ISRAEL JOURNAL OF MATHEMATICS 83 (1993), 97–127

POTENTIAL INEQUALITY*

BY

MURALI RAO

Department of Mathematics
University of Florida, Gainesville, FL 32611, USA

AND

HRVOJE ŠIKIĆ

Department of Mathematics
University of Florida, Gainesville, FL 32611, USA
and
Department of Mathematics
University of Zagreb, 41000 Zagreb, Croatia

ABSTRACT

We prove a general inequality for kernels satisfying the maximum principle. This is then used to derive a sufficient condition for the kernel to define a continuous map of Lebesgue spaces. Exactly this condition happens to be necessary and sufficient for the validity of Hardy's inequality with weights in one dimension. Some applications indicating the unifying nature of the potential inequality are given.

Introduction

In this article we describe a powerful potential theoretic inequality, which generates Hardy's inequality and many other inequalities as special cases. Our approach is very general and has potential for much further development. Some of these directions are indicated in the paper.

* A part of this paper was presented at the 865th AMS meeting, March 1991, Tampa, Florida.

Received July 30, 1991 and in revised form November 28, 1992

In 1920, Hardy [8] discovered an inequality while attempting to simplify the then-known proofs of Hilbert's inequality ([9], chapter IX). Since its inception, this inequality has found many applications: almost all books on PDE contain it. Many authors have extended and generalized this inequality. A major contribution was made by Talenti [19] in 1966, and later developed by Talenti [20], Tomaselli [21], Muckenhoupt [15], Bradley [7], etc. These developments concerned Hardy's inequality with weights. Later contributions are due to Kufner–Triebel [12], Maz'ja [14], and most recently due to Stepanov [18] (1987), and Ariño–Muckenhoupt [2] (1990).

Most of these works are one-dimensional in nature and rely on the special kernel involved, and the potential theoretic nature of the inequalities has gone unnoticed. In contrast, we introduce the maximum principle in section 1, and apply it to prove several new potential inequalities from which we obtain several classical applications. The main ingredient is the maximum principle - well-known in potential theory and satisfied by most of the classical kernels. We believe that our approach reveals new aspects of this principle, which has been succesfully applied in several endeavours such as capacity theory. In section 2 several sufficient conditions are given so that the "potential operator" is a continuous map of L^p-spaces. Some of the conditions are generalizations, and some are new. Section 3 deals with applications of results in section 2. Several classical cases are treated. Finally in section 4 we exhibit several interesting examples of kernels satisfying the maximum principle.

1. Maximum Principle and Potential Inequality

Let $(X, \mathcal{B}(X))$ be a measurable space, where X is a locally compact Hausdorff space with a countable base and $\mathcal{B}(X)$ the σ-algebra of Borel sets.

Let $N(x, dy)$ be a (positive) *kernel* on X in the sense that N: $X \times \mathcal{B}(X) \to [0, +\infty]$ is a mapping such that, for every $x \in X$, $A \to N(x, A)$ is a σ-finite measure, and, for every $A \in \mathcal{B}(X)$, $x \to N(x, A)$ is a Borel function. If f is a Borel function on X then we write

$$(1) \qquad (Nf)(x) = \int_X f(y) N(x, dy)$$

whenever the right side of the equation makes sense, and we refer to Nf as **the potential of** f. For example, for any nonnegative Borel function f the potential of f exists and is a nonnegative Borel function. Also, given such a kernel $N(x, dy)$

and a measure μ on $(X, \mathcal{B}(X))$ we shall write

$$(\hat{N}\mu)(dy) = \int_X N(x, dy)\mu(dx) \tag{2}$$

for the measure $\hat{N}\mu$ defined by (2).

In the special case, when $N(x, dy)$ has a density, i.e. when there exists a σ-finite measure Λ on $(X, \mathcal{B}(X))$ and a Borel function G: $X \times X \to [0, +\infty]$ such that $N(x, dy) = G(x, y)\Lambda(dy)$ we will write Gf instead of Nf and $\hat{G}\mu$ for the density of $\hat{N}\mu$, i.e.

$$(\hat{G}\mu)(y) = \int_X G(x, y)\mu(dx).$$

Hence $(\hat{N}\mu)(dy) = (\hat{G}\mu)(y)\Lambda(dy)$.

Let us denote by $\mathcal{POT} = \mathcal{POT}(N)$ the set of all Borel functions on X such that the potential of f exists. Let us denote by $\mathcal{X}^+$ the set of all non-negative Borel functions on X. Then $\mathcal{X}^+ \subseteq \mathcal{POT}(N)$, for every positive kernel N.

Definition 1: Let $N(x, dy)$ be a (positive) kernel on X and $\mathcal{R} \subseteq \mathcal{POT}(N)$. N satisfies *the strong maximum principle on* $\mathcal{R}$ (with constant $M \geq 1$), if for every $f \in \mathcal{R}$, and for every $\lambda \geq 0$,

$$s \leq M\lambda + N\left[f^+ 1_{\{s \geq \lambda\}}\right] \tag{3}$$

where $s = Nf$. In the special case, when $\mathcal{R} = \mathcal{X}^+$, we will simply say that N satisfies *the maximum principle.* ∎

We can rewrite (3) as

$$(Nf)^+ \leq M\lambda + N\left[f^+ 1_{\{(Nf)^+ \geq \lambda\}}\right]. \tag{4}$$

If (3) is true for every $\lambda \geq 0$, then (4) is true for every $\lambda > 0$, since for such λ is $\{Nf \geq \lambda\} = \{(Nf)^+ \geq \lambda\}$ and the right side of (3) is always nonnegative. For $\lambda = 0$ (4) is always true, since $\{(Nf)^+ \geq 0\} = X$ and N is a positive kernel. If (4) is true for every $\lambda \geq 0$, then (3) is true for every $\lambda > 0$, since $Nf \leq (Nf)^+$. Now, letting $\lambda \to 0$ in (3) we obtain

$$Nf \leq M \cdot 0 + N\left[f^+ 1_{\{Nf > 0\}}\right] \leq N\left[f^+ 1_{\{Nf \geq 0\}}\right].$$

Therefore, (3) and (4) are equivalent.

Notice also that, if both f and $-f$ are in $\mathcal{R}$, it follows from (4) that

$$|Nf| \leq 2M\lambda + N\left[|f| \cdot 1_{\{|Nf| \geq \lambda\}}\right], \tag{5}$$

since $[N(-f)]^+ = (Nf)^-$ and $\{(Nf)^+ \geq \lambda\} \cup \{(Nf)^- \geq \lambda\} \subseteq \{|Nf| \geq \lambda\}$. In the following text we will mainly use (4) and (5).

Remark 1: Let us recall (see [16], chapters 3 and 6) that the potential operator V of a transient Markov process satisfies the principle of positive maximum in the sense that $\sup(Vf)(x) = \sup\{(Vf)(y)\colon f(y) > 0\}$, for every continuous function f with compact support, for which Vf attains strictly positive values. If we generalize this notion, in the sense that we allow the constant M to be bigger than 1, then we can state the maximum principle as

$$(Nf \leq 1 \text{ on } \{f > 0\} \implies Nf \leq M \text{ everywhere })$$

for any bounded, measurable, nonnegative function f with compact support. Simple computation then shows that, for every $\lambda > 0$,

$$\begin{aligned} Nf &= N\left[f \cdot 1_{\{Nf<\lambda\}}\right] + N\left[f \cdot 1_{\{Nf\geq\lambda\}}\right] \\ &\leq \lambda M + N\left[f \cdot 1_{\{Nf\geq\lambda\}}\right], \end{aligned}$$

where the last inequality follows from the maximum principle above. Notice that every nonnegative Borel function f can be approximated from below by a sequence $f_n = \min(f \cdot 1_{K_n}, n)$, where $K_n \subseteq X$ is compact and, $K_n \subseteq K_{n+1}$, $\bigcup_{n=1}^{\infty} K_n = X$. The monotone convergence theorem and the fact that (4) is trivial for $\lambda = 0$, shows that (4) is satisfied for any $f \in \mathcal{X}^+$. Therefore, the maximum principle, as stated in Definition 1, is the generalization of the classical notion from potential theory. Since we allow, in general, that $\mathcal{R}$ is bigger than $\mathcal{X}^+$, it justifies the word "strong" in Definition 1.

In particular, the fact that the n-dimensional Brownian motion is a transient Markov process for $n \geq 3$ shows that the kernel $N(x, dy) = G(x, y)dy$, where dy is the n-dimensional Lebesgue measure, and

$$G(x, y) = \frac{1}{\| x - y \|^{n-2}}, \tag{6}$$

satisfies the maximum principle with constant $M = 1$. Similarly (see [16], page 3.32), the Riesz kernel I_α defined by

$$I_\alpha(x, dy) = \frac{dy}{\| x - y \|^{n-\alpha}}, \quad \text{where } n \geq 3, \quad 0 < \alpha \leq 2 \tag{7}$$

satisfies the maximum principle with constant $M = 1$.

We believe that this explains the "background" for the notion defined in Definition 1. Many other interesting examples (some of them need the generality of Definition 1) will be treated in sections 3 and 4. ∎

Before we state the main results of this paper, let us specify what we mean by convex function and concave function in this paper. We will denote the set $[0, +\infty)$ of nonnegative real numbers by $\mathbf{R}_+$.

