) \in M\}).$$

The last statement might look awkward, as the requirement that the set on the left side is nonempty seems redundant. After all, in this case the measure on the right would be zero anyways! If we were to drop this requirement though, then we would find that $(0, y) \in S^k(M)$. We do not want this, as then the symmetrized set contains unnecessary zero sets. For example, the symmetrization in 1 dimension of any 2-dimensional set would contain the whole y-axis.

(i) Monotonicity is clear. The function $f := \mathbb{1}_M$ is measurable because M is measurable. Then Fubini's theorem tells us that the function $\left[y \mapsto \int_{\mathbb{R}^k} f(x, y)\, dm^k(x)\right]$ is measurable. Note that

$$\int_{\mathbb{R}^k} f(x, y)\, dm^k(x) = m^k(\{x : (x, y) \in M\}).$$

Now let $(x, y) \in \mathbb{R}^d$ such that $\{x : (x, y) \in M\} \neq \varnothing$. Then

$$(x, y) \in S^k M \iff x \in SW_k(\{x : (x, y) \in M\}) \iff (x, y) \in g^{-1}([0, \infty)),$$

where $g(x, y) = \int_{\mathbb{R}^k} f(x', y)\, dm^k(x') - \omega_k |x|^k$ is a measurable function. It follows from this that $S^k M$ is measurable. Fubini's theorem also implies

$$m^d(M) = \int_{\mathbb{R}^{d-k}} \int_{\mathbb{R}^k} f(x, y)\, dm^k(x)\, dm^{d-k}(y)$$

$$= \int_{\mathbb{R}^{d-k}} m^k(\{x : (x, y) \in M\})\, dm^{d-k}(y)$$

$$= \int_{\mathbb{R}^{d-k}} m^k(\{x : (x, y) \in S^k M\})\, dm^{d-k}(y) = m^d(S^k M).$$

(ii) Let $(x, y), (x', y') \in S^k M$. Define $A = \{x'' : (x'', y) \in M\}$ and $B = \{x'' : (x'', y') \in M\}$. Then, clearly, $x \in SW_k(A)$ and $x' \in SW_k(B)$. We apply *Lemma 1.4* to A and B and get sequences $(a_n)_{n \in \mathbb{N}} \subset A$, $(b_n)_{n \in \mathbb{N}} \subset B$ such that

$$\sup_{n \in \mathbb{N}} |a_n - b_n| \geq |x - x'|.$$

Then

$$|(x, y) - (x', y')|^2 \leq \sup_{n \in \mathbb{N}} |y - y'|^2 + |a_n - b_n|^2 \leq \operatorname{diam}(M)^2,$$

since $(a_n, y), (b_n, y') \in M$ for every $n \in \mathbb{N}$.

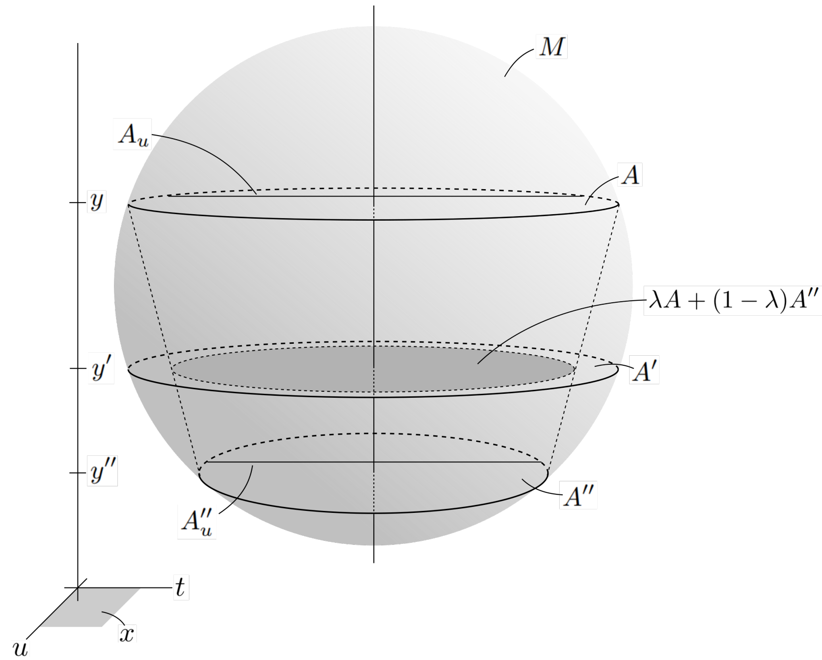


Figure 1: The variables used in the proof of (iii).

(iii) Let $y, y'' \in \{y \in \mathbb{R}^{d-k} : \exists x \in \mathbb{R}^k, (x, y) \in S^k M\}$. Let $\lambda \in (0, 1)$ and for this λ define $y' = \lambda y + (1-\lambda)y''$. Further, define $A = \{(x, y) \in M\}$ as well as A' and A'' analogously for y' and y''. Then, by convexity of M, we see that $M \supseteq \lambda A + (1 - \lambda)A''$. This implies

$$\{(x, y') \in M\} \supseteq \lambda A + (1 - \lambda)A''.$$

In particular

$$m^k(\{x : (x, y') \in M\}) \geq m^k(\{x : (x, y') \in \lambda A + (1 - \lambda)A''\}).$$

We now rewrite this as an integral over the last $k - 1$ dimensions of x.

$$= \int_{\mathbb{R}^{k-1}} m^1(\{t \in \mathbb{R} : ((t, u), y') \in \lambda A + (1 - \lambda)A''\})\, dm^{k-1}(u).$$

Now for all $u \in \mathbb{R}^{k-1}$, we define $A_u = \{t \in \mathbb{R} : ((t, u), y) \in A\}$ as well as A''_u analogously for y''. Since

$$\lambda A + (1 - \lambda)A'' \supseteq \lambda A_u \times \{u\} \times \{y\} + (1 - \lambda)A''_u \times \{u\} \times \{y''\},$$

we can continue the previous sequence of inequalities as follows:

$$\geq \int_{\mathbb{R}^{k-1}} m^1(\{t \in \mathbb{R} : (t, u) \in \lambda A_u + (1 - \lambda)A''_u\}\, dm^{k-1}(u).$$

Because M is convex the sets A and A'' are convex as well. It then follows that $A_u = (a_u, b_u)$, up to some m^1 zero set, for certain $a_u \le b_u \in \mathbb{R}$. The analogous is true for A''_u and certain numbers a''_u, b''_u. We can further transform the previous:

$$= \int_{\mathbb{R}^{k-1}} m^1(\lambda(a_u, b_u) + (1-\lambda)(a''_u, b''_u))\, dm^{k-1}(u)$$

$$= \int_{\mathbb{R}^{k-1}} \lambda m^1(a_u, b_u) + (1-\lambda) m^1(a''_u, b''_u)\, dm^{k-1}(u).$$

$$\begin{aligned} &= \lambda \int_{\mathbb{R}^{k-1}} m^1(\{t \in \mathbb{R} : (t, u) \in A_u\})\, dm^{k-1}(u) \\ &+(1-\lambda) \int_{\mathbb{R}^{k-1}} m^1(\{t \in \mathbb{R} : (t, u) \in A''_u\})\, dm^{k-1}(u) \end{aligned}$$

$$= \lambda m^k(\{x : (x, y) \in M\}) + (1-\lambda) m^k(\{x : (x, y'') \in M\}).$$

In particular, because $\left[x \to |x|^{\frac{1}{d}}\right]$ is concave, it is the case that $r' \ge \lambda r + (1-\lambda) r''$, where r is the radius of $\{x : (x, y) \in S^k M\}$ and r', r'' are defined analogously for y' and y''.

Finally, if $(x, y), (x'', y'') \in S^k M$, then $|x| \le r$ and $|x''| \le r''$, so by the above $|x'| \le r'$ where $x' = \lambda x + (1-\lambda) x''$. It follows that $(x', y') \in S^k M$.

(iv) With (ii) or a separate argument, we see that that boundedness of M implies boundedness of $S^k M$. Now let M be closed and $(x_n, y_n)_{N \in \mathbb{N}} \subseteq S^k(M)$ converge to a point $(x, y) \in \mathbb{R}^d$. M being closed implies that

$$\bigcap_{n \in \mathbb{N}} \bigcup_{m \ge n} \{x' : (x', y_m) \in M\} \subseteq \{x' : (x', y) \in M\}.$$

Applying the definition of $(x_n, y_n) \in S^k(M)$, we get that for any $n \in \mathbb{N}$ we know $\{x' : (x', y_n) \in M\} \ne \varnothing$ and

$$\begin{aligned} \omega_k |x|^k &= \liminf_{n \to \infty} \omega_k |x_n|^k \le \liminf_{n \to \infty} m^k(\{x' : (x', y_n) \in M\}) \\ &\le m^k\left(\bigcap_{n \in \mathbb{N}} \bigcup_{m \ge n} \{x' : (x', y_m) \in M\}\right) \le m^k(\{x' : (x', y) \in M\}). \end{aligned}$$

It remains to show that $\{x' : (x', y) \in M\} \ne \varnothing$. We can find an element in this set by first finding a sequence of x'_n so that $(x'_n, y_n)_{n \in \mathbb{N}} \subseteq M$, and then using compactness of M.

(v) Assume, without loss of generality, that $m^d(N) \le m^d(M)$. In the case $k = d$, i.e. for the Schwarz symmetrization, this implies

$$d_L(S^d M, S^d N) = m^d(SW_d M) - m^d(SW_d N) = m^d(M) - m^d(N) \le d_L(M, N).$$

For the case $k < d$, we use Fubini's theorem and the previous result:

$$d_L(S^k M, S^k N) = \int_{\mathbb{R}^d} |\mathbb{1}_{S^k M} - \mathbb{1}_{S^k N}|\, dm^d$$

$$= \int_{\mathbb{R}^{d-k}} \int_{\mathbb{R}^k} d_L(SW_k(\{x : (x,y) \in M\}), SW_k(\{x : (x,y) \in N\}))\, dy$$
$$\le \int_{\mathbb{R}^{d-k}} \int_{\mathbb{R}^k} d_L(\{x' : (x',y) \in M\}, \{x' : (x',y) \in N\})\, dy = d_L(M,N).$$

□

1.4 The Steiner symmetrization of a function

As described in the introduction, we now define the corresponding rearrangement of functions. Instead of defining a class of rearrangeable functions, we simply accept the cases where the rearranged function has value infinity as valid and say that the rearranged function does not exist.

Definition 1.7. *Let $\Omega \subset \mathbb{R}^d$ be open and let $f : \overline{\Omega} \longrightarrow [0,\infty)$.*
We define the Steiner Symmetrization in k dimensions of f as the function

$$\begin{aligned} S^k(f) : S^k\overline{\Omega} &\longrightarrow [0,\infty) \cup \{\infty\} \\ x &\longmapsto \sup\{c \ge 0 : x \in S^k\{f \ge c\}\}. \end{aligned}$$

We say that the Steiner symmetrization exists in x whenever the supremum is finite. As for the Steiner symmetrization of sets, we define symmetrization with respect to a k-dimensional subvector space V, and we adopt the same notations.

We have mentioned before that there are some slightly different ways in which one could define the rearrangement of a function. For example, one could rearrange each horizontal slice $\{f \ge c\}$ of the function and consider the resulting shape in $\mathbb{R}^{d+1}$ the subgraph of the new, rearranged function. This does not always work though, as the example below shows. Our new function would not be defined in the point 0. However, our *lemma 1.1* tells us that we would at least get the subgraph of f^* up to a set of measure zero, when rearranging each layer of f.

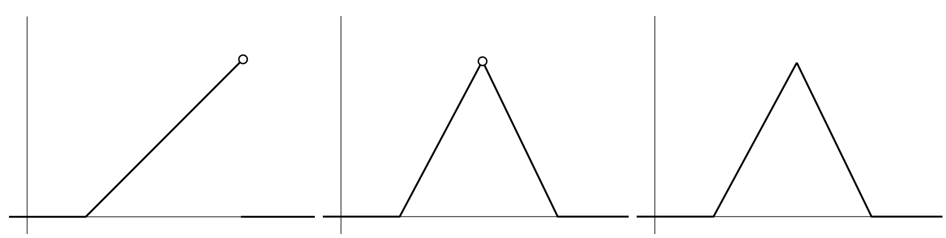


Figure 2: In order, a function f, the result of rearranging each layer, and f^*.

For our purposes, we are mostly interested in rearranging functions in L^p, so let us analyze how the Steiner symmetrization acts as a nonlinear operator on L^p.

Lemma 1.8. *Let $1 \le k \le d$ and set $* = S^k$. Let $1 \le p < \infty$ and $\Omega \subset \mathbb{R}^d$ be open. Then $* : L^p_{\ge}(\overline{\Omega}) \longrightarrow L^p_{\ge}(\overline{\Omega}^*)$ is a well-defined, continuous, and norm-preserving function.*

Proof. We have to show that the Steiner symmetrization exists almost everywhere and is zero outside of $\overline{\Omega}^*$. Let $(x, y) \in \mathbb{R}^d$ so that there exists a sequence $0 \leq a_n \nearrow \infty$ with the property that $(x, y) \in \{f \geq a_n\}^*$ for every $n \in \mathbb{N}$. Then it must be true that

$$\omega_k |x|^k \leq m^d(\{f \geq a_n\}) \searrow m^d \left(\bigcap_{n \in \mathbb{N}} \{f \geq a_n\} \right) = 0.$$

Therefore $x = 0$, and since the set of points $(0, y)$ with $y \in \mathbb{R}^{d-k}$ is a zero set, we know that the rearranged function exists almost everyhwere. Now let $x \notin \overline{\Omega}^*$. Then

$$f^*(x) = \sup\{a \geq 0 : x \in \{f \geq a\}^*\} = 0$$

because $x \notin \{f \geq 0\}^*$.
It remains to show that $*$ is continuous and preserves the norm. Let us establish a very useful fact that will help us deal with the cases $p > 1$:

$$f^{p*}(x) = \sup\{a \geq 0 : x \in \{f \geq a^{\frac{1}{p}}\}^*\} = f^{*p}(x).$$

Because the symmetrization is equimeasurable, it is clearly true that $\|f^*\|_{L^1} = \|f\|_{L^1}$. For the case $p > 1$, consider the following:

$$\|f^*\|_{L^p}^p = \int_{\Omega^*} f^{*p}\, dm^d = \int_{\Omega^*} f^{p*}\, dm^d = \int_{\Omega} f^p\, dm^d.$$

Here we used that $f^p \in L^1(\Omega)$. Again, we first show continuity for the case $p = 1$. Let $f_n \longrightarrow f$ in $L^1_{\geq}(\Omega)$. Then

$$\begin{aligned}
\int_{\Omega^*} |f_n^* - f^*|\, dm^d &= \int_{\Omega^* \cap \{f_n^* \geq f^*\}} f_n^* - f^*\, dm^d + \int_{\Omega^* \cap \{f_n^* < f^*\}} f^* - f_n^*\, dm^d \\
&= \int_0^\infty m^d(\{f_n^* \geq f^*\} \cap (\{f_n^* > t\} \setminus \{f^* > t\}))dt \\
&\quad + \int_0^\infty m^d(\{f_n^* < f^*\} \cap (\{f^* > t\} \setminus \{f_n^* > t\}))dt \\
&= \int_0^\infty d_L(\{f_n^* > t\}, \{f^* > t\})dt = \int_0^\infty d_L(\{f_n > t\}^*, \{f > t\}^*)dt \\
&\leq \int_0^\infty d_L(\{f_n > t\}, \{f > t\})dt = \ldots = \int_\Omega |f_n - f|\, dm^d.
\end{aligned}$$

Now we will deal with the case $1 < p < \infty$. Let $f_n \to f$ in $L^p(\mathbb{R}^d)$ and assume that there exists an $\epsilon > 0$ and a subsequence f_{n_k} such that $\|f_{n_k}^* - f^*\|_{L^p} \geq \epsilon$ for every $k \in \mathbb{R}^d$. We know that we can take another subsequence so that $f'_j := f_{n_{k_j}} \longrightarrow f$ pointwise almost everywhere. Note that $f^p, f_j'^p \in L^1_{\geq}(\mathbb{R}^d)$. With dominated convergence, we see that

$$\begin{aligned}
0 \longleftarrow \int_{\mathbb{R}^d} |f_j'^p - f^p|\, dm^d &\geq \int_{\mathbb{R}^d} |(f_j'^p)^* - (f^p)^*|\, dm^d \\
= \int_{\mathbb{R}^d} |(f_j'^*)^p - (f^*)^p|\, dm^d &\geq \int_{\mathbb{R}^d} |f_j'^* - f^*|^p\, dm^d \geq \epsilon.
\end{aligned}$$

This is a contradiction. □

Notably, it has been shown in [1] that the Steiner symmetrization for the case $k = 1$ is also continuous as an operator on $W^{1,p}(\mathbb{R}^d)$ for $1 \leq p < \infty$.

