GAFA, Geom. funct. anal.
Vol. 7 (1997) 1 – 38
1016-443X/97/050001-38 $ 1.50+0.20/0

GAFA Geometric And Functional Analysis

STEINER SYMMETRIZATION IS CONTINUOUS IN $W^{1,p}$

A. Burchard

Dedicated to the memory of Frederick J. Almgren Jr.

Abstract

We study the continuity, smoothing, and convergence properties of Steiner symmetrization in higher space dimensions. Our main result is that Steiner symmetrization is continuous in $W^{1,p}$ $(1 \leq p < \infty)$ in all dimensions. This implies that spherical symmetrization cannot be approximated in $W^{1,p}$ by sequences of Steiner symmetrizations. We also give a quantitative version of the standard energy inequalities for spherical symmetrization.

1 Introduction

Steiner symmetrization was invented as a tool for a geometric proof of the isoperimetric inequality. The isoperimetric inequality says that among all bodies of a given volume, the ball has the smallest perimeter; or, in the language of rearrangements, that the perimeter of a body can only decrease under spherical symmetrization. Steiner observed that the perimeter of a body is generally larger than the perimeter of a related body of the same volume which is symmetric at a hyperplane; or, as we would say, that the perimeter of the body decreases under Steiner symmetrization. Since the perimeter strictly decreases under symmetrization unless the body is symmetric at the hyperplane to begin with, it follows that a body which minimizes perimeter for a given volume must be symmetric at all hyperplanes, and hence a ball [Ste].

This simple and convincing argument, however, does not show that a perimeter-minimizing body exists. The problem can be overcome by constructing a sequence of Steiner symmetrizations that approximates spherical symmetrization, and then using continuity of the volume and lower semicontinuity of the perimeter with respect to that convergence [CStu].

In the first half of the century, rearrangements were used to find the optimal shape of a body of a given size in a variety of geometric and physical problems [Bl],[HaLitP],[Lu],[PSz],[R],[S]. In the 1970s, interest in rearrangements was renewed, as mathematicians began to look for geometric proofs of functional inequalities [Be1],[BrLi],[Cr],[FFl],[Li2],[Ma],[Sp],[T2] (see also [Ba],[BuZ],[Ch],[K],[Zi]). Rearrangements were generalized from smooth or convex bodies to measurable sets and to functions in Sobolev spaces [Hi],[T1]. Besides the characterization of the cases of equality of the new, more general rearrangement inequalities [BrotZi],[HO, Li1], the main technical point was the approximation of spherical symmetrization in function spaces by sequences of simpler rearrangements [BrLiLut],[H].

In this paper, we address two related questions about Steiner symmetrization as a transformation on the Sobolev spaces $W^{1,p}(\mathbf{R}^{n+1})$ for $1 \le p < \infty$ and $n \ge 1$. First, is it continuous? Secondly, how closely can sequences of Steiner symmetrizations approximate spherical symmetrization? We also discuss some refinements of the standard energy inequalities.

It is well known that Steiner and spherical symmetrization respect L^p and $W^{1,p}$ spaces, preserve L^p norms, decrease L^p distances, and are smoothing in the sense that they reduce $W^{1,p}$-norms and surface areas of graphs [AlLi],[BTa],[BrotZi],[Hi]. (Higher Sobolev spaces are not preserved [K] (see also [DSt])). However, since they are neither linear, nor bounded as transformations on $W^{1,p}$, nor spatially localized, continuity questions are subtle.

There are two results in the literature concerning the continuity of rearrangements in Sobolev spaces. Coron proved that symmetrization in one space dimension is continuous in $W^{1,p}$ [Co], and Almgren and Lieb proved that spherical symmetrization in all dimensions higher than one is discontinuous [AlLi]. Clearly, then, symmetrizations along subspaces of dimension greater than one, such as Schwarz symmetrization in dimensions three and above, cannot be continuous either. We resolve the remaining case here: we show that Steiner symmetrization is continuous as a transformation from $W^{1,p}(\mathbf{R}^{n+1})$ to itself for all $n \ge 0$.

It turns out that the question how well sequences of Steiner symmetrization can approximate spherical symmetrization is closely related to the continuity question. It is well known that the spherical symmetrization of a nonnegative function can be approximated in $L^p(\mathbf{R}^{n+1})$ by a sequence of Steiner symmetrizations and rotations [CStu] (for modern proofs see for example [BrLiLut],[BuZ]). The continuity of Steiner symmetrization together with the discontinuity result by Almgren and Lieb implies that the approximating sequence will in general not converge in $W^{1,p}$. For sufficiently smooth functions, however, the approximating sequence will converge in $W^{1,p}$.

Even for general functions in $W^{1,p}$, however, sequences of Steiner symmetrizations and rotations can approximate spherical symmetrization remarkably well: Under these sequences, nonnegative functions in $W^{1,p}$ converge to their spherical symmetrizations in L^p (and hence weakly in $W^{1,p}$) in such a way that the angular part of the gradients converges to zero in L^p; in other words, the level sets and their perimeters converge to balls and their perimeters.

The smoothing properties of rearrangements play a central role in our proofs. We use a local version, due to Alvino, Trombetti and Lions [AlvTrLio], of the well-known statement that gradient norms cannot increase under the standard rearrangements to show that energy cannot accumulate in small sets under sequences of symmetrizations. Finally, we strengthen the results of Brothers and Ziemer [BrotZi] by finding a lower bound for the difference between the gradient norms of a function and its rearrangement.

Acknowledgments. The problem was initiated by Thomas Lachand-Robert [L], who suggested the link between the continuity and approximation questions mentioned above. I am very grateful to Fred Almgren, Elliott Lieb, and Michael Loss for mentioning the problem to me, and for many useful discussions. They have carried me through this project. Special thanks to Michael Loss for a proof of the "only if" part of Theorem 2, which made Lachand-Robert's idea precise. Although I choose to present a different proof here, the original proof played an important role in the development of ideas. Finally, many thanks to Bernd Kawohl for pointing out an error in an earlier version, and to Friedemann Brock for drawing my attention to the results of Alvino, Trombetti, and Lions in [AlvTrLio].

2 Statement of the Results

We begin with some definitions. Let A be a measurable set in $\mathbf{R}^{n+1}$, $n \geq 1$. The *spherical symmetrization*, A^*, of A is the open ball centered at the origin which has the same Lebesgue measure as A. To define *Steiner symmetrization*, we write points in $\mathbf{R}^{n+1}$ as pairs (x, y) with $x \in \mathbf{R}$ and $y \in \mathbf{R}^n$. The Steiner symmetrization, $\mathcal{S}A$, of A is the set whose one-dimensional cross sections parallel to the x-axis are centered open intervals whose lengths equal the measures of the corresponding cross sections of A. In short, if the cross section

$$A(y) := \{x \in \mathbf{R} \mid (x,y) \in A\}$$

has finite measure, we set

$$(\mathcal{S}A)(y) = (A(y))^*\,,$$

where * denotes symmetrization of the cross section in $\mathbf{R}$; if $A(y)$ is not measurable or does not have finite measure, the corresponding cross section of $\mathcal{S}A$ is defined to be $\mathbf{R}$.

Assume that f is a function whose *level sets*

$$E(h) := \{x \in \mathbf{R}^n \mid f(x) > h\}$$

have finite measure for all positive heights h. We say that another function, g, is a *rearrangement* of f, or *equimeasurable* with f, if almost all level sets of g have the same measure as the corresponding level sets of f, that is, if f and g have the same *distribution function*

$$\rho(h) := |E(h)|\ . \tag{2.1}$$

We define the spherical symmetrization f^* and the Steiner symmetrization $\mathcal{S}f$ of f using the *layer-cake representation*

$$f(x) = \int \mathcal{X}_{E(h)}\,\mathrm{d}h\,; \tag{2.2}$$

that is, we set

$$f^*(x) = \int \mathcal{X}_{E(h)^*}\,\mathrm{d}h\,, \quad \mathcal{S}f(x) = \int \mathcal{X}_{\mathcal{S}E(h)}\,\mathrm{d}h\,. \tag{2.3}$$

By definition, f^* and $\mathcal{S}f$ are equimeasurable with f. We will often define a **slice** of a function f between two heights $h_1 < h_2$ by

$$\tilde{f}(x) := \big(\min\{f(x), h_2\} - h_1\big)_+\,. \tag{2.4}$$

If f is in $W^{1,p}$, so is $\tilde{f}$; it is bounded by $h_2 - h_1$, and vanishes outside a set of finite measure. By definition, $\widetilde{f^*} = (\tilde{f})^*$, and $\widetilde{\mathcal{S}f} = \mathcal{S}\tilde{f}$.

We say that a nonnegative measurable function f *vanishes at infinity*, if all level sets at positive heights h differ from bounded sets by sets of measure zero. We will need this assumption for our convergence results.

As mentioned in the introduction, Steiner symmetrization preserves L^p-norms and acts as a contraction on L^p for all positive p. Integrals of convex functionals of $|\nabla f|$ never increase under Steiner symmetrization; in particular, $W^{1,p}$ norms of functions $(1 \le p \le \infty)$, surface areas of graphs, and perimeters of sets can only decrease. However, since Steiner symmetrization is not a linear transformation, these statements imply nothing about continuity with respect to gradient norms – and, indeed, spherical symmetrization, which shares all these properties, is discontinuous as a transformation on $W^{1,p}$ in any dimension greater than one [AlLi]. Our main result is the following Theorem 1.

Theorem 1 (Continuity). *Steiner symmetrization is continuous in $W^{1,p}(\mathbf{R}^{n+1})$ for $1 \leq p < \infty$ in all dimensions $n \geq 1$. That is, for every sequence of nonnegative functions*

$$f_k \to f \text{ in } W^{1,p} \Longrightarrow \mathcal{S} f_k \to \mathcal{S} f \text{ in } W^{1,p}\,.$$

The key to the proof of Theorem 1 is the observation that Steiner symmetrization preserves the measure of the set of critical points (see Lemma 4.3 and Corollary 4.5) – in contrast with spherical symmetrization in dimensions two and above, which, in general, shrinks the set of critical points.

Almgren and Lieb proved that spherical symmetrization is discontinuous at a function f precisely if the set of critical points of f^* has smaller measure than the set of critical points of f [AlLi]. Such functions form a dense subspace of $W^{1,p}$ in all dimensions greater than one; however, spherical symmetrization is continuous at sufficiently smooth functions and at radial functions.

In view of the discontinuity result of Almgren and Lieb and the continuity statement of Theorem 1, it is natural to suspect that spherical symmetrization cannot be approximated in the strong $W^{1,p}$-topology by a sequence of Steiner symmetrizations and rotations. The following theorem confirms this suspicion.

Theorem 2 (Approximation of spherical symmetrization by Steiner symmetrizations). *Let f be a nonnegative function in $W^{1,p}(\mathbf{R}^{n+1})$ that vanishes at infinity. There exists a sequence of successive Steiner symmetrizations and rotations $\{f_k\}_{k\geq 0}$ of f which approximates f^* in $W^{1,p}$, if and only if spherical symmetrization is continuous at f.*

In spite of Theorem 2, sequences of Steiner symmetrizations can approximate spherical symmetrization quite well. For instance, the angular component of the gradient converges to zero.

Theorem 3 (Convergence of the angular component of the derivative). *Let f be a nonnegative measurable function in $W^{1,p}(\mathbf{R}^{n+1})$ for some $n > 1$*

and $p \geq 1$. Assume that f vanishes at infinity. There exists a sequence of functions $\{f_k\}_{k\geq 0}$ which is constructed from f by a sequence of Steiner symmetrizations and rotations so that

$$\big\| \, |\nabla f_k| - \partial f_k/\partial r \big\|_p \to 0 \qquad (n \to \infty)\,.$$

Theorem 3 is equivalent to the statement that the level sets of f converge to balls in such a way, that also their perimeters converge to the perimeters of the corresponding balls (see Proposition 7.1). In other words, for bounded measurable sets, the approximation of spherical symmetrization by a sequence of Steiner symmetrizations is as good as one may hope for.

We begin the main part of the paper in section 3 with a discussion of the smoothing properties of Steiner and spherical symmetrization. Proposition 3.1 says that these symmetrizations never increase the average energy density in a set of a given size. This is a useful reformulation of the standard energy inequalities. We use it in the proofs of all three theorems to show convergence of sequences of functions constructed by rearrangements; it serves as a substitute for a majorizing function in Lebesgue's dominated convergence theorem.

In section 4, we develop our principal tool for the proof of Theorem 1, a distribution function, and study its properties. We pay special attention to the effect of Steiner symmetrization on sets where partial derivatives vanish. In Corollary 4.5 we show that the set of critical points does not shrink under Steiner symmetrization. At the beginning of the section, we explain how the work of Almgren and Lieb [AlLi] motivates our approach. Proposition 4.1 gives a quantitative version of the strict rearrangement inequality for convex gradient integrals proved by Brothers and Ziemer [BrotZi]. The first half of the paper ends with the proof of Theorem 1 in section 5. Our proof strongly relies on Coron's continuity result for symmetrization in one dimension [Co].

We then turn to the approximation results. We begin section 6 by combining Theorem 1 with the discontinuity result of Almgren and Lieb (or alternately Corollary 4.5 with Proposition 4.1) to show that, in general, the spherical symmetrization cannot be approximated by Steiner symmetrizations. This establishes the "only if" part of Theorem 2. We also show that a sequence which approximates the spherical rearrangement of a given function must certainly satisfy the conclusions of Theorem 3 and Proposition 7.1, that is, the angular derivative converges to zero, and the perimeter of every level set converges to the perimeter of a ball. We also give some notation and define the sequences of Steiner symmetrizations and rotations that we use in the following two sections. In section 7, we prove Theorem 3 and related statements about $W^{1,1}$-norms and perimeters of level sets. We conclude with the proof of Theorem 2 in section 8.

