Regular Article

The Pólya-Szegő inequality for smoothing rearrangements ☆

Gabriele Bianchi [a,*], Richard J. Gardner [b], Paolo Gronchi [c], Markus Kiderlen [d]

[a] *Dipartimento di Matematica e Informatica "U. Dini", Università di Firenze, Viale Morgagni 67/A, Firenze, I-50134, Italy*
[b] *Department of Mathematics, Western Washington University, Bellingham, WA 98225-9063,USA*
[c] *Dipartimento di Matematica e Informatica "U. Dini", Università di Firenze, Piazza Ghiberti 27, Firenze, I-50122, Italy*
[d] *Department of Mathematical Sciences, University of Aarhus, Ny Munkegade, DK–8000 Aarhus C, Denmark*

A B S T R A C T

A basic version of the Pólya-Szegő inequality states that if Φ is a Young function, the Φ-Dirichlet energy—the integral of $\Phi(\|\nabla f\|)$—of a suitable function $f \in \mathcal{V}(\mathbb{R}^n)$, the class of non-negative measurable functions on $\mathbb{R}^n$ that vanish at infinity, does not increase under symmetric decreasing rearrangement. This fact, along with variants that apply to polarizations and to Steiner and certain other rearrangements, has numerous applications. Very general versions of the inequality are proved that hold for all smoothing rearrangements, those that do not increase the modulus of continuity of functions. The results cover all the main classes of functions previously considered: Lipschitz functions $f \in \mathcal{V}(\mathbb{R}^n)$, functions $f \in W^{1,p}(\mathbb{R}^n) \cap \mathcal{V}(\mathbb{R}^n)$ (when $1 \leq p < \infty$ and $\Phi(t) = t^p$), and functions $f \in W_{loc}^{1,1}(\mathbb{R}^n) \cap \mathcal{V}(\mathbb{R}^n)$. In addition, anisotropic versions of these results, in which the role of the unit ball is

played by a convex body containing the origin in its interior,
are established. Taken together, the results bring together all
the basic versions of the Pólya-Szegő inequality previously
available under a common and very general framework.
© 2024 The Authors. Published by Elsevier Inc. This is an
open access article under the CC BY license (http://
creativecommons.org/licenses/by/4.0/).

1. Introduction

A familiar version of the Pólya-Szegő inequality states that if $1 \leq p \leq \infty$ and $f \in W^{1,p}(\mathbb{R}^n) \cap \mathcal{V}(\mathbb{R}^n)$, then $f^{\#} \in W^{1,p}(\mathbb{R}^n)$ and

$$\int_{\mathbb{R}^n} \|\nabla f^{\#}(x)\|^p \, dx \leq \int_{\mathbb{R}^n} \|\nabla f(x)\|^p \, dx. \tag{1.1}$$

See, e.g., [3, Theorem 3.20 and p. 113]; when $p = \infty$, the integrals of pth powers are replaced by the essential suprema over $\mathbb{R}^n$. Here $f^{\#}$ denotes the *symmetric decreasing rearrangement* of f (another common notation is $f^{\star}$) and $\mathcal{V}(\mathbb{R}^n)$ is the class of non-negative measurable functions on $\mathbb{R}^n$ that vanish at infinity, a natural class for which this rearrangement is defined. Most of our definitions and terminology can be found in Sections 2 and 3. However, since we have no need for a precise definition of $f^{\#}$ (see, e.g., [3, Definition 1.29], [23, p. 9], [26, p. 80]), we lean on a vivid description of Sperner [38, Abstract]: Imagine the subgraph of f as a lump of clay on a potter's wheel, which on turning is molded into a perfectly symmetrical shape, maintaining the height of each particle of clay. The molded shape then represents the subgraph of $f^{\#}$. The map that takes f to $f^{\#}$ is the primary example of a rearrangement on $\mathcal{V}(\mathbb{R}^n)$. In general, if X is a class of measurable functions on $\mathbb{R}^n$ containing the characteristic functions of sets in $\mathcal{L}^n$, the $\mathcal{H}^n$-measurable sets of finite measure, a *rearrangement T on X* is an essentially monotonic (i.e., monotonic up to sets of $\mathcal{H}^n$-measure zero) and equimeasurable (preserving the $\mathcal{H}^n$-measure of superlevel sets) map $T : X \to X$.

Inequality (1.1) has its roots in studies of symmetrization of sets and rearrangements of functions that go back to Jakob Steiner's work on the isoperimetric inequality around 1836. (The isoperimetric inequality can actually be deduced from the case $p = 1$ of (1.1).) With extra assumptions on f, it was first proved for $n = p = 2$ by G. Faber and E. Krahn independently, and then for all n and $p = 2$ by Krahn, all in the 1920s. In fact, (1.1) was a key ingredient in the solution by Faber and Krahn of Lord Rayleigh's 1884 conjecture that the disk has the lowest fundamental frequency of vibration of all membranes of a given area. References are given by Daners [14] in his detailed commentary focusing on Krahn's solution, and by Mondino and Semola [30], who provide a lucid account of this early history. (The latter also describe extensions of (1.1) to non-Euclidean settings, but the present paper is set entirely in $\mathbb{R}^n$.) The many sources that outline the recent history

of (1.1) often contradict each other. Pólya and Szegő's classic text [32] on isoperimetric inequalities in mathematical physics is always cited, but (1.1) is not explicitly stated there, and the setting is different. They work with smooth surfaces A_0 and A_1 in $\mathbb{R}^3$ with A_0 in the interior of A_1, and functions equal to 0 on A_0, 1 on A_1, and between 0 and 1 in the region bounded by A_0 and A_1. Their arguments in [32, pp. 154–156], with additional work, lead to (1.3) below when f is smooth with compact support, $n = 3$, and T is the $(2,3)$-Steiner rearrangement (explained below) with respect to the xy-plane. This and the approximation of Schwarz rearrangement by a sequence of $(2,3)$-Steiner rearrangements sketched in [32, pp. 157] yield (1.3), and hence (1.1) for $1 < p < \infty$, with the same restrictions on f and n. We believe Baernstein [3, Section 3.8] is correct in giving credit to Sperner [38] for proving (1.1) when f is Lipschitz (though he cites the wrong paper) and Hildén [22] for the result as stated above when $p < \infty$.

Diverse variants and applications of the Pólya-Szegő inequality (often called the Pólya-Szegő principle) have generated a very substantial literature, surveyed by Talenti [40, Sections 1.3 and 1.5], [41, Section 5], who in [41, p. 126] provides over fifty references. The main themes are: Pólya-Szegő inequalities on spheres, hyperbolic, or other spaces, and for other functionals of the gradient; weighted versions involving other measures; versions invariant under affine transformations; anisotropic inequalities; the examination of equality cases; connections with capacitary inequalities; and applications to mathematical physics, PDEs, and function spaces.

Like (1.1), this paper has also arisen from earlier work on symmetrization and rearrangement, including our previous investigations [4–6]. As in those articles, the attention is less on particular symmetrizations or rearrangements than on general properties that allow results for those special cases to be extended and unified. For the classes X of measurable functions considered, such as $\mathcal{V}(\mathbb{R}^n)$, each rearrangement T is essentially determined by an associated map $\Diamond_T : \mathcal{L}^n \to \mathcal{L}^n$ defined by $\Diamond_T A = \{x : T1_A(x) = 1\}$, where 1_A is the characteristic function of A, satisfying

$$\{x : Tf(x) \geq t\} = \Diamond_T\{x : f(x) \geq t\},$$

essentially, for $t > \operatorname{ess\,inf} f$; see Proposition 3.5 below. Another such map may then be defined by $\Diamond_T^* A = (\Diamond_T A)^*$, where E^* denotes the set of density points of E. We focus here on *smoothing* rearrangements, those for which

$$(\Diamond_T^* A) + dB^n \subset \Diamond_T^*(A + dB^n), \tag{1.2}$$

essentially, for each $d > 0$ and bounded measurable set A. Several equivalent variations of this definition, which stems from that of Sarvas [35, p. 11], are given in Lemma 4.4. The use of density points on the left of (1.2) is crucial and a feature of our methods, which differ from those in related studies of rearrangements by Brock and Solynin [9] and Van Schaftingen and Willem [46]. See [6, Appendix] and the remarks around (3.13) and at the beginning of Section 4 below for commentary about the various approaches.

It turns out that for the main classes X of interest, smoothing rearrangements are precisely the rearrangements $T : X \to X$ that *reduce the modulus of continuity* of functions in X, that is, are such that $\omega_d(Tf) \leq \omega_d(f)$ for $d > 0$ and $f \in X$, where

$$\omega_d(f) = \operatorname*{ess\,sup}_{\|x-y\| \leq d} |f(x) - f(y)|.$$

This result, a consequence of Corollary 4.12, relies on several others, such as Theorems 4.8 and 4.11, which collectively generalize (even in the special case when $K = B^n$) the theorem of Brock and Solynin [9, Theorem 3.3]. Corollary 4.12 also shows that the equivalence of smoothing and reduction of the modulus of continuity is true when only the continuous functions, or indeed the contractions, in X, are considered. Via this equivalence and results in the literature, it can be seen that the class of smoothing rearrangements on $\mathcal{V}(\mathbb{R}^n)$ includes the symmetric decreasing rearrangement (see, e.g., [3, Theorem 2.12]); more generally, the *Schwarz* or (k,n)-*Steiner* rearrangement with respect to a k-dimensional subspace in $\mathbb{R}^n$ [3, Theorem 6.10], [9, Corollary 6.1] (here the axis of the potter's wheel is $(n-k+1)$-dimensional in $\mathbb{R}^{n+1}$, $1 \leq k \leq n$, with $k = n$ corresponding to the symmetric decreasing rearrangement); and polarization with respect to a hyperplane, defined by (1.5) below [3, Proposition 1.37], [9, Lemma 5.1]. Further examples, inspired by an idea of Pólya and Szegő [32, Note B] in collaboration with M. Shiffman, we shall call *Brock rearrangements* in view of the significant extension and analysis of Brock [7, Section 3], [8, Remark 2.3]. Others still, the SC 1-symmetrizations [37, Lemma 4.4 and Definition 4.4] and their generalizations [37, Section 9], which we refer to as *Solynin rearrangements*, originate from an idea of McNabb [29]. Brock and Solynin rearrangements result from processes called continuous symmetrization and partial symmetrization (neither of which are in general true symmetrizations), respectively, that provide infinite families of rearrangements indexed by a parameter.

The Pólya-Szegő inequality (1.1) holds for each of the just-mentioned rearrangements. (For the symmetric decreasing rearrangement, references were provided above, while proofs for the Schwarz rearrangement, polarization (when (1.1) becomes an equality), Brock rearrangements, and Solynin rearrangements can be found in [3, Theorem 6.19], [9, Lemma 5.3], [8, Theorem 3.2 and Remark 3.3], and [37, Theorem 10.2], respectively.) One of the main purposes of this paper is to prove (1.1) for all smoothing rearrangements on $\mathcal{V}(\mathbb{R}^n)$; see Corollary 6.5. The initial goal in this direction, achieved in Theorem 5.8, is to show that if $T : \mathcal{V}(\mathbb{R}^n) \to \mathcal{V}(\mathbb{R}^n)$ is a smoothing rearrangement, $\Phi : [0, \infty) \to [0, \infty]$ is left-continuous and convex with $\Phi(0) = 0$ (i.e., what we call a Young function), and $f \in \mathcal{V}(\mathbb{R}^n)$ is Lipschitz, then

$$\int_{\mathbb{R}^n} \Phi\left(\|\nabla Tf(x)\|\right) dx \leq \int_{\mathbb{R}^n} \Phi\left(\|\nabla f(x)\|\right) dx, \tag{1.3}$$

where the integrals may be infinite. (For real-valued Φ, this result for the special rearrangements discussed above can be found in [3, Theorem 3.11], [3, Theorem 6.16], [3,

Proposition 3.12], with polarization again giving an equality even when $\Phi : [0, \infty) \to [0, \infty)$ is an arbitrary Borel function, and [37, Theorem 10.4], respectively.) The core of the proof of (1.3) is an inequality (see (5.10) below) between the $(n + 1)$-dimensional measure of the Minkowski sum of the part of the subgraph K_{Tf} of Tf above a fixed height and a suitable convex body $C \subset \mathbb{R}^{n+1}$, on the one hand, and the corresponding quantity for the subgraph K_f of f, on the other. (The inequality follows from a containment relation in Lemma 5.2 between horizontal sections of these two sets, that comes from (1.2) and other properties of T.) This yields an inequality between the (upper) outer Minkowski contents of the two sets (Lemma 5.3). Some results from geometric measure theory, in particular a formula of Lussardi and Villa [27], allow us to express this inequality in terms of integrals over the graphs of Tf and of f of the support function of C of the outer unit normal (Lemma 5.5). The last main step is to prove that C can be chosen so that it represents Φ (Lemma 5.6), i.e., so that the mentioned inequalities transform into (1.3). We also use the McShane-Whitney extension theorem for Lipschitz functions (Lemma 5.7).

In Theorem 6.3, we present a far-reaching version of (1.3) in $W_{loc}^{1,1}(\mathbb{R}^n)$. Specifically, we show that if $T : \mathcal{V}(\mathbb{R}^n) \to \mathcal{V}(\mathbb{R}^n)$ is a smoothing rearrangement, Φ is a Young function, $f \in W_{loc}^{1,1}(\mathbb{R}^n) \cap \mathcal{V}(\mathbb{R}^n)$, and $\int_{\mathbb{R}^n} \Phi(\|\nabla f(x)\|) \, dx < \infty$, then $Tf \in W_{loc}^{1,1}(\mathbb{R}^n)$ and (1.3) holds. This generalizes the results for Schwarz and Solynin rearrangements in [9, Theorem 8.3] and [37, Theorem 10.4]. The passage from Theorem 5.8 to Theorem 6.3 requires overcoming some technical difficulties, made all the more challenging because we do not assume that Φ is an N-function. In particular, we approximate Φ by a real-valued Young function Φ_r such that the Orlicz space $L^{\Phi_r}(\mathbb{R}^n)$ is equivalent to $L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$, the largest Orlicz space, and use both the necessary and the sufficient condition of the so-called de La Vallée-Poussin criterion. The necessary background on Orlicz spaces is provided at the beginning of Section 6. Since $W^{1,p}(\mathbb{R}^n) \subset W_{loc}^{1,1}(\mathbb{R}^n)$ for $1 \le p < \infty$, Theorem 6.3 immediately yields Corollary 6.5, the classical version (1.1) of the Pólya-Szegő inequality, but now for every smoothing rearrangement.

Finally, anisotropic versions of Theorems 5.8 and 6.3, in which the role of the unit ball B^n is replaced by a convex body $K \subset \mathbb{R}^n$ containing the origin in its interior, are proved in Theorems 7.1 and 7.3. Here the rearrangement $T : \mathcal{V}(\mathbb{R}^n) \to \mathcal{V}(\mathbb{R}^n)$ is assumed to be K-smoothing (i.e., (1.2) holds with B^n replaced by K), and then, with Φ and f as in Theorem 5.8 or Theorem 6.3, respectively, the conclusion is that

$$\int_{\mathbb{R}^n} \Phi\left(h_{-K}(\nabla Tf(x))\right) \, dx \le \int_{\mathbb{R}^n} \Phi\left(h_{-K}(\nabla f(x))\right) \, dx, \tag{1.4}$$

where h_K is the support function of K. When $K = B^n$, (1.4) becomes (1.3). This type of Pólya-Szegő inequality was introduced by Alvino, Ferone, Trombetti, and Lions [1, Theorem 3.1] when K is o-symmetric and $\Phi(t) = t^p$, $p \ge 1$, but for all $f \in W_0^{1,p}(\mathbb{R}^n)$. In their anisotropic framework, Schwarz symmetrization is replaced by one they call convex symmetrization; in the potter's wheel description above, each horizontal slice of clay

would be molded into a dilate of the convex body K. See Example 4.3(ii) below, where we call the process when it is extended to functions a K-*Schwarz rearrangement* and note that it is K-smoothing. These concepts, which align with that of Wulff shape in crystallography (see, e.g., [36, Section 7.5]), were generalized by Van Schaftingen in a process he calls partial anisotropic symmetrization in [44], where he proves a corresponding generalization of [1, Theorem 3.1]. In the rearrangements resulting from partial anisotropic symmetrizations, which we may consistently also call K-Schwarz rearrangements, K is a k-dimensional convex body in $\mathbb{R}^k$ and the axis of the potter's wheel is $(n-k+1)$-dimensional in $\mathbb{R}^{n+1}$, $1 \leq k \leq n$. We also introduce in (3.4) below the K-modulus of continuity of a function, and show in Theorem 4.8 that when $X = \mathcal{V}(\mathbb{R}^n)$, for example, a rearrangement $T : X \to X$ that reduces the K-modulus of continuity of functions in X is K-smoothing. The converse is true when K is o-symmetric (see Theorem 4.11), but Example 4.13 shows that the K-Schwarz rearrangement does not generally reduce the K-modulus of continuity of functions in X when K is not o-symmetric.

Different anisotropic extensions of the Pólya-Szegő inequality were found by Klimov [24] and Van Schaftingen [44]. Our methods can be used to prove Klimov's inequality for the K-Schwarz rearrangement of symmetrizable functions.

Known proofs of (1.1) and its variants seem to follow one of two approaches. The first, adopted in the present paper, proceeds via isoperimetric inequalities applied to (super-) level sets, while the second uses approximation by special rearrangements, principally polarizations. The second approach does not provide information about the cases of equality and moreover does not help with the anisotropic case, but otherwise can be extremely efficient. The standard *polarization* process, sometimes called *two-point symmetrization*, with respect to an *oriented* $(n-1)$-dimensional (linear) subspace H, takes a function $f : \mathbb{R}^n \to \mathbb{R}$ and replaces it by

$$P_H f(x) = \begin{cases} \max\{f(x), f(x^\dagger)\}, & \text{if } x \in H^+, \\ \min\{f(x), f(x^\dagger)\}, & \text{if } x \in H^-, \end{cases} \tag{1.5}$$

where † denotes the reflection in H and where H^+, H^-, are the two closed half-spaces bounded by H and determined by its orientation. For background and references, see [6, Introduction], where it is explained in exactly what sense all Schwarz rearrangements, including the symmetric decreasing rearrangement, can be approximated by polarizations, a result due to Brock and Solynin [9] and refined by Van Schaftingen [43,45]. Solynin [37, Lemmas 7.4 and 9.2] proved that his rearrangements can also be approximated by polarizations. We conjecture that this is not true of all smoothing rearrangements, and have some partial results in this direction.

An obvious question is whether the smoothing or K-smoothing assumptions are necessary for our Pólya-Szegő inequalities, though Example 6.6 shows that this is not the case when $p = 1$. (Example 6.6 also shows that the smoothing assumption cannot generally be omitted in Theorems 5.8 and 6.3 and Corollary 6.5.) Also left for a future investigation are the cases of equality. Even for the symmetric decreasing rearrangement, this is

challenging; see [12] and the references given there, which go back to the initial study of Brothers and Ziemer [10].

When $p = 2$, the Pólya-Szegő inequality (1.1) can be derived from the Riesz-Sobolev inequality (also called the Riesz rearrangement inequality) for the symmetric decreasing rearrangement [3, Theorem 8.4], [26, Theorem 3.7]; for a proof of this fact, see [26, Lemma 7.17]. This inequality holds for all Schwarz rearrangements but not for polarizations (see [44, Corollary 4.3]), and therefore is not true for all smoothing rearrangements. On the other hand, it is pointed out in [9, p. 1763] that a very special case of the Riesz-Sobolev inequality, the Hardy-Littlewood inequality [3, p. 54], [26, Theorem 3.4], is a simple consequence of Proposition 3.7 below, and therefore holds for all rearrangements.

We are very grateful to a referee for drawing our attention to [7] and [8].

2. Preliminaries

As usual, S^{n-1} denotes the unit sphere and o the origin in Euclidean n-space $\mathbb{R}^n$. Unless stated otherwise, we assume throughout that $n \geq 2$. The standard orthonormal basis for $\mathbb{R}^n$ is $\{e_1, \dots, e_n\}$ and the Euclidean norm is denoted by $\|\cdot\|$. The term *ball* in $\mathbb{R}^n$ will always mean a *closed* n-dimensional ball unless otherwise stated. The unit ball in $\mathbb{R}^n$ will be denoted by B^n and $B(x, r)$ is the ball with center x and radius r. We write D^n for the open unit ball in $\mathbb{R}^n$. If $x, y \in \mathbb{R}^n$ we write $x \cdot y$ for the inner product and $[x, y]$ for the line segment with endpoints x and y. If $x \in \mathbb{R}^n \setminus \{o\}$, then $x^\perp$ is the $(n-1)$-dimensional subspace orthogonal to x and $\langle x \rangle$ is the 1-dimensional subspace spanned by x. Throughout the paper, the term *subspace* means a linear subspace.

If A is a set, we denote by $\operatorname{cl} A$, $\operatorname{int} A$, and $\dim A$ the *closure*, *interior*, and *dimension* (that is, the dimension of the affine hull) of A, respectively. If H is a subspace of $\mathbb{R}^n$, then $A|H$ is the (orthogonal) projection of A on H and $x|H$ is the projection of a vector $x \in \mathbb{R}^n$ on H.

If A and B are sets in $\mathbb{R}^n$ and $t \in \mathbb{R}$, then we denote by $tA = \{tx : x \in A\}$ the *dilate* of A by the factor t, and by

$$A + B = \{x + y : x \in A, y \in B\}$$

the *Minkowski sum* of A and B. We write $-A = (-1)A$ for the reflection of A in the origin and call A *origin symmetric* or *o-symmetric* if $-A = A$.

We write $\mathcal{H}^k$ for k-dimensional Hausdorff measure in $\mathbb{R}^n$, where $k \in \{1, \dots, n\}$. When dealing with relationships between sets in $\mathbb{R}^n$ or functions on $\mathbb{R}^n$, the term *essentially* means up to a set of $\mathcal{H}^n$-measure zero.

The Grassmannian of k-dimensional subspaces in $\mathbb{R}^n$ is denoted by $\mathcal{G}(n, k)$.

We denote by $\mathcal{C}^n$, $\mathcal{G}^n$, $\mathcal{B}^n$, $\mathcal{M}^n$, and $\mathcal{L}^n$ the class of nonempty compact sets, open sets, bounded Borel sets, $\mathcal{H}^n$-measurable sets, and $\mathcal{H}^n$-measurable sets of finite $\mathcal{H}^n$-measure, respectively, in $\mathbb{R}^n$.

Let $\mathcal{K}^n$ be the class of nonempty compact convex subsets of $\mathbb{R}^n$ and let $\mathcal{K}^n_n$ be the class of *convex bodies*, i.e., members of $\mathcal{K}^n$ with interior points. We write $\mathcal{K}^n_{(o)}$ for subclass of $\mathcal{K}^n_n$ whose members contain the origin in their interiors. If $K \in \mathcal{K}^n$, then

$$h_K(x) = \sup\{x \cdot y : y \in K\} \tag{2.1}$$

for $x \in \mathbb{R}^n$, defines the *support function h_K* of K. The texts by Gruber [21] and Schneider [36] contain a wealth of useful information about convex sets and related concepts such as the *intrinsic volumes V_j, $j \in \{1, \ldots, n\}$* (see also [20, Appendix A]). In particular, if $K \in \mathcal{K}^n$ and $\dim K = n$ then $2V_{n-1}(K)$ is the *surface area* of K. If $\dim K = k$, then $V_k(K) = \mathcal{H}^k(K)$ is the *volume* of K. By κ_n we denote the volume $\mathcal{H}^n(B^n)$ of the unit ball in $\mathbb{R}^n$.

If $K \in \mathcal{K}^n_{(o)}$, the *polar body K°* of K is defined by

$$K^\circ = \{x \in \mathbb{R}^n : x \cdot y \leq 1 \text{ for } y \in K\}. \tag{2.2}$$

Then $(K^\circ)^\circ = K$ and (see [36, (1.52), p. 57])

$$\rho_K(x)h_{K^\circ}(x) = h_K(x)\rho_{K^\circ}(x) = 1 \quad \text{for } x \in \mathbb{R}^n \setminus \{o\}, \tag{2.3}$$

where

$$\rho_K(x) = \max\{\lambda \geq 0 : \lambda x \in K\} \tag{2.4}$$

for $x \in \mathbb{R}^n \setminus \{o\}$, is the *radial function* of K. We shall also find use for the *gauge function* of K, defined by

$$\|x\|_K = \inf\{\lambda \geq 0 : x \in \lambda K\} = h_{K^\circ}(x) \tag{2.5}$$

for $x \in \mathbb{R}^n$. The previous equality follows from (2.3) and (2.4), or see [36, Lemma 1.7.13]. Despite the notation, $\|\cdot\|_K$ is a norm if and only if K is o-symmetric; in general it is sublinear but does not satisfy $\|-x\|_K = \|x\|_K$ for all $x \in \mathbb{R}^n$. When $K = B^n$, $\|\cdot\|_K$ is the Euclidean norm.