Definition 2: We will say that $\Phi: \mathbf{R}_+ \to \mathbf{R}_+$ is a *convex function* if there exists a (positive) Borel σ-finite measure η on $\mathbf{R}_+$ such that

$$\Phi(t) = \int_0^t \varphi(u)du, \text{ for every } t \in \mathbf{R}_+, \tag{8}$$

where

$$\varphi(u) = \eta([0,u]), \text{ for every } u \in \mathbf{R}_+. \quad ∎ \tag{9}$$

Definition 3: We will say that $\Phi: \mathbf{R}_+ \to \mathbf{R}_+$ is a **concave function** if there exists a (positive) Borel measure η on $\mathbf{R}_+$ such that

$$\eta([a,b]) \text{ is finite,} \quad \text{for every } 0 < a \leq b < +\infty \tag{10}$$

and

$$\Phi(t) = \int_0^t \varphi(u)du, \quad \text{for every } t \in \mathbf{R}_+, \tag{11}$$

where φ satisfies the following properties:

$$0 \leq \varphi(u) < +\infty, \quad \text{for every } u > 0, \tag{12}$$

$$\lim_{u \to 0+} u\varphi(u) = 0, \tag{13}$$

and

$$\varphi(v) - \varphi(u) = -\eta((u,v]), \quad \text{for every } 0 < u < v < +\infty. \quad ∎ \tag{14}$$

In the definition above $\varphi(0)$ is not required to be finite. However, the behaviour of φ at zero is controlled by (13), which, as pointed out by one of the referees, is redundant. Since we need this property later, we leave it as part of the definition. Notice that one consequence of (10) is that η is σ-finite.

Remark 2: Notice that $\Phi(t) = t^p$, $p \geq 1$, satisfies the conditions in Definition 2. For $p > 1$ we get $\eta(du) = p(p-1)u^{p-2}du$, $\varphi(u) = pu^{p-1}$. For $p = 1$ we get η = the point mass at zero, $\varphi \equiv 1$.

Notice also that $\Phi(t) = t^p$, $0 < p < 1$, satisfies the conditions in Definition 3. We get $\eta(du) = p(1-p)u^{p-2}du$, $\varphi(u) = pu^{p-1}$. Then $u\varphi(u) = pu^p \to 0$, when $u \to 0+$. ∎

We can state the main results now:

THEOREM 1: (The Potential Inequality for Convex Functions) *Let Φ be a convex function (in the sense of Definition 2) and $\varphi = \Phi'$. Let $N(x, dy)$ be a kernel on X which satisfies the strong maximum principle on $\mathcal{R} \subseteq \mathcal{POT}(N)$, with constant M. Then, for every $f \in \mathcal{R}$,*

$$\Phi\left[\frac{s}{M}\right] \leq \frac{1}{M} N\left[f^+ \varphi(s)\right] \tag{15}$$

where $s = (Nf)^+$.

Proof: Let us denote by η the measure, defined by (9), which corresponds to Φ. Applying Fubini's theorem several times, and the strong maximum principle, we obtain

$$\begin{aligned}
\Phi\left[\frac{s}{M}\right] &= \int_0^{s/M} dt \int_0^t \eta(du) = \int_0^{s/M} \eta(du) \int_u^{s/M} dt \\
&= \int_0^{s/M} [s/M - u]\eta(du) = \int_0^{s/M} [s/M - u]^+ \eta(du) \\
&= \int_0^{+\infty} [s/M - u]^+ \eta(du) \leq \text{ by } (4) \\
&\leq \int_0^{+\infty} \frac{1}{M} N\left[f^+ 1_{\{s \geq u\}}\right] \eta(du) \\
&= \frac{1}{M} \int_0^{+\infty} \eta(du) \int_X N(x, dy) f^+(y) 1_{\{s(y) \geq u\}} \\
&= \frac{1}{M} \int_X N(x, dy) f^+(y) \cdot \left[\int_0^{+\infty} 1_{\{s(y) \geq u\}} \eta(du)\right] \\
&= \frac{1}{M} N\left[f^+ \varphi(s)\right]. \qquad \blacksquare
\end{aligned}$$

COROLLARY 2: *Let Φ and $N(x, dy)$ be as in the previous theorem. Then for each σ-finite measure μ on $(X, \mathcal{B}(X))$, and for every $f \in \mathcal{R}$,*

$$\int_X \Phi\left(\frac{s}{M}\right) d\mu \leq \frac{1}{M} \int_X f^+ \varphi(s) d(\hat{N}\mu). \tag{16}$$

In particular, if $\Phi(t) = t^p$, $p \geq 1$, then

$$\int_X s^p d\mu \leq pM^{p-1} \int_X f^+ s^{p-1} d(\hat{N}\mu). \tag{17}$$

Proof: The second inequality is an obvious consequence of the first inequality and the fact that $\varphi(t) = pt^{p-1}$. The first inequality follows immediately from (15) by Fubini argument. ∎

Remark 3: If f and $-f$ are both in $\mathcal{R}$, then we can use (5), instead of using (4), in Theorem 1 and Corollary 2. Then we will obtain same inequalities as before, but with $|Nf|$ instead of $(Nf)^+$, $|f|$ instead of f^+, and $2M$ instead of M. ∎

THEOREM 3: (The Potential Inequality for Concave Functions) *Let Φ be a concave function (in the sense of Definition 3) and $\varphi = \Phi'$. Let $N(x, dy)$ be a kernel on X which satisfies the maximum principle (i.e., $\mathcal{R} = \mathcal{X}^+$), with constant M. Then, for every $f \in \mathcal{X}^+$,*

$$\Phi\left[\frac{s}{M}\right] \geq \frac{1}{M} N\left[f\varphi(s)\right] \tag{18}$$

where $s = (Nf)^+ = Nf$.

Proof: The proof of this theorem is similar to the proof of Theorem 1. As before, we denote the measure which corresponds to Φ by η.

If $s(x) = 0$ then $f(y) = 0$, $N(x, dy)$-(a.e.), since $f \geq 0$ (this is the only place in the proof where we cannot have general $\mathcal{R}$). Hence, $N[f\varphi(s)](x) = 0$, which proves the desired inequality in this case.

Suppose now that $s(x) > 0$. Let us denote $s(x)/M$ by τ. Then for every $0 < t < \tau$ we obtain

$$\Phi(\tau) - \Phi(t) = \int_t^\tau \varphi(u) du = \int_t^\tau \left[\varphi(u) - \varphi(t)\right] du + \varphi(t) \cdot (\tau - t) .$$

Applying Fubini's theorem and (14), as in Theorem 1, we obtain

$$\begin{aligned}\Phi(\tau) - \Phi(t) &= -\int_{(t,\tau]} (\tau - u)\eta(du) + \varphi(t) \cdot (\tau - t) \\ &= -\int_{(t,+\infty)} [\tau - u]^+ \eta(du) + \varphi(t) \cdot (\tau - t).\end{aligned}$$

The maximum principle shows that $\tau - u \leq (1/M)(N[f1_{\{s \geq u\}}])$. Since the right side is nonnegative, it follows that $[\tau - u]^+ \leq (1/M)(N[f1_{\{s \geq u\}}])$, which implies that

$$
\begin{aligned}
\Phi(\tau)-\Phi(t) &\geq -\int_{(t,+\infty)} \frac{1}{M} N\left[f 1_{\{s\geq u\}}\right] \eta(du) + \varphi(t)\cdot(\tau-t) \\
&= -\frac{1}{M} N\left[f \int_{(t,+\infty)} 1_{\{s\geq u\}} \eta(du)\right] + \varphi(t)\cdot(\tau-t) \\
&= \frac{1}{M} N\left[f\cdot(-\eta(t,s]))\right] + \varphi(t)\cdot(\tau-t) \\
&= \frac{1}{M} N\left[f\cdot(\varphi(s)-\varphi(t))\right] + \varphi(t)\cdot(\tau-t).
\end{aligned}
$$

Notice that $\varphi(t)$ is a constant, which implies that $(1/M)N[f\varphi(t)](x)$ equals $\varphi(t)[(1/M)\cdot s(x)] = \varphi(t)\tau$. Hence, it follows that, for every $t \in (0,\tau)$,

$$\Phi(\tau)-\Phi(t) \geq \frac{1}{M} N\left[f\varphi(s)\right] - t\varphi(t).$$

Letting $t \to 0+$ in the last inequality, and using (11) and (13) we finally obtain (18). ∎

An immediate consequence of Theorem 3, exactly as in the case of Theorem 1 and Corollary 2, is the following corollary:

COROLLARY 4: *Let Φ and $N(x,dy)$ be as in Theorem 3. Then for each σ-finite measure μ on $(X,\mathcal{B}(X))$, and for every $f \in \mathcal{X}^+$,*

$$\int_X \Phi\left(\frac{s}{M}\right) d\mu \geq \frac{1}{M}\int_X f\varphi(s)d(\hat{N}\mu). \tag{19}$$

In particular, if $\Phi(t)=t^p$, $0<p<1$, then

$$\int_X s^p d\mu \geq pM^{p-1}\int_X f\, s^{p-1} d(\hat{N}\mu). \tag{20}$$

Remark 4: Notice that (15), (16), (18), and (19) are very general forms of "the integration by parts"-type formula. This becomes even more transparent when we refer to the proofs of Hardy's inequalities for $p \geq 1$ and $0<p<1$ in [9] p. 242 and p. 251–252. Hence, the best possible constants will be obtained in inequalities, derived from potential inequalities, which are classicaly proven by "the integration by parts" method. A typical example is the Hardy's inequality (see section 3). ∎

Remark 5: Notice also that the maximum principle and the potential inequality are not stated in the most general form. Purely mathematically speaking we can repeat exactly the same proof in the case when f is even a vector function and $s = Tf$, where T can be any operator (not necessarily linear, and not necessarily single-valued), which maps vector functions f into one-dimensional functions s, and, instead of f^+, we can have any nonnegative real-valued function $\beta(f)$. Of course, the corresponding potential inequality will be valid for (T, β, N) whenever the proper form of the maximum principle is satisfied for (T, β, N).