3 Smoothing Properties

Proposition 3.1 is a local version of the well-known fact that gradient norms cannot increase under symmetrization. It was first proved in a more general form by Alvino, Trombetti, and Lions [AlvTrLio].

PROPOSITION 3.1 (Local smoothing property). *Let f be a nonnegative function in $W^{1,p}(\mathbf{R}^n)$ (where $1 \leq p < \infty$, and $n \geq 1$), and let $\mathcal{S}f$ and f^* be the Steiner and spherical symmetrizations of f. For every $\varepsilon > 0$ and every convex function F with $F(0) = 0$, $F(z) \geq 0$ for $z \geq 0$, we have the inequalities*

$$
\begin{aligned}
\sup_{|E|\leq\varepsilon} \int_E F\big(|\nabla f^*(x)|\big)\mathrm{d}x &\leq \sup_{|E|\leq\varepsilon} \int_E F\big(|\nabla \mathcal{S}f(x)|\big)\mathrm{d}x \\
&\leq \sup_{|E|\leq\varepsilon} \int_E F\big(|\nabla f(x)|\big)\mathrm{d}x\,,
\end{aligned} \tag{3.1}
$$

where $|E|$ denotes the Lebesgue measure of E. In particular, if the integral on the right is finite, then the other two integrals are also finite.

REMARK. (i) The usual smoothing statements [BrotZi],[Du],[K] are recovered by taking $\varepsilon \to \infty$; note that the integrals are infinite unless $F(0) = 0$. The choices $F(z) = z^p$ and $F(z) = \sqrt{1+z^2} - 1$ give the statements that $W^{1,p}$-norms and surface areas can only decrease under symmetrization.

(ii) Alvino, Trombetti, and Lions showed in Proposition 2.1 of [AlvTrLio] that inequality (3.1) is in fact equivalent to the standard global energy inequalities.

Proof. Following Ahlfors [A] and Baernstein and Taylor [BTa], we define a simple rearrangement $\mathcal{T}f$ of a measurable function f. Fix a hyperplane H that does not pass through the origin. Denote the half-space containing the origin by H^+, the other half space by H^-, and let $\mathcal{R}$ be the reflection at the hyperplane. Set

$$
\mathcal{T}f(x) := \begin{cases} \max\{f(x), f(\mathcal{R}x)\} & \text{if } \; x \in H^+ \\ \min\{f(x), f(\mathcal{R}x)\} & \text{if } \; x \in H^- \end{cases}
$$

Then it is easy to see that

$$
\sup_{|E|\leq\varepsilon} \int_E F\big(|\nabla \mathcal{T}f(x)|\big)\mathrm{d}x = \sup_{|E|\leq\varepsilon} \int_E F\big(|\nabla f(x)|\big)\mathrm{d}x\,, \tag{3.2}
$$

in fact, $|\nabla \mathcal{T} f|$ and $|\nabla f|$ are equimeasurable. This rearrangement appears under may names in the literature [A],[BTa],[Ben],[Be2],[BoLe],[BroSo], [FrFu],[HaLitP],[So]; we will refer to it as the *two-point rearrangement* of f at the hyperplane H. The two-point rearrangements of a nonnegative function f in $W^{1,p}$ at a suitable sequence of hyperplanes converge to the spherical symmetrization f^* strongly in L^p, and weakly in $W^{1,p}$ [BTa] (see also [B],[BroSo],[So]). Similarly, the two-point rearrangements at a suitable sequence of parallel hyperplanes converge to the Steiner symmetrization $\mathcal{S}f$. To show the inequality for f^*, fix $\varepsilon > 0$, and let E be a set for which the supremum on the left-hand side of (3.1) is achieved, and consider the restriction of the sequence of rearranged functions to E. Since the functional

$$f \mapsto \int_E F\big(|\nabla f(x)|\big) \mathrm{d}x$$

is convex because F is convex and nondecreasing by assumption, it is lower semiconinuous, and the claim follows from (3.2). □

It follows immediately from Proposition 3.1 that for every sequence $\{f_k\}$ obtained from a function f in $W^{1,p}$ by a sequence of Steiner symmetrizations and rotations, we have

$$\int |\nabla f_k|^p \mathcal{X}_{\{|\nabla f_k| > M\}} \mathrm{d}x \le \sup_{|E| \le (\|\nabla f_k\|_p / M)^p} \int_E |\nabla f_k|^p \, \mathrm{d}x \to 0 \quad (M \to \infty)$$

uniformly in k. Note that this property is certainly necessary for a sequence to converge in $W^{1,p}$: If $\{g_k\}_{k\ge 0}$ is a sequence of nonnegative functions in $L^p(\mathbf{R}^d)$, then

$$\text{(3.3)} \qquad \sup_{|E| \le \varepsilon} \int_E (g_k(x))^p \mathrm{d}x = \int_{E^*} (g_k^*(x))^p \mathrm{d}x \to 0 \qquad (\varepsilon \to 0)\,,$$

by the convergence of the sequence $\{g_k^*\}_{k\ge 0}$ in L^p. Here g_k^* is the spherical symmetrization of g_k, and E^* is the centered ball of measure ε.

The following Lemma 3.2 will be used several times in the proof of the continuity and convergence results. We only state it for Steiner symmetrization, but it holds equally for spherical symmetrization and any rearrangement that can be obtained as a limit of two-point rearrangements in L^p. The lemma is motivated by Theorems 7.2 and 7.5 of Almgren and Lieb [AlLi], which correspond to the special case with $f_k = \tilde{f}_k^*$. We prove it simply as a corollary of Proposition 3.1 .

LEMMA 3.2 (Equivalence of norms). *Let f_k ($k \ge 0$) and g be nonnegative functions in $W^{1,p}(\mathbf{R}^n)$ ($1 \le p < \infty$). Assume that each function f_k is the* *result of a (finite or L^p-convergent) sequence of Steiner symmetrizations and rotations of a function $\tilde f_k$, and let g be a measurable function with gradient ∇g. Then*

$$\left.\begin{array}{lll} \tilde f_k \to \tilde g & \text{in } W^{1,p} \\ f_k \to g & \text{in measure} \\ \nabla f_k \to \nabla g & \text{pointwise a.e.} \end{array}\right\} \implies f_k \to g \quad \text{in} \quad W^{1,p} .$$

In particular, convergence in $W^{1,q}$ for some q implies convergence in $W^{1,p}$.

Proof. If functions and their gradients are uniformly bounded and vanish outside a set of finite measure, the conclusion holds by dominated convergence

In general, fix $0 < h_1 < h_2$, let B be the set where $g > h_1/2$, and write each function f_k as the sum of three slices

$$f_k = \min\{f_k, h_1\} + \big(\min\{f_k, h_2\} - h_1\big)_+ + (f_k - h_2)_+ ,$$

and similarly for g. The bottom and top slices are small in $W^{1,p}$ uniformly in k for h_1 small, h_2 large enough, because the f_k are equimeasurable with the $\tilde f_k$ which form a convergent sequence in $W^{1,p}$ by assumption.

The restrictions of the middle slices to B converge in L^p by dominated convergence. The restrictions of the middle slices to the complement of B converge to zero in measure, and hence in L^p.

Similarly, the sequence of truncated gradients

$$\nabla f_k \mathcal{X}_{\{|\nabla f_k| < M\}} \mathcal{X}_B$$

converges in L^p. By Proposition 3.1, the error term satisfies

$$\begin{aligned} \big\| \nabla f_k \mathcal{X}_{\{|\nabla| f > M\}} \big\|_p^p &\le \sup_{|E| \le (\|\nabla f_k\|_p / M)^p} \int_E |\nabla f_k|^p \, \mathrm{d}x \\ &\le \sup_{|E| \le (\|\nabla \tilde f_k\|_p / M)^p} \int_E |\nabla \tilde f_k|^p \mathrm{d}x , \end{aligned}$$

which converges to zero uniformly in k as $M \to \infty$ by equation (3.3). Again, the contribution of the middle slices outside B is small if k is large enough by Proposition 3.1 and equation (3.3). □

4 Distribution Functions

In preparation for the proof of Theorem 1, we recall the techniques developed by Almgren and Lieb in their proof of the discontinuity of spherical symmetrization on $W^{1,p}$ [AlLi]. By definition, the spherical symmetrization of a function is determined by its distribution function (2.1). The properties of spherical symmetrization in L^p are easily understood from the layer-cake decompositions (2.2) and (2.3). However, it is not so easy use information about the level sets to answer questions concerning the gradient. The principal tool for such questions is the *co-area formula*, which says that for every measurable function g,

$$\int_0^\infty \int_{\partial E(h)} g \, \mathrm{d}s \, \mathrm{d}h = \int_{\mathbf{R}^n} g \, |\nabla f| \, \mathrm{d}x \, ,$$

where $\mathrm{d}x$ denotes integration in $\mathbf{R}^n$, and $\mathrm{d}s$ integration with respect to $n-1$-dimensional Hausdorff measure on the boundary of the level set of f at height h. In the general co-area formula ([F, Theorem 4.5.9]), which holds for functions of bounded variation, the domain of the inner integral is $\{x \in \mathbf{R}^n \mid \lambda(x) \leq h \leq \mu(x)\}$, where $\lambda(x)$ and $\mu(x)$ are the upper and lower approximate limits of f at x. We argue as Brothers and Ziemer (see [BrotZi, Section 2]), that, since for f in $W^{1,p}$ the approximate limits λ and μ differ only on a set of measure zero, we may integrate instead over $\partial E(h)$ (or, alternately, over $f^{-1}(h)$) if we choose f to coincide with $\lim_{\varepsilon \to 0} |B_\varepsilon|^{-1} \int_{B_\varepsilon(x)} f(x') \mathrm{d}x'$.

It follows with monotone convergence that

$$\int_0^\infty \int_{\partial E(h)} g \, |\nabla f|^{-1} \, \mathrm{d}s \, \mathrm{d}h = \lim_{\varepsilon \to 0} \int_0^\infty \int_{\partial E(h)} g \big(|\nabla f| + \varepsilon \big)^{-1} \mathrm{d}s \, \mathrm{d}h$$

$$= \int_{\mathbf{R}^n} g \, \mathcal{X}_{\{\nabla f \neq 0\}} \mathrm{d}x \, . \tag{4.1}$$

Note that the co-area formula gives no information about the values of g on the set of critical points of f.

Since the distribution function of f is a nonnegative nonincreasing function, it defines a positive measure on $\mathbf{R}^+$. Almgren and Lieb write the distribution function as a sum of three parts: The *co-area distribution function*

$$\rho_{reg}(h) := \big| \{x \in \mathbf{R}^n \mid f(x) > h \, , \; \nabla f(x) \neq 0\} \big| \tag{4.2}$$

is the contribution of the set of regular points of f to the distribution function. It is always absolutely continuous with respect to Lebesgue measure. The name is motivated by the formula

$$\rho_{reg}(h) = \int_h^\infty \int_{\partial E(h')} |\nabla f|^{-1} \, \mathrm{d}s \, \mathrm{d}h' \, . \tag{4.3}$$

The contribution of the critical points, the *residual distribution function*

$$\rho_{crit}(h) := \big| \{x \in \mathbf{R}^n \mid f(x) > h \, , \; \nabla f(x) = 0\} \big| \, , \tag{4.4}$$

is split into a singular component which grows on a set of measure zero whose pre-image has positive measure, and an absolutely continuous component which corresponds to the part of the set of critical values that is smeared out continuously over the heights. Spherical symmetrization is discontinuous at those functions whose residual distribution function contains an absolutely continuous component; these functions are called *co-area irregular*. Functions where spherical symmetrization is continuous are called *co-area regular*.

Sufficiently smooth functions are always co-area regular, because their critical values form a set of measure zero by the Morse-Sard-Federer theorem. The required smoothness, however, depends on the dimension. Functions of a single variable and radial functions in $W^{1,p}$ are always co-area regular (hence Coron's continuity result for $n = 1$), but for $n > 1$, co-area irregular functions are dense in $W^{1,p}(\mathbf{R}^n)$. Note that the co-area and residual distribution functions may change under equimeasurable rearrangements, even though the distribution function and its absolutely continuous and singular components are preserved. Spherical symmetrization is discontinuous at co-area irregular functions because it removes the critical points that produce the absolutely continuous component of the residual distribution function.

We illustrate the role of the different distribution functions with a variation of a sharp rearrangement result by Brothers and Ziemer ([BrotZi, Theorem 1.1]).

PROPOSITION 4.1 (Quantitative rearrangement inequality for convex gradient integrals). *Let F be a strictly convex nonnegative function on $\mathbf{R}^+$ with $F(0) = 0$. For every function f in $W^{1,p}$, equality in*

$$\int F(|\nabla f|)\mathrm{d}x \geq \int F(|\nabla f^*|)\mathrm{d}x$$

implies that f is co-area regular, and that the level sets of f at almost all heights are balls. In particular, we have for $F(z) = z^p$

$$\|\nabla f\|_p^p - \|\nabla f^*\|_p^p \geq \int_0^\infty \frac{\sigma(h)^p}{|\mathrm{d}/\mathrm{d}h\, \rho_{reg}(h)|^{p-1}} - \frac{\sigma^*(h)^p}{|\mathrm{d}/\mathrm{d}h\, \rho(h)|^{p-1}}\mathrm{d}h\,,$$

and for $F(z) = \sqrt{1+z^2} - 1$ $(1 \leq p \leq 2)$

$$\Psi(f) - \Psi(f^*)r \geq \int_0^\infty \left(\sqrt{\sigma(h)^2 + (\mathrm{d}/\mathrm{d}h\, \rho_{reg}(h))^2} - |\mathrm{d}/\mathrm{d}h\, \rho_{reg}(h)|\right) - \left(\sqrt{\sigma^*(h)^2 + (\mathrm{d}/\mathrm{d}h\, \rho(h))^2} - |\mathrm{d}/\mathrm{d}h\, \rho(h)|\right) \mathrm{d}h\ .$$

Here, $\sigma(h)$ and $\sigma^*(h)$ are the perimeters of the level sets of f and f^* at h, the distribution function ρ and ρ_{reg} are defined by (2.1) and (4.2), and $\Psi(f) = \int F(|\nabla f|)\,\mathrm{d}x$ is a substitute for the $W^{1,1}$-norm.

