Obstacle and Dirichlet problems on arbitrary nonopen sets in metric spaces, and fine topology

Anders Björn and Jana Björn

Abstract. We study the double obstacle problem for p-harmonic functions on arbitrary bounded nonopen sets E in quite general metric spaces. The Dirichlet and single obstacle problems are included as special cases. We obtain the Adams criterion for the solubility of the single obstacle problem and establish connections with fine potential theory. We also study when the minimal p-weak upper gradient of a function remains minimal when restricted to a nonopen subset. Many of the results are new even for open E (apart from those which are trivial in this case) and also on $\mathbb{R}^n$.

1. Introduction

Sobolev spaces $W^{1,p}(\Omega)$ are usually defined for open sets Ω, and it may be difficult to use the traditional approach to make reasonable sense of $W^{1,p}(E)$ for nonopen sets E. One possibility is to let $f \in W^{1,p}(E)$ if $f \in W^{1,p}(\Omega)$ for some open set $\Omega \supset E$ depending on f, but that defeats the purpose of the definition a bit. A more fruitful approach is to consider Sobolev spaces on finely open sets, as in Kilpeläinen–Malý [24] and Malý–Ziemer [31]. This is a part of fine potential theory in $\mathbb{R}^n$, which was begun in the linear case by Cartan in 1940 and has been further developed also in the nonlinear case by various authors. See the survey Lukeš–Malý [30], the notes to Chapter 12 in Heinonen–Kilpeläinen–Martio [21] and Section 2.6 in Malý–Ziemer [31], especially for the early nonlinear history.

Since the 1990s, inspired by various applications, one has been studying Sobolev spaces on metric measure spaces without any differentiable structure. Earlier, Sobolev spaces had been extended to manifolds, Heisenberg groups and other situations with a vector-field differentiable structure. Hajłasz [18] was the first to give a definition of Sobolev spaces, the so-called Hajłasz spaces, on general metric

spaces, while Shanmugalingam [36] and Cheeger [13] a little later introduced the so-called Newtonian spaces. We follow Shanmugalingam below but Cheeger's definition is more or less equivalent. Let us point out that we only consider first order Sobolev spaces in this discussion.

Since a measurable subset E of a metric measure space X can be considered as a metric measure space in its own right, these new definitions are well suited for defining Sobolev spaces on arbitrary nonopen measurable sets, e.g. of $\mathbb{R}^n$ and other smooth spaces.

In many situations, in particular on (unweighted) $\mathbb{R}^n$, both Hajłasz and Newtonian spaces coincide with the usual Sobolev space; see [36]. However on general open subsets of $\mathbb{R}^n$ it is only the Newtonian space that coincides with the usual Sobolev space. The Hajłasz space is in general smaller and the Hajłasz gradient is not local, i.e., it need not vanish on sets where the function is constant; see e.g. Shanmugalingam [36], Hajłasz [19] and the discussion in Appendix B.1 in Björn–Björn [6]. It therefore seems that the Newtonian approach is the most suitable, for example for solving partial differential equations and variational problems on metric spaces and on general subsets of $\mathbb{R}^n$. Other advantages of Newtonian spaces are that the equivalence classes are defined up to sets of capacity zero, rather than measure zero, and that all Newtonian functions are absolutely continuous on p-almost all curves. Under suitable assumptions, they are also finely continuous outside sets of zero capacity (see J. Björn [12] and Korte [27]), which provides another connection to the fine potential theory mentioned above.

In this paper we study the double obstacle problem on general bounded measurable subsets of a metric space X with a Borel regular measure μ, i.e., we minimize the p-energy functional

$$(1.1) \qquad \int_E g_{u,E}^p \, d\mu,$$

among all functions u lying (up to sets of capacity zero) between two obstacles $\psi_1, \psi_2 \colon E \to \overline{\mathbb{R}}$ and with prescribed boundary values f from the Newtonian space $N^{1,p}(E)$ on E. The Dirichlet problem is included as the special case with $\psi_1 \equiv -\infty$ and $\psi_2 \equiv \infty$.

Here $g_{u,E}$ is the minimal p-weak upper gradient of u (with respect to E), which is the metric space counterpart of the (modulus of) the usual gradient. It depends on the underlying metric space and it is therefore important for us to understand when a restriction of a minimal p-weak upper gradient from the underlying metric space X remains minimal on E. This is studied in Section 3. In particular, we show that $g_{u,E} = g_{u,X}$ if E is p-path almost open, which in unweighted $\mathbb{R}^n$ holds for all finely open sets E. In this case we have $g_{u,E} = g_{u,X} = |\nabla u|$ a.e., where ∇u is the distributional gradient of u. An interesting example of this phenomenon on a nowhere dense set $E \subset [0,1]^n \subset \mathbb{R}^n$ with almost full measure in $[0,1]^n$ is presented in Examples 9.5 and 9.6.

Existence and uniqueness (up to sets of capacity zero) of solutions to the above $\mathcal{K}_{\psi_1,\psi_2,f}(E)$-obstacle problem associated with (1.1) is proved in Section 4. The assumptions under which these results hold, and possibilities for relaxing them,

are discussed in Section 5. We have made an effort to consider the obstacle problem under minimal assumptions. In particular, we do not assume that the measure μ is doubling and we only use a very weak version of the Poincaré type inequality, namely the Friedrichs inequality, which moreover can be further relaxed in many situations. Note that there are infinite-dimensional spaces with nondoubling measures supporting a Poincaré inequality; see e.g. Rajala [34]. One existence result that we obtain is the following theorem which follows from Theorem 4.2 and Remark 5.6.

Theorem 1.1. *Let X be an arbitrary metric space, let $E \subset X$ be a bounded measurable set, whose complement has positive capacity, and let $\psi_1, \psi_2 \in L^p(E)$, $p > 1$. If $f \in N^{1,p}(E)$ is such that $\mathcal{K}_{\psi_1,\psi_2,f}(E) \neq \varnothing$, then the $\mathcal{K}_{\psi_1,\psi_2,f}(E)$-obstacle problem is soluble.*

Moreover, if the p-Friedrichs inequality holds on X then the assumption that $\psi_1, \psi_2 \in L^p(E)$ can be omitted and the solution is unique (up to sets of capacity zero).

The p-Friedrichs inequality holds for example if there is an increasing sequence of balls B_j covering X such that for each $j = 1, 2, \ldots$, and all $u \in N_0^{1,p}(B_j)$,

$$\int_{B_j} |u|^p \, d\mu \leq C_j \int_{B_j} g_u^p \, d\mu.$$

This is usually easier to verify than the classical Poincaré inequality; see Example 5.2.

Along the way, we also discuss alternative definitions of the obstacle problem and relations between them. In particular, we compare our obstacle problem with the obstacle problem defined by means of the global minimal p-weak upper gradient $g_{u,X}$ and with the classical obstacle problem on open sets. Another novelty here (apart from E being nonopen) is that we allow f to merely belong to the Dirichlet space $D^p(E)$ of measurable functions with an upper gradient in $L^p(E)$. A useful application of our theory to condenser capacities is given in Theorem 5.13.

In Section 6 we establish the Adams criterion for the solubility of the single obstacle problem with $\psi_2 \equiv \infty$. We also show by examples that the situation is much more subtle for the double obstacle problem.

A natural question is when all the competing functions in $\mathcal{K}_{\psi_1,\psi_2,f}(E)$ coincide (up to sets of capacity zero). In this case they are of course all solutions of the obstacle problem. This happens for example if $N_0^{1,p}(E)$ is trivial (i.e., all functions vanish outside a set of capacity zero). In Section 7 we characterize those sets where this occurs. It turns out that this problem has close connections with fine potential theory and that

$$N_0^{1,p}(E) = N_0^{1,p}(\text{fine-int } E).$$

We also give the following characterization of the fine interior, which is useful in applications and examples, as it is easier and more explicit to verify the thinness condition (7.2) for $X \setminus E$ than for $X \setminus \text{fine-int } E$; see Examples 9.5 and 9.6.

Theorem 1.2. *Assume that X is complete and supports a $(1, p)$-Poincaré inequality, that μ is doubling and that $p > 1$. Let $E \subset X$ be arbitrary. Then $x \in$ fine-int E if and only if $x \in E$ and $X \setminus E$ is thin at x.*

A corresponding characterization in $\mathbb{R}^n$ (formulated in terms of the fine closure) can be found in Theorem 2.136 in Malý–Ziemer [31], where it is obtained as a consequence of the Cartan property. Our proof in metric spaces is based on more elementary properties of the capacity and does not use any characterization of thin sets by superharmonic functions. It may therefore be of interest also in the classical setting.

On (unweighted) $\mathbb{R}^n$, our theory comes together in an elegant way, which we explain in Section 9. In particular, we have the following result, which is a special case of Theorem 8.3 (in view of the results in Section 9).

Theorem 1.3. *Let $E \subset \mathbb{R}^n$ (unweighted) be a bounded measurable set. Assume that $p > 1$, that $f \in D^p(E)$ and that $\mathcal{K}_{\psi_1, \psi_2, f}(E) \neq \varnothing$. Then the solutions of the $\mathcal{K}_{\psi_1, \psi_2, f}(E)$-problem coincide with the solutions of the $\mathcal{K}_{\psi_1, \psi_2, f}(E_0)$-problem, where E_0 is the fine interior of E.*

Moreover, $g_{u, E_0} = g_{u, E}$ a.e. in E_0 and if the Lebesgue measure of $E \setminus E_0$ is zero, then also the p-energies (1.1) associated with these two problems coincide.

If $f \in D^p(\Omega)$ for some open set $\Omega \supset E$, then $g_{u, E_0} = g_{u, E} = |\nabla u|$ a.e. in E_0 and the above solutions coincide with the solutions of the $\mathcal{K}_{\psi_1', \psi_2', f}(\Omega)$-problem, where $\psi_j' = \psi_j$ in E and $\psi_j' = f$ on $\Omega \setminus E$, $j = 1, 2$.

These results in turn justify the earlier studies of finely open sets and the fine obstacle problem on $\mathbb{R}^n$ in the literature, as in Kilpeläinen–Malý [24], Malý–Ziemer [31] and Latvala [29]. We hope to use the results from this paper for further development of fine potential theory in the setting of Newtonian spaces on metric spaces (with $\mathbb{R}^n$ as an important special case). Some aspects of fine potential theory in this setting have been studied by Kinnunen–Latvala [25], J. Björn [12] and Korte [27].

2. Notation and preliminaries

We assume throughout the paper that $X = (X, d, \mu)$ is a metric space equipped with a metric d and a measure μ such that

$$0 < \mu(B) < \infty$$

for all balls $B = B(x_0, r) := \{x \in X : d(x, x_0) < r\}$ in X (we make the convention that balls are nonempty and open). The σ-algebra on which μ is defined is obtained by completion of the Borel σ-algebra. We also assume that $1 \leq p < \infty$ and that $\Omega \subset X$ is a nonempty open set.

The measure μ is *doubling* if there exists a constant $C > 0$ such that

$$0 < \mu(2B) \leq C\,\mu(B) < \infty$$

for all balls $B \subset X$, where $\lambda B = B(x_0, \lambda r)$.

A *curve* is a continuous mapping from an interval. We will only consider curves which are nonconstant, compact and rectifiable. A curve can thus be parameterized by its arc length ds.

We follow Heinonen and Koskela [22] in defining upper gradients as follows (they called them very weak gradients).

Definition 2.1. A nonnegative Borel function g on X is an *upper gradient* of an extended real-valued function f on X if for all (nonconstant, compact and rectifiable) curves $\gamma : [0, l_\gamma] \to X$,

$$(2.1) \qquad |f(\gamma(0)) - f(\gamma(l_\gamma))| \leq \int_\gamma g\,ds,$$

where we make the convention that the left-hand side is ∞ whenever at least one of its terms is infinite. If g is a nonnegative measurable function on X and if (2.1) holds for p-almost every curve (see below), then g is a *p-weak upper gradient* of f.

Here and in what follows, we say that a property holds for *p-almost every curve* if it fails only for a curve family Γ with zero p-modulus, i.e., there exists $0 \leq \rho \in L^p(X)$ such that $\int_\gamma \rho\,ds = \infty$ for every curve $\gamma \in \Gamma$. It is easy to show that a countable union of curve families with zero p-modulus also has zero p-modulus. Moreover, if $\mathrm{Mod}_p(\Gamma) = 0$ and Γ' consists of all curves which have a subcurve in Γ, then $\mathrm{Mod}_p(\Gamma') = 0$.

Note that a p-weak upper gradient need not be a Borel function; it is only required to be measurable. It is implicitly assumed that $\int_\gamma g\,ds$ is defined (with a value in $[0, \infty]$) for p-almost every curve γ, although this is in fact a consequence of the measurability; see Section 3 in Björn–Björn [4] (which is not in Björn–Björn [5]).

The p-weak upper gradients were introduced in Koskela–MacManus [28]. It was also shown there that if $g \in L^p_{\mathrm{loc}}(X)$ is a p-weak upper gradient of f, then one can find a sequence $\{g_j\}_{j=1}^\infty$ of upper gradients of f such that $g_j - g \to 0$ in $L^p(X)$. If f has an upper gradient in $L^p_{\mathrm{loc}}(X)$, then it has a *minimal p-weak upper gradient* $g_f \in L^p_{\mathrm{loc}}(X)$ in the sense that for every p-weak upper gradient $g \in L^p_{\mathrm{loc}}(X)$ of f we have $g_f \leq g$ a.e.; see Shanmugalingam [37] and Hajłasz [19]. The minimal p-weak upper gradient is well defined up to an equivalence class in the cone of nonnegative functions in $L^p_{\mathrm{loc}}(X)$.

For proofs of the various facts in this section we refer to Björn–Björn [6]. (Some of the references we mention here may not provide a proof in the generality considered here, but such proofs are given in [6].)

Note that upper gradients and in particular the minimal p-weak upper gradient strongly depend on the underlying space. Any measurable $E \subset X$ can be considered as a metric space in its own right, thus giving rise to upper gradients with

respect to E. An upper gradient with respect to X is always an upper gradient with respect to E, but the converse need not be true; see Example 3.6. We denote the minimal p-weak upper gradient of u with respect to E by $g_{u,E}$, whereas g_u always denotes the minimal p-weak upper gradient with respect to X (also denoted $g_{u,X}$).

Following Shanmugalingam [36], we define a version of Sobolev spaces on the metric space X.

Definition 2.2. The *Newtonian space* on X is

$$N^{1,p}(X) = \{u : \|u\|_{N^{1,p}(X)} < \infty\},$$

where

$$\|u\|_{N^{1,p}(X)} = \left(\int_X |u|^p \, d\mu + \int_X g_u^p \, d\mu \right)^{1/p},$$

if $u : X \to \overline{\mathbb{R}} := [-\infty, \infty]$ is an everywhere defined measurable function having an upper gradient in $L_{\mathrm{loc}}^p(X)$.

We also say that an everywhere defined measurable function u on X belongs to the *Dirichlet space* $D^p(X)$ if it has an upper gradient in $L^p(X)$.

The local spaces $N_{\mathrm{loc}}^{1,p}(X)$ and $D_{\mathrm{loc}}^p(X)$ are defined by requiring that for every $x \in X$ there is a ball $B_x \subset X$ such that $u \in N^{1,p}(B_x)$ or $u \in D^p(B_x)$, respectively. For a measurable set $E \subset X$, the spaces $N^{1,p}(E)$, $D^p(E)$ and the corresponding local spaces are defined by considering E as a metric space in its own right. Note a subtle point here (recall that X is *proper* if all closed and bounded sets are compact): If X is not proper, then the above definition of the local spaces need not be equivalent to requiring that $u \in N^{1,p}(K)$ for all compact $K \subset X$. (See A. Björn–Marola [10] for a related definition on noncomplete spaces.) Note that if μ is doubling then X is proper if and only if it is complete.

The quotient space $N^{1,p}(X)/\sim$, where $u \sim v$ if and only if $\|u - v\|_{N^{1,p}(X)} = 0$, is a Banach space and a lattice; see Shanmugalingam [36]. Note that we assume functions in Newtonian and Dirichlet spaces to be defined everywhere, and not just up to an equivalence class in the corresponding function space. This is needed for the definition of upper gradients to make sense. Shanmugalingam [36] also showed that every $u \in D_{\mathrm{loc}}^p(X)$ is absolutely continuous on p-almost every curve γ in X, in the sense that $u \circ \gamma$ is a real-valued absolutely continuous function.

If $u, v \in D_{\mathrm{loc}}^p(X)$, then their minimal p-weak upper gradients coincide almost everywhere in the set $\{x \in X : u(x) = v(x)\}$; in particular $g_{\min\{u,c\}} = g_u \chi_{\{u<c\}}$ a.e. for $c \in \mathbb{R}$. Moreover, $g_{uv} \leq |u|g_v + |v|g_u$.

Definition 2.3. The *capacity* of a set $E \subset X$ is the number

$$C_p(E) = \inf \|u\|_{N^{1,p}(X)}^p,$$

where the infimum is taken over all $u \in N^{1,p}(X)$ such that $u = 1$ on E.

We say that a property holds *quasi-everywhere* (q.e.) if the set of points for which it fails has capacity zero.

This capacity was introduced and used for Newtonian spaces in Shanmugalingam [36]. It is countably subadditive and the correct gauge for distinguishing between two Newtonian functions. If $u \in N_{\text{loc}}^{1,p}(X)$ and $v : X \to \overline{\mathbb{R}}$, then $u \sim v$ if and only if $u = v$ q.e. Moreover, if $u, v \in D_{\text{loc}}^p(X)$ and $u = v$ a.e., then $u = v$ q.e. See also Appendix B where the variational capacity is defined. Note that if $C_p(E) = 0$, then p-almost every curve in X avoids E, by e.g. Lemma 3.6 in Shanmugalingam [36].

To be able to compare the boundary values of Newtonian functions we need a Newtonian space with zero boundary values. We let

$$N_0^{1,p}(E) = \{f|_E : f \in N^{1,p}(X) \text{ and } f = 0 \text{ on } X \setminus E\}.$$

One can replace the assumption "$f = 0$ on $X \setminus E$" with "$f = 0$ q.e. on $X \setminus E$" without changing the space $N_0^{1,p}(E)$. Functions from $N_0^{1,p}(E)$ can be extended by zero quasi-everywhere in $X \setminus E$ and we will regard them in that sense if needed. Note that if $C_p(X \setminus E) = 0$, then $N_0^{1,p}(E) = N^{1,p}(E) = N^{1,p}(X)$, since p-almost every curve in X avoids $X \setminus E$.

The following lemma is useful for proving that certain functions belong to $N_0^{1,p}(E)$. For open E, it was obtained in Björn–Björn [5]. The proof of the general case can be found in Björn–Björn [6].

Lemma 2.4. *Assume that $E \subset X$ is measurable. Let $u \in N^{1,p}(E)$ and $v, w \in N_0^{1,p}(E)$ be such that $v \leq u \leq w$ q.e. in E. Then $u \in N_0^{1,p}(E)$.*

The following Poincaré inequality is often assumed in the literature. Because of the dilation λ on the right-hand side, it is sometimes called a weak Poincaré inequality.

Definition 2.5. We say that X supports a (q, p)-*Poincaré inequality*, $q \geq 1$, if there exist constants $C > 0$ and $\lambda \geq 1$ such that for all balls $B \subset X$ and all integrable $u \in D_{\text{loc}}^p(X)$,

$$(2.2) \qquad \left(\fint_B |u - u_B|^q \, d\mu \right)^{1/q} \leq C \, (\text{diam } B) \left(\fint_{\lambda B} g_u^p \, d\mu \right)^{1/p},$$

where $u_B := \fint_B u \, d\mu := \int_B u \, d\mu / \mu(B)$.

Using the above-mentioned results on p-weak upper gradients from Koskela–MacManus [28], it is easy to see that (2.2) can equivalently be required for all upper gradients g of u. If X supports a $(1, p)$-Poincaré inequality and μ is doubling, then by Theorem 5.1 in Hajłasz–Koskela [20], it supports a (q, p)-Poincaré inequality for some $q > p$, and in particular a (p, p)-Poincaré inequality. Moreover, under these assumptions, Lipschitz functions are dense in $N^{1,p}(X)$; see Shanmugalingam [36]. If X is also complete then functions in $N^{1,p}(X)$ as well as those in $N^{1,p}(\Omega)$ are quasicontinuous; see Björn–Björn–Shanmugalingam [9]. It also follows that $N_0^{1,p}(\Omega)$ for open Ω can be defined equivalently as the closure of Lipschitz functions with compact support in Ω; see Shanmugalingam [37] or Theorem 5.45

in Björn–Björn [6]. For a general set E this is not always possible and the above definition of $N_0^{1,p}(E)$ seems to be the natural one.