The reason that we didn't state our theorems in this generality is that we didn't find interesting examples which can be formulated in (T, β, N) setting but can not be formulated in terms of theorems stated here. Also this more general setting obscures the "nature" of the potential inequality, which was explained in the previous remark. ∎

2. Inequality $\|Nf\|_{q,\mu} \leq C\|f\|_{p,\nu}$

Let $N(x, dy)$ be a (positive) kernel on X. Let μ be a σ-finite measure on $(X, \mathcal{B}(X))$. Let ν be a measure on $(X, \mathcal{B}(X))$. Let p and q be real numbers such that $1 \leq p, q < +\infty$, and let us denote their conjugates by p' and q', respectively. More precisely $1/p + 1/p' = 1$ and $1/q + 1/q' = 1$. We denote the standard q-norm (p-norm) with respect to μ (ν) by $\| \ \|_{q,\mu}$ ($\| \ \|_{p,\nu}$). We allow these norms to attain value $+\infty$ throughout this paper. Also, in this paper we accept the convention that $0 \cdot (+\infty) = 0$.

In the special case, when μ (ν) has a density $w(x)$ ($v(x)$) with respect to Λ (for the definition of Λ see the beginning of section 1), then we will write $\| \ \|_{q,w}$ for $\| \ \|_{q,\mu}$ ($\| \ \|_{p,v}$ for $\| \ \|_{p,\nu}$). As before, we will write G instead of N, when $N(x, dy)$ has a density. For simplicity in notation we will denote $\Lambda(dx)$ by dx. Also, (a.e.) means "almost everywhere with respect to Λ" (for other measures we will specify the measure in (a.e.) expressions). Notice that the fact that μ is σ-finite implies that $0 \leq w(x) < +\infty$ (a.e.). Since ν is not necessarily σ-finite, $0 \leq v(x) \leq +\infty$ in general.

In this section we will consider the question of finding conditions for which an inequality of the type given in the title of the section is valid for some set of functions f, assuming that the potential inequality holds. Since our setting is very general we will not be able to obtain sharp constants C, but we will be concerned only in establishing inequalities of the type above in this situation.

In the following theorems we present some arguments. Instead of explaining these arguments each time, we will explain them now, and quote them when necessary.

Remark 6: If $\|f\|_{p,v} = +\infty$, then there is nothing to prove. Hence, it is enough to prove the inequality when $f^p \cdot v$ is finite (a.e.), and $\|f\|_{p,v}$ is finite. In particular, it means that $\{v = +\infty\} \subseteq \{f = 0\}$. Therefore, in our proofs we can assume, without loss of generality, that $0 \leq v < +\infty$.

Another argument in these theorems will be that in some cases there will be a set $Y \subseteq X$, $Y \in \mathcal{B}(X)$, such that

$$\int_X |Nf|^q d\mu = \int_Y |N(f1_Y)|^q d\mu \quad .$$

We claim that in such a case we can restrict our attention on $(Y, \mathcal{B}(Y))$. Clearly, it is true for μ and f. For ν we can do that, because $\|f\|_{p,v}$ on Y is less than $\|f\|_{p,v}$ on X. The only part which is not immediately clear is if the restriction

$$\tilde{N} = N/(Y \times \mathcal{B}(Y))$$

satisfies the same maximum principle as N. But, since $\tilde{N}[f/Y] = N[f1_Y]$, it follows that $\tilde{N}$ satisfies the maximum principle, too. Hence, it is enough to prove our inequality on $(Y, \mathcal{B}(Y))$. ∎

The last possibility that may occur is when $\|f\|_{p,v} = 0$. It is not clear that in this case the left side, i.e., $\|Gf\|_{q,w}$ must be zero. Under certain technical assumptions, which are satisfied in all the theorems, this will be true.

LEMMA 5: *Let $G(x,y)$ satisfies either the strong maximum principle in the form (5) or the maximum principle on $\mathcal{X}^+$. Let us denote the set $\{x\colon\ v(x) = 0\}$ by B. Suppose that $B \subseteq C \cup D$, where*

$$C = \{x\colon\ 1_B(y) \cdot G(x,y) = 0, \text{ for almost all } y\} \tag{21}$$

and

$$D = \{x\colon\ w(y) \cdot G(y,x) = 0, \text{ for almost all } y\}. \tag{22}$$

If $f \in \mathcal{R}$ and $\|f\|_{p,v} = 0$, then $\|Gf\|_{q,w} = 0$, where $1 < q < +\infty$.

Proof: We denote the set $\{\, x\colon f(x) = 0 \,\}$ by A. If $\|f\|_{p,v} = 0$, then $X = A \cup B$ (a.e.). Applying the potential inequality for convex functions on G, we obtain

$$\begin{aligned}\int_X |Gf|^q w\, dx &\leq q(2M)^{q-1} \int_X |f|\, |Gf|^{q-1} (\hat{G}w)\, dx \\ &= q(2M)^{q-1} \int_B |f|\, |Gf|^{q-1} (\hat{G}w)\, dx,\end{aligned}$$

where the last equality follows from the fact that f is zero on A. Using $B \subseteq C \cup D$ and (22), which says that $(\hat{G}w)(x) = 0$, for every $x \in D$, we obtain that

$$\int_X |Gf|^q w\, dx \leq q(2M)^{q-1} \int_C |f|\, |Gf|^{q-1} (\hat{G}w)\, dx.$$

Since $q - 1 > 0$, it is enough to prove that $(Gf)(x) = 0$, for every $x \in C$. For $x \in C$, by (21),

$$|Gf|(x) \leq \int_X G(x,y)|f|(y)dy = \int_B G(x,y)|f|(y)dy = 0. \quad ∎$$

In the case when $p = q$, we can obtain the desired inequality by applying very simple estimate of the measure $\hat{N}\mu$.

Proposition 6: *Let $N(x, dy)$ satisfy the strong maximum principle on $\mathcal{R}$ in the form (5). If there exists a constant $K \geq 0$ such that*

$$\hat{N}\mu \leq K\mu, \tag{23}$$

then, for every $f \in \mathcal{R}$,

$$\|Nf\|_{p,\mu} \leq C\|f\|_{p,\mu}, \tag{24}$$

where $C = pK(2M)^{p-1}$.

Proof: Apply the condition (23) on the inequality (17), which is valid with $s = |Nf|$ (see Remark 3). We obtain that

$$\int_X |Nf|^p d\mu \leq pK(2M)^{p-1} \int_X |f|\, |Nf|^{p-1} d\mu.$$

The straightforward application of Hölder's inequality finishes the proof. ∎

Example 1: Consider the classical example, where $N(x, dy)$ is defined by

$$Nf(x) = \int_x^{+\infty} f(y)\frac{dy}{y}.$$

It is easy to see that every nondecreasing function w satisfies (23) for N defined above. Since $N(x, dy)$ satisfies the maximum principle (see Remark 11 in the next section), then Proposition 6 shows that for every nondecreasing function w is

$$\|Nf\|_{p,w} \leq C\|f\|_{p,w}. \quad ∎$$

The condition (23) is satisfied in some classical situations (see the example above), but we would like to have more general conditions. Some of them can be derived as the consequence of the potential inequality for convex functions, and others by applying the potential inequality for concave functions.

CONCAVE CASE. In this paragraph we assume that $1 < p \leq q < +\infty$, and that all of μ, ν, and $N(x, dy)$ have densities. For $r = q/p \geq 1$ we define a (positive) kernel $G_r(x, y)$ by

$$G_r(x,y) = [G(x,y)]^r.$$

Since we have densities G and G_r, $\hat{G}(x,y) = G(y,x)$ and $\hat{G}_r(x,y) = G_r(y,x)$ are (positive) kernels, too.

Suppose now that *both* G and $\hat{G}_r$ satisfy the maximum principle on $\mathcal{X}^+$, with constants M and $\hat{M}$, respectively. Then it follows from Theorem 3 that, for every $f \in \mathcal{X}^+$, and for every $0 < a < 1$,

$$\bar{G}\left[f \cdot (\bar{G}f)^{a-1}\right] \leq \frac{\bar{M}^{1-a}}{a} \cdot (\bar{G}f)^a, \tag{25}$$

where $\bar{G}$ is either G or $\hat{G}_r$, and $\bar{M}$ is either M or $\hat{M}$, respectively.

THEOREM 7: *Suppose that w and v satisfy the condition*

$$\left[Gv^{-p'/p}\right]^{1/p'}(z) \cdot \left[\hat{G}_r w\right]^{1/q}(z) \leq K < +\infty, \tag{26}$$

for almost all $z \in X$. Then, for every $f \in \mathcal{POT}(G)$,

$$\|Gf\|_{q,w} \leq C\|f\|_{p,v}, \tag{27}$$

where $C = Kq^{1/q}(q')^{1/p'}M^{1/(qp')}\hat{M}^{1/(qq')}$.