It will be convenient to call a function $f : \mathbb{R}^n \to \mathbb{R}$ a *K-contraction* if

$$|f(x) - f(y)| \leq \|x - y\|_K \tag{2.6}$$

for all $x, y \in \mathbb{R}^n$. Note that when $K = B^n$, a K-contraction is a contraction in the usual sense of the term. Note also that f is a K-contraction if and only if it is a $-K$-contraction, since (2.6) is equivalent to $|f(y) - f(x)| \leq \|x - y\|_K$ for all $x, y \in \mathbb{R}^n$, and hence to $|f(x) - f(y)| \leq \|y - x\|_K = \|x - y\|_{-K}$ for all $x, y \in \mathbb{R}^n$. Clearly, every K-contraction is continuous.

From (2.1) and (2.3), it is easy to see that

$$x \cdot y \le h_K(x) \, h_{K^\circ}(y) \tag{2.7}$$

for $x, y \in \mathbb{R}^n$ (see [36, (1.40), p. 54]) and that equality holds when $x, y \ne o$ if and only if x is an outer normal to K at $\rho_K(y)y = y/h_{K^\circ}(y) \in \partial K$.

Given $A \in \mathcal{M}^n$, let $\overline{\mathcal{M}}^*(A)$ and $\overline{\mathcal{M}}_*(A)$ denote, respectively, its *upper* and *lower outer Minkowski content*, i.e.,

$$\begin{aligned}
\overline{\mathcal{M}}^*(A) &= \limsup_{\varepsilon \to 0+} \frac{\mathcal{H}^n(A + \varepsilon B^n) - \mathcal{H}^n(A)}{\varepsilon} \quad \text{and} \\
\overline{\mathcal{M}}_*(A) &= \liminf_{\varepsilon \to 0+} \frac{\mathcal{H}^n(A + \varepsilon B^n) - \mathcal{H}^n(A)}{\varepsilon}.
\end{aligned} \tag{2.8}$$

See [11, p. 69] and [2], whose notation and terminology differs from ours, and note that the limits in (2.8) are unchanged if B^n is replaced by D^n. We shall also need the following generalization of these concepts. If $C \in \mathcal{K}^n_{(o)}$ and $A \in \mathcal{M}^n$, let $\overline{\mathcal{M}}^*_C(A)$ and $\overline{\mathcal{M}}_{*C}(A)$ denote, respectively, the *upper* and *lower anisotropic outer Minkowski content* of A with respect to C, obtained by replacing B^n in (2.8) by C. When the two limits coincide we denote them by $\overline{\mathcal{M}}_C(A)$, and again, the limits are unchanged if C is replaced by $\mathrm{int}\, C$.

Let $A \in \mathcal{M}^n$. We shall write $S(A)$ for the *perimeter* of A. For the definition of this widely-used term, see, for example, [17, p. 170], [25, p. 107], [28, p. 122], or [31, p. 34]. When K is a convex body, its perimeter is equal to its surface area, defined in that case as

$$S(K) = \lim_{\varepsilon \to 0+} \frac{\mathcal{H}^n(K + \varepsilon B^n) - \mathcal{H}^n(K)}{\varepsilon},$$

its outer Minkowski content. It is for this reason that we prefer not to use the more common $P(A)$ for the perimeter of A.

Let $\mathcal{M}(\mathbb{R}^n)$ (or $\mathcal{M}_+(\mathbb{R}^n)$) denote the set of real-valued (or nonnegative, respectively) measurable functions on $\mathbb{R}^n$ and let $\mathcal{S}(\mathbb{R}^n)$ denote the set of functions f in $\mathcal{M}(\mathbb{R}^n)$ such that $\mathcal{H}^n(\{x : f(x) > t\}) < \infty$ for $t > \mathrm{ess\,inf}\, f$. By $\mathcal{V}(\mathbb{R}^n)$, we denote the set of functions f in $\mathcal{M}_+(\mathbb{R}^n)$ such that $\mathcal{H}^n(\{x : f(x) > t\}) < \infty$ for $t > 0$. The four classes of functions satisfy $\mathcal{V}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset \mathcal{M}(\mathbb{R}^n)$ and $\mathcal{V}(\mathbb{R}^n) \subset \mathcal{M}_+(\mathbb{R}^n) \subset \mathcal{M}(\mathbb{R}^n)$. Members of $\mathcal{S}(\mathbb{R}^n)$ have been called *symmetrizable* (see, e.g., [9]) and those of $\mathcal{V}(\mathbb{R}^n)$ are often said to *vanish at infinity*. Note that the constant functions are symmetrizable but do not vanish at infinity unless they are identically zero.

We shall define a *Young function* as a left-continuous and convex function $\Phi : [0, \infty) \to [0, \infty]$ such that $\Phi(0) = 0$, and say that such a function is *nontrivial* if $\Phi \not\equiv 0$ and $\Phi \not\equiv \infty$ on $(0, \infty)$. Note that a real-valued Young function is both continuous and increasing (which will always mean non-decreasing in this paper). In [16, Definition 2.1.1], the term *Orlicz function* is used for a nontrivial Young function. Both terms have other definitions in the literature.

Our notation for Sobolev spaces such as $W^{1,p}(\mathbb{R}^n)$ is standard. Definitions can be found in many texts, such as [26].

If $f \in \mathcal{M}(\mathbb{R}^n)$, we denote its *graph* by G_f and define its *subgraph* $K_f \subset \mathbb{R}^{n+1}$ by

$$K_f = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : f(x) \geq t\}. \tag{2.9}$$

If $T : X \to X$, where X is one of the function classes given above, we shall usually write Tf instead of $T(f)$. If $T_0, T_1 : X \to X$ are maps, we say that T_0 is *essentially equal* to T_1 if for $f \in X$, $T_0 f(x) = T_1 f(x)$ for $\mathcal{H}^n$-almost all $x \in \mathbb{R}^n$, where the exceptional set may depend on f.

If f is a locally integrable function on $\mathbb{R}^n$, define

$$f^*(x) = \lim_{r \to 0^+} \frac{1}{\mathcal{H}^n(B(x,r))} \int_{B(x,r)} f(y)\, dy \tag{2.10}$$

when the limit exists and $f^*(x) = 0$ otherwise. The limit exists and equals $f(x)$ $\mathcal{H}^n$-almost everywhere in $\mathbb{R}^n$, by the Lebesgue differentiation theorem (see, e.g., [25, Proposition 3.5.4]). Evans and Gariepy [17, p. 46] call f^* the *precise representative* of f. If A is a measurable set,

$$\Theta(A,x) = 1_A^*(x) = \lim_{r \to 0^+} \frac{\mathcal{H}^n(A \cap B(x,r))}{\mathcal{H}^n(B(x,r))}, \tag{2.11}$$

is the *density* of A at x, provided the limit exists.

If $A \in \mathcal{M}^n$, define

$$A^* = \{x \in \mathbb{R}^n : \Theta(A,x) = 1_A^*(x) = 1\}.$$

Elements of A^* are called *Lebesgue density points*, or simply density points, of A. Note that $A^* = A$, essentially, by the Lebesgue density theorem (see, e.g., [31, Theorem 1.5.2]). Since it follows immediately from the definition of perimeter as a supremum of integrals of divergences (see, for example, [31, p. 34]) that if two measurable sets are essentially equal, their perimeters are equal, we have

$$S(A^*) = S(A) \tag{2.12}$$

for $A \in \mathcal{M}^n$.

Lemma 2.1. *Let $A, B \in \mathcal{M}^n$.*
(i) If $A \subset B$, essentially, then $A^* \subset B^*$.
(ii) If $A = B$, essentially, then $A^* = B^*$.
(iii) $(A^*)^* = A^*$.
(iv) If $K \subset \mathbb{R}^n$ is a convex body, then $A^* + K = A^* + \operatorname{int} K$ is open.

Proof. (i) If $A \subset B$, essentially, then $1_A \leq 1_B$, essentially. Therefore, if $x \in A^*$, then

$$1 = 1_A^*(x) = \lim_{r \to 0^+} \frac{1}{\mathcal{H}^n(B(x,r))} \int_{B(x,r)} 1_A(y)\,dy \leq \liminf_{r \to 0^+} \frac{1}{\mathcal{H}^n(B(x,r))} \int_{B(x,r)} 1_B(y)\,dy$$

$$\leq \limsup_{r \to 0^+} \frac{1}{\mathcal{H}^n(B(x,r))} \int_{B(x,r)} 1_B(y)\,dy \leq 1$$

and hence $1_B^*(x) = 1$. Therefore $A^* \subset B^*$.

Parts (ii) and (iii) follow easily.

(iv) Let $x \in A^* + K$ and choose $y \in A^*$ such that $x \in y + K$. Choose $r > 0$ and an open cone C with vertex at o such that $x + (C \cap rD^n) \subset y + \operatorname{int} K$. Note that if $w \in C \cap rD^n$, then $x \in y - w + \operatorname{int} K$. Since $\mathcal{H}^n(-C \cap rD^n) > 0$, $y \in A^*$, and $A = A^*$, essentially, there is a $w_0 \in C \cap rD^n$ such that $y - w_0 \in A^*$. Hence $x \in y - w_0 + \operatorname{int} K \subset A^* + \operatorname{int} K$ and it follows from the definition of A^* that $A^* + \operatorname{int} K$ is open since $\operatorname{int} K$ is open. $\quad\square$

A function $f \in \mathcal{M}(\mathbb{R}^n)$ is *approximately continuous* at $x \in \mathbb{R}^n$ if for each $\varepsilon > 0$,

$$\lim_{r \to 0+} \frac{\mathcal{H}^n\left(B(x,r) \cap \{y : |f(x) - f(y)| < \varepsilon\}\right)}{\mathcal{H}^n(B(x,r))} = 1. \tag{2.13}$$

We shall use the fact that each $f \in \mathcal{M}(\mathbb{R}^n)$ is approximately continuous at $\mathcal{H}^n$-almost all $x \in \mathbb{R}^n$; see, for example, [17, Theorem 3, Section 1.7.2].

Note that a measurable characteristic function 1_A is approximately continuous at x if and only if the limit in (2.10) with $f = 1_A$ exists and equals $1_A(x)$, that is, if and only if either $x \in A$ and $1_A^*(x) = 1$ or $x \notin A$, the limit in (2.10) exists, and $1_A^*(x) = 0$. This means that the set of points of approximate continuity can change even when a function is only changed on a set of measure zero. Moreover, if 1_A is approximately continuous at x, then $1_A(x) = 1_{A^*}(x)$ and 1_{A^*} is also approximately continuous at x. Hence, the set of approximate continuity points of 1_{A^*} is the largest set of approximate continuity points of any 1_B for which A and B essentially coincide. In particular, if $A = B^n$, essentially, then 1_A is not approximately continuous at any unit vector. This precludes the possibility of finding a representative of each $f \in \mathcal{M}(\mathbb{R}^n)$ that is approximately continuous everywhere and agrees with f for $\mathcal{H}^n$-almost all $x \in \mathbb{R}^n$.

3. Properties of maps

Let $i \in \{1, \ldots, n-1\}$ and let $H \in \mathcal{G}(n,i)$ be fixed. We consider a map $\Diamond : \mathcal{E} \subset \mathcal{L}^n \to \mathcal{L}^n$ and define

$$\Diamond^* A = (\Diamond A)^* \tag{3.1}$$

for each $A \in \mathcal{E}$. We assume (here and throughout the paper) that the properties listed below hold for all $A, B \in \mathcal{E}$ and that the class $\mathcal{E}$ is appropriate for the property concerned.

1. (*Monotonic* or *strictly monotonic*) $A \subset B \Rightarrow \Diamond A \subset \Diamond B$, essentially (or $\Rightarrow \Diamond A \subset \Diamond B$, essentially, and $A \neq B \Rightarrow \Diamond A \neq \Diamond B$, essentially, respectively).

2. (*Measure preserving*) $\mathcal{H}^n(\Diamond A) = \mathcal{H}^n(A)$.

3. (*Maps balls to balls*)   If $K = B(x, r)$, then $\Diamond K = B(x', r')$, essentially.

4. (*Continuous from the inside*)   If (A_m) is an increasing sequence of sets in $\mathcal{E}$ such that $\cup_{m \in \mathbb{N}} A_m \in \mathcal{E}$, then $\Diamond(\cup_{m \in \mathbb{N}} A_m) = \cup_{m \in \mathbb{N}} \Diamond A_m$, essentially.

5. (*Continuous from the outside*)   If (A_m) is a decreasing sequence of sets in $\mathcal{E}$ such that $\cap_{m \in \mathbb{N}} A_m \in \mathcal{E}$, then $\Diamond(\cap_{m \in \mathbb{N}} A_m) = \cap_{m \in \mathbb{N}} \Diamond A_m$, essentially.

6. (*Smoothing and K-smoothing*)   If $K \in \mathcal{K}^n_{(o)}$, we say that $\Diamond$ is *K-smoothing* if whenever $d > 0$,

$$(\Diamond^* A) + dK \subset \Diamond^*(A + dK) = \Diamond(A + dK), \tag{3.2}$$

essentially, for each bounded $A \in \mathcal{E}$ with $A + dK \in \mathcal{E}$, where $\Diamond^* A$ is defined by (3.1). Then $\Diamond$ is called *smoothing* if it is K-smoothing with $K = B^n$.

Information concerning relations between the first three properties listed above and others besides may be found in [6, Sections 3 and 6]. The terms "continuous from the inside," "continuous from the outside," and "smoothing" are employed by Sarvas [35, p. 11], although his definitions differ slightly from those above.

In the definition of K-smoothing, one can equivalently require a pointwise inclusion in (3.2). To see this, note that by Lemma 2.1(iv) with A and K replaced by $\Diamond A$ and dK, respectively, $(\Diamond^* A) + dK = (\Diamond^* A) + d \operatorname{int} K$ is open. Then the essential inclusion in (3.2) and parts (i) and (iii) of Lemma 2.1 give

$$(\Diamond^* A) + dK = ((\Diamond^* A) + dK)^* \subset (\Diamond^*(A + dK))^* = \Diamond^*(A + dK).$$

Lemma 3.1. *If $\Diamond : \mathcal{E} \subset \mathcal{L}^n \to \mathcal{L}^n$ is monotonic and measure preserving, then $\Diamond$ is continuous from the inside and from the outside.*

Proof. Let (A_m) be an increasing sequence of sets in $\mathcal{E}$ such that $\cup_{m \in \mathbb{N}} A_m \in \mathcal{E}$. Since $\Diamond$ is monotonic, we have $\Diamond A_m \subset \Diamond(\cup_{m \in \mathbb{N}} A_m)$ for $m \in \mathbb{N}$, essentially, and hence $\cup_{m \in \mathbb{N}} \Diamond A_m \subset \Diamond(\cup_{m \in \mathbb{N}} A_m)$, essentially. The continuity of measures of increasing sequences and the fact that $\Diamond$ preserves measure yield

$$\mathcal{H}^n\left(\Diamond(\cup_{m \in \mathbb{N}} A_m)\right) = \mathcal{H}^n\left(\cup_{m \in \mathbb{N}} A_m\right) = \lim_{m \to \infty} \mathcal{H}^n(A_m) = \lim_{m \to \infty} \mathcal{H}^n(\Diamond A_m)$$
$$= \mathcal{H}^n\left(\cup_{m \in \mathbb{N}} \Diamond A_m\right).$$

It follows that $\Diamond(\cup_{m \in \mathbb{N}} A_m) = \cup_{m \in \mathbb{N}} \Diamond A_m$, essentially, and hence that $\Diamond$ is continuous from the inside.

The proof of the continuity from the outside is similar. $\square$

Let $X \subset \mathcal{M}(\mathbb{R}^n)$, where we assume henceforth that X contains the characteristic functions of sets in $\mathcal{L}^n$. Let $T : X \to X$ and if $A \in \mathcal{L}^n$, let

$$\Diamond_T A = \{x : T1_A(x) = 1\}$$

and let

$$\Diamond_T^* A = (\Diamond_T A)^*. \tag{3.3}$$

By Proposition 3.3(i) below, the induced map $\Diamond_T : \mathcal{L}^n \to \mathcal{L}^n$ is well defined when $X = \mathcal{M}(\mathbb{R}^n)$, $\mathcal{M}_+(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$, or $\mathcal{V}(\mathbb{R}^n)$. Of course, $\Diamond_T^* : \mathcal{L}^n \to \mathcal{L}^n$ is well defined whenever $\Diamond_T : \mathcal{L}^n \to \mathcal{L}^n$ is.

If $X \subset \mathcal{M}(\mathbb{R}^n)$, we consider the following properties of a map $T : X \to X$, where the first four properties are assumed to hold for all $f, g \in X$:

1. (*Equimeasurable*)

$$\mathcal{H}^n(\{x : Tf(x) > t\}) = \mathcal{H}^n(\{x : f(x) > t\})$$

for $t \in \mathbb{R}$.

2. (*Monotonic*) $f \leq g$, essentially, implies $Tf \leq Tg$, essentially.

3. (*L_p-contracting*) $\|Tf - Tg\|_p \leq \|f - g\|_p$ when $f - g \in L^p(\mathbb{R}^n)$.

4. (*Modulus of continuity reducing and K-modulus of continuity reducing*) If $d > 0$ and $K \in \mathcal{K}_{(o)}^n$, we define the *K-modulus of continuity* of $f \in X$ by

$$\omega_{K,d}(f) = \underset{\|x-y\|_K \leq d}{\operatorname{ess\,sup}} |f(x) - f(y)| = \underset{x-y \in dK}{\operatorname{ess\,sup}} |f(x) - f(y)|. \tag{3.4}$$

The equivalence of these two expressions follows easily from the left-hand equality in (2.5). Then *T reduces the K-modulus of continuity* if $\omega_{K,d}(Tf) \leq \omega_{K,d}(f)$ for all $d > 0$ and $f \in X$. When $K = B^n$, we refer simply to the *modulus of continuity* of $f \in X$ and drop the suffix K, i.e.,

$$\omega_d(f) = \underset{\|x-y\| \leq d}{\operatorname{ess\,sup}} |f(x) - f(y)|,$$

and say that *T reduces the modulus of continuity* if $\omega_d(Tf) \leq \omega_d(f)$ for all $d > 0$ and $f \in X$.

5. (*Continuous from the inside (or outside)*)   The induced map $\Diamond_T$ is well defined on $\mathcal{L}^n$ and continuous from the inside (or outside, respectively) when $\mathcal{E} = \mathcal{L}^n$.

6. (*Smoothing and K-smoothing*)   If $K \in \mathcal{K}_{(o)}^n$, we say that *$T$ is K-smoothing* if the induced map $\Diamond_T$ is well defined on $\mathcal{L}^n$ and K-smoothing when $\mathcal{E} = \mathcal{L}^n$, i.e.,

$$(\Diamond_T^* A) + dK \subset \Diamond_T^*(A + dK) = \Diamond_T(A + dK), \tag{3.5}$$

essentially, for each $d > 0$ and bounded $A \in \mathcal{M}^n$. Then *T is called smoothing* if it is K-smoothing with $K = B^n$.

The map T is called a *rearrangement* if it is equimeasurable and monotonic.

In their somewhat different setting, versions of Properties 5 and 6 (for $K = B^n$) were also considered by Brock and Solynin [9, p. 1764]. In particular, their definition of a smoothing rearrangement $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ corresponds to requiring $(\Diamond_T A) + dD^n \subset \Diamond_T (A + dD^n)$, for each $d > 0$ and $A \in \mathcal{L}^n$. However, $A + dD^n \notin \mathcal{L}^n$, in general, when $A \in \mathcal{L}^n$. Moreover, their definition is sensitive to changing T on a set of $\mathcal{H}^n$-measure zero. For example, if $T_0 f = f$ is the identity map and $T_1 f = \max\{f, 1_{\mathbb{Q}^n}\}$, then $T_0 = T_1$, essentially, while $\Diamond_{T_0} A = A$ and $\Diamond_{T_1} A = A \cup \mathbb{Q}^n$ implies that T_0 is smoothing but T_1 is not under their definition.

Our definitions of smoothing and K-smoothing are examined further in Lemma 4.4 below. See also the remarks at the beginning of Section 4.

For the convenience of the reader, we now state five results proved in [6] as Lemmas 4.1, 4.5, 4.7, Theorem 4.8 and the remarks that follow it, and Theorem 4.9, respectively.

Proposition 3.2. (i) *If $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is equimeasurable, then $\operatorname{ess\,inf} Tf = \operatorname{ess\,inf} f$ for $f \in \mathcal{S}(\mathbb{R}^n)$.*
(ii) *If $T : \mathcal{M}(\mathbb{R}^n) \to \mathcal{M}(\mathbb{R}^n)$ is a rearrangement, then $\operatorname{ess\,inf} Tf \geq \operatorname{ess\,inf} f$ for $f \in \mathcal{M}(\mathbb{R}^n)$. Hence, $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$.*
(iii) *In either case, $T : \mathcal{V}(\mathbb{R}^n) \to \mathcal{V}(\mathbb{R}^n)$ and T is essentially the identity on constant functions.*

Proposition 3.3. *Let $X = \mathcal{M}(\mathbb{R}^n)$, $\mathcal{M}_+(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$, or $\mathcal{V}(\mathbb{R}^n)$, and let $T : X \to X$ be equimeasurable.*
(i) *The induced map $\Diamond_T : \mathcal{L}^n \to \mathcal{L}^n$ given by*

$$\Diamond_T A = \{x : T1_A(x) = 1\} \tag{3.6}$$

for $A \in \mathcal{L}^n$ is well defined and measure preserving.
(ii) *If $X = \mathcal{M}_+(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$, or $\mathcal{V}(\mathbb{R}^n)$, then T essentially maps characteristic functions of sets in $\mathcal{L}^n$ to characteristic functions of sets in $\mathcal{L}^n$, in the sense that for each $A \in \mathcal{L}^n$,*

$$T1_A = 1_{\Diamond_T A}, \tag{3.7}$$

essentially.

Proposition 3.4. *Let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$ and let $T : X \to X$ be a rearrangement. For $X = \mathcal{S}(\mathbb{R}^n)$, $A \in \mathcal{L}^n$, and $\alpha, \beta \in \mathbb{R}$ with $\alpha \geq 0$, we have*

$$T(\alpha 1_A + \beta) = \alpha T1_A + \beta, \tag{3.8}$$

essentially. When $X = \mathcal{V}(\mathbb{R}^n)$, (3.8) holds, essentially, if $\beta = 0$.

Proposition 3.5. *Let $X = \mathcal{M}(\mathbb{R}^n)$, $\mathcal{M}_+(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$, or $\mathcal{V}(\mathbb{R}^n)$ and let $T : X \to X$ be a rearrangement.*

(i) *The map $\Diamond_T : \mathcal{L}^n \to \mathcal{L}^n$ defined by (3.6) is monotonic.*

(ii) *If $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$ and $f \in X$, then*

$$\{x : Tf(x) \geq t\} = \Diamond_T\{x : f(x) \geq t\} \quad and \quad \{x : Tf(x) > t\} = \Diamond_T\{x : f(x) > t\}, \quad (3.9)$$

essentially, for $t > \operatorname{ess\,inf} f$. Moreover, T is essentially determined by $\Diamond_T$, since

$$Tf(x) = \max\left\{\sup\{t \in \mathbb{Q}, \, t > \operatorname{ess\,inf} f : x \in \Diamond_T\{z : f(z) \geq t\}\}, \operatorname{ess\,inf} f\right\}, \quad (3.10)$$

essentially.

Proposition 3.6. *Let $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ be a rearrangement and let $f \in \mathcal{S}(\mathbb{R}^n)$. If $\varphi : \mathbb{R} \to \mathbb{R}$ is right-continuous and increasing (i.e., non-decreasing), then $\varphi \circ f \in \mathcal{S}(\mathbb{R}^n)$ and*

$$\varphi(Tf) = T(\varphi \circ f), \quad (3.11)$$

essentially.

The following result was first proved, without the assumption that j is nonnegative, by Crowe, Zweibel, and Rosenbloom [13] for Schwarz rearrangement. Versions of it have been stated for general rearrangements in [9, Theorem 3.1], [42, Proposition 3.3.9], and [46, Corollary 1]; however, these works take a different approach to rearrangements, so we provide a proof and brief commentary in the Appendix.

Proposition 3.7. *Let $j : \mathbb{R} \to [0, \infty)$ be convex with $j(0) = 0$. If $T : \mathcal{V}(\mathbb{R}^n) \to \mathcal{V}(\mathbb{R}^n)$ is a rearrangement, then*

$$\int_{\mathbb{R}^n} j(Tf(x) - Tg(x)) \, dx \leq \int_{\mathbb{R}^n} j(f(x) - g(x)) \, dx \quad (3.12)$$

for $f, g \in \mathcal{V}(\mathbb{R}^n)$ such that either integral exists. In particular, T has the L^p-contracting property.

Lemma 3.8. *Let $X = \mathcal{M}(\mathbb{R}^n)$, $\mathcal{M}_+(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$, or $\mathcal{V}(\mathbb{R}^n)$. If $T : X \to X$ is a rearrangement, then the maps $\Diamond_T, \Diamond_T^* : \mathcal{L}^n \to \mathcal{L}^n$ are well defined, measure preserving, and monotonic (pointwise monotonic, in the case of $\Diamond_T^*$, i.e., $A \subset B \Rightarrow \Diamond_T^* A \subset \Diamond_T^* B$). Moreover, T and $\Diamond_T^*$ are continuous from the inside and from the outside.*

Proof. The induced map $\Diamond_T : \mathcal{L}^n \to \mathcal{L}^n$ is well defined, measure preserving, and monotonic by Propositions 3.3(i) and 3.5(i). It follows that $\Diamond_T^* : \mathcal{L}^n \to \mathcal{L}^n$ is well defined and measure preserving by the Lebesgue density theorem, and pointwise monotonic by Lemma 2.1(i).