Moreover, if X is unweighted $\mathbb{R}^n$ and $u \in D_{\mathrm{loc}}^p(X)$, then $g_u = |\nabla u|$ a.e., where ∇u is the distributional gradient of u. This means that for open $\Omega \subset \mathbb{R}^n$, $N^{1,p}(\Omega)$ is the refined Sobolev space as in Heinonen–Kilpeläinen–Martio [21], p. 96. See Hajłasz [19] or Appendix A.1 in [6] for a full proof of this fact for unweighted $\mathbb{R}^n$, and Appendix A.2 in [6] for a proof for weighted $\mathbb{R}^n$ (requiring $p > 1$).

For most results in this paper we will need some kind of Poincaré type inequality, but it is enough with a considerably weaker form than the one in Definition 2.5. Therefore we introduce the following Friedrichs inequality, which will be useful when proving the existence and uniqueness of the solutions of our obstacle problems. It is sometimes called the (p,p)-Poincaré inequality for $N_0^{1,p}$, and follows from, but does not imply, the usual Poincaré inequality as in Definition 2.5; see Lemma 5.1 and Example 5.2. It is also in general easier to verify. Friedrichs obtained it (for $p = 2$ on $\mathbb{R}^2$) in [16], p. 211, with an additional term on the right-hand side which vanishes for functions in $N_0^{1,p}$. We have chosen this name to avoid confusion with other Poincaré type inequalities, which are abundant in this field.

Definition 2.6. We say that X supports a p-*Friedrichs inequality* if for every bounded $E \subset X$ with $C_p(X \setminus E) > 0$ there exists $C_E > 0$ such that for all $u \in N_0^{1,p}(E)$ (extended by 0 outside E),

$$(2.3) \qquad \int_X |u|^p \, d\mu \le C_E \int_X g_u^p \, d\mu.$$

A direct consequence is that $\|u\|_{N^{1,p}(X)}^p \le \widetilde{C}_E \|g_u\|_{L^p(X)}^p$ for $u \in N_0^{1,p}(E)$. If E is measurable, then the integrals and the norms can equivalently be taken with respect to E. As in (2.2), one can equivalently verify (2.3) for all upper gradients g of u. If X is unbounded then the condition $C_p(X \setminus E) > 0$ is of course redundant. On the other hand, if X is bounded then it is essential, as otherwise $1 \in N_0^{1,p}(E)$ violates (2.3).

We will also need the space

$$D_0^p(E) = \{f|_E : f \in D^p(X) \text{ and } f = 0 \text{ on } X \setminus E\}.$$

As we shall now see, it will coincide with $N_0^{1,p}(E)$ in most cases, and when it does we prefer to write $N_0^{1,p}(E)$.

Proposition 2.7. *Assume that X supports the p-Friedrichs inequality* (2.3) *and that $E \subset X$ is bounded and $C_p(X \setminus E) > 0$. Then*

$$D_0^p(E) = N_0^{1,p}(E).$$

Proof. Let $u \in D_0^p(E)$ and extend u by 0 outside E. Let $g \in L^p(X)$ be an upper gradient of u, and let $u_k = \max\{\min\{u, k\}, -k\}$, $k = 1, 2, \ldots$, be the truncations of u at levels $\pm k$. Then g is an upper gradient of u_k as well. As E is bounded,

$u_k \in L^p(X)$ and thus $u_k \in N_0^{1,p}(E)$. Hence, by monotone convergence and the Friedrichs inequality,

$$\int_X |u|^p \, d\mu = \lim_{k \to \infty} \int_X |u_k|^p \, d\mu \leq C_E \int_X g^p \, d\mu < \infty.$$

Thus $u \in N^{1,p}(X)$ and hence $u \in N_0^{1,p}(E)$. This proves one inclusion, while the converse inclusion is trivial. $\qquad\qquad\square$

Finally, we make the convention that, unless otherwise stated, the letter C denotes various positive constants whose exact values are not important and may vary with each use.

3. Restrictions of minimal p-weak upper gradients

In Section 4, we will define and study the obstacle problem, in which we minimize the p-energy functional (1.1) on general sets. Since the energy functional is defined using the minimal p-weak upper gradient, it is natural to study how this notion depends on the underlying set. This will be done in this section. We point out that for this we do not impose any assumptions on X, such as the doubling property of μ or the Poincaré inequality.

If Ω is open and $f \in D_{\mathrm{loc}}^p(X)$ then the minimal p-weak upper gradient of f with respect to X remains minimal when restricted to Ω, i.e., with respect to $D_{\mathrm{loc}}^p(\Omega)$. This is folklore but the interested reader can find a proof in Björn–Björn [6], Lemma 2.23. We will need a generalization of this result to p-path almost open sets; see Proposition 3.5.

Definition 3.1. The set $G \subset X$ is *p-path open* (in X) if for p-almost every curve $\gamma \colon [0, l_\gamma] \to X$, the set $\gamma^{-1}(G)$ is (relatively) open in $[0, l_\gamma]$.

Furthermore, $G \subset X$ is *p-path almost open* (in X) if for p-almost every curve $\gamma \colon [0, l_\gamma] \to X$, the set $\gamma^{-1}(G)$ is the union of an open set and a set with zero one-dimensional Lebesgue measure.

The p-path open sets were introduced by Shanmugalingam [37], Remark 3.5. The name "p-path almost open" is perhaps a little misleading, as we *do not* allow $\gamma^{-1}(G)$ to be an open set *minus* a set of measure zero. For our purposes there are counterexamples showing that we cannot allow this; see Example 3.6 below.

Clearly, every open set is p-path open, and every p-path open set is p-path almost open. The following observation sheds some light on which sets are p-path (almost) open.

Lemma 3.2. *Let $E, G \subset X$.*

If G is p-path open and $C_p(E \setminus G) = C_p(G \setminus E) = 0$, then E is p-path open.

If G is p-path almost open and $C_p(G \setminus E) = \mu(E \setminus G) = 0$, then E is p-path almost open. In particular, if $\mu(E \cap \partial E) = 0$, then E is p-path almost open.

Proof. First, assume that G is p-path open and that $C_p(E \setminus G) = C_p(G \setminus E) = 0$. Then p-almost every curve γ avoids $(E \setminus G) \cup (G \setminus E)$ and hence $\gamma^{-1}(E) = \gamma^{-1}(G)$ is (relatively) open for p-almost every curve γ, i.e., E is p-path open.

Next, assume that G is p-path almost open and $C_p(G \setminus E) = \mu(E \setminus G) = 0$. Then p-almost every curve γ avoids $G \setminus E$ and is such that $\gamma^{-1}(E \setminus G)$ has zero one-dimensional Lebesgue measure, by e.g. Lemma 1.42 in Björn–Björn [6]. For all such curves we have $\gamma^{-1}(E) = \gamma^{-1}(G) \cup \gamma^{-1}(E \setminus G)$, i.e., E is p-path almost open. $\square$

Remark 3.3. The collection of all p-path open sets does not (in general) form a topology on X. Consider unweighted $\mathbb{R}^n$ with $n > 1$ and $1 \leq p \leq n$, in which case all singleton sets are p-path open since they have capacity zero. If the p-path open sets formed a topology it would follow that any set on $\mathbb{R}^n$ would be p-path open. However it is quite easy, using Lemma A.1 in Björn–Björn [6], to see that $\mathbb{R}^{n-1} \times \mathbb{Q}$ is not p-path open. If singletons have positive capacity (e.g. if $X = \mathbb{R}^n$ and $p > n$), then any p-path open set is open, and thus the family of p-path open sets does form a topology in this case.

Similarly, the p-path almost open sets do not (in general) from a topology on X. On $\mathbb{R}^n$, any singleton set is p-path almost open (but not p-path open if $p > n$). The set $\mathbb{R}^{n-1} \times \mathbb{Q}$ is p-path almost open, but $\mathbb{R}^{n-1} \times (\mathbb{R} \setminus \mathbb{Q})$ is not. Thus, the p-path almost open sets do not form a topology on $\mathbb{R}^n$.

If there are no nonconstant rectifiable curves in X, as for example on the von Koch snowflake curve, then all sets are p-path open, and thus in this case the p-path open sets form a topology, and so do the p-path almost open sets. This also shows that p-path open sets need not be measurable.

A consequence of Lemma 3.2 is that the union of a p-path open set and a set of measure zero is p-path almost open.

Open problem 3.4. Can every p-path almost open set be written as a union of a p-path open set and a set of measure zero?

The following result shows that p-path almost open sets preserve the minimal p-weak upper gradients in the same way as open sets do. Recall that by g_u we always mean the p-weak upper gradient of u with respect to X.

Proposition 3.5. *Let G be a p-path almost open measurable set and let $u \in D^p_{\mathrm{loc}}(X)$. Then $g_{u,G} = g_u$ a.e. in G, i.e., $g_u|_G$ is a minimal p-weak upper gradient of u with respect to G.*

Before proving this result it may be worth observing that some condition on G is necessary.

Example 3.6. Let $X = \mathbb{R}$ and $E = (0,1) \setminus \mathbb{Q}$. Since E contains no rectifiable curves, the minimal p-weak upper gradient taken with respect to E is zero for every function on E. On the other hand, the minimal p-weak upper gradient with respect to $\mathbb{R}$ is just the modulus of the distributional derivative. For example, if $u(x) = x$, then $g_u = 1$ a.e., while $g_{u,E} = 0$ a.e. Note also that E has full measure in the open interval $I = (0,1)$ for which $g_{u,I} = g_u = 1$ a.e.

Proof of Proposition 3.5. Clearly, $g_{u,G} \leq g_u$ a.e. in G. We shall show that the function

$$g = \begin{cases} g_{u,G} & \text{in } G, \\ g_u & \text{in } X \setminus G, \end{cases}$$

is a p-weak upper gradient of u in X. Let Γ_0 consist of all curves γ in X for which $\gamma^{-1}(G)$ is not a union of an open set and a set with zero one-dimensional Lebesgue measure. Also let Γ_1 be the collection of all curves in G on which (2.1) fails for u and $g_{u,G}$. Similarly, let Γ_2 consist of all those curves in X on which (2.1) fails for u and g_u. Finally, let Γ_3 consist of all those curves in X for which $\int_\gamma g_u \, ds = \infty$. Then $\mathrm{Mod}_p(\Gamma_0 \cup \Gamma_1 \cup \Gamma_2 \cup \Gamma_3) = 0$.

Let $\gamma \colon [0, l_\gamma] \to X$ be a curve having no subcurve in $\Gamma_0 \cup \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$. By Lemma 1.34 in Björn–Björn [6], p-almost every curve in X has this property. Then $\gamma^{-1}(G) = G' \cup A$, where G' is open in $(0, l_\gamma)$ and A has zero one-dimensional Lebesgue measure. The set G' can be written as a countable pairwise disjoint union $\bigcup_{j=1}^\infty I_j$ of open intervals $I_j = (a_j, b_j)$, $j = 1, 2, \ldots$ (Here we allow some of the intervals I_j to be empty.) We then have

$$|u(\gamma(0)) - u(\gamma(l_\gamma))| \leq |u(\gamma(0)) - u(\gamma(a_1))| + |u(\gamma(a_1)) - u(\gamma(b_1))|$$
$$+ |u(\gamma(b_1)) - u(\gamma(l_\gamma))| \leq \int_{\gamma|_{I_1}} g_{u,G} \, ds + \int_{\gamma|_{[0,l_\gamma] \setminus I_1}} g_u \, ds.$$

Continuing in this way, we obtain for all $j = 1, 2, \ldots$,

$$|u(\gamma(0)) - u(\gamma(l_\gamma))| \leq \int_{\gamma|_{\bigcup_{i=1}^j I_i}} g_{u,G} \, ds + \int_{\gamma|_{[0,l_\gamma] \setminus \bigcup_{i=1}^j I_i}} g_u \, ds.$$

Since $\int_\gamma g_u \, ds < \infty$, letting $j \to \infty$ and using monotone and dominated convergence shows that

$$|u(\gamma(0)) - u(\gamma(l_\gamma))| \leq \int_{\gamma|_{G'}} g_{u,G} \, ds + \int_{\gamma|_{[0,l_\gamma] \setminus G'}} g_u \, ds$$
$$= \int_{\gamma|_{G' \cup A}} g_{u,G} \, ds + \int_{\gamma|_{[0,l_\gamma] \setminus (G' \cup A)}} g_u \, ds = \int_\gamma g \, ds.$$

Thus, g is a p-weak upper gradient of u in X and hence $g_u \leq g$ a.e. in X. It follows that $g_u \leq g_{u,G}$ a.e. in G, which finishes the proof. □

Corollary 3.7. *Let $E \subset X$ be measurable and let $G \subset E$ be a p-path almost open (with respect to X) measurable set. If $u \in D_{\mathrm{loc}}^p(X)$, then*

$$g_{u,G} = g_{u,E} = g_u \quad \text{a.e. in } G.$$

Proof. Clearly, $g_{u,G} \leq g_{u,E} \leq g_u$ a.e. in G. Since G is p-path almost open, Proposition 3.5 shows that equality must hold almost everywhere in G. □

Remark 3.8. Note that Corollary 3.7 can also be applied to E instead of X, giving that for $u \in D^p_{\mathrm{loc}}(E)$, $g_{u,G} = g_{u,E}$ a.e. in G, whenever $G \subset E$ is measurable and p-path almost open with respect to E, in particular if it is measurable and p-path almost open with respect to X.

Another application of p-path open sets is the following sufficient condition for when a function belongs to $N^{1,p}_0(E)$. This generalizes Theorem 2.147 and Corollary 2.162 in Malý–Ziemer [31]. See also Lemma 2.4 and Theorems 1.2 and 7.3 for related results, and Proposition 9.4 where this is combined with fine topology on $\mathbb{R}^n$.

Lemma 3.9. *Let $E_1 \subset E_2 \subset X$ with E_1 and $X \setminus E_2$ being p-path open. If $u \in N^{1,p}(E_2)$ and $u = 0$ q.e. in $E_2 \setminus E_1$ then the zero extension of u belongs to $N^{1,p}(X)$ and in particular $u \in N^{1,p}_0(E_1)$.*

Note that "p-path open" in Lemma 3.9 cannot be replaced by "p-path almost open", as the example with $E_1 = E_2 = (0,1) \subset \mathbb{R} = X$ and $u = \chi_{E_1}$ shows. The most common use of Lemma 3.9 is perhaps when E_1 and E_2 are the interior and the closure of some set, respectively.

Proof. We shall show that g_{u,E_2} (extended by zero) is a p-weak upper gradient of u (extended by zero) in X. Let Γ be the family of curves in E_2 on which (2.1) fails for u and g_{u,E_2}. Then $\mathrm{Mod}_p(\Gamma) = 0$. Also let $A = \{x \in E_2 \setminus E_1 : u(x) \neq 0\}$. Since $C_p(A) = 0$, we conclude that p-almost every curve $\gamma \colon [0, l_\gamma] \to X$ avoids A, does not have a subcurve in Γ, and is such that both $\gamma^{-1}(E_1)$ and $\gamma^{-1}(X \setminus E_2)$ are relatively open.

Let γ be such a curve. We can assume that γ passes through both E_1 and $X \setminus E_2$. Otherwise there is nothing to prove, since g_{u,E_2} is a p-weak upper gradient in E_2 and $u = 0$ outside $E_1 \cup A$. By splitting γ into two parts and reversing the orientation, if necessary, we can assume that $\gamma(0) \in E_1$ and $\gamma(l_\gamma) \in X \setminus E_2$.

Let $c = \inf\{t \in [0, l_\gamma] : \gamma(t) \in X \setminus E_2\}$. Since both $\gamma^{-1}(E_1)$ and $\gamma^{-1}(X \setminus E_2)$ are relatively open in $[0, l_\gamma]$, we conclude that $\gamma(c) \in (E_2 \setminus E_1) \setminus A$, i.e. $u(\gamma(c)) = 0 = u(\gamma(l_\gamma))$. Hence

$$|u(\gamma(0)) - u(\gamma(l_\gamma))| = |u(\gamma(0)) - u(\gamma(c))| \leq \int_{\gamma|_{[0,c]}} g_{u,E_2}\, ds \leq \int_\gamma g_{u,E_2}\, ds. \qquad \square$$

As $N^{1,p}_0(E)$ is defined through $N^{1,p}(X)$, it is natural that the minimal p-weak upper gradients of functions in $N^{1,p}_0(E)$ are taken with respect to X. The following result is therefore important for our considerations. This result holds for $u \in D^p_0(E)$ even in situations when $N^{1,p}_0(E) \subsetneq D^p_0(E)$ so we formulate it in this generality. In fact it even holds for $u \in D^p_{\mathrm{loc},0}(E) := \{f|_E : f \in D^p_{\mathrm{loc}}(X) \text{ and } f = 0 \text{ on } X \setminus E\}$.

Proposition 3.10. *Let $E \subset X$ be measurable and let $u \in D^p_0(E)$ with a minimal p-weak upper gradient g_u (with respect to X, and with $u = 0$ outside E). Then $g_{u,E} = g_u|_E$ a.e. in E, i.e., $g_u|_E$ is a minimal p-weak upper gradient of u with respect to E.*

Note that the corresponding result for arbitrary $u \in N^{1,p}(X)$ is false; see Example 3.6.

Proof. Clearly, $g_u|_E$ is a p-weak upper gradient of u in E. To show that it is minimal, we shall show that the function

$$g = \begin{cases} g_{u,E} & \text{in } E, \\ 0 & \text{in } X \setminus E, \end{cases}$$

is a p-weak upper gradient of u in X. Let Γ be the set of curves in E on which (2.1) fails for u and $g_{u,E}$. Then $\mathrm{Mod}_p(\Gamma) = 0$.

Let $\gamma \colon [0, l_\gamma] \to X$ be a curve such that u is absolutely continuous along it and γ does not have any subcurve in Γ. As $u \in D_0^p(E)$ and $\mathrm{Mod}_p(\Gamma) = 0$, p-almost every curve in X has these properties. We can also assume that γ passes through both E and $X \setminus E$. Otherwise, there is nothing to prove, since $g_{u,E}$ is a p-weak upper gradient in E and $u = 0$ outside E. By splitting γ into two parts and reversing the orientation, if necessary, we can assume that $\gamma(0) \in E$ and $\gamma(l_\gamma) \in X \setminus E$.

Let $c = \inf\{t \in [0, l_\gamma] : \gamma(t) \in X \setminus E\}$. Since $u = 0$ in $X \setminus E$, the continuity of u along γ implies that $u(\gamma(c)) = 0 = u(\gamma(l_\gamma))$. If $c > 0$, then

$$|u(\gamma(0)) - u(\gamma(l_\gamma))| = |u(\gamma(0)) - u(\gamma(c))| = \lim_{t \to c^-} |u(\gamma(0)) - u(\gamma(t))|$$

$$\leq \lim_{t \to c^-} \int_{\gamma|_{[0,t]}} g_{u,E} \, ds \leq \int_\gamma g \, ds,$$

by the continuity of u along γ. For $c = 0$, these estimates are trivial. Thus g is a p-weak upper gradient of u in X, and hence $g \geq g_u$ a.e. in X. It follows that $g_{u,E} \leq g_u \leq g_{u,E}$ a.e. in E, which finishes the proof. $\qquad\square$

4. The obstacle problem

Throughout this section we assume that $p > 1$ and that X supports the p-Friedrichs inequality (2.3). We also assume that $E \subset X$ is a bounded measurable set such that $C_p(X \setminus E) > 0$.

In Section 5 we will discuss when these assumptions can be relaxed. Observe that we do not assume that μ is doubling or that X is complete, although we will need to add these assumptions for parts of the theory in Sections 7 and 8.

We shall now consider the obstacle and Dirichlet problems on general sets. We start by formulating the obstacle problem studied in this paper.

Definition 4.1. Let $A \subset X$ be a bounded measurable set such that $C_p(X \setminus A) > 0$. Let $f \in D^p(A)$ and $\psi_1, \psi_2 : A \to \overline{\mathbb{R}}$. Then we define

$$\mathcal{K}_{\psi_1, \psi_2, f}(A) = \{v \in D^p(A) : v - f \in N_0^{1,p}(A) \text{ and } \psi_1 \leq v \leq \psi_2 \text{ q.e. in } A\}.$$

Furthermore, a function $u \in \mathcal{K}_{\psi_1,\psi_2,f}(A)$ is a *solution of the* $\mathcal{K}_{\psi_1,\psi_2,f}(A)$-*obstacle problem* if

$$(4.1) \qquad \int_A g_{u,A}^p \, d\mu \leq \int_A g_{v,A}^p \, d\mu \quad \text{for all } v \in \mathcal{K}_{\psi_1,\psi_2,f}(A).$$

If $A = E$ we often omit the set from the notation and simply write $\mathcal{K}_{\psi_1,\psi_2,f} := \mathcal{K}_{\psi_1,\psi_2,f}(E)$. Similarly, we often omit ψ_2 from the notation when $\psi_2 \equiv \infty$, i.e., when there is no upper obstacle. Such an obstacle problem is called a *single obstacle problem*.

The Dirichlet problem is the special case of the obstacle problem with the obstacles $\psi_1 \equiv -\infty$ and $\psi_2 \equiv \infty$. Note that the boundary data f are only required to belong to $D^p(A)$ and need not be defined on ∂A.

