Illinois Journal of Mathematics
Volume 45, Number 3, Fall 2001, Pages 1021–1050
S 0019-2082

HARMONIC FUNCTIONS ON METRIC SPACES

NAGESWARI SHANMUGALINGAM

Abstract. This paper explores a Dirichlet type problem on metric measure spaces. The problem is to find a Sobolev-type function that minimizes the energy integral within a class of "Sobolev" functions that agree with the boundary function outside the domain of the problem. This is the analogue of the Euler-Lagrange formulation in the classical Dirichlet problem. It is shown that, under certain geometric constraints on the measure imposed on the metric space, such a solution exists. Under the condition that the space has many rectifiable curves, the solution is unique and satisfies the weak maximum principle.

1. Introduction

The classical Dirichlet boundary value problem arises from a partial differential equation; if Ω is a domain in $\mathbb{R}^n$ and $f : \partial\Omega \to \mathbb{R}$ is a continuous map, the problem is to find a continuous function u so that $\Delta u = 0$ on Ω, and $u = f$ on $\partial\Omega$. The function f is called the boundary value of u. Here $\Delta u = 0$ means that for every smooth function φ on $\mathbb{R}^n$ with compact support in Ω,

$$\int_\Omega \nabla u \cdot \nabla \varphi = 0,$$

where ∇u is the distributional gradient of u. By Weyl's lemma, such a function u is of class C^2, and the above equation is equivalent to $\Delta u = 0$ in the classical sense.

The more general (non-linear) Dirichlet problem corresponding to the index p, $1 < p < \infty$, is to find a function u, continuous on Ω, so that

$$\nabla \cdot (|\nabla u|^{p-2} \nabla u) = 0$$

on the domain Ω and $u = f$ on $\partial\Omega$. The above classical problem corresponds to the general problem when $p = 2$.

In the general setting of metric measure spaces, an alternate way of stating the general Dirichlet problem is necessary. Such an alternate problem is the

Received August 29, 2000; received in final form January 18, 2001.
2000 *Mathematics Subject Classification.* 31C05, 49J40.

energy minimizing problem, which is equivalent to the general Dirichlet problem in the Euclidean setting via the Euler-Lagrange equations. The general Dirichlet problem in this paper covers a wide variety of examples, from the well-studied Euclidean space setting to the more recently discovered spaces of Laakso. The Dirichlet problems on domains in Euclidean spaces and the Carnot-Carathéodory spaces studied in the papers [GN], [FLW], and [HK2, Section 11] are covered by this paper. However, the results in this paper cover even more spaces, such as the spaces of Laakso [L] and the spaces of Bordon and Pajot [BP]. These spaces are not as well-studied as the Carnot-Carathéodory spaces, and in this setting the results of this paper are new. The uniqueness and the weak maximum principle results of this paper are applicable to the Dirichlet problems on domains in MEC_p-spaces, of which Riemannian spaces are examples. This paper thus attempts to unify results about p-harmonic functions on such diverse spaces.

The two main ingredients in this problem are the Sobolev type function spaces, called the Newtonian spaces, and the energy operator. The theory of Newtonian spaces was developed in [Sh1] and [Sh2], and related topics can be found in [C], [H1], [H2], [HK1], [HK2], [HeK1], [KM1], [KM2], [KSh], and references therein. The aim of this paper is to set up the Dirichlet boundary value problem, and to explore some properties of solutions to such problems. It is shown that if the solution is unique, then it satisfies a weak maximum principle. It is also shown that a solution to a Dirichlet problem on a domain is a solution to a Dirichlet problem on every subdomain. Similar results have been obtained by Cheeger [C] when the domain of harmonicity and the boundary function f are bounded. The proof given in [C] is different from the proof given in this paper. The approach taken in the proofs given here provides a more geometric description of the concepts employed.

The second section of this paper catalogs the needed definitions, and the third section explores some properties of weak upper gradients. Weak upper gradients replace the role of gradients in the following arguments. The fourth section explores the relationships between alternative definitions of Newtonian spaces with zero boundary values. In the fifth section the existence of solutions is established under certain conditions on the metric measure space, and the uniqueness of such solutions is also explored. The last section investigates the maximum principle property of such solutions.

This paper is based on the last chapter of the author's Ph.D. thesis [Sh1].

2. Definitions

This section develops the definitions necessary for the rest of the paper.

Let X be a metric measure space equipped with a complete metric d and a Borel regular measure μ. It is assumed throughout this paper that the measure of every open set is positive, and that the measure of every bounded set is finite. Let $1 \le p < \infty$. The following definitions are from [HeK1], [HK2], and [Sh2].

Paths γ in X are continuous maps $\gamma : I \to X$, where I is some interval in $\mathbb{R}$; abusing terminology, the image $|\gamma| := \gamma(I)$ of γ is also called a path. For a discussion of rectifiable paths and path integration see [HeK1, Section 2] or [V, Chapter 1].

DEFINITION 2.1. Let u be a real-valued function on X. A non-negative Borel-measurable function ρ is said to be an *upper gradient* of u if, for all compact rectifiable paths γ,

$$|u(x) - u(y)| \le \int_\gamma \rho ds, \tag{1}$$

where x and y denote the endpoints of the path.

See [He], [KM2] and [HeK1, Section 2.9] for a discussion on upper gradients; [HeK1] uses the term *very weak gradients* for this concept.

The following definition is applicable to all families of paths, not just collections of compact rectifiable paths. In the remainder of the paper, however, we consider only families of non-constant compact rectifiable paths.

DEFINITION 2.2. Let Γ be a collection of paths in X. The *p-modulus* of the family Γ, denoted $\mathrm{Mod}_p\Gamma$, is defined to be the number

$$\inf_\rho \|\rho\|_{L^p}^p,$$

where the infimum is taken over the set of all non-negative Borel-measurable functions ρ such that for all rectifiable paths γ in Γ the path integral $\int_\gamma \rho ds$ is not smaller than 1. Such functions ρ are said to be *admissible* for the family Γ.

It is known from [Fu1] that the p-modulus is an outer measure on the collection of all paths in X. For additional information about p-moduli see [AO], [Fu1, Chapter 1], [Hs], [KSh], and [V]. A property relevant to paths in X is said to hold for *p-almost all paths* if the family of rectifiable compact paths on which the property does not hold has zero p-modulus. The paper [Fu1] shows that the p-modulus of a curve family is zero if, and only if, there is a non-negative, p-integrable, Borel function ρ so that for every rectifiable curve γ in this family, $\int_\gamma \rho = \infty$; see [Fu1, Theorem 2].

An extended real-valued Borel function f on X is said to be *p-integrable* if $\int_X |f|^p d\mu < \infty$.

DEFINITION 2.3. Let u be an arbitrary real-valued function on X, and let ρ be a non-negative Borel function on X. If inequality (1) holds for p-almost all paths γ, then ρ is said to be a *p-weak upper gradient* of u.

Definition 2.4. Let the set $\tilde{N}^{1,p}(X,d,\mu)$ be the collection of all real-valued p-integrable functions u on X that have a p-integrable p-weak upper gradient.

If u is a function in $\tilde{N}^{1,p}$, let

$$\|u\|_{\tilde{N}^{1,p}} = \|u\|_{L^p} + \inf_{\rho} \|\rho\|_{L^p},$$

where the infimum is taken over all p-integrable p-weak upper gradients of u. If u, v are functions in $\tilde{N}^{1,p}$, let $u \sim v$ if $\|u-v\|_{\tilde{N}^{1,p}} = 0$. Note that $\sim$ is an equivalence relation, partitioning $\tilde{N}^{1,p}$ into equivalence classes.

Definition 2.5. The *Newtonian space* corresponding to the index p, $1 \leq p < \infty$, is the normed space $\tilde{N}^{1,p}(X,d,\mu)/\sim$, with norm $\|u\|_{N^{1,p}} := \|u\|_{\tilde{N}^{1,p}}$, and is denoted $N^{1,p}(X)$.

It is shown in [Sh1] and [Sh2] that these Newtonian spaces are Banach spaces. The paper [C] has another definition of Sobolev spaces, and [Sh1] and [Sh2] show that when $p > 1$ the definition of [C] is isometrically equivalent to the definition given above.

Definition 2.6. A function u is said to be ACC_p or *absolutely continuous on p-almost every curve* if $u \circ \gamma$ is absolutely continuous on $[0, l(\gamma)]$ for p-almost every rectifiable arc-length parametrized path γ in X. Here $l(\gamma)$ denotes the length of γ.

In [Sh1] and [Sh2] it was shown that functions in $N^{1,p}(X)$ are ACC_p.

A more sensitive alternative to measure can be defined on subsets of X using the Newtonian spaces. Measure is the natural gauge for defining $L^p(X)$ functions; two functions in $L^p(X)$ are in the same L^p class if and only if the set on which they disagree is of measure zero. A finer gauge is needed to explore $N^{1,p}(X)$ function classes. One of the possible gauges for this property is p-capacity. Different definitions for a capacity of a set can be found in literature. For more information on p-capacity, see [AO], [KM1], [Sh1], [Sh2], and [KSh].

Definition 2.7. The *p-capacity* of a set $E \subset X$ with respect to the space $N^{1,p}(X)$ is defined by

$$\mathrm{Cap}_p E = \inf_u \|u\|^p_{N^{1,p}}, \tag{2}$$

where the infimum is taken over all the functions u in $N^{1,p}(X)$ whose restriction to E is bounded below by 1.

It is easy to see that sets of zero p-capacity have zero measure. In general, sets of zero measure need not have zero p-capacity. See [KM1] for more on capacity properties. It is shown in [Sh1] and [Sh2] that if u_1 and u_2 are two functions in $N^{1,p}(X)$ so that $\|u_1 - u_2\|_{L^p} = 0$, then $u_1 \sim u_2$ and hence both functions belong to the same equivalence class in $N^{1,p}(X)$. It is also easy to see in this case that the p-capacity of the set where u_1 and u_2 differ is zero.

DEFINITION 2.8. A property is said to hold *p-quasi everywhere,* or p-q.e., if that property holds for all points outside a set of zero p-capacity. A function is said to be *p-quasi continuous* if for every $\epsilon > 0$ there are sets F_ϵ, of p-capacity smaller than ϵ, so that the function is continuous outside of F_ϵ.

DEFINITION 2.9. The space X supports a *weak* $(1,p)$*-Poincaré inequality* if there are positive constants C, τ, so that for all open balls B in X and all pairs of functions u and ρ defined on τB, whenever ρ is an upper gradient of u in τB and u is integrable on B, then

$$\fint_B |u - u_B| \le C \operatorname{diam}(B) \left(\fint_{\tau B} \rho^p \right)^{1/p}, \tag{3}$$

where, for any a measurable function f on X,

$$f_B := \frac{1}{\mu(B)} \int_B f =: \fint_B f.$$

The space X supports a $(1,p)$*-Poincaré inequality* if it supports a weak $(1,p)$-Poincaré inequality with $\tau = 1$.