Lemma 3.1 with $\mathcal{E} = \mathcal{L}^n$ and $\Diamond$ replaced by $\Diamond_T$ shows that T is continuous from the inside and from the outside. The fact that $\Diamond_T^*$ is also continuous from the inside and from the outside is then an easy consequence of the Lebesgue density theorem. $\square$

It is convenient to state the following lemma for the induced maps $\Diamond_T$ of a rearrangement T, but it holds more generally for any monotonic map $\Diamond : \mathcal{E} \subset \mathcal{L}^n \to \mathcal{L}^n$ and $A, B \in \mathcal{E}$.

Lemma 3.9. *Let $X = \mathcal{M}(\mathbb{R}^n)$, $\mathcal{M}_+(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$, or $\mathcal{V}(\mathbb{R}^n)$, let $T : X \to X$ be a rearrangement, and let $A, B \in \mathcal{L}^n$.*
(i) If $A \subset B$, essentially, then $\Diamond_T^* A \subset \Diamond_T^* B$.
(ii) If $A = B$, essentially, then $\Diamond_T^* A = \Diamond_T^* B$.
(iii) $\Diamond_T^* A = \Diamond_T^* A^*$.
(iv) $\Diamond_T^* A = \Diamond_T A = \Diamond_T A^*$, essentially.
(v) For $f \in X$ and $s \geq t > \operatorname{ess\,inf} f$, we have $\Diamond_T^* \{z : f(z) \geq s\} \subset \Diamond_T^* \{z : f(z) \geq t\}$.

Proof. (i) Since $\Diamond_T$ is monotonic by Lemma 3.8, we have $\Diamond_T A \subset \Diamond_T B$, essentially. The conclusion follows from (3.3) and Lemma 2.1(i).

Parts (ii) and (iii) follow easily, the latter using the fact that $A = A^*$, essentially. The latter equality and the monotonicity of $\Diamond_T$ yield the second equality in (iv), while the first is a consequence of (3.3) and $\Diamond_T A = (\Diamond_T A)^*$, essentially.

Part (v) follows from (i) and the fact that $\{z : f(z) \geq s\} \subset \{z : f(z) \geq t\}$ for $s \geq t > \operatorname{ess\,inf} f$. $\square$

Let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$ and let $T : X \to X$ be a rearrangement. Using (3.10), the Lebesgue density theorem and the fact that the supremum is over a countable set of values, and Lemma 3.9(v), we obtain

$$
\begin{aligned}
Tf(x) &= \max\left\{\sup\{t \in \mathbb{Q}, \, t > \operatorname{ess\,inf} f : x \in \Diamond_T\{z : f(z) \geq t\}\}, \operatorname{ess\,inf} f\right\} \\
&= \max\left\{\sup\{t \in \mathbb{Q}, \, t > \operatorname{ess\,inf} f : x \in \Diamond_T^*\{z : f(z) \geq t\}\}, \operatorname{ess\,inf} f\right\} \\
&= \max\left\{\sup\{t \in \mathbb{R}, \, t > \operatorname{ess\,inf} f : x \in \Diamond_T^*\{z : f(z) \geq t\}\}, \operatorname{ess\,inf} f\right\}, \quad (3.13)
\end{aligned}
$$

essentially. This shows that by substituting $\Diamond_T^*$ for $\Diamond_T$ in (3.10), we may take the supremum over $\mathbb{R}$ and thus bring the formula into line with those in [9] and [46]; see the discussion in [6, Appendix].

4. Smoothing rearrangements and reduction of the modulus of continuity

Before embarking on the main goal of this section, we prove the inequalities (4.1) below for smoothing rearrangements. These will not be needed for the sequel, but seem interesting and follow fairly easily from what we know so far.

Lemma 4.1. *Suppose that $\mathcal{K}^n_n \subset \mathcal{E} \subset \mathcal{L}^n$ and that $\lozenge : \mathcal{E} \to \mathcal{L}^n$ is measure preserving and smoothing. Then $S(\lozenge K) \leq S(K)$ for $K \in \mathcal{K}^n_n$ and hence $\lozenge$ maps balls to balls.*

Proof. Let $K \in \mathcal{K}^n_n$. For $\varepsilon > 0$, the assumed properties of $\lozenge$ imply that

$$\mathcal{H}^n((\lozenge^* K) + \varepsilon B^n) \leq \mathcal{H}^n(\lozenge(K + \varepsilon B^n)) = \mathcal{H}^n(K + \varepsilon B^n)$$

and $\mathcal{H}^n(\lozenge^* K) = \mathcal{H}^n(K)$. It follows, using (2.12) with $A = \lozenge K$ and the relation between the (lower) outer Minkowski content (defined by (2.8)) and perimeter [11, Theorem 14.2.1], that

$$
\begin{aligned}
S(\lozenge K) = S(\lozenge^* K) \leq \overline{\mathcal{M}}_*(\lozenge^* K) &= \liminf_{\varepsilon \to 0+} \frac{\mathcal{H}^n((\lozenge^* K) + \varepsilon B^n) - \mathcal{H}^n(\lozenge^* K)}{\varepsilon} \\
&\leq \lim_{\varepsilon \to 0+} \frac{\mathcal{H}^n(K + \varepsilon B^n) - \mathcal{H}^n(K)}{\varepsilon} = S(K).
\end{aligned}
$$

Consequently, $\lozenge K$ has finite perimeter. The fact that $\lozenge$ maps balls to balls is now a direct consequence of the isoperimetric inequality for sets in $\mathcal{L}^n$ of finite perimeter and its equality condition (see [28, p. 165]). $\square$

Recall that $\overline{\mathcal{M}}^*(A)$ and $\overline{\mathcal{M}}_*(A)$ are the upper and lower outer Minkowski content of A, defined by (2.8).

Theorem 4.2. *Let $X = \mathcal{M}(\mathbb{R}^n)$, $\mathcal{M}_+(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$, or $\mathcal{V}(\mathbb{R}^n)$ and suppose that $T : X \to X$ is a rearrangement. If T is smoothing, then*

$$\overline{\mathcal{M}}^*(\lozenge^*_T A) \leq \overline{\mathcal{M}}^*(A^*) \quad \text{and} \quad \overline{\mathcal{M}}_*(\lozenge^*_T A) \leq \overline{\mathcal{M}}_*(A^*) \tag{4.1}$$

for bounded $A \in \mathcal{M}^n$. Moreover, $S(\lozenge_T K) \leq S(K)$ for $K \in \mathcal{K}^n_n$ and hence $\lozenge_T$ maps balls to balls.

Proof. Let $A \in \mathcal{M}^n$ be bounded and let $\varepsilon > 0$. Then the fact that $\lozenge^*_T A = \lozenge^*_T A^*$ by Lemma 3.9(iii), (3.5) with A replaced by A^*, Lemma 3.9(iv) with A replaced by $A^* + \varepsilon B^n$, and the equimeasurability of T imply that

$$\mathcal{H}^n((\lozenge^*_T A) + \varepsilon B^n) \leq \mathcal{H}^n(\lozenge^*_T(A^* + \varepsilon B^n)) = \mathcal{H}^n(\lozenge_T(A^* + \varepsilon B^n)) = \mathcal{H}^n(A^* + \varepsilon B^n) \tag{4.2}$$

and

$$\mathcal{H}^n(\lozenge^*_T A) = \mathcal{H}^n(\lozenge_T A) = \mathcal{H}^n(A) = \mathcal{H}^n(A^*). \tag{4.3}$$

We obtain the inequalities (4.1) directly from (2.8), (4.2), and (4.3).

The second statement in the lemma follows directly from Lemma 4.1 with $\lozenge$ replaced by $\lozenge_T$, which is valid by Lemma 3.8. $\square$

Example 4.3. (i) Let $H = e_n^\perp$ and define $\Diamond : \mathcal{K}_n^n \to \mathcal{K}_n^n$ by $\Diamond K = M_H K$, the Minkowski symmetral of K (see [4, Section 3]). Then $\Diamond$ is smoothing but not measure preserving and does not satisfy $S(\Diamond K) \le S(K)$ for each $K \in \mathcal{K}_n^n$. This shows that the measure-preserving assumption in Lemma 4.1 cannot be dropped.

(ii) Let $K \in \mathcal{K}_{(o)}^n$ and define $\Diamond : \mathcal{L}^n \to \mathcal{K}_n^n$ by $\Diamond A = r_A K$, where $r_A = (\mathcal{H}^n(A)/\mathcal{H}^n(K))^{1/n}$. This map, which corresponds to the convex symmetrization in [1], is monotonic and measure preserving and, using (3.2) and the Brunn-Minkowski inequality [19], it is easy to see that it is K-smoothing. Clearly, $\Diamond$ maps balls to balls if and only if K is a ball, and we claim that it is smoothing if and only if K is an o-symmetric ball. To see this, note that by definition, $\Diamond$ is smoothing if

$$\mathcal{H}^n(A)^{1/n} K + d\mathcal{H}^n(K)^{1/n} B^n \subset \mathcal{H}^n(A + dB^n)^{1/n} K$$

for all $A \in \mathcal{L}^n$. If $A = K$, this implies that

$$\mathcal{H}^n(K)^{1/n}(K + dB^n) \subset \mathcal{H}^n(K + dB^n)^{1/n} K.$$

Since the sets on both sides of this inclusion have the same volume, the inclusion must be an equality and hence

$$\mathcal{H}^n(K)^{1/n} K + \mathcal{H}^n(K)^{1/n} dB^n = \mathcal{H}^n(K)^{1/n} K + \left(\mathcal{H}^n(K + dB^n)^{1/n} - \mathcal{H}^n(K)^{1/n}\right) K.$$

The cancelation law [36, p. 139] for Minkowski addition yields

$$\mathcal{H}^n(K)^{1/n} dB^n = \left(\mathcal{H}^n(K + dB^n)^{1/n} - \mathcal{H}^n(K)^{1/n}\right) K,$$

which holds if and only if $K = rB^n$ for some $r \ge 0$. Thus, the smoothing assumption cannot be omitted in Lemma 4.1.

Now suppose that $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$ and $T : X \to X$ is the rearrangement given, for all $x \in \mathbb{R}^n$, by (3.10) or (3.13) with $\Diamond_T = \Diamond$. Then the superlevel sets of Tf are dilates of K, so T is the K-*Schwarz rearrangement* mentioned in the Introduction, and again, unless K is an o-symmetric ball, T is K-smoothing but not smoothing. Therefore the smoothing assumption in Theorem 4.2 cannot be replaced by K-smoothing for any non-spherical convex body K.

See Example 4.13 for more information about the rearrangement T. $\square$

In the rest of this section, we study the relationship between the K-smoothing and reduction of the K-modulus of continuity properties of a rearrangement $T : X \to X$, where $X = \mathcal{M}(\mathbb{R}^n)$, $\mathcal{M}_+(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$, or $\mathcal{V}(\mathbb{R}^n)$. When $K = B^n$, some information of this type was obtained by Brock and Solynin in [9, Theorem 3.3], which states that a rearrangement (in their sense of the term) that is continuous from the inside is smoothing if and only if it reduces the modulus of continuity of continuous functions in $\mathcal{S}(\mathbb{R}^n)$. A comparison of their approach to rearrangements and ours can be found in [6, Appendix].

When $K = B^n$, it is possible to use [9, Theorem 3.3] to obtain the same result for our rearrangements, i.e., the equivalence (ii)⇔(iii) of Corollary 4.12 below. To see this, note firstly that the continuity from the inside assumption is not necessary in our context, by Lemma 3.8. If $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a rearrangement, then by Lemma 3.8, the set transformation $\Diamond_T^* : \mathcal{L}^n \to \mathcal{L}^n$ is a rearrangement in the sense of [9, p. 1762], since it is pointwise monotonic. If $\overline{T} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ denotes the rearrangement map induced by $\Diamond_T^*$ via [9, (3.1), p. 1762], then (3.13) yields $\overline{T} = T$, so [9, Theorem 3.3] is valid for T.

However, even when $K = B^n$, Theorems 4.8 and 4.11 below are more general than [9, Theorem 3.3], since they apply to much wider classes of functions.

In (ii) of the following lemma, we assume that $A \in \mathcal{M}^n$ is bounded to ensure that $A + d \operatorname{int} K \in \mathcal{L}^n$ when $d > 0$. This seems unavoidable since $A + d \operatorname{int} K \notin \mathcal{L}^n$ when A is unbounded, in which case $\Diamond_T^*(A + d \operatorname{int} K)$ is not defined.

Lemma 4.4. *Let $X = \mathcal{M}(\mathbb{R}^n)$, $\mathcal{M}_+(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$, or $\mathcal{V}(\mathbb{R}^n)$, let $T : X \to X$ be a rearrangement, and let $K \in \mathcal{K}_{(o)}^n$. The following statements are equivalent.*
(i) T is K-smoothing.
(ii) For each $d > 0$ and bounded $A \in \mathcal{M}^n$, we have

$$(\Diamond_T^* A) + d \operatorname{int} K \subset \Diamond_T^*(A + d \operatorname{int} K). \tag{4.4}$$

(iii) For each $d > 0$ and bounded $A \in \mathcal{M}^n$, (4.4) holds essentially.
(iv) For each $d > 0$ and $A \in \mathcal{L}^n$, we have

$$(\Diamond_T^* A) + d \operatorname{int} K \subset \cup\{\Diamond_T^* E : E \in \mathcal{L}^n, \ E \subset A + d \operatorname{int} K\}, \tag{4.5}$$

essentially.

Proof. (i)⇒ (iii) Let $d > 0$, let $A \in \mathcal{M}^n$ be bounded, and choose $N \in \mathbb{N}$ so that $1/N < d$. Using the K-smoothing property of T and the continuity of $\Diamond_T^*$ from the inside provided by Lemma 3.8, we obtain

$$(\Diamond_T^* A) + d \operatorname{int} K = \cup_{m=N}^\infty ((\Diamond_T^* A) + (d - 1/m)K) \subset \cup_{m=N}^\infty \Diamond_T^*(A + (d - 1/m)K)$$
$$= \Diamond_T^*(\cup_{m=N}^\infty (A + (d - 1/m)K)) = \Diamond_T^*(A + d \operatorname{int} K),$$

essentially. This proves (4.4).

(iii)⇒ (ii) Let $d > 0$ and let $A \in \mathcal{M}^n$ be bounded. From the fact that $(\Diamond_T^* A) + d \operatorname{int} K$ is open, the essential inclusion (4.4), and parts (i) and (iii) of Lemma 2.1, we obtain

$$(\Diamond_T^* A) + d \operatorname{int} K = ((\Diamond_T^* A) + d \operatorname{int} K)^* \subset \Diamond_T^*(A + d \operatorname{int} K),$$

proving (ii).

(ii)⇒(i) Let $d > 0$ and let $A \in \mathcal{M}^n$ be bounded. We use (4.4) and the continuity of $\Diamond_T^*$ from the outside from Lemma 3.8 to get

$$(\lozenge_T^* A) + dK \subset \cap_{m \in \mathbb{N}} ((\lozenge_T^* A) + (d + 1/m) \operatorname{int} K) \subset \cap_{m \in \mathbb{N}} \lozenge_T^* (A + (d + 1/m) \operatorname{int} K)$$
$$= \lozenge_T^* (\cap_{m \in \mathbb{N}} (A + (d + 1/m) \operatorname{int} K)) = \lozenge_T^* (A + dK),$$

essentially, so T is K-smoothing.

(iii)$\Rightarrow$(iv) Let $d > 0$, let $A \in \mathcal{L}^n$, and define $A_m = A \cap m \operatorname{int} K$ for $m \in \mathbb{N}$. Since $\lozenge_T^*$ is continuous from the inside by Lemma 3.8, we obtain

$$(\lozenge_T^* A) + d \operatorname{int} K = (\lozenge_T^* (\cup_{m \in \mathbb{N}} A_m)) + d \operatorname{int} K$$
$$= (\cup_{m \in \mathbb{N}} \lozenge_T^* A_m) + d \operatorname{int} K = \cup_{m \in \mathbb{N}} (\lozenge_T^* A_m + d \operatorname{int} K)$$
$$\subset \cup_{m \in \mathbb{N}} \lozenge_T^* (A_m + d \operatorname{int} K) \subset \cup \{\lozenge_T^* E : E \in \mathcal{L}^n, \ E \subset A + d \operatorname{int} K\},$$

essentially, and (4.5) follows.

(iv)$\Rightarrow$(iii) Let $d > 0$ and let $A \in \mathcal{M}^n$ be bounded. Applying (4.5), we immediately obtain (4.4). $\square$

Recall the definition (2.13) of approximate continuity.

Lemma 4.5. *Let $d > 0$, let $f \in \mathcal{M}(\mathbb{R}^n)$, and let C be the set of points of approximate continuity of f. If $K \in \mathcal{K}_{(o)}^n$, then*

$$\omega_{K,d}(f) = \operatorname*{ess\,sup}_{\|x-y\|_K \leq d} |f(x) - f(y)| = \sup_{\|x-y\|_K \leq d; \ x,y \in C} |f(x) - f(y)|.$$

Proof. Let d, f, and C be as in the statement of the lemma. Since $C = \mathbb{R}^n$, essentially, we have

$$\omega_{K,d}(f) = \operatorname*{ess\,sup}_{\|x-y\|_K \leq d; \ x,y \in C} |f(x) - f(y)|.$$

Moreover, it is clear that

$$\operatorname*{ess\,sup}_{\|x-y\|_K \leq d; \ x,y \in C} |f(x) - f(y)| \leq \sup_{\|x-y\|_K \leq d; \ x,y \in C} |f(x) - f(y)| = s,$$

say. To prove the reverse of the previous inequality, we may assume that $s > 0$, since if $s = 0$, it is trivial. Let $0 < \varepsilon < s/2$. It suffices to show that

$$\mathcal{H}^{2n} (\{(x, y) \in C \times C : \|x - y\|_K \leq d \text{ and } |f(x) - f(y)| > s - \varepsilon\}) > 0. \qquad (4.6)$$

To this end, choose $\bar{x}, \bar{y} \in C$ with $\bar{d} = \|\bar{x} - \bar{y}\|_K \leq d$, such that

$$|f(\bar{x}) - f(\bar{y})| > s - \varepsilon/2,$$

and note that $\bar{d} > 0$. For $x \in \mathbb{R}^n$ and $r > 0$, let

$$A_{x,r} = \{y \in B(x,r) \cap C : |f(x) - f(y)| < \varepsilon/8\}.$$

If $(x,y) \in A_{\bar{x},r} \times A_{\bar{y},r}$, then

$$|f(x) - f(y)| \geq |f(\bar{x}) - f(\bar{y})| - \varepsilon/4 > s - \varepsilon. \tag{4.7}$$

Define $E_d = \{(x,y) \in \mathbb{R}^{2n} : \|x - y\|_K \leq d\}$. We aim to prove that $\mathcal{H}^{2n}\big((A_{\bar{x},r} \times A_{\bar{y},r}) \cap E_d\big) > 0$ for small $r > 0$. Since

$$(A_{\bar{x},r} \times A_{\bar{y},r}) \cap E_d = (A_{\bar{x},r} \times A_{\bar{y},r}) \setminus \big((A_{\bar{x},r} \times A_{\bar{y},r}) \setminus E_d\big)$$

and

$$(A_{\bar{x},r} \times A_{\bar{y},r}) \setminus E_d \subset (B(\bar{x},r) \times B(\bar{y},r)) \setminus E_d,$$

we have

$$\mathcal{H}^{2n}\left((A_{\bar{x},r} \times A_{\bar{y},r}) \cap E_d\right) \geq \mathcal{H}^{2n}(A_{\bar{x},r} \times A_{\bar{y},r})$$
$$- \mathcal{H}^{2n}(B(\bar{x},r) \times B(\bar{y},r)) + \mathcal{H}^{2n}\left((B(\bar{x},r) \times B(\bar{y},r)) \cap E_d\right). \tag{4.8}$$

The approximate continuity of f at $\bar{x}$ and $\bar{y}$ yields

$$\lim_{r \to 0} \frac{\mathcal{H}^{2n}(A_{\bar{x},r} \times A_{\bar{y},r})}{\mathcal{H}^{2n}(B(\bar{x},r) \times B(\bar{y},r))} = \lim_{r \to 0} \frac{\mathcal{H}^n(A_{\bar{x},r})}{\mathcal{H}^n(B(\bar{x},r))} \lim_{r \to 0} \frac{\mathcal{H}^n(A_{\bar{y},r})}{\mathcal{H}^n(B(\bar{y},r))} = 1. \tag{4.9}$$

Let $\Delta = \{(x,x) : x \in \mathbb{R}^n\}$ be the diagonal in $\mathbb{R}^{2n}$. Now $\|x - y\|_K \leq \bar{d}$ if and only if $x \in \bar{d}K + y$, which holds if and only if $(x,y) \in (\bar{d}K \times \{o\}) + (y,y)$. It follows that

$$E_{\bar{d}} = (\bar{d}K \times \{o\}) + \Delta \tag{4.10}$$

is a $2n$-dimensional convex cylinder in $\mathbb{R}^{2n}$ as $\bar{d} > 0$. Since $\bar{d} = \|\bar{x} - \bar{y}\|_K$, we have $(\bar{x}, \bar{y}) \in (\partial(\bar{d}K) \times \{o\}) + (\bar{y}, \bar{y})$ and hence $(\bar{x}, \bar{y}) \in \partial E_{\bar{d}}$. By the convexity of $E_{\bar{d}}$,

$$\lim_{r \to 0} \frac{\mathcal{H}^{2n}\left((B(\bar{x},r) \times B(\bar{y},r)) \cap E_{\bar{d}}\right)}{\mathcal{H}^{2n}\left(B(\bar{x},r) \times B(\bar{y},r)\right)} \geq \lim_{r \to 0} \frac{\mathcal{H}^{2n}\left(B((\bar{x},\bar{y}),r) \cap E_{\bar{d}}\right)}{\mathcal{H}^{2n}\left(B(\bar{x},r) \times B(\bar{y},r)\right)}$$
$$= a = a(\bar{x},\bar{y}) > 0, \tag{4.11}$$

as the density of $E_{\bar{d}}$ at the boundary point $(\bar{x}, \bar{y})$ is positive. Since $\bar{d} \leq d$, we have $E_{\bar{d}} \subset E_d$ and (4.11) implies that

$$\lim_{r \to 0} \frac{\mathcal{H}^{2n}\left((B(\bar{x},r) \times B(\bar{y},r)) \cap E_d\right)}{\mathcal{H}^{2n}(B(\bar{x},r) \times B(\bar{y},r))} \geq a. \tag{4.12}$$

From (4.8), (4.9), and (4.12), we conclude that

$$\lim_{r \to 0} \frac{\mathcal{H}^{2n}\left((A_{\bar{x},r} \times A_{\bar{y},r}) \cap E_d\right)}{\mathcal{H}^{2n}(B(\bar{x},r) \times B(\bar{y},r))} \geq a > 0. \tag{4.13}$$

Finally, (4.7) and (4.13) imply (4.6). □

Lemma 4.6. *Let $A \subset \mathbb{R}^n$, let $d > 0$, and let $K \in \mathcal{K}^n_{(o)}$. For $x \in \mathbb{R}^n$, define*

$$d_K(x, A) = \inf\{\|x - y\|_K : y \in A\} = \inf\{\lambda \geq 0 : x \in \lambda K + y \text{ for some } y \in A\}$$

and

$$f_A(x) = (d - d_K(x, A))^+, \tag{4.14}$$

where s^+ is the nonnegative part of $s \in \mathbb{R}$. Then f_A is a K-contraction as defined in (2.6).

Proof. Let $x, y \in \mathbb{R}^n$. If $x, y \notin A + dK$, then (4.14) implies that $|f_A(x) - f_A(y)| = 0 \leq \|x - y\|_K$. Otherwise, we may, by relabeling if necessary, assume that $x \in A + dK$ and $d_K(x, A) \leq d_K(y, A)$. Let $\delta > 0$ and choose $x' \in A$ such that $\|x - x'\|_K < d_K(x, A) + \delta$. Then

$$|f_A(x) - f_A(y)| \leq d_K(y, A) - d_K(x, A) < \|y - x'\|_K - \|x - x'\|_K + \delta \leq \|y - x\|_K + \delta.$$

Therefore $|f_A(x) - f_A(y)| \leq \|y - x\|_K = \|x - y\|_{-K}$. As was noted directly after (2.6), this proves that f_A is a K-contraction. □

Lemma 4.7. (i) *If $T : \mathcal{M}(\mathbb{R}^n) \to \mathcal{M}(\mathbb{R}^n)$ is a rearrangement, then* (3.7) *holds, essentially.*
(ii) *Let $X = \mathcal{M}(\mathbb{R}^n)$ or $\mathcal{M}_+(\mathbb{R}^n)$ and suppose that $T : X \to X$ is a rearrangement. If $\alpha > 0$ and $\beta = 0$, then* (3.8) *holds, essentially.*
(iii) *Let $X = \mathcal{M}(\mathbb{R}^n)$ or $\mathcal{M}_+(\mathbb{R}^n)$ and suppose that $T : X \to X$ is a rearrangement. Then* (3.8) *holds, essentially, if $\alpha = 0$ and $\beta \in \mathbb{R}$ when $X = \mathcal{M}(\mathbb{R}^n)$, and also if $\alpha = 0$ and $\beta \geq 0$ when $X = \mathcal{M}^+(\mathbb{R}^n)$ and T reduces the K-modulus of continuity for some $K \in \mathcal{K}^n_{(o)}$.*

Proof. (i) By Proposition 3.2(iii), $T : \mathcal{V}(\mathbb{R}^n) \to \mathcal{V}(\mathbb{R}^n)$. Let $A \in \mathcal{L}^n$. Since $1_A \in \mathcal{V}(\mathbb{R}^n)$, the result follows from Proposition 3.3(ii).

