A GENERAL REARRANGEMENT INEQUALITY

CRISTINA DRAGHICI

(Communicated by David Preiss)

ABSTRACT. We prove a general rearrangement inequality for multiple integrals, using polarization. We introduce a special class of kernels for which the product inequality holds, and then we prove that it also holds when the product is replaced by a so-called function $AL_m$.

1. INTRODUCTION

Let $f : \Omega \to \mathbb{R}^+$ be a nonnegative measurable function defined on a measure space $(\Omega, \mu)$, which satisfies the condition
\begin{equation}
\mu(\{f > t\}) < \infty, \quad \forall \, t > 0
\end{equation}
where we write
\[ \{f > t\} := \{x \in \Omega : f(x) > t\}. \]
Its distribution function $\lambda_f$ is defined to be
\[ \lambda_f(t) = \mu(\{f > t\}), \quad t \in [0, \infty). \]
Two functions $f$ and $g$ are said to be equimeasurable if they have the same distribution function and we write $f \sim g$. Functions which are equimeasurable are also said to be rearrangements of each other.

Let $X$ denote either $\mathbb{R}^n$, $S^n = \{x \in \mathbb{R}^{n+1} : |x| = 1\}$, or the $n$-dimensional hyperbolic space $\mathbb{H}^n$. We equip $X$ with its geodesic distance $d$ and canonical measure $\mu$, and we fix an origin $o$ in $X$. For $\mathbb{R}^n$, $d$ is the Euclidean distance, $\mu$ is the Lebesgue measure, and $o$ is $(0, \ldots, 0)$. For $S^n$, $d$ is the great circle distance on the sphere, $\mu$ is the surface area measure and $o$ is $(1, 0, \ldots, 0)$. We take as model for $\mathbb{H}^n$ the open unit ball $\{x \in \mathbb{R}^n : |x| < 1\}$ endowed with the distance element $ds = \frac{2}{1-|x|^2} |dx|$; $\mu$ is the volume measure associated to $ds$, and the origin $o$ is $(0, \ldots, 0)$.

Given $f : X \to \mathbb{R}^+$ satisfying (1.1) we define its symmetric decreasing rearrangement $f^\sharp : X \to \mathbb{R}^+$ as follows:
\[ f^\sharp(x) = \inf \{t : \lambda_f(t) \leq \mu(B(d(x)))\}, \]
where $B(d(x))$ is the ball centered at $o$ and with radius $d(x) = d(x, o)$. It follows easily that $f^\sharp$ is constant on each sphere centered at $o$ and decreases as $d(x)$ increases. Moreover, $f^\sharp \sim f$. See, e.g., [3].

We define now a very simple rearrangement, called polarization (see [2, 3, 6, 7]).
Let $\mathcal{H}(\mathbb{R}^n)$ be the collection of all $(n-1)$-dimensional affine subspaces $H$ of $\mathbb{R}^n$ such that $0 \notin H$, $\mathcal{H}(S^n)$ the collection of all intersections of $S^n$ with hyperplanes through the origin in $\mathbb{R}^{n+1}$ which do not contain $0$ and $\mathcal{H}(\mathbb{H}^n)$ the collection of all images under the group of hyperbolic motions of the hyperbolic $n-1$ plane $\{x \in \mathbb{R}^n : |x| < 1, x_n = 0\} \subset \mathbb{H}^n$ which do not contain the origin $0$. For $H \in \mathcal{H}(X)$, let $\bar{x}$ denote the reflection in $H$ (i.e. $\bar{x}$ is the reflection of $x$ in $H$), and let $H^+$ and $H^-$ be the two components of $X \setminus H$, such that $0 \in H^+$.

For $f : X \to \mathbb{R}$, we define a function $f^H : X \to \mathbb{R}$ by $f^H(x) = f(x)$ if $x \in H$, and

$$f^H(x) = \begin{cases} \max(f(x), f(\bar{x})) & \text{if } x \in H^+, \\ \min(f(x), f(\bar{x})) & \text{if } x \in H^-. \end{cases}$$

We explore, in what follows, some of the rich history of rearrangement inequalities. The most basic rearrangement inequality is due to Hardy and Littlewood [13]:

$$\int_X f(x)g(x) \, d\mu(x) \leq \int_X f^H(x)g^H(x) \, d\mu(x),$$

with $f, g : X \to \mathbb{R}^+$ satisfying (1.2). One can show that the inequality also holds for polarizations.