Remark 7: Recall that we follow the convention $0 \cdot (+\infty) = 0$, and that our kernel G is very general. The proofs of some inequalities (which are relatively simple for particular kernels) are more complicated than they may look at first sight. In proving the inequality (27) we will have to consider several subcases, and the role of the potential inequality, as well as (26), will be crucial. ∎

Proof: It is enough to prove (27) for $0 \leq v < +\infty$ (see Remark 6). Consider now the sets $A = \{f = 0\}$, $B = \{v = 0\}$, $C = \{Gv^{-p'/p} = 0\}$, $D = \{Gv^{-p'/p} = +\infty\}$, and E, defined by,

$$E = \{x \in X \mid G(x,y) = 0, \text{ for almost all } y \in A^c \cap B^c\}.$$

Let us prove first that the conditions of Lemma 5 are satisfied. Notice that $v^{-p'/p} = +\infty$ on B. Hence, if $x \in B$ and $1_B(y)G(x,y)$ is positive (as a function of y) on a set of positive measure, then $[Gv^{-p'/p}](x) = +\infty$. By (26) $[\hat{G}_r w](x) = 0$, which implies that $G(y,x)w(y) = 0$, for almost all y. Therefore, by Lemma 5, we can assume that $\|f\|_{p,v} > 0$.

Consider the set C. By the assumption above $v^{-p'/p}$ is (strictly) positive. Hence, for every $x \in C$, $G(x,y) = 0$, for almost all y. In particular $\{G1_C \geq \lambda\} \subseteq C^c$, for every $\lambda > 0$. Applying now the maximum principle for G on 1_C we obtain

$$G1_C \leq M\lambda + G[1_C \cdot 1_{C^c}] = M\lambda,$$

for every $\lambda > 0$. Then $(G1_C)(x) = 0$, for every $x \in C$, i.e., for every $x \in C$, $G(x,y)1_C(y) = 0$, for almost all y. Then, by Remark 6, we can restrict ourselves to C^c, or, in other words, we can assume that $C = \emptyset$ (a.e.).

Consider the set D. If $x \in D$, then, by (26), $[\hat{G}_r w](x) = 0$, and, therefore, $(\hat{G}w)(x) = 0$, too. Applying (17) on G and $f \in \mathcal{POT}(G)$, we obtain

$$\| G(f1_D) \|_{q,w}^q \leq \text{const.} \cdot \int_D |f| \, [G(|f|1_D)]^{q-1} \hat{G}w \, dx = 0,$$

which shows that we can assume, without loss of generality, that $f = 0$ (a.e.) on D, i.e., $D \subseteq A$ (a.e.).

Consider the set E. By the definition of E, it follows that

$$\int_E |Gf|^q w dx = \int_X |G(f1_{A\cup B})|^q (1_E w) dx.$$

The fact that $\hat{G}_r(1_E w) \leq \hat{G}_r w$ implies that we can consider v and $1_E w$ as weights, for the present, and the conditions of Lemma 5 are fulfilled in this case. Since $f1_{A\cup B} \cdot v = 0$, it follows that

$$\int_E |Gf|^q w dx = 0. \tag{28}$$

Using all these we obtain that for $x \in X \backslash E$

$$|Gf|(x) \leq \int_X G(x,y)|f|(y)dy$$

$$= \int_{A^c} G(x,y)|f|(y)dy \quad (\text{since } D \subseteq A, \ C = \emptyset)$$

$$= \int_{A^c\cap B^c} G(x,y)|f|(y)\left\{v^{1/p}\, v^{-1/p}\,[Gv^{-p'/p}]^{1/(p'q)}\,[Gv^{-p'/p}]^{-1/(p'q)}\right\}(y)\,dy$$
$$+\int_{A^c\cap B} G(x,y)|f|(y)dy$$
$$\leq \left\{\int_{A^c\cap B^c} G(x,y)\,|f|^p(y)\,v(y)\,[Gv^{-p'/p}]^{p/(p'q)}(y)\,dy\right\}^{1/p}$$
$$\times\left\{\int_{A^c\cap B^c} G(x,y)\,v^{-p'/p}(y)\,[Gv^{-p'/p}]^{-1/q}(y)\,dy\right.$$
$$\left.+\int_{A^c\cap B} G(x,y)\,v^{-p'/p}(y)\,[Gv^{-p'/p}]^{-1/q}(y)\,dy\right\}^{1/p'}.$$

To justify the last inequality notice that all the functions involved in the first integral on the right side are (strictly) positive on $A^c\cap B^c$, and, since $x\in X\backslash E$, this integral must be (strictly) positive. If $A^c\cap B=\emptyset$ (a.e.), then it is just Hölder's inequality. If $G(x,y)=0$, for almost all $y\in A^c\cap B$, then it is again just Hölder's inequality. If it is none of these two, then the third integral is $+\infty$, and, since the first integral is positive, the right side is $+\infty$.

Finally, using (28) and the formula above, we obtain

$$\begin{aligned}\int_X |Gf|^q w\,dx \leq \int_X w\cdot &\left[\int_X G(x,y)\,|f|^p\, v\,(Gv^{-p'/p})^{p/(p'q)}dy\right]^{q/p}\\ \times &\left[\int_X G(x,y)\,v^{-p'/p}(Gv^{-p'/p})^{-1/q}\,dy\right]^{q/p'} dx.\end{aligned}\tag{29}$$

We apply (25) on G, $v^{-p'/p}$, and $a=1/q'$ (notice $a-1=-1/q$) to get

$$\left\{\int_X |Gf|^q w\,dx\right\}^{1/r}\leq \frac{M^{1/(rp')}}{(1/q')^{q/(rp')}}$$
$$\times\left\{\int_X w(Gv^{-p'/p})^{q/(q'p')}\cdot\left[\int_X G(x,y)|f|^p v(Gv^{-p'/p})^{p/(p'q)}dy\right]^r dx\right\}^{1/r}$$

Using the general form of Minkowski's inequality (see, for example, Theorem 202 on p.148 in [9]) we have that

$$\|Gf\|^p_{q,w}\leq M^{1/(rp')}(q')^{q/(rp')}\int_X |f|^p v(Gv^{-p'/p})^{p/(p'q)}$$
$$\times\left[\int_X G_r(x,y)w(x)(Gv^{-p'/p})^{q/(q'p')}(x)\,dx\right]^{1/r} dy$$
$$= M^{1/(rp')}(q')^{q/(rp')}\int_X |f|^p v(Gv^{-p'/p})^{p/(p'q)}\left\{\hat{G}_r[w(Gv^{-p'/p})^{q/(p'q')}]\right\}^{1/r} dy.$$

We apply (26) on $(Gv^{-p'/p})^{q/(q'p')}$, and then (25) on $\hat{G}_r$, w, and $a = 1/q$, to obtain

$$\|Gf\|_{q,w}^p \le M^{\frac{1}{rp'}} \hat{M}^{\frac{1}{rq'}} (q')^{\frac{q}{rp'}} q^{\frac{1}{r}} K^{\frac{q}{rq'}} \int_X |f|^p\, v\, (Gv^{-p'/p})^{\frac{p}{qp'}} (\hat{G}_r w)^{\frac{p}{q^2}}\, dy.$$

Now, one more application of (26) completes the proof. ∎

Remark 8: The condition (26) has been treated in connection with Hardy's inequality with weights. More precisely, it is well-known that in the case of Hardy's inequality, i.e., in the case when $X = (0, +\infty)$ and

$$G(x,y) = \begin{cases} 1 & \text{if } y \le x \\ 0 & \text{otherwise} \end{cases} \tag{30}$$

(26) is necessary and sufficient condition for (27) to be true.

Surprisingly enough, as we have shown in Theorem 7, (26) is still a sufficient condition for such a general space as X, and such a general kernel G as in this article. Notice that the crucial steps were the applications of the potential inequality for concave functions (twice; for G and $\hat{G}_r$). These applications correspond to the applications of the integration by parts in Hardy's classical cases.

Notice also that in classical situation r cannot be observed on G, since the values of G are either 0 or 1. Even in (26) we have $\hat{G}$ instead of $\hat{G}_r$, in the case when $p = q$.

Hardy's inequality with weights is considered very early using methods via differential equations (see [6], [3], [4] ; see also historical remarks on p.14 in [11]). But, the condition (26) appears for the first time (as far as we know) in Talenti [19] (1966). He considered only the case $p = q = 2$ and weights related by formula $w(x) = x^{-2}v(x)$. These restrictions were removed in later papers by Talenti [20], and especially Tomaselli [21]. These theorems were reproved and extended to the case of measures by Muckenhoupt [15] in 1972. He also mentioned, in connection with the condition (26), the untitled and unpublished manuscript by M. Artola (unfortunately we were not able to see this manuscript). In contrast to previous authors, who still relied on differential equations, Muckenhoupt uses Hölder's inequality and integration by parts formula, which is a technique that we managed to apply in the generality of Theorem 7. In all these papers the case $p = q$ was treated. For the case $p \le q$ the proof was done by J.S. Bradley [7] in 1978, and, independently, a year later in [13] and [10]. A thorough treatment of these theorems can be found in [14] and in [11], where also detailed historical remarks are given. In all these papers the special kernel G from (30) (or some simple modification of G or $\hat{G}_r$) is treated. ∎

The fact that in the classical case the condition (26) is also necessary suggests that it will be interesting to examine the problem in the general situation. The next proposition gives a condition on a kernel G (which is satisfied in classical Hardy's case) under which (26) is a necessary condition. The example after the proposition shows that there are kernels for which (26) is not a necessary condition.