The Euclidean spaces and John domains support a $(1,p)$-Poincaré inequality, while quasi-discs in general support only a weak $(1,p)$-Poincaré inequality. Spaces without rectifiable curves, such as the Koch snowflake, do not support even a weak $(1,p)$-Poincaré inequality. For more on Poincaré inequalities, see [HeK1], [HeK2], [HK2], and the references therein. In the rest of this paper, the results requiring X to support a $(1,p)$-Poincaré inequality remain true if X supports only the weak version of this inequality, but for the sake of simplifying notation only the $(1,p)$-Poincaré inequality is assumed in those cases.

DEFINITION 2.10. A measure μ on a metric measure space is said to be *doubling* if there is a constant $C \ge 1$ so that for all x in X and all $r > 0$,

$$\mu(B(x,2r)) \le C\mu(B(x,r)).$$

REMARK 2.11. The paper [C] shows that if the measure on X is doubling and supports a $(1,p)$-Poincaré inequality with $1 < p < \infty$, then $N^{1,p}(X)$ is reflexive; see Section 5 for a use of this fact.

The following theorem gives a condition under which Newtonian functions are quasi continuous. The proof of the theorem is a modification of an idea due to Semmes [S] and is a generalization of [Sh2, Theorem 4.1]. The paper [C] proves a weaker version of this theorem, namely, that under the assumptions of the following theorem locally Lipschitz functions are dense in $N^{1,p}(X)$. The proof given in [C] uses a Calderon-Zygmund decomposition type argument applied to balls. The paper [HK2] proves a version of the following theorem for another Sobolev-type space, which yields the same Newtonian space provided X supports a $(1,q)$-Poincaré inequality for some $1 \leq q < p$.

THEOREM 2.12. *If X is a doubling space supporting a $(1,p)$-Poincaré inequality, then Lipschitz functions are dense in $N^{1,p}(X)$.*

The proof of this theorem uses the following lemmas.

LEMMA 2.13. *Let X be a doubling space, and let M^* be the non-centered maximal operator defined by*

$$M^* f(x) := \sup_B \fint_B |f| d\mu, \tag{4}$$

where the supremum is taken over balls B in X containing the point x. If g is a function in $L^1(X)$, then

$$\lim_{\lambda \to \infty} \lambda \ \mu(\{x \in X : M^* g(x) > \lambda\}) = 0.$$

The proof of this lemma is similar to the proof in the Euclidean case and can also be found in [Sh1].

Fix $x_0 \in X$, and for each positive integer j consider the function

$$\eta_j(x) = \begin{cases} 1 & \text{if } d(x_0,x) \leq j-1, \\ j - d(x_0,x) & \text{if } j-1 < d(x_0,x) < j, \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that this function is 1-Lipschitz.

LEMMA 2.14. *Let u be a function in $N^{1,p}(X)$. Then the function $v_j = u\eta_j$ is also in $N^{1,p}(X)$. Furthermore, the sequence v_j converges to u in $N^{1,p}(X)$.*

Proof. Let x, y be two points in X. Then,

$$\begin{aligned} |v_j(x) - v_j(y)| &\leq |\eta_j(x)u(x) - \eta_j(x)u(y)| + |\eta_j(x)u(y) - \eta_j(y)u(y)| \\ &\leq |u(x) - u(y)| + d(x,y)(|u(x)| + |u(y)|). \end{aligned}$$

Hence, by Lemma 3.1 of the next section, if g is an upper gradient of u, then $g + 4|u|$ is also an upper gradient of v_j. By Lemma 3.3 of the next section the function $g_j := (g + 4|u|)\chi_{B(x_0,j)}$ is a weak upper gradient of v_j

and $(g+g_j)\chi_{X\setminus B(x_0,j-1)}$ is an upper gradient of $u-v_j$. Now,

$$\|u-v_j\|_{L^p(X)} = \left(\int_{X\setminus B(x_0,j-1)} |u-v_j|^p\right)^{1/p} \le 2\left(\int_{X\setminus B(x_0,j-1)} |u|^p\right)^{1/p} \to 0,$$

since u is p-integrable. Moreover,

$$\left(\int_X [(g+g_j)\chi_{X\setminus B(x_0,j-1)}]^p\right)^{1/p} \le 2\left(\int_{X\setminus B(x_0,j-1)} g^p\right)^{1/p} + 4\left(\int_{X\setminus B(x_0,j-1)} |u|^p\right)^{1/p},$$

which tends to 0 as $j\to\infty$ since the two functions g and $|u|$ are p-integrable. Now it can be concluded that $\|u-v_j\|_{N^{1,p}(X)}\to 0$ as j tends to infinity. □

Proof of Theorem 2.12. Let u be a function in $N^{1,p}(X)$. By the above lemma, without loss of generality it can be assumed that u is zero outside of a bounded set. Let

$$E_\lambda = \{x\in X : M^* g^p(x) > \lambda^p\},$$

where g is a p-integrable upper gradient of u. Then by Lemma 2.13,

$$\lambda^p \mu(E_\lambda) \to 0 \qquad \text{as} \qquad \lambda\to\infty. \tag{5}$$

If x is a point in $X\setminus E_\lambda$, then by the same argument as in the proof of [Sh2, Theorem 4.1] it is seen that for all $r,s>0$,

$$|u_{B(x,s)} - u_{B(x,r)}| \le C\lambda r.$$

Hence any sequence $u_{B(x,r_i)}$ is a Cauchy sequence in $\mathbb{R}$. Therefore, on $X\setminus E_\lambda$ we can define

$$u_\lambda(x) := \lim_{r\to 0} u_{B(x,r)}.$$

Since the measure is doubling, almost every point in X is a Lebesgue point of u; see [Ma, Theorem 2.12] or [He]. Note that at Lebesgue points of u in $X\setminus E_\lambda$ we have $u_\lambda = u$, and that E_λ is an open set. Hence $u-u_\lambda$ satisfies the hypotheses of Lemma 3.3. By the same argument as in the proof of [Sh2, Theorem 4.1], if x and y are in $X\setminus E$, then,

$$|u_\lambda(x) - u_\lambda(y)| \le \sum_{i=-\infty}^{\infty} |u_{B_i} - u_{B_{i+1}}| \le C\lambda d(x,y).$$

Hence u_λ is $C\lambda$-Lipschitz on $X\setminus E_\lambda$. Extend u_λ as a $C\lambda$-Lipschitz extension to the entire space X; see [MS] for existence of such extensions. Choose an extension such that u_λ is bounded by $2C\lambda$. This can be done by truncating any Lipschitz extension at $C\lambda$. Such truncation will not affect the values of u_λ on the set $X\setminus E_\lambda$ for large enough λ, since u itself is zero outside a bounded set and hence the non-zero values of u_λ in $X \backslash E_\lambda$ lie within a bounded set which is independent of λ, and u_λ is $C\lambda$-Lipschitz.

Now, as in the proof of [Sh2, Theorem 4.1], we see that $u_\lambda \to u$ in the norm of $N^{1,p}(X)$. □

The following corollary follows from [Sh2, Corollary 3.9] and the above theorem.

Corollary 2.15. *If X is a doubling metric measure space supporting a $(1,p)$-Poincarè inequality, then for each function u in $N^{1,p}(X)$ there are open sets of arbitrarily small capacity such that u is continuous in the complement of these sets; that is, u is p-quasicontinuous.*

Definition 2.16. Let ρ be a p-integrable non-negative Borel function in X. For $x, y \in X$, define a relation $x \sim_\rho y$ if either $y = x$ or there exists a compact rectifiable path γ connecting x to y such that $\int_\gamma \rho ds < \infty$. Note that this is indeed an equivalence relation, and $\sim_\rho$ partitions X into equivalence classes.

A metric measure space X is said to admit the *Main Equivalence Class* property with respect to p, or MEC_p, if each p-integrable non-negative Borel function ρ generates an equivalence class G_ρ, called the *main equivalence class* of ρ, such that $\mu(X \backslash G_\rho) = 0$.

It was shown in [O] that $\mathbb{R}^n$ has the MEC_p-property for all p. In general equivalence classes may not be measurable sets. However, in MEC_p spaces, the main equivalence class, being of full measure, is necessarily measurable, and so are the other equivalence classes. See [Sh1] and [Sh2] for more on MEC_p spaces.

If $X = \mathbb{R}^n$ and the gradient of a Sobolev function is zero a.e., then that function is constant. This is not true in general; if X has no rectifiable curve, then every measurable function has 0 as a weak upper gradient, but not every measurable function is constant. If X is an MEC_p space and 0 is a p-weak upper gradient of a function u, then u is a constant function; see the proof of Theorem 5.6.

A metric space X is said to be φ-*convex* if there is a cover of X by open sets $\{U_\alpha\}$ together with homeomorphisms $\{\varphi_\alpha : [0,\infty) \to [0,\infty)\}$, such that each pair of distinct points x and y in U_α can be joined by a curve whose length does not exceed $\varphi_\alpha(d(x,y))$. The paper [HeK1] shows that if X is proper (that is, X is closed and bounded sets are compact), φ-convex, and Q-regular, then X is Q-Loewner if and only if X supports a $(1,Q)$-Poincaré inequality. Hence, by [Sh2, Theorem 6.2], if X is proper, φ-convex, Q-regular, and supports a $(1,Q)$-Poincaré inequality, then X is an MEC_Q space. The following theorem is a generalization of this result. Note that the proof of this theorem is different from the argument for [Sh2, Theorem 6.2]. The proof of [Sh2, Theorem 6.2] uses a path-family argument based on the Loewner property, and hence cannot be generalized to prove the theorem below.

THEOREM 2.17. *Let X be a proper, φ-convex metric measure space supporting a $(1,p)$-Poincaré inequality. Then X is an MEC_p space.*

Proof. Since X is proper and φ-convex, [KSh, Lemma 3.2 and Remark 3.3] shows that $\mathrm{Mod}_p(E,F) = N^{1,p}_{\mathrm{loc}} - \mathrm{Cap}_p(E,F)$ if E and F are disjoint compact subsets of X. Here $\mathrm{Mod}_p(E,F)$ is the p-modulus of the collection of curves connecting E to F, and $N^{1,p}_{\mathrm{loc}} - \mathrm{Cap}_p(E,F)$ is the infimum of $\|\rho\|^p_{L^p}$ over all functions ρ that are p-integrable upper gradients of functions $u \in L^p_{\mathrm{loc}}(X)$ such that $u\big|_E = 1$ and $u\big|_F = 0$.