We will prove inequalities from my thesis [10], of the form

$$\int_X \Psi(f_1(x_1), \ldots, f_m(x_m))K(x_1, \ldots, x_m) \, dx$$

$$\leq \int_X \Psi(f_1^H(x_1), \ldots, f_m^H(x_m))K(x_1, \ldots, x_m) \, dx,$$

where $dx = d\mu(x_1) \ldots d\mu(x_m)$, which generalize the Hardy-Littlewood inequality. See also [11].

The theory of rearrangement inequalities is well-developed in $\mathbb{R}^n$. The following inequality, known as Riesz-Sobolev, was proved for $n = 1$ by F. Riesz in 1930 and for $n \geq 1$ by S. L. Sobolev in 1938:

$$\int_{\mathbb{R}^n} f(x)g(x)h(x-y) \, dx \, dy \leq \int_{\mathbb{R}^n} f^H(x)g^H(y)h^H(x-y) \, dx \, dy.$$

The functions $f$, $g$, and $h$ are nonnegative.

A general rearrangement inequality involving nonnegative functions is due to H. J. Brascamp, E. H. Lieb, and J. M. Luttinger [5]:

$$(1.3) \quad \int_{\mathbb{R}^n} \prod_{j=1}^k f_j(\sum_{m=1}^p a_{jm}x_m) \, dx \leq \int_{\mathbb{R}^n} \prod_{j=1}^k f_j^H(\sum_{m=1}^p a_{jm}x_m) \, dx$$

where $a_{jm}$ are constants, $dx = dx_1 \ldots dx_p$, and each $x_i \in \mathbb{R}^n$. When $p = 2$, $k = 3$ and $a_{jm}$ are suitably chosen, (1.3) implies the Riesz-Sobolev inequality.

The proof of (1.3) uses the Brunn-Minkowski inequality [15], which states that if $K$ and $L$ are convex bodies (compact convex sets with nonempty interiors) in $\mathbb{R}^k$ and $0 < \lambda < 1$, then

$$V((1-\lambda)K + \lambda L)^{1/k} \geq (1 - \lambda)V(K)^{1/k} + \lambda V(L)^{1/k}.$$

$V$ denotes the $k$-dimensional volume and $+$ is the vector sum of $K$ and $L$. Further generalizations of (1.3) were given by M. Christ [9] and R. E. Pfiefer [16].
The proofs of the rearrangement inequalities on $\mathbb{R}^n$ require the product structure $\mathbb{R}^n = \mathbb{R} \times \ldots \times \mathbb{R}$, as well as the Brunn-Minkowski inequality. These tools are not available on $S^n$ or $H^n$, and for this reason complete analogues of (1.3) for $S^n$ and $H^n$ are not known. To get rearrangement inequalities on $X$, polarization proved to be an effective tool. A. Baernstein and B. A. Taylor [2] used it to derive inequalities for $S^n$, and W. Beckner [4] noticed that it can also be used for $H^n$.

Baernstein and Taylor showed that the symmetric decreasing rearrangement of a function can be approximated through a sequence of polarizations [2]. The polarization is easier to handle and if an inequality holds for polarizations in all hyperplanes the rest are fixed. This can be seen, for example, by taking the Baernstein-Taylor inequality:

\[
\int_{S^n} \int_{S^n} f(x)g(y)K(d(x,y)) \, d\mu(x) \, d\mu(y) \leq \int_{S^n} \int_{S^n} f^+(x)g^+(y)K(d(x,y)) \, d\mu(x) \, d\mu(y),
\]

for $K : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ a decreasing function. This inequality was proved for $f^H$ and $g^H$ in place of $f^+$ and $g^+$, for every $H \in \mathcal{H}(X)$, and then they used the approximation argument to deduce it for $f^+$ and $g^+$.

We introduce now a special class of functions of particular interest for our result. For $\mathbb{R}^n_+ = (x_1^+, \ldots, x_m^+)$, $\mathbb{R}^m_+ = (x_1^+, \ldots, x_m^+)$ we write $\mathbf{e}_i^+ = (\mathbf{e}_1^+, \ldots, \mathbf{e}_i^+, \ldots, \mathbf{e}_m^+)$ if $x_i^+ = x_i^+$, for every $i = 1, \ldots, m$.