Since we consider boundary values $f \in D^p(A)$ rather than $f \in N^{1,p}(A)$, it would be natural to consider the obstacle problem with $D_0^p(A)$ instead of $N_0^{1,p}(A)$. By Proposition 2.7, this is exactly what we do, even though we prefer to write $N_0^{1,p}(A)$. On the other hand, in the more general situations discussed in Section 5 the equality $D_0^p(A) = N_0^{1,p}(A)$ may not hold, and it will be essential to consider the obstacle problem with $N_0^{1,p}(A)$, at least for our proof of Theorem 4.2 (through the use of Lemma A.2).

The p-weak upper gradients $g_{u,A}$ and $g_{v,A}$ in Definition 4.1 are taken with respect to A, but quasi-everywhere is taken with respect to X. Below we comment on obstacle problems with quasi-everywhere taken with respect to E and with almost everywhere inequalities.

Obstacle and Dirichlet problems have traditionally been solved on open sets Ω, in which case $g_{u,\Omega} = g_u$ a.e. See, however, Kilpeläinen–Malý [24] and Malý–Ziemer [31], where they are studied on finely open sets in $\mathbb{R}^n$. In metric spaces the single obstacle problem was studied by Kinnunen–Martio [26], while the double obstacle problem was studied by Farnana [15]. In both cases they studied the obstacle problems for bounded open sets in a complete metric space X supporting a $(1,p)$-Poincaré inequality and with a doubling measure μ (and with boundary values in the Newtonian space).

The Dirichlet problem on metric spaces was first studied by Shanmugalingam [37]. She studied it on bounded, not necessarily open, sets in a complete metric space X with a doubling measure μ supporting a $(1,p)$-Poincaré inequality, under the stronger requirement that $f \in N^{1,p}(X)$.

In all the above cases, the p-energy functional was defined by means of the global minimal p-weak upper gradient g_u. Thus, the Dirichlet problem studied by Shanmugalingam [37] differs in general from the Dirichlet problem considered here. Similarly, for a nonopen set E, another possible generalization of the obstacle problem would be to require that the boundary data f belong to $D^p(\Omega)$ for some open set $\Omega \supset E$ and to minimize the energy $\int_E g_v^p \, d\mu$ among all $v \in \mathcal{K}'_{\psi_1,\psi_2,f}$, where

$$(4.2) \qquad \mathcal{K}'_{\psi_1,\psi_2,f} = \{v \in D^p(\Omega) : v - f \in N_0^{1,p}(E) \text{ and } \psi_1 \leq v \leq \psi_2 \text{ q.e. in } E\}.$$

As Ω is open, the minimal p-weak upper gradients and the notion of quasi-everywhere are taken with respect to Ω or equivalently X. If we let

$$(4.3) \qquad \psi'_j = \begin{cases} \psi_j & \text{in } E, \\ f & \text{in } \Omega \setminus E, \end{cases} \qquad j = 1, 2,$$

then $\mathcal{K}'_{\psi_1,\psi_2,f} = \mathcal{K}_{\psi'_1,\psi'_2,f}(\Omega)$, where we use our convention that $v - f \in N_0^{1,p}(E)$ can be extended by zero in $\Omega \setminus E$. Moreover, for any $v \in \mathcal{K}'_{\psi_1,\psi_2,f}$,

$$\int_\Omega g_v^p \, d\mu = \int_{\Omega \setminus E} g_f^p \, d\mu + \int_E g_v^p \, d\mu,$$

as $v = f$ q.e. in $\Omega \setminus E$. Hence the minimizers of the energies $\int_E g_v^p \, d\mu$ and $\int_\Omega g_v^p \, d\mu$ among $v \in \mathcal{K}'_{\psi_1,\psi_2,f} = \mathcal{K}_{\psi'_1,\psi'_2,f}(\Omega)$ coincide and the theory for this generalization follows directly from the theory for open sets. Observe however that we study the obstacle problem on more general metric measure spaces than have been considered previously, also for open sets (see e.g. Example 5.2 and Section 10) and that we require only that f belong to the Dirichlet space D^p.

Here we have ignored one subtle point, namely we require $C_p(X \setminus E) > 0$, but it is not clear if one can find an open set $\Omega \supset E$ such that $C_p(X \setminus \Omega) > 0$. This is always possible if X is unbounded, and also if $\mu(X \setminus E) > 0$, by the regularity of the measure and the measurability of E. Similarly, if E is a G_δ set, then Ω can be found using an analogue for the C_p-capacity of the property (iii) in Theorem B.3. Moreover, if X is a complete metric space supporting a $(1, p)$-Poincaré inequality, μ is doubling, and $X \setminus E$ is Suslin (in particular if E is Borel), then the same follows from the Choquet capacitability theorem; see Theorem 6.11 in Björn–Björn [6].

In our approach, we only assume that the boundary data belong to $D^p(E)$ and the minimal p-weak upper gradient is taken with respect to E. This leads to a different obstacle problem since $g_{u,E}$ is in general smaller than g_u; see Example 3.6. The two definitions of obstacle problems will be further compared in Section 8.

Note that even though we take the gradients with respect to E, we require the obstacle inequalities $\psi_1 \le u \le \psi_2$ to hold quasi-everywhere (taken with respect to X). This may seem unnatural, but there are several reasons for this choice. First, this is the natural condition for $N_0^{1,p}(E)$ and means that the uniqueness we obtain in Theorem 4.2 is precisely up to sets of capacity zero with respect to X (not E). It also turns out to be essential for the Adams criterion (Theorem 6.1).

Second, we could actually have developed the theory with quasi-everywhere taken with respect to E, which is a coarser condition, or with the even coarser condition "almost everywhere". The latter was used by Kinnunen–Martio [26]. See also the discussion on quasi-everywhere and almost everywhere obstacle problems in Farnana [15]. In particular, if $C_p(A) > 0 = C_p^E(A)$, where $C_p^E(A)$ is the capacity of A with respect to E, then the zero function belongs to $\mathcal{K}_{\chi_A,0}$ (and solves the obstacle problem) with quasi-everywhere taken with respect to E, but not when taken with respect to X. The E-quasi-everywhere theory would fall in between the quasi-everywhere and almost everywhere theories, and it is easy to adapt most of our results to this setting if need arises, but there is no direct counterpart of the Adams criterion.

We have the following existence and uniqueness theorem.

Theorem 4.2. *Let $f \in D^p(E)$ and $\psi_1, \psi_2 : E \to \overline{\mathbb{R}}$. If $\mathcal{K}_{\psi_1,\psi_2,f} \neq \varnothing$, then there is a unique solution (up to sets of capacity zero in X) of the $\mathcal{K}_{\psi_1,\psi_2,f}$-obstacle problem.*

Proof. Let

$$I = \inf_{v \in \mathcal{K}_{\psi_1,\psi_2,f}} \int_E g_{v,E}^p \, d\mu.$$

Since $\mathcal{K}_{\psi_1,\psi_2,f} \neq \varnothing$, we have $0 \leq I < \infty$. Let $\{u_j\}_{j=1}^\infty \subset \mathcal{K}_{\psi_1,\psi_2,f}$ be a minimizing sequence such that

$$\int_E g_{u_j,E}^p \, d\mu \searrow I, \quad \text{as } j \to \infty.$$

Then $\{g_{u_j,E}\}_{j=1}^\infty$ is bounded in $L^p(E)$. Remember that $u_j \in D^p(E)$ and that $g_{u_j,E}$ is taken with respect to E.

Using (2.3) and Proposition 3.10 we find that

$$\int_E |u_j - f|^p \, d\mu \leq C \int_E g_{u_j-f}^p \, d\mu = C \int_E g_{u_j-f,E}^p \, d\mu \leq C \int_E g_{u_j,E}^p \, d\mu + C \int_E g_{f,E}^p \, d\mu.$$

Thus $\{u_j - f\}_{j=1}^\infty$ is bounded in $N^{1,p}(E)$. By Lemma A.2 (with X replaced by E), we can find convex combinations $v_j = \sum_{k=j}^{N_j} a_{j,k} u_k$ with p-weak upper gradients $g_j = \sum_{k=j}^{N_j} a_{j,k} g_{u_k,E}$ on E and limit functions v and g such that $v - f \in N^{1,p}(E)$, both $v_j - v \to 0$ and $g_j \to g$ in $L^p(E)$, as $j \to \infty$, and such that g is a p-weak upper gradient of v with respect to E.

Furthermore, $w_j := v_j - f \in N_0^{1,p}(E)$ and we can thus consider w_j to be identically zero outside of E. Let $w = v - f$, $g_j' = g_j + g_{f,E}$ and $g' = g + g_{f,E}$, all three considered to be identically zero outside of E. Proposition 3.10 implies that

$$g_{w_j} = g_{w_j,E} \leq g_j + g_{f,E} = g_j' \quad \text{a.e. in } E.$$

As $g_{w_j} = 0$ a.e. in $X \setminus E$, we see that g_j' is a p-weak upper gradient of w_j in X, $j = 1, 2, \ldots$ We also have that $w_j \to w$ and $g_j' \to g'$ in $L^p(X)$, as $j \to \infty$. Proposition A.1 yields that there exists $\widetilde{w} \in N^{1,p}(X)$ such that $w = \widetilde{w}$ a.e. in X. Then $u := f + \widetilde{w} \in D^p(E)$ and $u = v$ a.e. in E. Since $u, v \in D^p(E)$, we have $u = v$ E-q.e. in E (i.e., quasi-everywhere with respect to E), and thus g is a p-weak upper gradient of u as well, with respect to E.

Proposition A.1 also implies that a subsequence of $\{w_j\}_{j=1}^\infty$ converges quasi-everywhere (with respect to X) to $\widetilde{w}$. As $\psi_1 \leq v_j \leq \psi_2$ q.e. in E, this implies that $\psi_1 \leq u \leq \psi_2$ q.e. in E. Moreover, it implies that $\widetilde{w} = 0$ q.e. in $X \setminus E$ and thus $u - f = \widetilde{w} \in N_0^{1,p}(E)$. Hence $u \in \mathcal{K}_{\psi_1,\psi_2,f}$. Since

$$I \leq \int_E g_{u,E}^p \, d\mu \leq \int_E g^p \, d\mu = \lim_{j \to \infty} \int_E g_j^p \, d\mu = I,$$

we conclude that u is the desired minimizer.

To prove the uniqueness, assume that u_1 and u_2 are solutions. Then also $u' = \frac{1}{2}(u_1 + u_2) \in \mathcal{K}_{\psi_1,\psi_2,f}$ and thus

$$I \leq \|g_{u',E}\|_{L^p(E)} \leq \left\|\tfrac{1}{2}(g_{u_1,E} + g_{u_2,E})\right\|_{L^p(E)}$$
$$\leq \tfrac{1}{2}\|g_{u_1,E}\|_{L^p(E)} + \tfrac{1}{2}\|g_{u_2,E}\|_{L^p(E)} = I.$$

Hence $g_{u_1,E} = g_{u_2,E}$ a.e. in E by the strict convexity of $L^p(E)$. We shall show that $g_{u_1-u_2,E} = 0$ a.e. in E. Since $u_1 - u_2 \in N_0^{1,p}(E)$, (2.3) and Proposition 3.10 then yield $\|u_1 - u_2\|_{L^p(E)} = 0$. From this it follows that $u_1 - u_2 = 0$ a.e. in E and thus in X (when we set $u_1 - u_2 := 0$ in $X \setminus E$). As $u_1 - u_2 \in N_0^{1,p}(E) \subset N^{1,p}(X)$, we obtain $u_1 - u_2 = 0$ q.e. in X, and hence $u_1 = u_2$ q.e. in E. (Note that since we consider quasi-everywhere with respect to X, we have to use the fact that $u_1 - u_2 \in N^{1,p}(X)$ rather than $u_1 - u_2 \in N^{1,p}(E)$.)

To show that $g_{u_1-u_2,E} = 0$ a.e. in E, let $c \in \mathbb{R}$ and

$$u = \max\{u_1, \min\{u_2, c\}\}.$$

Then $u - f \in N^{1,p}(E)$ and $\psi_1 \leq u \leq \psi_2$ q.e. in E. Also,

$$u - f \leq \max\{u_1, u_2\} - f = \max\{u_1 - f, u_2 - f\} \in N_0^{1,p}(E)$$

and $u - f \geq u_1 - f \in N_0^{1,p}(E)$. Lemma 2.4 shows that $u - f \in N_0^{1,p}(E)$ and hence $u \in \mathcal{K}_{\psi_1,\psi_2,f}$.

Let $V_c = \{x \in E : u_1(x) < c < u_2(x)\}$ and note that $V_c \subset \{x \in E : u(x) = c\}$ and hence $g_{u,E} = 0$ a.e. in V_c. The minimizing property of $g_{u_1,E}$ then implies that

$$(4.4) \qquad \int_E g_{u_1,E}^p \, d\mu \leq \int_E g_{u,E}^p \, d\mu = \int_{E \setminus V_c} g_{u,E}^p \, d\mu = \int_{E \setminus V_c} g_{u_1,E}^p \, d\mu,$$

since $g_{u,E} = g_{u_1,E} = g_{u_2,E}$ a.e. in $E \setminus V_c$. From (4.4) we conclude that $g_{u_2,E} = g_{u_1,E} = 0$ a.e. in V_c for all $c \in \mathbb{R}$. Now,

$$\{x \in E : u_1(x) < u_2(x)\} \subset \bigcup_{c \in \mathbb{Q}} V_c$$

and hence $g_{u_2,E} = g_{u_1,E} = 0$ a.e. in $\{x \in E : u_1(x) < u_2(x)\}$, and similarly in $\{x \in E : u_1(x) > u_2(x)\}$. It follows that

$$g_{u_1-u_2,E} \leq (g_{u_1,E} + g_{u_2,E})\chi_{\{x \in E : u_1(x) \neq u_2(x)\}} = 0 \quad \text{a.e. in } E,$$

and thus $u_1 = u_2$ q.e. by the above argument.

It remains to show that if u is a solution and $v = u$ q.e., then v is also a solution. Indeed, it follows directly that $v \in \mathcal{K}_{\psi_1,\psi_2,f}$. Moreover, $v = u$ E-q.e., and thus $g_{u,E} = g_{v,E}$ a.e., so that

$$\int_E g_{v,E}^p \, d\mu = \int_E g_{u,E}^p \, d\mu,$$

showing that v must also be a solution. $\qquad\square$

The following comparison principle follows from the uniqueness of the solutions and is useful in various applications. Note again that the boundary data f and f' are only defined on E. But if $f, f' \in N^{1,p}(\overline{E})$ and $f \leq f'$ q.e. on ∂E, then Lemma 3.9 implies that the condition $(f - f')_+ \in N_0^{1,p}(E)$ is satisfied. Recall that $f_+ = \max\{f, 0\}$ and $f_- = \max\{-f, 0\}$.

Corollary 4.3. (Comparison principle) *Assume that $f, f' \in D^p(E)$ and $\psi_j, \psi_j' :$ $E \to \overline{\mathbb{R}}$, $j = 1, 2$, are such that $\mathcal{K}_{\psi_1, \psi_2, f}$ and $\mathcal{K}_{\psi_1', \psi_2', f'}$ are nonempty. Let u and u' be solutions of the $\mathcal{K}_{\psi_1, \psi_2, f}$- and $\mathcal{K}_{\psi_1', \psi_2', f'}$-obstacle problems, respectively. If $\psi_j \leq \psi_j'$ q.e. in E, $j = 1, 2$, and $(f - f')_+ \in N_0^{1,p}(E)$, then $u \leq u'$ q.e. in E.*

In the next section we discuss relaxations of the conditions imposed in this section. For the comparison principle to hold it is enough that one of the obstacle problems is quasi-everywhere uniquely soluble (and the other soluble). (In the proof, the uniqueness of the $\mathcal{K}_{\psi_1, \psi_2, f}$-obstacle problem is used, but by symmetry one can equally well use the uniqueness of the $\mathcal{K}_{\psi_1', \psi_2', f'}$-obstacle problem.)

Proof. Let $w = \min\{u, u'\}$ and $h = u - f - (u' - f') \in N_0^{1,p}(E)$. It follows that

$$-(f - f')_+ - h_- = -(f' - f)_- - h_- \leq \min\{f' - f, h\} \leq h.$$

Lemma 2.4 then implies that $\min\{f' - f, h\} \in N_0^{1,p}(E)$ and hence

$$w - f = \min\{u' - f, u - f\} = u' - f' + \min\{f' - f, h\} \in N_0^{1,p}(E).$$

As $\psi_1 \leq w \leq \psi_2$ q.e. in E, we get $w \in \mathcal{K}_{\psi_1, \psi_2, f}$.

Similarly $v = \max\{u, u'\} \in \mathcal{K}_{\psi_1', \psi_2', f'}$. Let

$$A = \{x \in E : u(x) > u'(x)\}.$$

Since u' is a solution of the $\mathcal{K}_{\psi_1', \psi_2', f'}$-obstacle problem, we have

$$\int_E g_{u', E}^p \, d\mu \leq \int_E g_{v, E}^p \, d\mu = \int_A g_{u, E}^p \, d\mu + \int_{E \setminus A} g_{u', E}^p \, d\mu.$$

Thus

$$\int_A g_{u', E}^p \, d\mu \leq \int_A g_{u, E}^p \, d\mu.$$

It follows that

$$\int_E g_{w, E}^p \, d\mu = \int_A g_{u', E}^p \, d\mu + \int_{E \setminus A} g_{u, E}^p \, d\mu \leq \int_A g_{u, E}^p \, d\mu + \int_{E \setminus A} g_{u, E}^p \, d\mu = \int_E g_{u, E}^p \, d\mu.$$

Since u is a solution of the $\mathcal{K}_{\psi_1, \psi_2, f}$-obstacle problem, so is w. By uniqueness $u = w = \min\{u, u'\}$ q.e. in E, and thus $u \leq u'$ q.e. in E. $\qquad\square$

5. Assumptions and examples

Both in the existence and the uniqueness parts of the proof of Theorem 4.2 we used the "extra" assumptions that $p > 1$ (through the use of Lemma A.2 and the strict convexity of L^p), that $C_p(X \setminus E) > 0$ and that X supports a p-Friedrichs inequality. It is worthwhile discussing when these assumptions hold and whether they could possibly be omitted or weakened. We start by discussing the p-Friedrichs inequality. By the following lemma it follows from the (p, p)-Poincaré inequality on large balls.

Lemma 5.1. *Assume that for every ball $B \subset X$ there is a constant $C_B > 0$ such that for all $u \in N_0^{1,p}(B)$,*

$$(5.1) \qquad \int_B |u - u_B|^p \, d\mu \leq C_B \int_B g_u^p \, d\mu.$$

Then X supports the p-Friedrichs inequality (2.3).

Since $g_u = 0$ outside B there is no reason to have a dilation constant λ in (5.1), as in Definition 2.5. Note also that the doubling property of μ is not needed. The proof of Lemma 5.1 was inspired by Theorem 10.1.2 in Maz'ya [32] and Proposition 3.2 in J. Björn [11], but is slightly simpler and sufficient for our purpose.

For unbounded X we always have $C_p(X \setminus B) > 0$ and hence (5.1) follows from (2.3) by means of the Hölder and Minkowski inequalities. Thus, the p-Friedrichs inequality and (5.1) are equivalent in unbounded spaces. The case when X is bounded is more subtle, since we cannot take $E = X$ in (2.3). We do not know if the equivalence is true in this case.

It is also worth observing that contrary to the case of the classical Poincaré inequalities, here it is enough to require (5.1) or (2.2) for large balls, i.e., that for every ball B' there is a ball $B \supset B'$ such that (5.1) or (2.2) holds. (If X is bounded it suffices to assume that (5.1) or (2.2) holds for $B = X$.) The following example shows that this is not equivalent to (2.2) holding for all balls.

Example 5.2. Let $X \subset \mathbb{R}^2$ be the graph of the function $y = x^\alpha \sin(\pi \log_2 x)$, $0 < \alpha < 1$, $0 \leq x \leq 1$, with the $\mathbb{R}^2$-Euclidean metric and the arc length measure Λ_1. It is easily verified that $L := \Lambda_1(X) < \infty$. Let $\gamma \colon [0, L] \to X$ be an arc length parameterized curve such that $\gamma(0) = (0, 0)$ and $\gamma(L) = (1, 0)$. Since γ gives a natural bijection between X and $[0, L]$, every function in $N^{1,p}(X) = N_0^{1,p}(X)$ is absolutely continuous on X with $g_u(\gamma(t)) = |(u \circ \gamma)'(t)|$ a.e.

Let $z = (2^{-k}, 0) \in X$ and $2^{-k-1} < r < 2^{-k}$, $k = 1, 2, \ldots$ Then the ball $B = B(z, r)$ is not connected and B does not even belong to a single component of $\lambda_k B$, where $\lambda_k = 2^{k(1-\alpha)-1}$. Letting $k \to \infty$ shows that X cannot support any Poincaré inequality with the same dilation constant λ for all balls. At the same time, the p-Friedrichs inequality (2.3) holds on X, since

$$\int_X |u - u_X|^p \, d\Lambda_1 = \int_0^L |u(\gamma(t)) - u_X|^p \, dt \leq C \int_0^L |(u \circ \gamma)'(t)|^p \, dt = C \int_0^L g_u^p \, d\Lambda_1,$$

by the (p, p)-Poincaré inequality for $[0, L]$.