(ii) By (i), (3.7) holds, essentially, when $X = \mathcal{M}(\mathbb{R}^n)$. With this in hand, the second paragraph of the proof of Proposition 3.4 can be followed verbatim.

(iii) If $X = \mathcal{M}(\mathbb{R}^n)$, the result follows from Proposition 3.2(iii). The latter fails when $X = \mathcal{M}^+(\mathbb{R}^n)$, by [6, Example 4.3], but if $\beta \geq 0$, T reduces the K-modulus of continuity for some $K \in \mathcal{K}^n_{(o)}$, and $f = \beta$, essentially, then $\omega_{K,d}(Tf) = \omega_{K,d}(f) = 0$ for all $d > 0$. This and the equimeasurability of T give $Tf = \beta$, essentially. □

Theorem 4.8. *Let $X = \mathcal{M}(\mathbb{R}^n)$, $\mathcal{M}_+(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$, or $\mathcal{V}(\mathbb{R}^n)$, and let $K \in \mathcal{K}_{(o)}^n$. If $T : X \to X$ is a rearrangement that reduces the K-modulus of continuity of each K-contraction in X, then T is K-smoothing.*

Proof. Let $d > 0$, let $A \in \mathcal{M}^n$ be bounded, and let f_A be defined by (4.14). By Lemma 4.6, $f_A \in \mathcal{V}(\mathbb{R}^n)$ is a K-contraction, so $\omega_{K,d}(f_A) \leq d$ for all $d > 0$, by the definition (3.4) of the K-modulus of continuity. If N_0 is the complement of the set of points of approximate continuity of Tf_A, then $\mathcal{H}^n(N_0) = 0$. Lemma 4.5 implies that

$$|Tf_A(x) - Tf_A(y)| \leq \omega_{K, \|x-y\|_K}(Tf_A) \leq \|x - y\|_K$$

for $x, y \in \mathbb{R}^n \setminus N_0$, so Tf_A is a K-contraction on $\mathbb{R}^n \setminus N_0$. Since T is monotonic and $d1_{A+d\,\mathrm{int}\,K} \geq f_A \geq d1_A$, we have

$$T(d1_{A+d\,\mathrm{int}\,K})(x) \geq Tf_A(x) \geq T(d1_A)(x)$$

for $x \in \mathbb{R}^n \setminus N_1$, where $\mathcal{H}^n(N_1) = 0$.

By Proposition 3.3(ii) and Lemma 4.7(i), there is a set N_2 with $\mathcal{H}^n(N_2) = 0$ such that

$$T1_{A+d\,\mathrm{int}\,K}(x) = 1_{\Diamond_T(A+d\,\mathrm{int}\,K)}(x) \tag{4.15}$$

for $x \in \mathbb{R}^n \setminus N_2$. By Proposition 3.4 and Lemma 4.7(ii), there is a set N_3 with $\mathcal{H}^n(N_3) = 0$ such that (3.8) holds everywhere on $\mathbb{R}^n \setminus N_3$ when $\alpha = d$, $\beta = 0$, and $A = A$ or $A = A + d\,\mathrm{int}\,K$. Let $N = \cup\{N_i : i = 0, 1, 2, 3\}$.

Let $x_0 \in ((\Diamond_T^* A) + d\,\mathrm{int}\,K) \setminus N$. There is a $y_0 \in (\Diamond_T A) \setminus N$ such that $\|x_0 - y_0\|_K = d' < d$. As Tf_A is a K-contraction on $\mathbb{R}^n \setminus N$, $Tf_A(x_0) \geq Tf_A(y_0) - d' > Tf_A(y_0) - d$. Using (3.8), we obtain

$$dT1_{A+d\,\mathrm{int}\,K}(x_0) = T(d1_{A+d\,\mathrm{int}\,K})(x_0) \geq Tf_A(x_0)$$
$$> Tf_A(y_0) - d \geq T(d1_A)(y_0) - d = dT1_A(y_0) - d = d - d = 0,$$

since $T1_A(y_0) = 1$ due to $y_0 \in \Diamond_T A$ and (3.6). Thus $T1_{A+d\,\mathrm{int}\,K}(x_0) > 0$ and then $x_0 \in \Diamond_T(A+d\,\mathrm{int}\,K)$, by (4.15). This shows that $(\Diamond_T^* A) + d\,\mathrm{int}\,K \subset \Diamond_T(A+d\,\mathrm{int}\,K) = \Diamond_T^*(A+d\,\mathrm{int}\,K)$, essentially. Therefore (4.4) holds and T is K-smoothing by Lemma 4.4. $\square$

Lemma 4.9. *Let $d > 0$, let $f \in \mathcal{M}(\mathbb{R}^n)$, let $K \in \mathcal{K}_{(o)}^n$, and let $t \in \mathbb{R}$. Then*

$$\{x : f(x) \geq t + \omega_{K,d}(f)\}^* + d\,\mathrm{int}\,K \subset \{x : f(x) \geq t\}^*.$$

Proof. If N is the complement of the set of points of approximate continuity of f, then $\mathcal{H}^n(N) = 0$ and, by Lemma 4.5,

$$\omega_{K,d}(f) = \operatorname*{ess\,sup}_{\|x-y\|_K \leq d} |f(x) - f(y)| = \sup_{\|x-y\|_K \leq d;\ x,y \notin N} |f(x) - f(y)|$$

for any $d > 0$. Now fix $d > 0$ and let $y \in \{x : f(x) \geq t + \omega_{K,d}(f)\}^* + d \operatorname{int} K$. By (2.5), we can choose $z \in \{x : f(x) \geq t + \omega_{K,d}(f)\}^*$ and $\varepsilon = \varepsilon(y) > 0$ such that $\|y - z\|_K + 2\varepsilon < d$. Let $y' \in (\varepsilon K + y) \setminus N$ and $z' \in (\{x : f(x) \geq t + \omega_{K,d}(f)\} \cap (-\varepsilon K + z)) \setminus N$. Then $\|y' - y\|_K \leq \varepsilon$ and $\|z - z'\|_K \leq \varepsilon$, so $\|y' - z'\|_K \leq \|y - z\|_K + 2\varepsilon < d$. Since $y', z' \notin N$, we have $|f(y') - f(z')| \leq \omega_{K,d}(f)$. Hence

$$f(y') \geq f(z') - \omega_{K,d}(f) \geq t.$$

This implies that $(\varepsilon K + y) \subset \{x : f(x) \geq t\}$, essentially. Thus the latter set essentially contains an open neighborhood of y, and the desired conclusion follows easily. $\quad\square$

Lemma 4.10. *If $\alpha > 0$, $d > 0$, $g \in \mathcal{M}(\mathbb{R}^n)$, and an o-symmetric $K \in \mathcal{K}_{(o)}^n$ are such that*

$$\{x : g(x) \geq t + \alpha\}^* + d \operatorname{int} K \subset \{x : g(x) \geq t\}^*, \tag{4.16}$$

essentially, for $t > \operatorname{ess\,inf} g$, then $\omega_{K,d}(g) \leq \alpha$.

Proof. Note firstly that the inclusion in (4.16) actually holds pointwise, a fact we shall use later in the proof. This is because the set $G = \{x : g(x) \geq t + \alpha\}^* + d \operatorname{int} K$ is open, so $G^* = G$, and the pointwise inclusion then follows from parts (i) and (iii) of Lemma 2.1.

Suppose that $\omega_{K,d}(g) > \alpha$. Let $E_d = \{(y, z) \in \mathbb{R}^n \times \mathbb{R}^n : \|y - z\|_K \leq d\}$, let

$$F = \{(y, z) \in \mathbb{R}^n \times \mathbb{R}^n : g(y), \ g(z) \geq \operatorname{ess\,inf} g \ \text{ and } \ |g(y) - g(z)| > \alpha\},$$

and let $A = E_d \cap F$. Then $\omega_{K,d}(g) > \alpha$ implies that $\mathcal{H}^{2n}(A) > 0$. The o-symmetry of K yields $\|z - y\|_K = \|y - z\|_K$, so $(z, y) \in A$ if and only if $(y, z) \in A$. For $k \in \mathbb{N}$ and $q \in \mathbb{Q}$ with $q > \operatorname{ess\,inf} g$, let

$$V_{k,q} = E_{d-1/k} \cap \{(y, z) \in \mathbb{R}^n \times \mathbb{R}^n : g(y) \geq q + \alpha, \ \text{and} \ \operatorname{ess\,inf} g \leq g(z) < q\},$$

let $W_{k,q} = V_{k,q} \cup \{(z, y) : (y, z) \in V_{k,q}\}$, and let $W = \cup\{W_{k,q} : k \in \mathbb{N}, \ q \in \mathbb{Q}\}$. Clearly $W \subset A$. Let $Z = \{(y, z) \in \mathbb{R}^n \times \mathbb{R}^n : \|y - z\|_K = d\}$. By Fubini's theorem, $\mathcal{H}^{2n}(Z) = 0$. For each $(y, z) \in A \setminus Z$, we can find $k \in \mathbb{N}$ such that $\|y - z\|_K \leq d - 1/k$ and $|g(y) - g(z)| > \alpha$, and hence also a $q \in \mathbb{Q}$ such that $q > \operatorname{ess\,inf} g$ and

$$\operatorname{ess\,inf} g \leq g(z) < q < q + \alpha \leq g(y)$$

(or the same with y and z interchanged). This means that $A \setminus Z \subset W$ and therefore $\mathcal{H}^{2n}(W) = \mathcal{H}^{2n}(A) > 0$. It follows that there are $k_0 \in \mathbb{N}$ and $q_0 \in \mathbb{Q}$ with $q_0 > \operatorname{ess\,inf} g$ such that $\mathcal{H}^{2n}(W_{k_0,q_0}) > 0$ and hence, without loss of generality, $\mathcal{H}^{2n}(V_{k_0,q_0}) > 0$. By the Lebesgue density theorem, there exists $(y_0, z_0) \in V_{k_0,q_0}$ such that $y_0 \in \{x : g(x) \geq q_0 + \alpha\}^*$ and $z_0 \in \{x : g(x) < q_0\}^*$. Since $\|z_0 - y_0\|_K = \|y_0 - z_0\|_K < d$, we have $z_0 \in \operatorname{int} dK + y_0$. The pointwise inclusion in (4.16) with $t = q_0$ then implies that $z_0 \in \{x : g(x) \geq q_0\}^*$, a contradiction that completes the proof. $\quad\square$

Theorem 4.11. *Let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$, let $K \in \mathcal{K}_{(o)}^n$, and let $T : X \to X$ be a K-smoothing rearrangement.*
(i) *If K is o-symmetric, then T reduces the K-modulus of continuity of each $f \in X$.*
(ii) *If $d > 0$, $f \in X$, and $rB^n \subset K \subset RB^n$ for $0 < r \leq R$, then $\omega_{dr}(Tf) \leq \omega_{dR}(f)$.*

Proof. (i) Let $d > 0$, let $f \in X$, and let K be o-symmetric. If $\omega_{K,d}(f) = 0$, then clearly f is essentially constant, and hence, by Proposition 3.2(iii), Tf is also essentially constant. Then $\omega_{K,d}(Tf) = 0 = \omega_{K,d}(f)$. Therefore we may assume that $\omega_{K,d}(f) > 0$.

Let $t > \operatorname{ess\,inf} f$ and recall that $\operatorname{ess\,inf} Tf = \operatorname{ess\,inf} f$, by Proposition 3.2(i). Then, using (3.9), Lemma 3.9(iii), (4.5) with $A = \{x : f(x) \geq t + \omega_{K,d}(f)\}^*$, Lemma 3.9(i), Lemma 4.9, and Lemma 3.9(iii) and (3.9) again, we obtain

$$
\begin{aligned}
&\{x : Tf(x) \geq t + \omega_{K,d}(f)\}^* + d \operatorname{int} K \\
&= (\Diamond_T^* \{x : f(x) \geq t + \omega_{K,d}(f)\}) + d \operatorname{int} K \\
&= (\Diamond_T^* \{x : f(x) \geq t + \omega_{K,d}(f)\}^*) + d \operatorname{int} K \\
&\subset \cup \{\Diamond_T^* E : E \in \mathcal{L}^n, \ E \subset \{x : f(x) \geq t + \omega_{K,d}(f)\}^* + d \operatorname{int} K\} \\
&\subset \Diamond_T^* \{x : f(x) \geq t\}^* = \Diamond_T^* \{x : f(x) \geq t\} = \{x : Tf(x) \geq t\}^*,
\end{aligned}
$$

essentially. The conclusion $\omega_{K,d}(Tf) \leq \omega_{K,d}(f)$ follows from the o-symmetry of K and Lemma 4.10 with $g = Tf$ and $\alpha = \omega_{K,d}(f)$.

(ii) The proof is an easy modification of that of part (i). Let $d > 0$ and let $f \in X$. If $\omega_{dR}(f) = 0$, then as at the beginning of the proof of (i), we conclude that $\omega_{dr}(Tf) = 0 = \omega_{dR}(f)$. Therefore we may assume that $\omega_{dR}(f) > 0$. The inclusion $K \subset RB^n$ implies that $\omega_{K,d}(f) \leq \omega_{RB^n,d}(f) = \omega_{dR}(f)$, so Lemma 4.9 yields

$$\{x : f(x) \geq t + \omega_{dR}(f)\}^* + d \operatorname{int} K \subset \{x : f(x) \geq t\}^* \tag{4.17}$$

for $t \in \mathbb{R}$. Arguing as above with $A = \{x : f(x) \geq t + \omega_{dR}(f)\}^*$ and using (4.17) instead of Lemma 4.9, we obtain

$$\{x : Tf(x) \geq t + \omega_{dR}(f)\}^* + d \operatorname{int} K \subset \{x : Tf(x) \geq t\}^*,$$

essentially, as this does not require K to be o-symmetric. Since $rdD^n \subset d \operatorname{int} K$, the conclusion $\omega_{dr}(Tf) \leq \omega_{dR}(f)$ follows from Lemma 4.10 with $K = B^n$, $g = Tf$, $\alpha = \omega_{dR}(f)$, and d replaced by dr. $\quad\square$

Corollary 4.12. *Let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$, let $K \in \mathcal{K}_{(o)}^n$ be o-symmetric, and let $T : X \to X$ be a rearrangement. The following are equivalent.*
(i) *T reduces the K-modulus of continuity.*
(ii) *T reduces the K-modulus of continuity of each K-contraction in X.*
(iii) *T is K-smoothing.*

Proof. The implication (i)⇒(ii) is obvious on noting that each K-contraction is continuous, while (ii)⇒(iii) and (iii)⇒(i) follow from Theorems 4.8 and 4.11(i), respectively. □

The following example shows that the o-symmetry assumption on K in Theorem 4.11(i) and Corollary 4.12 (the implications (iii)⇒ (i) and (iii)⇒ (ii)) cannot be omitted.

Example 4.13. Let $K \in \mathcal{K}^n_{(o)}$, let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$, and let $T : X \to X$ be the K-Schwarz rearrangement defined in Example 4.3(ii). If $f \in X$, the superlevel sets of Tf are dilates of K. Since T is K-smoothing, it follows from Theorem 4.11 that when K is o-symmetric, T reduces the K-modulus of continuity of functions in X. However, this is not generally the case if K is not o-symmetric. To see this, let $o \neq x_0 \in D^n = \text{int } B^n$, let $K = B^n + x_0$, and let f be defined by $f(x) = 1 - \|x\|$ if $x \in B^n$ and $f(x) = 0$ otherwise. Define $v = x_0/\|x_0\|$ and

$$M = \max_{u \in S^{n-1}} \rho_K(u) = \rho_K(v) = 1 + \|x_0\|.$$

Then $\omega_{K,d}(f) = dM$ for small $d > 0$. (For $\omega_{K,d}(f) \geq dM$, it suffices to take $y = o$ and $x = dMv$; then $\|x - y\|_K = d$, because this is equivalent to $x \in y + d\partial K$, and $|f(x) - f(y)| = \|x\| = dM$ for small $d > 0$. The reverse inequality comes from the observation that f is Lipschitz with Lipschitz constant 1, and from $\|x - y\| \leq M\|x - y\|_K$.)

Now we claim that $\omega_{K,d}(Tf) > dM$. It suffices to prove $|Tf(o) - Tf(-dMv)| > dM$, since $\|o - (-dMv)\|_K = d$. For $t \in (0, 1)$, the statement

$$-dMv \in (1 - t)(B^n + x_0) = \Diamond_T\{z : f(x) \geq t\} = \Diamond_T^*\{z : f(x) \geq t\}$$

is true if and only if $t \leq 1 - dM/(1 - \|x_0\|)$. By (3.13), this implies that $Tf(-dMv) = 1 - dM/(1 - \|x_0\|)$. With similar but simpler arguments we argue that $Tf(o) = 1$. These two facts yield $|Tf(o) - Tf(-dMv)| = dM/(1 - \|x_0\|) > dM$, as required. □

When $X = \mathcal{M}(\mathbb{R}^n)$ or $\mathcal{M}_+(\mathbb{R}^n)$, the implication (ii)⇒(iii) in Corollary 4.12 remains true, by Theorem 4.8, but the following example, a modification of [6, Example 4.4], shows that (iii)⇒(ii) does not hold generally.

Example 4.14. If $K \in \mathcal{K}^n_{(o)}$ and $X = \mathcal{M}(\mathbb{R}^n)$ or $\mathcal{M}_+(\mathbb{R}^n)$, there are K-smoothing rearrangements $T : X \to X$ that do not reduce the K-modulus of continuity. To see this, call $f \in X$ *of type I* if $\mathcal{H}^n(\{x : f(x) > t\}) = \infty$ for $t \geq \text{ess inf } f$ and *of type II* otherwise, i.e., if there is a $t_0 \geq \text{ess inf } f$ such that $\mathcal{H}^n(\{x : f(x) > t\}) < \infty$ for $t > t_0$. Then define

$$Tf = \begin{cases} f + 1_{B^n}, & \text{if } f \text{ is of type I,} \\ f, & \text{if } f \text{ is of type II.} \end{cases}$$

Clearly, $T : X \to X$ is equimeasurable. If $f \le g$, then either f and g are of the same type, or f is of type II and g is of type I. It follows that $Tf \le Tg$ and hence that T is a rearrangement. The associated mapping $\Diamond_T$ is the identity on $\mathcal{L}^n$, so T is K-smoothing. The function $f_0(x) = e^{x_1}$ is continuous but its image Tf_0 is not. Hence, T does not reduce the K-modulus of continuity of continuous functions. $\quad\square$

5. The Pólya-Szegő inequality for Lipschitz functions

Recall that the subgraph $K_f \subset \mathbb{R}^{n+1}$ of a function $f \in \mathcal{M}(\mathbb{R}^n)$ is defined by (2.9).

Lemma 5.1. *If $f \in \mathcal{M}(\mathbb{R}^n)$ and $s \in \mathbb{R}$, then*

$$K_f^* \cap \{x_{n+1} = s\} \subset (K_f \cap \{x_{n+1} = s\})^*, \tag{5.1}$$

where the set of Lebesgue density points on the right is formed with respect to the hyperplane $\{x_{n+1} = s\} = \mathbb{R}^n + se_{n+1}$, identified with $\mathbb{R}^n$.

Proof. Let $x \in \{x_{n+1} = s\}$. If $x \notin (K_f \cap \{x_{n+1} = s\})^*$, there exists $0 < a < 1$ such that if $r_0 > 0$, there is an $0 < r < r_0$ such that

$$\mathcal{H}^n \left(K_f \cap \{x_{n+1} = s\} \cap B(x, r)\right) < (1 - a)\kappa_n r^n, \tag{5.2}$$

where $B(x, r)$ is the $(n+1)$-dimensional ball with center x and radius r. If $y \in e_{n+1}^\perp$ and $y + se_{n+1} \notin K_f \cap \{x_{n+1} = s\} \cap B(x, r)$, then by the definition (2.9) of K_f, $y + te_{n+1} \notin K_f$ for each $t > s$. Therefore, by (5.2) and Fubini's theorem, $\mathcal{H}^{n+1}(K_f \cap B(x, r))$ is largest when

$$K_f \cap \{x_{n+1} = s\} \cap B(x, r) \subset \{x_{n+1} = s\} \cap B\left(x, (1 - a)^{1/n}r\right),$$

in which case

$$K_f \cap B(x, r) \subset \left(\left(\left((1 - a)^{1/n}rB^n\right) \times [0, s)\right) \cup (rB^n \times (-\infty, s])\right) \cap B(x, r). \tag{5.3}$$

Let

$$E(a, n, r) = rB^{n+1} \setminus \left(\left((1 - a)^{1/n}rB^n\right) \times \mathbb{R}\right)$$

be the region in $\mathbb{R}^{n+1}$ between the sphere with center o and radius r and an infinite o-symmetric cylinder with radius $(1 - a)^{1/n}r$. From (5.3), we see that

$$\mathcal{H}^{n+1}(B(x, r) \setminus K_f) \ge \mathcal{H}^{n+1}(E(a, n, r))/2 \ge c(a, n)r^{n+1},$$

where $c(a, n) = \mathcal{H}^{n+1}(E(a, n, 1))/2 > 0$. It follows that $x \notin K_f^*$. $\quad\square$

Let $X = \mathcal{M}(\mathbb{R}^n)$, $\mathcal{M}_+(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$, or $\mathcal{V}(\mathbb{R}^n)$, let $T : X \to X$ be a rearrangement, and let E be a subset of the hyperplane $\{x_{n+1} = t\} = \mathbb{R}^n + te_{n+1}$ in $\mathbb{R}^{n+1}$, such that $E \,|\, \mathbb{R}^n \in \mathcal{L}^n$. Slightly abusing notation, we shall define

$$\Diamond_T E = (\Diamond_T(E \,|\, \mathbb{R}^n)) + te_{n+1}, \tag{5.4}$$

thereby extending the action of $\Diamond_T$ to horizontal hyperplanes in $\mathbb{R}^{n+1}$. The action of $\Diamond_T^*$ can be extended in a similar fashion. Note that by (3.9), (2.9), and (5.4), for $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$ we have

$$\begin{aligned}
K_{Tf} \cap \{x_{n+1} = t\} &= \{x \in \mathbb{R}^n : Tf(x) \geq t\} + te_{n+1} \\
&= (\Diamond_T\{x \in \mathbb{R}^n : f(x) \geq t\}) + te_{n+1} \\
&= \Diamond_T(K_f \cap \{x_{n+1} = t\}),
\end{aligned}$$

essentially, for $t > \operatorname{ess\,inf} f$. By Lemma 2.1(ii), this yields the pointwise identity

$$(K_{Tf} \cap \{x_{n+1} = t\})^* = \Diamond_T^*(K_f \cap \{x_{n+1} = t\}) \tag{5.5}$$

for $t > \operatorname{ess\,inf} f$, where here and below, sets of Lebesgue density points are taken with respect to the appropriate horizontal hyperplane identified with $\mathbb{R}^n$.

The following lemma is stated in a general form required for Section 7. The reader interested only in the results of this section may focus on the special case corresponding to $K = B^n$, $\operatorname{int} K = D^n$, when $C \subset \mathbb{R}^{n+1}$ is an o-symmetric convex body of revolution about the x_{n+1}-axis.

Lemma 5.2. *Let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$, let $T : X \to X$ be a rearrangement, and let $d > 0$. Let $K \in \mathcal{K}^n_{(o)}$ and let $C \subset \mathbb{R}^{n+1}$ be a convex body supported by the hyperplanes $\{x_{n+1} = \pm 1\}$ and all of whose sections $C \cap \{x_{n+1} = t\}$, $t \in [-1, 1]$, are dilates of K. If T is K-smoothing, $a > d + \operatorname{ess\,inf} f$, and $f \in X$ is such that $\{x : f(x) \geq a\}$ is bounded, then*

$$\begin{aligned}
&((K_{Tf} \cap \{x_{n+1} \geq a\})^* + d \operatorname{int} C) \cap \{x_{n+1} = t\} \\
&\subset \Diamond_T^* (((K_f \cap \{x_{n+1} \geq a\}) + d \operatorname{int} C) \cap \{x_{n+1} = t\}) \tag{5.6}
\end{aligned}$$

for $t > \operatorname{ess\,inf} f$.