**Definition 1.1.** Let $\Psi : \mathbb{R}^m_+ \rightarrow \mathbb{R}_+$ be a continuous function such that

\[
\Psi_{i|x_i=0} = 0, \text{ for every } i = 1, \ldots, m.
\]

We say that $\Psi \in AL_m(\mathbb{R}_+^m)$ if, for every $\mathbf{e}_1^+, \mathbf{e}_2^+ \in \mathbb{R}_+^m$, such that $\mathbf{e}_1^+ \leq \mathbf{e}_2^+$, the following condition holds:

\[
\sum_{i_1, \ldots, i_m \in \{1, 2\}} (-1)^{i_1+\ldots+i_m} \Psi(x_1^{i_1}, \ldots, x_m^{i_m}) \geq 0.
\]  

(1.4)

If $\Psi \in C^m(\mathbb{R}_+^m)$, then (1.4) is equivalent to $\frac{\partial^m \Psi}{\partial x_{1}\ldots \partial x_{m}} \geq 0$.

Consider now a family of $m$ positive random variables $X_1, \ldots, X_m$ on some probability space $(\Omega, \mathcal{P})$, and define

\[
\Psi(x_1, \ldots, x_m) = P(X_1 \leq x_1, \ldots, X_m \leq x_m),
\]

their joint distribution function. Then $\Psi$ satisfies (1.4).

One can show that $\Psi \in AL_m(\mathbb{R}_+^m)$ implies $\Psi$ is increasing in each argument, when the rest are fixed. This can be seen, for example, by taking $\mathbf{e}_1 = (x_1, 0, \ldots, 0)$ and $\mathbf{e}_2 = (y_1, x_2, \ldots, x_m)$ with $\mathbf{e}_1 \leq \mathbf{e}_2$, and using (1.4) and the fact that $\Psi$ vanishes on the boundary.

We say that a continuous function $\Psi : \mathbb{R}_+^m \rightarrow \mathbb{R}_+$ is in $AL_2(\mathbb{R}_+^m)$ if for every fixed $m-2$ variables $\Psi$ is in $AL_2(\mathbb{R}_+^2)$ in the remaining two arguments [6]. This class of functions was studied by Almgren and Lieb [1].

Brock [6] proved an extension of the Hardy-Littlewood inequality: If $\Psi \in AL_2(\mathbb{R}_+^m)$, then

\[
\int_X \Psi(f_1(x), \ldots, f_m(x)) \, d\mu(x) \leq \int_X \Psi(f_1^+(x), \ldots, f_m^+(x)) \, d\mu(x),
\]

for every $f_1, \ldots, f_m : X \rightarrow \mathbb{R}_+$ satisfying (1.1). As before, $X$ is any of the spaces $\mathbb{R}^n$, $S^n$, or $H^n$. 
In the same spirit, Morpurgo [14] proved the following inequality: For $K_{ij} : \mathbb{R}_+ \to \mathbb{R}_+$, $1 \leq i < j \leq m$, decreasing functions and $\Psi_0 : \mathbb{R} \to \mathbb{R}$ a continuous convex function:

$$\int_{X^m} \prod_{i<j} K_{ij}(d(x_i, x_j))\Psi_0\left(\sum_{l=1}^m f_l(x_l)\right) \, dx$$

$$\leq \int_{X^m} \prod_{i<j} K_{ij}(d(x_i, x_j))\Psi_0\left(\sum_{l=1}^m f_l^2(x_l)\right) \, dx,$$

where $dx = dx_1 \ldots dx_m$. A similar result was proved by Burchard and Schmuckenschläger, [5]. Both Morpurgo’s result (1.6) and Brock’s result (1.5) were first proved for polarization, then obtained for the symmetric decreasing rearrangement by the approximation argument. Also, Morpurgo’s inequality contains the Baernstein-Taylor inequality as a special case.