We denote the set $\{y\colon G(z,y)>0\}$ by P_z, and the set $\{x\colon G(x,z)>0\}$ by P^z.

PROPOSITION 8: *Suppose that there exists a constant $K>0$ such that, for every $z\in X$, for every $x\in P^z$, and for every $y\in P_z$*

$$G(x,y)\cdot\left[\frac{G(z,y)}{G(x,z)}\right]^{1/p}\geq K\cdot G(z,y). \tag{31}$$

If the inequality (27) is satisfied for some $C>0$, and for all $f\in\mathcal{X}^+\subseteq\mathcal{POT}(G)$, then (26) is satisfied for all $z\in X$, with constant C/K.

Proof: For every $z\in X$ we consider $f_z\in\mathcal{X}^+$, defined by

$$f_z(y)=\begin{cases}G(z,y)^{1/p}v^{-p'/p}(y) & \text{if } y\in P_z,\\ 0 & \text{otherwise.}\end{cases}$$

Then on the right side of the inequality (27) we have

$$C\|f_z\|_{p,v}=C\left(\int_X f_z^p(y)v(y)dy\right)^{1/p}$$

$$=C\left(\int_{P_z}G(z,y)v^{-p'}(y)v(y)dy\right)^{1/p}=C[(Gv^{-p'/p})(z)]^{1/p}.$$

On the left side of (27) we have (with $r=q/p$), using (31),

$$\|Gf_z\|_{q,w}\geq\left(\int_{P^z}(Gf_z)^q(x)w(x)dx\right)^{1/q}$$

$$=\left[\int_{P^z}G_r(x,z)w(x)\left(\int_{P_z}\frac{G(x,y)}{[G_r(x,z)]^{1/q}}[G(z,y)]^{1/p}v^{-p'/p}(y)dy\right)^q dx\right]^{1/q}$$

$$\geq K\left[\int_{P^z} G_r(x,z)w(x)\left(\int_{P_z} G(z,y)v^{-p'/p}(y)dy\right)^q dx\right]^{1/q}$$
$$= K[(\hat{G}_r w)(z)]^{1/q}\cdot[(Gv^{-p'/p})(z)].$$

The proof is finished by putting both sides of (27) together (the case 0, $+\infty$ works here by the standard argument). ∎

Remark 9: Notice that in the case when $G(x,y)$ is either 0 or 1 condition (31) simply says that, for every $z \in X$, and for every $x \in P^z$,

$$P_z \subseteq P_x,$$

and $K = 1$. In a classical Hardy's case $P^z = [z,+\infty)$ and $P_z = [0,z]$, which shows that (31) is satisfied. ∎

Example 2: Let $1 < p < +\infty$ and α such that $0 < p\alpha < 1$. Then $0 < \alpha < 1$ and $0 < 1 - p\alpha < 1$. Let us denote $1 - p\alpha$ by β. We define the kernel $G(x,y)$ on $(0,+\infty)$ by

$$G(x,y) = \begin{cases} (y-x)^{\alpha-1} & \text{if } 0 < x \leq y < +\infty, \\ 0 & \text{otherwise;} \end{cases} \tag{32}$$

G is known as the Weyl fractional kernel, and $\hat{G}$ as the Riemann–Liouville fractional kernel. Let us take $w(x) = x^{\beta-1}$, and $v(x) \equiv 1$. Then simple calculations lead to

$$(\hat{G}w)(y) = Ky^{\alpha+\beta-1} \quad \text{and} \quad (Gv^{-p'/p})(x) \equiv +\infty,$$

where K depends on α and β only. $\hat{G}$ and G both satisfy the maximum principle (see Proposition 12), but obviously (26) is not satisfied for v and w. We claim that (27) still holds for v and w, in the case $p = q$. We will show that v and w satisfy conditions in Theorem 10 below. Indeed

$$v^{-p'/p}\cdot(\hat{G}w)^{p'} = K^{p'}w,$$

since $p'(\alpha+\beta-1) = p'(\alpha+1-p\alpha-1) = \alpha p'(1-p) = -\alpha p = \beta - 1$. Riemann–Liouville and Weyl fractional kernels are treated in [1]. ∎

CONVEX CASE. In this paragraph we assume that $1 < p \leq q < +\infty$, and that measures μ and ν have densities. We assume also that $N(x,dy)$ satisfies the strong maximum principle on $\mathcal{R} \subseteq \mathcal{POT}(N)$ in the form (5). Therefore we will apply the potential inequality for convex functions as it was stated in Remark 3.

Our first theorem in the convex case is inspired by methods developed in Stepanov [18] for classical Hardy's case. Consider first the case when $p < q$. In this case we assume that $N(x, dy)$ has a density.

THEOREM 9: *Suppose there exist a Borel function w^* on X, which is (a.e.) finite, and a positive constant K, such that the following two conditions are satisfied:*

$$v^{-\frac{p'}{p}}(G_{p'}v^{-p'/p})^{\frac{q-p}{p}}(\hat{G}w^*)^{p'} = w^* \quad \text{(a.e.)}, \tag{33}$$

and

$$\hat{G}w \le K \cdot \hat{G}w^* \quad \text{(a.e.)}. \tag{34}$$

Then there exists a constant C such that, for every $f \in \mathcal{R}$,

$$\|Gf\|_{q,w} \le C\|f\|_{p,v} \ . \tag{35}$$

Proof: Let us prove the inequality (35) for w^*, instead of w, first. It is enough to prove it for $0 \le v < +\infty$ (see Remark 6). Consider now the sets $B = \{v = 0\}$, $C = \{G_{p'}v^{-p'/p} < +\infty\}$, and $D = \{\hat{G}w^* = 0\}$. If $x \in B$ then $v^{-p'/p}(x) = +\infty$. Since w^* is finite (a.e.) it follows that (33) can be fulfilled only if

$$G_{p'}v^{-p'/p} \cdot \hat{G}w^* = 0 \quad \text{(a.e.) on } B.$$

Thus $w^* = 0$ (a.e.) on B and

$$B \subseteq (C \cup D) \cap \{w^* = 0\} \quad \text{(a.e.)}. \tag{36}$$

Notice also that for every $x \in C$

$$1_B(y)G_{p'}(x,y) = 1_B(y)G(x,y) = 0 \qquad \text{for almost all } y. \tag{37}$$

In particular, it shows that Lemma 5 is satisfied. Therefore we can assume, without loss of generality, that $\|f\|_{p,v} > 0$. Notice also that $0 \le v < +\infty$, (33), and (36) imply

$$(G_{p'}v^{-p'/p})^{\frac{q-p}{p}}(\hat{G}w^*)^{p'} = v^{p'/p}w^* \quad \text{(a.e.)}. \tag{38}$$

We claim that

$$|Gf|(x) \le \|f\|_{p,v} \cdot \left(\int_X G_{p'}(x,y)v^{-p'/p}(y)dy\right)^{1/p'} \tag{39}$$

for every $f \in \mathcal{R}$.

Notice that if $G_{p'}v^{-p'/p} = +\infty$, then, since $\|f\|_{p,v} > 0$, (39) is trivial. If $G_{p'}v^{-p'/p} < +\infty$, i.e., if $x \in C$, then, using (37), the fact that $0 < v < +\infty$ on B^c, and Hölder's inequality, we obtain

$$\begin{aligned}
|Gf|(x) &\leq \int_{B^c} G(x,y)|f|(y)dy \\
&= \int_{B^c} G(x,y)v^{-1/p}(y)v^{1/p}(y)|f|(y)dy \\
&\leq \left(\int_{B^c} |f|^p(y)v(y)dy\right)^{1/p} \cdot \left(\int_{B^c} G_{p'}(x,y)v^{-p'/p}(y)dy\right)^{1/p'}
\end{aligned}$$

which proves (39).

We obtain the desired inequality for w^* using the potential inequality, (38), and (39) in the following way

$$\begin{aligned}
\int_X |Gf|^q w^* dx &\leq K_1 \int_X |f|\,|Gf|^{q-1}(\hat{G}w^*)dx \\
&= K_1 \int_X |f|\,|Gf|^{q/p'}|Gf|^{(q-p)/p}(\hat{G}w^*)dx \leq \text{ by (39)} \\
&\leq K_1 \|f\|_{p,v}^{\frac{q-p}{p}} \cdot \int_X |f|\,|Gf|^{\frac{q}{p'}}(G_{p'}v^{-p'/p})^{\frac{q-p}{pp'}}(\hat{G}w^*)dx \\
&= (\text{ by (38)})K_1\|f\|_{p,v}^{\frac{q-p}{p}} \cdot \int_X |f|v^{1/p}|Gf|^{q/p'}(w^*)^{1/p'}dx \\
&\leq K_1\|f\|_{p,v}^{\frac{q-p}{p}} \cdot \|f\|_{p,v} \cdot \left(\int_X |Gf|^q w^* dx\right)^{1/p'} \\
&= K_1\, \|f\|_{p,v}^{q/p}\, \|Gf\|_{q,w^*}^{q/p'}.
\end{aligned}$$

Finally, we will consider two possibilities to prove the inequality (35). Either

$$\int_X |Gf|^q w dx \leq \int_X |Gf|^q w^* dx$$

or

$$\int_X |Gf|^q w^* dx \leq \int_X |Gf|^q w dx.$$

In the first case it is simply

$$\|Gf\|_{q,w} \leq \|Gf\|_{q,w^*} \leq C\|f\|_{p,v}.$$

In the second case we apply the potential inequality on w first, then the condition (34), after that the inequality for w^*, and finally the assumption of the second case, to obtain

$$\int_X |Gf|^q w dx \leq K_1 \int_X |f|\,|Gf|^{q-1}(\hat{G}w)dx$$

$$\leq C_1 \int_X |f|\,|Gf|^{q-1}(\hat{G}w^*)dx \leq C_1\|f\|_{p,v}^{q/p} \cdot \left(\int_X |Gf|^q w^* dx\right)^{1/p'}$$

$$\leq C_1\|f\|_{p,v}^{q/p} \cdot \left(\int_X |Gf|^q w dx\right)^{1/p'}. \quad \blacksquare$$

In the case $p = q$ we can drop the assumption that $N(x, dy)$ has a density, and again obtain the same theorem as above. The proof is exactly the same as in the previous theorem, except that we are dealing with measures in conditions (33) and (34), instead of with functions. We can do that since for $p = q$ the condition (33) is much simpler than in the theorem above. Since in this case the standard argument on measures apply we will state the theorem only.