Proof. Let $d > 0$ and let $C = \{(x, x_{n+1}) \in \mathbb{R}^n \times \mathbb{R} : x \in g(x_{n+1})K, |x_{n+1}| \leq 1\}$, for a suitable concave function g defined on $[-1, 1]$. For $t \in \mathbb{R}$, denote by Π_t the orthogonal projection onto $\{x_{n+1} = t\}$. If L is any set in $\mathbb{R}^{n+1}$, then

$$(L + d \operatorname{int} C) \cap \{x_{n+1} = t\} = \bigcup_{t-d < s < t+d} \Pi_t \left((L \cap \{x_{n+1} = s\}) + r_s \operatorname{int} K\right), \tag{5.7}$$

where $r_s = d\,g((t-s)/d)$. Indeed, $p \in (L + d\operatorname{int} C) \cap \{x_{n+1} = t\}$ if and only if $p \,|\, \langle e_{n+1} \rangle = t e_{n+1}$ and there is a $z \in L$ such that $p \in z + d\operatorname{int} C$. If $z \,|\, \langle e_{n+1} \rangle = s e_{n+1}$, then this holds if and only if $t - d < s < t + d$ and

$$p - \Pi_t z \in \ d\,g\left(\frac{t-s}{d}\right)\operatorname{int} K,$$

that is, $p \in \Pi_t(z + r_s \operatorname{int} K)$.

Applying (5.7) with L replaced by $L \cap \{x_{n+1} \geq a\}$, we obtain

$$((L \cap \{x_{n+1} \geq a\}) + d\operatorname{int} C) \cap \{x_{n+1} = t\}$$
$$= \bigcup_{t-d<s<t+d} \Pi_t\left((L \cap \{x_{n+1} \geq a\} \cap \{x_{n+1} = s\}) + r_s \operatorname{int} K\right)$$
$$= \bigcup_{t-d<s<t+d,\ s\geq a} \Pi_t\left((L \cap \{x_{n+1} = s\}) + r_s \operatorname{int} K\right). \tag{5.8}$$

Let $f \in X$ satisfy the hypotheses of the lemma. By Lemma 2.1(i), we have

$$(K_{Tf} \cap \{x_{n+1} \geq a\})^* \subset K_{Tf}^* \cap \{x_{n+1} \geq a\}^* \subset K_{Tf}^* \cap \{x_{n+1} \geq a\}.$$

From this and (5.1) with f replaced by Tf, we obtain

$$(K_{Tf} \cap \{x_{n+1} \geq a\})^* \cap \{x_{n+1} = s\} \subset K_{Tf}^* \cap \{x_{n+1} = s\} \subset (K_{Tf} \cap \{x_{n+1} = s\})^*, \tag{5.9}$$

whenever $s \geq a$, while the set on the left is clearly empty if $s < a$. We use (5.7) with $L = (K_{Tf} \cap \{x_{n+1} \geq a\})^*$, (5.9), (5.5), (4.4) applied (via (5.4)) with A replaced by the bounded set $K_f \cap \{x_{n+1} = s\}$, $s \geq a$, the fact that the action of $\Diamond_T^*$ as extended by (5.4) is the same for each t, the pointwise monotonicity of $\Diamond_T^*$ provided by Lemma 3.9(i), and (5.8) with $L = K_f$, to obtain

$$((K_{Tf} \cap \{x_{n+1} \geq a\})^* + d\operatorname{int} C) \cap \{x_{n+1} = t\}$$
$$= \bigcup_{t-d<s<t+d} \Pi_t\left([(K_{Tf} \cap \{x_{n+1} \geq a\})^* \cap \{x_{n+1} = s\}] + r_s \operatorname{int} K\right)$$
$$\subset \bigcup_{t-d<s<t+d,\ s\geq a} \Pi_t\left([K_{Tf} \cap \{x_{n+1} = s\}]^* + r_s \operatorname{int} K\right)$$
$$= \bigcup_{t-d<s<t+d,\ s\geq a} \Pi_t\left([\Diamond_T^*(K_f \cap \{x_{n+1} = s\})] + r_s \operatorname{int} K\right)$$
$$\subset \bigcup_{t-d<s<t+d,\ s\geq a} \Pi_t\left(\Diamond_T^*[(K_f \cap \{x_{n+1} = s\}) + r_s \operatorname{int} K]\right)$$
$$= \bigcup_{t-d<s<t+d,\ s\geq a} \Diamond_T^*\left(\Pi_t[(K_f \cap \{x_{n+1} = s\}) + r_s \operatorname{int} K]\right)$$

$$\subset \Diamond_T^* \Big(\bigcup_{t-d<s<t+d,\ s\geq a} \Pi_t\left[(K_f \cap \{x_{n+1}=s\}) + r_s \operatorname{int} K\right] \Big)$$
$$= \Diamond_T^* \left(((K_f \cap \{x_{n+1} \geq a\}) + d \operatorname{int} C) \cap \{x_{n+1}=t\}\right). \quad \square$$

Recall that $\overline{\mathcal{M}}_C^*(A)$ is the upper anisotropic outer Minkowski content of $A \in \mathcal{M}(\mathbb{R}^n)$ with respect to a convex body $C \in \mathcal{K}_{(o)}^n$, obtained via the left-hand limit in (2.8) with B^n replaced by C. We will apply this notion in $\mathbb{R}^{n+1}$.

Lemma 5.3. *Let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$, let $T : X \to X$ be a rearrangement, and let C be as in Lemma 5.2. If T is smoothing, $a > \operatorname{ess\,inf} f$, and $f \in X$ is such that $\{x : f(x) \geq a\}$ is bounded, then*

$$\overline{\mathcal{M}}_C^*\left((K_{Tf} \cap \{x_{n+1} \geq a\})^*\right) \leq \overline{\mathcal{M}}_C^*(K_f \cap \{x_{n+1} \geq a\}).$$

Proof. Taking the $\mathcal{H}^n$-measures of both sides of (5.6), integrating with respect to t, and using Fubini's theorem and the fact that $\Diamond$ is measure preserving, we obtain

$$\mathcal{H}^{n+1}\left((K_{Tf} \cap \{x_{n+1} \geq a\})^* + d \operatorname{int} C\right) \leq \mathcal{H}^{n+1}\left((K_f \cap \{x_{n+1} \geq a\}) + d \operatorname{int} C\right) \quad (5.10)$$

for $0 < d \leq a - \operatorname{ess\,inf} f$. By the equimeasurability of T,

$$\mathcal{H}^{n+1}\left((K_{Tf} \cap \{x_{n+1} \geq a\})^*\right) = \mathcal{H}^{n+1}(K_{Tf} \cap \{x_{n+1} \geq a\}) = \mathcal{H}^{n+1}(K_f \cap \{x_{n+1} \geq a\}).$$

The desired inequality now follows directly from the definition of $\overline{\mathcal{M}}_C^*$ (in $\mathbb{R}^{n+1}$ and with C replaced by $\operatorname{int} C$). $\quad \square$

Lemma 5.4. *Let $f \in \mathcal{S}(\mathbb{R}^n)$ be Lipschitz. If $a > \operatorname{ess\,inf} f$, then $\{x : f(x) \geq a\}$ is bounded.*

Proof. Let $\varepsilon > 0$ be such that $a - \varepsilon > \operatorname{ess\,inf} f$ and let L be the Lipschitz constant of f.

Suppose that $\{x : f(x) \geq a\}$ is unbounded. Then there are points x_k in this set with $\|x_{k+1}\| > \|x_k\| + 2\varepsilon/(1+L)$ for $k \in \mathbb{N}$. The Lipschitz property implies that $f(x) \geq a - \varepsilon$ whenever $x \in B(x_k, \varepsilon/(1+L))$, $k \in \mathbb{N}$. As these balls are disjoint, $\mathcal{H}^n(\{x : f(x) \geq a - \varepsilon\}) = \infty$, contradicting $f \in \mathcal{S}(\mathbb{R}^n)$. $\quad \square$

Recall that h_C is the support function of C and G_f denotes the graph of $f \in \mathcal{M}(\mathbb{R}^n)$. A result in the spirit of the following lemma was proved by Zhang [48, Lemma 3.1].

Lemma 5.5. *Let $f \in \mathcal{S}(\mathbb{R}^n)$ be Lipschitz and let C be as in Lemma 5.2. Let $a > \operatorname{ess\,inf} f$ be such that $\mathcal{H}^n(\{x : f(x) = a\}) = 0$. Then*

$$\overline{\mathcal{M}}_C\left((K_f \cap \{x_{n+1} \geq a\})^*\right) = \overline{\mathcal{M}}_C\left(K_f \cap \{x_{n+1} \geq a\}\right)$$
$$= \int_{G_f \cap \{x_{n+1}>a\}} h_C(\nu(x))\, d\mathcal{H}^n(x) + \mathcal{H}^n(\{x : f(x) \geq a\}), \quad (5.11)$$

where $\nu(x)$ denotes the outer unit normal to K_f at x.

Proof. Recall that if $E \subset \mathbb{R}^{n+1}$ is $\mathcal{H}^{n+1}$-measurable, its density $\Theta(E, x)$ at x is defined by (2.11) with n replaced by $n + 1$. For $t \in [0, 1]$, define

$$E^t = \{x \in \mathbb{R}^{n+1} : \Theta(E, x) = t\}.$$

Let $\partial^e E = \mathbb{R}^{n+1} \setminus (E^0 \cup E^1)$ denote the essential boundary of E. If E has locally finite perimeter, then by Federer's theorem, $\nu(x)$ exists for $\mathcal{H}^n$-almost all $x \in \partial^e E$; see, for example, [28, Theorem 16.2].

Lussardi and Villa [27, Remark 4.2, Theorem 4.4, and Remark 4.5] prove the following result. If $E \subset \mathbb{R}^{n+1}$ is a Borel set whose boundary is countably $\mathcal{H}^n$-rectifiable and bounded, $\mathcal{H}^n(\partial E \cap E^0) = 0$, and E has the property that there exist $\gamma > 0$ and a probability measure μ in $\mathbb{R}^{n+1}$ absolutely continuous with respect to $\mathcal{H}^n$, such that for each $x \in \partial E$ and $r \in (0, 1)$,

$$\mu(B(x, r)) \geq \gamma r^n, \tag{5.12}$$

then E has finite perimeter, the anisotropic outer Minkowski content of E with respect to C is defined, and

$$\overline{\mathcal{M}}_C(E) = \int_{\partial^e E} h_C(\nu(x)) \, d\mathcal{H}^n(x). \tag{5.13}$$

Recall that a set $E \subset \mathbb{R}^{n+1}$ is *countably $\mathcal{H}^n$-rectifiable* if there exist countably many Lipschitz maps $g_i : \mathbb{R}^n \to \mathbb{R}^{n+1}$ such that $\mathcal{H}^n(E \setminus \cup_i g_i(\mathbb{R}^n)) = 0$.

Let $K_{f,a} = K_f \cap \{x_{n+1} \geq a\}$. We have

$$\partial K_{f,a} = (G_f \cap \{x_{n+1} > a\}) \cup (K_f \cap \{x_{n+1} = a\}). \tag{5.14}$$

If $x \in \mathbb{R}^{n+1}$, write $x' = (x_1, \dots, x_n)$. It is clear that if $x \in K_f \cap \{x_{n+1} = a\}$ and $f(x') > a$ (or $f(x') = a$), then $\Theta(K_{f,a}, x) = 1/2$ (or $\Theta(K_{f,a}, x) \leq 1/2$, respectively). We claim that if $x \in G_f \cap \{x_{n+1} > a\}$ and $\Theta(K_{f,a}, x)$ exists, then $\Theta(K_{f,a}, x) \in (0, 1)$. Indeed, let $x \in G_f \cap \{x_{n+1} > a\}$. If L denotes the Lipschitz constant of f and $\varepsilon > 0$ is sufficiently small, then

$$\{(y', y_{n+1}) : y_{n+1} \leq x_{n+1} - L\|y' - x'\|\} \cap B(x, \varepsilon) \subset K_{f,a},$$

from which it is easy to see that $\Theta(K_{f,a}, x) > 0$. Similarly, $\Theta(K_{f,a}, x) < 1$ follows easily from

$$\{(y', y_{n+1}) : y_{n+1} > x_{n+1} + L\|y' - x'\|\} \cap K_{f,a} = \emptyset.$$

This proves the claim.

Since $\mathcal{H}^n(\{x : f(x) = a\}) = 0$, by assumption, the observations in the previous paragraph imply that

$$\mathcal{H}^n\left(\partial K_{f,a} \cap (K_{f,a})^0\right) = 0 \quad \text{and} \quad \partial^e K_{f,a} = \partial K_{f,a}, \tag{5.15}$$

up to a set of $\mathcal{H}^n$-measure zero. The assumption that $\mathcal{H}^n(\{x : f(x) = a\}) = 0$ also implies that $K_{f,a}^* = \text{int } K_{f,a}$ and $\partial\left(K_{f,a}^*\right) = \partial K_{f,a}$, up to a set of $\mathcal{H}^n$-measure zero. Arguments similar to those used for $K_{f,a}$ prove that

$$\mathcal{H}^n\left(\partial\left(K_{f,a}^*\right) \cap (K_{f,a}^*)^0\right) = 0 \quad \text{and} \quad \partial^e\left(K_{f,a}^*\right) = \partial K_{f,a}, \tag{5.16}$$

up to a set of $\mathcal{H}^n$-measure zero.

We claim that $K_{f,a}$ and $K_{f,a}^*$ satisfy the hypotheses of Lussardi and Villa's result. Towards this goal, note firstly that by Lemma 5.4, $K_{f,a}$ and $K_{f,a}^*$ are bounded, and by definition, their boundaries are countably $\mathcal{H}^n$-rectifiable. Let

$$D = \{x' \in \mathbb{R}^n : f(x') \geq a\} + B^n \subset \mathbb{R}^n,$$

let $A_1 = \{(x', f(x')) : x' \in D\}$, and let $A_2 = \{(x', a) : x' \in D\}$. For $E \subset \mathbb{R}^{n+1}$, define

$$\mu(E) = \left(\mathcal{H}^n\left(\pi(E \cap A_1)\right) + \mathcal{H}^n(E \cap A_2)\right)/c,$$

where $c = \mathcal{H}^n(\pi(A_1)) + \mathcal{H}^n(A_2)$ and $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is defined by $\pi((x', x_{n+1})) = x'$. It is clear that μ is a probability measure in $\mathbb{R}^n$ and that since π is a contraction, μ is absolutely continuous with respect to $\mathcal{H}^n$.

It now suffices to prove that μ satisfies (5.12) for each $x \in \partial K_{f,a}$. To this end, let $y = (y', y_{n+1}) \in \partial K_{f,a}$ and let $r \in (0,1)$. If $y_{n+1} = a$ then

$$\mu(B(y,r)) \geq \mathcal{H}^n\left(B(y,r) \cap A_2\right)/c = \mathcal{H}^n\left(B(y,r) \cap \{x_{n+1} = a\}\right)/c = \kappa_n r^n/c.$$

If $y_{n+1} = f(y')$, then $\pi\left(B(y,r) \cap A_1\right)$ contains $\{x \in \mathbb{R}^n : \|x - y'\| \leq r/(1+L)\}$. Therefore

$$\mu(B(y,r)) \geq \mathcal{H}^n\left(\pi\left(B(y,r) \cap A_1\right)\right)/c \geq \mathcal{H}^n\left(\{x \in \mathbb{R}^n : \|x - y'\| \leq r/(1+L)\}\right)/c$$
$$= \frac{\kappa_n r^n}{c(1+L)^n}.$$

Thus μ satisfies (5.12) with $\gamma = \kappa_n/(c(1+L)^n)$.

Once we observe that

$$\int\limits_{K_f \cap \{x_{n+1}=a\}} h_C(\nu(x)) \, d\mathcal{H}^n(x) = h_C(-e_{n+1})\mathcal{H}^n(\{x : f(x) \geq a\}) = \mathcal{H}^n(\{x : f(x) \geq a\}),$$

where we used the fact that $h_C(-e_{n+1}) = 1$ for our choice of C, formula (5.13) with $E = K_{f,a}$ and $K_{f,a}^*$, (5.14), (5.15), and (5.16) give (5.11). $\quad\square$

The following lemma is stated in a general form necessary for Section 7. The reader interested only in the main results of this section may choose to focus on the special case when $K = B^n$, in which case $h_K(y) = \|y\|$ and $C \subset \mathbb{R}^{n+1}$ is a o-symmetric convex body of revolution about the x_{n+1}-axis.

Lemma 5.6. *Let $\Phi : [0, \infty) \to [0, \infty)$ be convex with $\Phi(0) = 0$ and $\Phi \not\equiv 0$, let $M > 0$, and let $K \in \mathcal{K}^n_{(o)}$. Then there exist $b > 0$ and a convex body $C \subset \mathbb{R}^{n+1}$, supported by the hyperplanes $\{x_{n+1} = \pm 1\}$ and all of whose sections $C \cap \{x_{n+1} = t\}$, $t \in [-1, 1]$, are dilates of K, such that*

$$h_C(y, 1) = 1 + b\,\Phi(h_K(y)), \tag{5.17}$$

for $y \in \mathbb{R}^n$ with $h_K(y) \leq M$. In particular, C satisfies the conditions in Lemma 5.2.

Proof. Define

$$\Psi(t) = \begin{cases} \Phi(t), & \text{if } 0 \leq t \leq M, \\ mt + q, & \text{if } t \geq M, \end{cases}$$

where $m > 0$ and $q \leq 0$ are such that $\Psi : [0, \infty) \to [0, \infty)$ is convex. Then, for $y \in \mathbb{R}^n$ and $t \in \mathbb{R}$, define

$$h(y, t) = \begin{cases} |t|\,(1 + b\,\Psi(h_K(y)/|t|)), & \text{if } t \neq 0, \\ b\,m\,h_K(y), & \text{if } t = 0, \end{cases} \tag{5.18}$$

$$= \begin{cases} |t|\,(1 + b\,\Phi(h_K(y)/|t|)), & \text{if } |t| \geq h_K(y)/M, \\ b\,m\,h_K(y) + (1 + b\,q)|t|, & \text{if } |t| \leq h_K(y)/M, \end{cases} \tag{5.19}$$

where $b > 0$. We show that b can be chosen so that $h = h_C$ is the support function of a convex body C. To this end, note that from (5.18), the positive homogeneity of h follows immediately and the subadditivity of h for $t > 0$ or for $t < 0$ is a routine exercise using the triangle inequality and the convexity of Ψ. It is then enough to observe that if b is small enough to ensure that $1 + b\,q > 0$, then the function $b\,m\,h_K(y) + (1 + b\,q)|t|$ in (5.19) coincides with the support function of the cylinder $b\,m\,K \times [-(1 + b\,q), 1 + b\,q] \subset \mathbb{R}^n \times \mathbb{R}$. This proves (5.17).

It remains to prove that all sections $C \cap \{x_{n+1} = t\}$, $t \in [-1, 1]$, are dilates of K. Clearly $o \in \text{int}\,C$, since $h_C > 0$. From (2.3) and (5.18), it follows easily that the sublevel sets of $h_C(y, 1)$, considered as a function of y, are dilates of the polar body K° of K. If $t \in (0, 1]$, and ρ_{C° denotes the radial function of the polar body C° of C (cf. (2.4)), then

$$C^\circ \cap \{x_{n+1} = t\} = \{(z, t) : \rho_{C^\circ}(z, t) \geq 1\} = \{t(z/t, 1) : \rho_{C^\circ}(z/t, 1) \geq t\}$$

$$= \{t(y, 1) : \rho_{C^\circ}(y, 1) \geq t\} = t\{(y, 1) : h_C(y, 1) \leq 1/t\}$$

is also a dilate of K°. This argument can be repeated for the sections of C° corresponding to $t \in [-1, 0)$ and, by continuity, for $C^\circ \cap \{x_{n+1} = 0\}$. Thus there exists a concave function f on $[-1, 1]$ such that

$$C^\circ = \{(f(t)z, t) \in \mathbb{R}^n \times \mathbb{R} : t \in [-1, 1], z \in K^\circ\}.$$

Let $s \in [-1, 1]$. A point $(y, s) \in \mathbb{R}^n \times \mathbb{R}$ belongs to C if and only if

$$f(t)(y \cdot z) + s\,t \le 1 \quad \forall z \in K^\circ, \quad \forall t \in [-1, 1],$$

that is, if and only if

$$y \cdot z \le \min_{t \in -[1,1]} \frac{1 - s\,t}{f(t)} \quad \forall z \in K^\circ.$$

The last formula shows that $C \cap \{x_{n+1} = s\}$ equals K dilated by the factor $\min_{t \in -[1,1]}(1 - s\,t)/f(t)$. $\quad \Box$

The special case $r = R$ of the following lemma will be needed for the main results of this section, while the general case is applied in Section 7.

Lemma 5.7. *Let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$, let $K \in \mathcal{K}^n$ satisfy $rB^n \subset K \subset RB^n$ for $0 < r \le R$, and let $T : X \to X$ be a K-smoothing rearrangement. If $f \in X$ is Lipschitz with Lipschitz constant L, then there is a Lipschitz function $F : \mathbb{R}^n \to \mathbb{R}$, with Lipschitz constant at most LR/r, such that $F(x) = Tf(x)$ for $\mathcal{H}^n$-almost all $x \in \mathbb{R}^n$.*

Proof. Let $f \in X$ be Lipschitz with Lipschitz constant L and let A be the set of points of approximate continuity of Tf. Since f is continuous, Lemma 4.5 and Theorem 4.11(ii) yield

$$\sup_{\|x-y\| \le dr;\ x,y \in A} |Tf(x) - Tf(y)| \le \sup_{\|x-y\| \le dR} |f(x) - f(y)|.$$

Let $x, y \in A$, $x \ne y$, and choose $d > 0$ so that $\|x - y\| = dr$. Then, by the previous inequality,

$$\begin{aligned}
|Tf(x) - Tf(y)| &\le \sup_{\|w-z\| \le dr;\ w,z \in A} |Tf(w) - Tf(z)| \\
&\le \sup_{\|w-z\| \le dR} |f(w) - f(z)| \le LdR = (LR/r)\|x - y\|.
\end{aligned}$$

Therefore Tf is Lipschitz on A with Lipschitz constant at most LR/r. By the McShane-Whitney extension theorem (see, e.g., [18, p. 202]) there is a function $F : \mathbb{R}^n \to \mathbb{R}$ with the same Lipschitz properties as Tf on the entire space $\mathbb{R}^n$, such that $F = Tf$ on A. Since $\mathcal{H}^n(\mathbb{R}^n \setminus A) = 0$, the proof is complete. $\quad \Box$

Recall that a Young function is a left-continuous and convex function $\Phi : [0, \infty) \to [0, \infty]$ with $\Phi(0) = 0$.

Theorem 5.8. *Let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$, let $T : X \to X$ be a rearrangement, and let Φ be a Young function. If T is smoothing and $f \in X$ is Lipschitz, then Tf coincides with a Lipschitz function $\mathcal{H}^n$-almost everywhere on $\mathbb{R}^n$, and*

$$\int\limits_{\{x:\ Tf(x) \geq a\}} \Phi\left(\|\nabla Tf(x)\|\right) dx \leq \int\limits_{\{x:\ f(x) \geq a\}} \Phi\left(\|\nabla f(x)\|\right) dx \tag{5.20}$$

for each $a > \operatorname{ess\,inf} f$. Hence

$$\int\limits_{\mathbb{R}^n} \Phi\left(\|\nabla Tf(x)\|\right) dx \leq \int\limits_{\mathbb{R}^n} \Phi\left(\|\nabla f(x)\|\right) dx, \tag{5.21}$$

where the integrals may be infinite.

Proof. Without loss of generality, we may assume that Φ is a nontrivial real-valued function. Indeed, the result is obvious if $\Phi \equiv 0$, and if Φ attains the value ∞, it does so on some maximal interval (t_0, ∞), $t_0 \geq 0$. Suppose first that $t_0 > 0$. For $0 < t < t_0$, the right derivative $\Phi'^+(t)$ of Φ at t is increasing, so we may define $c = \lim_{t \to t_0-} \Phi'^+(t)$. If $c = \infty$, let (t_k) be a strictly increasing sequence in $(0, t_0)$ converging to t_0, and define the real-valued Young functions

$$\Phi_k(t) = \begin{cases} \Phi(t), & \text{if } 0 \leq t < t_k, \\ \Phi'^+(t_k)(t - t_k) + \Phi(t_k), & \text{otherwise.} \end{cases}$$

If $c < \infty$, we must have $\Phi(t_0) < \infty$ by left continuity, and may define

$$\Phi_k(t) = \begin{cases} \Phi(t), & \text{if } 0 \leq t \leq t_0, \\ (c + k)(t - t_0) + \Phi(t_0), & \text{otherwise.} \end{cases}$$

When $t_0 = 0$, we put $\Phi_k(t) = kt$. In each case, we have $\Phi_k \leq \Phi$ and (Φ_k) is an increasing sequence of real-valued Young functions converging pointwise to Φ. Thus, if (5.20) holds for real-valued Young functions, it holds with Φ replaced by Φ_k, and hence, by the monotone convergence theorem, for Φ itself.