The following lemma shows that S^1 takes Lipschitz functions to Lipschitz functions, a fact which we will need in the proof of *theorem 3.5*. Once we have this theorem in hand, we can take the limit as p goes to infinity in the inequality $\|\nabla S^k f\|_{L^p} \leq \|\nabla f\|_{L^p}$ and see that $\|\nabla S^k f\|_{L^\infty} \leq \|\nabla f\|_{L^\infty}$. This implies the stronger statement that S^k for $k > 1$ maps Lipschitz functions to Lipschitz functions with smaller or equal constant. Showing Hölder continuity here for $k > 1$ is therefore technically superfluous but since the result naturally falls out of the proof for $k = 1$, we did not exclude it.

Lemma 1.9. *Let $1 \leq k \leq d$. Let $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ be Lipschitz continuous with Lipschitz constant $L > 0$ and compact support in B_R. Then the rearranged function $S^k(f)$ exists and is Hölder continuous with exponent $\frac{1}{k}$ and constant $2L(kR^{k-1})^{\frac{1}{k}}$. In particular $S^1(f)$ is Lipschitz continuous.*

Proof. We first define a constant $C > 0$ by $C^k := kR^{k-1}$. Now let $(x, y), (x', y') \in S^k(\overline{\Omega})$ and define $\epsilon^k := |(x, y) - (x', y')|$. We want to show that

$$|S^k(f)(x, y) - S^k(f)(x', y')| < \epsilon 2LC.$$

Assuming without loss of generality that $S^k(f)(x, y) \geq S^k(f)(x', y')$, this is equivalent to

$$\sup\{a \geq 0 : (x', y') \in S^k\{f \geq a\}\} \geq \sup\{a \geq 0 : (x, y) \in S^k\{f \geq a\}\} - \epsilon 2LC.$$

Let $a \geq 0$ so that $(x, y) \in S^k(\{f \geq a\})$. This means, by definition, that

$$\omega_k |x|^k \leq m^k(\{x'' : (x'', y) \in \{f \geq a\}\}).$$

If we can show that $(x', y') \in S^k\{f \geq a - \epsilon 2LC\}$, then the inequality that we want to prove will follow. Again, this is equivalent to

$$\omega_k |x'|^k \leq m^k(\{x'' : (x'', y') \in \{f \geq a - \epsilon 2LC\}\}).$$

For all $r > 0$ and $M \subseteq \mathbb{R}^d$, we define $M_r := \{x \in \mathbb{R}^d : \mathrm{dist}(x, M) < r\}$. It then follows from Lipschitz continuity of f that $\{f \geq a\}_r \subseteq \{f \geq a - Lr\}$. In particular, for $r = \epsilon 2C$, this implies

$$\{x'' : (x'', y') \in \{f \geq a - \epsilon 2LC\}\} \supseteq \{x'' : (x'', y') \in \{f \geq a\}_{\epsilon 2C}\}.$$

If $(x'', y) \in \{f \geq a\}_{\epsilon C}$, then $(x'', y') \in \{f \geq a\}_{\epsilon 2C}$ because $|y - y'| < \epsilon C$. Therefore

$$\{x'' : (x'', y') \in \{f \geq a\}_{\epsilon 2C}\} \supseteq \{x'' : (x'', y) \in \{f \geq a\}_{\epsilon C}\}.$$

Now note that for bounded, measurable sets $A \subseteq \mathbb{R}^k$ and $t > 0$ one can show

$$m^k(A_t) \geq m^k(A) + m^k(B_t).$$

This is done by "sandwiching" the set inbetween the hyperplanes

$$H_- := \{\inf\{x \in \mathbb{R} : \exists y \in \mathbb{R}^{k-1} : (x, y) \in A\}\} \times \mathbb{R}^{k-1}$$
$$\text{and } H_+ := \{\sup\{x \in \mathbb{R} : \exists y \in \mathbb{R}^{k-1} : (x, y) \in A\}\} \times \mathbb{R}^{k-1},$$

and finding a half ball's worth of measure contained in $A_t \setminus A$, once to the negative and once to the positive side of the hyperplanes.

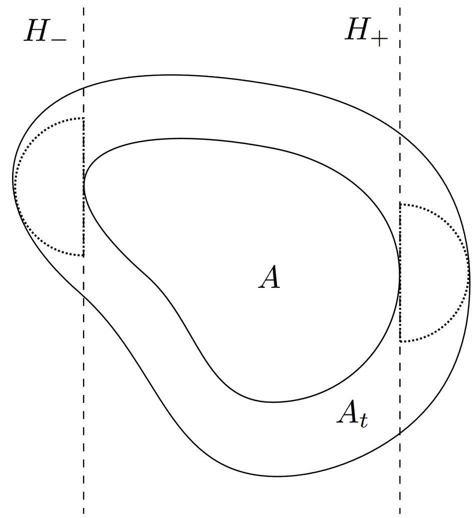


Figure 3: A sketch of the proof.

Using this together with our previous results, we can deduce the following chain of inequalities:

$$\begin{aligned} m^k(\{x'' : (x'', y') \in \{f \geq a - \epsilon 2LC\}\}) &\geq m^k(\{x'' : (x'', y) \in \{f \geq a\}_{\epsilon C}\}) \\ &\geq m^k(\{x'' : (x'', y) \in \{f \geq a\}\}) + m^k(B_{\epsilon C}) \\ &\geq \omega_k \left(|x|^k + (\epsilon C)^k\right). \end{aligned}$$

Finally, observe that

$$\left||x|^k - ||x'|^k\right| = \left|(|x| - |x|') \sum_{j=0}^{k-1} |x|^j |x'|^{k-1-j}\right| \leq \epsilon \sum_{j=0}^{k-1} R^{k-1} = \epsilon^k C^k,$$

This concludes the proof, as it implies that

$$\omega_k |x'|^k \leq m^k(\{x'' : (x'', y') \in \{f \geq a - \epsilon 2LC\}\}).$$

$\square$

2 Steiner Symmetrizations can reshape a set into a ball

Recall that our goal is to prove *theorem 3.5*, which the reader can find at the very beginning of section 1 and the end of section 3. To do this, we will first consider the case $k = 1$. Here the argument will be essentially one-dimensional. Then we will generalize to the cases $k > 1$ by approximating the Steiner symmetrization in k dimensions with the symmetrization in 1 dimension.

For this, we will show that for any compact set, one can find a sequence of directions in $\{0\}^k \times \mathbb{R}^{d-k}$ so that the result, after symmetrizing in these directions, converges to a ball. Since the Steiner symmetrization in k-dimensions essentially applies the Schwarz symmetrization in k-dimensional space to affine k-dimensional hyperplanes, this will enable us to perform the previously mentioned approximation.

We will use compactness of a certain space of sets, which contains our sequence of Steiner symmetrizated sets, to find a limit. The obvious choice for a notion of distance would be the L^1 metric on indicator functions, which we have given the name "Lebesgue metric" in the first chapter.

In this case, we might be able to use the Fréchet–Kolmogorov theorem, also known as the Kolmogorov compactness theorem, to find that our space of sets is compact in this metric. The theorem requires that this space, in a certain sense, is both bounded and has limited variation or oscillation.

We will, however, follow a different approach and use the Hausdorff metric as our measure of distance. Under this metric, the space of compact subsets of a compact set can be quite naturally seen to be compact. The proof of *theorem 2.13*, the main result of this section, is based on a proof by Federer in [3] (theorem 2.10.31, page 195). He proves that one can approximate a ball through Steiner symmetrizations in the Hausdorff metric, and he does this with a compactness argument. Since we want to end up with an approximation in the Lebesgue metric, we have to first do some additional work and can then emulate his proof.

2.1 The Hausdorff metric

Definition 2.1 (Hausdorff metric). *Let (X, d) be a metric space. For $\varnothing \neq A, B \subseteq X$, we define the Hausdorff metric by*

$$d_H(A, B) = \max\left\{\sup_{a\in A} \operatorname{dist}(a, B), \sup_{b\in B} \operatorname{dist}(b, A)\right\}.$$

We also define $C(X)$ to be the compact sets in X.

Lemma 2.2. *For any metric space (X, d) the following holds:*

(i) $(C(X), d_H)$ *is a metric space.*

(ii) If (X, d) *is compact, then* $(C(X), d_H)$ *is compact.*

Proof. (i) Let $A, B \in C(X)$. The sets being bounded implies $d_H(A, B) < \infty$. Symmetry directly follows from the symmetry in the definition. Using that the sets are closed, we see that

$$d_H(A,B) = 0 \iff dist(a,B) = 0 = dist(b,A) \quad \forall a \in A\, \forall b \in B \iff A = B.$$

Now let $C \in \mathcal{C}(X)$. Then for any $b \in B$,

$$\sup_{a\in A} \operatorname{dist}(a,C) = \sup_{a\in A} \inf_{c\in C} d(a,c) \leq \sup_{a\in A} \inf_{c\in C} d(a,b) + d(b,c).$$

It follows that

$$\sup_{a\in A} \operatorname{dist}(a,C) \leq \sup_{a\in A} \inf_{c\in C} \inf_{b\in B} d(a,b) + d(b,c)$$

$$= \sup_{a\in A} \left(\inf_{b\in B} d(a,b) + \inf_{c\in C} d(b,c) \right) \leq d_H(A,B) + d_H(B,C).$$

We can do the same with A and C being exchanged and arrive at the triangle inequality.

(ii) Let (X,d) be compact. We will show that $\mathcal{C}(X)$ is totally bounded and complete. Let $\epsilon > 0$ and use compactness of X to find $x_1, ..., x_n \in X$ with $X \subseteq \bigcup_{i=1}^n B_\epsilon(x_i)$. Define $\mathcal{J} := \mathcal{P}(\{1,...,n\})$ and, for $I \in \mathcal{J}$, define $M_I := \bigcup_{i\in I} B_\epsilon(x_i)$. Now let $A \in \mathcal{C}(X)$ and consider the set

$$I := \{i \in \{1,...,n\} : A \cap B_{r_i}(x_i) \neq \varnothing\}.$$

We want to show that $d_H(A, M_I) < \epsilon$. Observe that

$$\sup_{a\in A} \operatorname{dist}(a, M_I) = 0 < 2\epsilon$$

because A is covered by the balls $B_\epsilon(x_i)$. Furthermore,

$$\sup_{m\in M_I} \operatorname{dist}(m,A) = \max_{i\in I} \sup_{m\in B_\epsilon(x_i)} \operatorname{dist}(m,A) < 2\epsilon.$$

Here we used that every one of the balls must contain an element of A. We have shown that for an arbitrary $\epsilon > 0$, the set $\mathcal{C}(X)$ is covered by the open 2ϵ balls with centers $\{M_I : I \in \mathcal{J}\}$.Therefore $\mathcal{C}(X)$ being totally bounded.
For completeness, let $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{C}(X)$ be a Cauchy sequence and define

$$A := \{a \in X : \text{ there exists a sequence } a_n \in A_n \text{ with } a_n \to a\}.$$

Now let $\epsilon > 0$ and choose an $N \in \mathbb{N}$ such that $d_H(A_k, A_n) < \frac{1}{2}\epsilon$ for all $k \geq n \geq N$. Let $a \in A$ and let a_n be the sequence from the definition. If $|a - a_n| \geq \epsilon$ for some $n \geq N$, then there must exist a $k > n$ such that $d_H(A_k, A_n) \geq |a_k - a_n| > \frac{1}{2}\epsilon$, which contradicts the previous. Therefore $\operatorname{dist}(a, A_n) < |a - a_n| < \epsilon$.
Now let $n \geq N$ and $a_n \in A_n$. For every $k \geq n$, there must exist an $a_k \in A_k$ such that $|a_k - a_n| < \epsilon$. By compactness of $\overline{B_\epsilon(a_n)}$ and the definition of A, this sequence must converge to a limit $a \in A$. Then $|a - a_n| \leq \epsilon$, so $\operatorname{dist}(a_n, A) \leq \epsilon$.
In summary, $d_H(A, A_n) \leq \epsilon$. It is easy to show that A is closed by taking a diagonal sequence. Therefore $A \in \mathcal{C}(X)$ and $A_n \longrightarrow A$ in $(\mathcal{C}(X), d_H)$, so we know that $(\mathcal{C}(X), d_H)$ is totally bounded and complete.

□

We have now gained a notion of compactness with less effort than the Kolmogorov compactness theorem would have required. There was a price to pay though, as convergence in the Hausdorff metric does not imply convergence in the Lebesgue metric. In fact, any bounded set can, in the Hausdorff metric, be arbitrarily approximated by a finite set of points which clearly have measure zero. This is a problem as our goal is to have a sequence of symmetrized sets converge to a ball in the *Lebesgue* metric.
To solve this, we have to introduce a notion of bounded variation or regularity for our sets. A requirement for the Kolmogorov compactness theorem is that for C, some space of functions that we want to be compact, the following is true:

$$\|f(\cdot - \delta) - f(\cdot)\|_{L^p} \overset{\delta \to 0}{\longrightarrow} 0 \text{ uniformly on } C.$$

In the special case of our situation f would be the indicator function of a set and then the L^1 distance is the Lebesgue distance.
On a heuristic level we realize that if all of our sets were actually convex and subsets of some bounded set, then the above requirement should be fulfilled. At the same time, if the Hausdorff distance between two sets is small and those sets are convex, then their Lebesgue distance ought to be small too! It seems like convexity is the right way of limiting variation to gain convergence in the Lebesgue metric from convergence in the Hausdorff metric. Furthermore, we have shown before that Steiner symmetrization maps convex sets into convex sets!
Unfortunately, we do not know if we can turn any set into a convex set via finitely many Steiner symmetrizations. Actually the author could not even find an answer to this question elsewhere and would like, not claiming originality, to formulate it as an open problem.

Open problem 2.3. *For $1 \leq k < d \in \mathbb{N}$, can every compact set in $\mathbb{R}^d$ be turned into a convex set via finitely many Steiner symmetrizations in k dimensions? If yes, is the number of necessary symmetrizations bounded by a constant $C(k, d)$?*

We can avoid dealing with the difficulty of this question by introducing a different property of sets, which can be seen as a weakened version of convexity. It is easy to achieve with Steiner symmetrizations and also yields the necessary "boundedness of variation".