Proof. With the co-area formula, we write

$$\begin{aligned}\int F(|\nabla f|)\mathrm{d}x &= \int_0^\infty \int_{\partial E(h)} F(|\nabla f|)\,|\nabla f|^{-1}\,\mathrm{d}s\,\mathrm{d}h\\ &= \int_0^\infty \int_{\partial E(h)} G(|\nabla f|^{-1})\mathrm{d}s\,\mathrm{d}h\,,\end{aligned}$$

where $G(z) = zF(z^{-1})$ is positive, strictly convex, and decreasing, and $F(z) = zG(z^{-1})$. Jensen's inequality gives

$$\begin{aligned}\int_{\partial E(h)} G(|\nabla f|^{-1})\mathrm{d}s &\geq \sigma(h)G\Big(\sigma(h)^{-1}\int_{\partial E(h)} |\nabla f|^{-1}\,\mathrm{d}s\Big)\\ &= \sigma(h)G\big(\sigma(h)^{-1}(-\mathrm{d}/\mathrm{d}h\,\rho_{reg}(h))\big)\\ &\geq \sigma^*(h)G\big(\sigma^*(h)^{-1}(-\mathrm{d}/\mathrm{d}h\,\rho(h))\big)\,,\end{aligned}$$

where we have used (4.3) in the second line. The last inequality follows with the monotonicity and convexity properties of G from $\sigma^*(h) \leq \sigma(h)$ (with equality only if the level set of f at h is a ball), and $|\mathrm{d}/\mathrm{d}h\,\rho_{reg}(h)| \leq |\mathrm{d}/\mathrm{d}h\,\rho(h)|$ (with equality for almost all h only if f is co-area regular). The claim follows since f^* produces equality in Jensen's inequality. □

In the proof of Theorem 1, we adapt the approach of Almgren and Lieb to Steiner symmetrization. Let f be a function on $\mathbf{R}^{n+1}$ with arguments (x, y), where $x \in \mathbf{R}$, and $y \in \mathbf{R}^n$. We view the Steiner symmetrization as the symmetrization of the family of one-dimensional cross sections $f(\cdot, y)$ parameterized by the transverse coordinate y. The relevant distribution function is

$$\tilde\rho(y, h) = \big|\{x \in \mathbf{R} \mid f(x, y) > h\}\big|\,,$$

where $|\cdot|$ denotes Lebesgue measure in $\mathbf{R}$; the Steiner symmetrization of f is determined by $\tilde\rho$. In other words, any property of f that can be formulated in terms of $\tilde\rho$ is preserved under Steiner symmetrization. We will frequently use the one-dimensional analogue of (4.1)

$$\int_0^\infty \sum_{f(\xi,y)=h} g(\xi, y) = \int_{\mathbf{R}} g(x, y)\mathcal{X}_{\{\partial/\partial x\, f(x,y)\neq 0\}}\mathrm{d}x\,. \tag{4.5}$$

Since for each fixed y, the distribution function $\tilde\rho(\cdot, y)$ is a nonincreasing function, it defines a positive measure on $\mathbf{R}^+$. We decompose this measure into a part that is absolutely continuous with respect to Lebesgue measure, and a singular part which is supported on the critical values of $f(\cdot, y)$. In contrast with the decomposition by Almgren and Lieb, this decomposition depends only on the distribution function $\tilde{\rho}$, and is accordingly preserved by Steiner symmetrization. We will see in section 5 that the difficulty in proving continuity of Steiner symmetrization is concentrated on the support of the singular part of the measure. We describe this support in the following two lemmas.

LEMMA 4.2 (Critical values). *Let f be a nonnegative function in $W^{1,p}(\mathbf{R})$, and and let μ^s be the singular part of the measure induced by its distribution function ρ on $\mathbf{R}^+$. Then, for every Borel set B,*

$$\mu^s(B) = \left|\{x \in f^{-1}(B) \mid \mathrm{d}/\mathrm{d}x\, f(x) = 0\}\right| = \left|f^{-1}(B \cap C)\right|,$$

where C is the support of the singular measure μ^s.

REMARK. (i) We will often refer to C as the set of *critical values* of f. Note that the usual definition as the image of the set of critical points makes sense only for differentiable functions.

(ii) The co-area regular functions on $\mathbf{R}^n$ are characterized by the analogous property that

$$\mu^s(B) = \left|\{x \in f^{-1}(B) \mid \nabla f(x) = 0\}\right| = \left|f^{-1}(B \cap C)\right|$$

for all Borel sets B, where C is the support of the singular part of the measure induced by the distribution function of f.

Proof. The claim follows immediately from two facts: a) The critical values form a Borel set of measure zero by the Morse-Sard-Federer theorem ([F, Theorem 3.4.3]); b) The derivative of a continuously differentiable function in $W^{1,p}$ vanishes almost everywhere on the inverse image of a set of measure zero ([AlLi, Theorem 3.1]; see also [LiLo, Theorem 6.19]). □

The essential step in our proof of Theorem 1 will be to understand the transverse derivatives $\nabla_y f$ on the set where $\partial f/\partial x$ vanishes. It is easy to see that the transverse derivatives are constant almost everywhere on a set where $f(\cdot, y)$ is constant, that is, for almost all $y \in \mathbf{R}^n$, the set

$$\{(x, x') \mid f(x, y) = f(x', y)\,,\ \nabla_y f(x, y) \neq \nabla_y f(x', y)\}$$

has measure zero, because the gradient of $F(x, x', y) = f(x, y) - f(x', y)$ vanishes almost everywhere on $F^{-1}(0)$. In the proof of Theorem 1, we will need the following stronger version of this observation.

LEMMA 4.3 (Key lemma). *Let f be a nonnegative function in $W^{1,p}(\mathbf{R}^{n+1})$, B a Borel set in $\mathbf{R}$, and C_1 and C_2 disjoint Borel sets in $\mathbf{R}^n$. For each*

$y \in \mathbf{R}^n$, define two measures on $\mathbf{R}^+$ by

$$\mu_1(y,B) := \big|\{x \in \mathbf{R} \mid f(x,y) \in B\,,\ \nabla_y f(x,y) \in C_1\}\big|$$
$$\mu_2(y,B) := \big|\{x \in \mathbf{R} \mid f(x,y) \in B\,,\ \nabla_y f(x,y) \in C_2\}\big|$$

where $\nabla_y f$ consists of the last n components of the gradient of f, and let $\mu_1^s(y)$ and $\mu_2^s(y)$ be the singular parts of these measures. Then, for almost all y, $\mu_1^s(y)$ and $\mu_2^s(y)$ are mutually singular.

Proof. The idea is that the map $(x,y) \to (f(x,y),y)$ transforms a neighborhood of a point where $\partial f/\partial x = 0$ into a set that has a hyperplane as a tangent space. The hyperplane is determined by the transverse derivative of f. Two such hyperplanes can meet at more than a point only if the transverse derivatives coincide.

Taking countable intersections and unions, we may assume that C_1 and C_2 are disjoint coordinate half-spaces. Integrating out $n-1$ variables, we may also assume that $n=1$, so that

$$\mu_1(y,B) := \big|\{x \in \mathbf{R} \mid f(x,y) \in B\,,\ \partial/\partial y\, f(x,y) > \alpha\}\big|$$
$$\mu_2(y,B) := \big|\{x \in \mathbf{R} \mid f(x,y) \in B\,,\ \partial/\partial y\, f(x,y) < \beta\}\big|$$

with $\alpha > \beta$. We consider only the case $\alpha = 1$, $\beta = -1$. The general case follows by scaling and changing $f(\cdot,y)$ to $f(\cdot,y) + \gamma y$ with a constant γ.

We will show that the measures defined on $\mathbf{R}^2$ by

$$\mu_1^s(B) := \int_{\mathbf{R}} \mu_1^s(y, B(y))\mathrm{d}y\,, \quad \mu_2^s(B) := \int_{\mathbf{R}} \mu_2^s(y, B(y))\mathrm{d}y$$

are mutually singular. Here, B is a measurable set in $\mathbf{R}^2$, and $B(y)$ is its cross section at y. In general, two measures μ_1 and μ_2 are mutually singular, if their overlap

$$\mu_1 \wedge \mu_2 := \inf_{\cup B_i = \mathbf{R}} \sum_i \min\big\{\mu_1(B_i), \mu_2(B_i)\big\}$$

vanishes (the infimum is taken over all finite disjoint unions). In our special case, we have

$$\mu_1^s \wedge \mu_2^s = \int \mu_1^s(y) \wedge \mu_2^s(y)\mathrm{d}y\,.$$

Assume for the moment that f is continuously differentiable. Then the sets

$$D_1 := \big\{(x,y) \in \mathbf{R}^2 \mid \partial/\partial y\, f(x,y) > 1\big\}$$
$$D_2 := \big\{(x,y) \in \mathbf{R}^2 \mid \partial/\partial y\, f(x,y) < -1\big\}$$

are open. We can cover D_1 and D_2 with countably many closed squares that satisfy the following compatibility condition: Any pair of squares can be subdivided into a finite number of smaller squares, so that the projections of each pair of smaller squares to the y-axis either coincide, or intersect in at most one point. (Such a covering can be constructed with dyadic subdivisions of a fixed grid on $\mathbf{R}^2$.)

We will show that the singular part of the measures induced by the restrictions of f to any pair of squares $Q_1 \subset D_1$ and $Q_2 \subset D_2$ are mutually singular. This will prove the claim when f is continuously differentiable.

Let Q_1 and Q_2 be a pair of squares from the covering of D_1 and D_2 that have the same projection to the y-axis. (For pairs of squares whose projections intersect in at most one point there is nothing to show.) By Lemma 4.2, the singular parts of the measures induced by the restrictions of f to Q_1 and Q_2 are just the measures induced by the restrictions of f to the subsets Q_1^s and Q_2^s, respectively, on which $\partial f/\partial x$ vanishes.

Fix $\varepsilon > 0$. Since Q_1^s and Q_2^s are compact, we can bound $|\partial f/\partial x|$ in a δ-neighborhood of Q_1^s and Q_2^s by ε if δ is small enough. We cut the squares Q_1 and Q_2 into vertical strips of equal height δ. Fix a pair of such strips, one in Q_1, and one in Q_2. By definition of Q_1 and Q_2, the images of these strips under $(x, y) \mapsto (y, f(x, y))$ satisfy the hypotheses of Lemma 4.4 proved below. If the intersection of the image strips contains a point (y_0, h) where y is a critical value for the restriction of $f(\cdot, y_0)$ to the strip in Q_1 as well as for the restriction to the strip in Q_2, then, by our choice of δ, the width of the image strips at y_0 is at most $\varepsilon\delta$. By Lemma 4.4, two strips can intersect only for y in an interval of length at most $\varepsilon\delta$. The contribution of any pair of strips at a given y to the overlap $\mu_1^s(y) \wedge \mu_2^s(y)$ is at most δ. Hence, that the total contribution of the strips to $\mu_1^s \wedge \mu_2^s$ is at most $\varepsilon\delta^2$. If the intersection of the image strips contains no such point, then the two strips do not contribute to $\mu_1^s \wedge \mu_2^s$ by Lemma 4.2. Summing over the strips, we see that the contribution of Q_1 and Q_2 to the overlap is bounded above by εl^2, where l is the side length of Q_1 and Q_2. This completes the proof for differentiable f.

To finish the proof for general $f \in W^{1,p}$, note that for any $\varepsilon > 0$, we can find a continuously differentiable function $\tilde{f}$ that differs from f on a set of measure less than ε (see [F, Theorem 3.1.16]). The measures $\tilde{\mu}_1^s$ and $\tilde{\mu}_2^s$ induced by $\tilde{f}$ differ from μ_1^s and μ_2^s by at most ε, and the same is true for their overlap. Since ε was arbitrary, the claim is proven. □

LEMMA 4.4 (Intersection of strips). *Let* $[a_1(y), b_1(y)]$ *and* $[a_2(y), b_2(y)]$ *be* *two families of intervals, where* $a_1(y) - y$ *and* $b_1(y) - y$ *are monotonically increasing, and* $a_2(y) + y$ *and* $b_2(y) + y$ *are monotonically decreasing for* y *in a given interval. Then* $[a_1(y), b_1(y)]$ *and* $[a_2(y), b_2(y)]$ *intersect for* y *in an interval of length at most* $(b_1(y_0) - a_1(y_0) + b_2(y_0) - a_2(y_0))/2$*, where* y_0 *is any point for which the intersection is not empty.*

Proof. We may assume that $y_0 = 0$. The intersection is empty unless $b_2(y) \geq a_1(y)$ and $b_1(y) \geq a_2(y)$. The monotonicity assumptions imply that for the intersection to be nonempty, we need $b_2(0) - y \geq a_1(0) + y$ for $y \geq 0$, and $b_1(0) + y \geq a_2(0) - y$ for $y \leq 0$. □

Lemma 4.3 immediately implies that the set of critical points cannot shrink under Steiner symmetrization.