Proof of Lemma 5.1. Let $E \subset X$ be bounded and such that $C_p(X \setminus E) > 0$. Let $u \in N_0^{1,p}(E)$, extended by zero in $X \setminus E$. We can assume that the left-hand side in (2.3) is nonzero.

If X is unbounded, let $B \supset E$ be a ball such that $\mu(E) < \mu(B)$. Then

$$(5.2) \qquad \Big(\int_B |u|^p \, d\mu \Big)^{1/p} \leq \Big(\int_B |u - u_B|^p \, d\mu \Big)^{1/p} + |u_B| \, \mu(B)^{1/p}.$$

The first term on the right-hand side is estimated using (5.1) and for the second term we have, using the Hölder inequality and the fact that u vanishes outside E, that

$$|u_B| \, \mu(B)^{1/p} \leq \frac{1}{\mu(B)^{1-1/p}} \int_B |u| \, d\mu \leq \Big(\frac{\mu(E)}{\mu(B)} \Big)^{1-1/p} \Big(\int_B |u|^p \, d\mu \Big)^{1/p}.$$

Since $\mu(E) < \mu(B)$, inserting this into (5.2) and subtracting the last term from both sides of (5.2) proves (2.3) for unbounded X.

If X is bounded, let

$$\bar{u} = \Big(\fint_X |u|^p \, d\mu \Big)^{1/p}.$$

Then $v := 1 - u/\bar{u}$ is admissible in the definition of $C_p(X \setminus E)$ and hence

$$(5.3) \quad 0 < C_p(X \setminus E) \leq \int_X v^p \, d\mu + \int_X g_v^p \, d\mu \leq \frac{1}{\bar{u}^p} \Big(\int_X |u - \bar{u}|^p \, d\mu + \int_X g_u^p \, d\mu \Big).$$

The first integral on the right-hand side can be estimated as

$$\|u - \bar{u}\|_{L^p(X)} \leq \|u - u_X\|_{L^p(X)} + \|\bar{u} - u_X\|_{L^p(X)},$$

where for the second term we have

$$\|\bar{u} - u_X\|_{L^p(X)} = \big| \|u\|_{L^p(X)} - \|u_X\|_{L^p(X)} \big| \leq \|u - u_X\|_{L^p(X)}.$$

Inserting this into (5.3) and using (5.1) with $B = X$ we obtain

$$\bar{u}^p \leq \frac{C}{C_p(X \setminus E)} \int_X g_u^p \, d\mu. \qquad \qquad \Box$$

The following two examples show that neither the existence nor the uniqueness of solutions remains valid for $p = 1$. Note that in Examples 5.3–5.7 we have $f \in N^{1,p}(E)$ and E is open.

Example 5.3. Let $X = \mathbb{R}$, $p = 1$, $E = (0, 1)$, $f(x) = x$, $\psi = -\infty$ and $d\mu = w \, dx$, where

$$w(x) = \begin{cases} 1 + x, & 0 < x < 1, \\ 1, & \text{otherwise,} \end{cases}$$

i.e., we consider a weighted Dirichlet problem. Note that μ is a doubling measure supporting a $(1,1)$-Poincaré inequality. Let $u_j(x) = \min\{jx, 1\} \in \mathcal{K}_{\psi,f}$, $j = 1, 2, \ldots$, so that

$$\int_E g_{u_j} \, d\mu = \int_0^{1/j} j \, d\mu = \int_0^{1/j} j(1 + x) \, dx = 1 + \frac{1}{2j} \to 1, \quad \text{as } j \to \infty.$$

On the other hand, for any $v \in \mathcal{K}_{\psi,f}$ we have that

$$1 = v(1) - v(0) = \int_0^1 v' \, dx \leq \int_0^1 |v'| \, dx = \int_0^1 g_v \, dx < \int_0^1 g_v \, d\mu,$$

since g_v cannot vanish almost everywhere. This shows that the minimum is not attained and thus there are no minimizers. Hence the assumption $p > 1$ cannot be removed for the existence part.

Example 5.4. Let $X = \mathbb{R}$ (unweighted), $p = 1$, $E = (0,1)$, $f(x) = x$ and $\psi = -\infty$. In this case any increasing absolutely continuous function $u : [0,1] \to [0,1]$ with $u(0) = 0$ and $u(1) = 1$ will be a solution of the $\mathcal{K}_{\psi,f}$-obstacle problem (i.e., of the Dirichlet problem with f as boundary values). Thus the assumption $p > 1$ cannot be omitted for the uniqueness part either.

It is more complicated to see when the Friedrichs inequality is essential. First we consider the question of existence of solutions.

Example 5.5. Let $1 < p < 2$ and

$$\begin{aligned}
X &= \{(x,y) \in [-2,2]^2 : xy \geq 0\}, \\
X_+ &= \{(x,y) \in X : x \geq 0\} \setminus \{(0,0)\} = [0,2]^2 \setminus \{(0,0)\}, \\
X_- &= \{(x,y) \in X : x \leq 0\} = [-2,0]^2.
\end{aligned}$$

Then there are p-almost no curves between X_+ and X_- (since $C_p(\{(0,0)\}) = 0$) which means that in this context they can be thought of as disconnected; see Example 5.6 in Björn–Björn [6]. In particular, $u = \chi_{X_+} \in N_0^{1,p}(X_+)$ with $g_u = 0$, showing that the p-Friedrichs inequality (2.3) is violated. (This also shows that the zero p-weak upper gradient property, introduced below, fails at $(0,0)$.)

Let $f = 0$, $1 - 2/p < \alpha < 0$, $E = X_+$,

$$\psi(x,y) = \begin{cases} |(x,y) - (1,1)|^\alpha, & (x,y) \in X_+, \\ 0, & (x,y) \in X_-, \end{cases}$$

and $u_j = \max\{\psi, j\}\chi_{X_+}$.

Then $\psi \in N^{1,p}(X)$ and $u_j - f \in N_0^{1,p}(X_+)$, i.e., $u_j \in \mathcal{K}_{\psi,f}(X_+)$. Moreover,

$$\int_{X_+} g_{u_j}^p \, d\mu \to 0, \quad \text{as } j \to \infty.$$

On the other hand, if $N^{1,p}(X) \ni v \geq \psi$ q.e. in X_+ then necessarily $\int_{X_+} g_v^p \, d\mu > 0$, and there does not exist any minimizer for the $\mathcal{K}_{\psi,f}(X_+)$-obstacle problem.

A problem with Example 5.5 is that $C_p(\partial X_+) = 0$ even though $C_p(X \setminus X_+) > 0$, allowing for $u_j \in N_0^{1,p}(X_+)$. Similarly, the same functions u_j show that the $\mathcal{K}_{\psi,f}(\Omega)$-obstacle problem with $\Omega = \{(x,y) \in X : x > -1\}$ is also not soluble. Here, the problem is that Ω is essentially disconnected and thus the boundary values f have no influence in X_+, even though $C_p(\partial\Omega) > 0$. (In fact, Ω need not be connected, but it should not have a component which is essentially disconnected from $X \setminus \Omega$.)

A Poincaré inequality of some kind prevents these difficulties and guarantees solubility of the obstacle problem. The above $\mathcal{K}_{\psi,f}(\Omega)$-obstacle problem also shows that it is not enough to just replace the assumption $C_p(X \setminus E) > 0$ with $C_p(\partial E) > 0$. Under a Poincaré inequality and for open E, these two conditions are equivalent by Lemma 4.5 in Björn–Björn [6]. For open E in general spaces, the latter condition is stronger, as seen above. On the other hand, the former condition can be stronger for nonopen sets. We have therefore chosen to use the condition $C_p(X \setminus E) > 0$, as it is closely related to $N_0^{1,p}(E)$.

Remark 5.6. On the other hand, if the data f, ψ_1 and ψ_2 are bounded then we can omit both the assumption of a Friedrichs inequality and the assumption $C_p(X \setminus E) > 0$. This will be important for Theorem 5.13. In the existence part of the proof, they were only used to deduce that $\{u_j - f\}_{j=0}^\infty$ is bounded in $L^p(E)$, and this can be deduced more directly if the data are bounded. More precisely, consider the following two cases:

(a) $\psi_j \in L^p(E)$, which in particular holds if $\psi_j \in L^\infty(E)$, $j = 1,2$;

(b) there exists $M < \infty$ such that $|f| \leq M$ a.e. in E, $\psi_1 \leq M$ q.e. in E and $\psi_2 \geq -M$ q.e. in E.

In case (a), the L^p-boundedness of $\{u_j\}_{j=1}^\infty$ follows directly since $\psi_1 \leq u_j \leq \psi_2$ a.e. In case (b) we may replace u_j by the truncations

$$u_j' := \max\{\min\{u_j, M\}, -M\}$$

at levels M and $-M$. Then $g_{u_j',E} \leq g_{u_j,E}$ and the sequence $\{u_j'\}_{j=1}^\infty$ is bounded in $N^{1,p}(E)$. (In both cases one uses the L^p-boundedness of $\{u_j\}_{j=1}^\infty$ (or $\{u_j'\}_{j=1}^\infty$) rather than the L^p-boundedness of $\{u_j - f\}_{j=1}^\infty$, in the proof of Theorem 4.2. This also makes the proof a little easier.)

We now turn to the question of uniqueness. The following example shows that we cannot drop the Friedrichs inequality entirely.

Example 5.7. Let X be the von Koch snowflake curve. For $a, b \in X$, $a \neq b$, let E be one of the two components of $X \setminus \{a, b\}$. Set $f = 0$ and $\psi = -\infty$. Since there are no rectifiable curves in X, we have $N_0^{1,p}(E) = L^p(E)$ and $g_u \equiv 0$ for all $u \in N_0^{1,p}(E)$, which means that any $u \in L^p(E)$ is a solution of the $\mathcal{K}_{\psi,f}$-obstacle problem (i.e., of the Dirichlet problem with f as boundary values). Thus the assumption that X supports some kind of Poincaré or Friedrichs inequality cannot be omitted for the uniqueness part. Similar arguments apply to other spaces without rectifiable curves, or with p-almost no rectifiable curves.

Remark 5.8. Even though the Poincaré or Friedrichs inequality cannot be omitted for the uniqueness part, it can be weakened. In the uniqueness part of the proof, it was only used to deduce that $v := u_1 - u_2 = 0$ a.e. in E from the fact that $g_v = 0$ in X and $v = 0$ outside E.

In A. Björn [3] the following weaker property was introduced: X has the *zero p-weak upper gradient property* if every measurable function f, which has 0 as a p-weak upper gradient in some ball $B(x, r)$, is essentially constant in some (possibly smaller) ball $B(x, \delta)$, which can depend both on f and $B(x, r)$. By considering the bounded function $h = \arctan f$ with $g_h = g_f/(1 + f^2)$, we easily conclude that one can equivalently consider only bounded measurable functions in the definition of the zero p-weak upper gradient property.

Thus, Lemma 3.2 in [3] shows that, when proving uniqueness of the solutions, we may replace the Friedrichs inequality by the zero p-weak upper gradient property, together with the fact that $C_p(G \setminus E) > 0$ for every component G of X. The latter is essential since there are nonconnected spaces having the zero p-weak upper gradient property, e.g. $X = [0, 1]^2 \cup [2, 3]^2$ in $\mathbb{R}^2$ and $X = [0, 1] \cup [2, 3]$ in $\mathbb{R}$.

The zero p-weak upper gradient property is a weaker assumption than a $(1, p)$-Poincaré inequality (as the two examples above show). On the other hand, the following example shows that X can support a p-Friedrichs inequality and at the same time fail to have the zero p-weak upper gradient property.

Example 5.9. Let $1 < p \leq 2$ and let

$$X = \{(x, y) \in \mathbb{R}^2 : xy \geq 0\},$$
$$X_+ = \{(x, y) \in X : x \geq 0\},$$
$$X_- = \{(x, y) \in X : x \leq 0\}.$$

As in Example 5.5, the function χ_{X_+} shows that the zero p-weak upper gradient property fails for all balls centred at the origin.

On the other hand, as both X_+ and X_- support (p, p)-Poincaré inequalities, they support the p-Friedrichs inequality (2.3), by Lemma 5.1. Considering $u|_{X_+}$ and $u|_{X_-}$ separately shows that for all bounded $E \subset X$ and all $u \in N_0^{1,p}(E)$ we have

$$\int_{X_\pm} |u|^p \, d\mu \leq C_{E_\pm} \int_{X_\pm} g_u^p \, d\mu,$$

where $E_\pm = E \cap X_\pm$. The p-Friedrichs inequality on X then follows by adding the L^p-norms on X_+ and X_-.

Finally, we discuss the assumption $C_p(X \setminus E) > 0$. If it fails (and thus necessarily X is bounded) then we lose existence in general. This is easily seen by letting $X = [0, 2]^2 \subset \mathbb{R}^2$ and using the construction in Example 5.5. However, we do have solubility if we assume boundedness of the data as in (a) or (b) of Remark 5.6. Moreover, the Dirichlet problem (i.e., the obstacle problem without obstacles) is always soluble if $C_p(X \setminus E) = 0$ since the zero function is a solution with any boundary data.

On the other hand, uniqueness always fails if $C_p(X \setminus E) = 0$ in the single obstacle problem (when it is soluble), by the following result. In particular it fails for the Dirichlet problem.

Proposition 5.10. *Let X be bounded and let $E \subset X$ be measurable and such that $C_p(X \setminus E) = 0$. Let $f \in N^{1,p}(E)$ and $\psi \colon E \to \overline{\mathbb{R}}$. Let u be a solution of the $\mathcal{K}_{\psi,f}$-obstacle problem and let $a \in \mathbb{R}$. Then $v := \max\{u, a\}$ is also a solution of the $\mathcal{K}_{\psi,f}$-obstacle problem.*

Proof. As $N_0^{1,p}(E) = N^{1,p}(X)$ we see that $v \in \mathcal{K}_{\psi,f}$. Moreover, $g_v \leq g_u$ a.e. in E, and thus v must also be a solution. □

In fact it follows from this proof that the $\mathcal{K}_{\psi,f}(X)$-obstacle problem for bounded X has a solution only if there is some function $u \in \mathcal{K}_{\psi,f}(X)$ with $g_u = 0$ a.e. If X supports the p-Friedrichs inequality (2.3), then this happens if and only if there is some constant (real-valued) function $u \in \mathcal{K}_{\psi,f}(X)$, which in turn happens (independently of f) if and only if there exists $M < \infty$ such that $\psi \leq M$ q.e.

We end this discussion with a comment on the case when E is unbounded. In this case we may also lose existence, as the following example shows.

Example 5.11. Let $X = \mathbb{R}$ (unweighted), $p > 1$, $E = (0, \infty)$, $f(x) = (1 - x)_+$ and $\psi = 0$. Set $f_j(x) = (1 - x/j)_+$, $j = 1, 2, \dots$ Then $f_j \in \mathcal{K}_{\psi,f}$ and

$$\int_E g_{f_j}^p \, d\mu = \int_0^j \frac{1}{j^p} \, dx = j^{1-p} \to 0, \quad \text{as } j \to \infty.$$

This shows that a solution of the $\mathcal{K}_{\psi,f}$-obstacle problem must have zero energy, and thus must be constant almost everywhere. The boundary condition would require a solution u to satisfy $u = 1$ a.e., but then $u \notin \mathcal{K}_{\psi,f}$.

We conclude this section with an application of our theory to condenser capacities. On metric spaces, such capacities have been used and studied under various assumptions; e.g. by Heinonen–Koskela [22], Kallunki–Shanmugalingam [23] and Adamowicz–Björn–Björn–Shanmugalingam [1].

Definition 5.12. Let $\Omega \subset X$ be a nonempty bounded open set, and let $A_0, A_1 \subset \Omega$ be disjoint. Then the *capacity* of the condenser (A_0, A_1, Ω) is

$$\operatorname{cap}_p(A_0, A_1, \Omega) = \inf_u \int_\Omega g_u^p \, d\mu,$$

where the infimum is taken over all $u \in N^{1,p}(\Omega)$ satisfying $0 \leq u \leq 1$ in Ω, $u = 0$ in A_0 and $u = 1$ in A_1.

Note that $\operatorname{cap}_p(A_0, A_1, \Omega) = \operatorname{cap}_p(A_1, A_0, \Omega)$. Since the equivalence classes in $N^{1,p}(\Omega)$ are defined up to sets of capacity zero, we can equivalently require the equalities in A_0 and A_1 to hold quasi-everywhere. This is thus a double obstacle problem in Ω but without boundary values. We obtain the following consequences of the results in this and the previous section.

Theorem 5.13. *Assume that $p > 1$. Let $\Omega \subset X$ be a nonempty bounded open set, and let $A_0, A_1 \subset \Omega$ be disjoint sets such that $\mathrm{cap}_p(A_0, A_1, \Omega) < \infty$ (which in particular happens if $\mathrm{dist}(A_0, A_1) > 0$).*

Then there is a minimizer for the condenser (A_0, A_1, Ω), i.e., a function $u \in N^{1,p}(\Omega)$ such that $0 \leq u \leq 1$ in Ω, $u = 0$ in A_0, $u = 1$ in A_1 and

$$(5.4) \qquad \mathrm{cap}_p(A_0, A_1, \Omega) = \int_\Omega g_u^p \, d\mu.$$

If X has the zero p-weak upper gradient property, $C_p(A_0 \cup A_1) > 0$, and Ω is connected, then the minimizer is unique (up to sets of capacity zero).

By Lemma 3.4 in A. Björn [3], the zero p-weak upper gradient property for X holds e.g. if X supports a $(1, p)$-Poincaré inequality. For the uniqueness in Theorem 5.13 it is actually enough that Ω have the zero p-weak upper gradient property, as can be seen from the proof below.

Observe that if $C_p(A_0) = C_p(A_1) = 0$, then any constant function with a value in $[0, 1]$ is a minimizer (after redefinition on $A_0 \cup A_1$), which is thus not unique.

Proof. Existence. Let $f = 0$, $\psi_1 = \chi_{A_1}$ and $\psi_2 = \chi_{\Omega \setminus A_0}$. It is then easy to see that every solution of the $\mathcal{K}_{\psi_1, \psi_2, f}(\Omega)$-obstacle problem taken with respect to the ambient space Ω is a minimizer for the condenser (after redefinition on a subset of $A_0 \cup A_1$ of capacity zero). The existence thus follows from Theorem 4.2 and Remark 5.6.

Uniqueness. By symmetry, we may assume that $C_p(A_0) > 0$. Assume that u and u' are two minimizers of the condenser and let

$$Z = \{x \in \Omega : u(x) = u'(x) = 0\},$$

which is a measurable set containing A_0. Let $E = \Omega \setminus Z$, $f = 0$ and $\psi = \chi_{A_1}$. It is again easy to see that both $u|_E$ and $u'|_E$ are solutions of the $\mathcal{K}_{\psi, f}(E)$-obstacle problem taken with Ω as the ambient space. (Recall that for $u \in N_0^{1,p}(E; \Omega)$ we have $g_{u,E} = g_{u,\Omega} = g_u$ a.e., by Proposition 3.10, and hence the energies considered for the condenser (A_0, A_1, Ω) and in the $\mathcal{K}_{\psi, f}(E)$-obstacle problem coincide. Here $N_0^{1,p}(E; \Omega)$ is $N_0^{1,p}(E)$ taken with respect to the ambient space Ω.) Since X has the zero p-weak upper gradient property, so does Ω, as this is a local property. Since $C_p(\Omega \setminus E) = C_p(Z) \geq C_p(A_0) > 0$ and Ω is connected, the uniqueness thus follows from Remark 5.8. $\qquad \square$

Observe that in the existence part of the proof f plays no role as the boundary is empty. This is allowed by Remark 5.6. The uniqueness, however, cannot be deduced using the obstacle problem without boundary values, and hence a different obstacle problem needs to be considered in the second part of the proof.

Next, we prove another application of our results, and in particular of Theorem 5.13. It turns out to be useful in connection with ends and prime ends on metric spaces in the paper Adamowicz–Björn–Björn–Shanmugalingam [1]; cf. Lemma A.11 therein.

Proposition 5.14. *Assume that X is complete and supports a $(1,p)$-Poincaré inequality, that μ is doubling and that $p > 1$. Let Ω be a nonempty bounded connected open set, and let $\{E_k\}_{k=1}^\infty$ be a decreasing sequence of subsets of Ω such that $\bigcap_{k=1}^\infty \overline{E}_k \subset \partial\Omega$.*

Then $\lim_{j\to\infty} \operatorname{cap}_p(E_j, K, \Omega) = 0$ for every compact $K \subset \Omega$ if and only if $\lim_{j\to\infty} \operatorname{cap}_p(E_j, K_0, \Omega) = 0$ for some compact $K_0 \subset \Omega$ with $C_p(K_0) > 0$.