Suppose X is not an MEC_p space. Then there exists a non-negative Borel measurable p-integrable function ρ on X and disjoint subsets A, B of X such that $X = A \cup B$, $\mu(A) > 0$, $\mu(B) > 0$, and for each point x in A and each point y in B the points x and y cannot be connected by a path γ such that $\int_\gamma \rho$ is finite. Hence, by [Fu1, Theorem 2] (see Definition 2.2), the p-modulus of the collection of curves joining A to B is zero. Since the measure is Borel regular, there exist closed and bounded (and hence compact) sets $E \subset A$ and $F \subset B$ whose measures are positive. As $\mathrm{Mod}_p(E,F) \le \mathrm{Mod}_p(A,B)$, the p-modulus of curves connecting the compact sets E and F is zero. Therefore $N^{1,p}_{\mathrm{loc}} - \mathrm{Cap}_p(E,F) = 0$. Hence for each positive integer n there is a function u_n and its upper gradient ρ_n so that $\|\rho_n\|^p_{L^p} < 2^{-n}$ and $u_n \in L^p_{\mathrm{loc}}(X)$ with $u_n\big|_E = 1$, $u_n\big|_F = 0$, and $0 \le u_n \le 1$. But there exists a ball B containing $E \cup F$, and on this ball,

$$0 < \frac{\mu(E)}{\mu(B)} \le u_{n_B} = \fint_B u_n \le \frac{\mu(B) - \mu(F)}{\mu(B)} < 1.$$

Therefore

$$\fint_B |u_n - u_{n_B}| \ge \frac{\mu(F)\mu(E)}{\mu(B)^2} > 0,$$

but

$$\left(\fint_B \rho_n^p \right)^{1/p} \le \frac{2^{-n/p}}{\mu(B)^{1/p}} \to 0 \quad \text{as } n \to \infty.$$

Therefore X cannot support a $(1,p)$-Poincaré inequality. $\square$

REMARK 2.18. In the above theorem the hypothesis of φ-convexity was used only to show that $\mathrm{Mod}_p(E,F) = N^{1,p}_{\mathrm{loc}} - \mathrm{Cap}_p(E,F)$. Hence this hypothesis can be replaced by the hypothesis that if E and F are disjoint compact subsets of X, then $\mathrm{Mod}_p(E,F) = N^{1,p}_{\mathrm{loc}} - \mathrm{Cap}_p(E,F)$. As mentioned in [HeK1] and in [KSh], whether this hypothesis holds for all metric measure spaces is not known. However, the above theorem holds even when the condition of φ-convexity is replaced by the more easily verifiable conditions of path-connectedness and doubling in measure, since doubling, proper, and path-connected metric measure spaces supporting a $(1, p)$-Poincaré inequality are quasiconvex and are a fortiori φ-convex; see [HK2].

The converse of the above theorem does not hold true. The following example shows that the MEC_p condition does not imply the presence of a $(1, p)$-Poincaré inequality.

Example 2.19. Let X be the open unit ball in $\mathbb{R}^2$ with the radial segment $(0, 1) \times \{0\}$ removed. Then X, equipped with the Euclidean metric and Lebesgue measure, is a domain in $\mathbb{R}^2$ and hence, by [Sh1, Lemma 4.2.10], is an MEC_p space for all finite $p \geq 1$. However, X does not support a $(1, p)$-Poincaré inequality.

In the above example X is not φ-convex, but X can be modified to be φ-convex by including the rational points on the radius $(0, 1) \times \{0\}$. This modified space is still MEC_p since the collection of new points has measure zero. For $1 \leq p \leq 2$ the collection of paths in $\mathbb{R}^2$ that pass through any given point has p-modulus zero. Hence the collection of paths in X that pass through these rational points on the radius has modulus zero. Hence the ball mentioned in the above example does not support a Poincaré inequality for the same reason as above.

3. Some properties of weak upper gradients

This section is devoted to exploring some crucial properties of weak upper gradients. Recall that the gradient of a Sobolev function u on Euclidean spaces has the property of being zero almost everywhere on sets $\{x : u(x) = \text{constant}\}$. This property, shared by weak upper gradients, is called the truncation property in [HK2], and enables one to "paste" two Newtonian functions along a set where they are equal; see Lemma 6.5. For more uses of this property, see [KiSh].

The proof of [Sh2, Lemma 4.7] yields the following lemma.

Lemma 3.1. *If u is a function on X such that there exist non-negative Borel-measurable functions g, h on X with the property that*

$$|u(x) - u(y)| \leq \int_\gamma g ds + d(x, y)(h(x) + h(y))$$

whenever γ is a compact rectifiable path in X with end points x, y, then $g + 4h$ is an upper gradient of u.

LEMMA 3.2. *Let u_1 and u_2 be ACC_p functions on X, with weak upper gradients g_1 and g_2 respectively. If u is another ACC_p function in X such that there is an open set $O \subset X$ with the property that $u = u_1$ on O and $u = u_2$ on $X \backslash O$, then $g_1\chi_O + g_2$ and $g_1 + g_2\chi_{X\backslash O}$ are weak upper gradients of u.*

Proof. Let Γ be the collection of paths γ on which at least one of u, u_1, and u_2 is not absolutely continuous, or for which on some subpath of γ the upper gradient inequality (1) is not satisfied by at least one of (u_1, g_1) and (u_2, g_2). Then $\mathrm{Mod}_p\Gamma = 0$. First consider $h = g_1\chi_O + g_2$. Suppose γ is a compact rectifiable path not in Γ connecting points x, y in X. If $|\gamma| \subset O$ or if $x, y \in X\backslash O$, then clearly $|u(x) - u(y)| \leq \int_\gamma h$. Suppose $|\gamma| \not\subset O$ and that x is in O. If y also belongs to O, it is possible to break γ into two pieces at a point $z \in |\gamma| \cap (X\backslash O)$, and consider the two pieces separately. Hence, without loss of generality it can be assumed that y is not in O. Since O is open, $\gamma^{-1}O$ is an open subset of $I = [a, b]$, where $\gamma : I \to X$. Therefore $\gamma^{-1}O$ is a countable disjoint union of relatively open intervals in I. Thus there is a number $a_1 \in I$ such that x is in $\gamma([a, a_1)) \subset O$ and $\gamma(a_1) \in X\backslash O$, $a < a_1 \leq b$. On $\gamma([a, a_1))$, by hypothesis $u = u_1$. Since u, u_1, and u_2 are continuous on γ and $\gamma(a_1) \in X\backslash O$, and on $X\backslash O$, $u = u_2$, we have $u(\gamma(a_1)) = u_1(\gamma(a_1)) = u_2(\gamma(a_1))$. Thus

$$|u(\gamma(a)) - u(\gamma(a_1))| \leq \int_{\gamma([a,a_1))} g_1 = \int_{\gamma([a,a_1))} g_1\chi_O.$$

Hence

$$\begin{aligned}|u(x) - u(y)| &\leq |u(x) - u(\gamma(a_1))| + |u(\gamma(a_1)) - u(y)| \\ &\leq \int_{\gamma([a,a_1))} g_1\chi_O + \int_{\gamma([a_1,b])} g_2 \leq \int_\gamma h.\end{aligned}$$

Therefore h is a weak upper gradient of u, since $\mathrm{Mod}_p\Gamma = 0$.

Now consider $g = g_1 + g_2\chi_{X\backslash O}$. If either $|\gamma| \subset X\backslash O$ or if $x, y \in O$, then clearly $|u(x) - u(y)| \leq \int_\gamma g$. Suppose $|\gamma| \cap O$ is not empty and that $x \in X\backslash O$. As before, without loss of generality it can be assumed that y is in O. Let $\gamma : I = [a, b] \to X$. The set $\partial(\gamma^{-1}(O))$ corresponds to some of the points on $|\gamma|$ that lie on the boundary of O. As a closed subset of the compact set I, this set is compact, and hence has a point a_1 such that there does not exist a number $t \in [a, a_1)$ with $t \in \partial(\gamma^{-1}(O))$; that is, $\gamma([a, a_1])$ does not intersect O (if $a = a_1$, then a is on the boundary of O, and the rest of the proof remains valid in this case as well), but there is a sequence of points in $\gamma^{-1}(O)$ that converge to a_1. On $\gamma([a, a_1])$, by hypothesis $u = u_2$. Furthermore, as a_1 is a limit point of a sequence of points in $\gamma^{-1}(O)$ with $u = u_1$ on O and u is absolutely continuous on γ, we have $u(\gamma(a_1)) = u_1(\gamma(a_1))$. Hence

$$|u(\gamma(a_1)) - u(y)| \leq \int_{\gamma([a_1,b])} g_1.$$

Therefore
$$|u(x) - u(y)| \le \int_\gamma (g_1 + g_2\chi_{X\setminus O}),$$
and hence g is also a weak upper gradient of u. □

The following lemma is from [Sh2, Lemma 4.3].

LEMMA 3.3. *Suppose that u is an ACC_p function on X and that there exists an open set $O \subset X$ with $u = 0$ μ-almost everywhere on $X\setminus O$. Then if g is an upper gradient of u, then $g\chi_O$ is also a weak upper gradient of u.*

LEMMA 3.4. *Let u be an ACC_p function in X, and suppose g and h are two p-integrable weak upper gradients of u. If F is a closed subset of X, then the function $\rho = g\chi_F + h\chi_{X\setminus F}$ is also a weak upper gradient of u.*

Proof. Since both g and h are p-integrable, the collection of paths γ on which either u is not absolutely continuous, or the integral $\int_\gamma (g+h)$ is infinite, or the upper gradient relation does not hold for either g or h, has p-modulus zero. Denote this collection by Γ_0. Let Γ be the collection of all rectifiable paths that have a subpath in Γ_0. The p-modulus of Γ is zero. Let γ be a compact rectifiable path in X that does not belong to the family Γ. If γ lies entirely in F or entirely outside F, then inequality (1) is satisfied by ρ and u. Suppose that γ passes through F and goes outside F as well. Since $X\setminus F$ is open, $\gamma^{-1}(X\setminus F)$ is a countable disjoint union of open intervals in the domain of γ. Let the images of these disjoint intervals be denoted γ_i. Denoting the end points of γ_i by x_i and y_i, and noting that u is continuous on γ, by induction it can be shown that ρ satisfies inequality (1):
$$\begin{aligned}|u(x) - u(y)| \le& |u(x) - u(x_1)| + |u(x_1) - u(y_1)| + |u(y_1) - u(y)|\\ \le& \int_{\gamma\setminus\gamma_1} g + \int_{\gamma_1} h \;... \le \int_{\gamma\setminus\bigcup_{1\le i\le n}\gamma_i} g + \int_{\bigcup_{1\le i\le n}\gamma_i} h\end{aligned}$$
for each positive integer n. By Lebesgue dominated convergence theorem applied to the path-integral on γ the result follows. □