THEOREM 10: *Suppose there exist a Borel function w^* on X, which is finite (a.e.), and a positive constant K, such that the following two conditions are satisfied:*

$$v^{-1/p}(x)(\hat{N}w^*)(dx) = (w^*)^{1/p'}(x)(dx) \tag{40}$$

and

$$(\hat{N}w)(dx) \leq K \cdot (\hat{N}w^*)(dx). \tag{41}$$

Then there exists a constant C such that, for every $f \in \mathcal{R}$,

$$\|Nf\|_{p,w} \leq C\,\|f\|_{p,v}. \tag{42}$$

Remark 10: Notice that the classical Hardy's case and Example 2 show that there are cases in which Theorem 7 is satisfied, but not Theorems 9 and 10, and there are cases in which Theorem 10 (Theorem 9) is satisfied, but not Theorem 7. $\blacksquare$

So far we have considered only the cases related to Hardy's inequality with weights. Let us show that direct applications of the potential inequality for convex functions and Hölder's inequality lead to theorems which are not satisfied even for the original Hardy's inequality from 1920. We believe that these theorems can be applied on some other kernels on bounded domains. The following theorem is a typical example of such results. We assume that $p = q$ and that $N(x, dy)$ has a density.

THEOREM 11: *Suppose that v and w satisfy the following condition:*

$$\sup_{y\in X}\left[\hat{G}\left[v^{-p'/p}(\hat{G}w)^{p'}(Gv^{-p'/p})^{p/p'}\right]\right](y)\leq M<+\infty. \tag{43}$$

Then there exists a constant C such that, for every $f\in\mathcal{R}$,

$$\|Gf\|_{p,w}\leq C\,\|f\|_{p,v}. \tag{44}$$

Proof: Let us denote $|Gf|$ by s. Using Hölder's inequality and arguments on 0, $+\infty$ cases as in the theorems above we obtain

$$s(x)\leq\int_X G(x,y)|f|(y)dy=\int_X G(x,y)|f|(y)v^{-1/p}(y)v^{1/p}(y)dy$$

$$\leq\left(\int_X G(x,y)|f|^p(y)v(y)dy\right)^{1/p}\left(\int_X G(x,y)v^{-p'/p}(y)dy\right)^{1/p'}$$

so that

$$s^p\leq[G(|f|^pv)][Gv^{-p'/p}]^{p/p'}.$$

Applying the potential inequality for convex functions we obtain

$$\begin{aligned}\int_X s^p(x)w(x)dx&\leq K\int_X|f|(x)s^{p-1}(x)(\hat{G}w)(x)dx\\&=K\int_X|f|v^{1/p}s^{p-1}(\hat{G}w)v^{-1/p}dx\\&\leq\|f\|_{p,v}\cdot\left(\int_X s^p(\hat{G}w)^{p'}v^{-p'/p}dx\right)^{1/p'}.\end{aligned}$$

On the other hand

$$\int_X s^p(\hat{G}w)^{p'}v^{-p'/p}dx\leq\int_X[G(|f|^pv)][Gv^{-p'/p}]^{p/p'}v^{-p'/p}(\hat{G}w)^{p'}dx$$

$$=\text{by Fubini Theorem}=\int_X|f|^pv\hat{G}\left[v^{-p'/p}(\hat{G}w)^{p'}(Gv^{-p'/p})^{p/p'}\right]dy$$

$$\leq\text{by (43)}\leq M\int_X|f|^pvdy.$$

Finally, it shows that there exists a constant C such that

$$\int_X s^pwdx\leq C\|f\|_{p,v}\|f\|_{p,v}^{p/p'}=C\int_X|f|^pvdx. \qquad\blacksquare$$

3. Applications: Analytic Inequalities

The following proposition gives a large class of kernels satisfying the maximum principle.

PROPOSITION 12: *Let $G(x,y)$ be a nonnegative function on $\mathbf{R}_+ \times \mathbf{R}_+$ satisfying*

$$G(x,y) = 0 \quad \textit{for } y > x, \tag{45}$$

$$\textit{for each } y,\ x \longrightarrow G(x,y) \textit{ is decreasing in } (y,+\infty), \tag{46}$$

and

$$\textit{for each bounded measurable } f,\ x \to (Gf)(x) \textit{ is continuous.} \tag{47}$$

Then G satisfies the maximum principle with constant 1.

Proof: Let $\lambda > 0$, $f \geq 0$, and $Gf(x) > \lambda$. Since Gf is continuous, the set $\{Gf > \lambda\}$ is open. Let a be the left end point of the component interval containing x. We have then

$$Gf(x) = \int_0^a G(x,y)f(y)dy + \int_a^x G(x,y)f(y)dy.$$

In the interval (a,x), $Gf > \lambda$, and, by property (46) of G, in the first integral if we replace x by a we increase the integral, i.e.,

$$\int_0^a G(x,y)f(y)dy \leq \int_0^a G(a,y)f(y)dy = \lambda$$

by choice of a. Using these remarks we obtain

$$\begin{aligned} Gf(x) &\leq \lambda + \int_a^x G(x,y)f(y)1_{\{Gf\geq\lambda\}}(y)dy \\ &\leq \lambda + \int_0^x G(x,y)f(y)1_{\{Gf\geq\lambda\}}(y)dy, \end{aligned}$$

proving the result. ∎

Remark 11: Similarly we may show that if $G(x,y)$ is a nonnegative function on $\mathbf{R}_+ \times \mathbf{R}_+$ such that

$$G(x,y) = 0 \quad \text{for } y < x, \tag{48}$$

(49) Gf is continuous for bounded measurable f with bounded support,

$$G(x,y) \leq M \cdot G(a,y) \quad \text{if } x < a \leq y, \tag{50}$$

then $G(x,y)$ satisfies the maximum principle with constant M. ∎

Example 3: Let α, β be positive numbers and let

$$A(x,y) = \begin{cases} (x^\alpha + y^\alpha)^{-\beta} & \text{if } x \leq y, \\ 0 & \text{if } 0 < y < x. \end{cases} \tag{51}$$

Then A satisfies the maximum principle with constant 2^β. And

$$B(x,y) = \begin{cases} (x^\alpha + y^\alpha)^{-\beta} & \text{if } 0 < y \leq x \\ 0 & \text{if } x < y \end{cases} \tag{52}$$

satisfies the maximum principle with constant 1. And $G = A + B$ satisfies the maximum principle with constant 2^β. ∎

HILBERT'S INEQUALITY. When $\alpha = \beta = 1$ in Example 3 then $H = A + B$ is the Hilbert kernel $(x+y)^{-1}$: $x, y > 0$. An inequality due to Hilbert states:

$$\|Hf\|_p \leq \frac{\pi}{\sin(\pi/p)} \|f\|_p \tag{53}$$

where H is the Hilbert kernel. Actually (53) was proved for $p = 2$ by Hilbert and by M. Riesz for $p > 1$. Attempts to simplify the then available proofs led to Hardy's inequality. $\pi/\sin(\pi/p)$ is the best constant in (53) (see [9]). Our methods do not give the best constant, but we can prove that H maps L^p into L^p continuously. Here is a more general result:

Let A be defined as in (51), and suppose $\alpha\beta = 1$. Then with $d\mu = dx$ we find $\hat{A}\mu = C\,dx$ where

$$C = \int_0^1 \frac{1}{(z^\alpha + 1)^\beta} dz.$$

So from (17) we get

$$\|Af\|_p \leq pC2^{p-1}\|f\|_p.$$

Now duality arguments show that if a kernel G satisfies $\|Gf\|_p \leq C_p\|f\|_p$ then the dual kernel $\hat{G}$ satisfies the same inequality with the same constant. Since $\hat{A}$ is just B of (52) we see that $A + B$ is a continuous map of L^p into L^p. As said before when $\alpha = \beta = 1$ we get Hilbert–Riesz result.