The set of values t such that $\mathcal{H}^n(\{x : f(x) = t\}) = 0$ is dense in $(\operatorname{ess\,inf} f, \infty)$, so there is an increasing sequence $\{a_m\}$ contained in $(\operatorname{ess\,inf} f, a)$ and converging to a such that $\mathcal{H}^n(\{x : f(x) = a_m\}) = 0$ for each m. Fix $m \in \mathbb{N}$. By Lemma 5.4, $\{x : f(x) \geq a_m\}$ is bounded, and the equimeasurability of T implies that

$$\mathcal{H}^n(\{x : Tf(x) = a_m\}) = \mathcal{H}^n(\{x : f(x) = a_m\}) = 0.$$

Assume that L is the Lipschitz constant for f. As T reduces the modulus of continuity by Corollary 4.12, Lemma 5.7 with $r = R$ implies that there is a Lipschitz function F on $\mathbb{R}^n$ with Lipschitz constant at most L such that $F(x) = Tf(x)$ for $\mathcal{H}^n$-almost all $x \in \mathbb{R}^n$. Weak derivatives and the remainder of this proof are unaffected by changing Tf on a set of measure zero, so we may assume that Tf itself is Lipschitz on $\mathbb{R}^n$ with Lipschitz constant at most L. Then, by Lemmas 5.3 and 5.5 (the latter applied to both f and Tf) with a replaced by a_m, and the equimeasurability of T, we obtain

$$\int_{G_{Tf} \cap \{x_{n+1} > a_m\}} h_C(\nu_{Tf}(x)) \, d\mathcal{H}^n(x) \le \int_{G_f \cap \{x_{n+1} > a_m\}} h_C(\nu_f(x)) \, d\mathcal{H}^n(x),$$

where C is any convex body as in Lemma 5.2, and $\nu_g(x)$ is the outer unit normal of K_g at x for a Lipschitz function g. The integral on the right can be written as

$$\int_{\{y : f(y) > a_m\}} h_C\left(\frac{(-\nabla f(y), 1)}{\sqrt{1 + \|\nabla f(y)\|^2}}\right) \sqrt{1 + \|\nabla f(y)\|^2} \, dy = \int_{\{y : f(y) > a_m\}} h_C(-\nabla f(y), 1) \, dy,$$

where we used the 1-homogeneity of h_C. Similarly, the integral on the left can be rewritten in the same form, with f replaced by Tf. Consequently,

$$\int_{\{y : Tf(y) > a_m\}} h_C(-\nabla Tf(y), 1) \, dy \le \int_{\{y : f(y) > a_m\}} h_C(-\nabla f(y), 1) \, dy,$$

which also yields

$$\int_{\{y : Tf(y) > a_m\}} h_C(-\nabla Tf(y), 1) - 1 \, dy \le \int_{\{y : f(y) > a_m\}} h_C(-\nabla f(y), 1) - 1 \, dy, \qquad (5.22)$$

since $\mathcal{H}^n(\{y : Tf(y) > a_m\}) = \mathcal{H}^n(\{y : f(y) > a_m\})$. If ℓ is the Lipschitz constant for some Lipschitz function g, we have $\|\nabla g(x)\| \le \ell$ for $x \in \mathbb{R}^n$ where the derivative of g exists. Applying this to $g = f$ and $g = Tf$ implies that

$$\max\left\{\|\nabla f(x)\|, \|\nabla Tf(x)\|\right\} \le L \qquad (5.23)$$

for $\mathcal{H}^n$-almost all $x \in \mathbb{R}^n$. As Φ is real-valued and not identically 0, we may define the convex body C as in Lemma 5.6 corresponding to $K = B^n$, $M = L$, and Φ. Then C satisfies the conditions stated in Lemma 5.2 and by (5.17),

$$h_C(y, 1) - 1 = b \, \Phi(\|y\|),$$

for $y \in \mathbb{R}^n$ with $\|y\| \le M$ and some $b > 0$. Substituting in (5.22) and taking limits as $m \to \infty$, we obtain (5.20).

By Proposition 3.2, we have $\operatorname{ess\,inf} Tf = \operatorname{ess\,inf} f$. Letting $a \to \operatorname{ess\,inf} f$ in (5.20), we arrive at (5.21). $\square$

Theorem 5.8 yields Pólya-Szegő inequalities with $\Phi(t) = t^p$, $1 \le p < \infty$, and it is possible to extend these to functions in $W^{1,p}(\mathbb{R}^n) \cap \mathcal{V}(\mathbb{R}^n)$ by means of standard techniques, utilized for example in the proof of (1.1) in [3, Theorem 3.20] (though the case $p = 1$ needs extra work). We shall not state this extension here, however, but instead derive it from a still more general result, Theorem 6.3, proved in Section 6. See Corollary 6.5. We also postpone to Section 6 a proof that the assumption in Theorem 5.8 that T is smoothing cannot be dropped in general; see Example 6.6. This example also shows that when $\Phi(t) = t$ in Theorem 5.8, the smoothing property is not necessary for the stated inequalities to hold.

Corollary 5.9. *Let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$, and let $T : X \to X$ be a rearrangement. If T is smoothing, $f \in X$ is Lipschitz, and $a > \operatorname{ess\,inf} f$, then*

$$\mathcal{H}^n(G_{Tf} \cap \{x_{n+1} \ge a\}) \le \mathcal{H}^n(G_f \cap \{x_{n+1} \ge a\}).$$

Proof. Since f is Lipschitz, we have the familiar formula (see, e.g., [17, p. 101])

$$\mathcal{H}^n(G_f \cap \{x_{n+1} \ge a\}) = \int\limits_{\{x:f(x)\ge a\}} \left(1 + \|\nabla f(x)\|^2\right)^{1/2} \, dx.$$

Noting that T reduces the modulus of continuity, by Corollary 4.12, we may, as in the proof of Theorem 5.8, assume that Tf is also Lipschitz, so the same formula holds with f replaced by Tf. The result now follows from (5.20) on setting $\Phi(t) = \sqrt{1+t^2} - 1$ for $t \ge 0$, and recalling that $\mathcal{H}^n(\{x : Tf(x) \ge a\}) = \mathcal{H}^n(\{x : f(x) \ge a\})$. $\square$

6. Extension of Theorem 5.8 to $W^{1,1}_{loc}(\mathbb{R}^n)$

We begin by recalling some notions from the theory of Orlicz spaces. Much of the literature is based on N-functions (nice Young functions), that is, functions $\Phi : [0,\infty) \to [0,\infty)$ that are continuous, convex, and such that $\Phi(t) = 0$ if and only if $t = 0$, $\lim_{t\to 0} \Phi(t)/t = 0$, and $\lim_{t\to\infty} \Phi(t)/t = \infty$. However, this restriction is sometimes unnecessary, so we prefer to follow [16] and [47] by working with nontrivial Young functions Φ, those such that $\Phi \not\equiv 0$ and $\Phi \not\equiv \infty$ on $(0,\infty)$.

Let $\Phi : [0,\infty) \to [0,\infty]$ be a nontrivial Young function. The *Orlicz space* $L^\Phi(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is an open set, is the set of all real-valued measurable functions f on Ω such that

$$\|f\|_{L^\Phi(\Omega)} = \inf\left\{\lambda > 0 : \int\limits_\Omega \Phi\left(\frac{|f(x)|}{\lambda}\right) \, dx \le 1\right\} < \infty. \qquad (6.1)$$

The norm defined by (6.1) is called the *Luxemburg norm.*

Note that $f \in L^{\Phi}(\Omega)$ if and only if there is a $c > 0$ such that $\int_{\Omega} \Phi(c|f(x)|)\, dx$ is finite. Indeed, if the latter condition holds, we can choose $\lambda > 0$ large enough that $1/\lambda < c$ and $\int_{\Omega} \Phi(c|f(x)|)\, dx \leq c\lambda$. Then, since $\Phi(t)/t$ is increasing, we have $\lambda \Phi(|f(x)|/\lambda) \leq \Phi(c|f(x)|)/c$ and hence

$$\int_{\Omega} \Phi\left(\frac{|f(x)|}{\lambda}\right)\, dx \leq \frac{1}{c\lambda} \int_{\Omega} \Phi(c|f(x)|)\, dx \leq 1.$$

If $f, f_k \in L^{\Phi}(\Omega)$, $k \in \mathbb{N}$, then

$$\lim_{k \to \infty} \|f_k - f\|_{L^{\Phi}(\Omega)} = 0 \iff \lim_{k \to \infty} \int_{\Omega} \Phi\left(c|f_k(x) - f(x)|\right)\, dx = 0, \text{ for all } c > 0, \quad (6.2)$$

by [16, Proposition 2.1.10(5)].

The *complementary function* $\Psi : [0, \infty) \to [0, \infty]$ *to* Φ, defined by $\Psi(t) = \sup_{s \geq 0}(st - \Phi(s))$, is also a nontrivial Young function, sometimes called the conjugate function of Φ. Another norm on $L^{\Phi}(\Omega)$ is the *Orlicz norm*

$$\|g\|'_{L^{\Phi}(\Omega)} = \sup \left\{ \int_{\Omega} |g(x)h(x)|\, dx : \int_{\Omega} \Psi(|h(x)|)\, dx \leq 1 \right\};$$

by [47, Theorem 132.2], this norm is equivalent to the Luxemburg norm.

The set

$$H^{\Phi}(\Omega) = \left\{ f \in L^{\Phi}(\Omega) : \int_{\Omega} \Phi(c|f(x)|)\, dx < \infty \text{ for all } c > 0 \right\} \quad (6.3)$$

is called the *heart* of $L^{\Phi}(\Omega)$.

We say that the Young function Φ *satisfies a Δ_2 condition at infinity* if there exist $c, t_0 > 0$ such that

$$\Phi(2t) \leq c\, \Phi(t)$$

for $t \geq t_0$. If Φ is in addition nontrivial and real-valued, and if $\mathcal{H}^n(\Omega) < \infty$, then $H^{\Phi}(\Omega) = L^{\Phi}(\Omega)$; see [16, Theorem 2.1.17(2)].

The *Orlicz-Sobolev space* $W^{1,\Phi}(\Omega)$ is defined as

$$W^{1,\Phi}(\Omega) = \{f : f \in L^{\Phi}(\Omega), f \text{ is weakly differentiable in } \Omega, \text{ and } \|\nabla f\| \in L^{\Phi}(\Omega)\}.$$

The norm of $f \in W^{1,\Phi}(\Omega)$ is defined by

$$\|f\|_{W^{1,\Phi}(\Omega)} = \|f\|_{L^{\Phi}(\Omega)} + \big\|\|\nabla f\|\big\|_{L^{\Phi}(\Omega)}. \tag{6.4}$$

Let Φ_1 and Φ_2 be nontrivial Young functions. We say that Φ_1 *dominates* Φ_2 *globally* and write $\Phi_1 \succ \Phi_2$ if there are constants $a, b > 0$ such that

$$b\,\Phi_1(at) \geq \Phi_2(t) \tag{6.5}$$

for all $t \geq 0$. If $\Phi_1 \succ \Phi_2$, then $L^{\Phi_1}(\Omega) \subset L^{\Phi_2}(\Omega)$ by [16, Theorem 2.2.3(1)]. If both $\Phi_1 \succ \Phi_2$ and $\Phi_2 \succ \Phi_1$, we say that Φ_1 and Φ_2 are *equivalent*. In this case, $L^{\Phi_1}(\Omega) = L^{\Phi_2}(\Omega)$ and by [16, Proposition 2.2.1], the norms $\|\cdot\|_{L^{\Phi_1}(\Omega)}$ and $\|\cdot\|_{L^{\Phi_2}(\Omega)}$ are equivalent.

Consider the complementary pair of nontrivial Young functions

$$\Phi_{\min}(t) = \begin{cases} 0, & \text{if } 0 \leq t \leq 1, \\ t - 1, & \text{if } 1 < t < \infty, \end{cases} \qquad \Phi_{\max}(t) = \begin{cases} t, & \text{if } 0 \leq t \leq 1, \\ \infty, & \text{if } 1 < t < \infty; \end{cases}$$

see [16, p. 54]. The following facts are gathered in [16, Proposition 2.2.4] and the remarks that follow it. We have

$$L^{\Phi_{\min}}(\mathbb{R}^n) = \{f_1 + f_\infty : f_1 \in L^1(\mathbb{R}^n), f_\infty \in L^\infty(\mathbb{R}^n)\},$$

with Luxemburg norm

$$\|f\|_{L^{\Phi_{\min}}(\mathbb{R}^n)} = \inf\{\max\{\|f_1\|_{L^1(\mathbb{R}^n)}, \|f_\infty\|_{L^\infty(\mathbb{R}^n)}\} : f = f_1 + f_\infty\} \tag{6.6}$$

for $f \in L^{\Phi_{\min}}(\mathbb{R}^n)$, and $L^{\Phi_{\max}}(\mathbb{R}^n) = L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$, with Luxemburg norm

$$\|f\|_{L^{\Phi_{\max}}(\mathbb{R}^n)} = \max\{\|f\|_{L^1(\mathbb{R}^n)}, \|f\|_{L^\infty(\mathbb{R}^n)}\} \tag{6.7}$$

for $f \in L^{\Phi_{\max}}(\mathbb{R}^n)$. We will often write $L^{\Phi_{\min}}(\mathbb{R}^n) = L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$. Any nontrivial Young function Φ satisfies $\Phi_{\max} \succ \Phi \succ \Phi_{\min}$ and hence all Orlicz spaces with nontrivial Young functions contain $L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ and are contained in $L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$.

The proof of the following lemma is a variant of arguments in [33, p. 77].

Lemma 6.1. *Let Z be the closure of*

$$\left\{ \sum_{j=1}^k a_j 1_{A_j} : a_1, \ldots, a_k \in \mathbb{R}, A_1, \ldots, A_k \in \mathcal{L}^n, k \in \mathbb{N} \right\} \tag{6.8}$$

in $L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$. Then $Z = L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$.

Proof. The step functions in the set (6.8) are clearly bounded and integrable, so we only have to show that $L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n) \subset Z$.

Let $f \in L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$ and suppose that $|f(x)| < M$ for $\mathcal{H}^n$-almost all $x \in \mathbb{R}^n$. For $k \in \mathbb{N}$ define $I_j = [jM/k, (j+1)M/k)$, $A_j = f^{-1}(I_j)$, and

$$a_j = \begin{cases} \min I_j, & \text{if } j \geq 0, \\ \sup I_j, & \text{otherwise,} \end{cases}$$

for $j = -k, \ldots, k-1$. For $j \in \{-1, 0\}$ we have $a_j = 0$, and for all other j,

$$|a_j| \mathcal{H}^n(A_j) \leq \int_{A_j} |f(x)| dx < \infty,$$

as $f \in L^1(\mathbb{R}^n)$. Therefore

$$f_k = \sum_{j=-k}^{k-1} a_j 1_{A_j} \in Z.$$

By construction, $|f_k(x) - f(x)| \leq M/k$ for $\mathcal{H}^n$-almost all $x \in \mathbb{R}^n$, so $\|f_k - f\|_\infty \to 0$ as $k \to \infty$. Since $|f_k(x)| \leq |f(x)|$ for $\mathcal{H}^n$-almost all $x \in \mathbb{R}^n$, the dominated convergence theorem yields $\|f_k - f\|_1 \to 0$ as $k \to \infty$. It follows that $f_k \to f$ as $k \to \infty$ in the norm (6.7), proving the lemma. $\square$

By definition, a subset S of $L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$ is relatively sequentially compact in the weak topology $\sigma\left(L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n), L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)\right)$ if any sequence (f_j) of functions in S has a subsequence (f_{j_k}) converging to some $f \in L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$, in the sense that

$$\lim_{k \to \infty} \int_{\mathbb{R}^n} (f_{j_k}(x) - f(x)) h(x) \, dx = 0$$

for all $h \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. As $\sigma\left(L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n), L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)\right)$ is the only weak topology used in this section, henceforth weak convergence and compactness will always refer to this topology. If S is bounded in the norm (6.6), it is relatively weakly sequentially compact if and only if there is a real-valued Young function $\widetilde{\Phi}$ with $\lim_{t \to \infty} \widetilde{\Phi}(t)/t = \infty$ and

$$\sup_{f \in S} \int_{\mathbb{R}^n} \widetilde{\Phi}(f(x)) \, dx < \infty. \tag{6.9}$$

This is essentially the criterion of de La Vallée Poussin [16, Theorem 2.3.5] combined with [16, Proposition 2.3.16], where the latter uses the σ-finiteness of the underlying measure in the proof.

For the reader's convenience we provide a proof of the following approximation lemma, mainly to stress that Φ need not be an N-function.

Lemma 6.2. *If $\Phi : [0, \infty) \to [0, \infty)$ is a real-valued Young function and $f \in H^\Phi(\mathbb{R}^n)$, there is a sequence (f_j) of $C^\infty(\mathbb{R}^n)$ functions with compact supports that converges to f in $L^\Phi(\mathbb{R}^n)$.*

If, in addition, f is nonnegative, or has bounded support, or $\|\nabla f\| \in H^\Phi(\mathbb{R}^n)$, then f_j, $j \in \mathbb{N}$, can be chosen to be nonnegative, or have uniformly bounded supports, or such that $\lim_{j \to 0} \|f_j - f\|_{W^{1,\Phi}(\mathbb{R}^n)} = 0$, respectively.

Proof. We may assume that $f \in H^\Phi(\mathbb{R}^n)$ has compact support. Indeed, let $0 \leq \phi \leq 1$ be a $C^\infty(\mathbb{R}^n)$ function with support in $2B^n$ and $\phi(x) = 1$ for all $x \in B^n$, and let $\phi_m(x) = \phi(x/m)$ for $m \in \mathbb{N}$ and $x \in \mathbb{R}^n$. Then $\phi_m f \in H^\Phi(\mathbb{R}^n)$ has compact support, and using (6.2) and (6.3), it is easy to see that $\phi_m f$ converges to f as $m \to \infty$ in $L^{\Phi_r}(\mathbb{R}^n)$. If $\|\nabla f\| \in H^\Phi(\mathbb{R}^n)$ is assumed in addition, this convergence even holds in $W^{1,\Phi}(\mathbb{R}^n)$, in view of (6.2) and (6.4). Indeed, since Φ is convex, we have, for $s, t \geq 0$,

$$\Phi(s + t) \leq (1/2)\left(\Phi(2s) + \Phi(2t)\right).$$

Therefore

$$\Phi\left(c\|\nabla(f(1 - \phi_m))\|\right) \leq (1/2)\left(\Phi(2c\|\nabla f\|) + \Phi\left((2c/m)|f|\sup_{\mathbb{R}^n}\|\nabla\phi\|\right)\right).$$

Since $\|\nabla f\| \in H^\Phi(\mathbb{R}^n)$, the integrals of both terms on the right-hand side are finite, and the desired conclusion follows from the dominated convergence theorem. Whenever f is nonnegative, we also have $\phi_m f \geq 0$. Thus if the result holds for functions with compact support, a standard diagonal-type argument shows that it holds generally.

Suppose that $f \in H^\Phi(\mathbb{R}^n)$ has compact support. Let $\rho : \mathbb{R}^n \to [0, \infty)$ be a $C^\infty(\mathbb{R}^n)$ function with support in B^n and integrating to 1, and define $\rho_j(x) = j^n \rho(jx)$ for $j \in \mathbb{N}$ and $x \in \mathbb{R}^n$. The convolution $f_j = f * \rho_j$ is $C^\infty(\mathbb{R}^n)$, satisfies $f_j \geq 0$ when $f \geq 0$, and has support in $M + B^n$ when the support of f is $M \subset \mathbb{R}^n$. We claim that $\|f - f_j\|_{L^\Phi(\mathbb{R}^n)} \to 0$ as $j \to \infty$. To see this, note that since Φ is real-valued, there is a sequence of integrable simple functions (i.e., finite weighted sums of integrable characteristic functions) converging to f in the $L^\Phi(\mathbb{R}^n)$ norm, by [16, Theorem 2.1.14(b)]. Using monotone convergence, the proof of the latter theorem can easily be modified to show that these simple functions can be assumed to have bounded supports. Thus, for each $\varepsilon > 0$ there is a simple function h with $|h(x)| \leq b1_{RB^n}(x)$ for some $b, R > 0$ and all $x \in \mathbb{R}^n$, such that $\|f - h\|_{L^\Phi(\mathbb{R}^n)} \leq \varepsilon/4$. If $h_j = h * \rho_j$, Jensen's inequality (see, e.g., [33, p. 62, Proposition 5]) implies that

$$\|f_j - h_j\|_{L^\Phi(\mathbb{R}^n)} = \|(f - h) * \rho_j\|_{L^\Phi(\mathbb{R}^n)} \leq \|f - h\|_{L^\Phi(\mathbb{R}^n)} \leq \varepsilon/4,$$

so

$$\|f - f_j\|_{L^\Phi(\mathbb{R}^n)} \leq \|f - h\|_{L^\Phi(\mathbb{R}^n)} + \|h - h_j\|_{L^\Phi(\mathbb{R}^n)} + \|h_j - f_j\|_{L^\Phi(\mathbb{R}^n)}$$

$$\leq \frac{\varepsilon}{2} + \|h - h_j\|_{L^\Phi(\mathbb{R}^n)}. \tag{6.10}$$

We have

$$|h(x) - h_j(x)| \leq \int_{\mathbb{R}^n} |h(x) - h(y)|\rho_j(x - y)dy \leq \|\rho_j\|_\infty \int_{B(x,1/j)} |h(x) - h(y)|\, dy.$$

As h is integrable, the Lebesgue differentiation theorem (see, e.g., [25, Proposition 3.5.4]) implies that $h_j(x) \to h(x)$ for $\mathcal{H}^n$-almost all $x \in \mathbb{R}^n$. Since $|h_j(x)| \leq b1_{(R+1)B^n}(x)$ for all $x \in \mathbb{R}^n$ and Φ is real-valued, the dominated convergence theorem gives $\lim_{j\to\infty} \int_{\mathbb{R}^n} \Phi(c|h(x) - h_j(x)|)\, dx = 0$ for all $c > 0$. Therefore, by (6.2), there is a $j_0 \in \mathbb{N}$ such that $\|h - h_j\|_{L^\Phi(\mathbb{R}^n)} \leq \varepsilon/2$ for all $j \geq j_0$, and inserting this into (6.10) proves the claim.

Since $f_j \geq 0$ when $f \geq 0$, it only remains to deal with the statement involving the extra assumption that $\|\nabla f\| \in H^\Phi(\mathbb{R}^n)$. For this, we may proceed exactly as in the proof of [15, Theorem 2.1]. If f and $\|\nabla f\|$ are functions in $H^\Phi(\mathbb{R}^n)$ with compact support, we have $\partial f_j/\partial x_i = (\partial f/\partial x_i) * \rho_j$ (where the weak derivative is used on the right-hand side). Applying the above, both to f and with f replaced by $\partial f/\partial x_i$, $i = 1, \ldots, n$, yields $\lim_{j\to 0} \|f - f_j\|_{W^{1,\Phi}(\mathbb{R}^n)} = 0$, as required. $\quad\square$

Theorem 6.3. *Let $T : \mathcal{V}(\mathbb{R}^n) \to \mathcal{V}(\mathbb{R}^n)$ be a smoothing rearrangement and let Φ be a Young function. If $f \in W^{1,1}_{loc}(\mathbb{R}^n) \cap \mathcal{V}(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n} \Phi(\|\nabla f(x)\|)\, dx < \infty$, then $Tf \in W^{1,1}_{loc}(\mathbb{R}^n)$ and*

$$\int_{\mathbb{R}^n} \Phi(\|\nabla Tf(x)\|)\, dx \leq \int_{\mathbb{R}^n} \Phi(\|\nabla f(x)\|)\, dx. \tag{6.11}$$

Proof. As at the beginning of the proof of Theorem 5.8, we may assume that Φ is a nontrivial real-valued function. We first aim to prove (6.11) when $\mathcal{H}^n(\operatorname{supp} f) < \infty$. Let Ω be an open set such that $\operatorname{supp} f \subset \Omega$ and $\mathcal{H}^n(\Omega) < \infty$. The proof will proceed via a succession of claims.

For $r > \sup\{t \geq 0 : \Phi(t) = 0\}$, define

$$\Lambda_r(t) = \begin{cases} \Phi(t), & \text{if } 0 \leq t < r, \\ \Phi'^+(r)(t - r) + \Phi(r), & \text{otherwise,} \end{cases}$$

where $\Phi'^+(r)$ is the right derivative of Φ at r. Then $\Phi_r = \max\{0, \Lambda_r - 1/r\}$ is a nontrivial real-valued (and hence continuous) Young function with $\Phi_r \leq \Phi$. By construction, there is a $\delta > 0$ such that $\delta \leq \Phi'^+_r(t) \leq 1/\delta$ for all $t > t_0 = \sup\{t \geq 0 : \Phi_r(t) = 0\} > 0$. It follows that for $t \geq t_0$, we have

$$\delta(t - t_0) = \int_{t_0}^t \delta\, ds \leq \Phi_r(t) \leq \int_{t_0}^t \frac{1}{\delta}\, ds = \frac{1}{\delta}(t - t_0).$$

Comparing (6.5), we see that Φ_r is equivalent to $\Phi_{\min}$. Therefore, as we remarked before Lemma 6.2, the $L^{\Phi_r}(\mathbb{R}^n)$ norm is equivalent to the norm (6.6) and hence $L^{\Phi_r}(\mathbb{R}^n) = L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)$. If Ψ_r denotes the complementary Young function to Φ_r, then $L^{\Psi_r}(\mathbb{R}^n) = L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ and the $L^{\Psi_r}(\mathbb{R}^n)$ norm is equivalent to (6.7).