Our result gives a larger class of kernels $K$ for which the polarization inequality holds in Morpurgo’s theorem, when $\Psi_0(\sum_{l=1}^m f_l(x_l))$ is replaced by $f_1(x_1) \cdots f_m(x_m)$ and hence by $\Psi(f_1(x_1), \ldots, f_m(x_m))$, where $\Psi \in AL_m(\mathbb{R}^+_m)$. We begin with a discrete result.

2. A DISCRETE REARRANGEMENT INEQUALITY

Let $Q = \{1, 2\}$ be a set with two points, and let $K_0 : Q^m \to \mathbb{R}_+$ be a nonnegative function. We think of $Q^m$ as the vertices of a cube in $\mathbb{R}^m$. Take $k \in \{1, \ldots, m\}$ and $S \subset \{1, \ldots, m\}$, with $|S| = k$. We write $S = \{i_1, \ldots, i_k\}$, with $i_1 < \ldots < i_k$. For every $y$ and $y = (y_1, \ldots, y_k) \in Q^k$, let

$$A(y, S) = \{x \in Q^m : x_{i_1} = y_1, \ldots, x_{i_k} = y_k\}.$$ 

We can think of $A(y, S)$ as an $(m - k)$-dimensional subcube of $Q^m$.

Let $1_k = (1, \ldots, 1) \in Q^k$. Then $A(1_k, S)$ is the set of points $x$ in $Q^m$ for which $x_{i_1} = \ldots = x_{i_k} = 1$.

We define

$$K_0(y, S) = \sum_{x \in A(y, S)} K_0(x), \quad y \in Q^k.$$ 

**Definition 2.1.** A function $K_0 : Q^m \to \mathbb{R}_+$ is called a “good” cube kernel if

$$K_0(y, S) \leq K_0(1_k, S),$$

for every $y \in Q^k$, every $S \subset \{1, \ldots, m\}$ with $|S| = k$, and every $k \in \{1, \ldots, m\}$.

Let us consider some particular cases. For $k = m$, we have $A(y, S) = \{y\}$ and $S = \{1, \ldots, m\}$. Then (2.2) says that

$$K_0(y) \leq K_0(1_m), \quad \forall y \in Q^m.$$ 

For $k = m - 1$ and $S = \{1, \ldots, m\} \setminus \{j\}$, (2.2) asserts that

$$K_0(y_1, \ldots, y_{j-1}, 1, y_{j+1}, \ldots, y_m) + K_0(y_1, \ldots, y_{j-1}, 2, \ldots, y_m) \leq K_0(1, 1, \ldots, 1) + K_0(1, 1, \ldots, 2, \ldots, 1), \quad \forall y \in Q^{m-1}.$$ 

One may picture the left-hand side of (2.3) as the sum of $K_0$ over an edge of the $m$-cube $Q^m$, and then (2.3) says that among all parallel edges, the maximal edge sum is achieved when the edge contains $(1, \ldots, 1)$.
Likewise, for every $S$, with $|S| = k$, $A(y, S)$ represents the sum of $K_0$ over the vertices of an $m - k$ subcube of $Q^m$ and (2.2) requires that among all parallel $m - k$ subcubes, the maximal sum is attained when the subcube contains $(1, \ldots, 1)$.

For $g : Q \to \mathbb{R}_+$ we define $g^*$, its decreasing rearrangement, as follows:

$$g^*(1) = \max(g(1), g(2)) \quad \text{and} \quad g^*(2) = \min(g(1), g(2)).$$

The following lemma is immediate from the definition of $K_0$.

**Lemma 2.2.** Let $K_0 : Q^m \to \mathbb{R}_+$ be a nonnegative function. Then, the following inequality is true:

$$\int_{Q^m} g_1(\varepsilon_1) \cdots g_m(\varepsilon_m)K_0(\varepsilon_1, \ldots, \varepsilon_m)d\varepsilon_1 \cdots d\varepsilon_m \leq \int_{Q^m} g_1^*(\varepsilon_1) \cdots g_m^*(\varepsilon_m)K_0(\varepsilon_1, \ldots, \varepsilon_m)d\varepsilon_1 \cdots d\varepsilon_m$$

for every $g_1, \ldots, g_m : Q \to \mathbb{R}_+$, if and only if $K_0$ is a good cube kernel.