2.2 The symmetric cuboid property and the space $Q(X)$

Definition 2.4 (symmetric cuboid property)**.** *We say that a set $M \subseteq \mathbb{R}^d$ has the symmetric cuboid property if $x \in M$ implies*

$$\operatorname{rect}(x) := \big[-|x_1|, |x_1|\big] \times \ldots \times \big[-|x_d|, |x_d|\big] \subseteq M.$$

Alternatively,

$$M = \bigcup_{x \in M} \operatorname{rect}(x).$$

For a set $X \subseteq \mathbb{R}^d$, we define

$$Q(X) := \{M \in C(X) : M \text{ has the symmetric cuboid property }\}.$$

The following two lemmas might be of interest, although the rest of the chapter does not strictly depend on them. We can extend the statement that the boundary of a convex set is always a zero set to sets with the symmetric cuboid property. Then we see that these sets are unions of only countable many cuboids up to a set of measure zero, which is useful whenever one wants to estimate the measure of such a union of cuboids by a sum.

Lemma 2.5. *Let $X \subseteq \mathbb{R}^d$ and $M \in Q(X)$. Then ∂M is a zero set, and the interior of M is the union of only countably many symmetric cuboids.*

Proof. The proof is pretty much the same as for convex sets. Define $Y = \{x \in X : \exists 1 \leq i \leq d, |x_i| = 0\}$ and note that it is a zero set. Now let $x \in \partial M \setminus Y$ and $0 < \lambda < 1$. If $x \in \lambda M$, then there must exist a $z \in M$ such that $x \in \mathrm{rect}(\lambda z)$. Since $x \notin Y$, it can't be the case that $z \in Y$.
This means that $\mathrm{rect}(\lambda z)$ is not a degenerate cuboid (nonzero measure), and so x is in the interior of $\mathrm{rect}(z)$. Then $x \in M^\circ$, which contradicts $x \in \partial M$. Therefore $x \notin \lambda M$.
At the same time $x \in \mathrm{rect}(\lambda^{-1}x)$, so $x \in \lambda^{-1}M$. We see that

$$m^d(\partial M) = m^d(\partial M \setminus Y) \leq m^d(\lambda^{-1}M \setminus \lambda M) = (\lambda^{-d} - \lambda^d)m^d(M) \xrightarrow{\lambda \to 1} 0.$$

Lastly, if $B_r(x) \subseteq M$, then we can find a $q \in \mathbb{Q}^d \cap B_r(x)$ such that $x \in \mathrm{rect}(q)$. This implies

$$M^\circ = \bigcup_{x \in M^\circ} \mathrm{rect}(x) = \bigcup_{x \in \mathbb{Q}^d \cap M^\circ} \mathrm{rect}(x).$$

□

Lemma 2.6. *Let $X \subseteq \mathbb{R}^d$ be compact. Then $(Q(X), d_H)$ is complete.*

Proof. Because $(C(X), d_H)$ is complete and contains $Q(X)$, we only have to show that $M_n \longrightarrow M$ implies $M \in Q(X)$ for any sequence $(M_n)_{n \in \mathbb{N}} \subseteq Q(X)$ and any set $M \in C(X)$. Let $x \in M$ and $y \in \mathrm{rect}(x)$. Convergence in the Hausdroff norm implies that there exists a sequence of $x_n \in M_n$ with $x_n \longrightarrow x$. Then $d_H(\mathrm{rect}(x_n), \mathrm{rect}(x)) \longrightarrow 0$, so y must be the limit of a sequence of $y_n \in \mathrm{rect}(x_n)$. In particular $y_n \in M_n$, which by our characterization of the limit in the Hausdorff metric given in the proof of *Lemma 2.2* implies $y \in M$. Therefore M has the symmetric cuboid property. □

Let us now show that it is easy to turn a set into one that has the symmetric cuboid property by applying d Steiner symmetrizations in 1 dimension.

Lemma 2.7. *Let $X \subseteq \mathbb{R}^d$ be bounded, $M \in C(X)$ and $S_1, ..., S_d$ be the Steiner symmetrizations with respect to the coordinate axes. Then $S_d...S_1M \in Q(X)$.*

Proof. Step 1. We inductively show that $M_k = S_k...S_1M$ for $1 \leq k \leq d$ is symmetric in axes 1 to k. Let a $1 \leq i \leq k$ be fixed. If $i = k$, then we are done because the last symmetrization was in direction k, and so the result is symmetric in direction i.
Now let $1 \leq i < k$. For ease of notation we prove the case $i = k - 1$, but the proofs for the other cases are identical. Let $(x, y, z) \in \mathbb{R}^{k-2} \times \mathbb{R} \times \mathbb{R}$ so that $(x, y, z) \in M_k$. We have to show $(x, -y, z) \in M_k$. Because $(x, y, z) \in S_k(M_{k-1})$, we know that both $\{y' : (x, y', z) \in M_{k-1}\} \neq \varnothing$ and

$$|z| \leq m^1(\{z' : (x, y, z') \in M_{k-1}\}).$$

By the inductive assumption, the set M_{k-1} is symmetric in axis i, so

$$\{z' : (x, y, z') \in M_{k-1}\} = \{z' : (x, -y, z') \in M_{k-1}\}.$$

Therefore $\{z' : (x, -y, z') \in M_{k-1}\} \neq \emptyset$ and

$$|z| \leq m^1(\{z' : (x, -y, z') \in M_{k-1}\}),$$

which is the definition of $(x, -y, z) \in M_k$.

Step 2. Now we inductively show that M_k for $1 \leq k \leq d$ is convex in axes 1 to k. Let a $1 \leq i \leq k$ be fixed. If $i = k$, then, as before, we are done because the last symmetrization was in direction k, so the set is convex in direction i.
Now let $1 \leq i < k$ and assume again that $i = k - 1$. With the same notation as before, let $(x, y, z), (x, y'', z) \in M_k$ and set $y' := \lambda y + (1 - \lambda)y''$ for some $\lambda \in (0, 1)$.
We have to show $(x, y', z) \in M_k$. This is true if $\{z' : (x, y', z') \in M_k\} \neq \emptyset$ and

$$|z| \leq m^1(\{z' : (x, y', z') \in M_k\}).$$

By the inductive assumption and Step 1, M_{k-1} is convex and symmetric in direction i. This implies that for any $z' \in \mathbb{R}$,

$$(x, y, z'), (x, y'', z') \in M_{k-1} \Longrightarrow (x, y', z') \in M_{k-1}.$$

Assuming without loss of generality that $|y| \leq |y''|$, the above implies

$$\{z' : (x, y', z') \in M_{k-1}\} \supseteq \{z' : (x, y, z'), (x, y'', z') \in M_{k-1}\}$$

$$\supseteq \{z' : (x, y'', z') \in M_{k-1}\}.$$

However, since $(x, y'', z) \in S_k(M_{k-1})$, we know that $\{z' : (x, y''z') \in M_{k-1}\} \neq \emptyset$ and

$$|z| \leq m^1(\{z' : (x, y'', z') \in M_{k-1}\}) \leq m^1(\{z' : (x, y', z') \in M_k\}).$$

Step 3. Let $x \in M'$. By symmetry in every axis
$\{(\sigma_1 x_1, ..., \sigma_d x_d) : \sigma_i \in \{1, -1\}, 1 \leq i \leq d\} \subseteq M'$. Since a set that is both convex in every direction and also contains the vertices of a cuboid must contain the whole cuboid, we know that $\text{rect}(x) \subseteq M'$. □

We now prove two lemmas that connect the Hausdorff metric to the Lebesgue metric on our space $Q(X)$. Note that the Lebesgue metric is not a metric on this space because we have not identified zero sets.

Lemma 2.8. *Let $X \subset B_R \subset \mathbb{R}^d$ and $M, M' \in Q(X)$. Then*

$$d_L(M, M') \leq d2^d R d_H(M, M')^{d-1}.$$

Corollary 2.9. *For a bounded set $X \subseteq \mathbb{R}^d$, convergence in $(Q(X), d_H)$ implies convergence in $(Q(X), d_L)$.*

Proof. For all $\delta > 0$, define

$$\phi_\delta(x) = x + \delta \cdot (\operatorname{sign}(x_1), ..., \operatorname{sign}(x_d)) \quad \forall x \in \mathbb{R}^d.$$

Here we choose $\operatorname{sign}(0) = 0$, but for our purposes it will not matter as these cases make up only a set of measure zero. Note that ϕ_δ is a translation on each quadrant, so it does not change the measure of a set. For an $M \in \mathcal{Q}(X)$, we also define

$$\Phi_\delta(M) = \bigcup_{x \in M} \operatorname{rect}(\phi_\delta(x)).$$

Beware that $\Phi_\delta(M) \neq \phi_\delta(M)$ in general. The figure should illustrate the difference.

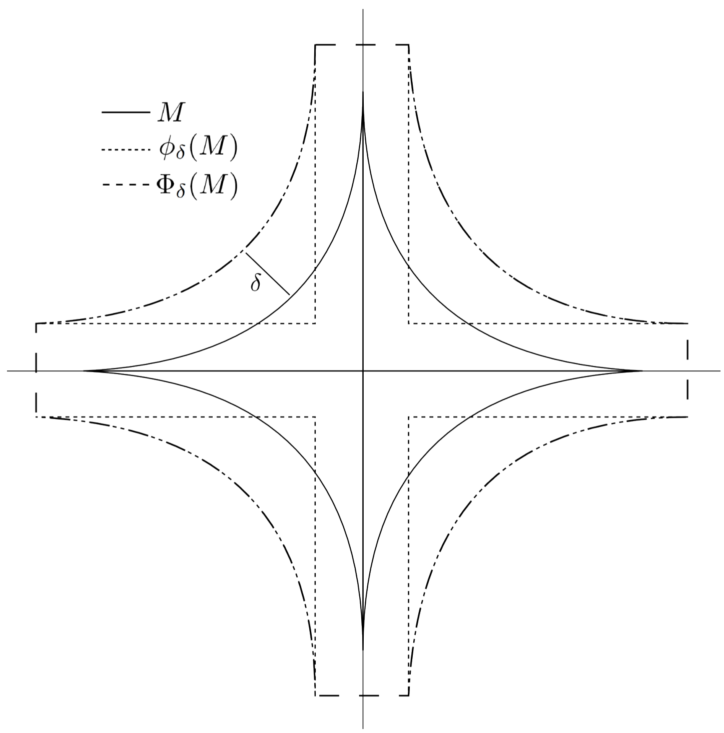


Figure 4: An illustration of ϕ and Φ.

To elaborate on this, define $Y_\delta := \{x \in X : \exists 1 \leq i \leq d, |x_i| < \delta\}$. Then

$$\operatorname{rect}(\phi_\delta(x)) \setminus Y_\delta = \phi_\delta(\operatorname{rect}(x)) \text{ up to a set of measure zero .}$$

A direct consequence of the definition is $M \subset \Phi_\delta(M)$. On the other hand, the following is true up to a set of measure zero:

$$\Phi_\delta(M) \setminus Y_\delta = \bigcup_{x \in M} \phi_\delta(\operatorname{rect}(x)) = \phi_\delta(M).$$

We arrive at the inequality

$$m^d(\Phi_\delta(M) \setminus M) = m^d(\Phi_\delta(M)) - m^d(M) \leq m^d(Y_\delta) < d2^d R\delta^{d-1}.$$

Now let $d_H(M, M') < \delta$ for $M, M' \in Q(X)$. It follows from the definition of the Hausdorff metric that

$$M \subseteq \bigcup_{x' \in M'} B_\delta(x') \text{ and } M' \subseteq \bigcup_{x \in M} B_\delta(x).$$

If $x \in M$, then there must exist some $y \in M'$ such that $|x - y| < \delta$. Because M' is symmetric in each axis, we can assume without loss of generality that y is in the same quadrant of $\mathbb{R}^d$. To simplify notationn we now assume that this is the first quadrant, so $x_i, y_i > 0$ for all $1 \leq i \leq d$. The arguments work in the same way for the other quadrants, and the cases where a coordinate of x or y is zero make up only a zero set.
Let $1 \leq i \leq d$. Then $|x - y| < \delta$ implies $|x_i - y_i| < \delta$, so $\phi_\delta(y)_i = y_i + \delta > x_i$. As a result, $x \in \operatorname{rect}(\phi_\delta(y)) \subseteq \Phi_\delta(M')$. Therefore $M \subseteq \Phi_\delta(M')$.
With the inequalities we previously derived, we get that

$$m^d(M \setminus M') \leq m^d(\Phi_\delta(M') \setminus M') < d2^d R\delta^{d-1}.$$

We can make the analogous argument with M and M' being interchanged, so we also know that $m^d(M' \setminus M) < d2^d R^{d-1}\delta$. In summary:

$$d_L(M, M') < d2^d R\delta^{d-1}.$$

The proof is now finished because δ was arbitrary. □

Unfortunately, we can not get a reverse bound of the Hausdorff distance by the Lebesgue distance. The problematic sets here are those that have thin spikes along the coordinate axes (these are the only directions in which a set with the symmetric cuboid property can have arbitrarily thin spikes). Then the Hausdorff distance between two sets can be large, due to the presence of such spikes, while the Lebesgue distance remains small. To solve this problem, we introduce a form of weak convergence in the Hausdorff norm, which ignores these spikes while still being strong enough to ultimately prove the convergence of a set to a ball through Steiner symmetrizations.

Definition 2.10. *Let $X \subseteq \mathbb{R}^d$, and recall the definition*

$$Y_t = \{x \in X : \exists 1 \leq i \leq d, |x_i| < t\}.$$

Let $M_n, M \in Q(X)$ for all $n \in \mathbb{N}$. We say that M_n converges weakly in $(Q(X), d_H)$ to M if, for every $\epsilon > 0$ and $t > 0$, there exists an $N \in \mathbb{N}$ such that

$$d_H(M_n \setminus Y_t, M \setminus Y_t) < \epsilon \text{ for all } n \geq N.$$

Lemma 2.11. *For a bounded set $X \subseteq \mathbb{R}^d$, convergence in $(Q(X), d_L)$ is equivalent to weak convergence in $(Q(X), d_H)$.*

Proof. First, we show the direction "$\Longrightarrow$".
Let $M_n, M \in Q(X)$ for all $n \in \mathbb{N}$ such that $d_L(M_n, M) \longrightarrow 0$. We assume that weak convergence in $(Q(X), d_H)$ fails. Then there must exist an $\epsilon > 0$ and a $t > 0$ such that $d_H(M_n \setminus Y_t, M \setminus Y_t) \geq 2\epsilon$ infinitely often. Choose an $N \in \mathbb{N}$ such that for all $n \geq N$,

$$d_L(M_n, M) < \frac{\min\{m^d(B_t(x)), m^d(B_\epsilon(x))\}}{2^d}.$$

There must exist such an $n \geq N$ where $d_H(M_n \setminus Y_t, M \setminus Y) \geq 2\epsilon$. By the definition of the Hausdorff metric this implies that there exists either a point $x \in M_n \setminus Y_t$ such that $B_\epsilon(x) \cap M = \emptyset$, or a point $x \in M \setminus Y_t$ such that $B_\epsilon(x) \cap M_n = \emptyset$. Note that $x \notin Y_t$ implies $B_t(x)$ is fully contained in the quadrant of $\mathbb{R}^d$ which contains x.
Consider the first case, where $x \in M_n$. If $\epsilon \leq t$, then $B_\epsilon(x)$ is also fully contained in the quadrant that x is in. Since $M_n \supseteq \mathrm{rect}(x)$, we can geometrically see

$$m^d(B_\epsilon(x) \cap M_n) \geq \frac{m^d(B_\epsilon(x))}{2^d}.$$

On the other hand, if $\epsilon > t$, then

$$m^d(B_\epsilon(x) \cap M_n) \geq \frac{m^d(B_t(x))}{2^d}.$$

For the case $x \in M$ the same argument holds. Therefore

$$d_L(M_n, M) \geq \frac{\min\{m^d(B_t(x)), m^d(B_\epsilon(x))\}}{2^d},$$

which is a contradiction.