REMARK 3.5. Lemmas 3.2 and 3.3 remain true if the requirement of openness for the set O is replaced by the requirement that for p-modulus almost every compact rectifiable curve γ the set $\gamma^{-1}(O)$ is open. Any set satisfying this requirement is said to be *p-path open*. This property is closely related to the property of p-quasi open set property defined by Fuglede; see [Fu2]. A set $O \subset X$ is said to be *p-quasi open* if for each positive ϵ there is a set $F_\epsilon \subset X$ so that the p-capacity of F_ϵ is no more than ϵ and the set $O \cup F_\epsilon$ is open in the metric topology of X. Every p-quasi open set is p-path open. To see this, suppose $O \subset X$ is p-quasi open. Then, by definition, for each positive integer n there is a set F_n with $\mathrm{Cap}_p(F_n) \le 2^{-n}$ so that $O \cup F_n$ is open. It can be assumed that $F_{n_1} \subset F_{n_2}$ if $n_1 \geq n_2$. Let $u_n \in N^{1,p}(X)$ so that $u_n\big|_{F_n} = 1$, $0 \leq u_n \leq 1$, and $\|u_n\|^p_{N^{1,p}} \leq 2^{-n+1}$. Let $v_i = \sum_{n=1}^i u_n$. Then the sequence v_i is a Cauchy sequence in $N^{1,p}(X)$, and hence, by [Sh2, Theorem 3.7], this sequence converges to a function $v \in N^{1,p}(X)$ in the $N^{1,p}(X)$. Suppose γ is a rectifiable curve. Clearly $\gamma^{-1}(F_n \cup O)$ is open in the domain of γ, for each n. If there is an integer n_0 so that $\gamma^{-1}(F_{n_0} \cup O) = \gamma^{-1}(O)$, then $\gamma^{-1}(O)$ is open in the domain of γ. If there is no such integer n_0, then γ intersects F_n for each n. Let x_n be a point in this intersection. Then, as $F_m \subset F_n$ for each $m \leq n$ and $u_m\big|_{F_m} = 1$, and since by the proof of [Sh2, Lemma 3.6] the function v can be taken to be the function $v(x) = \lim_{i\to\infty} v_i(x)$, it follows that $v(x_n) \geq n$. Hence $v \circ \gamma$ is not bounded on γ, and so v is not absolutely continuous on γ. Since $v \in N^{1,p}(X)$, by [Sh2, Proposition 3.1] the family of curves on which function v is not absolutely continuous has p-modulus zero. Therefore the collection of such γ is of zero p-modulus.

If X has no rectifiable curves, every set is p-path open, but not every set is p-quasi open. The author does not know whether in the Euclidean setting p-path open sets are p-quasi open.

Lemma 3.4 remains true if the condition that the set F be closed is replaced by the corresponding property that for p-modulus almost every compact rectifiable curve γ the set $\gamma^{-1}(F)$ be closed.

The following lemma is from [KSh]. As a corollary, it shows the existence of a "smallest" weak upper gradient in the case $p > 1$. For a different approach see [C], which shows the reflexivity of $N^{1,p}(X)$ if X supports a Poincaré inequality and the measure on X is doubling. However, for many problems in practice one does not need reflexivity and can instead invoke Mazur's lemma. The following lemma serves nearly the same purpose as Mazur's lemma, and can be used even if $N^{1,p}(X)$ is not reflexive.

Lemma 3.6. *Let Y be a metric measure space and $p > 1$. If $\{f_j\}_{j\in\mathbb{N}}$ is a sequence of functions in $L^p(Y)$ with upper gradients $\{g_j\}_{j\in\mathbb{N}}$ in $L^p(Y)$ such that f_j converges weakly to f in L^p and g_j converges weakly to g in L^p, then g is a weak upper gradient of f and there is a sequence of convex combinations $\tilde{f}_j = \sum_{k=j}^{n_j} \lambda_{kj} f_k$ and $\tilde{g}_j = \sum_{k=j}^{n_j} \lambda_{kj} g_k$ with $\sum_{k=j}^{n_j} \lambda_{kj} = 1$, $\lambda_{kj} > 0$, so that $\tilde{f}_j$ converges in L^p to f and $\tilde{g}_j$ converges in L^p to the function g.*

See [KSh, Lemma 3.1] for a proof of the above lemma.

One of the consequences of the above lemma is that every ACC_p function with a p-integrable weak upper gradient has a minimal p-integrable weak upper gradient. This fact was proved in [C]. An alternative and more geometric proof is given below. In [C] weak upper gradients are called generalized upper gradients.

Corollary 3.7. *Let X be a metric measure space equipped with a Borel regular measure. If u is a function with a p-integrable weak upper gradient, $p > 1$, then there exists a p-integrable weak upper gradient ρ_u of u with the following property: if ρ is another weak upper gradient of u, then $\rho_u \leq \rho$ almost everywhere.*

Proof. Since $p > 1$, by the reflexivity of $L^p(X)$ and by the above lemma there is a p-integrable weak upper gradient ρ_u with the smallest L^p-norm amongst all the p-integrable weak upper gradients of u. Let ρ be another weak upper gradient of u, and let E be the set on which ρ_u is larger than ρ. Suppose the measure of E is positive. Then there is a closed subset F of E of positive measure, since the measure is Borel regular. By Lemma 3.4 the function $\rho_u \chi_{X \setminus F} + \rho \chi_F$ is a weak upper gradient, of strictly smaller L^p-norm than ρ_u, contradicting the choice of ρ_u. Hence the measure of E must be zero. □

4. Newtonian spaces with zero boundary values

In order to solve a Dirichlet boundary value problem on a set $E \subset X$, it should be possible to compare two Sobolev functions on the set $X \setminus E$. The paper [KKM] defines and analyzes some properties of the space $M_0^{1,p}(E)$ of Sobolev type functions (called the Hajłasz functions) on X whose trace on $X \setminus E$ vanishes. This paper follows their approach to this generalization.

As pointed out in [KKM], there are many approaches to defining Sobolev spaces of functions with zero boundary values. For example, one can consider the set of Lipschitz functions on X that vanish on $X \setminus E$, and close that set under an appropriate norm, or the space of all Newtonian functions u on X that are zero p-q.e. in $X \setminus E$. In this section a third space is obtained by considering the closure of the set of compactly supported Lipschitz functions with support in E. In general the three approaches yield three different spaces, but for a broad class of metric spaces it will be shown that these definitions agree.

Definition 4.1. Let E be a subset of a metric measure space X. Let $\tilde{N}_0^{1,p}(E)$ be the set of all functions u from E to $[-\infty, \infty]$ for which there exists a function $\tilde{u}$ in $\tilde{N}^{1,p}(X)$ such that $\tilde{u} = u$ μ-a.e. on E and the p-capacity of the set $\{x \in X \setminus E : \tilde{u}(x) \neq 0\}$ is zero. If u and v are two functions in $\tilde{N}_0^{1,p}(E)$, define $u \sim v$ if $u = v$ μ-a.e. on E. Note that $\sim$ is an equivalence relation. Let $N_0^{1,p}(E) = \tilde{N}_0^{1,p}(E)/\sim$, equipped with the norm

$$\|u\|_{N_0^{1,p}(E)} = \|\tilde{u}\|_{N^{1,p}(X)}.$$

The norm on $N_0^{1,p}(E)$ is unambiguously defined. Since zero p-capacity sets have zero measure (for every $A \subset X$ it is easy to see that $\mu(A) \leq \mathrm{Cap}_p(A)$), by [Sh2, Corollary 3.3], if $\tilde{u}$ and $\tilde{u}'$ both correspond to u as in the above definition, then $\|\tilde{u} - \tilde{u}'\|_{N^{1,p}(X)} = 0$.

In a similar definition in [KKM], $\tilde{u}$ is required to be p-quasicontinuous. Functions in Newtonian spaces are already p-quasicontinuous if X satisfies certain criteria; see Corollary 2.15.

DEFINITION 4.2. Let $\mathrm{Lip}_0^{1,p}(E)$ be the collection of all Lipschitz functions in $N^{1,p}(X)$ that vanish on $X \setminus E$, and let $\mathrm{Lip}_{C,0}^{1,p}(E)$ be the collection of functions in $\mathrm{Lip}_0^{1,p}(E)$ that have compact support in E. Let $H_0^{1,p}(E)$ be the closure of $\mathrm{Lip}_0^{1,p}(E)$ in the norm of $N^{1,p}(X)$, and let $H_{C,0}^{1,p}(E)$ be the closure of $\mathrm{Lip}_{C,0}^{1,p}(E)$ in the norm of $N^{1,p}(X)$.

PROPOSITION 4.3. *The space $H_{C,0}^{1,p}(E)$ embeds isometrically into $H_0^{1,p}(E)$, and $H_0^{1,p}(E)$ embeds isometrically into $N_0^{1,p}(E)$.*

Proof. Let u be a function in $H_0^{1,p}(E)$. Then there is a sequence of Lipschitz functions $\{u_i\}_{i\in\mathbb{N}}$ from $N^{1,p}(X)$ that converge to u in $N^{1,p}(X)$, and such that, for each integer i, $u_i\big|_{X\setminus E} = 0$. As in the proof of [Sh2, Theorem 3.7], passing to a subsequence if necessary, we can write

$$\tilde{u} = \frac{1}{2}\left(\limsup_i u_i + \liminf_i u_i\right)$$

outside a set B of p-capacity zero. Then $\tilde{u}$ is a function in $N^{1,p}(X)$ and $u\big|_E = \tilde{u}\big|_E$ μ-almost everywhere. Furthermore, $\tilde{u}\big|_{(X\setminus E)\setminus B} = 0$. Hence $\tilde{u}\big|_E$ is in $N_0^{1,p}(E)$, with the two norms being equal. Since $\mathrm{Lip}_{C,0}^{1,p}(E)$ is a subset of $\mathrm{Lip}_0^{1,p}(E)$, it is easy to see that $H_{C,0}^{1,p}(E)$ embeds into $H_0^{1,p}(E)$ isometrically. $\square$

The definitions of $H_0^{1,p}(E)$ and $H_{C,0}^{1,p}(E)$ guarantee that these spaces are Banach spaces. The following theorem shows that $N_0^{1,p}(E)$ is a Banach space as well.