The proof is achieved by considering each $g_i$ to be the characteristic function of a point in $Q$. This result implies the following:

**Proposition 2.3.** Let $f_i : Q \to \mathbb{R}$ be $m$ nonnegative functions, let $\Psi \in AL_m(\mathbb{R}_+^m)$ and let $K_0 : Q^m \to \mathbb{R}_+$ be a good cube kernel. We define

$$I[g_1, \ldots, g_m] = \int_{Q^m} \Psi(\varepsilon_1, \ldots, \varepsilon_m)K_0(\varepsilon_1, \ldots, \varepsilon_m)d\varepsilon_1 \cdots d\varepsilon_m.$$

Sometimes we write $I(g_1, \ldots, g_m; \Psi, K)$ if we want to emphasize $\Psi$ and $K$. Then

$$I[g_1, \ldots, g_m] \leq I[g_1^*, \ldots, g_m^*].$$

Conversely, if for some $K_0 : Q^m \to \mathbb{R}_+$ we assume that the inequality holds for every $\Psi \in AL_m(\mathbb{R}_+^m)$ and all choices of $g_i$, then $K_0$ is a good cube kernel.

**Proof.** For every $a > 0$ we define the set

$$C(a) = \mathbb{R}_+^m \setminus \{ (x_1, \ldots, x_m) : x_i > a, \forall i = 1, \ldots, m \}.$$

We assume first that $\Psi \in C^m$ ($\Psi$ is $m$ times continuously differentiable) and that it vanishes on $C(a)$, for some $a$. We already have the inequality for the case when $\Psi$ is the product function, $\Psi(y_1, \ldots, y_m) = y_1 \cdots y_m$. So, we need to express $\Psi$ in terms of a product. To condense notation, let

$$\Phi(s) = \frac{\partial^m \Psi}{\partial s_1 \cdots \partial s_m}(s),$$

d$\varepsilon = d\varepsilon_1 \cdots d\varepsilon_m$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)$, $ds = ds_1 \cdots ds_m$, $s = (s_1, \ldots, s_m)$. Then $\Phi \geq 0$ since $\Psi \in AL_m$.

Since all the boundary terms are zero, we can write

$$\Psi(y_1, \ldots, y_m) = \int_0^{y_1} \int_0^{y_2} \cdots \int_0^{y_m} \Phi(s)ds_1 \cdots ds_m$$

$$= \int_0^\infty \cdots \int_0^\infty \Phi(s)\chi_{\{s_1 < y_1\}}(s_1) \cdots \chi_{\{s_m < y_m\}}(s_m)ds.$$

Setting $y_i = g_i(\varepsilon_i)$, multiplying by $K_0(\varepsilon)$ and integrating, we obtain

$$I[g_1, \ldots, g_m] = \int_{Q^m} \int_{(0, \infty)^m} \Phi(s)\chi_{\{g_1 > s_1\}}(s_1) \cdots \chi_{\{g_m > s_m\}}(s_m)K_0(\varepsilon)ds d\varepsilon.$$
Now reverse the order of integration. For fixed \( s_1, \ldots, s_m, \) positive,
\[
\chi_{\{s_1 > s_1\}} = \chi_{\{s_1 > s_1\}},
\]as functions on \( Q, \)
and we can apply Lemma 2.2 to show that
\[
\int_{Q^m} \chi_{\{s_1 > s_1\}}(\varepsilon_1) \cdots \chi_{\{s_m > s_m\}}(\varepsilon_m) K_0(\varepsilon) \, d\varepsilon 
\leq 
\int_{Q^m} \chi_{\{s_1 > s_1\}}(\varepsilon_1) \cdots \chi_{\{s_m > s_m\}}(\varepsilon_m) K_0(\varepsilon) \, d\varepsilon.
\]
It follows that \( I(g_1, \ldots, g_m, \Psi, K_0) \leq I(g_1^*, \ldots, g_m^*, \Psi, K_0) \) when \( \Psi \in C^m \) and \( \Psi \)
vanishes on \( C(a). \)

If \( \Psi \) vanishes on \( C(a) \) and is not smooth, then we can approximate it by smooth
functions by convolving it with \( C_\infty \) for \( \Psi \in C_\infty \).