Now we will show the direction "$\Longleftarrow$".
Let M_n converge weakly to M in $(Q(X), d_H)$ and let $\epsilon > 0$. Using *lemma 2.8*, we can find a $\delta > 0$ such that $d_H(A, B) < \delta$ implies $d_L(A, B) < \frac{1}{3}\epsilon$ for any two sets $A, B \in Q(X)$. Now pick a $t > 0$, so that $m^d(Y_t) < \frac{1}{3}\epsilon$. We can then find an $N \in \mathbb{N}$ so that for all $n \geq N$ the inequality $d_H(M_n \setminus Y_t, M \setminus Y_t) < \delta$ holds. We define

$$M_n' = \bigcup_{x \in M_n \setminus Y_t} \mathrm{rect}(x) \text{ and } M' = \bigcup_{x \in M \setminus Y_t} \mathrm{rect}(x).$$

This ensures that $M_n', M' \in Q(x)$. Note that $M_n \setminus M_n' \subseteq Y_t$ and the same is true for M and M'. It is also the case that

$$d_H(M_n', M') \leq d_H(M_n \setminus Y_t, M \setminus Y_t) =: r.$$

To prove this, we assume that $d_H(M_n', M') > r$. Then, after possibly but without loss of generality swapping M and M_n, there must exist an $x \in M_n'$ such that $|x - y| > r$ for every $y \in M'$.
If $x \notin Y_t$ then $M_n' \subseteq M_n$ implies $x \in M_n \setminus Y_t$. We can then find a $y \in M \setminus Y_t$ with $|x - y| \leq r$. Since $M \setminus Y_t = M' \setminus Y_t$ we have that $y \in M'$, a contradiction.
The remaining case is $x \in Y_t$. We assume without loss of generality that x is in the closed first quadrant, so $x_i \geq 0$ for all $1 \leq i \leq d$. This is possible because all involved sets are symmetric in each axis. We also know that there exists some $1 \leq i \leq d$ with $x_i < t$.
Now we define x' by setting $x_j' = x_j$ for all $i \neq j$ and $x_i' = t$. It is still true that $x' \in M_n'$ because the set is a union of cuboids $\mathrm{rect}(a)$ where $a \notin Y_t$, in particular $a_i \geq t$.
The point x' also still has the property that $|x' - y| > r$ for every $y \in M'$. This is true because if we could find some $y \in M' \setminus Y_t$ with $|x' - y| \leq r$, then by defining y' with $y_j' = y_j$ for all $j \neq i$ and $y_j' = x_i$, we would get a point which fulfills $|x' - y'| \leq |x' - y| \leq r$, but also $y' \in M'$. This would contradict what we know about x.

We now repeatedly replace x by x' until $x \in M_n' \backslash Y_t$. Then $x \notin Y_t$ and we get a contradiction as before. Now we know that

$$d_H(M_n', M') \leq d_H(M_n \setminus Y_t, M \setminus Y_t) < \delta.$$

By the way we chose δ this implies $d_L(M_n', M') < \frac{1}{3}\epsilon$ and we can conclude

$$d_L(M_n, M) \leq d_L(M_n', M') + 2m^d(Y_t) \leq d_H(M_n', M') + \frac{2}{3}\epsilon < \frac{2}{3}\epsilon + \frac{1}{3}\epsilon = \epsilon.$$

□

2.3 Convergence through Steiner symmetrizations

We are now ready to prove that any compact set can be reshaped into a closed ball of the same measure by applying Steiner symmetrizations in 1 dimension. The following theorem and proof are based on a theorem by Federer [**4**] (theorem 2.10.31, page 195) but have significant differences as he shows convergence in d_H, while we show it in d_L. The similarity is that we use compactness in the Hausdorff metric as well as the argument that will appear in *lemma 2.13*.

Theorem 2.12 (based on [3])**.** *Let $M \subseteq \mathbb{R}^d$ be compact and let $\epsilon > 0$. Let $R > 0$ such that $m^d(B_R) = m^d(M)$. Then there exist directions $v_1, ..., v_n$ such that*

$$d_L(S_{v_n} ... S_{v_1} M, \overline{B_R}) < \epsilon.$$

Proof. We let $X \subset \mathbb{R}$ be some large compact set whose interior contains M. Then we define $\mathcal{S}$ as the set of symmetrizations of the form

$$\mathcal{S} = \{S_{e_d} \circ ... \circ S_{e_1} \circ S_{v_m} ... \circ S_{v_1} : v_1, ..., v_m \in \mathbb{R}^d, m \in \mathbb{N}\}.$$

Here the vectors e_i are the canonical basis. Symmetrizing in all directions e_i at the end ensures that $S(A) \in Q(X)$ for any $A \in C(X)$ and $S \in \mathcal{S}$ (*lemma 2.7*). We define $\overline{\mathcal{S}(M)}$ as the closure under weak convergence in $(Q(X), d_H)$ of $\mathcal{S}(M)$, the sets that can be reached by symmetrizations in $\mathcal{S}$. Now we define

$$r := \inf\{r' > 0 : \overline{\mathcal{S}(M)} \text{ contains a subset of } \overline{B_{r'}}\}.$$

We want to show that $r = R$. If $r < R$, then there exists a sequence $(M_n)_{n \in \mathbb{N}} \subset \mathcal{S}(M)$ and a set $A \in Q(X)$ such that A is contained in $\overline{B_r}$ and M_n converges weakly in $(Q(X), d_H)$ to A. Then M_n converges in d_L to A, so we get the following contradiction:

$$m^d(M) = m^d(M_n) \longrightarrow m^d(A) \leq m^d(B_r) < m^d(B_R).$$

Therefore $r \geq R$. Let $r_n \searrow r$. Then, by definition of the infimum, for every $n \in \mathbb{N}$ there exists a sequence $(M_{n,k})_{k \in \mathbb{N}} \subset \mathcal{S}(\mathcal{M})$ and a set $A_n \in Q(X)$ such that A_n is contained in $\overline{B_{r_n}}$ and $M_{n,k} \overset{k\to\infty}{\longrightarrow} A_n$ weakly in $(Q(X), d_H)$.
Now we will define a certain diagonal sequence M_n. Let $n \in \mathbb{N}$ and set $t_n = \frac{1}{n}$. By the definition of weak convergence there exists a $K_n \in \mathbb{N}$ such that $k \geq K_n$ implies $d_H(M_{n,k} \setminus Y_{t_n}, A_n \setminus Y_{t_n}) < r_n - r$.
For a technical reason we also require that $K_n \geq K_{n-1}$. We then define $M_n := M_{n,K_n}$. What we gain from this is that, for a given $\epsilon > 0$ and $t > 0$, we can choose a large $N \in \mathbb{N}$ such that for any $n \geq N$ the following is true:

- $t_n < t$
- $d_H(M_n \setminus Y_{t_n}, A_n \setminus Y_{t_n}) < r_n - r < \epsilon$.

Because $A_n \subseteq \overline{B_{r_n}}$, this implies $M_n \setminus Y_t \subseteq M_n \setminus Y_{t_n} \subseteq \overline{B_{r+\epsilon}}$.
Our new sequence M_n is contained in the compact space $(Q(X), d_H)$, so there exists a (strongly in d_H) convergent subsequence. We replace M_n by that subsequence and call the limit $E \in Q(X)$. Note that *corollary 2.9* implies that M_n also converges to E in $(Q(X), d_L)$, and that by applying further Steiner symmetrizations this does not change. We will now show that $E \subseteq \overline{B_r}$.
To prove this, assume that there exists a point $x \in E \setminus \overline{B_r}$ which does not lie on any of the axes. In other words, there exists a $t > 0$ such that $x \notin Y_t$. Then we define

$$\epsilon := \frac{1}{2}(\min\{|x| - |r|, t\}) > 0.$$

Because of the way we chose the sequence M_n, we can, as explained above, find an $N \in \mathbb{N}$ such that $n \geq N$ implies $M_n \setminus Y_{t-\epsilon} \subseteq \overline{B_{r+\epsilon}}$. By convergence in d_H we can choose an even larger n so that $d_H(M_n, E) < \frac{1}{2}\epsilon$. Then it must be the case that there exists a point $y \in M_n$ with $|x - y| < \epsilon$.
We see, however, that then $y \notin \overline{B_{r+\epsilon}}$ because $|x| > r + 2\epsilon$. It also follows that $y \notin Y_{t-\epsilon}$, because $x \notin Y_t$. This contradicts $M_n \setminus Y_{t-\epsilon} \subseteq \overline{B_{r+\epsilon}}$. Therefore $E \subseteq \overline{B_r}$ except for some zero set which lies on the coordinate axes. It is easy to show that after a few Steiner symmetrizations $T \in \mathcal{S}$, which are not in the directions of these coordinates axes, the result TE is fully contained in $\overline{B_r}$.
If $r = R$, then we use the fact that convergence in $(Q(X), d_H)$ implies convergence in $(Q(X), d_L)$ to see $m^d(TE) = m^d(M_n) = m^d(B_R)$, so $TE = \overline{B_R}$ up to a set of measure zero. Because TE is closed this implies $TE = \overline{B_R}$. Then, by continuity of Steiner symmetrization in d_L, we get that $(T(M_n))_{n\in\mathbb{N}} \subseteq \mathcal{S}(M)$ is a sequence that converges to $\overline{B_R}$ in $(Q(X), d_L)$. This finishes the proof for this case.
Lastly, we have to treat the case $r > R$. For improved readability I have exported the proof of the following statement to a lemma below this proof: We can show that TE being closed and $m^d(TE) < m^d(B_r)$ implies that we can find a combination of Steiner symmetrizations $T' \in \mathcal{S}$ such that $T'TE \subseteq \overline{B_c}$ for some $c < r$.
Then the sequence $(T'TM_n)_{n\in\mathbb{N}}$ converges in $(Q(X), d_L)$ and equivalently weakly in $(Q(X), d_H)$ to $T'TE$. We have shown that $\overline{\mathcal{S}(M)}$ contains $T'TE$, which is a subset of $\overline{B_c}$. This is a contradiction to the definition of r. □

Lemma 2.13. *Let $M \subseteq \overline{B_r}$ be closed and $m^d(M) < m^d(B_r)$. Then there exist directions $v_1, ..., v_n \in \mathbb{R}^d$ such that $S_{v_n}...S_{v_1}M \subseteq \overline{B_c}$ for some $c < r$.*

Proof. Since $m^d(M) < m^d(B_r)$ and M is closed there must exist a point $x \in \overline{B_r}$ and a radius $\delta > 0$ such that $B_\delta(x) \cap M = \varnothing$. Because we can assume without loss of generality that M is the result of a Steiner symmetrization, there also exists an $x_1 \in \partial B_r$ with $B_\delta(x_1) \cap M = \varnothing$. Now we use compactness of ∂B_r to find pairwise different points $x_2, ..., x_n$ such that

$$\partial B_r \subset \bigcup_{i=1}^{n} B_\delta(x_i).$$

We set $v_i = \frac{x_{i+1}-x_i}{|x_{i+1}-x_i|}$ for $1 \le i \le n-1$. Since $B_\delta(x_1) \cap M = \varnothing$ one can show that for every $y \in B_\delta(x_1) \cap \partial B_r$

$$m^d(\{t \in \mathbb{R} : y + tv_1 \in M\}) < m^d(\{t \in \mathbb{R} : y + tv_1 \in B_r\}).$$

This is sufficient to show that $S_{v_1}(M) \cap B_\delta(x_i) \cap \partial B_r = \varnothing$ for both $i = 1$ and $i = 2$.
Now we inductively repeat this argument. Note that if $A \cap B_\delta(x) \cap \partial B_r = \varnothing$ for some closed $A \subseteq \overline{B_r}$ and $x \in \partial B_r$, then after applying any symmetrization S_v, it is still true that $S_v(A) \cap B_\delta(x) \cap \partial B_r = \varnothing$. This means that sets $B_\delta(x_i) \cap \partial B_r$, which have in the process of induction already been shown to be disjoint of $S_{v_k} ... S_{v_1} M$, will not contain any points of $S_{v_k+1} S_{v_k} ... S_{v_1} M$. After n steps of this process we know that $S_{v_n} ... S_{v_1} M \cap \partial B_r = \varnothing$. Since the set is closed, this implies that it is contained in some strictly smaller ball. □

With this important theorem in hand, we can now show that the Steiner symmetrization in k dimensions of a set can be approximated by applying Steiner symmetrizations in 1 dimension. The proof is conceptually easy but technical, if we want to avoid somewhat handwavey arguments.
Picture a set in three dimensions that consists of finitely many layers of small height, stacked on top of one another in the z-direction. Using the previous theorem, we can approximately reshape the bottom layer into a circle. Then we can reshape the second layer into a circle. This will change the first layer again, but applying more symmetrizations only decreases the Lebesgue distance to the circle further. Applying this procedure to all layers will then lead to the intended approximation.

Theorem 2.14. *Let $1 \le k \le d$, $M \in \mathcal{L}(\mathbb{R}^d)$ have finite measure and let $\epsilon > 0$. Then there exist directions $v_1, ..., v_n \in \mathbb{R}^k \times \{0\}^{d-k}$ such that*

$$d_L(S_{v_n} ... S_{v_1} M, S^k M) < \epsilon.$$

Proof. Define $V := \mathbb{R}^k \times \{0\}^{d-k}$ and $\tau : \mathbb{R}^k \longrightarrow V$ as the canonical injection. This proof is in essence a proof by induction. We will apply *theorem 2.12* to "V-parallel slices" of M, one after the other.
Define $Q^d := (0,1) \times ... \times (0,1) \subset \mathbb{R}^d$ and Q^k analogously. First we will show the theorem for the case that M is a finite union of open cubes in the following sense: There exists a $\delta > 0$ and pairwise different tuples $z_{i,j} = (z_{i,j,1}, ..., z_{i,j,d}) \in \mathbb{Z}^d$, for $1 \le i \le N$ and $1 \le j \le N_i$, such that

$$M = \bigcup_{i=1}^{N} \bigcup_{j=1}^{N_i} Q^d(z_{i,j}),$$

where

$$Q^d(z) := (Q^d + z) \cdot \delta \quad \forall z \in \mathbb{Z}^d.$$

Furthermore, we require that i determines the last $d-k$ components of $z_{i,j}$, meaning that

$$z_{i,j,m} = z_{i,j',m} \quad \forall 1 \le i \le N \quad \forall 1 \le j, j' \le N_i \quad \forall k+1 \le m \le d.$$

In other words i determines the V-parallel layer in which the cube $Q^d(z_{i,j})$ lies. As a result, for every $1 \le i \le N$ there exists a $z_i \in \mathbb{R}^{d-k}$ such that

$$z_i = (z_{i,j,k+1}, ..., z_{i,j,d}) \quad \forall 1 \le j \le N_i.$$

For ease of notation, we define $z_{i,j}^k = (z_{i,j,1}^k, ..., z_{i,j,k}^k)$ for $1 \leq j \leq N_i$. We can then write

$$z_{i,j} = \left(z_{i,j}^k, z_i\right) \in \mathbb{Z}^k \times \mathbb{Z}^{d-k} \quad \forall 1 \leq j \leq N_i.$$

Our naming scheme for the cubes allows us to easily see the following: Fix a $1 \leq m \leq N$. Then

$$M \cap \left(\mathbb{R}^k \times Q^{d-k}(z_m)\right) = \bigcup_{i=1}^{N} \bigcup_{j=1}^{N_i} Q^d(z_{i,j}) \cap \left(\mathbb{R}^k \times Q^{d-k}(z_m)\right).$$