Our first claim is that f can be approximated in the $W^{1,\Phi_r}(\mathbb{R}^n)$ norm by a sequence (f_j) of nonnegative $C^\infty(\mathbb{R}^n)$ functions with compact support. To see this, note that since $0 \le \Phi_r \le \Phi$, our assumptions give $\int_{\mathbb{R}^n} \Phi_r(\|\nabla f(x)\|)\, dx < \infty$. By [39, Lemma 3], this yields the existence of a $c > 0$ such that $\int_{\mathbb{R}^n} \Phi_r(c|f(x)|)\, dx$ is finite, and hence $f \in L^{\Phi_r}(\mathbb{R}^n)$, as was explained in the remarks at the beginning of this section. Therefore $f \in L^{\Phi_r}(\Omega)$ and our assumptions and $0 \le \Phi_r \le \Phi$ imply that $\|\nabla f\| \in L^{\Phi_r}(\Omega)$. As was mentioned before Lemma 6.1, $L^{\Phi_r}(\Omega) = H^{\Phi_r}(\Omega)$ because Φ_r satisfies a Δ_2 condition at infinity and $\mathcal{H}^n(\Omega) < \infty$. Since f vanishes on $\mathbb{R}^n \setminus \Omega$, we have $f, \|\nabla f\| \in H^{\Phi_r}(\mathbb{R}^n)$. By Lemma 6.2, there exists a sequence (f_j) of nonnegative $C^\infty(\mathbb{R}^n)$ functions with compact support that converges to f in the $W^{1,\Phi_r}(\mathbb{R}^n)$ norm. This completes the proof of the first claim.

Our second claim is that (Tf_j) converges to Tf in $L^{\Phi_r}(\mathbb{R}^n)$, $Tf \in W^{1,1}_{loc}(\mathbb{R}^n)$, and a subsequence of (∇Tf_j) converges weakly to the weak gradient of Tf. Indeed, Proposition 3.7 implies that $\|Tf_j - Tf\|_{L^{\Phi_r}(\mathbb{R}^n)} \le \|f_j - f\|_{L^{\Phi_r}(\mathbb{R}^n)}$. Thus $Tf_j \to Tf$ in $L^{\Phi_r}(\mathbb{R}^n)$ and, in particular, $Tf \in L^{\Phi_r}(\mathbb{R}^n)$ and $Tf \in L^1_{loc}(\mathbb{R}^n)$. We also have

$$\lim_{j \to \infty} \big\|\|\nabla f_j\|\big\|_{L^{\Phi_r}(\mathbb{R}^n)} = \big\|\|\nabla f\|\big\|_{L^{\Phi_r}(\mathbb{R}^n)}. \tag{6.12}$$

The map T and each f_j satisfy the hypotheses of Theorem 5.8, so for $j \ge 1$, the latter theorem implies that the nonnegative function Tf_j agrees with a Lipschitz function $\mathcal{H}^n$-almost everywhere, and

$$\big\|\|\nabla Tf_j\|\big\|_{L^{\Phi_r}(\mathbb{R}^n)} \le \big\|\|\nabla f_j\|\big\|_{L^{\Phi_r}(\mathbb{R}^n)}. \tag{6.13}$$

Due to (6.12) and (6.13), the set $\{\|\nabla Tf_j\|, j \ge 1\}$ is bounded in $L^{\Phi_r}(\mathbb{R}^n)$ and therefore also in $L^{\Phi_{\min}}(\mathbb{R}^n)$ by norm equivalence. Since $\|\nabla f_j\|$ converges in $L^{\Phi_r}(\mathbb{R}^n)$, it also converges in $L^{\Phi_{\min}}(\mathbb{R}^n)$, and the criterion of de La Vallée Poussin stated before Lemma 6.2 yields the existence of a real-valued Young function $\widetilde{\Phi}$ with $\lim_{t \to \infty} \widetilde{\Phi}(t)/t = \infty$ such that (6.9) holds with $S = \{\|\nabla f_j\| : j \ge 1\}$. Theorem 5.8, applied with $\widetilde{\Phi}$ instead of Φ, shows that (6.9) also holds for $S = \{\|\nabla Tf_j\| : j \ge 1\}$ with the same $\widetilde{\Phi}$. De La Vallée Poussin's criterion now shows the relative weak compactness of $\{\|\nabla Tf_j\| : j \ge 1\}$. Hence, there is a subsequence of (∇Tf_j), also denoted (∇Tf_j), and a vector field $g \in \big(L^1(\mathbb{R}^n) + L^\infty(\mathbb{R}^n)\big)^n = \big(L^{\Phi_r}(\mathbb{R}^n)\big)^n$, such that ∇Tf_j converges weakly to g, i.e., for each $h \in \big(L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)\big)^n = \big(L^{\Psi_r}(\mathbb{R}^n)\big)^n$,

$$\lim_{j \to \infty} \int_{\mathbb{R}^n} (\nabla Tf_j)(x)h(x)\, dx = \int_{\mathbb{R}^n} g(x)h(x)\, dx.$$

If $h \in C^\infty(\mathbb{R}^n)$ has compact support, then $h, \partial h/\partial x_i \in L^{\Psi_r}(\mathbb{R}^n)$ and

$$
\int_{\mathbb{R}^n} g_i(x) h(x)\, dx = \lim_{j \to \infty} \int_{\mathbb{R}^n} \frac{\partial T f_j}{\partial x_i}(x) h(x)\, dx = -\lim_{j \to \infty} \int_{\mathbb{R}^n} T f_j(x) \frac{\partial h}{\partial x_i}(x)\, dx
$$

$$
= -\int_{\mathbb{R}^n} T f(x) \frac{\partial h}{\partial x_i}(x)\, dx, \tag{6.14}
$$

where we have used the fact that the convergence of Tf_j to Tf in $L^{\Phi_r}(\mathbb{R}^n)$ also implies weak convergence. From (6.14) we see that g is the weak gradient of Tf, and the fact that $\partial Tf/\partial x_i$ is locally integrable is a direct consequence of the fact that it belongs to $L^{\Psi_r}(\mathbb{R}^n)$. The second claim is proved.

Next, we claim that

$$
\big\| \|\nabla Tf\| \big\|_{L^{\Phi_r}(\mathbb{R}^n)} \leq \liminf_{j \to \infty} \big\| \|\nabla T f_j\| \big\|_{L^{\Phi_r}(\mathbb{R}^n)}. \tag{6.15}
$$

With this goal in mind, we first prove that if $u \in (L^{\Phi_r}(\mathbb{R}^n))^n$, then

$$
\big\| \|u\| \big\|_{L^{\Phi_r}(\mathbb{R}^n)} = \sup \left\{ \left| \int_{\mathbb{R}^n} u(x) \cdot h(x)\, dx \right| : h \in (L^{\Psi_r}(\mathbb{R}^n))^n, \big\| \|h\| \big\|'_{L^{\Psi_r}(\mathbb{R}^n)} \leq 1 \right\}. \tag{6.16}
$$

A similar formula in the scalar case is proved in [33, (10), Proposition 10, Section 3.4]; when applied with f and Φ there replaced by $\|u\|$ and Φ_r, respectively, it becomes

$$
\big\| \|u\| \big\|_{L^{\Phi_r}(\mathbb{R}^n)} = \sup \left\{ \left| \int_{\mathbb{R}^n} \|u(x)\| v(x)\, dx \right| : v \in \mathcal{M}^{\Psi_r}, \|v\|'_{L^{\Psi_r}(\mathbb{R}^n)} \leq 1 \right\}, \tag{6.17}
$$

where $\mathcal{M}^{\Psi_r}$ is the closure of the span of all step functions in $L^{\Psi_r}(\mathbb{R}^n)$. (We warn the reader that in [33], $\|u\|_{L^\Phi(\mathbb{R}^n)}$ and $\|u\|'_{L^\Phi(\mathbb{R}^n)}$ are denoted by $N_\phi(u)$ and $\|u\|_\Phi$, respectively.) We have already seen that $L^{\Psi_r}(\mathbb{R}^n) = L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ and that the two spaces have equivalent norms, so $\mathcal{M}^{\Psi_r} = L^{\Psi_r}(\mathbb{R}^n)$ due to Lemma 6.1. Thus, (6.17) becomes

$$
\big\| \|u\| \big\|_{L^{\Phi_r}(\mathbb{R}^n)} = \sup \left\{ \left| \int_{\mathbb{R}^n} \|u(x)\| v(x)\, dx \right| : v \in L^{\Psi_r}(\mathbb{R}^n), \|v\|'_{L^{\Psi_r}(\mathbb{R}^n)} \leq 1 \right\}. \tag{6.18}
$$

Let S_1 and S_2 denote the right-hand sides of (6.16) and (6.18), respectively. The Cauchy-Schwartz inequality yields $S_1 \leq S_2$. We can restrict the supremum in (6.18) to nonnegative v. If $v \in L^{\Psi_r}(\mathbb{R}^n)$, $v \geq 0$, $\|v\|'_{L^{\Psi_r}(\mathbb{R}^n)} \leq 1$, and

$$
h(x) = \begin{cases} v(x) \dfrac{u(x)}{\|u(x)\|}, & \text{if } u(x) \neq 0, \\ 0, & \text{if } u(x) = 0, \end{cases}
$$

then $h \in (L^{\Psi_r}(\mathbb{R}^n))^n$, $\big\| \|h\| \big\|'_{L^{\Psi_r}(\mathbb{R}^n)} \le 1$, and

$$
S_1 \ge \left| \int_{\mathbb{R}^n} u(x) \cdot h(x)\, dx \right| = \left| \int_{\mathbb{R}^n} \|u(x)\| v(x)\, dx \right|.
$$

This proves that $S_1 \ge S_2$ and concludes the proof of (6.16). Now (6.16) with u replaced by $\nabla T f$ and $\nabla T f_j$ gives

$$
\big\| \|\nabla T f\| \big\|_{L^{\Phi_r}(\mathbb{R}^n)}
$$

$$
= \sup \left\{ \left| \int_{\mathbb{R}^n} \nabla T f(x) \cdot h(x)\, dx \right| : h \in (L^{\Psi_r}(\mathbb{R}^n))^n,\ \big\| \|h\| \big\|'_{L^{\Psi_r}(\mathbb{R}^n)} \le 1 \right\}
$$

$$
= \sup \left\{ \lim_{j \to \infty} \left| \int_{\mathbb{R}^n} \nabla T f_j(x) \cdot h(x)\, dx \right| : h \in (L^{\Psi_r}(\mathbb{R}^n))^n,\ \big\| \|h\| \big\|'_{L^{\Psi_r}(\mathbb{R}^n)} \le 1 \right\}
$$

$$
\le \liminf_{j \to \infty} \sup \left\{ \left| \int_{\mathbb{R}^n} \nabla T f_j(x) \cdot h(x)\, dx \right| : h \in (L^{\Psi_r}(\mathbb{R}^n))^n,\ \big\| \|h\| \big\|'_{L^{\Psi_r}(\mathbb{R}^n)} \le 1 \right\}
$$

$$
= \liminf_{j \to \infty} \big\| \|\nabla T f_j\| \big\|_{L^{\Phi_r}(\mathbb{R}^n)}.
$$

This proves (6.15).

From (6.12), (6.13), and (6.15), we conclude that

$$
\big\| \|\nabla T f\| \big\|_{L^{\Phi_r}(\mathbb{R}^n)} \le \big\| \|\nabla f\| \big\|_{L^{\Phi_r}(\mathbb{R}^n)}. \tag{6.19}
$$

Our fourth claim is that (6.11) holds when Φ is replaced by Φ_r. To see this, note that (6.19) holds if Φ_r is replaced by $a\,\Phi_r$ for any $a > 0$, because all the preceding arguments are valid with this replacement, due to $\int_{\mathbb{R}^n} a\,\Phi_r(\|\nabla f(x)\|)\, dx < \infty$. If we choose a so that $\int_{\mathbb{R}^n} a\,\Phi_r(\|\nabla f(x)\|)\, dx = 1$, then $\big\| \|\nabla f\| \big\|_{L^{a\,\Phi_r}(\mathbb{R}^n)} = 1$, by [16, Proposition 2.1.10(4)] and the fact that $\|\nabla f\| \in H^{a\Phi_r}(\mathbb{R}^n)$. Thus (6.19) becomes

$$
\big\| \|\nabla T f\| \big\|_{L^{a\Phi_r}(\mathbb{R}^n)} \le 1.
$$

By [16, Proposition 2.1.10(2)], the previous inequality holds if and only if

$$
\int_{\mathbb{R}^n} a\,\Phi_r(\|\nabla T f(x)\|)\, dx \le 1 = \int_{\mathbb{R}^n} a\,\Phi_r(\|\nabla f(x)\|)\, dx. \tag{6.20}
$$

This proves the fourth claim.

Since the nonnegative function Φ_r increases to Φ as $r \to \infty$, one may apply the monotone convergence theorem to both sides of (6.20) and obtain (6.11) when

$\mathcal{H}^n(\operatorname{supp} f) < \infty$. To remove the latter restriction, let $f \in W^{1,1}_{loc}(\mathbb{R}^n) \cap \mathcal{V}(\mathbb{R}^n)$ and let $f_c = \max\{f - c, 0\}$ for $c > 0$. Then f_c satisfies the hypotheses of the theorem and $\mathcal{H}^n(\operatorname{supp} f_c) < \infty$. Moreover, $\nabla f_c = \nabla f$ on $\{x : f(x) > c\}$ and $T f_c = \max\{T f - c, 0\}$, by Proposition 3.6 with $\varphi(t) = (t - c)^+$, so $\nabla T f_c = \nabla T f$ on $\{x : T f(x) > c\}$. By (6.11) with f replaced by f_c, we have

$$\int\limits_{\{x:Tf(x)>c\}} \Phi(\|\nabla Tf(x)\|)\, dx = \int\limits_{\mathbb{R}^n} \Phi(\|\nabla Tf_c(x)\|)\, dx$$

$$\leq \int\limits_{\mathbb{R}^n} \Phi(\|\nabla f_c(x)\|)\, dx = \int\limits_{\{x:f(x)>c\}} \Phi(\|\nabla f(x)\|)\, dx.$$

Letting $c \to 0$ and applying the monotone convergence theorem, we obtain (6.11). $\quad\square$

Lemma 6.4. *Let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$ and let $T : X \to X$ be a smoothing rearrangement. If $f \in W^{1,\infty}(\mathbb{R}^n) \cap X$, then $Tf \in W^{1,\infty}(\mathbb{R}^n)$ and*

$$\operatorname*{ess\,sup}_{x \in \mathbb{R}^n} \|\nabla Tf(x)\| \leq \operatorname*{ess\,sup}_{x \in \mathbb{R}^n} \|\nabla f(x)\|. \tag{6.21}$$

Proof. If $f \in W^{1,\infty}(\mathbb{R}^n) \cap X$, then by [3, Proposition 3.17], f coincides $\mathcal{H}^n$-almost everywhere with a Lipschitz function in X. We may therefore assume that f is Lipschitz, and then, as in the proof of Theorem 5.8, also assume that Tf is Lipschitz. By [3, Corollary 3.4], (6.21) is equivalent to $L_1 \leq L_2$, where L_1 and L_2 are the Lipschitz constants of Tf and f, respectively. Since

$$L_1 = \sup_{d>0} \omega_d(Tf)/d \leq \sup_{d>0} \omega_d(f)/d = L_2,$$

the proof is complete. $\quad\square$

Corollary 6.5. *Let $T : \mathcal{V}(\mathbb{R}^n) \to \mathcal{V}(\mathbb{R}^n)$ be a smoothing rearrangement and let $1 \leq p \leq \infty$. If $f \in W^{1,p}(\mathbb{R}^n) \cap \mathcal{V}(\mathbb{R}^n)$, then $Tf \in W^{1,p}(\mathbb{R}^n)$ and*

$$\left\|\|\nabla Tf(x)\|\right\|_p \leq \left\|\|\nabla f(x)\|\right\|_p, \tag{6.22}$$

where $\|\cdot\|_p$ denotes the L^p norm when $1 \leq p < \infty$ and the essential supremum over $\mathbb{R}^n$ when $p = \infty$.

Proof. The case when $p = \infty$ corresponds to Lemma 6.4. Suppose that $1 \leq p < \infty$ and let $f \in W^{1,p}(\mathbb{R}^n) \cap \mathcal{V}(\mathbb{R}^n)$. As $W^{1,p}(\mathbb{R}^n) \subset W^{1,1}_{loc}(\mathbb{R}^n)$, Theorem 6.3 with $\Phi(t) = t^p$ gives (6.22). Our assumptions on f show that the right-hand side of (6.22) is finite, implying that $Tf \in W^{1,p}(\mathbb{R}^n)$. $\quad\square$

The following example shows that the assumption that T is smoothing in Theorems 5.8 and 6.3 and Corollary 6.5 cannot be dropped in general. It also shows that when $\Phi(t) = t$ in Theorem 5.8 or 6.3, or $p = 1$ in Corollary 6.5, the smoothing property is not necessary for the stated inequalities to hold.

Example 6.6. Let $K \in \mathcal{K}^n_{(o)}$ and $\mathcal{H}^n(K) = \kappa_n$ and let $T : \mathcal{V}(\mathbb{R}^n) \to \mathcal{V}(\mathbb{R}^n)$ be the rearrangement defined in Example 4.3(ii). We make the following two claims.

(i) Inequalities (5.21) and (6.11) for $\Phi(t) = t$, and (6.22) for $p = 1$, each hold if and only if $K = x + B^n$ for some $x \in D^n$.
(ii) Inequalities (5.21) and (6.11) for strictly convex real-valued Φ, and (6.22) for $1 < p \le \infty$, each hold if and only if $K = B^n$.

Suppose these claims are true. If K is not a ball, then T is not smoothing by Example 4.3(ii) and Theorems 5.8 and 6.3 (for real-valued strictly convex Φ), and Corollary 6.5 fail by (i) and (ii). If $K = x + B^n$ for some $o \ne x \in D^n$, then T is not smoothing by Example 4.3(ii) but nevertheless Theorems 5.8 and 6.3 with $\Phi(t) = t$, and Corollary 6.5 with $p = 1$, hold by (i).

To prove the two claims, we note first that if $f \in \mathcal{V}(\mathbb{R}^n)$ and $\alpha_{f,t} = \mathcal{H}^n(\{x : f(x) \ge t\})$ for $t \ge 0$, the definition of T yields

$$\{x : Tf(x) \ge t\} = \left(\frac{\alpha_{f,t}}{\kappa_n} \right)^{1/n} K \tag{6.23}$$

for $t > 0$.

Now let $M \in \mathcal{K}^n_{(o)}$ and $\mathcal{H}^n(M) = \kappa_n$, and let $f_M(x) = (1 - h_{M^\circ}(x))^+$ for $x \in \mathbb{R}^n$, where M° is the polar body of M and s^+ is the nonnegative part of $s \in \mathbb{R}$. Then $\{x : f_M(x) \ge t\} = (1 - t)^+ M$ for $t > 0$. Since $\alpha_{f_M,t} = ((1 - t)^+)^n \kappa_n$, (6.23) with $f = f_M$ implies that

$$Tf_M = f_K. \tag{6.24}$$

When $M = B^n$, we have $f_{B^n}(x) = (1 - \|x\|)^+$ for $x \in \mathbb{R}^n$ and hence

$$\|\nabla f_{B^n}(x)\| = \begin{cases} 1, & \text{if } x \in D^n \setminus \{o\}, \\ 0, & \text{if } x \notin B^n. \end{cases} \tag{6.25}$$

The coarea formula for Lipschitz functions f on $\mathbb{R}^n$ (see [3, Theorem 4.19], [17, Theorem 1, p. 112]) states that

$$\int_{\mathbb{R}^n} \|\nabla f(x)\| \, dx = \int_0^\infty \mathcal{H}^{n-1}(\{x : f(x) = t\}) \, dt. \tag{6.26}$$

Suppose that T satisfies (5.21) or (6.11) with $\Phi(t) = t$, or (6.22) with $p = 1$. Then, using (6.26) with $f = f_K$, (6.24) with $M = B^n$, either (5.21), or (6.11), or (6.22) with $f = f_{B^n}$, and (6.25), we obtain

$$
\mathcal{H}^{n-1}(\partial K) = n \int_0^1 (1-t)^{n-1} \mathcal{H}^{n-1}(\partial K)\, dt = n \int_0^\infty \mathcal{H}^{n-1}(\{x : f_K(x) = t\})\, dt \quad (6.27)
$$

$$
= n \int_{\mathbb{R}^n} \|\nabla f_K(x)\|\, dx = n \int_{\mathbb{R}^n} \|\nabla T f_{B^n}(x)\|\, dx
$$

$$
\leq n \int_{\mathbb{R}^n} \|\nabla f_{B^n}(x)\|\, dx = n\kappa_n = \mathcal{H}^{n-1}(\partial B^n).
$$

Because $\mathcal{H}^n(K) = \mathcal{H}^n(B^n)$, equality must hold in the isoperimetric inequality and consequently $K = x + B^n$ for some $x \in D^n$. Conversely, suppose that $K = x + B^n$ for some $x \in D^n$ and $f \in W^{1,1}(\mathbb{R}^n) \cap \mathcal{V}(\mathbb{R}^n)$. From (6.23) and the fact that the decreasing function $t \mapsto \alpha_{f,t}$ can only have countably many discontinuities, we get

$$
\{x : Tf(x) = t\} = \{x : Tf(x) \geq t\} \setminus \bigcup_{s > t} \{x : Tf(x) \geq s\} \subset \left(\frac{\alpha_{f,t}}{\kappa_n}\right)^{1/n} \partial K
$$

for almost all $t > 0$. Since K is a translate of B^n, this and (6.26) imply that

$$
\int_{\mathbb{R}^n} \|\nabla Tf(x)\|\, dx = \int_0^\infty \mathcal{H}^{n-1}(\{x : Tf(x) = t\})\, dt \leq \int_0^\infty \mathcal{H}^{n-1}\left(\left(\frac{\alpha_{f,t}}{\kappa_n}\right)^{1/n} \partial K\right) dt
$$

$$
= \int_0^\infty \mathcal{H}^{n-1}\left(\left(\frac{\alpha_{f,t}}{\kappa_n}\right)^{1/n} \partial B^n\right) dt = \int_{\mathbb{R}^n} \|\nabla f^\#(x)\|\, dx.
$$

Hence, T satisfies (5.21) and (6.11) with $\Phi(t) = t$, and (6.22) with $p = 1$, as the Schwarz rearrangement does so. This proves (i).

For (ii), assume first that Φ is real-valued and strictly convex, and that (5.21) or (6.11) holds. Let $\Phi(1) = c > 0$. Note that the measure on K with differential dx/κ_n is a probability measure, since $\mathcal{H}^n(K) = \kappa_n$. We use (6.25), either (5.21) or (6.11), Jensen's inequality (see, e.g. [33, p. 62, Proposition 5]), (6.27), and the isoperimetric inequality, to obtain

$$
c = \int_{\mathbb{R}^n} \Phi\left(\|\nabla f_{B^n}(x)\|\right) \frac{dx}{\kappa_n} \geq \int_{\mathbb{R}^n} \Phi\left(\|\nabla T f_{B^n}(x)\|\right) \frac{dx}{\kappa_n} \geq \int_K \Phi\left(\|\nabla T f_{B^n}(x)\|\right) \frac{dx}{\kappa_n}
$$

$$
\geq \Phi\left(\int_K \|\nabla T f_{B^n}(x)\| \frac{dx}{\kappa_n}\right) = \Phi\left(\frac{1}{n\kappa_n} \mathcal{H}^{n-1}(\partial K)\right) \geq \Phi\left(\frac{1}{n\kappa_n} \mathcal{H}^{n-1}(\partial B^n)\right)
$$

$$
= \Phi(1) = c.
$$

It follows that there is equality in the isoperimetric inequality, giving $K = x + B^n$ for some $x \in D^n$, as before. But now equality also holds in Jensen's inequality with the strictly convex function Φ, so $\|\nabla T f_{B^n}(\cdot)\| = \|\nabla f_K(\cdot)\|$ must be constant $\mathcal{H}^n$-almost everywhere on K. This is only possible when $x = o$. If we assume instead that (6.22) holds with $1 < p < \infty$, we can apply the same argument with $\Phi(t) = t^p$. Conversely, when $K = B^n$, the rearrangement T is the Schwarz rearrangement and therefore the Pólya-Szegő inequality holds.

Finally, for (ii) when $p = \infty$, take $f = f_{B^n}$ in (6.22). We have $T f_{B^n} = f_K$ by (6.24), and it is clear that $\operatorname{ess\,sup}_{x \in \mathbb{R}^n} \|\nabla f_{B^n}(x)\| = 1$, while $\operatorname{ess\,sup}_{x \in \mathbb{R}^n} \|\nabla f_K(x)\| > 1$ if and only if $K \neq B^n$. This completes the proof of (ii).

7. The anisotropic case

The following result generalizes Theorem 5.8, which corresponds to the case when $K = B^n$.