THEOREM 4.4. *The function space $N_0^{1,p}(E)$ is complete, i.e., is a Banach space.*

Proof. Let $\{u_i\}_{i\in\mathbb{N}}$ be a Cauchy sequence in $N_0^{1,p}(E)$. Then there is a corresponding Cauchy sequence $\{\tilde{u}_i\}_{i\in\mathbb{N}}$ in $N^{1,p}(X)$, where $\tilde{u}_i$ is the function in $N^{1,p}(X)$ corresponding to u_i as in the definition of $N_0^{1,p}(E)$. Then, as $N^{1,p}(X)$ is a Banach space, there exists a function $\tilde{u}$ in $N^{1,p}(X)$, and a subsequence, also denoted by $\{\tilde{u}_i\}_{i\in\mathbb{N}}$ for the sake of brevity, so that, as in the proof of [Sh2, Theorem 3.7], $\tilde{u}_i$ converges to $\tilde{u}$ pointwise outside a set B of zero p-capacity, and also in the norm of $N^{1,p}(X)$. Let A_i be the set of points in $X\setminus E$ on which $\tilde{u}_i$ is non-zero. Then the p-capacity of $\cup_i A_i$ is zero. Moreover, on $(X\setminus E)\setminus(\cup_i A_i \cup B)$,

$$\tilde{u}(x) = \lim_{i\to\infty} \tilde{u}_i(x) = 0.$$

Since the p-capacity of $B\cup(\bigcup_i A_i)$ is zero, the function $u = \tilde{u}\big|_E$ is in $N_0^{1,p}(E)$. Since

$$\|u_i - u\|_{N_0^{1,p}(E)} = \|\tilde{u}_i - \tilde{u}\|_{N^{1,p}(X)} \to 0 \text{ as } i\to\infty,$$

the space $N_0^{1,p}(E)$ is complete. $\square$

In the rest of this paper we will make no distinction between the function u in $N_0^{1,p}(E)$ and its Newtonian extension $\tilde{u}$.

In the remainder of this section we explore the relationships between the three spaces defined above, and we give examples of spaces X and $E\subset X$ for which the three function spaces $N_0^{1,p}(E)$, $H_0^{1,p}(E)$, and $H_{C,0}^{1,p}(E)$ are different. Some of these examples are modifications of the examples in [KKM].

Example 4.5. Let $X = B$ be the unit ball in $\mathbb{R}^2$, endowed with Euclidean metric and 2-dimensional Lebesgue measure, and let E be the set of all points in B, at least one of whose coordinates is irrational. Then $B\setminus E$ is countable and hence is of zero measure. Since $B\setminus E$ is dense in B, the collections $\mathrm{Lip}_0^{1,p}(E)$ and $\mathrm{Lip}_{C,0}^{1,p}(E)$ are the same collection containing only the zero function, and therefore $H_{C,0}^{1,p}(E)$ and $H_0^{1,p}(E)$ are the same trivial space $\{0\}$. Let u be defined by

$$u(x) = \mathrm{dist}(x,\partial B) = 1-|x|.$$

It is easy to see that this function is in $N^{1,p}(B) = W^{1,p}(B)$ when $1<p\le 2$. For these values of p, the p-capacity of $B\setminus E$ is zero (as it is countable, and points have zero p-capacity; see [V] or [He]), and hence u is in $N_0^{1,p}(E)$.

Example 4.6. Let X be the open unit ball $B(0,1)$ in the Euclidean space $\mathbb{R}^n$, endowed with the Euclidean metric and Lebesgue measure. Let $E = X$. Then $H_0^{1,p}(E)$, $N_0^{1,p}(E)$, $N^{1,p}(B(0,1))$, and $W^{1,p}(B(0,1))$ are all the same function space, since Lipschitz functions are dense in $W^{1,p}(B(0,1))$. However, the function space $H_{C,0}^{1,p}(E)$ is the same space as the classical space $W_0^{1,p}(B(0,1))$, which is smaller than $W^{1,p}(B(0,1))$.

The above example shows that the spaces $H_0^{1,p}(E)$ and $N_0^{1,p}(E)$ depend on the ambient space X more than does $H_{C,0}^{1,p}(E)$.

The following example shows that it is possible for all three formulations of Sobolev functions with zero boundary values to be different.

Example 4.7. Let $X = \mathbb{R}^2\setminus([0,2]\times\{0\})$ and $E = B(0,1)\cap X$. Then functions in $\mathrm{Lip}_{C,0}^{1,p}(E)$ with respect to the ambient space X are also in $\mathrm{Lip}_0^{1,p}(E)$

with respect to $\mathbb{R}^2$, since these functions can be extended to be zero on $[0,1]\times\{0\}$. Hence the function space $H^{1,p}_{C,0}(E)$ with respect to X is the same as the function space $H^{1,p}_0(E)$ with respect to $\mathbb{R}^2$. However, there are functions in $\mathrm{Lip}^{1,p}_0(E)$ with respect to X that have only non-zero Lipschitz extensions to a part of $[0,1]\times\{0\}$, and, in fact, $H^{1,p}_0(E)$ with respect to X is the same as $H^{1,p}_0(B(0,1))$ with respect to $\mathbb{R}^2$.

Fix $0<\epsilon<1/4$. Let $\varphi:[0,\infty)\to\mathbb{R}$ be a non-negative Lipschitz map so that $\varphi\big|_{[0,\epsilon]}=1$ and $\varphi\big|_{[1-\epsilon,\infty)}=0$. Then the function $u(x)=u((r,\theta))=r\theta\varphi(r)$ is in $N^{1,p}_0(E)$, but since it is not an ACL function on $B(0,1)$, it is not in the function space $H^{1,p}_0(E)=W^{1,p}_0(B(0,1))$.

The following theorem gives a sufficient condition under which the three formulations agree.

Theorem 4.8. *Let X be a proper (that is, X is closed and bounded subsets of X are compact) doubling metric measure space supporting a $(1,p)$-Poincaré inequality. If E is an open subset of X, then $H^{1,p}_{C,0}(E)=H^{1,p}_0(E)=N^{1,p}_0(E)$.*

The proof of this theorem uses the following lemmas.

Lemma 4.9. *Let Y be a metric measure space equipped with a Borel regular measure that is finite on bounded sets. If u is a non-negative function in $N^{1,p}(Y)$, then the sequence of functions $u_k=\min\{u,k\}$, $k\in\mathbb{N}$, converges in the norm of $N^{1,p}(Y)$ to u.*

Proof. Let $E_k=\{x\in Y: u(x)>k\}$. If the measure of this set is zero, then $u_k=u$ almost everywhere, and as u_k is also in $N^{1,p}(Y)$, by [Sh2, Corollary 3.3] the $N^{1,p}(Y)$-norm of $u-u_k$ is zero for sufficiently large k. So suppose that the measure of E_k is positive. Since the measure is Borel regular, there is an open set O_k with $E_k\subset O_k$ such that $\mu(O_k)\le\mu(E_k)+2^{-k}$. Since $\mu(E_k)\le(1/k^p)\|u\|^p_{L^p(E_k)}$, we have

$$\mu(O_k)\le\mu(E_k)+2^{-k}\le\frac{1}{k^p}\|u\|^p_{L^p(Y)}+2^{-k}.$$

Hence the measure of O_k tends to zero as k tends to ∞. Note that $u=u_k$ on $X\setminus O_k$. Thus $2g\chi_{O_k}$ is a p-weak upper gradient of $u-u_k$ whenever g is an upper gradient of u, and hence also of u_k; see Lemma 3.3. Hence u_k converges to u in $N^{1,p}(Y)$. $\square$

Lemma 4.10. *Let X satisfy the hypotheses of Theorem 4.8, and let u be a function in $N^{1,p}(X)$. Suppose also that $0\le u\le M$, and that the set A of points in X at which u is not zero is a bounded subset of X. For each positive* *integer k, consider the function*

$$\varphi_k = (1 - w_k) \max\left\{u - \frac{1}{k}, 0\right\},$$

where w_k is a function in $N^{1,p}(X)$ such that $0 \le w_k \le 1$, $\|w_k\|_{N^{1,p}(X)} \le 2^{-k}$, and $w_k\big|_{F_k} = 1$, with F_k an open subset of X such that u is continuous on $X \backslash F_k$. Then $\varphi_k \to u$ in $N^{1,p}(X)$.

By Corollary 2.15, such F_k and w_k exist whenever u is in $N^{1,p}(X)$.

Proof. Let E_k be the set of points x in X such that $u(x) < 1/k$. Then by Corollary 2.15, and by the choice of F_k, there is an open set U_k such that $E_k \backslash F_k = U_k \backslash F_k$. Let $V_k = U_k \cup F_k$. Since the restriction of w_k to F_k is identically 1 and $u\big|_{E_k} < 1/k$, the set $\{x : \varphi_k(x) \ne 0\}$ is a subset of $A \backslash V_k \subset A$. Let $v_k = u - \varphi_k$. Then $0 \le v_k \le M$ since $0 \le \varphi_k \le u$. On the set $A \backslash V_k$ it is easy to see that $\varphi_k = (1 - w_k)(u - 1/k)$, and on V_k it is clear that $\varphi_k = 0$. Therefore, on $A \backslash V_k$,

$$v_k = w_k u + (1 - w_k)/k, \tag{6}$$

and on V_k,

$$v_k = u. \tag{7}$$

If x and y are two points in X, then as $|u| \le M$,

$$|w_k(x)u(x) - w_k(y)u(y)| \le w_k(x)|u(x) - u(y)| + M|w_k(x) - w_k(y)|.$$

Let ρ_k be an upper gradient of w_k such that $\|\rho_k\|_{L^p} \le 2^{-k+1}$, and let ρ be a p-integrable upper gradient of u. If γ is a path connecting two points x and y, then

$$|w_k(x)u(x) - w_k(y)u(y)| \le w_k(x) \int_\gamma \rho + M \int_\gamma \rho_k.$$

Hence, if z is a point in $|\gamma|$, then

$$\begin{aligned} |w_k(x)u(x) - w_k(y)u(y)| \le& |w_k(x)u(x) - w_k(z)u(z)| \\ &+ |w_k(z)u(z) - w_k(y)u(y)| \\ \le& w_k(z) \int_\gamma \rho + M \int_\gamma \rho_k. \end{aligned}$$

Hence

$$|w_k(x)u(x) - w_k(y)u(y)| \le \int_\gamma (w_k\rho + M\rho_k).$$

Therefore $w_k\rho + M\rho_k$ is an upper gradient of $w_k u$. Clearly $M\rho_k$ is p-integrable, and as $\|w_k\|_{N^{1,p}(X)} \to 0$, the L^p-norm of $M\rho_k$ also tends to zero as k goes to infinity, and by choosing a subsequence if necessary, $w_k \to 0$ a.e. Note that

$w_k\rho \le \rho$ everywhere on X. Hence, by the dominated convergence theorem and the fact that $w_k \to 0$ almost everywhere,

$$\int_X (w_k\rho)^p \to 0 \qquad \text{as } k \to \infty.$$

By Lemma 3.2 and equations (6) and (7),

$$g_k := \left(w_k\rho + M\rho_k + \frac{1}{k}\rho_k \right) \chi_{O_k} + \rho\chi_{O_k\cap V_k}$$

is a weak upper gradient of v_k, where, with B a fixed bounded open set containing the bounded set A, $O_k = (A \cup F_k) \cap B$ is a bounded open subset of X such that $A = \{x : u(x) > 0\} \subset O_k$. Now,

$$\|g_k\|_{L^p(X)} \le \left(\int_{O_k} (w_k\rho)^p \right)^{1/p} + (M + 1/k) \left(\int_{O_k} \rho_k^p \right)^{1/p} + \left(\int_{O_k\cap V_k} \rho^p \right)^{1/p}.$$