The intersection on the right hand side is empty unless $(z_{i,j,k+1}, ..., z_{i,j,d}) = z_m$, in which case $i = m$ and the intersection is $Q^d(z_{i,j})$. We get that

$$M \cap \left(\mathbb{R}^k \times Q^{d-k}(z_m)\right) = \bigcup_{j=1}^{N_m} Q^d(z_{i,j}).$$

This will help us understand $S^k M$. For a given $y \in \mathbb{R}^{d-k}$, the definition implies that

$$S^k M \cap \left(\mathbb{R}^k \times \{y\}\right) = SW_k(\{x \in \mathbb{R}^k : (x, y) \in M\}) \times \{y\}.$$

Now note that for any $1 \leq i \leq N$ and $1 \leq j \leq N_j$,

$$Q^d(z_{i,j}) = \left(Q^k \times Q^{d-k} + (z_{i,j}^k, z_i)\right) \cdot \delta = Q^k(z_{i,j}^k) \times Q^{d-k}(z_i).$$

Using this, we get that for any $y \in Q^{d-k}(z_i)$ it is the case that

$$\{x \in \mathbb{R}^k : (x, y) \in M\} = \bigcup_{j=1}^{N_i} Q^k(z_{i,j}^k),$$

and for any other y this set is empty. Therefore we can, for $y \in Q^{d-k}(z_i)$ and using the definition of Schwarz symmetrization, conclude that

$$S^k M \cap \left(\mathbb{R}^k \times \{y\}\right) = \overline{B_{r_i}^k} \times \{y\},$$

where $r_i > 0$ is chosen so that

$$\omega_k |r_i|^k = m^k \left(\bigcup_{j=1}^{N_i} Q^k(z_{i,j}^k)\right) = N_i \cdot \delta^k.$$

We can summarize the above results as

$$S^k \left(M \cap (\mathbb{R}^k \times Q^{d-k}(z_i))\right) = S^k \left(\bigcup_{j=1}^{N_i} Q^d(z_{i,j})\right) = \overline{B_{r_i}^k} \times Q^{d-k}(z_i)$$

for all $1 \leq i \leq N$. Now we will inductively show the following statement:
For every $\epsilon > 0$ and for every $1 \leq m \leq N$, there exist directions $v_1, ..., v_l \in V$ such that

$$d_L\left(S_{v_l} ... S_{v_1}\left(\bigcup_{i=1}^{m} \bigcup_{j=1}^{N_i} Q^d(z_{i,j})\right), S^k\left(\bigcup_{i=1}^{m} \bigcup_{j=1}^{N_i} Q^d(z_{i,j})\right)\right) < \frac{\epsilon}{N} \delta^{d-k} m.$$

$\underline{m=1.}$ The set $\bigcup_{j=1}^{N_1} \overline{Q^k(z_{1,j}^k)} \subseteq \mathbb{R}^k$ is compact, so according to *theorem 2.12* we can find directions $a_1, ..., a_l \in \mathbb{R}^k$ such that

$$d_L\left(S_{a_l}...S_{a_1}\left(\bigcup_{j=1}^{N_1} \overline{Q^k(z_{1,j}^k)}\right), B_{r_1}^k\right) < \frac{\epsilon}{N}.$$

Because the boundary of the cubes is a zero set the inequality still holds without taking the cubes' closures.
Now we embedd these sets in $\mathbb{R}^d$. We define $v_n := \tau(a_n) = (a_n, 0) \in V$ for all $1 \leq n \leq l$. The above statement gives us

$$d_L\left(S_{v_l}...S_{v_1}\left(\bigcup_{j=1}^{N_1} Q^k(z_{1,j}^k) \times Q^{d-k}(z_1)\right), B_{r_1}^k \times Q^{d-k}(z_1)\right) < \frac{\epsilon}{N}\delta^{d-k},$$

which can be seen with previous identities to be exactly the inequality that had to be shown.

$\underline{m>1.}$ We assume the statement to be true for $m-1$ with the directions $v_l, ..., v_1 \in V$. Denote the collective symmetrization by $S = S_{v_l} \circ ... \circ S_{v_1}$ and let $S' = S_{\tau^{-1}(v_l)}...S_{\tau^{-1}(v_1)}$ be the analogous symmetrization in $\mathbb{R}^k$. Again, the set $S'\left(\bigcup_{j=1}^{N_m} \overline{Q^k(z_{m,j}^k)}\right) \subseteq \mathbb{R}^k$ is compact, so according to *theorem 2.12* we can find directions $a_1, ..., a_t \in \mathbb{R}^k$ such that

$$d_L\left(S_{a_t}...S_{a_1}S'\left(\bigcup_{j=1}^{N_m} \overline{Q^k(z_{m,j}^k)}\right), B_{r_m}^k\right) < \frac{\epsilon}{N}.$$

As in the case $m = 1$, we now want to embedd this into $\mathbb{R}^d$ and therefore define directions $v_{l+n} := \tau(a_n) = (a_n, 0) \in V$ for all $1 \leq n \leq t$. Note that with this naming scheme $S_{v_{l+t}}...S_{v_{l+1}}S = S_{v_{l+t}}...S_{v_1}$. Then the above statement gives us

$$d_L\left(S_{v_{l+t}}...S_{v_1}\left(\bigcup_{j=1}^{N_m} Q^k(z_{m,j}^k) \times Q^{d-k}(z_m)\right), B_{r_m}^k \times Q^{d-k}(z_m)\right) < \frac{\epsilon}{N}\delta^{d-k}.$$

Now let us remember our inductive assumption:

$$d_L\left(S_{v_l}...S_{v_1}\left(\bigcup_{i=1}^{m-1}\bigcup_{j=1}^{N_i} Q^d(z_{i,j})\right), S^k\left(\bigcup_{i=1}^{m-1}\bigcup_{j=1}^{N_i} Q^d(z_{i,j})\right)\right) < \frac{\epsilon}{N}\delta^{d-k}(m-1).$$

We can apply $S_{v_{l+t}} \circ ... \circ S_{v_{l+1}}$ to both sides and the inequality remains true (*lemma 1.6*). Conveniently,

$$S_{v_{l+t}}...S_{v_{l+1}}S^k\left(\bigcup_{i=1}^{m-1}\bigcup_{j=1}^{N_i} Q^d(z_{i,j})\right) = S^k\left(\bigcup_{i=1}^{m-1}\bigcup_{j=1}^{N_i} Q^d(z_{i,j})\right),$$

because, informally, from the perspective of each $S_{v_{l+n}}$ the set is already a ball. Taking both results together yields

$$d_L\left(S_{v_{l+t}}...S_{v_1}\left(\bigcup_{i=1}^{m}\bigcup_{j=1}^{N_i} Q^d(z_{i,j})\right), S^k\left(\bigcup_{i=1}^{m}\bigcup_{j=1}^{N_i} Q^d(z_{i,j})\right)\right) < \frac{\epsilon}{N}\delta^{d-k}m.$$

This finishes our proof by induction. Our result for $m = N$ is then

$$d_L\left(S_{v_{l+t}}...S_{v_1}M, S^kM\right) < \epsilon\delta^{d-k}.$$

For the case of a more general measurable set $M \subseteq \mathbb{R}^d$, we first find a set M' which is a finite union of open cubes in the same sense as before. By the definition of the Lebesgue measure we can find such an M' with $d_L(M, M') < \frac{1}{3}\epsilon$. Then, according to the first part of the proof, we can find directions $v_1, ..., v_l \in V$ such that $d_L(S_{v_l}...S_{v_1}M, S_V(M)) < \frac{1}{3}\epsilon$. We use that Steiner symmetrization decreases the Lebesgue distance (*lemma 1.6*) to conclude

$$d_L(S_{v_l}...S_{v_1}M', S^k(M')) \leq 2d_L(M, M') + d_L(S_{v_l}...S_{v_1}M, S^k(M)) < \frac{2}{3}\epsilon + \frac{1}{3}\epsilon = \epsilon.$$

□

A simple function can be imagined as a stack of layers. Then, in a similar fashion as for the previous theorem, we can apply 1-dimensional symmetrizations to these layers one after the other, until the result approximates the k-dimensional symmetrization.

Theorem 2.15. *Let $1 < k \leq d$ and $1 \leq p < \infty$. Let $f \in L^p_{\geq}(\mathbb{R}^d)$. Then, for every $\epsilon > 0$, there exist directions $v_1, ..., v_n \in \mathbb{R}^k \times \{0\}^{d-k}$ such that*

$$\|S_{v_l}...S_{v_1}f, S^kf\|_{L^p} < \epsilon.$$

Proof. Because simple functions are dense and Steiner Symmetrization is continuous in L^p (*lemma 1.8*) it is sufficient to prove the theorem for the case that

$$f = \sum_{i=1}^{N} \lambda_i \mathbb{1}_{A_i},$$

where $A_{i+1} \subseteq A_i, \lambda_i \geq 0$ for all $1 \leq i \leq N$ and the sets have finite measure. We will first show this for the case $p = 1$. Note that for any rearrangement $*$ the following relation holds:

$$f^*(x) = \sup\{c \geq 0 : x \in \{f \geq c\}^*\} = \sup\left\{\sum_{i=1}^{N} \lambda_i : x \in A_N^*\right\} = \sum_{i=1}^{N} \lambda_i \mathbb{1}_{A_i^*}(x).$$

Now we will inductively show the following:
For every $\epsilon > 0$ and for every $1 \leq m \leq N$, there exist directions $v_1, ..., v_n \in V$ such that

$$\left\|S_{v_n}...S_{v_1}\left(\sum_{i=1}^{m} \lambda_i \mathbb{1}_{A_i}\right) - S^k\left(\sum_{i=1}^{m} \lambda_i \mathbb{1}_{A_i}\right)\right\|_{L^1} < \epsilon.$$

Before we start with the proof by induction, note that

$$\int_{\mathbb{R}^d} \left|S_{v_n}...S_{v_1}\left(\sum_{i=1}^{m} \lambda_i \mathbb{1}_{A_i}\right) - S^k\left(\sum_{i=1}^{m} \lambda_i \mathbb{1}_{A_i}\right)\right| dm^d$$

$$= \int_{\mathbb{R}^d} \left| \sum_{i=1}^{m} \lambda_i \mathbb{1}_{S_{v_n}...S_{v_1} A_i} - \sum_{i=1}^{m} \lambda_i \mathbb{1}_{S^k A_i} \right| dm^d = \sum_{i=1}^{m} \lambda_i d_L(S_{v_n}...S_{v_1} A_i, S^k A_i).$$

$\underline{m = 1}$. The *theorem 2.14* gives us directions $v_1, ..., v_n \in \mathbb{R}^k \times \{0\}^{d-k}$ such that

$$d_L(S_{v_n}...S_{v_1} A_1, S^k A_1) < \frac{\epsilon}{\lambda_1}.$$

$\underline{m > 1}$. The inductive assumption is that we have already solved the case $m - 1$. Let $v_1, ..., v_n$ be the directions that achieve this, i.e.:

$$\sum_{i=1}^{m-1} \lambda_i d_L(S_{v_n}...S_{v_1} A_i, S^k A_i) < \frac{\epsilon}{2}$$

Then, again, *theorem 2.14* gives us directions $v_{l+1}, ..., v_{l+t}$ such that

$$d_L(S_{v_{n+t}}...S_{v_{n+1}} S_{v_l}...S_{v_1} A_m, S^k A_m) < \frac{\epsilon}{2\lambda_m}.$$

We can conclude the proof by using (*lemma 1.8*), where we have shown that applying further transformations only decreases the distance in the lebesgue metric:

$$\sum_{i=1}^{m} \lambda_i d_L(S_{v_{n+t}}...S_{v_{n+1}} S_{v_l}...S_{v_1} A_i, S^k A_i)$$

$$\leq \sum_{i=1}^{m-1} \lambda_i d_L(S_{v_n}...S_{v_1} A_i, S^k A_i) + \lambda_m d_L(S_{v_{n+t}}...S_{v_1} A_m, S^k A_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

The case $m = N$ is what had to be shown to finish the proof for $p = 1$.
Now let $1 < p < \infty$ and $\epsilon > 0$. Note that $f^p \in L^1_{\geq}(\mathbb{R}^d)$. We apply the case $p = 1$ to f^p with this ϵ and find vectors $v_1, ..., v_n$. The proof is finished as follows:

$$\int_{\mathbb{R}^d} |S_{v_n}...S_{v_1} f - S^k f|^p dm^d \leq \int_{\mathbb{R}^d} |(S_{v_n}...S_{v_1} f)^p - (S^k f)^p| dm^d$$

$$= \int_{\mathbb{R}^d} |S_{v_n}...S_{v_1} f^p - S^k f^p| dm^d < \epsilon.$$

□

3 The Pólya–Szegő inequality

In section 2 we have proven *Theorem 2.15*, which allows us to approximate the Steiner symmetrization in k dimensions with the one in 1 dimension. We will use this to deduce the Pólya–Szegő inequality for the case $k > 1$ from the case $k = 1$. Hence, we will now show a restricted version of the Pólya–Szegő inequality for the case $k = 1$. Note that we do not need any of the results from section 2 for the following theorem and proof.

3.1 Pólya–Szegő for piecewise affine functions and $k = 1$

In this section, we will prove *theorem 1.2*, whose proof contains the main body of work necessary for the proof of *theorem 3.5*. We consider a certain class of functions which we call piecewise affine functions. This name is misleading in the sense that our definition of piecewise affine functions is slightly more restrictive than what one would normally consider to be a piecewise affine function.

Definition 3.1. *Let $\Omega \subseteq \mathbb{R}^d$ be open and bounded. We call a function $f : \overline{\Omega} \longrightarrow \mathbb{R}$ piecewise affine with support in $\overline{\Omega}$ if:*

(i) $f \in C(\overline{\Omega})$

(ii) $f\big|_{\partial\Omega} = 0$

(iii) There exist open, bounded, convex, nonempty, and disjoint sets $U_i \subseteq \Omega$, and affine functions $f_i : \mathbb{R} \times \mathbb{R}^{d-1} \longrightarrow \mathbb{R}, 1 \leq i \leq n$ such that $f\big|_{U_i} = f_i\big|_{U_i}$, $\partial_x f_i \neq 0$ and $\overline{\Omega} = \bigcup_{i=1}^n \overline{U_i}$.

The additional restriction here is the requirement that $\partial_x f_i \neq 0$. We require this because it makes symmetrization in direction x easier. At the same time, we do not lose much as our piecewise affine functions still lie dense in $W_0^{1,p}(\Omega)$. We will sketch a proof of this in *lemma 3.3*.
It is also worth mentioning that the sets U_i do indeed, as the reader might suspect, have to be polyhedral for the function to be continuous. Since we do not need this property in the proof of the following theorem, we do not mind our definition being ”unnecessarily weak” in this regard.

The following theorem and proof is based on a theorem in [4] (theorem 2.31 on page 83). We have modified the proof to avoid some more complex arguments and focus on the geometric aspects.

Theorem 3.2. *Let $\Omega \subset \mathbb{R} \times \mathbb{R}^{d-1}$ be open and bounded. Let $F : \mathbb{R} \times \mathbb{R}^{d-1} \longrightarrow [0, \infty)$ such that*

(i) F is convex

(ii) $F(x, y) = F(-x, y)$

(iii) $|x| \leq |x'| \Rightarrow F(x, y) \leq F(x', y)$.