Proof. Assume that $\lim_{j\to\infty} \operatorname{cap}_p(E_j, K_0, \Omega) = 0$ for some compact set $K_0 \subset \Omega$ with positive capacity, and let $K \subset \Omega$ be compact. By Lemma 4.49 in Björn–Björn [6], there is an open connected set $G \Subset \Omega$ such that $K_0 \cup K \subset G$. We can also find k_0 such that $E_{k_0} \cap G = \varnothing$.

For $k \geq k_0$, let u_k be a minimizer for $\operatorname{cap}_p(E_k, K_0, \Omega)$, which exists and is unique (up to sets of capacity zero) by Theorem 5.13. Note that $u_k = 0$ on E_k and $u_k = 1$ on K_0. Moreover, u_k is a superminimizer in $\Omega \setminus \overline{E}_k \supset G$ (see Kinnunen–Martio [26] or [6] for the definitions of superminimizers and superharmonic functions). Indeed, if $0 \leq \varphi \in N^{1,p}(X)$ and $\varphi = 0$ outside $\Omega \setminus \overline{E}_k$, then $v = \min\{u_k + \varphi, 1\}$ is admissible for $\operatorname{cap}_p(E_k, K_0, \Omega)$ and hence

$$\int_{\Omega \setminus \overline{E}_k} g_{u_k}^p \, d\mu \leq \int_{\Omega \setminus \overline{E}_k} g_v^p \, d\mu \leq \int_{\Omega \setminus \overline{E}_k} g_{u_k + \varphi}^p \, d\mu.$$

By Theorem 5.1 in [26] (or Theorem 8.22 in [6]),

$$u_k^*(x) := \lim_{r \to 0} \operatorname*{ess\,inf}_{B(x,r)} u_k$$

equals u_k quasi-everywhere in G, and by Proposition 7.4 in [26] (or Proposition 9.4 in [6]) u_k^* is superharmonic in G. As u_k^* is lower semicontinuous, the minimum $\delta_k := \min_K u_k$ is attained at some point in K. Since $u_k^*(x) = 1$ for some $x \in K_0$ (as $C_p(K_0) > 0$) we see that $u_k^* \not\equiv 0$ in G. Hence, as G is connected, the strong minimum principle in G (Theorem 9.13 in [6]) shows that $\delta_k > 0$.

By Corollary 4.3, we have $u_k \geq u_{k_0}$ q.e., and thus $\delta_k \geq \delta_{k_0} > 0$. It follows that $\min\{u_k/\delta_{k_0}, 1\}$ is admissible for $\operatorname{cap}_p(E_k, K, \Omega)$ as $u_k/\delta_{k_0} \geq 1$ on K. The monotonicity of cap_p then yields that

$$\operatorname{cap}_p(E_k, K, \Omega) \leq \frac{1}{\delta_{k_0}^p} \int_\Omega g_{u_k}^p \, d\mu = \frac{1}{\delta_{k_0}^p} \operatorname{cap}_p(E_k, K_0, \Omega) \to 0, \quad \text{as } k \to \infty.$$

The converse implication is trivial. $\qquad\square$

6. The Adams criterion for $\mathcal{K}_{\psi,f} \neq \varnothing$

As in Section 4, we assume that $p > 1$ and that X supports the p-Friedrichs inequality (2.3). We also assume that $E \subset X$ is a bounded measurable set such that $C_p(X \setminus E) > 0$.

In this section, we study when the single obstacle problem is soluble, i.e., when $\mathcal{K}_{\psi,f}$ is nonempty. In the characterization, we shall use the variational capacity with respect to nonopen sets; see Appendix B.

Theorem 6.1. (The Adams criterion) *Let $f \in D^p(E)$ and $\psi : E \to \overline{\mathbb{R}}$. Then $\mathcal{K}_{\psi,f} \neq \varnothing$ if and only if*

$$(6.1) \qquad \int_0^\infty t^{p-1} \operatorname{cap}_p(\{x : \psi(x) - f(x) > t\}, E)\, dt < \infty.$$

In the linear case on unweighted $\mathbb{R}^n$ and with E open and $f \in N^{1,2}(E)$ (or rather $f \in W^{1,2}(E)$ quasicontinuous) this result was obtained by Adams [2]. For open E in metric spaces and $f \in N^{1,p}(E)$, it is included in Björn–Björn [6].

By the Cavalieri principle, if $f : X \to [0, \infty]$ is a ν-measurable function then

$$\int_X f^p \, d\nu = p \int_0^\infty t^{p-1} \nu(\{x : f(x) > t\})\, dt.$$

By analogy, it is natural to write (6.1) as

$$\int_E (\psi - f)_+^p \, d\operatorname{cap}_p(\,\cdot\,, E) < \infty,$$

even though $\operatorname{cap}_p(\,\cdot\,, E)$ is *not* a measure. Such integrals are called *Choquet integrals* and their study goes back to Choquet [14].

Note that for Theorem 6.1 to hold it is important that the obstacle problem is defined by requiring the obstacle inequality to hold quasi-everywhere (with respect to X). If the inequality is only required to hold almost everywhere, as e.g. in Heinonen–Kilpeläinen–Martio [21] or Kinnunen–Martio [26], only one implication in Theorem 6.1 is true. To see this let $E = B(0,1) \subset \mathbb{R}^n$, $f \equiv 0$ and $\psi = \infty \chi_F$, where $F \subset E$ is a set such that $\mu(F) = 0 < C_p(F)$. By Lemma B.2, $\operatorname{cap}_p(F, E) > 0$, and thus by the Adams criterion, $\mathcal{K}_{\psi,f} = \varnothing$. On the other hand, 0 is a solution of the almost everywhere obstacle problem.

The same is true if we had used E-quasi-everywhere in the definition of the obstacle problem. In this case, we let $E = B(0,1) \setminus \mathbb{Q} \subset \mathbb{R}$, $f = 0$ and $\psi = \infty \chi_F$, where F is a nonempty set with $C_p^E(F) = 0 < C_p(F)$, which is easily accomplished, as in this case $C_p^E(A) = \mu(A)$ for all sets $A \subset E$. Again, $\operatorname{cap}_p(F, E) > 0$, by Lemma B.2, and thus $\mathcal{K}_{\psi,f} = \varnothing$, by the Adams criterion, while 0 is a solution of the E-quasi-everywhere (and also the almost everywhere) obstacle problem.

For the double obstacle problem it is much more difficult to obtain a characterization of when $\mathcal{K}_{\psi_1,\psi_2,f} \neq \varnothing$. The following two examples demonstrate this.

Example 6.2. Let $X = \mathbb{R}$, $p > 1$, $E = (0,1)$ and $f(x) = x$. Let $\psi_1, \psi_2 : \mathbb{R} \to \overline{\mathbb{R}}$ be defined by

$$\psi_1 = \begin{cases} x^{1-1/p}, & 0 < x < 1, \\ -\infty, & \text{otherwise,} \end{cases} \quad \text{and} \quad \psi_2 = \begin{cases} x^{1-1/p}, & 0 < x < 1, \\ \infty, & \text{otherwise.} \end{cases}$$

Then $\mathcal{K}_{\psi_1,\psi_2,f} = \varnothing$, as the function $x \mapsto x^{1-1/p}$ does not belong to $N^{1,p}(E)$.

In the above example, we have $\psi_1 = \psi_2$ on a large set. Next we shall see that it is possible to have $\mathcal{K}_{\psi_1,\psi_2,f} = \varnothing$ even if $\psi_2 - \psi_1 \equiv \infty$.

Example 6.3. Let $X = \mathbb{R}$, $p > 1$, $\Omega \subset \mathbb{R}$ be open,

$$\psi_1 = -\infty\chi_{\mathbb{Q}} \quad \text{and} \quad \psi_2 = \infty(1 - \chi_{\mathbb{Q}}).$$

Note that $\psi_2 - \psi_1 \equiv \infty$. Let $u \in N^{1,p}(\Omega)$ be such that $\psi_1 \leq u \leq \psi_2$ q.e. Since every function in $N^{1,p}(\Omega)$ is (absolutely) continuous, this implies that $u \geq 0$ a.e. (and hence everywhere) in Ω. On the other hand, as $\mathbb{Q}$ is dense in Ω, the continuity of u and the fact that $u \leq 0$ on $\mathbb{Q} \cap \Omega$ yield that $u \leq 0$ in Ω.

Hence $u = 0$ in Ω and $\mathcal{K}_{\psi_1,\psi_2,f} = \varnothing$ whenever $f \notin N_0^{1,p}(\Omega)$. Moreover, similar arguments show that if $\psi_1' = \psi_1 + 1$, then $\mathcal{K}_{\psi_1',\psi_2,f} = \varnothing$ for all $f \in D^p(\Omega)$.

To prove Theorem 6.1 we will use the following lemma.

Lemma 6.4. *Let* $a > 1$, $u \in N_0^{1,p}(E)$ *and* $E_t = \{x \in E : |u(x)| > t\}$, $t > 0$. *Then*

$$(6.2) \qquad \int_0^\infty t^{p-1} \operatorname{cap}_p(E_{at}, E_t)\, dt \leq \frac{\log a}{(a-1)^p} \int_E g_u^p\, d\mu.$$

Equivalently, with $b = 1/a \in (0,1)$,

$$(6.3) \qquad \int_0^\infty t^{p-1} \operatorname{cap}_p(E_t, E_{bt})\, dt \leq \frac{-\log b}{(1-b)^p} \int_E g_u^p\, d\mu.$$

Proof. As $g_u = g_{|u|}$ a.e., we may assume that $u \geq 0$. For $t > 0$, let

$$u_t = \min\{(u-t)_+, (a-1)t\}$$

be the truncations of u at levels t and at, $t > 0$. Then the function $v_t := u_t/(a-1)t$ is admissible in the definition of $\operatorname{cap}_p(E_{at}, E_t)$ and $g_{v_t} = g_u\chi_{\{t<u<at\}}/(a-1)t$ a.e. Using the Fubini theorem we get that

$$\int_0^\infty t^{p-1} \operatorname{cap}_p(E_{at}, E_t)\, dt \leq \int_0^\infty \left(\frac{1}{(a-1)t}\right)^p t^{p-1} \int_X g_u^p \chi_{\{t<u<at\}}\, d\mu\, dt$$

$$= \frac{1}{(a-1)^p} \int_X g_u(x)^p \int_{u(x)/a}^{u(x)} \frac{dt}{t}\, d\mu(x)$$

$$= \frac{\log a}{(a-1)^p} \int_X g_u^p\, d\mu,$$

which proves (6.2). To get the last equality we used the fact that $g_u = 0$ a.e. in $\{x : u(x) = 0\}$.

The inequality (6.3) follows from (6.2) by the substitution $s = bt$. $\qquad\square$

It follows directly from the definition that $\operatorname{cap}_p(E_t, E) \leq \operatorname{cap}_p(E_t, E_{bt})$ and hence the capacity on the left-hand side of (6.3) can be replaced by $\operatorname{cap}_p(E_t, E)$. Letting $b = 1/p$ yields the following result.

Corollary 6.5. (The Maz'ya capacitary inequality) *If* $u \in N_0^{1,p}(E)$, *then*

$$(6.4) \qquad \int_0^\infty t^{p-1} \operatorname{cap}_p(\{x : |u(x)| > t\}, E)\, dt \leq \frac{p^p \log p}{(p-1)^p} \int_E g_u^p\, d\mu.$$

Using the notation introduced above, (6.4) can be written as

$$\int_E |u|^p \, d \operatorname{cap}_p(\,\cdot\,, E) \leq \frac{p^{p+1} \log p}{(p-1)^p} \int_E g_u^p \, d\mu.$$

By minimizing the constant on the right-hand side in (6.3) for $b \in (0,1)$ one can optimize the result. An easy calculation shows that the minimum is attained when $1/b - 1 = -p \log b$.

In Section 2.3.1 in Maz′ya [32], the inequality (6.4) was proved with the optimal constant $p^{p-1}/(p-1)^{p-1}$ (for unweighted $\mathbb{R}^n$). See also Maz′ya [33].

Proof of Theorem 6.1. As $\mathcal{K}_{\psi,f} = f + \mathcal{K}_{\psi-f,0}$ we can assume, without loss of generality, that $f \equiv 0$.

Assume that there is some $\tilde{u} \in \mathcal{K}_{\psi,0}$. Then $u := \max\{\tilde{u}, \psi\} = \tilde{u}$ q.e. in E, and thus also $u \in \mathcal{K}_{\psi,0}$. Hence, by Corollary 6.5 we have

$$\int_0^\infty t^{p-1} \operatorname{cap}_p(\{x : \psi(x) > t\}, E) \, dt \leq \int_0^\infty t^{p-1} \operatorname{cap}_p(\{x : u(x) > t\}, E) \, dt$$
$$\leq C \int_E g_u^p \, d\mu < \infty.$$

Conversely, assume that (6.1) holds. As $\operatorname{cap}_p(\{x : \psi(x) > t\}, E)$ is nonincreasing with respect to t, it follows that $\operatorname{cap}_p(\{x : \psi(x) > t\}, E) < \infty$ for all $t > 0$. Thus we can find $u_k \in N_0^{1,p}(E)$, for $k \in \mathbb{Z}$, such that $\chi_{\{\psi > 2^k\}} \leq u_k \leq 1$ and

$$(6.5) \qquad \int_E g_{u_k}^p \, d\mu < \operatorname{cap}_p(\{x : \psi(x) > 2^k\}, E) + 2^{-|k|-(k+1)p}.$$

Let

$$v_N = \sup_{k \leq N} 2^{k+1} u_k, \qquad\qquad g_N = \sup_{k \leq N} 2^{k+1} g_{u_k}, \qquad\qquad N \in \mathbb{Z},$$
$$v = \sup_{k \in \mathbb{Z}} 2^{k+1} u_k = \sup_{N \in \mathbb{Z}} v_N, \qquad g = \sup_{k \in \mathbb{Z}} 2^{k+1} g_{u_k} = \sup_{N \in \mathbb{Z}} g_N.$$

(Here we take the same representative of g_{u_k} in all expressions.) Then $v \geq 2^{k+1}$ when $\psi > 2^k$, in particular when $2^k < \psi \leq 2^{k+1}$, $k \in \mathbb{Z}$, from which it follows that $v \geq \psi$ in E.

By Lemma 1.52 in Björn–Björn [6], g_N is a p-weak upper gradient of v_N. Moreover

$$\int_E g^p \, d\mu = \int_E \left(\sup_{k \in \mathbb{Z}} 2^{k+1} g_{u_k} \right)^p d\mu$$
$$\leq \int_E \sum_{k=-\infty}^\infty (2^{k+1} g_{u_k})^p \, d\mu = \sum_{k=-\infty}^\infty 2^{(k+1)p} \int_E g_{u_k}^p \, d\mu.$$

Using (6.5) we obtain

$$\int_E g^p \, d\mu < \sum_{k=-\infty}^{\infty} 2^{(k+1)p}(\mathrm{cap}_p(\{x : \psi(x) > 2^k\}, E) + 2^{-|k|-(k+1)p})$$

$$\leq 3 + \sum_{k=-\infty}^{\infty} 2^{(k+1)p} \int_{2^{k-1}}^{2^k} \left(\frac{t}{2^{(k-1)}}\right)^{p-1} \mathrm{cap}_p(\{x : \psi(x) > t\}, E) \, dt$$

$$= 3 + 4^p \int_0^{\infty} t^{p-1} \mathrm{cap}_p(\{x : \psi(x) > t\}, E) \, dt.$$

The assumption (6.1) now yields that $\int_E g^p \, d\mu < \infty$. Since $g_N \nearrow g$ pointwise in X, dominated convergence implies that $g_N \to g$ in $L^p(X)$. Monotone convergence and (2.3) then yield

$$(6.6) \qquad \int_E |v|^p \, d\mu = \lim_{N \to \infty} \int_E |v_N|^p \, d\mu \leq C_E \int_E g_{v_N}^p \, d\mu \leq C_E \int_E g^p \, d\mu < \infty.$$

Thus $v_N \to v$ both pointwise and in $L^p(X)$, by dominated convergence. Proposition A.1 shows that $v \in N^{1,p}(X)$. As $v = 0$ in $X \setminus E$, we get $v \in N_0^{1,p}(E)$ and therefore $v \in \mathcal{K}_{\psi,0}$.   □

If the obstacle $\psi \in N^{1,p}(E)$, then there is a much easier criterion for when $\mathcal{K}_{\psi,f} \neq \varnothing$.

Proposition 6.6. *Let* $f, \psi \in N^{1,p}(E)$ *(or, more generally, let* $f, \psi \in D^p(E)$ *be such that* $f - \psi \in N^{1,p}(E)$*). Then* $\mathcal{K}_{\psi,f} \neq \varnothing$ *if and only if* $(\psi - f)_+ \in N_0^{1,p}(E)$*.*

Proof. Assume that there is $u \in \mathcal{K}_{\psi,f}$. Then

$$0 \leq (\psi - f)_+ \leq (u - f)_+ \quad \text{q.e.}$$

Hence, $(\psi - f)_+ \in N_0^{1,p}(E)$, by Lemma 2.4.

Conversely, assume that $(\psi - f)_+ \in N_0^{1,p}(E)$ and let $u = \max\{\psi, f\}$. Then $u - f = (\psi - f)_+ \in N_0^{1,p}(E)$. As $u \geq \psi$ in E, it follows that $u \in \mathcal{K}_{\psi,f}$.   □

Remark 6.7. In this section, we only used the Friedrichs inequality and the assumption $C_p(X \setminus E) > 0$ in the proof of Theorem 6.1 (apart from some examples). More precisely, these assumptions were used in (6.6), where it is enough if (2.3) holds for the specific E under consideration. Neither of these two assumptions can be omitted in Theorem 6.1, which is seen by letting $\psi \equiv \infty$ and $f \equiv 0$ and considering either $E = X_+$ in Example 5.5 or an arbitrary E such that $C_p(X \setminus E) = 0$ (and $\mu(X) > 0$). Note that in both cases $\mathrm{cap}_p(E, E) = 0$ so that the integral in (6.1) converges even though $\mathcal{K}_{\psi,f} = \varnothing$.

All other results in this section hold without the Friedrichs or Poincaré inequalities.

7. Nontriviality of the obstacle problem and of $N_0^{1,p}$

In this section we assume that X is complete and supports a $(1,p)$-Poincaré inequality, that μ is doubling and that $p > 1$.

These assumptions are needed to be able to use some results from fine potential theory.

In the obstacle problem it is natural to ask when the obstacle problem is trivial, i.e., when all functions $v \in \mathcal{K}_{\psi_1,\psi_2,f}$ agree quasi-everywhere. This happens in particular when $N_0^{1,p}(E)$ is trivial. In the double obstacle problem it can happen also in other cases, e.g. if $\psi_1 \equiv \psi_2$ or in Examples 6.2 and 6.3. For the single obstacle problem the situation is simpler and we have the following characterization.

Proposition 7.1. *Let $E \subset X$ be a bounded measurable set with $C_p(X \setminus E) > 0$, $f \in D^p(E)$ and $\psi \colon E \to \overline{\mathbb{R}}$. Then $\mathcal{K}_{\psi,f}$ is trivial (in the sense that $u = v$ q.e. whenever $u,v \in \mathcal{K}_{\psi,f}$) if and only if either $\mathcal{K}_{\psi,f} = \varnothing$ or $N_0^{1,p}(E)$ is trivial (i.e., $u = 0$ q.e. for all $u \in N_0^{1,p}(E)$).*

Observe that the Adams criterion (Theorem 6.1) shows when $\mathcal{K}_{\psi,f} = \varnothing$. Note also that if $\mathcal{K}_{\psi,f}$ is nonempty but trivial, then $\mathcal{K}_{\psi,f} = \{u : u = f \text{ q.e.}\}$.

Proof. If $\mathcal{K}_{\psi,f} = \varnothing$, then the equivalence is clear. Assume therefore that $\mathcal{K}_{\psi,f} \neq \varnothing$. If $N_0^{1,p}(E)$ is trivial, then all $v \in \mathcal{K}_{\psi,f}$ agree with f quasi-everywhere, and thus $\mathcal{K}_{\psi,f}$ is trivial.

Conversely assume that $N_0^{1,p}(E)$ is nontrivial. Then there is $u \in N_0^{1,p}(E)$ such that $C_p(\{x : u(x) \neq 0\}) > 0$. Let $v \in \mathcal{K}_{\psi,f}$ and $w = v + |u|$. Then $w \in \mathcal{K}_{\psi,f}$ and as w and v do not agree quasi-everywhere, the nontriviality of $\mathcal{K}_{\psi,f}$ follows. $\square$

Our aim is now to characterize when $N_0^{1,p}(E)$ is trivial. We get the following result. (Definitions of the involved concepts are given below.)