COROLLARY 4.5 (Steiner symmetrization preserves the measure of the set of critical points). *Let* f *be a nonnegative function in* $W^{1,p}(\mathbf{R}^{n+1})$*, and let* B *be Borel set in* $\mathbf{R}^+$*. Then*

$$\begin{aligned} &\left|\{x \in \mathbf{R} \mid \mathcal{S}f(x,y) \in B\,,\ \nabla \mathcal{S}f(x,y) = 0\}\right| \\ &\qquad = \left|\{x \in \mathbf{R} \mid f(x,y) \in B\,,\ \nabla f(x,y) = 0\}\right| \end{aligned}$$

for almost all $y \in \mathbf{R}^n$*. In particular, Steiner symmetrization preserves the co-area distribution function* ρ_{reg} *and the residual distribution function function* ρ_{crit} *of* f *defined by equations* (4.2) *and* (4.4)*.*

Proof. Let C be a Borel set in $\mathbf{R}^n$, and let μ_C^s be the singular part of the measure induced by the restriction of f to the set where $\nabla_y f \in C$. Lemma 4.3 is equivalent to the more general statement that μ_C^s does not change under Steiner symmetrization, that is,

$$\begin{aligned} &\left|\{x \in \mathbf{R} \mid f(x,y) \in B\,,\ \partial/\partial x\, f(x,y) = 0\,,\ \nabla_y f(x,y) = 0\}\right| \\ &\qquad = \left|\{x \in \mathbf{R} \mid \mathcal{S}f(x,y) \in B\,,\ \partial/\partial x\, \mathcal{S}f(x,y) = 0\,,\ \nabla_y \mathcal{S}f(x,y) = 0\}\right| \end{aligned}$$

for almost all y. □

5 Continuity

Our strategy for the proof of Theorem 1 is to reduce all considerations to calculations with functions of a single variable. We think of the given function $f(x, y^1, \ldots, y^n)$, its partial derivatives, as well as $\mathcal{S}f$ and its partial derivatives as functions of x that depend on the parameters $y = (y^1, \ldots, y^n)$. We use Coron's result to show that the partial derivative in the x-direction transforms continuously under Steiner symmetrization.

For the partial derivatives in the transverse directions, we find a formula that can be calculated for a fixed value of y directly from the restriction of f and its partial derivatives to the cross section at y. At regular points of the restriction of $\mathcal{S}f$ to the cross section, this becomes an explicit formula, and we can easily prove continuity. To understand the behavior at critical points, we use our results from section 4.

Since the Steiner symmetrization of f is defined by rearranging the cross sections $f(\cdot,y)$, the partial derivative $\partial/\partial x\,\mathcal{S}f(\cdot,y)$ can be calculated from $f(\cdot,y)$ without any information on neighboring cross sections. The formula is

$$\frac{\partial}{\partial x}\mathcal{S}f(x,y)=2\Big(\sum_{\xi:f(\xi,y)=f(x,y)}\left(\partial/\partial x\,f(\xi,y)\right)^{-1}\Big)^{-1}, \tag{5.1}$$

with the convention that $\partial/\partial x\,\mathcal{S}f(x,y)=0$ whenever $\partial/\partial x\,f(\xi,y)=0$ for one of the terms in the sum (see [Co]). By the rearrangement inequality for the derivative in one dimension,

$$\int\left|\frac{\partial}{\partial x}\mathcal{S}f(x,y)\right|^p\mathrm{d}x\le\int\left|\frac{\partial}{\partial x}f(x,y)\right|^p\mathrm{d}x\,.$$

In the next lemma, we prove the less obvious fact that also the transverse derivatives $\nabla_y\mathcal{S}f(\cdot,y)$ are determined by the values of f and its partial derivatives on the cross section. Our formula says that, for almost all cross sections y, the transverse derivatives of $\mathcal{S}$ at a point (x,y) are the averages of the transverse derivatives of f over the set where $f(\cdot,y)=f^*(x,y)$; in other words, it is the expected value of the transverse derivative on the cross section, conditioned on the value of f.

LEMMA 5.1 (The transverse derivative of $\mathcal{S}f$). *Let f be a nonnegative function in $W^{1,p}$ with Steiner symmetrization $\mathcal{S}f$. Then*

$$\int\nabla_y\mathcal{S}f(x,y)\mathcal{X}_{\{\mathcal{S}f(x,y)\in B\}}\mathrm{d}x=\int\nabla_yf(x,y)\mathcal{X}_{\{f(x,y)\in B\}}\mathrm{d}x \tag{5.2}$$

and

$$\int|\nabla_y\mathcal{S}f(x,y)|^p\,\mathcal{X}_{\{\mathcal{S}f(x,y)\in B\}}\mathrm{d}x\le\int|\nabla_yf(x,y)|^p\,\mathcal{X}_{\{f(x,y)\in B\}}\mathrm{d}x \tag{5.3}$$

for almost all y, and all Borel sets $B\in\mathbf{R}^+$. Moreover, $\nabla_y(\mathcal{S}f)(x,y)$ depends on x only through $f(x,y)$, that is, there exists a function $\alpha:\mathbf{R}\to\mathbf{R}^n$

so that

$$\nabla_y \mathcal{S}f(x,y) = \alpha(y, f(x,y)) \text{ almost everywhere}. \tag{5.4}$$

Properties (5.2) *and* (5.4) *uniquely determine* $\nabla_y \mathcal{S}f$ *in* $L^p(\mathbf{R}^n)$.

Proof. We may again assume that $n = 1$. To show formula (5.2), it is sufficient to consider the case where B is an interval, $B = (h_1, h_2)$ with $0 < h_1 < h_2$. The slice $\tilde{f}$ defined by (2.4) satisfies

$$\int \frac{\partial}{\partial y} \tilde{f}(x,y) \mathrm{d}x = \int \frac{\partial}{\partial y} f(x,y) \mathcal{X}_{\{f(x,y) \in B\}} \mathrm{d}x \,,$$

and similarly for the rearrangements. But for almost every y,

$$\int \frac{\partial}{\partial y} \tilde{f}(x,y) \mathrm{d}x = \lim_{\delta \to 0} \int \frac{\tilde{f}(x,y+\delta) - \tilde{f}(x,y-\delta)}{2\delta} \mathrm{d}x \,,$$

which is unchanged under Steiner symmetrization since $\mathcal{S}f(\cdot,y)$ is equimeasurable with $f(\cdot,y)$. Inequality (5.3) holds because $\mathcal{S}$ acts as a contraction on $L^p(\mathbf{R})$.

To see the formula (5.4), note that the set where $\mathcal{S}f(\cdot,y)$ takes a given value h consists either of two points, or of two intervals of equal length. If the set consists of two points, (5.4) automatically holds by the reflection symmetry of $\mathcal{S}f$. If the set has positive measure, the transverse derivative $\partial f/\partial y$ is equal to a constant almost everywhere on the set by Lemma 4.3. Inequality (5.3) and identity (5.2) imply that $\partial \mathcal{S}f/\partial y$ on the set where $\mathcal{S}f(\cdot,y) = h$ equals the same constant.

Uniqueness follows from formulas (5.2) and (5.4) by choosing B to be a small interval of the form $(h-\varepsilon, h+\varepsilon)$ and taking the limit $\varepsilon \to 0$. □

As discussed above, the point of Lemma 5.1 is that the transverse derivatives $\nabla_y \mathcal{S}f(\cdot,y)$ can be defined for fixed y using only information on $f(\cdot,y)$ and $\nabla_y f(\cdot,y)$. We introduce new notation to suppress the dependence on y. For every $\phi \in W^{1,p}(\mathbf{R})$, define an operator $\mathcal{T}(\phi,\cdot)$ on $L^p(\mathbf{R})$ by

$$\int_{(\mathcal{S})\phi^{-1}(B)} \mathcal{T}(\phi,\psi)(x) \mathrm{d}x = \int_{\phi^{-1}(B)} \psi(x) \mathrm{d}x \tag{5.5}$$

for all Borel sets $B \subset \mathbf{R}^+$, and the requirement that for every ϕ and ψ there exists a measurable function α so that

$$\mathcal{T}(\phi,\psi) = \alpha \circ \phi \quad \text{in } W^{1,p}. \tag{5.6}$$

By Lemma 5.1, $\mathcal{T}(f(\cdot,y), \partial/\partial y^i f(\cdot,y))$ coincides with $\partial/\partial y^i \mathcal{S}f(\cdot,y)$ for almost every y. By construction, $\mathcal{T}(\phi,\cdot)$ is linear and order-preserving on L^p. It is a contraction, because the value of $\mathcal{T}(\phi,\psi)$ on the set where of

$\mathcal{S}\phi = h$ is a weighted average of the values of ψ on the set where $\phi = h$. Moreover, if $\psi = \alpha \circ \phi$, then $\mathcal{T}(\phi, \psi) = \psi$.

LEMMA 5.2 (Continuity of $\mathcal{T}$). *The operator $\mathcal{T}$ defined by equations* (5.5) *and* (5.6) *has the property that for every pair of sequences $\phi_k \to \phi$ in $W^{1,p}(\mathbf{R})$, $\psi_k \to \psi$ in $L^p(\mathbf{R})$, there exists a subsequence (again denoted by (ϕ_k, ψ_k)) so that*

$$\mathcal{T}(\phi_k, \psi_k)\mathcal{X}_{\{\mathrm{d}\mathcal{S}\phi/\mathrm{d}x \neq 0\}} \to \mathcal{T}(\phi, \psi)\mathcal{X}_{\{\mathrm{d}\mathcal{S}\phi/\mathrm{d}x \neq 0\}} \qquad (k \to \infty)$$

pointwise almost everywhere.

REMARK. It is not hard to see that the sequence $\{\mathcal{T}(\phi_k, \psi_k)\mathcal{X}_{\{\mathrm{d}\mathcal{S}\phi/\mathrm{d}x \neq 0\}}\}_{k \geq 0}$ actually converges in L^p. In particular, if the derivative of ϕ vanishes on a set of measure zero, then $\mathcal{T}$ is continuous at (ϕ, ψ) for any ψ. If it vanishes on a set of positive measure, then $\mathcal{T}$ is continuous at (ϕ, ψ) only if the two functions satisfy certain compatibility conditions. We will use Lemma 4.3 to show that these compatibility conditions are always satisfied if ψ is the transverse derivative of ϕ.

Proof. Since $\mathcal{T}(\phi, \cdot)$ is a contraction on L^p, we need to prove only that

$$\mathcal{T}(\phi_k, \psi)\mathcal{X}_{\{\mathrm{d}\mathcal{S}\phi/\mathrm{d}x \neq 0\}} \to \mathcal{T}(\phi, \psi)\mathcal{X}_{\{\mathrm{d}\mathcal{S}\phi/\mathrm{d}x \neq 0\}} \text{ in } L^p \quad (k \to \infty)\,,$$

where ψ is a smooth function with compact support. Fix a regular point x of $\mathcal{S}\phi$, that is, a point where the derivative is defined and nonzero. Then $\phi^{-1}(h)$, where $h = \mathcal{S}\phi(x)$, is discrete. Taking the limit $\varepsilon \to 0$ in formulas (5.5) and (5.6) with $B = (h - \varepsilon, h + \varepsilon)$ gives

(5.7)

$$\mathcal{T}(\phi, \psi)(x) = \Bigg(\sum_{\xi \in \phi^{-1}(h)} \big(\mathrm{d}/\mathrm{d}x\, \phi(\xi)\big)^{-1} \Bigg)^{-1} \sum_{\xi \in \phi^{-1}(h)} \big(\mathrm{d}/\mathrm{d}x\, \phi(\xi)\big)^{-1} \psi(\xi)\,.$$

Let $\xi_1, \ldots, \xi_l$ be the points in $\phi^{-1}(h)$, and let $I_1, \ldots, I_l$ be small open intervals containing them. Choose δ so small that $\phi^{-1}(h - \delta, h + \delta)$ is contained in those intervals. For the given sequence $\{\phi_k\}$ that converges to ϕ in $W^{1,p}$ as $k \to \infty$, let $h_k := \mathcal{S}\phi_k(x)$; this converges to h as $k \to \infty$. Moreover, by the locally uniform continuity and the locally uniform convergence of the ϕ_k, we see that $\phi_k^{-1}(h_k)$ is contained in the intervals $I_1, \ldots, I_l$ for k large enough.