As shown above, $\int_{O_k} (w_k\rho)^p \to 0$ and $(M+1/k)^p \int_{O_k} \rho_k^p \to 0$ as k tends to infinity. Since $O_k \cap V_k$ is a subset of $(E_k \cap A) \cup F_k$, the measure of $O_k \cap V_k$ is bounded above by $\mu(E_k \cap A) + \mu(F_k)$, that is, by $\mu(\{x \in X : 0 < u(x) < 1/k\}) + \mathrm{Cap}_p(F_k)$, and hence tends to zero as k becomes larger, since bounded sets have finite measure and therefore the measure of $\{x \in X : 0 < u(x) < 1/k\}$ tends to $\mu(\emptyset) = 0$. Hence as ρ is p-integrable, the integral $\int_{O_k\cap V_k} \rho^p$ tends to 0 as $k \to \infty$. Also,

$$\begin{aligned} \|v_k\|_{L^p(X)} = \|u - \varphi_k\|_{L^p} &\le \left(\int_{A\setminus V_k} (w_k u)^p \right)^{1/p} + \frac{1}{k} \left(\int_{A\setminus V_k} |1 - w_k|^p \right)^{1/p} \\ &\quad + \left(\int_{O_k\cap V_k} |u|^p \right)^{1/p} \\ &\le M\|w_k\|_{N^{1,p}(X)} + \frac{1}{k}\mu(A)^{1/p} + \left(\int_{O_k\cap V_k} |u|^p \right)^{1/p}. \end{aligned}$$

The right hand side of the above inequality tends to zero as k tends to infinity. Hence φ_k converges to u in $N^{1,p}(X)$. $\square$

Proof of Theorem 4.8. By Proposition 4.3, it is sufficient to prove that, as Banach spaces, $N_0^{1,p}(E) \subset H_{C,0}^{1,p}(E)$. Let u be a function in $N_0^{1,p}(E)$. Identify u with its extension $\tilde{u}$. By the lattice properties of $N^{1,p}(X)$ itself, it is easy to see that then $u^+ = \max\{u, 0\}$ and $u^- = -\min\{u, 0\}$ are both in $N_0^{1,p}(E)$, and $u = u^+ - u^-$. Hence it suffices to show that u^+ and u^- are in $H_{C,0}^{1,p}(E)$. Thus, without loss of generality, $u \ge 0$. Since $H_{C,0}^{1,p}(E)$ is a Banach space isometrically embedded in $N^{1,p}(X)$, if $\{u_n\}_{n\in\mathbb{N}}$ is a sequence in $N_0^{1,p}(E)$ that is Cauchy in $N^{1,p}(X)$, then the function to which the sequence converges, in fact, lies in $H_{C,0}^{1,p}(E)$. Hence, by Lemma 4.9, it suffices to consider functions u

satisfying $0 \leq u \leq M$ for some constant M, and by Lemma 2.14, it suffices to consider functions u such that the set $A = \{x \in X : u(x) \neq 0\}$ is a bounded set. By Lemma 4.10, it suffices to show that for each positive integer k the function

$$\varphi_k = (1 - w_k) \max\{u - 1/k, 0\}$$

is in $H_0^{1,p}(E)$.

As in the proof of Lemma 4.10, we get $A \cup F_k = O_k \cup F_k$, where O_k and F_k are open subsets of X and $\mathrm{Cap}_p F_k \leq 2^{-k}$. Moreover, it is possible to choose O_k as bounded sets (since u has bounded support), contained in E. By the choice of w_k, we have $w_k\big|_{F_k} = 1$ and hence $\varphi_k\big|_{F_k} = 0$. Again, as in the proof of Lemma 4.10, if E_k is the set of all points in E on which the value of u is less than $1/k$, there exists an open set $U_k \subset E$ such that $E_k \backslash F_k = U_k \backslash F_k$. As $\varphi_k\big|_{E_k} = 0$, we have $\varphi_k\big|_{U_k \cup F_k} = 0$. Therefore,

$$\overline{\{x : \varphi_k(x) \neq 0\}} \subset \{x \in E : u(x) \geq 1/k\} \backslash F_k = O_k \backslash (E_k \cup F_k) \subset O_k \subset E.$$

Since X is proper, the support of φ_k is a compact subset of E, and hence the distance δ between the support of φ_k and $X \backslash E$ is positive. Now, as X supports a $(1, p)$-Poincaré inequality and the measure is doubling, by Theorem 2.12, the function φ_k is approximated by Lipschitz functions in $N^{1,p}(X)$. Furthermore, if x is a point in $X \backslash E$, then if g_k is an upper gradient of φ_k (assuming that $g_k\big|_{X \backslash O_k} = 0$ by Lemma 3.3), we have

$$M^* g_k^p(x) = \sup_{x \in B} \fint_B g_k^p = \sup_{x \in B, \mathrm{rad} B > \delta/2} \fint_B g_k^p \leq C_0 \frac{\|g_k\|_{L^p}^p}{(\delta/2)^s} < \infty,$$

where $s = \log_2 C$, with C the doubling constant, and C_0 is a constant depending only on the doubling constant and A. Hence, in the proof of Theorem 2.12, choosing $\lambda > C_0 \frac{\|g_k\|_{L^p}^p}{(\delta/2)^s}$ ensures that the corresponding Lipschitz approximations agree with the functions φ_k on $X \backslash E$. Hence these Lipschitz approximations are in $H_0^{1,p}(E)$, and therefore so is φ_k. Moreover, these Lipschitz approximations have compact support in E, and hence φ_k is in $H_{C,0}^{1,p}(E)$, completing the proof of the theorem. $\square$

5. The energy integral minimizer

As stated in the first section, one approach to solving the classical Dirichlet problem is to find a minimizer for the energy operator within a certain function space. The energy operator, however, is dependent on the boundary value function.

Fix the boundary value function $w \in W^{1,2}(\mathbb{R}^n)$. The energy integral to be minimized is

$$I(u) = \inf_g \|g\|_{L^2}^2,$$

where the infimum is over all upper gradients (or, equivalently, over all weak upper gradients) of the function $u+w$, where $u \in W_0^{1,2}(\Omega)$, the Sobolev space of functions with zero boundary values. The unique minimizer in this case turns out to be smooth and solves the problem

$$\triangle u = 0 \quad \text{on } \Omega, \qquad\qquad u\big|_{\mathbb{R}^n \setminus \Omega} = w\big|_{\mathbb{R}^n \setminus \Omega};$$

see, for example, [E, Section 5.13] and [KiMa].

While the above differential equation is not defined for Newtonian functions, finding a minimizer for the energy integral is an interesting problem in this setting.

Definition 5.1. Let $w \in N^{1,p}(X)$ and $E \subset X$. The *energy operator* corresponding to the boundary value function w is the operator, acting on the function space $N_0^{1,p}(E)$, defined by

$$I_{w,p}^E(u) := \inf_g \|g\|_{L^p}^p, \tag{8}$$

where the infimum is taken over p-weak upper gradients of the function $u+w$.

The generalized problem is to find a function u that minimizes the above operator.

The following is a well known theorem in functional analysis; see, for example, [HeKM], [H3], and [KSt, Chapter 2, Theorem 2.1].

Theorem 5.2. *Let B be a reflexive Banach space. If $I : B \to \mathbb{R}$ is a convex, lower semicontinuous, coercive operator, then there is an element $\tilde{u}$ in B that minimizes I.*

Here I is said to be *convex* if, for all t in $[0,1]$ and for each pair u, v in B, $I(tu + (1-t)v) \le tI(u) + (1-t)I(v)$. The operator I is *lower semicontinuous* if $I(u) \le \liminf_{n\to\infty} I(u_n)$ whenever u_n is a sequence of elements in B converging to u, and is *coercive* if $I(u_n) \to \infty$ whenever u_n is a sequence of elements in B such that $\|u_n\| \to \infty$.

Let $B = N_0^{1,p}(E)$ and let w be a function in $N^{1,p}(X)$. If u is in $N_0^{1,p}(E)$, then $(u+w)\big|_{X\setminus E} = w$. Actually, for arbitrary representatives u and w, it is only known that $(u+w)\big|_{X\setminus E} = w$ p-q.e. However, by [Sh2, Lemma 3.6] and the fact that sets of zero p-capacity have zero measure, sets of zero p-capacity are removable for Newtonian functions, so the representative u can be adjusted so that $(u+w)\big|_{X\setminus E} = w$. The aim of this section is to show that, under certain conditions on X and E, for $p > 1$ the minimizer for the operator $I_{w,p}^E$ exists and is unique. This is the content of Theorem 5.2. The next series of lemmas show that under, certain conditions, the operator $I_{w,p}^E$ satisfies the hypotheses of this theorem.

Lemma 5.3. *Suppose X is a doubling metric measure space supporting a $(1,p)$-Poincaré inequality. Let E be a bounded subset of X such that the interior of $X\backslash E$ is not empty. Then the energy operator defined in (8) is coercive on $N^{1,p}_0(E)$.*

Proof. Let u_n be a sequence of functions in $N^{1,p}_0(E)$ whose norms tend to infinity. Identifying u_n with its representative in $N^{1,p}(X)$, it follows that $\|u_n + w\|_{N^{1,p}(X)} \to \infty$. For each integer n let g_n be an upper gradient of $u_n + w$. Let B_0 be a ball in $X\backslash E$, of radius 4δ. By the assumption on E, a point y can be chosen from $\frac{1}{4}B_0$ so that $w(y)$ is finite, $Jg_w(y)$ is finite, y is a Lebesgue point of $u_n + w$ for each integer n, $u_n(y) = 0$, and there exists a positive number δ so that $B(y,\delta) \cap E = \emptyset$. Here g_w is a p-integrable upper gradient of w, and J is the modified Riesz operator defined in [HK2, Section 5] by

$$Jg(x) := \sum_{i\in\mathbb{N}} 2^{-i}R \left(\fint_{B(x,2^{-i}R)} g^p \right)^{1/p},$$

with $R = 2\operatorname{diam} E + 2\operatorname{dist}(E, B_0) + \operatorname{diam} B_0$.