Theorem 7.2. *Let $E \subset X$ be arbitrary. Then the following are equivalent*:

(a) $N_0^{1,p}(E)$ *is nontrivial*;

(b) E *contains a nonempty finely open set, or in other terms* fine-int $E \neq \varnothing$;

(c) *there is a point $x \in E$ such that $X \setminus E$ is thin at x*;

(d) *there are a point $x \in E$ and a number $s > 0$ such that*

$$\mathrm{cap}_p(B(x,s) \setminus E, B(x,2s)) < \mathrm{cap}_p(B(x,s), B(x,2s)).$$

Note that if $\mu(E) = 0$ then all the statements are false, since in this case $f \in N_0^{1,p}(E)$ implies that $f = 0$ a.e. in X, and hence $f = 0$ q.e. in X, i.e., $N_0^{1,p}(E)$ is trivial.

The following result gives a more precise description of $N_0^{1,p}(E)$ and will be used to establish Theorem 7.2.

Theorem 7.3. *Let $E \subset X$ be arbitrary. Then*

$$N_0^{1,p}(E) = N_0^{1,p}(\text{fine-int } E).$$

Here we follow our convention that functions in $N_0^{1,p}$ can be extended by zero quasi-everywhere. Observe that we do not require E to be measurable in Theorems 7.2 and 7.3. See also Section 9 for some further consequences of Theorem 7.3 in the special case $X = \mathbb{R}^n$.

Corollary 7.4. *Let $E, E_0 \subset X$ be measurable sets such that*

$$\text{fine-int } E \subset E_0 \subset E.$$

If $f \in D^p(E)$ and $\mathcal{K}_{\psi_1, \psi_2, f}(E) \neq \varnothing$, then

$$\mathcal{K}_{\psi_1, \psi_2, f}(E) = \mathcal{K}_{\psi_1, \psi_2, f}(E_0).$$

Of course, the main interest is when $E_0 = \text{fine-int } E$, but here, contrary to Theorem 7.3, we also need measurability and we do not know in general if fine-int E is always measurable; cf. Section 9.

Remark 7.5. Note that it is possible to have $\mathcal{K}_{\psi_1, \psi_2, f}(E) = \varnothing \neq \mathcal{K}_{\psi_1, \psi_2, f}(E_0)$. Indeed, this happens exactly if $\mathcal{K}_{\psi_1, \psi_2, f}(E_0) \neq \varnothing$ and

$$(7.1) \qquad C_p(\{x \in E \setminus E_0 : \psi_1(x) > f(x) \text{ or } \psi_2(x) < f(x)\} > 0.$$

To see this, note that since $N_0^{1,p}(E) = N_0^{1,p}(E_0)$ it follows that any function in $N_0^{1,p}(E)$ is 0 quasi-everywhere in $E \setminus E_0$. Hence, if $u \in \mathcal{K}_{\psi_1, \psi_2, f}(E)$, then $u = f$ q.e. in $E \setminus E_0$ which is impossible if $\psi_1 \leq u \leq \psi_2$ q.e. at the same time as (7.1) holds. Conversely, if $u \in \mathcal{K}_{\psi_1, \psi_2, f}(E_0)$ and (7.1) fails, then we extend u as f in $E \setminus E_0$, so that $u \in \mathcal{K}_{\psi_1, \psi_2, f}(E)$, showing that $\mathcal{K}_{\psi_1, \psi_2, f}(E)$ is nonempty.

To make the above results precise we need a few more definitions. See Appendix B for the definition and some properties of the variational capacity cap_p.

Definition 7.6. A set A is *thin* at x if

$$(7.2) \qquad \int_0^1 \left(\frac{\text{cap}_p(A \cap B(x,r), B(x,2r))}{\text{cap}_p(B(x,r), B(x,2r))} \right)^{1/(p-1)} \frac{dr}{r} < \infty.$$

A set A is *finely open* if $X \setminus A$ is thin at all $x \in A$. Using the monotonicity and subadditivity of the capacity, it is easy to verify that the finely open sets form a topology on X; see Proposition 11.36 in Björn–Björn [6]. The *fine interior* fine-int E of E is the largest finely open set contained in E.

Since our variational capacity is the same as the one in Heinonen–Kilpeläinen–Martio [21] (see Björn–Björn [7] for a proof of this fact), we see that this definition coincides with the definition in [21], p. 221, when X is weighted $\mathbb{R}^n$ with a p-admissible weight. If $X = \mathbb{R}^n$ (unweighted) then it is also equivalent to Definition 2.47 in Malý–Ziemer [31].

In the definition of thinness we make the convention that the integrand is 1 whenever $\text{cap}_p(B(x,r), B(x,2r)) = 0$. This happens for example if $X = B(x, 2r)$ is bounded, but never if $r < \frac{1}{2} \text{diam } X$. Note that thinness is a local property, i.e., if $\delta > 0$, then E is thin at x if and only if $E \cap B(x, \delta)$ is thin at x.

To prove Theorem 7.3, we shall use the following result which was obtained by J. Björn (see Theorem 4.6 in [12]), and independently by Korte (Corollary 4.4 in [27]) (the result can also be found in Björn–Björn [6], Theorem 11.40).

Theorem 7.7. *Every $u \in N^{1,p}(X)$ is finely continuous at quasi-every $x \in X$.*

A function u, defined on a finely open set U, is *finely continuous* at $x \in U$ if for every $\varepsilon > 0$ there exists a finely open set $V \ni x$ such that $|u(y) - u(x)| < \varepsilon$ for all $y \in V$ (in particular $u(x) \in \mathbb{R}$).

Proof of Theorem 7.3. Let $u \in N_0^{1,p}(E)$ and extend u by 0 on $X \setminus E$, so that $u \in N^{1,p}(X)$. Let $G = \{x \in E : u(x) \neq 0\}$. By Theorem 7.7, there exists a set F with $C_p(F) = 0$ such that u is finely continuous at every $x \in X \setminus F$. Hence, for every $x \in G \setminus F$, there exists a finely open neighbourhood V_x of x such that $|u - u(x)| < |u(x)|$ in V_x. Note that $u \neq 0$ in V_x and hence $V_x \subset G \subset E$.

Letting $V = \bigcup_{x \in G \setminus F} V_x$, we obtain a finely open set V such that $G \setminus F \subset V \subset E$. As $X \setminus V \subset (X \setminus G) \cup F$, we see that $u = 0$ q.e. in $X \setminus V$, and hence $u \in N_0^{1,p}(V) \subset N_0^{1,p}(\text{fine-int } E)$. Since $u \in N_0^{1,p}(E)$ was arbitrary, this shows that $N_0^{1,p}(E) \subset N_0^{1,p}(\text{fine-int } E)$.

The converse inclusion is obvious.   □

Proof of Theorem 7.2. ¬ (b) ⇒ ¬ (a) By Theorem 7.3,

$$N_0^{1,p}(E) = N_0^{1,p}(\text{fine-int } E) = N_0^{1,p}(\varnothing),$$

and thus $N_0^{1,p}(E)$ is trivial.

(b) ⇒ (c) Let $G \subset E$ be a nonempty finely open set. Then $X \setminus E \subset X \setminus G$ is thin at every $x \in G$.

(c) ⇒ (d) For simplicity, let $B_r = B(x, r)$. Since

$$\int_0^1 \left(\frac{\text{cap}_p(B_r \setminus E, B_{2r})}{\text{cap}_p(B_r, B_{2r})} \right)^{1/(p-1)} \frac{dr}{r} < \infty,$$

we see that

$$\liminf_{r \to 0+} \frac{\text{cap}_p(B_r \setminus E, B_{2r})}{\text{cap}_p(B_r, B_{2r})} = 0$$

and (d) follows. (Actually the limit exists and equals 0, but we will not need that here.)

(d) ⇒ (a) For simplicity, let $B_r = B(x, r)$. Theorem B.3 (iii) implies that

$$\text{cap}_p(B_s, B_{2s}) = \sup_{t < s} \text{cap}_p(B_t, B_{2s}).$$

Hence, there exists $t < s$ such that

$$\text{cap}_p(B_s \setminus E, B_{2s}) < \text{cap}_p(B_t, B_{2s}).$$

Thus there exists a function $h \in N^{1,p}(X)$ such that $0 \le h \le 1$, $h = 1$ on $B_s \setminus E$, $h = 0$ on $X \setminus B_{2s}$ and $\|g_h\|_{L^p(X)}^p < \mathrm{cap}_p(B_t, B_{2s})$. Let $F = \{x \in B_t : h(x) < 1\}$. If $C_p(F)$ were 0, then we would have

$$\mathrm{cap}_p(B_t, B_{2s}) \le \|g_{h+\chi_F}\|_{L^p(X)}^p = \|g_h\|_{L^p(X)}^p < \mathrm{cap}_p(B_t, B_{2s}),$$

a contradiction. Thus $C_p(F) > 0$.

Now let f be a Lipschitz function such that $0 \le f \le 1$, $f = 1$ on B_t and $f = 0$ on $X \setminus B_s$. Set $k = (f - h)_+ \in N^{1,p}(X)$. It follows directly that $k = 0$ on $(B_s \setminus E) \cup (X \setminus B_s) \supset X \setminus E$, and thus $k \in N_0^{1,p}(E)$. Since $F = \{x \in B_t : k(x) > 0\}$ and $C_p(F) > 0$, we see that $k \not\sim 0$, i.e., $N_0^{1,p}(E)$ is nontrivial. $\qquad\square$

We conclude this section by proving the characterization of fine interior stated in the introduction.

Proof of Theorem 1.2. Let $E_0 = $ fine-int $E \subset E$. If $x \in E_0$ then by definition, $X \setminus E_0$ (and hence also $X \setminus E$) is thin at x.

Conversely, assume that $X \setminus E$ is thin at $x \in E$, i.e.,

$$\int_0^1 \Big(\frac{\mathrm{cap}_p(B_r \setminus E, B_{2r})}{\mathrm{cap}_p(B_r, B_{2r})} \Big)^{1/(p-1)} \frac{dr}{r} < \infty,$$

where we abbreviate $B_r = B(x, r)$. For $0 < r < 1$, let F_r be the fine closure of $B_r \setminus E$, i.e., the smallest finely closed set containing $B_r \setminus E$. Then $B_r \setminus F_r$ is finely open and contained in E. To conclude the proof, it suffices to show that F_r is thin at x, as then $(B_r \setminus F_r) \cup \{x\}$ is also finely open and contained in E, which implies that

$$(B_r \setminus F_r) \cup \{x\} \subset E_0,$$

and in particular $x \in E_0$.

We shall show that $B_r \cap F_r$ is thin at x. Since $X \setminus E$ is thin at x, it suffices to show that

$$\mathrm{cap}_p(B_\rho \cap F_r, B_{2\rho}) \le \mathrm{cap}_p(B_\rho \setminus E, B_{2\rho}) \quad \text{for } 0 < \rho \le r.$$

First, we note that $B_\rho \cap F_r \subset F_\rho$. Indeed, $F_\rho \cup (X \setminus B_\rho)$ is finely closed and contains $X \setminus E$ (and hence also F_r). It follows that

$$B_\rho \cap F_r \subset B_\rho \cap (F_\rho \cup (X \setminus B_\rho)) \subset F_\rho.$$

This and Corollary 4.5 in J. Björn [12] (or Corollary 11.39 in Björn–Björn [6]) now yield that

$$\mathrm{cap}_p(B_\rho \cap F_r, B_{2\rho}) \le \mathrm{cap}_p(F_\rho, B_{2\rho}) = \mathrm{cap}_p(B_\rho \setminus E, B_{2\rho}).$$

From this and the thinness of $X \setminus E$ at x we conclude that F_r is thin at x, which finishes the proof. $\qquad\square$

The following direct consequence of Theorem 1.2 characterizes fine closures and fine boundaries; cf. the first half of Theorem 2.136 in Malý–Ziemer [31].

Corollary 7.8. *Let $E \subset X$ be arbitrary. Then the fine closure of E is the set*

$$E \cup \{x \in X \setminus E : E \text{ is not thin at } x\}$$

and the fine boundary of E is

$$\{x \in E : X \setminus E \text{ is not thin at } x\} \cup \{x \in X \setminus E : E \text{ is not thin at } x\}.$$

In particular, the fine boundary of E is a subset of ∂E.

8. Comparing obstacle problems

In this section we assume that $E \subset X$ is a bounded measurable set such that $C_p(X \setminus E) > 0$.

If the boundary data belong to $D^p(\Omega)$ for some open $\Omega \supset E$, then we have two possible definitions of obstacle problems on E, namely Definition 4.1 and (4.2). We begin by comparing the admissible sets in these two definitions.

Lemma 8.1. *If $f \in D^p(\Omega)$ for some open set $\Omega \supset E$, then $\mathcal{K}'_{\psi_1,\psi_2,f} = \mathcal{K}_{\psi_1,\psi_2,f}$.*

Recall that $\mathcal{K}'_{\psi_1,\psi_2,f}$ was defined in (4.2). By saying that $\mathcal{K}'_{\psi_1,\psi_2,f} = \mathcal{K}_{\psi_1,\psi_2,f}$ we really mean that

$$\{f|_E : f \in \mathcal{K}'_{\psi_1,\psi_2,f}\} = \mathcal{K}_{\psi_1,\psi_2,f},$$

and that every $f \in \mathcal{K}_{\psi_1,\psi_2,f}$ corresponds to a unique (up to sets of capacity zero) $\tilde{f} \in \mathcal{K}'_{\psi_1,\psi_2,f}$. Note that already in Section 4 we observed that $\mathcal{K}'_{\psi_1,\psi_2,f} = \mathcal{K}_{\psi'_1,\psi'_2,f}(\Omega)$, where ψ'_1 and ψ'_2 are given by (4.3).

Proof. Clearly, $\mathcal{K}'_{\psi_1,\psi_2,f} \subset \mathcal{K}_{\psi_1,\psi_2,f}$. To prove the other inclusion, let $v \in \mathcal{K}_{\psi_1,\psi_2,f}$, i.e., $v \in D^p(E)$ and $v - f = w \in N_0^{1,p}(E)$. Then w (extended by zero outside of E) belongs to $N^{1,p}(\Omega)$ and hence $v = f + w \in D^p(\Omega)$, from which the result follows.   □

Note that even though $\mathcal{K}'_{\psi_1,\psi_2,f} = \mathcal{K}_{\psi_1,\psi_2,f}$ for $f \in D^p(\Omega)$, the minimal p-weak upper gradients considered in these two obstacle problems are different. The minimal p-weak upper gradient in the $\mathcal{K}_{\psi_1,\psi_2,f}$-obstacle problem is taken with respect to E and is in general smaller than the minimal p-weak upper gradient with respect to Ω or X, considered in the $\mathcal{K}'_{\psi_1,\psi_2,f}$-obstacle problem.

Example 8.2. As in Example 3.6, let $X = \mathbb{R}$ and $E = (0,1) \setminus \mathbb{Q}$, and recall that the minimal p-weak upper gradient (and thus the p-energy integral) taken with respect to E is zero for every function on E, while the minimal p-weak upper gradient with respect to $\mathbb{R}$ is just the modulus of the distributional derivative.

However, since $(0,1) \setminus E$ is dense in $(0,1)$ and all functions in $N^{1,p}(X)$ are absolutely continuous, the space $N_0^{1,p}(E)$ is trivial and so is $\mathcal{K}_{\psi_1,\psi_2,f}$; cf. Proposition 7.1. Hence, the only solution (if it exists) of both the $\mathcal{K}_{\psi_1,\psi_2,f}$- and the $\mathcal{K}'_{\psi_1,\psi_2,f}$-obstacle problem is f itself.

The last observation in Example 8.2 holds in much more generality, as we shall now see. Recall that p-path almost open sets were introduced in Definition 3.1.

Theorem 8.3. *Assume that X is complete and supports a $(1,p)$-Poincaré inequality, that μ is doubling and that $p > 1$. Let E_0 be a p-path almost open measurable set such that* fine-int $E \subset E_0 \subset E$, *and let $f \in D^p(E)$ and $\psi_j : E \to \overline{\mathbb{R}}$, $j = 1, 2$, be such that $\mathcal{K}_{\psi_1,\psi_2,f}(E) \neq \varnothing$. Then the solutions of the $\mathcal{K}_{\psi_1,\psi_2,f}(E)$-problem coincide with the solutions of the $\mathcal{K}_{\psi_1,\psi_2,f}(E_0)$-problem.*

Moreover, if $\mu(E \setminus E_0) = 0$ then also the p-energies associated with these two problems coincide. In particular, this holds if $\mu(\partial E) = 0$.

If $f \in D^p(\Omega)$ for some open set $\Omega \supset E$, then the above solutions coincide with the solutions of the $\mathcal{K}'_{\psi_1,\psi_2,f}(E)$-problem.

Of course, the main interest is when $E_0 =$ fine-int E, but as we do not know whether fine-int E is always measurable and p-path almost open, we have given the formulation above. See, however, Section 9 and Theorem 1.3 for an improvement in the case $X = \mathbb{R}^n$.

Note that even if the solutions coincide, the corresponding p-energies can in general be different for these obstacle problems. Indeed, even though

$$g_{u,E_0} = g_{u,E} \quad \text{a.e. in } E_0$$

for every $u \in \mathcal{K}_{\psi_1,\psi_2,f}(E)$, by Corollary 3.7, we only get

$$\int_{E_0} g_{u,E_0}^p \, d\mu = \int_{E_0} g_{u,E}^p \, d\mu \leq \int_E g_{u,E}^p \, d\mu$$

with strict inequality unless $g_{u,E} = 0$ a.e. in $E \setminus E_0$ (which holds in particular if $\mu(E \setminus E_0) = 0$).

If $f \in D^p(\Omega)$ for some open $\Omega \supset E$, then

$$g_{u,E_0} = g_{u,E} = g_u \quad \text{a.e. in } E_0$$

for every $u \in \mathcal{K}_{\psi_1,\psi_2,f}(E)$, by Corollary 3.7, but we have only

$$g_{u,E} = g_{f,E} \leq g_f = g_u \quad \text{a.e. in } E \setminus E_0$$

for those u, and the inequality in the middle can be strict; see Example 8.2 where E_0 is empty. Thus, the two p-energies $\int_E g_{u,E}^p \, d\mu$ and $\int_E g_u^p \, d\mu$ will coincide only if $g_{f,E} = g_f$ a.e. in $E \setminus E_0$, in particular if $\mu(E \setminus E_0) = 0$.

Proof of Theorem 8.3. To simplify the notation, we omit the subscripts ψ_1, ψ_2 and f and write only $\mathcal{K}(E)$, $\mathcal{K}(E_0)$ and $\mathcal{K}'(E)$ in this proof.

By Corollary 7.4, we have $\mathcal{K}(E) = \mathcal{K}(E_0)$. Since E_0 is p-path almost open, Corollary 3.7 (with X replaced by E) yields that for all $v \in \mathcal{K}(E) = \mathcal{K}(E_0)$,

$$(8.1) \qquad\qquad g_{v,E} = g_{v,E_0} \quad \text{a.e. in } E_0.$$

Moreover, as $v - f \in N_0^{1,p}(E_0)$, we have $v = f$ q.e. in $E \setminus E_0$ and hence

$$g_{v,E} = g_{f,E} \quad \text{a.e. in } E \setminus E_0.$$

Similarly, if $f \in D^p(\Omega)$ for some open $\Omega \supset E$, then $\mathcal{K}'(E) = \mathcal{K}(E)$, by Lemma 8.1, and for all $v \in \mathcal{K}(E)$,

$$(8.2) \qquad g_{v,E} = g_v \text{ a.e. in } E_0 \quad \text{and} \quad g_v = g_f \text{ a.e. in } \Omega \setminus E_0,$$

again by Corollary 3.7 (with X replaced by Ω).

Let u be a solution of the $\mathcal{K}(E_0)$-problem. Then (8.1) implies that for all $v \in \mathcal{K}(E_0) = \mathcal{K}(E)$,

$$(8.3) \qquad \int_{E_0} g_{u,E}^p \, d\mu = \int_{E_0} g_{u,E_0}^p \, d\mu \leq \int_{E_0} g_{v,E_0}^p \, d\mu = \int_{E_0} g_{v,E}^p \, d\mu.$$

Similarly, if $f \in D^p(\Omega)$ for some open $\Omega \supset E$, and u' is a solution of the $\mathcal{K}'(E)$-problem, then (8.2) implies that for all $v \in \mathcal{K}'(E) = \mathcal{K}(E)$,

$$
\begin{aligned}
(8.4) \qquad \int_{E_0} g_{u',E}^p \, d\mu &= \int_{E_0} g_{u'}^p \, d\mu = \int_E g_{u'}^p \, d\mu - \int_{E \setminus E_0} g_f^p \, d\mu \\
&\leq \int_E g_v^p \, d\mu - \int_{E \setminus E_0} g_f^p \, d\mu = \int_{E_0} g_v^p \, d\mu = \int_{E_0} g_{v,E}^p \, d\mu.
\end{aligned}
$$

Adding $\int_{E \setminus E_0} g_{f,E}^p \, d\mu$ to both sides in (8.3) and (8.4) shows that both u and u' are solutions of the $\mathcal{K}(E)$-problem (the latter assuming that $f \in D^p(\Omega)$). By uniqueness, they coincide quasi-everywhere in E and are the only (up to sets of capacity zero) solutions of the $\mathcal{K}(E)$-obstacle problem. $\square$

9. The Euclidean case

The situation simplifies in $\mathbb{R}^n$ (unweighted). In this case Theorem 2.144 in Malý–Ziemer [31] (which goes back to Fuglede [17]) shows that every finely open set G is *quasi-open*, i.e., for every $\varepsilon > 0$ there exists an open set V with $C_p(V) < \varepsilon$ such that $G \cup V$ is open. In particular, the fine interior fine-int E of every set $E \subset \mathbb{R}^n$ is quasi-open.