Let $f : \overline{\Omega} \longrightarrow [0,\infty)$ be piecewise affine with support in $\overline{\Omega}$, and let $f^* : \overline{\Omega}^* \longrightarrow [0,\infty)$ be the rank 1 Steiner symmetrization of f in direction x. Then

$$\int_{\overline{\Omega}^*} F(\nabla f^*)\, dm^d \leq \int_{\overline{\Omega}} F(\nabla f)\, dm^d.$$

Proof. The proof of this theorem is rather geometric, so we will make use of figures and urge the reader to try to visualize the presented statements on these figures.
Picture the function f as the one given below on the left. Then the rearranged function f^* is the one given below on the right. In each figure, a line in direction x intersecting

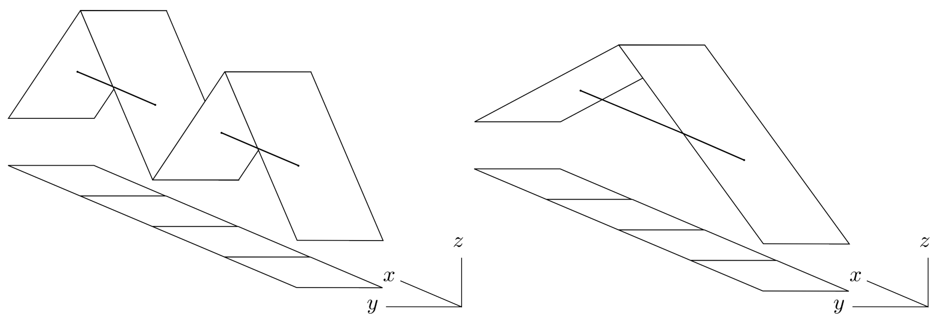


Figure 5: An example of f and f^*.

the graph is drawn. Due to the definition of the rearrangement, the total amount of line segment in both figures has to be the same.
We now define the variables that we will work with. These definitions should not be read too closely, and instead one should make sure to understand the figure which follows.

Let $1 \leq i \leq n$. First we define $b_i \in \mathbb{R}$ and $(u_i, v_i) \in \mathbb{R}^{d-1} \times \mathbb{R}$ so that

$$f_i(x,y) = b_i + (x,y)\cdot(u_i,v_i) \quad \forall (x,y) \in \mathbb{R}^{d-1}\times\mathbb{R}.$$

Then we define

$$\begin{aligned}
g_i : \mathbb{R}^{d-1}\times\mathbb{R} &\longrightarrow \mathbb{R} \\
(y,z) &\longmapsto \frac{1}{u_i}(z - yv_i - b_i) \\
\eta_i : \mathbb{R}\times\mathbb{R}^{d-1} &\longrightarrow \mathbb{R}^{d-1}\times\mathbb{R} \\
(x,y) &\longmapsto (y, f_i(x,y)) \\
(g_i(y,z), y) &\longleftarrow (y,z)
\end{aligned}$$

and see that

$$\begin{aligned}
f_i(g_i(y,z),y) &= b_i + (g_i(y,z),y)\cdot(u_i,v_i) = b_i + g_i(y,z)u_i + yv_i = z \\
g_i(y, f_i(x,y)) &= \frac{1}{u_i}(f_i(x,y) - yv_i - b_i) = x,
\end{aligned}$$

so the given inverse of η_i is actually correct, and η_i is an invertible affine function. Note that this means η_i is a homeomorphism, and by computing the determinant of its linear part, we get

$$m^d(\eta_i(M)) = |u_i| m^d(M) = |\partial_x f_i| m^d(M) \quad \forall M \subseteq \mathbb{R}^{d-1} \times \mathbb{R} \text{ measurable.}$$

This will be of use later in the proof. We also define

$$\eta(x, y) = (y, f(x, y)) \text{ and } \eta^*(x, y) = (y, f^*(x, y))$$

as functions on $\overline{\Omega}$ and $\overline{\Omega}^*$ respectively.

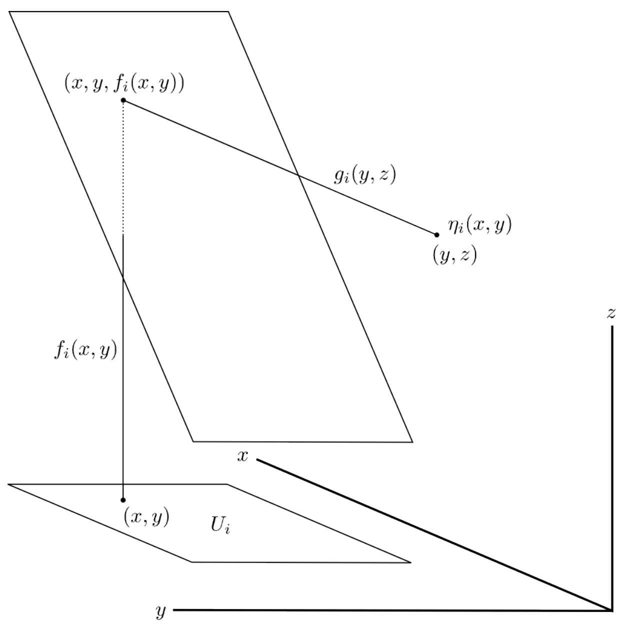


Figure 6: A demonstration of the variables that were introduced.

The way the proof will work is to switch from thinking about the function in terms of the segments f_i, to thinking about it in terms of the functions g_i. Rearrangement is a complicated process when we think of f in terms of f_i, as there is a supremum in the definition of f^* that has to be computed. Thinking of the rearrangement in terms of g_i is simpler though: The "g" that corresponds to the rearranged function will be a sort alternating of sum of the functions g_i.
We will use the functions η, η^* and η_i to map sets from the (x, y)-plane to the (y, z)-plane and back. The "η"-functions help us work with the domains of f_i and g_i.

The proof is split into two steps. In step 1, we show the following inequality:

$$\int_{\eta^{-1}(B)} F(\nabla f) dm^d \geq \int_{\eta^{*-1}(B)} F(\nabla f^*) dm^d,$$

where $B \subset \mathbb{R}^{d-1} \times [0, \infty)$ is an open ball with the property that for $1 \le i \le n$,

$$\big(B \subseteq \eta_i(U_i) \text{ or } B \cap \eta_i(U_i) = \varnothing\big),$$

and there exists a $1 \le i \le n$ such that $B \subseteq \eta_i(U_i)$.
In step 2 we generalize this to the complete domains $\overline{\Omega}$ and $\overline{\Omega}^*$.

The following two figures are the pictures that one should have in mind when reading the proof. We pick a ball B of which we can be sure that it only intersects straight pieces of the graph of f. Then we integrate over the set $\eta^{-1}(B)$, which is the set of points that f maps into the cylinder corresponding to the ball. With these simplifications, the part of the graph of f that we are interested in is merely a collection of circular "cutouts" of hyperplanes.

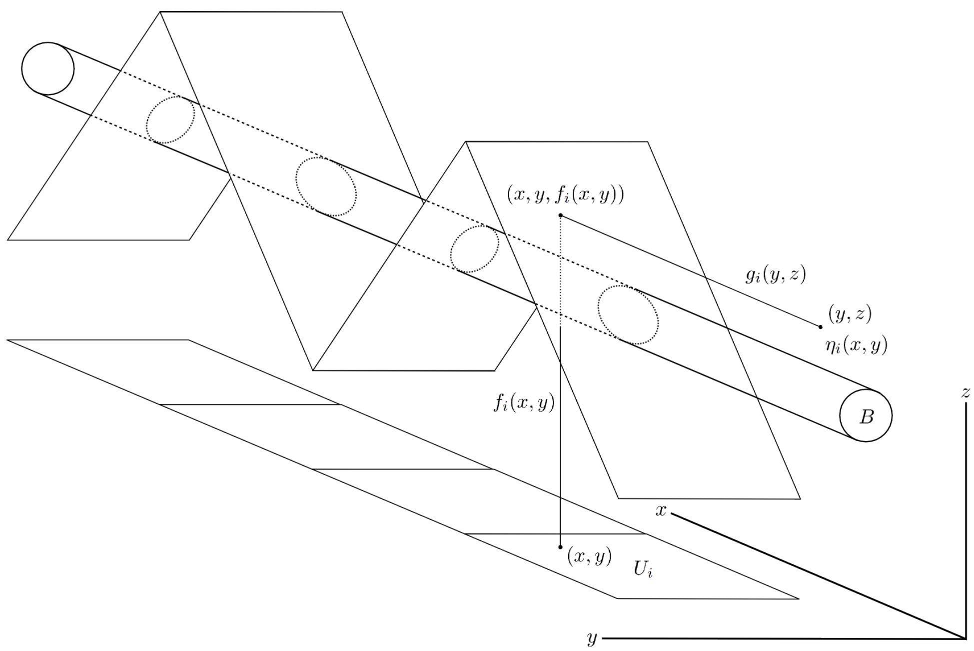


Figure 7: The function f and the cylinder with base B.

Step 1. First we see that because of our requirements for B,

$$\eta^{-1}(B) = \{(x, y) \in \overline{\Omega} : \eta(x, y) \in B\} = \bigcup_{i=1}^{n} \{(x, y) \in \overline{U_i} : (y, f(x, y)) \in B\}$$

$$= \bigcup_{i=1}^{n} \overline{U_i} \cap \eta_i^{-1}(B) = \bigcup_{i \in I} \eta_i^{-1}(B) = \bigcup_{i \in I} \{(g_i(y, z), y) : (y, z) \in B\},$$

where $I = \{1 \le i \le n : B \subseteq \eta_i(U_i)\}$.
The final union is disjoint because the sets $\eta_i^{-1}(B) \subseteq U_i$ are disjoint. Therefore, for any

$(y, z) \in B$, the values $g_i(y, z)$ with $i \in I$ are pairwise nonequal.
If $g_i(y, z) > g_j(y, z)$ and $g_i(y', z') < g_j(y', z')$ were true for some $(y, z) \neq (y', z') \in B$ and $i \neq j \in I$, then, because B is convex and $g_i - g_j$ is continuous, we could find a point $(y'', z'') \in B$ such that $g_i(y'', z'') = g_j(y'', z'')$. This would contradict the above pairwise nonequality.
Hence the functions $g_i\big|_B, i \in I$ are strictly ordered. We call these $g_1' < \ldots < g_m'$ for $m = |I|$. Analogously, we define f_i', η_i', U_i' for $1 \leq i \leq m$ to correspond to g_i'.

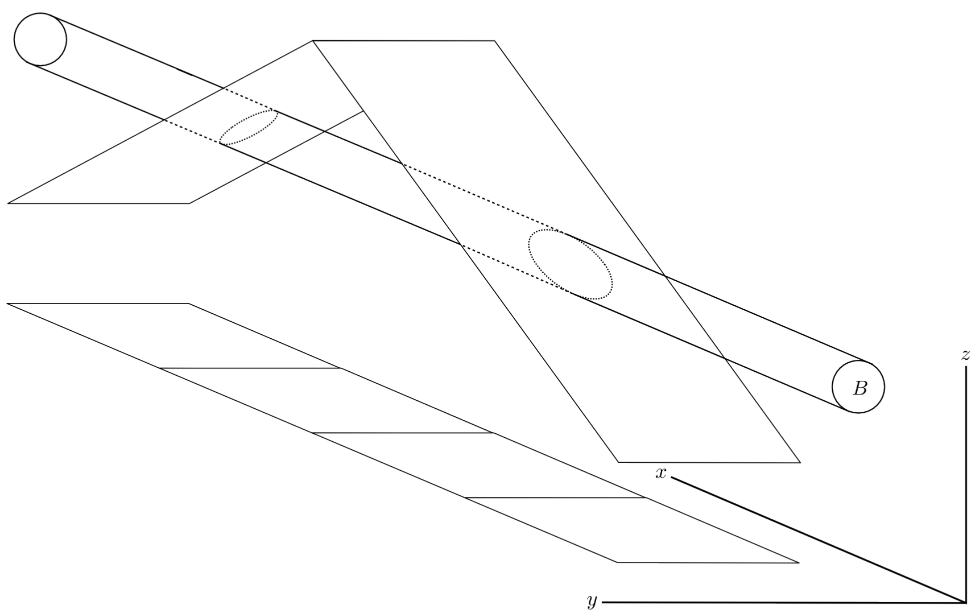


Figure 8: The function f^* and the cylinder with base B.

Let $(y, z) \in B$. Then, because of the above,

$$f(x, y) = z \iff (y, f(x, y)) = (y, z) \in B$$

$$\iff \exists i \in I : (x, y) \in U_i \text{ and } \eta_i(x, y) = (y, f_i(x, y)) = (y, z) \in B$$

$$\iff \exists 1 \leq i \leq m : g_i'(y, z) = x.$$

The function $\left[x \mapsto f(x, y)\right] \in C(\{x : (x, y) \in \overline{\Omega}\})$ is zero on the boundary and positive, so as a consequence of the intermediate value theorem and the fact that $\partial_x f(x, y) \neq 0$ for all but finitely many x, the points $g_i'(y, z), 1 \leq i \leq m$ are exactly where f passes from above to below z, and vice versa:

$$\{x : (x, y) \in \overline{\Omega} \wedge f(x, y) \geq z\} = \bigcup_{i=1}^{m} \left[g_{2i-1}'(y, z), g_{2i}'(y, z)\right]$$

$$\implies m^1(\{x : (x, y) \in \overline{\Omega} \wedge f(x, y) \geq z\}) = \sum_{i=1}^{m} (-1)^i g_i'(y, z) =: g(y, z).$$

Recall that we still let $(y, z) \in B$ be arbitrary. In addition, we now let $x \in \mathbb{R}$ with $(x, y) \in \overline{\Omega}^*$. There exists an $r > 0$ such that $|z - z'| < r$ implies $(y, z') \in B$.
We want to find out precisely when it is the case that $f^*(x, y) = z$, and we start with the following equivalence:

$$z = \sup\{z' \in (z - r, z + r) : (x, y) \in \{(x' : f(x', y) \geq z'\}^*\}$$

$$\iff z = \sup\{z' \in (z - r, z + r) : \{(x' : f(x', y) \geq z'\}\} \neq \varnothing$$
$$\text{and } |x| \leq \frac{1}{2} m^1(\{(x' : f(x', y) \geq z'\})\}.$$

Since $f_i(g_i(y, z'), y) = z'$ for any $1 \leq i \leq n$ such that $B \subseteq U_i$, the first part is always true. Then the above is equivalent to

$$z = \sup\{z' \in (z - r, z + r) : |x| \leq \frac{1}{2} g(y, z')\}.$$

Because g is decreasing in the z component and continuous, this is equivalent to

$$z = \sup\{z' \geq 0 : |x| \leq \frac{1}{2} g(y, z')\},$$

so we can conclude

$$f^*(x, y) = z \iff |x| = \frac{1}{2} g(y, z).$$

We can now analyze the following set:

$$\eta^{*-1}(B) = \{(x, y) \in \overline{\Omega}^* : (y, f^*(x, y)) \in B\}$$

$$= \{(x, y) \in \overline{\Omega}^* : \exists z \in \mathbb{R} : (y, z) \in B \wedge f^*(x, y) = z\}$$

$$= \left\{ \left(-\frac{1}{2} g(y, z), y \right) : (y, z) \in B \right\} \cup \left\{ \left(\frac{1}{2} g(y, z), y \right) : (y, z) \in B \right\}$$