Note that $w(y) + u_n(y) = w(y)$. Choose a Lebesgue point $x \in E$ of $u_n + w$. Let $B_i = B(a_i, 2^{-|i|}d(x,y))$, with $a_i = x$ if $i \geq 0$ and $a_i = y$ if $i < 0$. By the $(1,p)$-Poincaré inequality,

$$\begin{aligned} |u_n(x) + w(x) - w(y)| &\leq \sum_{i\in\mathbb{Z}} |u_{n_{B_i}} + w_{B_i} - u_{n_{B_{i+1}}} - w_{B_{i+1}}| \\ &\leq C \sum_i \operatorname{diam} B_i \left(\fint_{B_i} g_n^p \right)^{1/p} \\ &\leq C \left(Jg_n(x) + Jg_n(y) \right). \end{aligned}$$

Therefore,

$$|u_n(x) + w(x)| \leq C \left(Jg_n(x) + Jg_n(y) \right) + |w(y)|.$$

By [HK2, Theorem 5.3], if B is a ball containing E and B_0, then for every $t < Qp/(Q-p)$ (where Q is the local lower mass bound from the doubling property),

$$\|Jg_n\|_{L^t(B)} \leq C \operatorname{diam} B \mu(B)^{1/t-1/p} \|g\|_{L^p(B)}. \tag{9}$$

Since $u_n(z) + w(z) = w(z)$ on $B(y,\delta)$, by Lemma 3.2 we can assume that $g_n\big|_{B(y,\delta)} = g_w$. Therefore,

$$\begin{aligned} Jg_n(y) \le & Jg_w(y) + C \sup_{r>\delta} \left(\fint_{B(y,r)} g_n^p \right)^{1/p} \\ \le & Jg_w(y) + C \frac{1}{\mu(B(y,\delta))} \left(\int_B g_n^p \right)^{1/p} \\ \le & Jg_w(y) + C \|g_n\|_{L^p(B)}, \end{aligned}$$

with $Jg_w(y)$ finite by the choice of y. Note that the measure of E is finite since E is bounded. Hence, taking $t = p$ in (9),

$$\left(\int_E |u_n(x) + w(x)|^p d\mu(x) \right)^{1/p} \le C\|g_n\|_{L^p(B)} + C,$$

where C depends on E and y (and hence on g_w), but not on n. Since $u_n + w$ tends to ∞ in the norm of $N^{1,p}(X)$, it follows that $g_n \to \infty$ in the L^p norm. Thus $I^E_{w,p}(u_n)$ tends to infinity as $n \to \infty$. $\square$

LEMMA 5.4. *If $1 \le p < \infty$, then $I^E_{w,p}$ is convex.*

Proof. This follows from the convexity of the function $F(x) = |x|^p$ and the fact that if g_1 and g_2 are weak upper gradients of two functions u_1 and u_2, then $tg_1 + (1-t)g_2$ is a weak upper gradient of $tu_1 + (1-t)u_2$ whenever $0 \le t \le 1$. $\square$

LEMMA 5.5. *The operator $I^E_{w,p}$ is lower semicontinuous.*

Proof. Suppose u_n is a sequence of functions in $N^{1,p}_0(E)$ that converge to u in the norm of $N^{1,p}_0(E)$. Choose upper gradients g_n of $u_n + w$ so that $I^E_{w,p}(u_n)^{1/p} \ge \|g_n\|_{L^p} - 2^{-n}$. Since $\|u_n - u\|_{N^{1,p}(X)} \to 0$, there exist upper gradients ρ_n of $u_n - u$ such that $\|\rho_n\|_{L^p} \to 0$. It is easy to see that $\rho_n + g_n$ is an upper gradient of $u + w$. Hence,

$$I^E_{w,p}(u)^{1/p} \le \|\rho_n\|_{L^p} + \|g_n\|_{L^p} \le I^E_{w,p}(u_n)^{1/p} + \|\rho_n\|_{L^p} + 2^{-n}.$$

Therefore, taking the limit infimum as n tends to infinity,

$$I^E_{w,p}(u)^{1/p} \le \liminf_{n\to\infty} I^E_{w,p}(u_n)^{1/p}. \qquad \square$$

Using these lemmas, we can now prove the main theorem of this section; a different proof is given in [C].

THEOREM 5.6. *Let $p > 1$, and let X be a path-connected proper doubling metric measure space supporting a $(1,p)$-Poincaré inequality. Suppose also* *that E is a bounded subset of X so that the interior of $X\backslash E$ is not empty. Then for each function w in $N^{1,p}(X)$ there is a function $\tilde{u}$ in $N_0^{1,p}(E)$ satisfying*

$$I^E_{w,p}(\tilde{u}) = \inf_{u \in N_0^{1,p}(E)} I^E_{w,p}(u).$$

Furthermore, if X is also path-connected, such a function is unique up to sets of p-capacity zero.

Proof. By [C], $N^{1,p}(X)$ is reflexive. Since $N_0^{1,p}(E)$ is isometric with a closed subspace of $N^{1,p}(X)$, $N_0^{1,p}(E)$ is reflexive as well. Lemmas 5.3, 5.4, and 5.5, together with Theorem 5.2 yield the existence of the minimizing function.

Now suppose that u_1 and u_2 both are minimizing functions. Then, by the argument in [C, Theorem 7.15], $g_{u_1-u_2} = 0$, where $g_{u_1-u_2}$ is the minimal weak upper gradient of $u_1 - u_2$; see Corollary 3.7. By Theorem 2.17 and by the fact that X is path-connected and hence is quasiconvex (by the $(1,p)$-Poincaré inequality; see [HK2]), X satisfies the MEC_p property. Hence, for almost all points x, y in X,

$$|(u_1 - u_2)(x) - (u_1 - u_2)(y)| \le \int_\gamma g_{u_1-u_2} = 0.$$

Since by hypothesis there is a set B of positive measure on which u_1 and u_2 agree with the boundary function w, choosing such a point y that also satisfies the above inequality yields

$$|u_1(x) - u_2(x)| = 0.$$

Therefore $u_1 = u_2$ almost everywhere and hence quasi-everywhere. Thus the minimizer is unique up to sets of p-capacity zero. □

Remark 5.7. The Dirichlet type problem solved by the above theorem yields a solution $f = \tilde{u} + w$, where the boundary condition of the problem is $f\big|_{X\setminus E} = w\big|_{X\setminus E}$. If $w \in N^{1,p}(X)$ is the zero function $w = 0$, then $\tilde{u} = 0$, since then $I^E_{w,p}(u) = \inf_g \|g\|^p_{L^p}$, with the infimum being taken over weak upper gradients g of $u = u + w$ itself, and $I^E_{w,p}(0) = 0 \le I^E_{w,p}(u)$ for each function u in $N_0^{1,p}(E)$.

Henceforth, $N_0^{1,p}(E) + w$ will denote the set of all functions of form $v + w$ with v in the space $N_0^{1,p}(E)$.

6. The maximum principle

In the classical theory of harmonic functions, it is known that the minimizing function obtained in the previous section satisfies a maximum principle. A natural question is whether a maximum principle also holds in the setting of the previous section. It will be shown in this section that, under certain conditions on X, such a maximum principle does indeed hold.

Following the footsteps of [C], the previous section yields a natural definition of p-harmonic functions.

DEFINITION 6.1. Let $1 < p < \infty$. A function u in $N^{1,p}(X)$ is said to be a *relaxed p-harmonic function* on a set $E \subset X$ if for each function v in $N_0^{1,p}(E) + u$ it is true that

$$\|g_u\|_{L^p}^p \leq \|g_v\|_{L^p}^p,$$

where g_u and g_v are the minimal weak upper gradients of u and v respectively; see Corollary 3.7.

If $F \subset E$ then $N_0^{1,p}(F) \subset N_0^{1,p}(E)$. Therefore, if u is a relaxed p-harmonic function on E then for every subset F of E that function is p-harmonic on F.

DEFINITION 6.2. Let E be a subset of X. A function u on X for which there exists a function w in $N^{1,p}(X)$ with the properties that $u - w$ is in $N_0^{1,p}(E)$ and $u - w$ satisfies the conclusion of Theorem 5.6, is called a *relaxed solution of a Dirichlet problem on $E \subset X$ with boundary condition* w, where w is given by (8).

LEMMA 6.3. *Let $1 < p < \infty$. A function $u \in N^{1,p}(X)$ is a relaxed p-harmonic function on $E \subset X$ if and only if it is a relaxed solution to a Dirichlet problem on E.*

Proof. Suppose that u is a relaxed p-harmonic function on E. Then u satisfies Theorem 5.6 with $w = u$, and hence is a relaxed solution to a Dirichlet problem on E.

Conversely, if u is a relaxed solution to a Dirichlet problem on E, then u agrees with its boundary condition w on $X \backslash E$. If v is in $N_0^{1,p}(E) + u$, then it is in $N_0^{1,p}(E) + w$, and by the energy minimizing property of u it is true that

$$\|g_u\|_{L^p}^p \leq \|g_v\|_{L^p}^p.$$

Hence u is a relaxed p-harmonic function on E. □

It is known that classical Dirichlet solutions are Hölder continuous; see, for example, [HeKM]. In the above definitions the word “relaxed” is included to indicate that it is not known whether such functions are continuous in the generality of the setting considered. The paper [KiSh], however, proves that such relaxed solutions are locally Hölder continuous if X is doubling and supports a $(1, q)$-Poincaré inequality, with $1 \leq q < p$.

The following result is the main theorem of this section and establishes a maximum principle for such relaxed solutions. In this theorem the two relaxed solutions u and v correspond to two different boundary conditions, since if the boundary conditions are the same, by the uniqueness of the solution, the maximum principle is trivial. The hypotheses in this theorem are more general than the hypotheses needed for the related results in [C], which require that the measure to be doubling and that X support a $(1,p)$-Poincaré inequality. Many manifolds, such as hyperbolic spaces, do not have doubling measures and do not support a Poincaré inequality in general, but are nevertheless MEC_p.

THEOREM 6.4. *Suppose X is an MEC_p-space, where $1 < p < \infty$. Let $U \subset X$ and suppose u is a relaxed solution to a Dirichlet problem on U. If v is a relaxed solution to a Dirichlet problem on $V \subset U$ such that $v \le u$ quasi-everywhere in $\overline{U}\setminus V$, then almost everywhere on V it is true that $v \le u$.*

The proof of this result requires the following lemma.

LEMMA 6.5. *Let X be a metric measure space, and $1 \le p < \infty$. If u and v are two functions in $N^{1,p}(X)$ and E is a subset of X such that $u \le v$ p-q.e. on E and $u > v$ p-q.e. on $X\setminus E$, then $w = u\chi_E + v\chi_{X\setminus E}$ is also in $N^{1,p}(X)$.*

Proof. Since both u and v are continuous on p-modulus almost every curve, it is clear that E is p-path closed; see Remark 3.5. Let g and h be weak upper gradients of u and v, respectively. Let Γ_0 be the collection of curves γ on which at least one of u, v, is not absolutely continuous or at least one of the function-weak upper gradients pairs (u, g), (v, h) does not satisfy inequality (1) on some subpath of γ or γ passes through a point in E, where $u > v$ or through a point in $X\setminus E$ where $u \le v$. The p-modulus of this collection is zero by [Sh2, Proposition 3.1] and [Sh2, Lemma 3.6].