Remark 9.1. If $p > n$, then quasi-open sets are open in $\mathbb{R}^n$, and thus the quasi-open, finely open and open sets coincide. There are immediate consequences of this for the results in Section 8 which we leave to the reader to formulate explicitly.

If $1 < p \leq n$, then for every $x > 0$ and $\varepsilon > 0$ there is an open set $V \ni x$ with $C_p(V) < \varepsilon$, and thus $\{x\}$ is quasi-open. Since not all sets are quasi-open, by Lemma 9.2 and Remark 3.3, it follows that the quasi-open sets do not form a topology.

To be able to state Theorem 8.3 without additional assumptions on E_0, we recall the following results which hold in general metric spaces.

Lemma 9.2 (Shanmugalingam, Remark 3.5 in [37]). *Every quasi-open set is p-path open.*

Lemma 9.3. *Every quasi-open set G is measurable.*

Proof. For every $j = 1, 2, \ldots$, there is an open set V_j such that $C_p(V_j) < 1/j$ and $G_j := G \cup V_j$ is open. Let $A_j = G_j \setminus V_j \subset G$, $A = \bigcup_{j=1}^{\infty} A_j \subset G$ and $E = \bigcap_{j=1}^{\infty} G_j \supset G$, all of which are Borel sets. Then $A \subset G \subset E$ and

$$\mu(E \setminus A) \leq \mu(G_j \setminus A_j) = \mu(V_j) \leq C_p(V_j) < 1/j \quad \text{for } j = 1, 2, \ldots$$

Letting $j \to \infty$ shows that G is measurable. $\square$

Hence, if $E \subset \mathbb{R}^n$ then fine-int E is measurable and p-path open, and Theorem 8.3 turns into Theorem 1.3 in the introduction. We also have the following consequence of Lemma 3.9 and Theorem 7.3, which generalizes Theorem 2.147 in Malý–Ziemer [31]. See also Remark 2.148 in [31] for another description of $W_0^{1,p}(\Omega)$ in $\mathbb{R}^n$.

Proposition 9.4. *Let $E \subset \mathbb{R}^n$ be arbitrary and let $u \in N^{1,p}(\overline{E}^p)$, where $\overline{E}^p$ is the fine closure of E. Then $u \in N_0^{1,p}(E)$ if and only if $u = 0$ q.e. on the fine boundary $\overline{E}^p \setminus$ fine-int E of E.*

Proof. By the discussion at the beginning of this section, both fine-int E and $\mathbb{R}^n \setminus \overline{E}^p$ are p-path open. Lemma 3.9 with $E_1 = $ fine-int E and $E_2 = \overline{E}^p$ then yields that $u \in N_0^{1,p}($fine-int $E)$ if and only if $u = 0$ q.e. on the fine boundary of E. Theorem 7.3 concludes the proof. $\square$

In general metric spaces, the missing link is the implication that finely open sets are quasi-open. This is a part of fine potential theory on metric spaces which we plan to develop further in the future.

The following two examples illustrate some of the results in this paper, in particular the special situation in $\mathbb{R}^n$. They provide us with a closed nowhere dense set $E \subset [0,1]^n \subset \mathbb{R}^n$ with almost full measure in $[0,1]^n$, but whose fine interior has full measure in E. In particular, the fine boundary of E has zero measure even though the Euclidean boundary $\partial E = E$. This implies that for every $u \in D^p(\mathbb{R}^n)$,

$$g_{u,\text{fine-int } E} = g_{u,E} = g_{u,\mathbb{R}^n} = |\nabla u| \quad \text{a.e. in } E,$$

where ∇u is the distributional gradient of u, and that the energies and obstacle problems on E and its fine interior coincide. Examples 9.5 and 9.6 are for $1 < p < n$ and $p = n$, respectively. By Remark 9.1 there are no similar examples for $p > n$.

Recall that for $q, x \in \mathbb{R}^n$ and $r, s > 0$,

$$(9.1) \quad \text{cap}_p(B(q,s) \cap B(x,r), B(x,2r)) \leq \begin{cases} C(n,p)s^{n-p}, & \text{if } 1 < p < n, \\ C(n)\left(\log \dfrac{2r}{s}\right)^{1-n}, & \text{if } p = n, \end{cases}$$

and that

$$(9.2) \quad \text{cap}_p(B(x,r), B(x,2r)) = C(n,p)r^{n-p}, \quad 1 < p \leq n;$$
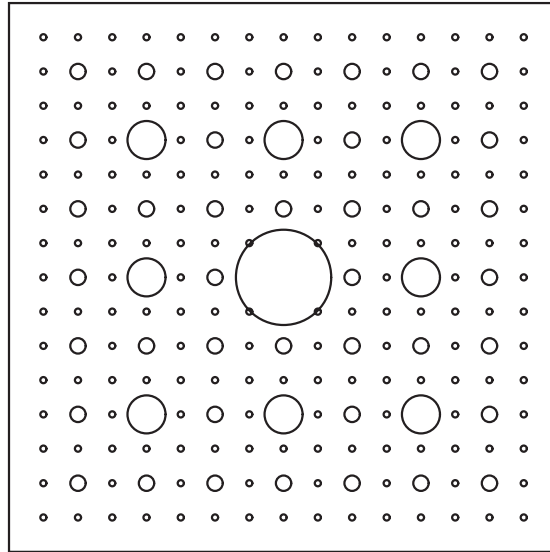
see Example 2.12 in Heinonen–Kilpeläinen–Martio [21].

FIGURE 1. The set E in Examples 9.5 and 9.6.

Example 9.5. Let $Q_k = ((0,1) \cap 2^{-k}\mathbb{N})^n$ be a bounded lattice in unweighted $\mathbb{R}^n$, $n \geq 2$, $k = 1, 2, \ldots$ Also let $a_k = 2^{-k}$ and $r_k = \delta a_k^\alpha$, $k = 1, 2, \ldots$, for some $0 < \delta < 1/2$ and $\alpha > n/(n-p)$, where $1 < p < n$. Note that for a fixed k, the balls $\{B(q, r_k)\}_{q \in Q_k}$ are disjoint. Finally, let

$$E = [0,1]^n \setminus \bigcup_{k=1}^\infty \bigcup_{q \in Q_k} B(q, r_k);$$

see Figure 1. Then $E \subset \mathbb{R}^n$ is a closed set with empty interior and

$$m([0,1]^n \setminus E) \leq C \sum_{k=1}^\infty (2^k - 1)^n r_k^n \leq C\,\delta^n \sum_{k=1}^\infty 2^{kn(1-\alpha)} \leq C\,\delta^n,$$

where m denotes the Lebesgue measure in $\mathbb{R}^n$. Thus, for small $\delta > 0$, E has almost full measure in $[0,1]^n$. We shall show that the set E has nonempty fine interior, and that $m(E \setminus \text{fine-int } E) = 0$.

For a fixed $0 < \theta < 1 - 1/\alpha$ and all $0 < \varepsilon < \delta$, we define

$$E_\varepsilon = [0,1]^n \setminus \bigcup_{k=1}^\infty \bigcup_{q \in Q_k} B(q, r_k + \varepsilon a_k^{1+\theta}).$$

Note that by the mean value theorem,

$$m(E \setminus E_\varepsilon) \leq \sum_{k=1}^{\infty} 2^{kn}\, m(B(0, r_k + \varepsilon a_k^{1+\theta}) \setminus B(0, r_k))$$

$$\leq C \sum_{k=1}^{\infty} 2^{kn}\, r_k^n \Big(\Big(1 + \frac{\varepsilon a_k^{1+\theta}}{r_k} \Big)^n - 1 \Big)$$

$$\leq C \sum_{k=1}^{\infty} 2^{kn}\, r_k^n \frac{\varepsilon a_k^{1+\theta}}{r_k}\, n \Big(1 + \frac{\varepsilon a_k^{1+\theta}}{r_k} \Big)^{n-1}.$$

As $\varepsilon < 1/2$ and $a_k^{1+\theta}/r_k > 1$, the last estimate can be simplified as

$$m(E \setminus E_\varepsilon) \leq C\,\varepsilon \sum_{k=1}^{\infty} 2^{kn}\, r_k^{n-1} a_k^{1+\theta} \Big(\frac{2a_k^{1+\theta}}{r_k} \Big)^{n-1}$$

$$= C\,\varepsilon \sum_{k=1}^{\infty} 2^{kn}\, a_k^{n(1+\theta)} = C\,\varepsilon \sum_{k=1}^{\infty} 2^{-kn\theta} \to 0, \quad \text{as } \varepsilon \to 0.$$

It follows that $m(E \setminus \bigcup_{\varepsilon > 0} E_\varepsilon) = 0$.

We claim that $\bigcup_{\varepsilon > 0} E_\varepsilon \subset \text{fine-int}\, E$. For this, it suffices to show that for every $x \in E_\varepsilon$, the set $X \setminus E$ is thin at x, in view of Theorem 1.2. Therefore, let $0 < \varepsilon < \delta$ and $x \in E_\varepsilon$ be fixed. We need to show that

$$(9.3) \qquad \sum_{j=j_\varepsilon}^{\infty} \Big(\frac{\text{cap}_p(B(x, 2^{-j}) \setminus E, B(x, 2^{1-j}))}{\text{cap}_p(B(x, 2^{-j}), B(x, 2^{1-j}))} \Big)^{1/(p-1)} < \infty$$

for some j_ε (possibly depending on x, α, θ and ε). Therefore, we let $r = 2^{-j}$ and estimate $\text{cap}_p(B(x, r) \setminus E, B(x, 2r))$.

We shall first estimate how many balls $B(q, r_k)$, with $q \in Q_k$ and $k < j$ (i.e., $a_k \geq 2r$), can intersect $B(x, r)$. Since for every $q \in Q_k$, $k = 1, 2, \ldots$, we have

$$\text{dist}(x, B(q, r_k)) \geq \varepsilon a_k^{1+\theta},$$

the intersection will be nonempty only if $\varepsilon 2^{-k(1+\theta)} < 2^{-j}$. This is equivalent to

$$(9.4) \qquad k > \frac{1}{1+\theta}\,(j + \log_2 \varepsilon) \geq \frac{(1-\theta^2)j}{1+\theta} = (1-\theta)j,$$

provided that

$$(9.5) \qquad j \geq \frac{1}{\theta^2}\,(-\log_2 \varepsilon).$$

In particular, for each ε and θ there exists j_ε such that (9.5) holds for all $j \geq j_\varepsilon$.

Moreover, for each $k < j$ as in (9.4), there are at most 2^n balls $B(q, r_k)$, $q \in Q_k$, intersecting $B(x, r)$, since $a_k \geq 2r$. By (9.1) their total capacity is at

most $C\, 2^n r_k^{n-p}$. Summing over all $k \in \mathbb{N}$ such that $(1-\theta)j < k < j$, yields the estimate for the capacity

$$(9.6) \qquad \sum_{(1-\theta)j<k<j} C\, r_k^{n-p} = C\, \delta^{n-p} \sum_{(1-\theta)j<k<j} 2^{-k\alpha(n-p)} \le C\, \delta^{n-p}\, 2^{-j\alpha(1-\theta)(n-p)}.$$

Now let $k \ge j$, i.e., $a_k \le r$. For each such k, there are at most $(4r/a_k)^n$ balls $B(q, r_k)$, $q \in Q_k$, intersecting $B(x,r)$. Their total capacity is at most

$$C\Big(\frac{4r}{a_k}\Big)^n r_k^{n-p} \le C\, 2^{n(k-j)}\, \delta^{n-p}\, 2^{-k\alpha(n-p)}.$$

Summing over $k \ge j$ and combining this with (9.6) yields for $r = 2^{-j}$ with $j \ge j_\varepsilon$,

$$\mathrm{cap}_p(B(x,r) \setminus E, B(x,2r)) \le C\delta^{n-p}\Big(2^{-j\alpha(1-\theta)(n-p)} + 2^{-jn}\sum_{k=j}^{\infty} 2^{-k(\alpha(n-p)-n)}\Big).$$

As $\alpha > n/(n-p)$, the last series sums to $C\, 2^{jn-j\alpha(n-p)}$ and we conclude that

$$\mathrm{cap}_p(B(x,r) \setminus E, B(x,2r)) \le C\, \delta^{n-p}\, 2^{-j\alpha(1-\theta)(n-p)}.$$

Inserting this and (9.2) into (9.3) shows that for each $x \in E_\varepsilon$ the sum in (9.3) is bounded from above by

$$\sum_{j=j_\varepsilon}^{\infty} \Big(\frac{C\delta^{n-p}2^{-j\alpha(1-\theta)(n-p)}}{2^{-j(n-p)}}\Big)^{1/(p-1)} = C\delta^{(n-p)/(p-1)} \sum_{j=j_\varepsilon}^{\infty} 2^{-j(\alpha(1-\theta)-1)(n-p)/(p-1)},$$

which is finite since $\alpha(1-\theta) > 1$.

Thus, $X \setminus E$ is thin at each $x \in E_\varepsilon$ and Theorem 1.2 shows that

$$\bigcup_{\varepsilon>0} E_\varepsilon \subset \text{fine-int } E.$$

Hence $m(E \setminus \text{fine-int } E) = 0$ and Theorem 1.3 implies that the minimal p-weak upper gradients with respect to E and $\mathbb{R}^n$ coincide, i.e., for every $u \in D^p(\mathbb{R}^n)$,

$$g_{u,\text{fine-int } E} = g_{u,E} = g_{u,\mathbb{R}^n} = |\nabla u| \quad \text{a.e. in } E.$$

Moreover, by Theorem 7.2, $N_0^{1,p}(E)$ is nontrivial and solutions of obstacle and Dirichlet problems on E are in general not equal to their boundary data. By Theorem 1.3 again, the solutions of the $\mathcal{K}_{\psi_1,\psi_2,f}(E)$- and $\mathcal{K}_{\psi_1,\psi_2,f}(\text{fine-int } E)$-obstacle problems coincide and have the same energies.

The following example is a modification of Example 9.5 for $p = n$. In particular, it covers the classical situation $p = n = 2$.

Example 9.6. If $p = n$, let E and E_ε be as in Example 9.5 but with $r_k = \delta\, 2^{-2^{k\alpha}}$ for some $\alpha > n/(n-1)$. As in Example 9.5, $E \subset \mathbb{R}^n$ is a nowhere dense closed set and

$$m([0,1]^n \setminus E) \le C\, \delta^n \sum_{k=1}^{\infty} 2^{kn-n2^{k\alpha}} \le C\, \delta^n.$$

That $m(E \setminus E_\varepsilon) \leq C\varepsilon \to 0$, as $\varepsilon \to 0$, is shown exactly as in Example 9.5. (Here it is enough to require that $0 < \theta < 1$.)

To show that $\bigcup_{\varepsilon > 0} E_\varepsilon \subset$ fine-int E, fix $x \in E_\varepsilon$ and let $r = 2^{-j}$, $j = 1, 2, \ldots$ As in Example 9.5, the ball $B(q, r_k)$ with $k < j$ intersects $B(x, r)$ only if

$$k > \frac{1}{1 + \theta}\,(j + \log_2 \varepsilon) \geq (1 - \theta)j \quad \text{provided that } j \geq j_\varepsilon \geq \frac{1}{\theta^2}\,(-\log_2 \varepsilon),$$

and for each such k there are at most 2^n such balls. By (9.1) each of these balls has capacity at most

$$C\Big(\log \frac{2r}{r_k}\Big)^{1-n} = C\,(1 - j - \log_2 \delta + 2^{k\alpha})^{1-n} \leq C\,2^{-k\alpha(n-1)}.$$

The total capacity of all such balls with $(1-\theta)j < k < j$ and $B(q, r_k) \cap B(x, r) \neq \varnothing$ is therefore at most

$$(9.7) \qquad \sum_{(1-\theta)j < k < j} C\,2^{-k\alpha(n-1)} \leq C\,2^{-j\alpha(1-\theta)(n-1)}.$$

Now, for each $k \geq j$, there are at most $(4r/a_k)^n$ balls $B(q, r_k)$, $q \in Q_k$, intersecting $B(x, r)$ and their total capacity is at most

$$C\Big(\frac{4r}{a_k}\Big)^n \Big(\log \frac{2r}{r_k}\Big)^{1-n} \leq C\,2^{n(k-j)}2^{-k\alpha(n-1)}.$$

Summing over $k \geq j$ and combining this with (9.7) yields for $r = 2^{-j}$, $j \geq j_\varepsilon$,

$$\operatorname{cap}_p(B(x, r) \setminus E, B(x, 2r)) \leq C\,2^{-j\alpha(1-\theta)(n-1)} + 2^{-jn}\sum_{k=j}^\infty 2^{-k(\alpha(n-1)-n)}.$$

As $\alpha > n/(n-1)$, the last series sums to $C\,2^{jn-j\alpha(n-1)}$ and we conclude that

$$\operatorname{cap}_p(B(x, r) \setminus E, B(x, 2r)) \leq C\,2^{-j\alpha(1-\theta)(n-1)}.$$

Inserting this and (9.2) into (9.3) shows that for each $x \in E_\varepsilon$ the sum in (9.3) is bounded from above by

$$\sum_{j=j_\varepsilon}^\infty (C\,2^{-j\alpha(1-\theta)(n-1)})^{1/(n-1)} = C\sum_{j=j_\varepsilon}^\infty 2^{-j\alpha(1-\theta)} < \infty.$$

Thus, $X \setminus E$ is thin at each $x \in E_\varepsilon$ and Theorem 1.2 shows that $\bigcup_{\varepsilon > 0} E_\varepsilon \subset$ fine-int E. Hence $m(E \setminus \text{fine-int}\,E) = 0$ and Theorem 1.3 implies that the minimal p-weak upper gradients with respect to E and $\mathbb{R}^n$ coincide almost everywhere on E. By Theorem 7.2, $N_0^{1,p}(E)$ is nontrivial and the solutions of the obstacle and Dirichlet problems on E are in general not equal to their boundary data. By Theorem 1.3 again, the solutions of the $\mathcal{K}_{\psi_1, \psi_2, f}(E)$- and $\mathcal{K}_{\psi_1, \psi_2, f}(\text{fine-int}\,E)$-obstacle problems coincide and have the same energies.

10. Further examples

Let $X = \mathbb{R}^2$ be equipped with $d\mu = dx + \alpha\,dx_1$, where dx is the two-dimensional Lebesgue measure on $\mathbb{R}^2$, dx_1 is the one-dimensional Lebesgue measure on $\mathbb{R}$ (extended as the zero measure on $\mathbb{R}^2 \setminus \mathbb{R}$), and α is a positive real constant.

Proposition 10.1. *Let $u \in N^{1,p}(X)$. Then the function*

$$(10.1) \qquad \tilde{g}_u = \begin{cases} |\nabla u| & in\ \mathbb{R}^2 \setminus \mathbb{R}, \\ |\partial_1 u| & in\ \mathbb{R}, \end{cases}$$

is a minimal p-weak upper gradient of u with respect to μ. Here ∇u is the distributional gradient on $\mathbb{R}^2$ and $\partial_1 u$ is the distributional derivative on $\mathbb{R}$.

Observe that $u \in N^{1,p}(\mathbb{R}^2, dx) \subset W^{1,p}(\mathbb{R}^2)$, and thus u has a distributional gradient. Similarly, $u|_{\mathbb{R}} \in N^{1,p}(\mathbb{R}, dx_1)$ is absolutely continuous on $\mathbb{R}$ and has a distributional derivative there. To prove Proposition 10.1 we need the following two auxiliary results which hold for arbitrary metric spaces X.

Lemma 10.2. *Any (rectifiable) curve $\gamma\colon [0, l_\gamma] \to X$ has an associated loop-erased simple curve $\tilde{\gamma}\colon [0, l_{\tilde{\gamma}}] \to X$.*

A loop along the curve γ is a part $\gamma|_{[t_0, t_1]}$ such that $0 \le t_0 < t_1 \le l_\gamma$ and $\gamma(t_0) = \gamma(t_1)$. Such a part can be removed by redefining $\gamma(t) = \gamma(t_0)$ for $t_0 < t < t_1$. By doing this iteratively in an appropriate way and then reparameterizing (see below) we can obtain a loop-free (i.e., simple) curve $\tilde{\gamma} \subset \gamma$ such that in particular $\int_{\tilde{\gamma}} g\,ds \le \int_\gamma g\,ds$ for all nonnegative Borel functions g. Note that a curve can have several different associated loop-erased simple curves, and that the loop-erased simple curve is a single point if $\gamma(0) = \gamma(l_\gamma)$.