Theorem 7.1. *Let $X = \mathcal{S}(\mathbb{R}^n)$ or $\mathcal{V}(\mathbb{R}^n)$, let $K \in \mathcal{K}_{(o)}^n$, let $T : X \to X$ be a rearrangement, and let Φ be a Young function. If T is K-smoothing and $f \in X$ is Lipschitz, then Tf coincides with a Lipschitz function $\mathcal{H}^n$-almost everywhere on $\mathbb{R}^n$, and*

$$\int_{\mathbb{R}^n} \Phi\left(h_{-K}(\nabla T f(x))\right) \, dx \leq \int_{\mathbb{R}^n} \Phi\left(h_{-K}(\nabla f(x))\right) \, dx, \tag{7.1}$$

where the integrals may be infinite.

Proof. We can argue as in the proof of Theorem 5.8, with very few changes. With arguments as at the beginning of that proof, we may assume without loss of generality that Φ is a nontrivial real-valued Young function. Let $f \in X$ be Lipschitz with Lipschitz constant L. Choosing $0 < r \leq R$ such that $rB^n \subset K \subset RB^n$, we may use Lemma 5.7 to conclude that Tf coincides $\mathcal{H}^n$-almost everywhere with a Lipschitz function with Lipschitz constant at most LR/r. Inequality (5.22) follows as before. Instead of (5.23) we have

$$\max\left\{\|\nabla f(x)\|, \|\nabla T f(x)\|\right\} \leq LR/r$$

for $\mathcal{H}^n$-almost all $x \in \mathbb{R}^n$. Then, if C is the convex body from Lemma 5.6 corresponding to $M = LR/r$ and Φ, we have

$$h_C(y, 1) - 1 = b\,\Phi(h_K(y)) = b\,\Phi(h_{-K}(-y))$$

for $y \in \mathbb{R}^n$ with $h_K(y) \leq M$ and some $b > 0$. As before, this leads to (7.1). $\quad\square$

Finally, we present in Theorem 7.3 an anisotropic version of Theorem 6.3, which again corresponds to the case when $K = B^n$. We shall need the following lemma. Recall that a convex body is *smooth* if all its boundary points are regular and *strictly convex* if it does not contain a line segment in its boundary; see [36, pp. 83, 87].

Lemma 7.2. *Let $L \in \mathcal{K}_{(o)}^n$ be a smooth and strictly convex body, and let $u : \mathbb{R}^n \to \mathbb{R}^n$ and $v : \mathbb{R}^n \to \mathbb{R}$ be measurable. Then there is a measurable $w : \mathbb{R}^n \to \mathbb{R}^n$ such that*
(i) $h_{L^\circ}(w(x)) = v(x)$ *for all x such that $u(x) \neq o$, and*
(ii) $u(x) \cdot w(x) = h_L(u(x)) \, h_{L^\circ}(w(x))$ *for all $x \in \mathbb{R}^n$.*

Proof. If $A = \{x : u(x) \neq 0\}$, then A is $\mathcal{H}^n$-measurable. As is observed in [36, Remark 1.7.14], it follows easily from (2.3) that since L is smooth and strictly convex, the same is true of L°. Let $n(L^\circ, y)$ denote the unit outer normal to L° at $y \in \partial L^\circ$. Note that $u/h_L(u) = \rho_{L^\circ}(u)u \in \partial L^\circ$ by (2.3), and define $\eta : S^{n-1} \to S^{n-1}$ by

$$\eta(u) = n\left(L^\circ, u/h_L(u)\right).$$

The map η is continuous, since L° is a convex body of class C^1 by [36, Theorem 2.2.4]. The map $f : \mathbb{R}^n \setminus \{o\} \to S^{n-1}$ defined by $f(z) = z/\|z\|$ is also continuous. The composition $\eta \circ f \circ u : A \to S^{n-1}$ of measurable functions is therefore also measurable, and so is its composition with the continuous support function $h_{L^\circ} : \mathbb{R}^n \to \mathbb{R}$ (see [36, p. 115]). Define

$$w(x) = \begin{cases} v(x) \frac{\eta(f(u(x)))}{h_{L^\circ}(\eta(f(u(x))))}, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}$$

Then $w : \mathbb{R}^n \to \mathbb{R}^n$ is measurable and clearly satisfies (i). By its definition, $w(x)$ is an outer normal to L° at $f(u(x))/h_L(f(u(x))) = u(x)/h_L(u(x))$, so (ii) holds due to the equality condition for (2.7) stated immediately after it. $\square$

Theorem 7.3. *Let $K \in \mathcal{K}_{(o)}^n$, let $T : \mathcal{V}(\mathbb{R}^n) \to \mathcal{V}(\mathbb{R}^n)$ be a K-smoothing rearrangement, and let Φ be a Young function. If $f \in W_{loc}^{1,1}(\mathbb{R}^n) \cap \mathcal{V}(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n} \Phi\left(h_{-K}(\nabla f(x))\right) dx < \infty$, then $Tf \in W_{loc}^{1,1}(\mathbb{R}^n)$ and*

$$\int\limits_{\mathbb{R}^n} \Phi\left(h_{-K}(\nabla Tf(x))\right) dx \leq \int\limits_{\mathbb{R}^n} \Phi\left(h_{-K}(\nabla f(x))\right) dx. \tag{7.2}$$

Proof. The proof follows that of Theorem 6.3, up to the end of the proof of the second claim, verbatim except that the role of Theorem 5.8 is now played by Theorem 7.1 and that $\|\nabla f\|$ and $\|\nabla Tf\|$ are replaced by $h_{-K}(\nabla f)$ and $h_{-K}(\nabla Tf)$, respectively. Instead of the third claim, that (6.15) holds, we claim that

$$\left\|h_{-K}(\nabla Tf)\right\|_{L^{\Phi_r}(\mathbb{R}^n)} \leq \liminf_{j \to \infty} \left\|h_{-K}(\nabla Tf_j)\right\|_{L^{\Phi_r}(\mathbb{R}^n)}. \tag{7.3}$$

The argument is similar to the one given for (6.15). We first prove that if $u \in (L^{\Phi_r}(\mathbb{R}^n))^n$, then

$$\left\| h_{-K}(u) \right\|_{L^{\Phi_r}(\mathbb{R}^n)} = \sup \left\{ \left| \int_{\mathbb{R}^n} u(x) \cdot w(x) \, dx \right| : w \in (L^{\Psi_r}(\mathbb{R}^n))^n, \left\| h_{-K^\circ}(w) \right\|'_{L^{\Psi_r}(\mathbb{R}^n)} \le 1 \right\},$$
(7.4)

where $-K^\circ = (-K)^\circ$ is the polar body of $-K$; see (2.2). For this, we apply [33, (10), Proposition 10, Section 3.4] to $h_{-K}(u)$, which yields

$$\left\| h_{-K}(u) \right\|_{L^{\Phi_r}(\mathbb{R}^n)} = \sup \left\{ \left| \int_{\mathbb{R}^n} h_{-K}(u(x)) \, v(x) \, dx \right| : v \in L^{\Psi_r}(\mathbb{R}^n), \|v\|'_{L^{\Psi_r}(\mathbb{R}^n)} \le 1 \right\}.$$
(7.5)

(Here we have used the fact that $\mathcal{M}^{\Psi_r} = L^{\Psi_r}(\mathbb{R}^n)$, where $\mathcal{M}^{\Psi_r}$ is the closure of the span of all linear step functions in $L^{\Psi_r}(\mathbb{R}^n)$; this was explained after (6.17), along with a warning about the different notation employed in [33].) Let S_1 and S_2 denote the right-hand sides of (7.4) and (7.5), respectively. From (2.7) with K, x, and y replaced by $-K$, u, and w, respectively, we obtain $S_1 \le S_2$.

For the converse, we can restrict the supremum in (7.5) to nonnegative v. Let $u \in (L^{\Phi_r}(\mathbb{R}^n))^n$ and let $v \in L^{\Psi_r}(\mathbb{R}^n)$, $v \ge 0$, and $\|v\|'_{L^{\Psi_r}(\mathbb{R}^n)} \le 1$. By [36, Theorem 2.7.1], the set of smooth and strictly convex bodies is dense in $\mathcal{K}^n$ with the Hausdorff metric, while the compact convex sets strictly contained in K and strictly containing aK for a fixed $0 < a < 1$ form an open set in $\mathcal{K}^n$. We may therefore choose a sequence (K_m) of smooth and strictly convex bodies converging to K as $m \to \infty$ in the Hausdorff metric and such that

$$\frac{1}{2}K \subset K_m \subset K$$
(7.6)

for $m \in \mathbb{N}$. Let $w_m : \mathbb{R}^n \to S^{n-1}$ be the measurable vector field supplied by Lemma 7.2 with $L = -K_m$. Since $v \in L^{\Psi_r}(\mathbb{R}^n)$, we have $h_{-K_m^\circ}(w_m) \in L^{\Psi_r}(\mathbb{R}^n)$; this is equivalent to $\|w_m\| \in L^{\Psi_r}(\mathbb{R}^n)$ and hence $w_m \in (L^{\Psi_r}(\mathbb{R}^n))^n$. By (7.6), $-K^\circ \subset -K_m^\circ$, so from Lemma 7.2(i), we get

$$\left\| h_{-K^\circ}(w_m) \right\|'_{L^{\Psi_r}(\mathbb{R}^n)} \le \left\| h_{-K_m^\circ}(w_m) \right\|'_{L^{\Psi_r}(\mathbb{R}^n)} = \|v\|'_{L^{\Psi_r}(\mathbb{R}^n)} \le 1.$$

Using (i) and (ii) of Lemma 7.2 with $L = -K_m$, we obtain

$$S_1 \ge \left| \int_{\mathbb{R}^n} u(x) \cdot w_m(x) \, dx \right| = \left| \int_{\mathbb{R}^n} h_{-K_m}(u(x)) \, h_{-K_m^\circ}(w_m(x)) \, dx \right|$$

$$= \left| \int_{\mathbb{R}^n} h_{-K_m}(u(x)) \, v(x) \, dx \right|.$$
(7.7)

By (7.6),

$$0 \leq h_{-K_m}(u(x))\, v(x) \leq h_{-K}(u(x))\, v(x)$$

and the function on the right-hand side is integrable. Taking the limit as $m \to \infty$ in (7.7), the dominated convergence theorem yields

$$S_1 \geq \lim_{m \to \infty} \int_{\mathbb{R}^n} h_{-K_m}(u(x))\, v(x)\, dx = \int_{\mathbb{R}^n} h_{-K}(u(x))\, v(x)\, dx.$$

This proves that $S_1 \geq S_2$ and concludes the proof of (7.4). Now (7.3) follows from (7.4) by the same argument that showed that (6.15) follows from (6.16).

The remainder of the proof of the theorem is a repetition of the last part of the proof of Theorem 6.3, from the point where (6.15) has been established onwards. □

Suppose that $K \subset \mathbb{R}^n$ is an o-symmetric convex body. Then

$$h_K(x) = h_{-K}(x) = \|x\|_{K^\circ}$$

for $x \in \mathbb{R}^n$, where $\| \cdot \|_{K^\circ}$ is the norm for which the unit ball is K°, the polar body of K, defined by (2.2). In this case, (7.1) and (7.2) may be rewritten in the form

$$\int_{\mathbb{R}^n} \Phi\left(\|\nabla T f(x)\|_{K^\circ}\right)\, dx \leq \int_{\mathbb{R}^n} \Phi\left(\|\nabla f(x)\|_{K^\circ}\right)\, dx.$$

Data availability

No data was used for the research described in the article.

Appendix A

The purpose of this appendix is to provide a proof of Proposition 3.7. This follows easily from Lemma 8.1 below, first proved by Crowe, Zweibel, and Rosenbloom [13, Theorem 3] for Schwarz rearrangement without the assumption that $F(s, 0)$ and $F(0, t)$ decrease with $s \geq 0$ and $t \geq 0$, respectively. Variants of Proposition 3.7 are stated for general rearrangements by Brock and Solynin [9, Theorem 3.1] and by Van Schaftingen and Willem [46, Corollary 1], whose approaches to rearrangements differ from ours; see [6, Appendix] for a comparison. Brock and Solynin refer to [13] for a proof, but do not explain why it should apply to general rearrangements, while [46, Corollary 1] is stated with the extra assumption that the function j is even. The proof of [46, Corollary 1] is based on that of [42, Proposition 3.3.9], which does not assume that j is even, or that it is nonnegative, but which requires a considerable amount of preliminary observations and terminology. For this reason, we prefer to follow the argument in [13].

Lemma 8.1. *Let $F : \mathbb{R}^2 \to \mathbb{R}$ be continuous with $F(0,0) = 0$ and such that $F(s,0)$ and $F(0,t)$ decrease with $s \geq 0$ and $t \geq 0$, respectively. Suppose that for all coordinate rectangles $R = [a,b] \times [c,d]$, where $a \leq b$, $c \leq d$,*

$$G(R) = F(b,d) + F(a,c) - F(b,c) - F(a,d) \geq 0. \tag{8.1}$$

Let $T : \mathcal{V}(\mathbb{R}^n) \to \mathcal{V}(\mathbb{R}^n)$ be a rearrangement. If $f, g \in \mathcal{V}(\mathbb{R}^n)$, then

$$\int_{\mathbb{R}^n} F(f(x), g(x)) \, dx \leq \int_{\mathbb{R}^n} F(Tf(x), Tg(x)) \, dx. \tag{8.2}$$

Proof. The function G is additive on coordinate rectangles, that is, if R, S, and $R \cup S$ are non-overlapping coordinate rectangles, then $G(R \cup S) = G(R) + G(S)$. This allows G to be extended to a measure ν on $\mathbb{R}^2$ such that each coordinate rectangle R is ν-measurable and $\nu(R) = G(R)$; see [34, pp. 64–68] (where the union of sets is denoted by $+$).

Let H denote the Heaviside function, i.e., $H(x) = 1$ if $x \geq 0$ and $H(x) = 0$ if $x < 0$. Then for $b, d, s, t \geq 0$,

$$1_{[0,b] \times [0,d]}(s,t) = H(b-s)H(d-t).$$

It follows that

$$\int_0^\infty \int_0^\infty H(b-s)H(d-t) \, d\nu(s,t) = \int_0^\infty \int_0^\infty 1_{[0,b] \times [0,d]}(s,t) \, d\nu(s,t)$$

$$= \nu([0,b] \times [0,d]) = G([0,b] \times [0,d])$$

$$= F(b,d) - F(b,0) - F(0,d).$$

From this we obtain

$$F(b,d) = F(b,0) + F(0,d) + \int_0^\infty \int_0^\infty H(b-s)H(d-t) \, d\nu(s,t).$$

On setting $b = f(x)$ and $d = g(x)$ and integrating, this gives

$$\int_{\mathbb{R}^n} F(f(x), g(x)) \, dx = \int_{\mathbb{R}^n} F(f(x), 0) \, dx + \int_{\mathbb{R}^n} F(0, g(x)) \, dx + I(f,g), \tag{8.3}$$

where by Fubini's theorem,

$$I(f,g) = \int_{\mathbb{R}^n} \int_0^\infty \int_0^\infty H(f(x) - s)H(g(x) - t) \, d\nu(s,t) \, dx$$

$$= \int\limits_0^\infty \int\limits_0^\infty \int\limits_{\mathbb{R}^n} H(f(x) - s) H(g(x) - t) \, dx \, d\nu(s, t). \tag{8.4}$$

Our assumptions on F imply that $-F(r, 0) \geq 0$ increases with $r \geq 0$. We can therefore apply (3.11), with $\varphi(r) = -F(r, 0)$ for $r \geq 0$, to obtain $-F(Tf(x), 0) = T(-F(f(x), 0))$, for each $x \in \mathbb{R}^n$. With this, the layer-cake representation formula, and the equimeasurability of T, we conclude that

$$\int\limits_{\mathbb{R}^n} F(Tf(x), 0) \, dx = -\int\limits_{\mathbb{R}^n} T(-F(f(x), 0)) \, dx$$

$$= -\int\limits_0^\infty \mathcal{H}^n(\{x : T(-F(f(x), 0)) > t\}) \, dt$$

$$= -\int\limits_0^\infty \mathcal{H}^n(\{x : -F(f(x), 0) > t\}) \, dt = \int\limits_{\mathbb{R}^n} F(f(x), 0) \, dx. \tag{8.5}$$

Similarly,

$$\int\limits_{\mathbb{R}^n} F(0, Tg(x)) \, dx = \int\limits_{\mathbb{R}^n} F(0, g(x)) \, dx. \tag{8.6}$$

Since $\lozenge_T$ is monotonic, we have

$$\mathcal{H}^n(\lozenge_T(A \cap B)) \leq \mathcal{H}^n((\lozenge_T A) \cap \lozenge_T B) \tag{8.7}$$

whenever $A, B \in \mathcal{L}^n$. Consequently, using the measure-preserving property of $\lozenge_T$, (8.7), and (3.9), we obtain

$$\int\limits_{\mathbb{R}^n} H(f(x) - s) H(g(x) - t) \, dx = \mathcal{H}^n(\{x : f(x) \geq s\} \cap \{x : g(x) \geq t\})$$

$$= \mathcal{H}^n(\lozenge_T(\{x : f(x) \geq s\} \cap \{x : g(x) \geq t\}))$$

$$\leq \mathcal{H}^n((\lozenge_T\{x : f(x) \geq s\}) \cap \lozenge_T\{x : g(x) \geq t\})$$

$$= \mathcal{H}^n(\{x : Tf(x) \geq s\} \cap \{x : Tg(x) \geq t\})$$

$$= \int\limits_{\mathbb{R}^n} H(Tf(x) - s) H(Tg(x) - t) \, dx.$$

By (8.4), this yields $I(f, g) \leq I(Tf, Tg)$. Then (8.2) follows from (8.3), (8.3) with f and g replaced by Tf and Tg, respectively, (8.5), and (8.6). $\square$

Proof of Proposition 3.7. Let $F(s,t) = -j(s-t)$ for $s, t \geq 0$. Since j is convex, for $r \in \mathbb{R}$ and $s, t \geq 0$, we have

$$j(r) - j(r-s) \leq j(r+t) - j(r-s+t).$$

If $r = b - d$, $s = b - a$, and $t = d - c$, this gives

$$j(b-d) - j(a-d) \leq j(b-c) - j(a-c),$$

yielding (8.1). Moreover, $F(r,0) = -j(r)$ and $F(0,r) = -j(-r)$ both decrease with $r \geq 0$ since $j \geq 0$ and $j(0) = 0$. Applying Lemma 8.1 with this choice of F, we obtain (3.12).

The L^p-contracting property results from taking $j(r) = |r|^p$, $p \geq 1$. $\quad\square$

References

[1] A. Alvino, V. Ferone, G. Trombetti, P.-L. Lions, Convex symmetrization and applications, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 14 (1997) 275–293.
[2] L. Ambrosio, A. Colesanti, E. Villa, Outer Minkowski content for some classes of closed sets and applications to stochastic geometry, Math. Ann. 342 (2008) 727–748.
[3] A. Baernstein II, Symmetrization in Analysis (with D. Drasin and R.S. Laugesen), Cambridge University Press, Cambridge, 2019.
[4] G. Bianchi, R.J. Gardner, P. Gronchi, Symmetrization in geometry, Adv. Math. 306 (2017) 51–88.
[5] G. Bianchi, R.J. Gardner, P. Gronchi, Convergence of symmetrization processes, Indiana Univ. Math. J. 71 (2022) 785–817.
[6] G. Bianchi, R.J. Gardner, P. Gronchi, M. Kiderlen, Rearrangement and polarization, Adv. Math. 374 (2020) 107380, 51 pp.
[7] F. Brock, Continuous Steiner-symmetrization, Math. Nachr. 172 (1995) 25–48.
[8] F. Brock, Continuous rearrangement and symmetry of solutions of elliptic problems, Proc. Indian Acad. Sci. Math. Sci. 110 (2000) 157–204.
[9] F. Brock, A.Y. Solynin, An approach to symmetrization via polarization, Trans. Am. Math. Soc. 352 (2000) 1759–1796.
[10] J.E. Brothers, W.P. Ziemer, Minimal rearrangements of Sobolev functions, J. Reine Angew. Math. 384 (1988) 153–179.
[11] Y.D. Burago, V.A. Zalgaller, Geometric Inequalities, Springer, New York, 1980.
[12] A. Burchard, A. Ferone, On the extremals of the Pólya-Szegő inequality, Indiana Univ. Math. J. 64 (2015) 1447–1463.
[13] J.A. Crowe, J.A. Zweibel, P.C. Rosenbloom, Rearrangements of functions, J. Funct. Anal. 66 (1986) 432–438.
[14] D. Daners, Krahn's proof of the Rayleigh conjecture revisited, Arch. Math. (Basel) 96 (2011) 187–199.
[15] T.K. Donaldsen, N.S. Trudinger, Orlicz–Sobolev spaces and imbedding theorems, J. Funct. Anal. 8 (1971) 52–75.
[16] G.A. Edgar, L. Sucheston, Stopping Times and Directed Processes, Cambridge University Press, New York, 1992.
[17] L.C. Evans, R.F. Gariepy, Measure Theory and Fine Properties of Functions, CRC Press, Boca Raton, FL, 1992.
[18] H. Federer, Geometric Measure Theory, Springer, New York, 1969.
[19] R.J. Gardner, The Brunn-Minkowski inequality, Bull. Am. Math. Soc. 39 (2002) 355–405.
[20] R.J. Gardner, Geometric Tomography, second edition, Cambridge University Press, New York, 2006.
[21] P.M. Gruber, Convex and Discrete Geometry, Springer, Berlin, 2007.
[22] K. Hildén, Symmetrization of functions in Sobolev spaces and the isoperimetric inequality, Manuscr. Math. 18 (1976) 215–235.
[23] B. Kawohl, Rearrangements and Convexity of Level Sets in PDE, Lecture Notes in Mathematics, vol. 1150, Springer, Berlin, 1985.

[24] V.S. Klimov, On the symmetrization of anisotropic integral functionals (Russian), Izv. Vysš. Učebn. Zaved., Mat. 8 (1999) 26–32; translation in Russ. Math. (Izv. VUZ) 43 (1999) 23–29.

[25] S.G. Krantz, H.R. Parks, The Geometry of Domains in Space, Birkhäuser, Boston, MA, 1999.

[26] E.H. Lieb, M. Loss, Analysis, second edition, American Mathematical Society, Providence, RI, 2001.

[27] L. Lussardi, E. Villa, A general formula for the anisotropic outer Minkowski content of a set, Proc. R. Soc. Edinb., Sect. A 146 (2016) 393–413.

[28] F. Maggi, Sets of Finite Perimeter and Geometric Variational Problems. An Introduction to Geometric Measure Theory, Cambridge University Press, Cambridge, 2012.

[29] A. McNabb, Partial Steiner symmetrization and some conduction problems, J. Math. Anal. Appl. 17 (1967) 221–227.

[30] A. Mondino, D. Semola, Polya-Szego inequality and Dirichlet p-spectral gap for non-smooth spaces with Ricci curvature bounded below, J. Math. Pures Appl. 137 (2020) 238–274.

[31] W.F. Pfeffer, Derivation and Integration, Cambridge University Press, New York, 2001.

[32] G. Pólya, G. Szegő, Isoperimetric Inequalities in Mathematical Physics, Princeton University Press, Princeton, NJ, 1951.

[33] M.M. Rao, Z.D. Ren, Theory of Orlicz Spaces, Marcel Dekker, Inc., New York, 1991.

[34] S. Saks, Theory of the Integral, second revised edition, Dover Publications Inc., New York, 1964.

[35] J. Sarvas, Symmetrization of condensers in n-space, Ann. Acad. Sci. Fenn., Ser. A I Math. (522) (1972), 44 pp.

[36] R. Schneider, Convex Bodies: The Brunn-Minkowski Theory, second edition, Cambridge University Press, Cambridge, 2014.

[37] A.Y. Solynin, Continuous symmetrization via polarization, Algebra Anal. 24 (2012) 157–222; translation in St. Petersburg Math. J. 24 (2013) 117–166.

[38] E. Sperner Jr., Symmetrisierung für Funktionen mehrerer reeler Variablen, Manuscr. Math. 11 (1974) 159–170.

[39] G. Talenti, Boundedness of minimizers, Hokkaido Math. J. 19 (1990) 259–279.

[40] G. Talenti, Inequalities in rearrangement invariant function spaces, in: Nonlinear Analysis, Function Spaces and Applications, vol. 5, Prague, 1994, Prometheus, Prague, 1994, pp. 177–230.

[41] G. Talenti, The art of rearranging, Milan J. Math. 84 (2016) 105–157.

[42] J. Van Schaftingen, Réarrangements et problèmes elliptiques non linéaires, Thesis (Ph.D.), Université Catholique de Louvain, 2002, 94 pp.

[43] J. Van Schaftingen, Universal approximation of symmetrizations by polarizations, Proc. Am. Math. Soc. 134 (2006) 177–186.

[44] J. Van Schaftingen, Anisotropic symmetrization, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 23 (2006) 539–565.

[45] J. Van Schaftingen, Explicit approximation of the symmetric rearrangement by polarizations, Arch. Math. 93 (2009) 181–190.

[46] J. Van Schaftingen, M. Willem, Set transformations, symmetrizations and isoperimetric inequalities, in: Nonlinear Analysis and Applications to Physical Sciences, Springer Italia, Milan, 2004, pp. 135–152.

[47] A.C. Zaanen, Riesz Spaces II, North-Holland Publishing Company, New York, 1983.

[48] G. Zhang, The affine Sobolev inequality, J. Differ. Geom. 53 (1999) 183–202.