HARDY'S INEQUALITY. The well-known and very useful Hardy's inequality states: if $\varepsilon < p - 1$ then

$$\int_0^\infty F^p x^{\varepsilon - p} dx \leq \left(\frac{p}{|\varepsilon - p + 1|}\right)^p \int_0^\infty f^p x^\varepsilon dx \tag{54}$$

where $f \geq 0$ and $F(x) = \int_0^x f(t)dt$. To derive this from our results let $G(x,y)$ be defined by (30). By Proposition 12, G satisfies the maximum principle with constant 1, and $Gf = F$. From (17) with $d\mu = x^{\varepsilon - p}dx$ we get, if $\varepsilon < p-1$,

$$\int_0^\infty F^p x^{\varepsilon - p} dx \leq \frac{p}{|\varepsilon - p + 1|} \int_0^\infty f\, F^{p-1} x^{\varepsilon - p + 1} dx.$$

Now use Hölder's inequality with $x^\varepsilon dx$ as the base measure to get (54).

The generalized Hardy's inequality due to A. Kufner and H. Triebel [12] states:

$$\int_0^\infty F^p \sigma_0 \, dx \leq q^p \int_0^\infty f^p \sigma_1 \, dx, \tag{55}$$

where F is as before, $q = p/(p-1)$, and σ_0 and σ_1 are functions such that

$$\sigma_1(t) = (p-1)^p \sigma_0^{1-p}(t) \left[\int_t^\infty \sigma_0(s) ds \right]^p. \tag{56}$$

To derive this take in (17) $d\mu = \sigma_0(x)dx$. Then $d(\hat{G}\mu) = \sigma_1^{1/p}(p-1)^{-1}\sigma_0^{1/q}dx$ and

$$\int_0^\infty F^p \sigma_0 \, dx \leq \frac{p}{p-1} \int_0^\infty f F^{p-1} \sigma_1^{1/p} \sigma_0^{1/q} dx.$$

Using Hölder's inequality as before we obtain (55).

We may regard Hardy's inequality as an estimate of the size (in L^p) of the "averages" $(1/x)\int_0^x f(y)dy$. This interpretation permits the following generalizations. In $\mathbf{R}^n$ define

$$G(x,y) = \begin{cases} 1 & \text{if } \|y\| \leq \|x\|, \\ 0 & \text{if } \|y\| > \|x\|. \end{cases} \tag{57}$$

It may be verified that the kernel $G(x,y)$ satisfies the maximum principle with constant 1. Taking

$$d\mu = \|x\|^{-np} dx$$

we get from (17), using $d(\hat{G}\mu) = [n(p-1)]^{-1}\, \omega_n \, \|y\|^{-n(p-1)}\, dy$,

$$\int_{\mathbf{R}^n} F^p \, \|x\|^{-np} dx \leq \frac{\omega_n p}{n(p-1)} \int_{\mathbf{R}^n} f\, F^{p-1} \|y\|^{-n(p-1)} dy,$$

where ω_n is the area of the surface of the unit sphere, and $F = Gf$, i.e.,

$$F(x) = \int_{B(0,\|x\|)} f(y)dy,$$

where $B(0, \|x\|)$ is the ball of radius $\|x\|$ around 0. Using Hölder's inequality again we obtain *an n-dimensional version of Hardy's inequality*

$$\|s\|_p \leq \frac{\omega_n p}{n(p-1)} \|f\|_p, \tag{58}$$

where

$$s(x) = \frac{1}{\|x\|^n} \int_{B(0,\|x\|)} f(y)dy.$$

Remark 12: Notice that in the literature different inequalities appear under the name "n-dimensional Hardy's inequality". For example, weighted Sobolev inequality (its special case is Friedrichs' inequality below) is considered in [11] as "n-dimensional Hardy's inequality". Our inequality (58) is of a different type. It doesn't involve the gradient of the function, but the "average" of the function with respect to the expanding balls. Notice that the generality of our approach enables us to also consider the other types of expanding sets. ∎

FRIEDRICHS' INEQUALITY. Let Ω be a bounded domain in $\mathbf{R}^n$. For every $u \in C^1(\Omega)$, such that $\text{supp}(u) \subseteq \Omega$, the well-known Friedrichs' inequality is valid, i.e.,

$$\int_\Omega |u|^p dx \leq C \int_\Omega \|\nabla u\|^p dx, \tag{59}$$

where $\nabla u = (\partial u/\partial x_1, \ldots, \partial u/\partial x_n)$ and $\|\nabla u\| = [\sum_{i=1}^n (\partial u/\partial x_i)^2]^{1/2}$.

Let consider this inequality carefully. Notice that, since Ω is open, u can be interpreted as in $C^1(\mathbf{R}^n)$, where u and ∇u are zero outside of Ω. We have the following well-known formula:

$$u(x) = \frac{1}{\omega_n} \int_{\mathbf{R}^n} \frac{\nabla u(y) \bullet (x-y)}{\| x - y \|^n} dy, \tag{60}$$

where ω_n is the area of the unit sphere $S^n(1)$ in $\mathbf{R}^n$. It follows that

$$|u| \leq v$$

where v is the "potential"

$$v(x) = \frac{1}{\omega_n} \int_{\mathbf{R}^n} \frac{\|\nabla u\|}{\| x - y \|^{n-1}} dy. \tag{61}$$

The kernel $\| x - y \|^{-n+1}$ satisfies the maximum principle with constant 1 (see Proposition 13 and (7)). Since in (61) we can put Ω instead of $\mathbf{R}^n$, and

$$\frac{1}{\omega_n} \int_\Omega \frac{dx}{\| x - y \|^{n-1}} \leq \text{diam}(\Omega) < +\infty,$$

for every $y \in \Omega$, we apply (17) on (61) to get

$$\int_\Omega v^p dx \le C \int_\Omega \|\nabla u\|^p dx. \tag{62}$$

Since $|u| \le v$, (59) is immediate from (62). However, our discussion below yields (as a particular case)

$$\int_\Omega |u|^p dx \le C \int_\Omega \|\nabla u\| \, |u|^{p-1} dx, \tag{63}$$

which also leads to (59).

Let us show that a slight modification in the argument leads to a better inequality than (59). We claim that u satisfies the following form of the maximum principle:

$$|u(x)| \le \lambda + \frac{1}{\omega_n} \int_{\mathbf{R}^n} \frac{|\nabla u(y) \bullet (x-y)|}{\| x - y \|^n} 1_{\{u(y) \ge \lambda\}} dy, \tag{64}$$

for every $\lambda \ge 0$. Again, we can replace $\mathbf{R}^n$ in (64) with Ω, and then repeat the proof of Theorem 1; using (64) instead of (4) (see also Remark 5), to get

$$\Phi(|u(x)|) \le \frac{1}{\omega_n} \int_\Omega \frac{|\nabla u(y) \bullet (x-y)|}{\| x - y \|^n} \varphi(|u(y)|) dy, \tag{65}$$

where Φ is the convex function (see Definition 2). Finally, let us prove that (64) holds. It is enough to prove that (64) holds for $x \in \Omega$ such that $u(x) > \lambda$, where $\lambda > 0$. Notice that the set $O = \{u - \lambda > 0\}$ is open, the set $K = \{u - \lambda \ge 0\}$ is bounded and closed, and $x \in O \subseteq K$. Then $u - \lambda = 0$ on ∂O, and, for every $\omega \in S^n(1)$, there exists $T(\omega) > 0$, such that $x - T(\omega) \cdot \omega \in \partial O$, and $x - t\omega \in O$, for every $0 \le t < T(\omega)$. Hence

$$-u(x) + \lambda = \int_0^{T(\omega)} \frac{d}{dt} u(x - t\omega)\, dt = \int_0^{T(\omega)} \nabla u(x - t\omega) \bullet (-\omega)\, dt.$$

It follows that

$$u(x) = \lambda + \int_0^{T(\omega)} \nabla u(x - t\omega) \bullet \omega \, dt,$$

which leads to

$$\begin{aligned} |u(x)| &\le \lambda + \int_0^{T(\omega)} |\nabla u(x - t\omega) \bullet \omega| \, dt \\ &\le \lambda + \int_0^{\infty} |\nabla u(x - t\omega) \bullet \omega| \, 1_{\{u(x-t\omega) \ge \lambda\}} dt. \end{aligned}$$

Now we integrate both sides with respect to the surface measure on $S^n(1)$, and change the variable $z = t\omega$, to get

$$\omega_n|u(x)| \le \omega_n\lambda + \int_{\mathbf{R}^n} \frac{|\nabla u(x-z)\bullet z|}{\| z \|^n} 1_{\{u(x-z)\ge\lambda\}}dz,$$

which is equivalent to (64).

4. Miscellaneous Examples

The following known proposition gives a wide class of kernels satisfying the maximum principle. We provide a proof for the reader's convenience.

PROPOSITION 13: *Let ψ be a nonnegative decreasing function on $[0,+\infty)$. Define G on $\mathbf{R}^n \times \mathbf{R}^n$ by*

$$G(x,y) = \psi(\| x-y \|). \tag{66}$$

Then G satisfies the maximum principle.