Now we apply Coron's result that $\mathcal{S}\phi_k$ converges to $\mathcal{S}\phi$ as $k \to \infty$. By passing to a subsequence, we may assume that the gradients converge pointwise almost everywhere, and in particular, that almost every regular point of $\mathcal{S}\phi$ is a regular point of $\mathcal{S}\phi_k$ for k large enough. Choosing functions

$\phi_{kj} \in W^{1,p}$ $(j=1,\dots,l)$ to coincide with ϕ_k on I_j and with ϕ on I_i $(i \neq j)$, and to take values outside $(h-\delta, h+\delta)$ on the complement of the I_i, Coron's result says that

$$\Big(\sum_i \sum_{\xi \in \phi_{kj}^{-1}(h_k)\cap I_i} \big(\mathrm{d}/\mathrm{d}x\, \phi_k(\xi)\big)^{-1}\Big)^{-1} \to \Big(\sum_i \big(\mathrm{d}/\mathrm{d}x\, \phi(x_i)\big)^{-1}\Big)^{-1} \ (k \to \infty)$$

(see formula (5.1)), that is,

$$\sum_{\xi \in \phi_k^{-1}(h_k)\cap I_j} \big(\mathrm{d}/\mathrm{d}x\, \phi_k(\xi)\big)^{-1} \to \big(\mathrm{d}/\mathrm{d}x\, \phi(x_j)\big)^{-1} \qquad (k \to \infty)\,.$$

Since ψ is continuous by assumption, it follows that

$$\sum_{\xi \in \phi_k^{-1}(h_k)\cap I_j} \big(\mathrm{d}/\mathrm{d}x\, \phi_k(\xi)\big)^{-1}\psi(\xi) \to \big(\mathrm{d}/\mathrm{d}x\, \phi(\xi_j)\big)^{-1}\psi(\xi_j) \qquad (k \to \infty)\,.$$

This proves pointwise convergence almost everywhere on the set of regular points of $\mathcal{S}\phi$. □

Proof of Theorem 1. Let $\{f_k\}_{k\geq 0}$ be a sequence of nonnegative functions on $\mathbf{R}^{n+1}$ that converges to a function f in $W^{1,p}$. We want to show that $\mathcal{S}f_k$ converges to $\mathcal{S}f$ in $W^{1,p}$. Since it is well known that $\mathcal{S}f_k$ converges to $\mathcal{S}f$ in L^p as $k \to \infty$, we see from Lemma 3.2 that we need to show only that there exists a subsequence along which $\nabla \mathcal{S}f_k(\cdot, y)$ converges to $\nabla \mathcal{S}f_k(\cdot, y)$ pointwise almost everywhere for almost every y. Coron's continuity result for symmetrization in $W^{1,p}(\mathbf{R})$ implies the desired convergence for the partial derivative $\partial/\partial x\, \mathcal{S}f(\cdot, y)$ in the direction of symmetrization. To prove the convergence of the transverse partial derivative $\partial/\partial y^i\, f(\cdot, y)$, we need to consider only x and y^i as variables and may treat the other variables y^j $(j \neq i)$ as parameters. This shows that Theorem 1 holds for all $n \geq 1$ provided it holds for $n=1$.

Assume now that $n=1$, that is, f is a function on the plane. Lemma 5.2 says that $\partial/\partial y\, \mathcal{S}f_k(\cdot, y)$ converges to $\partial/\partial y\, \mathcal{S}f(\cdot, y)$ pointwise almost everywhere on the set of regular points of $\mathcal{S}f(\cdot, y)$. We now turn to the convergence on the set of critical points. We can write with the layer-cake principle

$$\frac{\partial}{\partial y} f(x,y) = \int_0^\infty \mathcal{X}_{\{\frac{\partial}{\partial y} f(x,y) > \alpha\}} - \mathcal{X}_{\{\frac{\partial}{\partial y} f(x,y) < -\alpha\}}\, d\alpha$$

so we need to prove

$$\mathcal{T}(f_k, \mathcal{X}_M)\mathcal{X}_{\{\partial \mathcal{S}\phi/\partial x = 0\}} \to \mathcal{T}(f, \mathcal{X}_M)\mathcal{X}_{\{\partial \mathcal{S}\phi/\partial x = 0\}} \text{ in } L^p \qquad (k \to \infty)$$

where M is a level set of $\partial f/\partial y$. Clearly,

$$0 \le \mathcal{T}(f_k, \mathcal{X}_M) \le 1$$

by the properties (5.5) and (5.6). Choose B in (5.2) to be the support of the singular part of the measure induced by the restriction to M. Using Lemma 5.1 and Lemma 4.3, which says that this component of the measure is singular with respect to the components induced by the restriction to the complement of M, we see that

$$\begin{aligned}\int \mathcal{T}(f(\cdot,y),\mathcal{X}_M)\mathcal{X}_{\{\partial/\partial x\, f(x,y)=0\}}\mathrm{d}x &= \int \mathcal{T}(f(\cdot,y),\mathcal{X}_M)\mathcal{X}_{\{\mathcal{S}f(x,y)\in B\}}\mathrm{d}x \\ &= \int \mathcal{X}_M(x)\mathcal{X}_{\{f(x,y)\in B\}}\mathrm{d}x \\ &= \left|\{x \mid f(x,y)\in B\}\right|,\end{aligned}$$

in other words, $\mathcal{T}(f(\cdot,y),\mathcal{X}_M)$ restricted to the set of critical points is a characteristic function. The following Lemma 5.3 with

$$g_k = \mathcal{T}(f_k(\cdot,y),\mathcal{X}_M)\mathcal{X}_{\{\partial/\partial x\, f(\cdot,y)=0\}}$$

completes the proof. □

LEMMA 5.3 (Convergence to characteristic functions). *Let $\{g_k\}_{k\ge 1}$ be a sequence of functions in $L^p(R)$ $(1 \le p < \infty)$ which converges weakly in L^p to a characteristic function $\mathcal{X}_A$. Assume moreover, that $0 \le g(x) \le 1$. Then there exists a subsequence that converges pointwise almost everywhere to $\mathcal{X}_A$. Equivalently, for all sets B of finite measure, $g_k\mathcal{X}_B \to \mathcal{X}_{A\cap B}$ strongly in L^p.*

Proof. Weak convergence of $g_k\mathcal{X}_B$ to $\mathcal{X}_{A\cap B}$ implies that

$$\liminf_{k\to\infty}\int_B (g_k(x))^p\mathrm{d}x \ge |A\cap B| \ ,$$

and the other assumption that

$$\limsup_{k\to\infty}\int_B (g_k(x))^p\mathrm{d}x \le \lim_{k\to\infty}\int g_k(x)\mathcal{X}_B(x)\mathrm{d}x = |A\cap B| \ .$$

For $p>1$, convergence of the norm together with weak convergence implies strong convergence. For $p=1$, we use that, since the g_k are uniformly bounded, and B has finite measure, all L^p- norms are essentially equivalent. □

6 Sequences of Steiner Symmetrizations

We now turn to the problem of approximating spherical symmetrization by a sequence of Steiner symmetrizations. We were led to consider this problem by an argument of Michael Loss [Lo] which showed that the continuity of Steiner symmetrization alone implies that such an approximation is not possible in $W^{1,p}$ for functions at which the spherical rearrangement is discontinuous.

Another argument for the same fact goes as follows. Let $\{f_k\}_{k\geq 0}$ be a sequence of functions obtained from a nonnegative function f in $W^{1,p}$ by a sequence of Steiner symmetrizations and rotations. Since Steiner symmetrization does not change the co-area regular distribution function ρ_{reg} defined by (4.2), Proposition 4.1 implies that

$$\begin{aligned}\|\nabla f_k\|_p^p - \|\nabla f^*\|_p^p &= \int_0^\infty \frac{\sigma_k(h)^p}{|\mathrm{d}/\mathrm{d}h\,\rho_{reg}(h)|^{p-1}} - \frac{\sigma^*(h)^p}{|\mathrm{d}/\mathrm{d}h\,\rho(h)|^{p-1}}\,\mathrm{d}h \\ &\geq \int_0^\infty \sigma^*(h)^p\Big(\frac{1}{|\mathrm{d}/\mathrm{d}h\,\rho_{reg}(h)|^{p-1}} - \frac{1}{|\mathrm{d}/\mathrm{d}h\,\rho(h)|^{p-1}}\Big)\mathrm{d}h\,,\end{aligned}$$

which is strictly positive if $p > 1$ unless f is co-area regular. (For $p = 1$, we use the convex functional Ψ in place of the norm.) This implies, first, that the spherical rearrangement of a co-area irregular function cannot be approximated by Steiner symmetrizations, and secondly, that it is certainly necessary for convergence of an approximating sequence that the level sets converge to balls.

It is known that most sequences of Steiner symmetrizations and rotations produce sequences of functions that converge to the spherical rearrangement in L^p [M], the only exceptions being those that violate an ergodicity condition. We suspect that, similarly, most sequences should give the convergence properties claimed in Theorems 2 and 3, but, for technical reasons, we will choose sequences of symmetrizations that depend on the function being symmetrized.

We find it convenient to change notation (as compared with section 2). We write points in $\mathbf{R}^{n+1}$ as $x = (x^0, \ldots, x^n) = (x^0, \hat{x})$ with $x^0 \in \mathbf{R}$ and $\hat{x} \in \mathbf{R}^n$, and denote Steiner symmetrization with respect to the coordinate axes by $\mathcal{S}^0, \ldots, \mathcal{S}^n$. We define a sequence of symmetrizations corresponding to a given sequence of rotations $\{\mathcal{R}_k\}_{k\geq 1}$ by

$$\mathcal{S}_k := \mathcal{S}^n \cdots \mathcal{S}^0 \mathcal{R}_k\,. \tag{6.1}$$

Let f be a nonnegative measurable function whose level sets have finite measure. We define a sequence of symmetrizations $\{f_k\}_{k\geq 0}$ of f by

$$f_0 = f\,, \quad f_k = \mathcal{S}_k f_{k-1}\,. \tag{6.2}$$

We use the co-area formula to reduce the statements of Theorems 2 and 3 to statements about level sets. If A is a level set of f, then the corresponding level set of f_k is determined by

$$A_0 = f\,, \quad A_k = \mathcal{S}_k A_{k-1}\,. \tag{6.3}$$

By construction,

$$\mathcal{S}^n \cdots \mathcal{S}^0 A = \left\{ x \in \mathbf{R}^{n+1} \;\middle|\; |x^0| < \alpha(\hat{x}) \right\} \tag{6.4}$$

for some function α on $\mathbf{R}^n$ which is nonnegative, symmetric under reflection through coordinate planes, and nonincreasing in the variables $(|x_1|,\ldots,|x_n|)$. If A is contained in a ball of radius R, then α is bounded above by R, and has support in the ball of radius R.

It is well known that there are many sequences of rotations for which the sequence of sets defined by (6.3) converges to a ball with respect to both symmetric difference and Hausdorff distance [H]. For instance if, for all k, $\mathcal{R}_k$ equals a fixed rotation $\mathcal{R}$ which satisfies the ergodicity condition that together with the reflections at the coordinate hyperplanes it generates a dense subset of the rotation group $O(n)$, then then there exists for each $\delta > 0$ a number k_0 (which depends only on the ratio R/r) so that for every subset A of the ball of radius R,

$$B_{r(1-\delta)} \subset A^n \subset B_{r(1+\delta)} \tag{6.5}$$

for all $k \geq k_0$. Here, r denotes the radius of the ball equimeasurable with A.

Similarly, it is clear that for $f \in W^{1,p}$, the sequence $\{f_k\}$ is bounded, and hence has a subsequence that converges weakly in $W^{1,p}$. If we pick the sequence of rotations so that f_k converges to f^* in L^p, the weak $W^{1,p}$-limit must be f^*. Another consequence of the L^p-convergence is the following lemma.

LEMMA 6.1. *Let f be a nonnegative function in L^p, and let $\{f_k\}_{k\geq 0}$ be a sequence of rearrangements of f so that the level sets of f_k converge to the level sets of f^* in the sense of* (6.5). *Let F be a Borel measurable function for which $\int F(f)\mathrm{d}x$ is finite. Then*

$$F(f_k) \;\to\; F(f^*) \quad \text{in} \quad L^1 \qquad (k \to \infty)\,.$$

Proof. Decomposing F into its positive and negative parts and using the layer-cake principle, we see that it is sufficient to prove that

$$f_k^{-1}(B) \;\to\; {f^*}^{-1}(B) \quad \text{in symmetric difference} \quad (k \to \infty) \tag{6.6}$$

for all Borel sets B. Our assumption on the convergence of the level sets implies that (6.6) holds for intervals $B = (a, b]$; it holds for the complement of a set if it holds for the set, and it holds for unions of chains of sets by monotone convergence. Hence it holds for the algebra of finite unions of intervals and their complements. We extend (6.6) to all Borel sets with the monotone class theorem (see, for example, Theorem 1.3 in [LiLo]). $\square$

LEMMA 6.2 (Lipschitz parameterization). *Let A be a subset of the ball of radius R whose boundary is given by monotone functions as in* (6.4). *For $\varepsilon > 0$, let $\hat{C}_\varepsilon$ be the complement of an ε-neighborhood of the coordinate hyperplanes in S^n, and C_ε the corresponding cone in $\mathbf{R}^{n+1}$. There exists a number L (which depends only on ε) so that the boundary of A is given in polar coordinates by a function which is continuous on S^n, and Lipschitz continuous with constant LR on $\hat{C}_\varepsilon$.*

Proof. We may assume that $R = 1$. The angle between the normal of A at a boundary point x with the position vector is at most $\pi/2 - \varepsilon$. The Lipschitz constant is $L = (1 + \delta)\tan(\pi/2 - \varepsilon)$. $\square$

Equation (6.4) also leads to an upper bound for the perimeter.

LEMMA 6.3 (Surface area of the graph of a monotone function). *The surface area of the graph of a monotone function α from the n-dimensional box $[0, a_1] \times \cdots \times [0, a_n]$ to $[0, a_{n+1}]$ is at most $\sum_{i=1}^{n+1} \prod_{j \neq i} a_j$.*

Proof (by induction over the dimension). For $n = 0$, the graph is a single point, and there is nothing to show. Assume the claim has been proven for dimension $n-1$, and let α be a function on the n-dimensional box. We may assume by approximation that α is differentiable and strictly decreasing in the sense that

$$x_i' \le x_i \ (i = 1, \ldots, n), \quad x' \neq x \Longrightarrow \alpha(x') > \alpha(x).$$

With the co-area formula (4.1) we estimate the surface area of the graph of α by

$$\begin{aligned}\int_0^{a_1} \cdots \int_0^{a_n} \sqrt{1 + |\nabla \alpha|^2}\, \mathrm{d}x_1, \ldots, \mathrm{d}x_n &\le \int_0^{a_1} \cdots \int_0^{a_n} 1 + |\nabla \alpha|\ \mathrm{d}x_1 \ldots, \mathrm{d}x_n \\ &= \prod_{i=1}^{n} a_i + \int_0^{a_{n+1}} \int_{\alpha^{-1}(h)} 1\, \mathrm{d}s\, \mathrm{d}h.\end{aligned}$$

Since the boundaries of the level sets of α are again graphs of monotone functions, we can apply the inductive hypothesis to complete the proof. $\square$

7 Convergence of $W^{1,1}$-norms and Perimeters

In this section we prove that for a given nonnegative function f that vanishes at infinity, there are sequences of Steiner symmetrizations and rotations that approach the space of radial functions in the sense that the angular component of the gradient converges to zero. We first reformulate the statement of Theorem 3.