Let \( G : \mathbb{R}^m \to \mathbb{R}^+ \) be a \( C^\infty \) function such that \( \int_{\mathbb{R}^m} G = 1 \). Define \( G_\delta(x) = \delta^{-m} G(x/\delta) \) and
\[
\Psi_\delta(x) = (\Psi * G_\delta)(x) = \int_{\mathbb{R}^m} \Psi(y) G_\delta(x - y) \, dy.
\]
Then \( \Psi_\delta \) is smooth and in \( AL_m(\mathbb{R}^m) \), and since \( \Psi = 0 \) on \( C(a) \), \( \Psi_\delta \) is zero on \( C(a_\delta) \), when \( \delta \) is small enough, for some \( a_\delta > 0 \). Thus we can apply our inequality for \( \Psi_\delta \) to get
\[
I(g_1, \ldots, g_m, \Psi_\delta, K_0) \leq I(g_1^*, \ldots, g_m^*, \Psi_\delta, K_0).
\]

Since the \( \Psi_\delta \) are uniformly close to \( \Psi \) on the discrete set \( f_1(Q) \times \cdots \times f_m(Q), \) we let \( \delta \to 0 \) to get
\[
I(g_1, \ldots, g_m, \Psi, K_0) \leq I(g_1^*, \ldots, g_m^*, \Psi, K_0),
\]
for \( \Psi \in AL_m(\mathbb{R}^m) \) and \( \Psi = 0 \) on \( C(a) \).

For general \( \Psi \in AL_m(\mathbb{R}^m), \) let
\[
\Psi_k(x) = \Psi(x - (1/k)1_m), \quad \text{with} \quad 1_m = (1, \ldots, 1) \in \mathbb{R}^m, \quad x \in \mathbb{R}^m \setminus C(1/k).
\]
Then we can define \( \Psi_k \) to be zero on \( C(1/k) \). The new \( \Psi_k \)'s are in \( AL_m(\mathbb{R}^m) \).
Since \( \Psi_k \uparrow \Psi \) on \( R^m, \) and since
\[
I(g_1, \ldots, g_m, \Psi_k, K_0) \leq I(g_1^*, \ldots, g_m^*, \Psi_k, K_0),
\]
we can pass to the limit under the integral sign, and we obtain
\[
I(g_1, \ldots, g_m, \Psi, K_0) \leq I(g_1^*, \ldots, g_m^*, \Psi, K_0).
\]
The converse statement of Proposition 2.3 follows from Lemma 2.2 by taking
\( \Psi(y_1, \ldots, y_m) = y_1 \cdots y_m. \)

3. MAIN RESULT

We return now to analysis on \( X. \) We consider a measurable function \( K : \mathbb{X}^m \to \mathbb{R}^+ \), and take \( H \in \mathcal{H}(X). \) For each \( (x_1, \ldots, x_m) \in (\mathbb{H}^+)^m, \) we define \( K_0 : Q^m \to \mathbb{R}^+ \)
as follows:
\[
K_0(\varepsilon_1, \ldots, \varepsilon_m) := K(y_1, \ldots, y_m), \quad \text{with} \quad y_i = \begin{cases} x_i & \text{if } \varepsilon_i = 1, \\ \bar{x}_i & \text{if } \varepsilon_i = 2. \end{cases}
\]
As usual, \( \bar{x}_i \) is the reflection of \( x_i \) in the hyperplane \( H. \) We note that \( K_0 \) depends
on \( (x_1, \ldots, x_m). \)
Definition 3.1. We call \( K : X^m \to \mathbb{R}_+ \) a good polarization kernel if each \( K_0 \) is a good cube kernel for almost every \((x_1, \ldots, x_m) \in (H^+)^m\) and every \( H \in \mathcal{H}(X)\) with \( 0 \in H^+\).