Proof: Let μ be a measure on $\mathbf{R}^n$ with compact support F. We show below that there is a C such that

$$G\mu(x) \le C \sup_{y\in F} G\mu(y). \tag{67}$$

We claim that there is an integer m depending only on n with the following property: For any $x \in F^c$, we can find m closed subsets $V_i = V_i(x) \subseteq F$ such that if x_i is a point nearest to x in V_i, then $\| x_i - y \| \le \| x-y \|$, for all $y \in V_i$. Indeed let $F_1,\dots,F_m$ be closed subsets of the unit sphere $\{y:\ \|y\| = 1\}$ such that $\operatorname{diam}F_i \le 1$ and with $\bigcup_{i=1}^m F_i = \{y:\ \|y\|=1\}$. This number m depends only on the dimension n. Put $R_i = \{tx:\ x\in F_i,\ t>0\}$, and note that for $\xi,\eta \in R_i$ with $\|\xi\| \le \|\eta\|$, $\|\xi-\eta\| \le \|\eta\|$. Given $x \in F^c$ let $V_i = F\cap(x-R_i)$. If x_i in V_i is any point nearest to x, then for any $y\in V_i$, the points $z_i = x-x_i$ and $z = x-y$ both belong to R_i and $\| x-x_i \| \le \| x-y \|$. Hence remembering the note above $\| z_i - z \| \le \|z\|$, i.e., $\| x_i - y \| \le \| x-y \|$ as claimed.

Now to prove (67) we have

$$\begin{aligned} G\mu(x) &\le \sum_{i=1}^m \int_{V_i} \psi(\| x-y \|)\,\mu(dy) \\ &\le \sum_{i=1}^m \int_{V_i} \psi(\| x_i-y \|)\,\mu(dy) \\ &\le \sum_{i=1}^m G\mu(x_i) \le m \sup_{y\in F} G\mu(y). \end{aligned}$$

Recall (see Remark 1) that (67) implies the maximum principle. Therefore, the proof is complete. ∎

Example 4: Using the above proposition we see the kernels

$$G(x,y) = \| x - y \|^{-n+\alpha},$$

$0 < \alpha < n$, $x, y \in \mathbf{R}^n$ satisfy the maximum principle. ∎

Example 5: Potential kernels of strong Markov processes satisfy the strong maximum principle. Let us just show this for the Green functions corresponding to the Laplacian on bounded domains. Let G be such a function and $h = Gf$ with say f bounded. Let (X_t) be the Brownian motion killed upon exit from D and T the stopping time $T = \inf\{ t\colon\ h(X_t) < \lambda \}$. Then if $h(x) > \lambda$,

$$\begin{aligned} h(x) &= \lambda + E^x \left[\int_0^T f(X_t) dt \right] \\ &= \lambda + E^x \left[\int_0^T f(X_t) 1_{\{ h \geq \lambda \}}(X_t)\, dt \right] \\ &\leq \lambda + E^x \left[\int_0^T f^+(X_t) 1_{\{ h \geq \lambda \}}(X_t)\, dt \right] \\ &\leq \lambda + G[f^+ 1_{\{ h \geq \lambda \}}], \end{aligned}$$

which gives the desired result. ∎

The following example gives a method of constructing a class of kernels satisfying the maximum principle.

Example 6: Let $v(t,x)$ be a semigroup of maps on $\mathbf{R}_+$ into $\mathbf{R}^n$ such that

$$\begin{aligned} & v(t,x) \in \mathbf{R}^n, \quad 0 \leq t < +\infty, \\ & v(0,x) = x, \\ & v(t+s,x) = v[t, v(s,x)]. \end{aligned} \tag{68}$$

Let μ be a potential kernel of a convolution semigroup on $\mathbf{R}_+$. This means there is a convolution semigroup (μ_t) of probability measures on $\mathbf{R}_+$ such that $\int_0^\infty \mu_t dt = \mu$. Define

$$Tf(x) = \int_0^\infty f(v(t,x))\, \mu(dt). \tag{69}$$

We claim T satisfies the strong maximum principle. To show this let h denote the function Tf. Using the semigroup property (68) we have

$$h(v(s,x)) = \int_0^\infty f[v(t+s,x)]\,\mu(dt). \tag{70}$$

If we write $g(s) = f[v(s,x)]$ and $U(t) = h(v(t,x))$, (70) may be written $U = g * \mu$, i.e., U is the potential of g. We infer $U \le \lambda + \left(g\,1_{\{U \ge \lambda\}}\right) * \mu$. This last inequality may be rewritten to get the maximum principle for T.

Two instances of the above example occurs in the operators P_n and Q_n considered in G. Sinnamon [17]. P_n and Q_n are defined as:

$$P_n f(x) = \int_0^1 f(\lambda x) d\lambda, \tag{71}$$

$$Q_n f(x) = \int_1^\infty f(\lambda x)\frac{d\lambda}{\lambda}. \tag{72}$$

The change of variables $\lambda \to \exp(-t)$ in the integral defining P_n and $\lambda \to \exp(t)$ in the integral defining Q_n give

$$P_n f(x) = \int_0^\infty f(\mathrm{e}^{-t}x)\mathrm{e}^{-t}dt,$$

$$Q_n f(x) = \int_0^\infty f(\mathrm{e}^{t}x)dt;$$

dt and $\exp(-t)dt$ are both potential kernels. ∎

One other method of obtaining kernels satisfying the maximum principle is the following:

Example 7: Let $\mathbf{G}$ be a group which acts on X. Suppose μ is a potential kernel of a convolution semigroup on $\mathbf{G}$ (see [5] for the examples). Similarly as in the previous example this means that there is a convolution semigroup of probability measures (μ_t) on $\mathbf{G}$ such that $\mu = \int_0^\infty \mu_t dt$. Define for f on X, Tf by

$$Tf(x) = \int_{\mathbf{G}} f(gx)\,\mu(dg). \tag{73}$$

We claim T satisfies the maximum principle with constant 1. For the proof let $\lambda > 0$. Fix $\xi \in X$. For each $h \in \mathbf{G}$:

$$h \to Tf(h\xi) = \int_{\mathbf{G}} f(gh\xi)\,\mu(dg) = \tilde{f} * \mu$$

where $\tilde{f}(h_1) = f(h_1\xi)$, $h_1 \in \mathbf{G}$. Thus, for fixed ξ, $h \to Tf(h\xi)$ is the potential of $\tilde{f}$. Hence

$$\begin{aligned} Tf(h\xi) &\leq \lambda + \int_{\mathbf{G}} \tilde{f}(g)\, 1_{\{Tf(g\xi)\geq\lambda\}}\,\mu(dg) \\ &= \lambda + \int_{\mathbf{G}} f(g\xi)\, 1_{\{Tf(g\xi)\geq\lambda\}}\,\mu(dg) \\ &= \lambda + T\left[f\, 1_{\{Tf\geq\lambda\}}\right] \end{aligned}$$

which finishes the proof. ∎

References

[1] K. F. Andersen and E. T. Sawyer, *Weighted norm inequalities for the Riemann–Liouville and Weyl fractional integral operators,* Trans. Amer. Math. Soc. **308** (1988), 547–558.

[2] M. A. Ariño and B. Muckenhoupt, *Maximal functions on classical Lorentz spaces and Hardy's inequality with weights for nonincreasing functions,* Trans. Amer. Math. Soc. **320** (1990), 727–735.

[3] P. R. Beesack, *Hardy's inequality and its extensions,* Pacific J. Math. **11** (1961), 39–61.

[4] P. R. Beesack, *Integral inequalities involving a function and its derivatives,* Amer. Math. Monthly **78** (1971), 705–741.

[5] C. Berg and G. Forst, *Potential Theory on Locally Compact Abelian Groups,* Springer-Verlag, Berlin–Heidelberg, 1975.

[6] G. A. Bliss, *An integral inequality,* J. London Math. Soc. **5** (1930), 40–46.

[7] J. S. Bradley, *Hardy inequalities with mixed norms,* Canad. Math. Bull. **21** (1978), 405–408.

[8] G. H. Hardy, *Note on a theorem of Hilbert,* Math. Zeitschr. **6** (1920), 314–317.

[9] G. Hardy, G., J. E. Littlewood and G. Pólya, *Inequalities,* 1st paperback edition, Cambridge University Press, Cambridge–New York, 1989.

[10] V. M. Kokilashvili, *On Hardy's inequalities in weighted spaces* (Russian), Soobšč. Akad. Nauk Gruzin. SSR (1) **96** (1979), 37–40.

[11] A. Kufner and B. Opic, *Hardy-type Inequalities,* Longman Scientific, Essex, 1990.

[12] A. Kufner and H. Triebel, *Generalizations of Hardy's inequality,* Confer. Sem. Mat. Univ. Bari **156** (1978), 1–21.

[13] V. G. Maz'ja, *Einbettungssätze für Sobolewsche Räume 1,* Teubner, Leipzig, 1979.

[14] V. G. Maz'ja, *Sobolev Spaces,* Springer-Verlag, Berlin–Heidelberg, 1985.

[15] B. Muckenhoupt, *Hardy's inequality with weights*, Studia Math. **44** (1972), 31–38.

[16] M. Rao, *Brownian Motion and Classical Potential Theory*, Lecture Notes Series No.47. Aarhus University, 1977.

[17] G. Sinnamon, *A weighted gradient inequality*, Proc. Roy. Soc. Edin. **111A** (1989), 329–335.

[18] V. D. Stepanov, *Weighted Hardy inequality*, Siber. Math. J. **28** (1987), 515–517.

[19] G. Talenti, *Una diseguaglianza integrale*, Boll. Un. Mat. Ital. (3) **21** (1966), 25–34.

[20] G. Talenti, *Osservazioni sopra una classe di diseguaglianze*, Rend. Sem. Mat. e Fis. di Milano **39** (1969), 171–185.

[21] G. Tomaselli, *A class of inequalities*, Boll. Un. Mat. Ital.(4) **2** (1969), 622–631.