PROPOSITION 7.1 (Convergence of the $W^{1,1}$-norm). *Let f be a nonnegative measurable function in $W^{1,p}(\mathbf{R}^{n+1})$ for some $n \geq 1$ and some $p \geq 1$. Assume that f vanishes at infinity. There exists a sequence of functions $\{f_k\}_{k\geq 0}$ which is obtained from f by a sequence Steiner symmetrizations and rotations as in* (6.2) *so that for every $0 < h_1 < h_2$, the slices of f^* and f_k defined by* (2.4) *satisfy*

$$\tilde{f}_k \to f^* \text{ in } L^1 \,, \quad \tilde{f}_k \rightharpoonup f^* \text{ weakly in } W^{1,1} \,, \quad \|\tilde{f}_k\|_{1,1} \to \|\tilde{f}^*\|_{1,1} \quad (k \to \infty) \,.$$

Proof of Theorem 3 (given Proposition 7.1). Replacing f by a slice as in (2.4), we may assume that it is a bounded function with compact support. By Lemma 3.2, we only need to show that $|\nabla f_k| - \partial f_k/\partial r$ converges to zero in L^1. Using the one-dimensional co-area formula in polar coordinates, we see that

$$\begin{aligned} \big\| |\nabla f_k| - \partial f_k/\partial r \big\|_1 &= \int_{S^n} \int_0^\infty \big(|\nabla f_k| - \partial f_k/\partial r \big) r^n \mathrm{d}r \, \mathrm{d}\theta \\ &= \|\nabla f_k\|_1 - \int_0^\infty \int_{S^n} r(\theta, h)^n \mathrm{d}\theta \, \mathrm{d}h \,, \end{aligned}$$

where $r(\theta, h)$ is the parameterization of the boundary of the level set at height h in polar coordinates. By Proposition 7.1, the first term converges to $\|\nabla f^*\|_1$. But since the function $r(\theta, h)$ converges uniformly to the radius of the level set of f^* at height h, the second term converges to the same limit. □

We need some more notation for the proof of Proposition 7.1. We denote the Lebesgue measure of a measurable set $A \subset R^{n+1}$ by $|A|$, and the perimeter of A by $\sigma(A)$. To be specific, our definition of the *perimeter* of A will be

$$\sigma(A) := \lim_{\varepsilon \to 0} \varepsilon^{-1} \big| (A + B_\varepsilon) \setminus A \big| \,,$$

where B_ε is the centered ball of radius ε, and $A + B_\varepsilon$ is the set where the convolution of the characteristic function is positive (that is, the ε-neighborhood of the set of density points of A). However, since the sequence of symmetrizations defined below quickly transforms A into a set with Lipschitz boundary by Lemma 6.2, any of the common definitions will do. The Brunn-Minkowski inequality implies that the perimeter never increases under Steiner symmetrization; it must strictly decrease unless A is (up to a translation, and except for a set of measure zero) already symmetric under reflection at the hyperplane of symmetrization. If C is another measurable set, we will denote by

$$\sigma(A|C) := \lim_{\varepsilon \to 0} \varepsilon^{-1} \left| ((A + B_\varepsilon) \cap C) \setminus A \right|$$

the *perimeter of A in C*.

By the co-area formula, the proof of Proposition 7.1 amounts to showing that the perimeter of all level sets of the f_n converges to the perimeter of f^*. Conversely, if A is a bounded star-shaped set, we can write $\sigma(A) = 2 \|\nabla f\|_1$, where f is defined by

$$f(x) := \begin{cases} 1 - \inf\{\lambda > 0 \mid \lambda^{-1} x \in A\} & x \in A \\ 0 & x \notin A\,. \end{cases}$$

(If A is not star-shaped, we replace it by $\mathcal{S}^n \cdots \mathcal{S}^0 A$). Since f satisfies the assumptions of Proposition 7.1, there exists a sequence of symmetrizations $\{A_k\}$ of A whose perimeters converge to the perimeter of A^*.

The main idea in the proof of Proposition 7.1 is the observation that, although Steiner symmetrization is not a local transformation, its effect on a set that is close to a ball in the sense of (6.5) can be localized in polar coordinates.

LEMMA 7.2 (Localization). *Let $C_1 \subset C_2$ be two cones in $\mathbf{R}^{n+1}$ which are symmetric under the reflection $(x^0, x^1, \dots, x^n) \mapsto (-x^0, x^1, \dots, x^n)$, and let $\hat{C}_1$ and $\hat{C}_2$ be their intersections with S^n. Assume that $\hat{C}_1$ has distance at least ε from the complement of $\hat{C}_2$.*

For every $\varepsilon > 0$ there exists $\delta > 0$ (which does not depend on C_1, C_2, r, and n) and a set of the form

$$D = \{x \in \mathbf{R}^{n+1} \mid x^0 \in R\,,\ \hat{x} \in \hat{D}\}$$

where $\hat{D}$ is a subset of $\mathbf{R}^n$, so that for every measurable set A satisfying (6.5), we have the inclusions

$$A \cap C_1 \ \subset \ A \cap D \ \subset \ A \cap C_2\,.$$

Proof. It is sufficient to prove the claim in the case where $\hat{C}_1$ is a pair of points and $\hat{C}_2$ is an ε-neighborhood of C_1. Performing a rotation about the

x^0-axis, we may assume that the pair of points lies in the $x^0 - x^1$-plane. Suppressing the coordinates $x^2, \dots, x^d$, we write

$$\hat{C}_1 := \left\{ (\pm r \sin\phi, \cos\phi) \right\}, \quad \hat{C}_2 := \left\{ (\sin\theta, \cos\theta) \;\middle|\; ||\theta| - |\phi|| < \varepsilon \right\},$$

where ϕ is a fixed angle in $[0, \pi/2]$. If δ is so small that

$$(1+\delta)\cos(\phi+\varepsilon) \; < \; (1-\delta)\cos(\phi), \quad (1+\delta)\cos(\phi) \; < \; (1-\delta)\cos(\phi-\varepsilon),$$

we can choose $\hat{D} = [(1-\delta) r \cos\phi, (1+\delta) r \cos\phi]$. □

LEMMA 7.3 (Local perimeter estimate). *Let A be a measurable set in $\mathbf{R}^n$, let C_1, C_2, and δ be as in Lemma 7.2, and assume that A satisfies relation* (6.5). *Then*

$$\begin{aligned} &\sigma(\mathcal{S}^0 A | C_1) \le \sigma(A|C_2) \\ &\sigma(\mathcal{S}^0 A | C_2) \ge \sigma(A|C_1) - \sigma(A) + \sigma(\mathcal{S}^0 A). \end{aligned}$$

Similarly, if f is a function in $W^{1,1}$ so that all level sets of f satisfy (6.5), *then*

$$\begin{aligned} &\int_{C_1} |\nabla \mathcal{S}^0 f| \mathrm{d}x \le \int_{C_1} |\nabla f| \, \mathrm{d}x \\ &\int_{C_2} |\nabla \mathcal{S}^0 f| \mathrm{d}x \ge \int_{C_1} |\nabla f| \, \mathrm{d}x - \|\nabla f\|_1 + \|\nabla \mathcal{S}^0 f\|_1 . \end{aligned}$$

Proof. The first line follows directly from Lemma 7.2 and the fact that

$$\mathcal{S}^0 (A \cap D) = (\mathcal{S}^0 A) \cap D,$$

to prove the second line we make the same argument for the complement of D. The claims for f follow with the co-area formula. □

We next prove that perimeter of a set that has been symmetrized sufficiently often to be close to a ball in the sense of (6.5) is close to the perimeter of a ball at least in some sectors of space. We use the fact that the boundary of A_k can be written as the graph of a monotone function as in (6.4).

LEMMA 7.4 (Area of a polar cap). *Let*

$$\hat{C}_\phi := \left\{ x \in S^n \;\middle|\; x^0 > 0, \; |\hat{x}| \le \sin\phi \right\}$$

be the polar cap of opening angle ϕ on the unit sphere in $\mathbf{R}^n$ $(n \ge 2)$, let C_ϕ the corresponding cone, and let A be a set of the form (6.4) *that is close to the ball B_r in the sense of* (6.5). *If ϕ is small enough and $\delta \le \phi^2$, then*

$$\sigma(A|C_\phi) \; \le \; (1+\varepsilon) \, \sigma(A^*|C_\phi). \tag{7.1}$$

Similarly, if f is a nonnegative function in $W^{1,1}$ so that all level sets satisfy (6.4) *and* (6.5) *with δ and ϕ as above, then*

$$\int_{C_\phi} |\nabla f|\,\mathrm{d}x \le (1+\varepsilon)\int_{C_\phi} |\nabla f^*|\,\mathrm{d}x\,. \tag{7.2}$$

Proof. The boundary of A in C_ϕ can be written as the graph of a monotone function over the disc of radius $(1+\delta)r\sin\phi$. We can continue this function monotonically to the smallest square containing the disc without changing the total variation. Lemma 6.3 gives

$$\sigma(A|C_\phi) \le |S^{n-1}|\frac{((1+\delta)r\phi)^n}{n}\left(1+n2^n\frac{(1+\delta-(1-\delta)\cos\phi}{\phi}\right),$$

which together with

$$\sigma(A^*|C_\phi) \ge |S^{n-1}|\frac{(r\sin\phi)^n}{n}$$

implies the claim (7.1). Inequality (7.2) follows with the co-area formula. □

Proof of Proposition 7.1. Let f be a nonnegative function in $W^{1,1}$ that vanishes at infinity. Fix $\varepsilon > 0$. We will construct a finite sequence of rotations $\{\mathcal{R}_j\}_{1\le j\le k}$ so that all level sets of f_k satisfy

$$\|\nabla f_j\|_1 \le (1+\varepsilon)\,\|\nabla f^*\|_1 \tag{7.3}$$

for $j \ge k$. Since the norm cannot increase under symmetrization, we need to prove the claim only for $j = k$.

By Lemma 7.4, there exist positive numbers δ_0 and ϕ_0, so that, if the level sets of f_k satisfy (6.5), then

$$\int_{C_\phi} |\nabla f_k|\,\mathrm{d}x \le (1+\varepsilon/3)\int_{C_\phi} |\nabla f^*|\,\mathrm{d}x \tag{7.4}$$

for all $\delta \le \delta_0$, $\phi \le \phi_0$, where C_ϕ is the cone of opening angle ϕ centered at the x^0-axis. We choose two such cones $C_1 \subset C_2$ whose opening angles are close enough so that

$$(1+\varepsilon)\int_{C_1} |\nabla f^*|\,\mathrm{d}x \ge (1+2\varepsilon/3)\int_{C_2} |\nabla f^*|\,\mathrm{d}x\,. \tag{7.5}$$

Then we choose $\delta \le \delta_0$ so small that Lemma 7.3 holds with C_1, C_2, and δ.

We would like to perform a sequence of Steiner symmetrizations and rotations so that all level sets of f_k satisfy (6.5) with the chosen δ. This will in general not be possible because the convergence of level sets to balls is uniform only for sets so that R/r is bounded above, where R is the radius of the smallest ball containing the set, and r the radius of its spherical symmetrization. However, we may approximate f in $W^{1,1}$ by a slice $\tilde{f}$ as in (2.4) with an error that is as small as we please. Since we assumed that f vanishes at infinity, we can choose the heights h_1 and h_2 so that $\tilde{f}$ has compact support, and the sizes of level sets are bounded away from zero. If k is large enough, all level sets of $\tilde{f}_k$ satisfy (6.5) and consequently (7.4). If $\tilde{f}_k$ already satisfies (7.3), we are done. Otherwise, there exists a rotation $\mathcal{R}$ so that

$$\int_{C_1} |\nabla \mathcal{R} \tilde{f}_k| \mathrm{d}x > (1+\varepsilon) \int_{C_1} |\nabla \tilde{f}^*| \mathrm{d}x \,. \tag{7.6}$$

We set $\mathcal{R}_{k+1} = \mathcal{R}$, and define $\mathcal{S}_{k+1}$ and f_{k+1} by (6.1) and (6.2). Combining Lemma 7.3 with inequalities (7.4), (7.5), and (7.6), we see that

$$\begin{aligned} \|\nabla f_k\|_1 - \|\nabla f_{k+1}\|_1 &\geq \|\nabla \mathcal{R} \tilde{f}_k\|_1 - \|\nabla \mathcal{S}^0 \mathcal{R} \tilde{f}_k\|_1 \\ &\geq \int_{C_1} |\nabla \mathcal{R} \tilde{f}_k| \mathrm{d}x - \int_{C_2} |\nabla \mathcal{S}^0 \mathcal{R} \tilde{f}_k| \mathrm{d}x \\ &\geq \varepsilon/3 \int_{C_2} |\nabla \tilde{f}^*| \mathrm{d}x > 0 \,. \end{aligned}$$

The claim (7.3) will hold after repeating the last step finitely many times. Our construction also guarantees the convergence in L^1 and the weak convergence in $W^{1,1}$, since (6.5) holds with δ as small as we please for k large enough. □

REMARK. The sequence $\{f_k\}$ constructed in the proof of Proposition 7.1 satisfies

$$\int_C |\nabla f_k| \ \mathrm{d}x \ \to \ \int_C |\nabla f^*| \ \mathrm{d}x \qquad (k \to \infty)$$

for all open cones C in $\mathbf{R}^{n+1}$. Similarly, if A is a bounded set, there exists a sequence $\{A_k\}$ as in (6.3) so that

$$\sigma(A_k|C) \to \sigma(A^*|C) \qquad (k \to \infty) \,.$$

8 Symmetrization of Co-area Regular Functions

Let f be a nonnegative co-area regular function in $W^{1,p}$ that vanishes at infinity. We will construct a sequence of Steiner symmetrizations and rotations by (6.2) that approximates f^* in $W^{1,p}$.