Note that the last union is disjoint because we know that $g\big|_B > 0$.
The function $\left[(y, z) \mapsto (x, \frac{1}{2} g(y, z)) \right]$ is affine, so by computing the determinant of the linear part we can find that

$$m^d(D_\pm) = \frac{1}{2} |\partial_z g| m^d(B),$$

where $D_\pm := \{(\pm \frac{1}{2} g(y, z), y) : (y, z) \in B\}$. Since g is affine, there exists a $b \in \mathbb{R}$ such that $g(y, z) = +(\partial_y g, \partial_z g) \cdot (y, z)$. Then for every $(x, y) \in D_+$,

$$f^*(x, y) = \frac{1}{\partial_z g} (2x - b - \partial_y g \cdot y),$$

and the analogous is true for D_-. We can now write down the gradient of f^* as

$$\nabla f^*\big|_{D_+} = \left(\frac{2}{\partial_z g}, \frac{-\partial_y g}{\partial_z g} \right) \text{ and } \nabla f^*\big|_{D_-} = \left(\frac{-2}{\partial_z g}, \frac{-\partial_y g}{\partial_z g} \right).$$

If we look back at the definitions of f_i and g_i, we can calculate, for all $1 \le i \le n$, that

$$\nabla f_i = \left(\frac{1}{\partial_z g_i}, \frac{-\partial_y g_i}{\partial_z g_i} \right).$$

The last pieces that we need to finish step are the signs of $\partial_z g_i'$ and $\partial_z g$. Assume it is the case for some $1 \le i < m$ that $\partial_z g_i'$ and $\partial_z g_{i+1}'$ have the same sign. Then $\partial_x f_i'$ and $\partial_x f_{i+1}'$ are nonzero and have the same sign as well. Earlier, we had discovered that the points $g_i'(y,z)$, for $1 \le i \le m$, are exactly where f passes from above to below z, and vice versa. Using this, we get that

$$\begin{aligned} &\left(g_i'(y,z), g_{i+1}'(y,z)\right) \subseteq \{x : (x,y) \in \overline{\Omega} \wedge f(x,y) \ge z\} \\ \text{or } &\left(g_i'(y,z), g_{i+1}'(y,z)\right) \subseteq \{x : (x,y) \in \overline{\Omega} \wedge f(x,y) < z\}. \end{aligned}$$

At the same time $(g_i'(y,z), y) \in \eta_i^{-1}(B) \subseteq U_i$, so $\partial_x f(g_i'(y,z), y)$ and $\partial_x f(g_{i+1}'(y,z), y)$ have the same sign (and are nonzero). This is not possible together with the previous. Because $g_1'(x,y)$ is the smallest value for x where $f(x,y) = z$, and for all smaller x is is the case that $f(x,y) < z$, we know that $\partial_x f_1' > 0$. Finally, note that $g(y,z)$ is the measure of a set that shrinks with z. Combining all these results yields the following signs:

$$\operatorname{sign}(\partial_z g) = -1 \text{ and } \operatorname{sign}(\partial_z g_i') = (-1)^{i-1} \quad \forall 1 \le i \le m.$$

Now we show the inequality.

$$\begin{aligned} \int_{\eta^{*-1}(B)} F(\nabla f^*) dm^d &= \int_{D_+} F(\nabla f^*) + \int_{D_-} F(\nabla f^*) \\ &= m^d(D_+) F\left(\frac{2}{\partial_z g}, \frac{-\partial_y g}{\partial_z g} \right) + m^d(D_-) F\left(\frac{-2}{\partial_z g}, \frac{-\partial_y g}{\partial_z g} \right) \\ &= -m^d(B) \frac{1}{2} \partial_z g \left(F\left(\frac{2}{\partial_z g}, \frac{-\partial_y g}{\partial_z g} \right) + F\left(\frac{-2}{\partial_z g}, \frac{-\partial_y g}{\partial_z g} \right) \right). \end{aligned}$$

We use properties (*ii*) and (*iii*) of F, with $m \ge 2$:

$$\begin{aligned} &\le -m^d(B) \partial_z g F\left(\frac{m}{\partial_z g}, \frac{-\partial_y g}{\partial_z g} \right) \\ &= -m^d(B) \partial_z g F\left(\sum_{i=1}^m (-1)^i \frac{(-1)^i}{\partial_z g}, \sum_{i=1}^m (-1)^i \frac{-\partial_y g_i'}{\partial_z g} \right) \\ &= -m^d(B) \partial_z g F\left(\sum_{i=1}^m (-1)^i \frac{\partial_z g_i'}{\partial_z g} \frac{(-1)^i}{\partial_z g_i'}, \sum_{i=1}^m (-1)^i \frac{\partial_z g_i'}{\partial_z g} \frac{-\partial_y g_i'}{\partial_z g_i'} \right). \end{aligned}$$

Finally, we use that F is convex:

$$\le -m^d(B) \partial_z g \sum_{i=1}^m (-1)^i \frac{\partial_z g_i'}{\partial_z g} F\left(\frac{(-1)^i}{\partial_z g_i'}, \frac{-\partial_y g_i'}{\partial_z g_i'} \right)$$

$$= \sum_{i=1}^{m} (-1)^{i-1} \partial_z g_i' m^d(B) F\left(\frac{1}{\partial_z g_i'}, \frac{-\partial_y g_i'}{\partial_z g_i'}\right)$$

$$= \sum_{i=1} m^d(\eta_i'^{-1}(B)) F(\nabla f_i') = \int_{\eta^{-1}(B)} F(\nabla f) dm^d.$$

Step 2. Since each U_i is convex, the sets ∂U_i have measure zero (in particular their boundary has measure zero). Now recall that every η_i is a homeomorphism that changes the measure of a set by only a factor. Then

$$\bigcup_{i=1}^{n} \eta_i(\partial U_i) = \bigcup_{i=1}^{n} \partial(\eta_i(U_i))$$

is a zero set and

$$\eta(\overline{\Omega}) = \bigcup_{i=1}^{n} \eta(\overline{U_i}) = \bigcup_{i=1}^{n} \eta_i(U_i) \cup \eta_i(\partial U_i).$$

The set

$$\bigcup_{i=1}^{n} \eta_i(U_i) \setminus \bigcup_{i=1}^{n} \partial(\eta_i(U_i)).$$

is open, so we can write it as a countable union of balls $\mathcal{B} \subset \mathcal{P}(\mathbb{R}^{d-1} \times \mathbb{R})$ to get

$$\eta(\overline{\Omega}) = \bigcup_{B \in \mathcal{B}} B \text{ up to a set of measure zero.}$$

These balls fulfill the requirements for the balls in Step 1.
We will show later η^{*-1} takes zero sets to zero sets. For now we assume that this is the case. Then

$$\eta^{*-1}(\eta(\overline{\Omega})) = \eta^{*-1}\left(\bigcup_{B \in \mathcal{B}} B\right) = \bigcup_{B \in \mathcal{B}} \eta^{*-1}(B) \text{ up to a set of measure zero.}$$

We now show that $\eta^*(\overline{\Omega}^*) \subseteq \eta(\overline{\Omega})$. Let $(y, z) \in \eta^*(\overline{\Omega}^*)$. Then there exists an $x \in \mathbb{R}$ with $(x, y) \in \overline{\Omega}^*$ such that

$$z = f^*(x, y) = \sup\{a \in \mathbb{R} : (x, y) \in \{f \geq a\}^*\}.$$

Let $a_n \searrow z$ be a sequence that attains the supremum. Then

$$\{x' \in \mathbb{R} : (x', y) \in \{f \geq a_n\}\} \neq \emptyset \quad \forall n \in \mathbb{N}.$$

By using compactness of $\{x \in \mathbb{R} : (x, y) \in \overline{\Omega}\}$, we can find a convergent sequence $x_n' \to x'$ such that $f(x_n', y) \geq a_n$ for every $n \in \mathbb{N}$. Now continuity implies $f(x', y) \geq z$, and since $f\big|_{\partial\Omega} = 0$, there must be an $x'' \in \mathbb{R}$ such that $f(x'', y) = z$.

We can now show the inequality:

$$\int_{\overline{\Omega}^*} F(\nabla f^*) dm^d = \int_{\eta^{*-1}(\eta^*(\overline{\Omega}^*))} F(\nabla f^*) dm^d$$

$$\leq \sum_{B\in\mathcal{B}} \int_{\eta^{*-1}(B)} F(\nabla f^*)dm^d \leq \sum_{B\in\mathcal{B}} \int_{\eta^{-1}(B)} F(\nabla f)dm^d \leq \int_{\overline{\Omega}} F(\nabla f)dm^d.$$

As the last part of the proof, it remains to show that η^{*-1} takes zero sets to zero sets. By the nature of the rearrangement, there are two values for x which could fulfill $\eta^*(x,y) = (y,z)$. We will show that these two parts of η^* are Lipschitz functions and hence each map zero sets to zero sets. We need to subdivide Ω in a certain way:

$$J \in \mathcal{J} = \left\{ I \subseteq \{1, ..., n\} : \bigcap_{i\in I} \eta_i(\overline{U_i}) \cap \eta_k(\overline{U_k}) = \varnothing \quad \forall k \notin I \right\}.$$

With this definition, every ball $B \subseteq \bigcap_{j\in J} \eta_j(U_j)$ has the properties that we specified in step 1. Repeating the arguments from step 1, we get that

$$f^*(x,y) = z \iff |x| = \frac{1}{2}g(y,z) \quad \forall (y,z) \in \bigcup_{j\in J} \eta_j(U_j),$$

where g is some affine function that depends on J.
Because $\bigcup_{i=1}^n \eta_i(\partial U_i)$ is a zero set, we have

$$\bigcup_{J\in\mathcal{J}} \bigcap_{j\in J} \eta_j(U_j) = \bigcup_{i=1}^n \eta_i(U_i) = \eta(\overline{\Omega}) \text{ up to a set of measure zero}$$

$$\implies \bigcup_{J\in\mathcal{J}} \bigcap_{j\in J} \eta_j(U_j) \text{ is dense in } \eta(\overline{\Omega}) \supseteq \eta^*(\overline{\Omega}^*).$$

If $(y,z) \in \eta^*(\overline{\Omega}^*)$, then there exists a sequence $(y_k, z_k) \in \bigcup_{J\in\mathcal{J}} \bigcap_{j\in J} \eta_j(U_j)$ that converges to it. Because $\mathcal{J}$ is finite, we pick a subsequence and, without loss of generality, assume that $(y_k, z_k) \in \bigcap_{j\in J} \eta_j(U_j)$ for all $k \in \mathbb{N}$.
Let g correspond to this J as before. We use continuity of g to see that

$$f^*(x,y) = z \iff f^*(x_k, y_k) \longrightarrow z$$

$$\iff \frac{1}{2}g(y_k, f^*(x_k, y_k)) \longrightarrow |x| \iff |x| = \frac{1}{2}g(y,z).$$

Therefore we can define g on all of $\eta^*(\overline{\Omega}^*)$ and this function is continuous.
In particular, since g is Lipschitz continuous on each of the finitely many $\bigcap_{j\in J} \eta_j(U_j)$ and continuous as a whole, it is Lipschitz continuous. Then $\left[(y,z) \mapsto (\frac{1}{2}g(y,z), y)\right]$ is Lipschitz continuous as well. As mentioned before, we will use that Lipschitz continuous functions take zero sets to zero sets.
To complete the proof, observe that for any $M \subseteq \mathbb{R}^{d-1} \times \mathbb{R}$,

$$\eta^{*-1}(M) = \{(x,y) \in \overline{\Omega}^* : \exists (y', z') \in M : (y', z') = (y, f^*(x,y))\}$$

$$= \{(x,y) \in \overline{\Omega}^* : \exists z \in \mathbb{R} : (y,z) \in M \wedge |x| = \frac{1}{2}g(y,z)\}$$

$$= \{(\frac{1}{2}g(y,z), y) : (y,z) \in M\} \cup \{(-\frac{1}{2}g(y,z), y) : (y,z) \in M\}.$$

Therefore η^{*-1} takes zero sets to zero sets. $\square$

3.2 The Pólya–Szegő inequality for $k \geq 1$

As a first step, we will generalize *theorem 3.2* to $W^{1,p}_{0,\geq}(\Omega)$ for some open Ω. This is possible because the piecewise affine functions are dense.

Lemma 3.3. *Let Ω be open. Then the piecewise affine functions with compact support in $\overline{\Omega}$ are dense in $W^{1,p}_{0,\geq}(\overline{\Omega})$ for $1 \leq p < \infty$.*

Proof. We will only sketch this proof because it is a standard argument. We already know that $C^\infty_c(\mathbb{R}^d)$ is a dense subspace, so we assume f to be from this space. Now we pick a fine rectilinear grid so that the cells are hypercubes. By connecting the vertices of the hypercube to its midpoint, we can split each cell into $2d$ simplices. Then we define the approximating function on each simplex as the linear interpolation, where each vertex of the simplex is assigned a value sufficiently close to the value of f at this vertex to ensure that the total approximation error is small. It is also possible to do this while ensuring that the interpolating function's graph does not lie flat in the x-direction. $\square$

The following corollary and its proof closely resemble Kawohl's corollary 2.32 in [4]. The key property that we need for our transition from piecewise affine functions to $W^{1,p}_{0,\geq}(\overline{\Omega})$ is for the functional $\int_{\overline{\Omega}} F(\nabla f)dm^d$ to be sequentially weakly lower semicontinuous. We gain this through convexity of F. This is a standard result in the Calculus of Variations and can be found in [2]. More precisely: If $F : \mathbb{R}^d \longrightarrow [0, \infty)$ is a convex function such that J_F is a well-defined functional on $W^{1,p}_0(\mathbb{R}^d)$ for $1 \leq p < \infty$, then J_F is sequentially weakly lower semicontinuous. Note that this remains true when we use J_F as a functional on $W^{1,p}_{0,\geq}(\overline{\Omega})$ for some open Ω.

Corollary 3.4 (modified from [4])**.** *Let $*$ denote the Steiner symmetrization in* 1 *dimension, and let Ω be open. Let F be as in* theorem 3.2 *with the further requirement that the functional $J_F : W^{1,p}(\mathbb{R}^d) \longrightarrow [0, \infty)$, defined by the formula*

$$J_F(f) := \int_{\mathbb{R}^d} F(\nabla f)dm^d,$$

is well-defined and continuous. Finally, let $f \in W^{1,p}_{0,\geq}(\overline{\Omega})$ for some $1 \leq p < \infty$. Then the symmetrized function exists and $f^* \in W^{1,p}_{0,\geq}(\overline{\Omega}^*)$. Furthermore, we have the inequality

$$J_F(f^*) \leq J_F(f).$$

Proof. By the aforementioned result, we know that J_F is sequentially weakly lower semicontinuous. Let $f_n \overset{W^{1,p}}{\longrightarrow} f$ be a sequence of piecewise linear functions with compact support in $\overline{\Omega}$. Continuity of the Steiner symmetrization in L^p directly implies $f_n^* \overset{L^p}{\longrightarrow} f^*$. Note that $f_n^*, f^* \in L^p(\overline{\Omega}^*)$. Furthermore, the functions f_n^* are still Lipschitz continuous, so they are in $W^{1,p}_0(\overline{\Omega}^*)$. Then *theorem 3.2* applied to $|\cdot|^p$ and F implies that both

$$\|\nabla f_n^*\|_{L^p} \leq \|\nabla f_n\|_{L^p} \text{ and } J_F(f^*) \leq J_F(f).$$

From the bound on the norm of the gradient we gets a bound on each derivative, so $\|f_n^*\|_{W^{1,p}}$ is bounded. Then Banach-Alaoglu implies that there exists some $\tilde{f} \in W^{1,p}_0(\overline{\Omega}^*)$

such that, after replacing the sequence by a subsequence, $f_n^* \overset{W^{1,p}}{\rightharpoonup} \tilde{f}$. In particular $f_n^* \overset{L^p}{\rightharpoonup} \tilde{f}$, and as a consequence $f^* = \tilde{f} \in W^{1,p}_{0,\geq}(\overline{\Omega}^*)$ almost everywhere. Because the functional is sequentially weakly lower semicontinuous and $f_n^* \overset{W^{1,p}}{\rightharpoonup} f^*$, we can conclude that

$$J_F(f^*) \leq \liminf J_F(f_n^*) \leq \lim J_F(f_n) = J_F(f).$$

□

We are now, at last, able to prove the theorem that was our initial goal. The proof is similar to the proof of the previous corollary. All we need to do is use *corollary 3.4* together with *theorem 2.15*, the approximation of S^k by several applications of S^1. Since we need to apply *corollary 3.4* to Steiner symmetrizations in 1 dimension in arbitrary directions, F has to fulfill the requirements of this theorem in every direction. In other words, F conjugated with arbitrary rotations has to fulfill the requirements. This implies that F has to be radially symmetric.