Let γ be a compact rectifiable curve that is not in Γ_0. Denote the two endpoints of γ by x and y. If both of these points lie in E or if both lie in $X\setminus E$, then w satisfies inequality (1) together with the function $g+h$. Hence, without loss of generality, x is in E and y lies outside of E. Since $\gamma^{-1}(E)$ is a closed subset of a compact interval, there is a point x_0 on ∂E at which γ passes from E to outside of E for the last time. Now, as $w(x_0) = u(x_0) = v(x_0)$,

$$|w(x) - w(x_0)| = |u(x) - u(x_0)| \le \int_\gamma g\,ds,$$

and

$$|w(x_0) - w(y)| = |v(x_0) - v(y)| \le \int_\gamma h\,ds.$$

Hence $|w(x) - w(y)| \le \int_\gamma (g+h)ds$. Therefore $g+h$ is a weak upper gradient of w, and hence w is in $N^{1,p}(X)$. □

Proof of Theorem 6.4. Because X is an MEC_p space and $p > 1$, by the proof of uniqueness in Theorem 5.6, relaxed solutions to a specific Dirichlet problem are unique up to sets of zero p-capacity. Let E be the set of points x in V for which $v(x) > u(x)$. The aim is to show that this set has measure zero.

Let Γ_0 be the family of all compact rectifiable paths γ on which at least one of u, v is not absolutely continuous or at least one of the function-weak upper gradient pairs (u, g_u), (v, g_v) does not satisfy inequality (1). Here g_u and g_v are the minimal weak upper gradients of u and v, respectively; see Corollary 3.7. The p-modulus of Γ_0 is zero.

Define a function w on V by

$$w(x) = \begin{cases} u(x) & \text{if } x \in X \backslash E, \\ v(x) & \text{if } x \in E. \end{cases}$$

Since by assumption $v \leq u$ on $U \backslash E$ and $v > u$ on E, by Lemma 6.5 the function w is in $N^{1,p}(X)$. Now by Lemma 3.2 and Lemma 3.4, the function $g_0 = g_v \chi_E + g_u \chi_{X \backslash E}$ is a weak upper gradient of w. Therefore, by the energy minimizing property of u and by the fact that w is in $N_0^{1,p}(U) + u$,

$$\|g_v\|^p_{L^p(E)} + \|g_u\|^p_{L^p(X \backslash E)} = \|g_0\|^p_{L^p} \geq \|g_u\|^p_{L^p(X)},$$

and hence $\|g_v\|_{L^p(E)} \geq \|g_u\|_{L^p(E)}$. Using the fact that v is a Dirichlet solution on V and setting

$$f(x) = \begin{cases} u(x) & \text{if } x \in E, \\ v(x) & \text{otherwise }, \end{cases}$$

we conclude that $\|g_v\|_{L^p(E)} \leq \|g_u\|_{L^p(E)}$, since the function f is in $N_0^{1,p}(V) + v$. Hence $\|g_v\|_{L^p(E)} = \|g_u\|_{L^p(E)}$, and therefore, since $g_w \leq g_0$ almost everywhere, $\|g_w\|_{L^p(E)} = \|g_u\|_{L^p(E)}$. Thus w is also a relaxed solution to the same Dirichlet problem as u. Hence, by the uniqueness of Dirichlet solutions, $w = u$ quasi-everywhere on X; in particular, $\mu(E) = 0$. □

As a corollary to Theorem 6.4 it is seen that a relaxed solution to a Dirichlet problem achieves its extrema on the boundary of the domain of the problem.

If U is an open subset of X, then for $\eta > 0$, define the set

$$U_\eta = \{x \in U : B(x, 2\eta) \subset U\}.$$

COROLLARY 6.6. *Let $p > 1$ and let X be an MEC_p space. Let $U \subset X$ be an open set and suppose that u is a relaxed solution to a Dirichlet problem on U. Then, for all $\eta > 0$, we have*

$$\sup_{U \backslash U_\eta} u = \sup_U u \quad \text{and} \quad \inf_{U \backslash U_\eta} u = \inf_U u.$$

That is, u cannot be larger in the interior of U than at the boundary of U, nor can it be smaller.

Proof. This corollary follows from Theorem 6.4 by noting that constant functions are relaxed solutions to some Dirichlet problem on U and that if u is a relaxed solution to a Dirichlet problem on U with boundary condition

w, then $-u$ is a relaxed solution to a Dirichlet problem on U with boundary condition $-w$. $\square$

The above corollary is a well-known fact for the classical harmonic functions in Euclidean spaces. Theorem 6.4 is also well-known for harmonic functions in Euclidean spaces, and indicates that such relaxed solutions are both subharmonic and superharmonic; see [HeKM] and [KiMa].

Remark 6.7. The paper [KiSh] shows that if X is doubling in measure and supports a $(1, q)$-Poincaré inequality for some $q < p$, then:

- A stronger form of the maximum principle holds, stating that p-harmonic functions on a domain do not achieve their extrema in the interior of that domain.
- A p-harmonic function on a domain in X is locally Hölder continuous on that domain.
- The Harnack inequality is satisfied by positive p-harmonic functions u on a domain Ω, i.e.,
$$\sup_{B_R} u \le C \inf_{B_R} u,$$
where $C > 0$ is independent of u and the ball B_R of radius R, with $2B_R \subset \Omega$.

These results are proved in [KiSh] using the DeGiorgi class argument, based on the argument given in [G] for the Euclidean case.

Acknowledgement. The author thanks Juha Heinonen for his patient guidance and helpful discussions. The author also wishes to thank Jana Björn for her helpful comments, and the referee for helpful suggestions on improving the exposition.

References

[AO] H. Aikawa and M. Ohtsuka, *Extremal length of vector measures*, Ann. Acad. Sci. Fenn. Math. **23** (1998).

[BP] M. Bourdon and H. Pajot, *Poincaré inequalities and quasiconformal structure on the boundary of some hyperbolic buildings*, Proc. Amer. Math. Soc. **127** (1999), 2315–2324.

[C] J. Cheeger, *Differentiability of Lipschitz functions on measure spaces*, Geom. Funct. Anal. **9** (1999), 428–517.

[E] R. E. Edwards, *Functional analysis: theory and applications*, Dover Publications, New York, 1994.

[FLW] B. Franchi, G. Lu, and R. L. Wheeden, *Weighted Poincaré inequalities for Hörmander vector fields and local regularity for a class of degenerate elliptic equations*, Potential Anal. **4** (1995), 361–375.

[Fu1] B. Fuglede, *Extremal length and functional completion*, Acta Math. **98** (1957), 171–218.

[Fu2] ——— *The quasi topology associated with a countable subadditive set function*, Ann. Institut Fourier (Grenoble) **21** (1971), 123–169.

[GN] N. Garofalo and D. M. Nhieu, *Isoperimetric and Sobolev inequalities for Carnot-Carathéodory spaces and the existence of minimal surfaces*, Comm. Pure Appl. Math. **49** (1996), 1081–1144.

[G] M. Giaquinta, *Introduction to regularity theory for nonlinear elliptic systems*, Birkhäuser Verlag, Zürich, 1993.

[H1] P. Hajłasz, *Sobolev spaces on an arbitrary metric space*, Potential Anal. **5** (1996), 403–415.

[H2] ———, *Geometric approach to Sobolev spaces and badly degenerated elliptic equations*, Nonlinear analysis and applications (Warsaw, 1994), GAKUTO Internat. Ser. Math. Sci. Appl. 7, Gakkotosho, Tokyo, 1996, pp. 141–168.

[H3] ———, *Sobolev spaces: theory and applications*, Lecture notes, Summer School, Jyväskylä, 1998.

[HK1] P. Hajłasz and P. Koskela, *Sobolev meets Poincaré*, C.R. Acad. Sci. Paris Sér. I Math. **320** (1995), 1211–1215.

[HK2] ———, *Sobolev met Poincaré*, Mem. Amer. Math. Soc. **145** (2000).

[He] J. Heinonen, *Lectures on analysis on metric spaces*, Springer-Verlag, New York, 2001.

[HeK1] J. Heinonen and P. Koskela, *Quasiconformal maps in metric spaces with controlled geometry*, Acta Math. **181** (1998), 1–61.

[HeK2] ———, *A note on Lipschitz functions, upper gradients, and the Poincaré inequality*, New Zealand J. Math. **28** (1999), 37–42.

[HeKM] J. Heinonen, T. Kilpeläinen, and O. Martio, *Nonlinear potential theory of degenerate elliptic equations*, Oxford University Press, Oxford, 1993.

[Hs] J. Hesse, *A p-extremal length and p-capacity equality*, Ark. Mat. **13** (1975), 131–144.

[KSh] S. Kallunki and N. Shanmugalingam, *Modulus and continuous capacity*, Ann. Acad. Sci. Fenn. Math. **26** (2001), 455–464.

[KKM] T. Kilpeläinen, J. Kinnunen, and O. Martio, *Sobolev spaces with zero boundary values on metric spaces*, Potential Anal. **12** (2000), 233–247.

[KiMa] T. Kilpeläinen and J. Malý, *Supersolutions to degenerate elliptic equation on quasi open sets*, Comm. Partial Differential Equations **17** (1992), 371–405.

[KSt] D. Kinderlehrer and G. Stampacchia, *Introduction to variational inequalities and their applications*, Academic Press, New York, 1980.

[KM1] J. Kinnunen and O. Martio, *The Sobolev capacity on metric spaces*, Ann. Acad. Sci. Fenn. Math. **21** (1996), 367–382.

[KiSh] J. Kinnunen and N. Shanmugalingam, *Regularity of quasi-minimizers on metric spaces*, Manuscripta Math. **105** (2001), 401–423.

[KM2] P. Koskela and P. MacManus, *Quasiconformal mappings and Sobolev spaces*, Studia Math. **131** (1998), 1–17.

[L] T. J. Laakso, *Ahlfors q-regular spaces with arbitrary q admitting weak Poincaré inequality*, Reports of the Department of Mathematics, University of Helsinki, No. 180, 1998.

[Ma] P. Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge University Press, Cambridge, 1995.

[MS] E. J. McShane, *Extension of range of functions*, Bull. Amer. Math. Soc. **40** (1934), 837–842.

[O] M. Ohtsuka, *Extremal lengths and precise functions in 3-space*, manuscript.

[S] S. Semmes, Personal communication.

[Sh1] N. Shanmugalingam, *Newtonian spaces: an extension of Sobolev spaces to metric measure spaces*, Ph.D. Thesis, University of Michigan, 1999.

[Sh2] ______, *Newtonian spaces: an extension of Sobolev spaces to metric measure spaces*, Revista Mat. Iberoamericana **16** (2000), 243–280.

[V] J. Väisäla, *Lectures on n-dimensional quasiconformal mapping*, Lecture Notes in Mathematics, vol. 229, Springer-Verlag, Heidelberg, 1971.

Department of Mathematics, University of Texas at Austin, Austin, TX 78712-1082

E-mail address: nageswari@math.utexas.edu