Following Almgren and Lieb [AlLi], we use the functional

$$\Psi(f) := \int \left(\sqrt{1 + |\nabla f|^2} - 1 \right) \mathrm{d}x$$

in place of $W^{1,p}$-norms to indicate convergence or divergence of a sequence in $W^{1,p}$. If f has compact support, $\Psi(f)$ gives the surface area of the graph of f, minus the measure of its support. Note that

$$F(z) = \sqrt{1+z^2} - 1$$

is strictly convex and increasing for all positive x, and $F(0) = 0$. In particular, $\Psi(f)$ can only decrease under Steiner or spherical symmetrization of f.

Our strategy is to symmetrize a given function f until we arrive at a function f_k for which $\Psi(f_k)$ can only decrease very little under any Steiner symmetrization. Using an explicit lower bound for the drop in Ψ under Steiner symmetrization, we show that the gradient of such a function restricted to a level surface is close to a constant.

This is not enough to ensure convergence in general, because the co-area formula gives no information about the set of critical points. If f is co-area regular, however, convergence of f_k to f^* in L^p implies that the set of critical points of f_k converges to the set of critical points of f^* in symmetric difference, since the sets of critical points of f_k and f^* differ by sets of measure zero from and $f_k^{-1}(C)$ and ${f^*}^{-1}(C)$, where C is the set of critical values discussed in Remark (ii) after Lemma 4.2, and Lemma 6.1 applies.

LEMMA 8.1 (Convergence of surface area). *Let f be a function in $W^{1,1}$. There exists a sequence of rotations $\{\mathcal{R}_k\}_{k\geq 1}$ so that for f_k defined by* (6.2),

$$\sup_{\mathcal{R}} \big(\Psi(f_k) - \Psi(\mathcal{S}^0\mathcal{R}f_k)\big) \to 0 \qquad (k\to\infty)\,,$$

and, moreover,

$$f_k \to f^* \text{ in } L^1\,, \quad \|f_k\|_{1,1} \ \to \ \|f^*\|_{1,1} \qquad (k\to\infty)\,.$$

Proof. We construct the sequence recursively. Set $k_1 = 1$. Since the surface area can only decrease under symmetrization, any sequence of the form (6.2) satisfies

$$\Psi(f_k) - \Psi(\mathcal{S}^0\mathcal{R}f_k) \leq \Psi(f_k) - \inf_{\{\mathcal{R}_l\}} \lim_{l\to\infty} \Psi(f_{k+l}) \tag{8.1}$$

for every rotation $\mathcal{R}$. Assume that $f_1, \ldots, f_{k_j}$ are already given. If j is even, we choose the next terms in the sequence of rotations $\mathcal{R}k_j + 1, \ldots, \mathcal{R}_{k_{j+1}}$ so that

$$\Psi(f_{k_{j+1}}) \leq \tfrac{1}{2}\Big(\Psi(f_{k_j}) + \inf_{\{\mathcal{R}_k\}} \lim_{l\to\infty} \Psi(f_{k_j+l})\Big)\,,$$

that is,

$$\Psi(f_{k_{j+1}}) - \inf_{\{\mathcal{R}_k\}} \lim_{l\to\infty} \Psi(f_{k_{j+1}+l}) \le \tfrac{1}{2}\Big(\Psi(f_{k_j}) - \inf_{\{\mathcal{R}_k\}} \lim_{l\to\infty} \Psi(f_{k_j+l})\Big)\,,$$

where the infimum is taken over all possible continuations of the given finite sequence of rotations. If j is odd, we apply Proposition 7.1, and choose the next terms in the sequence of rotations so that

$$\|f_{k_{j+1}}\|_{1,1} \le \|f^*\|_{1,1} + 2^{-j}\,.$$

By the monotonicity of the norm, this inequality will hold for all later terms in the sequence. Repeating the last two steps, we can produce a sequence of rotations along which the right hand side of (8.1) converges to zero, and which satisfies the conclusion of Proposition 7.1. □

LEMMA 8.2 (The drop in surface area). *Let f be a nonnegative function in $W^{1,p}(\mathbf{R}^{n+1})$. Assume that for $\hat{x}$ in some open set B and all $h_1 < h < h_2$, the equations $f(x) = h$, $(x^1, \ldots, x^n) = \hat{x}$ have exactly two solutions, $x^+(\hat{x}, h)$ and $x^-(\hat{x}, h)$, that $\partial/\partial x^0\, f(x^+) \le 0$, $\partial/\partial x^0\, f(x^-) \ge 0$, and that the angles of $\nabla f(x^+(\hat{x}, h))$ and $\nabla f(x^+(\hat{x}, h))$ with the line $\hat{x} =$ const are uniformly bounded away from $\pi/2$. There exists a nonnegative continuous so that*

$$\Psi(f) - \Psi(\mathcal{S}^0 f)$$
$$\ge \int_{h_1}^{h_2} \int_B G\big(|\nabla f(x^+)|, |\nabla f(x^-)|\big) \mathcal{X}_{\mathrm{d}/\mathrm{d}x^0\, f(x^+)\neq 0} \mathcal{X}_{\mathrm{d}/\mathrm{d}x^0 f(x^-)\neq 0} \mathrm{d}\hat{x}\, \mathrm{d}h\,.$$

The function G has the properties that $G(z, w) = G(w, z)$, $G(z, w) = 0 \Leftrightarrow z = w$, and for fixed z, $G(z, \cdot)$ increases with w for $w > z$, and decreases for $w < z$.

Proof. By formulas (5.1) and (5.7), the derivatives of the Steiner symmetrization satisfy

$$\big|\mathrm{d}/\mathrm{d}x^0 \mathcal{S}^f(\tilde{x}^\pm)\big|^{-1} = \tfrac{1}{2}\big(|\mathrm{d}/\mathrm{d}x^0 f(x^+)|^{-1} + |\mathrm{d}/\mathrm{d}x^0 f(x^-)|^{-1}\big)$$

and

$$\big|\mathrm{d}/\mathrm{d}x^0 \mathcal{S}^f(\tilde{x}^\pm)\big|^{-1} \nabla_{\hat{x}} \mathcal{S}^0 f(\xi^\pm, \hat{x})$$
$$= \tfrac{1}{2}\big(|\mathrm{d}/\mathrm{d}x^0 f(x^+)|^{-1} \nabla_{\hat{x}} f(x^+) + |\mathrm{d}/\mathrm{d}x^0 f(x^-)|^{-1} \nabla_{\hat{x}} f(x^-)\big)\,,$$

where $\tilde{x}^+$ and $\tilde{x}^-$ are the points on the level surface of $\mathcal{S}^0 f$ corresponding to x^+ and x^-, that is, the two solutions of the equations $\mathcal{S}^0 f(x) = h$, $(x^1, \ldots, x^n) = \hat{x}$. By the one-dimensional co-area formula (4.5), we obtain

$$\Psi(f) - \Psi(\mathcal{S}^0 f) \ge \int_{h_1}^{h_2} \int_B \big(F(0, \nabla f(x^+), \nabla f(x^-)) + F(1, \nabla f(x^+), \nabla f(x^-))$$
$$-2F(1/2, \nabla f(x^+), \nabla f(x^-))\big) \mathcal{X}_{\nabla f(x^+)\neq 0} \mathcal{X}_{\nabla f(x^-)\neq 0} \mathrm{d}\hat{x}\, \mathrm{d}h\,,$$

where for $u, v \in \mathbf{R}^{n+1}$ with $\hat{u}^0 < 0 < \hat{v}^0$,

$$F(\lambda, u, v) = \left| \lambda - (u^0)^{-1} \begin{pmatrix} -u^0 \\ \hat{u} \\ 1 \end{pmatrix} + (1-\lambda)(v^0)^{-1} \begin{pmatrix} v^0 \\ \hat{v} \\ 1 \end{pmatrix} \right|_{\mathbf{R}^{n+2}} .$$

Note that F is strictly convex in λ unless $(-u^0, \hat{u}) = (v^0, \hat{v})$. Setting

$$G(z, w) := \inf_{|u|=z, |v|=w, u^0 \geq c|\hat{u}|, v^0 \geq c|\hat{v}|} \big(F(0, u, v) + F(1, u, v) - 2F(1/2, u, v) \big)$$

where c is a sufficiently small positive constant, we see that G has the claimed properties. □

LEMMA 8.3 (Change of coordinates). *Assume that S is a surface that can be parameterized in polar coordinates by a positive function on S^n which is bounded above by R and Lipschitz continuous with constant LR. Choose c_1 so small that for $|\xi| < c_1 R$, the line determined by (ξ, ω) intersects S transversally in a pair of points which we denote by $x^+(\xi, \omega)$ and $x^-(\xi, \omega)$.*

There exists a constant c_2 which depends only on L and the choice of c_1, so that for every nonnegative measurable function H on $S \times S$ which is symmetric under exchange of the two variables, the inequality

$$\begin{aligned} &\int_{S^n} \int_{\xi \perp \omega} H(x^+, x^-) \mathcal{X}_{\{|\xi| < c_1 R\}} \mathrm{d}\xi \, \mathrm{d}\omega \\ &\qquad \geq \frac{c_2}{R^n} \int_S \int_S H(x(s), x(t)) \mathcal{X}_{\{\xi(s,t) < c_1 R\}} \mathrm{d}s \, \mathrm{d}t \end{aligned}$$

holds. Here, $x(s)$ and $x(t)$ are the points in $\mathbf{R}^{n+1}$ corresponding to s and t, and $\xi(s,t)$ is the distance of the line through $x(s)$ and $x(t)$ from the origin.

Proof. The claim follows immediately from the fact that the transformation $(\xi, \omega) \to (x^+(\xi, \omega), x^-(\xi, \omega))$ is Lipschitz continuous with Lipschitz inverse by our transversality assumption. □

LEMMA 8.4 (Convergence on level surfaces). *Let $\{f_k\}_k \geq 0$ be a sequence of functions constructed by rearrangements as in (6.2), and let G be the function of Lemma 8.2. If*

(8.2)

$$\int_{h_1}^{h_2} \int_{\partial E_(h)} \int_{\partial E_k(h)} G\big(|\nabla f_k|_k (s, h), |\nabla f_k| (t, h) \big) \mathrm{d}s \, \mathrm{d}t \, \mathrm{d}h \to 0 \quad (k \to \infty),$$

then

$$\int_{h_1}^{h_2} \int_{\partial E_{f_k}(h)} \big| F(|\nabla f_k(s, h)|) - c_1(h) \big| \, |\nabla f_k|^{-1} \, \mathrm{d}s \, \mathrm{d}h \to 0 \qquad (k \to \infty),$$

where $E_k(h)$ is the level set of f_k at height h, $c_1(h)$ is defined by

$$c_1(h) = \frac{2c_2(h)^2}{1 - c_2(h)^2}\,, \quad c_2(h) := \lim_{k\to\infty} \sigma_k^{-1} \int_{\partial E_k(h)} F(|\nabla f_k|)\, |\nabla f_k|^{-1}\, \mathrm{d}s\,,$$

and $\sigma_k(h)$ is the perimeter of the level set of f_k at height h.

Proof. The limits

$$\sigma(h) := \lim_{k\to\infty} \sigma_k(h)\,, \quad c_3(h) := \lim_{k\to\infty} \int_{\partial E_k(h)} F(|\nabla f_k|)\, |\nabla f_k|^{-1}\, \mathrm{d}s$$

exist for every h, since both sequences decrease monotonically with k by the usual rearrangement inequalities for the perimeter and for Ψ. By the properties of G,

$$\inf_{|\phi(\alpha)-\phi(\beta)|\geq z} G(\alpha,\beta)$$

is positive for $z > 0$, and nondecreasing; multiplying it with a strictly increasing function less than one, we see that

$$G(\alpha,\beta) \geq \tilde{G}\big(\,|\phi(\alpha) - \phi(\beta)|\,\big)$$

where $\tilde{G}$ is a strictly increasing nonnegative function. Assumption (8.2) implies that $\phi(|\nabla f_k(s,h)|) - \phi(|\nabla f_k(t,h)|)$ (as a function of s, t, and h) converges to zero in measure. Since ϕ is bounded, it follows that

$$\begin{aligned}
&\int_{h_1}^{h_2} \int_{\partial E_k(h)} \big|\phi(|\nabla f_k|) - c_2(h)\big|^2 \mathrm{d}s\, \mathrm{d}h \\
&\quad = \int_{h_1}^{h_2} \sigma_k^{-1}(h) \int_{\partial E_k(h)} \int_{\partial E_k(h)} \big|\phi(|\nabla f_k(s,h)| - \phi(|\nabla f_k(t,h)|)\big|^2 \mathrm{d}s\, \mathrm{d}t\, \mathrm{d}h \\
&\quad \to 0 \qquad\qquad (k \to \infty)\,.
\end{aligned}$$

Hence, on almost all level surfaces of f_k, $\phi(|\nabla f_k|)$ converges in measure to the constant $c_2(h)$. It follows that on each level surface, $|\nabla f_k|$ converges in measure to a constant $c_4(h)$, which may be infinity. Calculating $c_1(h) = F(c_4(h))$, and $c_2(h) = F(c_4(h))/c_4(h)$ proves the claim. □

Proof of Theorem 2. We already showed at the beginning of section 6 that the spherical symmetrization of a co-area irregular function cannot be approximated by a sequence of Steiner symmetrizations and rotations.