Here is an equivalent definition of a good polarization kernel, in terms of sums. Consider \( k \in \{1, \ldots, m\} \) and \( S \subset \{1, \ldots, m\}, S = \{i_1, \ldots, i_k\} \) with \( i_1 < \ldots < i_k \), and \( H \in \mathcal{H}(X) \) with \((x_1, \ldots, x_m) \in (H^+)^m\) being fixed. For every \( y = (y_1, \ldots, y_k) \in \{x_{i_1}, \bar{x}_{i_1}\} \times \cdots \times \{x_{i_k}, \bar{x}_{i_k}\} \), let

\[
A_H(y, S) = \{x \in \{x_1, \bar{x}_1\} \times \cdots \times \{x_m, \bar{x}_m\} : x_{i_1} = y_1, \ldots, x_{i_k} = y_k\}.
\]

We define

\[
K(y, S) = \sum_{x \in A_H(y, S)} K(x).
\]

Let \( y^* = (x_1, \ldots, x_k) \). Then we see that our definition is equivalent to the condition \( K(y, S) \leq K(y^*, S) \), for every \( S \subset \{1, \ldots, m\}, H \in \mathcal{H}(X) \) and a.e. \((x_1, \ldots, x_m) \in (H^+)^m\).

Now, consider \( \Psi : \mathbb{R}_+^m \to \mathbb{R}_+ \) and \( f_i : X \to \mathbb{R}_+ \), \( 1 \leq i \leq m \). Define

\[
I(f_1, \ldots, f_m, K, \Psi) = \int_{X^m} \Psi(f_1(x_1), \ldots, f_m(x_m)) K(x_1, \ldots, x_m) dx_1 \cdots dx_m.
\]

If we split each integral \( \int_X = \int_{H^+} + \int_{H^-} \) and use reflection in the second integral and collect terms, we get in (3.2):

\[
I(f_1, \ldots, f_m, K, \Psi) = \int_{(H^+)^m} I(g_1, \ldots, g_m, \Psi, K_0) dx_1 \cdots dx_m.
\]

Here the \( g_i : Q \to \mathbb{R}_+ \) depend on \( x_i \) and are defined as follows:

\[
g_i(1) = f_i(x_i), \quad g_i(2) = f_i(\bar{x}_i), \quad x_i \in X.
\]

Theorem 3.2. Let \( X = \mathbb{R}^n, \mathbb{S}^n \) or \( \mathbb{H}^n \), let \( f_i : X \to \mathbb{R}_+ \) be nonnegative functions, let \( \Psi \in AL_m(\mathbb{R}_+^m) \), and let \( K \) be a good polarization kernel. We define

\[
I[f_1, \ldots, f_m] = \int_{X^m} \Psi(f_1(x_1), \ldots, f_m(x_m)) K(x_1, \ldots, x_m) dx_1 \cdots dx_m.
\]

Then, the following inequalities hold:

\[
I[f_1, \ldots, f_m] \leq I[f_1^H, \ldots, f_m^H] \quad \forall H \in \mathcal{H}(X)
\]

and

\[
I[f_1, \ldots, f_m] \leq I[f_1^t, \ldots, f_m^t].
\]

If \( K : X^m \to \mathbb{R}_+ \) is such that the polarization inequality (3.4) holds for every \( H \in \mathcal{H}(X) \) and every \( \Psi \in AL_m(\mathbb{R}_+^m) \), then \( K \) is a good polarization kernel.

Proof. The polarization inequality in Theorem 3.2 follows from Proposition 2.3 applied to the integrand in (3.3), using the fact that if \( g_1, \ldots, g_m \) correspond to \( f_1, \ldots, f_m \) for a certain \( (x_1, \ldots, x_m) \), then \( g_1^1, \ldots, g_m^1 \) correspond to \( f_1^H, \ldots, f_m^H \). Also, inequality (3.5) can be deduced using an approximation argument similar to the one presented in [2].

Conversely, to prove that \( K \) must be a good polarization kernel if (3.7) always holds, let \( \Psi(y_1, \ldots, y_m) = y_1 \cdots y_m \). Let us fix \( H \in \mathcal{H}(X) \) and consider
(x_1, \ldots, x_m) \in (H^+)^m$. For $S = \{i_1, \ldots, i_k\}$ and $y = (y_1, \ldots, y_k) \in \{x_{i_1}, \bar{x}_{i_1}\} \times \ldots \times \{x_{i_k}, \bar{x}_{i_k}\}$, define

$$f_{i_1} = (1/c_\varepsilon)\chi_{B(y_1, \varepsilon)}, \quad \ldots, \quad f_{i_k} = (1/c_\varepsilon)\chi_{B(y_k, \varepsilon)}$$

and $f_j = (1/c_\varepsilon)\chi_{B(x_{i_j}, \varepsilon)} + (1/c_\varepsilon)\chi_{B(\bar{x}_{i_j}, \varepsilon)}$, for $j \in S^c$. Here $c_\varepsilon$ represents the measure of any ball of radius $\varepsilon$ in $X$.