Theorem 3.5. *Let $1 \leq k \leq d$, $1 \leq p < \infty$, and let Ω be open. Let $F : \mathbb{R}^d \longrightarrow [0,\infty)$ be a convex, radially symmetric, and radially nondecreasing function which fulfills the requirement that the functional $J_F : W^{1,p}(\mathbb{R}^d) \longrightarrow [0,\infty)$, defined by the formula*

$$J_F(f) = \int_{\overline{\Omega}} F(\nabla f)dm^d,$$

is well-defined and continuous. Finally, let $f \in W^{1,p}_{0,\geq}(\overline{\Omega})$. Then the symmetrized function exists and $S^k f \in W^{1,p}_{0,\geq}(S^k\overline{\Omega})$. Furthermore, we have the inequality

$$J_F(S^k f) \leq J_F(f).$$

Proof. We use *theorem 2.15* to find a sequence of concatenations of Steiner symmetrizations in one dimension S_n such that $S_n f =: f_n \overset{L^p}{\longrightarrow} S^k f$. We now apply *corollary 3.4* with the functionals defined by $|\cdot|^p$ and F to this sequence. This yields the inequalities

$$\|\nabla f_n\|_{L^p} \leq \|\nabla f\|_{L^p} \text{ and } J_F(f_n) \leq J_F(f).$$

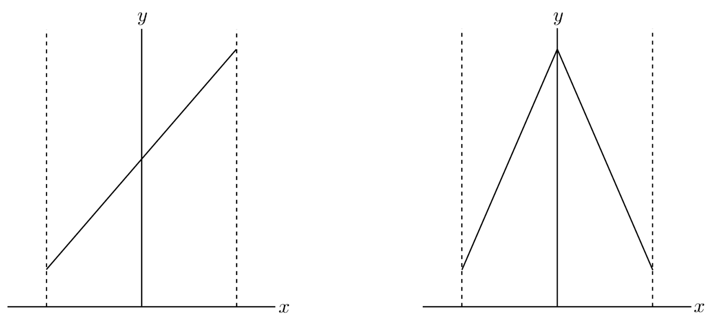
It then follows that $\|f_n\|_{W^{1,p}}$ is bounded. The theorem of Banach-Alaoglu implies that there exists some $\tilde{f} \in W^{1,p}_0(\mathbb{R}^d)$ such that, after passing to a subsequence, we have $f_n \overset{W^{1,p}}{\rightharpoonup} \tilde{f}$. In particular, $f_n^* \overset{L^p}{\rightharpoonup} \tilde{f}$ and therefore $S^k f = \tilde{f} \in W^{1,p}_{0,\geq}(S^k\overline{\Omega})$ almost everywhere. Since the functional J_F is sequentially weakly lower semicontinuous, $f_n \overset{W^{1,p}}{\rightharpoonup} S^k f$ implies

$$J_F(S^k f) \leq \liminf J_F(f_n) \leq J_F(f).$$

□

Remark 3.6. *It is somewhat vacuous to consider the Sobolev space $W^{1,p}_0(\overline{\Omega})$ in the previous corollary and theorem instead of just extending all functions by zero and working with $W^{1,p}(\mathbb{R}^d)$. The only thing we gain from introducing a domain $\overline{\Omega}$ is the statement that $S^k f$ will be supported in $S^k\overline{\Omega}$, which is not a particular strong or surprising observation.*

It should, however, be noted that there is *a case where the domain of the functions matters: One might try to extend the Pólya–Szegő inequality to the case of functions in* $W^{1,p}_{\geq}(\overline{\Omega})$.
What happens here is that we can not just extend functions onto all of $\mathbb{R}^d$*, as the shape of the domain influences the symmetrized function. It turns out that there are easy counterexamples, when no futher restrictions are made, for the Pólya–Szegő inequality in this case. The figures given below sketch a situation where the rearranged function clearly contains more gradient, in the sense that* $\|\nabla f^*\|_{L^p} > \|\nabla f\|_{L^p}$*, than the original function.*



(a) A function f and its domain. (b) The function f^* and its domain.

3.3 An application: The Nash inequality

The Pólya–Szegő inequality can be used to reduce the proofs Sobolev- and Poincaré-like inequalities to the case of functions which are the result of Steiner symmetrizations (and therefore have symmetries). We formulate this method in a general way in the following lemma.

Lemma 3.7. *Let* $n \in \mathbb{N}$ *and* $1 \leq p_1, ..., p_n < \infty$. *Consider an inequality of this shape: For all* $f \in W^{1,p_1}(\mathbb{R}^d) \cap ... \cap W^{1,p_n}(\mathbb{R}^d)$*, the following inequality holds:*

$$A\left(\|f\|_{L^{p_1}}, ..., \|f\|_{L^{p_n}}\right) \leq B\left(\|\nabla f\|_{L^{p_1}}, ..., \|\nabla f\|_{L^{p_n}}, \|f\|_{L^{p_1}}, ..., \|f\|_{L^{p_n}}\right).$$

Here A *and* B *are real-valued, nonnegative functions, and* B *is monotonous in the first* n *components. It is then sufficient to prove the inequality for the case of radially symmetric and radially decreasing functions.*

Proof. Set $* = S^d$. It is sufficient to consider nonnegative functions because neither the norm of the gradient, nor the function change when taking the absolute value. We assume that the inequality has been proven for the case of radially symmetric and radially decreasing functions, in particular for f^* whenever $f \in W^{1,p_i}(\mathbb{R}^d)$, $1 \leq i \leq n$. Because the functionals $\|\cdot\|_{L^p}$, for $1 \leq p < \infty$, fulfill the requirements of *theorem 3.5*, we can apply it:

$$\begin{aligned} A\left(\|f\|_{L^{p_1}}, ..., \|f\|_{L^{p_n}}\right) &= A\left(\|f^*\|_{L^{p_1}}, ..., \|f^*\|_{L^{p_n}}\right) \\ &\leq B\left(\|\nabla f^*\|_{L^{p_1}}, ..., \|\nabla f^*\|_{L^{p_n}}, \|f^*\|_{L^{p_1}}, ..., \|f^*\|_{L^{p_n}}\right) \\ &\leq B\left(\|\nabla f\|_{L^{p_1}}, ..., \|\nabla f\|_{L^{p_n}}, \|f\|_{L^{p_1}}, ..., \|f\|_{L^{p_n}}\right). \end{aligned}$$

□

In this situation, we only use the Pólya–Szegő inequality for the case $k = d$, which is probably its most useful version. Since we want to make the proof as easy as possible, it only makes sense to apply the "strongest" one of our symmetrizations, turning the function into a radially symmetric and decreasing one. The fact that we lose a lot of information about the function does not matter here.
There might be other situations, however, where one does want to retain certain features of the original function when rearranging it. In this case applying a Steiner symmetrization in less than d dimensions might simplify the function while keeping the desired properties. To say it briefly, a Steiner symmetrization in 1 dimension introduces only 1 new symmetry, but little information about the function is lost. Conversely, a symmetrization in d dimensions yields a strong symmetry, but a lot of information is lost.

Let us apply this method to the example of the Nash inequality with a sharp constant. The formulation of the theroem and proof are taken from [5] (theorem 8.13, page 220), with modifications such as omitting the proof of sharpness of the inequality. Before we get to the theorem, we need to define a certain constant. Recall that we use the notation $f_A := \frac{1}{m^d(A)} \int_A f dm^d$. We define

$$\lambda := \min_{f \in W^{2,2}(B_1)} \frac{\|\nabla f\|_{L^2}^2}{\|f - f_{B_1}\|_{L^2}^2} > 0.$$

The fact that the minimum is positive follows from the classical Poincaré inequality. It can be found in [5] as theorem 8.11 on page 218.

Theorem 3.8 (Nash inequality, modified from [5]). *Let $f \in W^{1,2}(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. Then*

$$\|f\|_{L^2}^{1+\frac{2}{d}} \leq C_d \|\nabla f\|_{L^2} \|f\|_{L^1}^{\frac{2}{d}},$$

where

$$C_d = \left(\left(\frac{2}{d} \right)^{\frac{1}{2}} \right) \lambda^{-\frac{1}{2}} \omega_d^{-\frac{1}{d}}.$$

Equality only occurs when f is, after possibly scaling and rotating it, a specific function related to the minimizer of the previously mentioned minimization problem.

Proof. The method which we formulated in *lemma 4.1* can be used in the proof of this inequality. We therefore assume f to be nonnegative, as well as radially symmetric and radially decreasing. Let $R > 0$. We will later choose this R to be optimal. Consider the following:

$$\|f\|_{L^2}^2 = \|f\|_{L^2(B_R)}^2 + \|f\|_{L^2(B_R^c)}^2 = \|f - f_{B_R} + f_{B_R}\|_{L^2(B_R)}^2 + \int_{B_R^c} f^2 dm^d$$

$$= \|f - f_{B_R}\|_{L^2(B_R)}^2 + 2 \int_{B_R} (f - f_{B_R}) f_{B_R} dm^d + \|f_{B_R}\|_{L^2(B_R)}^2 + \int_{B_R^c} f^2 dm^d.$$

First of all, note that the second term is zero. Further, it follows from the definition of λ and a scaling argument, to get from B_R to B_1, that

$$\|f - f_{B_R}\|^2_{L^2(B_R)} \leq \frac{R^2}{\lambda}\|\nabla f\|^2_{L^2(B_R)}.$$

We will apply this to the first term. Lastly, f being radially decreasing implies that $f\big|_{B_R^c} \leq f_{B_R}$. We apply these results and get

$$\|f\|^2_{L^2} \leq \frac{R^2}{\lambda}\|\nabla f\|^2_{L^2(B_R)} + m^d(B_R) f^2_{B_R} + f_{B_R}\int_{B_R^c} f\, dm^d.$$

Now we will add something nonnegative on the right. This might seem like this stops the inequality from being sharp, but in the sharp case what we add is actually zero.

$$\leq \frac{R^2}{\lambda}\|\nabla f\|^2_{L^2} + \frac{1}{m^d(B_R)}\left(\left(\int_{B_R} f\, dm^d\right)^2 + 2\int_{B_R} f\, dm^d \int_{B_R^c} f\, dm^d + \left(\int_{B_R^c} f\, dm^d\right)^2\right)$$

$$= \frac{R^2}{\lambda}\|\nabla f\|^2_{L^2} + \frac{R^{-d}}{\omega_d}\|f\|^2_{L^1}.$$

The only thing that is left to do is to find an R which minimizes the right hand side, thereby making our inequality as strong as possible. For both large and small R the value goes to infinity, so the minimum is at the point where the derivative with respect to R is zero. This is the case when

$$0 = 2\frac{R}{\lambda}\|\nabla f\|^2_{L^2} - d\frac{R^{-(d+1)}}{\omega_d}\|f\|^2_{L^1}$$

$$\iff R^{d+2} = \frac{d\lambda}{2\omega_d}\frac{\|f\|^2_{L^1}}{\|\nabla f\|^2_{L^2}}.$$

Plugging this into our previous inequality yields

$$\|f\|^2_{L^2} \leq \left(\frac{d\lambda}{2\omega_d}\frac{\|f\|^2_{L^1}}{\|\nabla f\|^2_{L^2}}\right)^{\frac{2}{d+2}} \frac{\|\nabla f\|^2_{L^2}}{\lambda} + \left(\frac{d\lambda}{2\omega_d}\frac{\|f\|^2_{L^1}}{\|\nabla f\|^2_{L^2}}\right)^{\frac{-d}{d+2}} \frac{\|f\|^2_{L^1}}{\omega_d}.$$

$$= \left(\frac{d}{2\omega_d}\right)^{\frac{2}{d+2}} \lambda^{\frac{-d}{d+2}} \|f\|^{\frac{4}{d+2}}_{L^1} \|\nabla f\|^{\frac{2d}{d+2}}_{L^2} + \left(\frac{d\lambda}{2}\right)^{\frac{-d}{d+2}} \omega_d^{\frac{-2}{d+2}} \|f\|^{\frac{4}{d+2}}_{L^1} \|\nabla f\|^{\frac{2d}{d+2}}_{L^2}.$$

Now we take both sides to the power $\frac{d+2}{2d}$ and arrive at the Nash inequality:

$$\|f\|^{\frac{d+2}{d}}_{L^2} \leq \left(\left(\frac{d}{2}\right)^{\frac{1}{d}} + \left(\frac{2}{d}\right)^{\frac{1}{2}}\right) \lambda^{-\frac{1}{2}} \omega_d^{-\frac{1}{d}} \|\nabla f\|_{L^2} \|f\|^{\frac{1}{d}}_{L^1}.$$

As mentioned before, we omit the proof of sharpness. □

4 Conclusion

We did not choose the fastest path that arrives at some useful version of the Pólya–Szegő inequality. This would have probably been some proof of the inequality for the Schwarz symmetrization with the functional being the L^p-norm of the gradient. Instead, we took a slower, geometric approach which lead to a more general statement. In our version of the inequality the rearrangement is some intermediary between the standard Steiner symmetrization and the Schwarz symmetrization, and the functional is from a larger class. Along the way, we have shown that any compact set can be turned into a circle by a sequence of Steiner symmetrizations. To do this, we studied sets with the "symmetric cuboid property", which could be seen as a weakened form of convexity that can be easily gained through Steiner symmetrizations.

We aimed to present the proofs in a detailed manner so that the reader does not have to do the majority of the work himself. The downside of this is that some conciseness was lost. In particular, our proof of *theorem 3.2* is certainly far longer than other proofs of this statement. There are two reasons for this: Firstly, we wanted the proof to be easy to visualize, allowing the reader to develop more intuition as to why the proof works. Secondly, by looking at small pieces of the piecewise affine function where "everything is linear", we avoided a certain argument with implicit differentiation. It is up to the reader to decide if this was worth the added verbosity compared to the original proof in [4] (theorem 2.31, page 83).

References

[1] A. Buchard. "Steiner Symmetrization is Continuous in $W^{1,p}$". In: *Geometric & Functional Analysis GAFA* 7 (1997), pp. 823–860. DOI: https://doi.org/10.1007/s000390050027.

[2] Bernard Dacorogna. *Direct Methods in the Calculus of Variations*. Applied Mathematical Sciences. Springer-Verlag New York, 2001. ISBN: 978-1-4419-2259-5. URL: https://www.springer.com/gp/book/9780387357799.

[3] Herbert Federer. *Geometric Measure Theory*. Classics in Mathematics. Springer-Verlag Berlin Heidelberg, 2001. ISBN: 978-3-540-60656-7. URL: https://www.springer.com/de/book/9783540606567.

[4] Bernhard Kawohl. *Rearrangements and Convexity of Level Sets in PDE*. Lecture Notes in Mathematics. Springer-Verlag, 1985. ISBN: 3-540-15693-3.

[5] E.H. Lieb et al. *Analysis*. Crm Proceedings & Lecture Notes. American Mathematical Society, 2001. ISBN: 9780821827833. URL: https://books.google.de/books?id=Eb%5C_7oRorXJgC.