To prove the converse, let f be a co-area regular function in $W^{1,p}(\mathbf{R}^{n+1})$ that vanishes at infinity. We will construct a sequence of rearrangements $\{f_k\}$ of f by (6.2) which converges to f^* in $W^{1,p}$. Replacing f by a slice as in (2.4), we may assume without loss of generality that f is a bounded function with compact support, and that the measures of the nontrivial level sets of f are bounded away from zero.

By Lemma 8.1 there exists a sequence of symmetrizations so that

$$\sup_{\mathcal{R}}\big(\Psi(f_k) - \Psi(\mathcal{S}^0\mathcal{R}f_k)\big) \to 0 \quad (k \to \infty) \tag{8.3}$$

and

$$f_k \to f^* \text{ in } L^1, \quad \|f_k\|_{1,1} \to \|f^*\|_{1,1} \quad (k \to \infty).$$

This implies that the perimeters of almost all level sets converge to the perimeters of the level sets of f^*, that is,

$$\lim_{k\to\infty} \sigma_k(h) = \sigma^*(h) \tag{8.4}$$

for almost all h.

By assumption, all level sets of f_k satisfy $B_{r_1} \subset E_k(h) \subset B_{r_2}$ for some $0 < r_1 < r_2$. Fix a small number $\varepsilon > 0$. By Lemma 6.2, there exists a constant c so that for $\omega \in \hat{C}_\varepsilon$ and ξ perpendicular to ω in $\mathbf{R}^{n+1}$ with $|\xi| \le c$, the line defined by $x(t) = \xi + t\omega$ intersects the boundary of the level set $E_k(h)$ transversally in exactly two points $x^+(\xi,\omega)$ and $x^-(\xi,\omega)$. By Lemma 8.2,

$$\sup_{\omega\in\hat{C}_\varepsilon} \int_{h_1}^{h_2} \int_{\xi\perp\omega} G\big(|\nabla f_k(x^+)|, |\nabla f_k(x^-)|\big)\mathcal{X}_{\mathrm{d}/\mathrm{d}x^0 f_k(x^+)\neq 0}\mathcal{X}_{\mathrm{d}/\mathrm{d}x^0 f_k(x^-)\neq 0}\mathcal{X}_{|\xi|\le c}\mathrm{d}\xi\,\mathrm{d}h \to 0 \qquad (k \to \infty).$$

Note that the left-hand side of (8.3) depends on $\mathcal{R}$ only through $\omega = \mathcal{R}(1,0,\dots,0)$. Integrating over ω and using Lemma 8.3 with

$$\begin{aligned} &H(x(s,h), x(t,h)) \\ &\quad = G\big(|\nabla f_k(x(s,h))|, |\nabla f_k(x(t,h))|\big)\mathcal{X}_{C_\varepsilon}(x(s,h))\mathcal{X}_{C_\varepsilon}(x(t,h)) \end{aligned}$$

we see that

$$\int_0^\infty \int_{\partial E_k(h)\cap C_\varepsilon} \int_{\partial E_k(h)\cap C_\varepsilon} G\big(|\nabla f_k(x(s,h))|, |\nabla f_k(x(t,h))|\big)\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}h \longrightarrow 0 \qquad (k \to \infty).$$

Set

$$c_{k,\varepsilon}(h) : \sigma_{k,\varepsilon}(h)^{-1} \lim_{\delta\to 0} \int_{h-\delta}^{h+\delta} \int_{E_k(h')\cap C_\varepsilon} F(|\nabla f_k|)\,|\nabla f_k|^{-1}\,\mathrm{d}s\,\mathrm{d}h'$$

where $\sigma_{k,\varepsilon}(h)$ is the perimeter of the level set $E_k(h)$ in C_ε. By Lemma 8.4,

we have that
(8.5)
$$\int_{h_1}^{h_2}\int_{E_k(h)\cap C_\varepsilon}\big|F(|\nabla f_k|)-c(h)\big|\,|\nabla f_k|^{-1}\,\mathrm{d}s\,\mathrm{d}h\to 0\quad (k\to\infty)\,,$$
where
$$c(h):=\lim_{k\to\infty}c_{k,\varepsilon}(h)\,,$$
which clearly cannot depend on ε. As a pointwise limit of the $c_{k,\varepsilon}$, c is Borel measurable and vanishes if h is in the support of the singular part of the measure induced by the distribution function of f.

We next show that $F(|\nabla f_k|)$ converges to $c(f^*)$ in L^1. We write
$$\begin{aligned}\int\big|F(|\nabla f_k(x)|)-c(f^*(x))\big|\mathrm{d}x\\ \le\int\big|F(|\nabla f_k(x)|)-c(f_k(x))\big|\mathrm{d}x+\int\big|c(f_k(x))-c(f^*(x))\big|\mathrm{d}x\,.\end{aligned}$$
If f is co-area regular, then $c(f_k(x))$ vanishes almost everywhere on the set of critical points of f_k by the remark after Lemma 4.2, and we can rewrite the first term as
$$\begin{aligned}\int\big|F(|\nabla f_k(x)|)-c(f_k(x))\big|\mathcal{X}_{|\nabla f_k|>0}\mathrm{d}x\\ =\int_{h_1}^{h_2}\int_{E_k(h)\cap C_\varepsilon}\big|F(|\nabla f_k|)-c(h)\big|\,|\nabla f_k|^{-1}\,\mathrm{d}s\,\mathrm{d}h\,,\end{aligned}$$
which converges to zero by (8.5). The second term converges to zero by Lemma 6.1.

By Lemma 3.2, the sequence $|\nabla f_k|$ converges to some limiting function in L^p. By Theorem 3, the angular part of the gradient converges to zero in L^p. The limit of the sequence of gradients must coincide with the weak limit ∇f^*. □

References

[A] L.V. AHLFORS, Conformal Invariants: Topics in Geometric Function Theory, McGraw-Hill Series in Higher Mathematics, 1973.

[AlLi] F. ALMGREN, E.H. LIEB, Continuity and discontinuity of the spherically decreasing rearrangement, Journal of the AMS 2 (1989), 683–773.

[AlvTrLio] A. ALVINO, G. TROMBETTI, P.-L. LIONS, On optimization problems with prescribed rearrangements, Nonlinear Analysis 13 (1989), 185–220.

[B] A. BAERNSTEIN, A unified approach to symmetrization, Symposia Matematica 35 (1995), 47–91.

[BTa] A. BAERNSTEIN, B.A. TAYLOR, Spherical rearrangements, subharmonic functions, and *-functions in n-space, Duke Math. J. 43 (1976), 245–268.

[Ba] C. BANDLE, Isoperimetric Inequalities and Applications, Pitman, Boston, 1986.

[Be1] W. BECKNER, Inequalities of Fourier analysis, Ann. Math. (2) 102 (1975), 159–182.

[Be2] W. BECKNER, Sobolev inequalities, the Poisson semigroup, and analysis on the sphere S^n, Proc. Natl. Acad. Sci. USA 89 (1992), 4816–4819.

[Ben] Y. BENYAMINI, Two-point symmetrization, the isoperimetric inequality on the sphere and some applications, In "Texas Functional Analysis Seminar 1983/84, Austin, Texas", Longhorn Notes, University of Texas Press, Austin, TX (1984), 53–76.

[Bl] W. BLASCHKE, Eine isoperimetrische Eigenschaft des Kreises, Math. Z. 3 (1918), 52–57.

[BoLe] B. BOLLOBÁS, I.B. LEADER, Compressions and isoperimetric inequalities, J. Combin. Theory Ser. A 56 (1991), 47–62.

[BrLi] H.J. BRASCAMP, E.H. LIEB, Best constants in Young's inequality, its converse, and its generalization to more than three functions, Advances in Mathematics 20 (1976), 151–173.

[BrLiLut] H.J. BRASCAMP, E.H. LIEB, J.M. LUTTINGER, A general rearrangement inequality for multiple integrals, J. Funct. Anal. 17 (1974), 227–237.

[BroSo] F. BROCK, A.YU. SOLYNIN, An approach to symmetrization via polarization, Technical report, Universität zu Köln, 1996.

[BrotZi] J. BROTHERS, W.P. ZIEMER, Minimal rearrangements of Sobolev functions, Crelle J. Reine Angew. Math. 384 (1988), 153–179.

[BuZ] YU.D. BURAGO, V.A. ZALGALLER, Geometric Inequalities (English translation of the Russian original), Springer Grundlehren 285, 1980.

[CStu] C. CARATHEODORY, E. STUDY, Zwei Beweise des Satzes, daß der Kreis unter allen Figuren gleichen Umfangs den größten Inhalt hat, Math. Annalen 68 (1909), 133–144.

[Ch] I. CHAVEL, Eigenvalues in Riemannian Geometry, Academic Press, Orlando, Fla, 1984.

[Co] J.-M. CORON, The continuity of the rearrangement in $W^{1,p}(\mathbf{R})$, Ann. Scuola Norm. Sup. Pisa, Cl. Sci (4) 11 (1984), 57–85.

[Cr] P.S. CROOKE, An isoperimetric bound for a Sobolev constant, Colloq. Math. 38 (1977/78), 263–267.

[DSt] D.L. DONOHO, P.B. STARK, A note on rearrangements, spectral concentration, and the zero-order prolate spheroidal wavefunction, IEEE Trans. Inform. Theory 39 (1993), 257–260.

[Du] G.F.D. DUFF, A general inequality for the derivative of an equimeasurable rearrangement, Canad. J. Math. 28 (1976), 793–804.

[F] G. FEDERER, Geometric Measure Theory, Springer Grundlehren 153, 1969.

[FFl] H. FEDERER, W.H. FLEMING, Normal and integral currents, Ann. of Math. (2) 72 (1960), 458–520.

[FrFu] P. FRANKL, Z. FÜREDI, A short proof for a theorem of Harper about Hamming spheres, Discrete Math. 34 (1981), 311–313.

[H] H. HADWIGER, Einfache Herleitung der isoperimetrischen Ungleichung für abgeschlossene Punktmengen, Math. Ann. 124 (1952), 158–160.

[HO] H. HADWIGER, D. OHMANN, Brunn-Minkowski Satz und Isoperimetrie, Math. Z. 66 (1956), 1–8.

[HaLitP] G.E. HARDY, J.E. LITTLEWOOD, G. PÓLYA, Inequalities, Cambridge University Press, London and New York, 1952.

[Hi] K. HILDÉN, Symmetrization of functions in Sobolev spaces and the isoperimetric inequality, Manuscripta Math. 18 (1976), 215–235.

[K] B. KAWOHL, Rearrangements and convexity of level sets in PDE, LNM 1150, Springer Verlag, Berlin Heidelberg, 1985, PAGENOS???.

[L] T. LACHAND-ROBERT, Letter to E.H. Lieb, May 1995. Noted as "open problem" at the meeting on qualitative aspects of PDE, organized by H. Berestycki, B. Kawohl, and G. Talenti, Oberwolfach, Germany February 12–18, 1995.

[Li1] E.H. LIEB, Existence and uniqueness of the minimizing solution of Choquard's nonlinear equation, Studies in Applied Mathematics 57 (1977), 93–105.

[Li2] E.H. LIEB, Sharp constants in the Hardy-Littlewood-Sobolev and related inequalities, Ann. Math. 118 (1983), 349–374.

[LiLo] E.H. LIEB, M. LOSS, Analysis, AMS Graduate Studies in Mathematics, 1996.

[Lo] M. LOSS, Personal communication, July 1995.

[Lu] L. LUSTERNIK, Die Brunn-Minkowskische Ungleichung für beliebige meßbare Mengen, Dokl. Acad. Sci. URSS 3 (1935), 55–58.

[M] P. MANI-LEVITSKA, Random Steiner symmetrizations, Studia Sci. Math. Hungar. 21 (1986), 373–378.

[Ma] W. MAZJA, Einbettungssätze für Sobolewsche Räume. Teil 1 (German translation of the Russian original), Teubner Verlagsgesellschaft, Leipzig, 1979.

[PSz] F. PÓLYA AND G. SZEGŐ, Isoperimetric Inequalities, Princeton University Press, Princeton, 1951.

[R] F. RIESZ, Sur une inégalité intégrale, J. London Math. Soc. 5 (1930), 162–168.

[S] S.L. SOBOLEV, On a theorem of functional analysis, Mat. Sb. (N.S.) 4 (1938), 471–497 (Russian original); AMS Transl. (2), 34 (1963), 39–68

(English translation).

[So] A.Yu. Solynin, Application of the polarization for proofs of functional inequalities (in Russian), Russian Academy of Sciences, St. Petersburg, PDMI preprint, 10, 1995.

[Sp] E. Sperner, Symmetrisierung für Funktionen mehrerer reeller Variablen, Manuscripta Math. 11 (1974), 159–170.

[Ste] J. Steiner, Einfacher Beweis der isoperimetrischen Hauptsätze, Crelle J. Reine Angew. Math. 18 (1838) 281–296; Ges. Werke 2 (Berlin, 1882), 77–91 (in German).

[T1] G. Talenti, The standard isoperimetric theorem, Handbook of Convex Geometry, Vol. A, B, 73–123, North Holland, Amsterdam, YEAR???

[T2] G. Talenti, Best constant in Sobolev inequality, Ann. Mat. Pura Appl (4) 110 (1976), 353–372.

[Zi] W. Ziemer, Weakly differentiable functions. Sobolev spaces and functions of bounded variation, Graduate Texts in Mathematics 120. Springer-Verlag, New York, 1989.

Almut Burchard
Department of Mathematics
Princeton University
Princeton, NJ 08544
USA
E-mail: burchard@math.princeton.edu

Submitted: September 1996