For $\varepsilon$ sufficiently small,

$$f_{i_1}^H = (1/c_\varepsilon)\chi_{B(x_{i_1}, \varepsilon)}, \quad \ldots, \quad f_{i_k}^H = (1/c_\varepsilon)\chi_{B(x_{i_k}, \varepsilon)}$$

and $f_j^H$ for $j \in S^c$.

Without loss of generality, we can assume that $K$ is integrable ($K \in L^1(X^m)$). Since

$$I(f_1, \ldots, f_m) \leq I(f_1^H, \ldots, f_m^H),$$

and by passing to the limit as $\varepsilon \to 0$ we get that LHS $\to K(y, S)$, RHS $\to K(y^*, S)$ for a.e. $(x_1, \ldots, x_m) \in (H^+)^m$ (see [12, p. 93]), it follows that

$$K(y, S) \leq K(y^*, S), \quad \text{for a.e. } (x_1, \ldots, x_m) \in (H^+)^m.$$

Since $H \in \mathcal{H}(X)$ was arbitrarily chosen, this proves that $K$ is a good polarization kernel.

We conclude this section with a few remarks. Firstly, when $m = 2$, $K(x, y) = K_1(d(x, y))$, with $K_1$ a decreasing function, is a good polarization kernel.

Kernels of the form $K(x_1, \ldots, x_m) = k(|x_1|^\alpha_1 + |x_2|^\alpha_2 + \ldots + |x_m|^\alpha_m)$, where $\alpha_1, \ldots, \alpha_m \geq 0$ are positive real numbers and $k$ is a decreasing function of one variable, are good polarization kernels. Here $|x|$ denotes the distance to the origin, $d(x, 0)$. To see this, let $H \in \mathcal{H}(X)$ and $(x_1, \ldots, x_m) \in (H^+)^m$. With the notation in (3.1), $K(x) \leq K(z)$ for $x \in A_H(y, S)$ and $z \in A_H(y^*, S)$ such that $x_i = z_i$ for every $i \in S^c$, since any component of $z$ has distance to the origin less than or equal to the distance to the origin of the corresponding component of $x$ and since $k$ is decreasing. There is a one-to-one correspondence between elements in $A_H(y, S)$ and $A_H(y^*, S)$, and thus

$$K(y, S) = \sum_{x \in A_H(y, S)} K(x) \leq \sum_{x \in A_H(y^*, S)} K(x) = K(y^*, S),$$

for every $H \in \mathcal{H}(X)$, and $(x_1, \ldots, x_m) \in (H^+)^m$ and $y$. This proves that $K$ is a good polarization kernel.

More generally, $K(x_1, \ldots, x_m)$, where $K$ is symmetric decreasing in each argument, is a good polarization kernel.

We show now that Morpurgo’s kernels are good polarization kernels. Let $K(x_1, \ldots, x_m) = \prod_{i<j} K_{ij}(d(x_i, x_j))$ where $x_1, \ldots, x_m \in X$, and $K_{ij} : \mathbb{R}_+ \to \mathbb{R}_+$ are decreasing. We have $I(f_1, \ldots, f_m, \Psi, K) \leq I(f_1^H, \ldots, f_m^H, \Psi, K)$ for

$$\Psi(y_1, \ldots, y_m) = y_1 \cdots y_m \in AL_m(\mathbb{R}_+^m)$$

and all nonnegative $f_i$’s and $H \in \mathcal{H}(X)$. This follows from Morpurgo’s result (1.6), which was proved using polarization, and by taking $\Psi_0(y) = e^{y}$ and using the fact that $e^y \leq (e^y)^H$. The proof of the necessity part of Theorem 3.2 required only the case when $\Psi$ is the product function. Thus, we can conclude that $K$ is a good polarization kernel.
The author acknowledges that the results presented in this paper are part of her dissertation written under the direction of Albert Baernstein II.

REFERENCES


Department of Mathematics, Western Michigan University, Kalamazoo, Michigan 49008

E-mail address: cristina.draghici@wmich.edu