Proof. As the length of γ is finite there is a longest loop (it may not be unique), unless γ is already loop-free. Remove it, as described above, and call the resulting curve γ_1. Repeat the procedure to produce γ_2, γ_3, etc. This can terminate after a finite number of steps with γ_n, which is then (after reparameterization with respect to arc length) the desired loop-erased simple curve $\tilde{\gamma}$.

Otherwise, we get curves $\gamma_j : [0, l_\gamma] \to X$, $j = 1, 2, \ldots$, which by the Ascoli theorem converge to a curve $\tilde{\gamma}$ with the same endpoints. (Note that here we need a version of the Ascoli theorem valid for metric space valued equicontinuous functions; see e.g. Royden [35], p. 169.) The resulting curve is a 1-Lipschitz map which (after reparameterization with respect to arc length) is the desired loop-erased simple curve $\tilde{\gamma}$. $\square$

Lemma 10.3. *Let X be equipped with measures μ_1 and μ_2 such that $\mu_1 \le \mu_2$. Then $N^{1,p}(X, \mu_2) \subset N^{1,p}(X, \mu_1)$ and for every $u \in N^{1,p}(X, \mu_2)$, the minimal p-weak upper gradients with respect to μ_1 and μ_2 satisfy $g_{u,\mu_1} \le g_{u,\mu_2}$ μ_1-a.e.*

Proof. The inclusion $N^{1,p}(X, \mu_2) \subset N^{1,p}(X, \mu_1)$ follows directly from the fact that upper gradients do not depend on the underlying measure and that $N^{1,p}(X, \mu_j)$, $j = 1, 2$, can be defined using only upper gradients.

To compare the minimal p-weak upper gradients, let $u \in N^{1,p}(X, \mu_2)$. It is easily verified that $\mathrm{Mod}_{p,\mu_1}(\Gamma) = 0$ whenever $\mathrm{Mod}_{p,\mu_2}(\Gamma) = 0$. Hence, the minimal p-weak upper gradient g_{u,μ_2} of u with respect to μ_2 is a p-weak upper gradient of u with respect to μ_1 and we conclude that $g_{u,\mu_1} \leq g_{u,\mu_2}$ μ_1-a.e. in X. $\square$

Corollary 10.4. *Let μ_1 and μ_2 be measures on X which support the p-Friedrichs inequality* (2.3). *Then so does the measure $\mu = \mu_1 + \mu_2$.*

Proof. Lemma 10.3 shows that $g_{u,\mu_j} \leq g_{u,\mu}$ μ_j-a.e. in X, $j = 1, 2$. Hence

$$\int_X g_{u,\mu_j}^p \, d\mu_j \leq \int_X g_{u,\mu}^p \, d\mu_j \leq \int_X g_{u,\mu}^p \, d\mu, \quad j = 1, 2.$$

The p-Friedrichs inequality with respect to μ_1 and μ_2, together with

$$\|u\|_{L^p(X,\mu)}^p = \|u\|_{L^p(X,\mu_1)}^p + \|u\|_{L^p(X,\mu_2)}^p,$$

then finishes the proof. $\square$

Proof of Proposition 10.1. Let

$$\tilde{u}(x) = \begin{cases} 0, & \text{if } u(x) = \pm\infty, \\ u(x), & \text{otherwise.} \end{cases}$$

Then $u = \tilde{u}$ q.e., by Proposition 1.30 in Björn–Björn [6], and thus Corollary 1.49 in [6] shows that u and $\tilde{u}$ have the same p-weak upper gradients. Hence, we may assume that u is real-valued for the rest of the proof.

Since $d\mu \geq dx$ on $\mathbb{R}^2$ and $d\mu \geq dx_1$ on $\mathbb{R}$, Lemma 10.3 implies that the minimal p-weak upper gradient with respect to μ satisfies $g_u \geq \tilde{g}_u$ μ-a.e. It is therefore enough to show that $\tilde{g}_u$ itself is also a p-weak upper gradient of u with respect to μ. This will be done by showing that it belongs to the $L^p(X)$-closure of the set of upper gradients of u. Proposition 2.10 in [6] then shows that $\tilde{g}_u$ is a p-weak upper gradient of u with respect to μ.

Let $\varepsilon > 0$. As $|\nabla u|$ is a minimal p-weak upper gradient of u with respect to dx, we can find an upper gradient $\tilde{g} \in L^p(\mathbb{R}^2, dx)$ of u such that $\|\tilde{g} - |\nabla u|\|_{L^p(\mathbb{R}^2, dx)} < \varepsilon$. Let

$$g = \begin{cases} \tilde{g} & \text{in } \mathbb{R}^2 \setminus \mathbb{R}, \\ |\partial_1 u| & \text{in } \mathbb{R}. \end{cases}$$

Then $\|g - \tilde{g}_u\|_{L^p(X)} = \|\tilde{g} - g_u\|_{L^p(\mathbb{R}^2, dx)} < \varepsilon$. We shall show that g is an upper gradient of u in $\mathbb{R}^2$. We can require $\partial_1 u$ above to be a Borel function on $\mathbb{R}$, by Proposition 1.2 in Björn–Björn [6]. Since $\tilde{g}$ is a Borel function, so is g.

Let $\gamma \colon [0, l_\gamma] \to X$ be a curve. If $\gamma \subset \mathbb{R}$, then

$$|u(\gamma(0)) - u(\gamma(l_\gamma))| \leq \int_\gamma |\partial_1 u| \, ds = \int_\gamma g \, ds.$$

Similarly, if $\{t : \gamma(t) \in \mathbb{R}\}$ is a finite set, then

$$|u(\gamma(0)) - u(\gamma(l_\gamma))| \leq \int_\gamma \tilde{g} \, ds = \int_\gamma g \, ds.$$

After possibly splitting any other curve into at most three parts we may assume that $\gamma(0), \gamma(l_\gamma) \in \mathbb{R}$, but $\gamma \not\subset \mathbb{R}$. If $\gamma(0) = \gamma(l_\gamma)$, there is nothing to prove (as u is real-valued). Otherwise, let $G = \{t : \gamma(t) \in \mathbb{R}^2 \setminus \mathbb{R}\}$ which is a nonempty open subset of $(0, l_\gamma)$. It can thus be written as a pairwise disjoint union $\bigcup_{i=1}^\infty I_i$ of open intervals. (Here we allow some of the intervals I_i to be empty.) For a fixed n let $G_n = \bigcup_{i=1}^n I_i$. Let $\pi(x, y) = (x, 0)$ be the orthogonal projection of $\mathbb{R}^2$ onto $\mathbb{R}$, and let

$$\gamma_n(t) = \begin{cases} \gamma(t), & t \in G_n, \\ \pi \circ \gamma(t), & t \in [0, l_\gamma] \setminus G_n. \end{cases}$$

Then γ_n is a rectifiable curve. The given parameterization need not be by arc length, but it is a 1-Lipschitz map. Let $\tilde{\gamma}_n$ be an associated loop-erased simple curve of γ_n, given by Lemma 10.2. Then $\tilde{\gamma}_n$ can be split into at most $2n + 1$ subcurves such that each subcurve either is completely in $\mathbb{R}$, or it intersects $\mathbb{R}$ only at its endpoints. Denote the union of the former by $\tilde{\gamma}_n \cap \mathbb{R}$, and the union of the latter by $\tilde{\gamma}_n \setminus \mathbb{R}$. Note that $\tilde{\gamma}_n \setminus \mathbb{R} \subset \gamma|_G$. Using that these subcurves have already been treated above, we conclude that

$$(10.2) \quad |u(\gamma(0)) - u(\gamma(l_\gamma))| \le \int_{\tilde{\gamma}_n \setminus \mathbb{R}} g \, ds + \int_{\tilde{\gamma}_n \cap \mathbb{R}} g \, ds \le \int_{\gamma|_G} g \, ds + \int_{\tilde{\gamma}_n \cap \mathbb{R}} g \, ds.$$

Since $\tilde{\gamma}_n$ is a simple curve we obtain that

$$\liminf_{n \to \infty} \int_{\tilde{\gamma}_n \cap \mathbb{R}} g \, ds = \liminf_{n \to \infty} \int_{\mathbb{R}} g \chi_{\tilde{\gamma}_n \cap \mathbb{R}} \, dx \le \int_{\mathbb{R}} g \chi_{\gamma \cap \mathbb{R}} \, dx \le \int_{\gamma|_{[0, l_\gamma] \setminus G}} g \, ds.$$

Here we have used dominated convergence, which is justified by the fact that the integrands in the second integral are dominated by $g \chi_{[-a, a]}$ for some $a > 0$, and $g \in L^p(\mathbb{R}) \subset L^1_{\mathrm{loc}}(\mathbb{R})$. (It is to justify the use of dominated convergence here that we need to use loop-erased simple curves.) Moreover, we have used that the arc length of the projection is bounded from above by the arc length of the original curve.

Inserting this into (10.2) shows that

$$|u(\gamma(0)) - u(\gamma(l_\gamma))| \le \int_{\gamma|_G} g \, ds + \int_{\gamma|_{[0, l_\gamma] \setminus G}} g \, ds = \int_\gamma g \, ds. \qquad \square$$

Remark 10.5. The same proof as in Proposition 10.1 shows that if ν is any positive Borel measure on $\mathbb{R}$ satisfying $0 < \nu(I) < \infty$ for every finite interval I, then the function

$$g_{u, \mu} = \begin{cases} |\nabla u| & \text{in } \mathbb{R}^2 \setminus \mathbb{R}, \\ g_{u, \nu} & \text{in } \mathbb{R}, \end{cases}$$

is a minimal p-weak upper gradient of u with respect to $d\mu = dx + d\nu$. Here ∇u is the distributional gradient on $\mathbb{R}^2$ and $g_{u, \nu}$ is the minimal p-weak upper gradient of u on $\mathbb{R}$ with respect to ν. (In this case, g in the proof of Proposition 10.1 consists of $\tilde{g}$ and an upper gradient approximating $g_{u, \nu}$ in $L^p(\mathbb{R}, \nu)$.) See Proposition 10.6

below and the comments after it for some results on one-dimensional minimal
p-weak upper gradients for different measures.

Note also that by Corollary 10.4, the p-Friedrichs inequality (2.3) holds for μ,
provided it holds for ν on $\mathbb{R}$. Combined with Proposition 10.6, this provides us
with many examples of nonstandard measures on $\mathbb{R}^2$ to which a large part of our
theory applies.

With a little bit more work we can show that the measure $d\mu = dx + \alpha\, dx_1$
on $\mathbb{R}^2$ supports a (q, p)-Poincaré inequality as in Definition 2.5, and not only a
p-Friedrichs inequality as in the above remark. Here $q = 2p/(2 - p)$ (for $p < 2$)
or $q < \infty$ (for $p \geq 2$) is the usual Sobolev exponent on $\mathbb{R}^2$. We can clearly assume
that $q \geq p$. Note however that μ is not doubling and therefore we cannot conclude
the (q, p)-Poincaré inequality directly from the $(1, 1)$-Poincaré inequality which
would have been somewhat simpler to derive.

To derive the (q, p)-Poincaré inequality, let $u \in N^{1,p}(\mathbb{R}^2, \mu)$ and $Q = I \times I' \subset \mathbb{R}^2$,
where $I, I' \subset \mathbb{R}$ are finite intervals of length R. We can assume that $0 \in I'$, as
otherwise $\mu|_Q$ is just the Lebesgue measure on Q. Also let $u_{Q,d\mu}$, $u_{Q,dx}$ and u_{I,dx_1}
be the integral averages of u over Q with respect to $d\mu$, dx and dx_1, respectively.
Split the left-hand side in the (q, p)-Poincaré inequality as

$$\left(\int_Q |u - u_{Q,d\mu}|^q\, d\mu \right)^{1/q} \leq \|u - u_{Q,dx}\|_{L^q(Q,dx)} + |Q|^{1/q}|u_{Q,dx} - u_{Q,d\mu}|$$

$$(10.3) \qquad\qquad + \alpha^{1/q}\|u - u_{I,dx_1}\|_{L^q(I,dx_1)} + (\alpha|I|)^{1/q}|u_{I,dx_1} - u_{Q,d\mu}|,$$

where $|Q|$ and $|I|$ are the two- and one-dimensional Lebesgue measures of Q and I,
respectively. The first and the third term are estimated using the usual Sobolev–
Poincaré inequalities on $\mathbb{R}^2$ and $\mathbb{R}$, respectively. For the second term we have
(using the fact that u is absolutely continuous on almost every line parallel to the
x_2-axis) that

$$|u_{Q,dx} - u_{Q,d\mu}| = \left| \frac{\mu(Q) - |Q|}{\mu(Q)} \fint_Q u\, dx - \frac{\alpha}{\mu(Q)} \int_I u\, dx_1 \right|$$

$$\leq \frac{\alpha}{\mu(Q)} \fint_{I'} \int_I |u(x_1, x_2) - u(x_1, 0)|\, dx_1\, dx_2$$

$$\leq \frac{\alpha}{\mu(Q)} \int_I \int_{I'} |\partial_{x_2} u(x_1, t)|\, dt\, dx_1 \leq \frac{\alpha|Q|^{1-1/p}}{\mu(Q)} \left(\int_Q |\nabla u|^p\, dx \right)^{1/p}.$$

Similarly,

$$|u_{I,dx_1} - u_{Q,d\mu}| = \left| \frac{\mu(Q) - \alpha|I|}{\mu(Q)} \fint_I u\, dx_1 - \frac{1}{\mu(Q)} \int_Q u\, dx \right|$$

$$\leq \frac{|Q|}{\mu(Q)} \fint_{I'} \fint_I |u(x_1, 0) - u(x_1, x_2)|\, dx_1\, dx_2$$

$$(10.4) \qquad\qquad \leq \frac{|I|\, |Q|^{1-1/p}}{\mu(Q)} \left(\int_Q |\nabla u|^p\, dx \right)^{1/p}.$$

Since $|\nabla u| \leq g_u$ a.e. on $\mathbb{R}^2$, $|u'| \leq g_u$ a.e. on $\mathbb{R}$, $dx \leq d\mu$ and $\alpha \, dx_1 \leq d\mu$, inserting (10.4) into (10.3) yields

$$\left(\int_Q |u - u_{Q,d\mu}|^q \, d\mu\right)^{1/q} \leq C(R) \left(\int_Q g_u^p \, d\mu\right)^{1/p},$$

where

$$C(R) = CR(|Q|^{1/q - 1/p} + |I|^{1/q - 1/p}) + \frac{\alpha |Q|^{1+1/q-1/p}}{\mu(Q)} + \frac{(\alpha |I|)^{1/q} |I| \, |Q|^{1-1/p}}{\mu(Q)}.$$

As $|Q| \leq \mu(Q)$, $|I| = R \leq \mu(Q)$ and $\alpha|Q| \leq R\mu(Q)$, this proves the (q,p)-Poincaré inequality on squares $Q \subset \mathbb{R}^2$. For balls, using the circumscribed squares, and Lemma 4.17 in Björn–Björn [6], gives a weak Poincaré inequality with dilation $\sqrt{2}$.

Similar arguments can be used in other situations, in particular on Euclidean spaces. Here we give a rather general one-dimensional result.

Proposition 10.6. *Let μ be a positive locally finite Borel measure on $\mathbb{R}$ with the Lebesgue–Radon–Nikodym decomposition $d\mu = w \, dx + d\sigma$, where $0 \leq w \in L^1_{\text{loc}}(\mathbb{R})$ is locally essentially bounded away from zero and $\sigma \perp dx$. Then for all $u \in N^{1,p}(\mathbb{R}, \mu)$, all $q \geq 1$ and all finite intervals $I \subset \mathbb{R}$,*

$$\left(\fint_I |u - u_{I,\mu}|^q \, d\mu\right)^{1/q} \leq 2 \, |I|^{1-1/p} \left(\frac{\mu(I)}{\operatorname{ess\,inf}_I w}\right)^{1/p} \left(\fint_I g_{u,\mu}^p \, d\mu\right)^{1/p},$$

where $|I|$ denotes the Lebesgue measure of I. In particular, $(\mathbb{R}, \mu)$ supports the p-Friedrichs inequality (2.3).

Moreover, for every $u \in N^{1,p}(\mathbb{R}, \mu)$, the minimal p-weak upper gradient of u with respect to μ is the function

$$(10.5) \qquad \tilde{g}_u = \begin{cases} |u'| & in \ A, \\ 0 & in \ \mathbb{R} \setminus A, \end{cases}$$

where u' is the distributional derivative, and A is any maximal null set of the singular part σ of μ with respect to the Lebesgue measure, i.e., $\sigma(A) = 0$ and $|\mathbb{R} \setminus A| = 0$.

Remark 10.7. Since $\int_E g_{u,\mu}^p \, d\mu = \int_E g_{u,\mu}^p w \, dx$, Proposition 10.6 shows that there is no need to consider measures with a singular part when solving the Dirichlet problem on $\mathbb{R}$, provided that the measure is locally bounded from below by a positive multiple of the Lebesgue measure. On the other hand, for obstacle problems it still makes sense to distinguish between μ and its absolutely continuous part $w \, dx$, since the presence of the singular part σ may influence the capacity C_p and hence the obstacle condition $\psi_1 \leq u \leq \psi_2$ q.e.

If μ is not bounded from below by a positive multiple of the Lebesgue measure, then Proposition 10.6 can fail, as shown by the following examples.

Example 10.8. Let $d\mu = |x|^\alpha \, dx$ with $\alpha > 2p - 1$ and $u(x) = |x|^{-\beta}$, where $1 \leq \beta < (\alpha + 1)/p - 1$. Then $u \in N^{1,p}_{\text{loc}}(\mathbb{R}, \mu)$ and $g_{u,\mu} = \beta|x|^{-\beta-1}$, but u is not a distribution, so $g_{u,\mu}$ cannot be its distributional derivative.

Example 10.9. Let $\{q_j\}_{j=1}^\infty$ be a countable dense subset of $\mathbb{R}$ and $\{a_j\}_{j=1}^\infty$ be a sequence of positive numbers such that $\sum_{j=1}^\infty a_j < \infty$. Let $p \geq 1$, $\alpha > p - 1$, $0 < \varepsilon < 1/\alpha$ and $1 \leq \beta < (\alpha + 1)/p$. Then the function

$$f(x) = 1 + \sum_{j=1}^\infty a_j |x - q_j|^{-\alpha\varepsilon}$$

belongs to $L^1_{\mathrm{loc}}(\mathbb{R})$ and is thus finite almost everywhere. Since $f \geq 1$ on $\mathbb{R}$, it follows that $w := f^{-1/\varepsilon} \in L^1_{\mathrm{loc}}(\mathbb{R})$ is positive almost everywhere and $w(x) < |x - q_j|^\alpha / a_j^{1/\varepsilon}$ for all $j = 1, 2, \ldots$

Let $d\mu = w\,dx$ and $u(x) = \sum_{j=1}^\infty a_j^{1+1/p\varepsilon} |x - q_j|^{-\beta}$. Since

$$\int_{-R}^R (a_j^{1/p\varepsilon} |x - q_j|^{-\beta})^p \, d\mu \leq \int_{-R}^R (a_j^{1/p\varepsilon} |x|^{-\beta})^p \frac{|x|^\alpha \, dx}{a_j^{1/\varepsilon}} = \int_{-R}^R |x|^{\alpha - \beta p} \, dx < \infty,$$

we see that $u \in L^p_{\mathrm{loc}}(\mathbb{R}, \mu)$. As $\int_a^b u(x)\,dx = \infty$ for every nonempty interval $(a, b) \subset \mathbb{R}$, Proposition 1.37 (c) in Björn–Björn [6] implies that the family of all rectifiable curves on $\mathbb{R}$ has zero $\mathrm{Mod}_{p,\mu}$-modulus. It follows that the zero function is a p-weak upper gradient with respect to μ of every function and hence $N^{1,p}(\mathbb{R}, \mu) = L^p(\mathbb{R}, \mu)$.

Proof of Proposition 10.6. Lemma 10.3 implies that $u \in N^{1,p}_{\mathrm{loc}}(\mathbb{R}, dx)$ and

$$g_{u,\mu} \geq g_{u,dx} = |u'| = \tilde{g}_u \quad dx\text{-a.e. in } \mathbb{R},$$

and hence $g_{u,\mu} \geq \tilde{g}_u$ μ-a.e. in A. Since $\tilde{g}_u = 0$ in $\mathbb{R} \setminus A$, we see that $g_{u,\mu} \geq \tilde{g}_u$ μ-a.e. in $\mathbb{R}$. Conversely, as u is absolutely continuous on $\mathbb{R}$, the fundamental theorem of calculus and the fact that $\tilde{g}_u = |u'|$ dx-a.e. show that for all $x \leq y \in \mathbb{R}$,

$$|u(x) - u(y)| \leq \int_x^y |u'(t)| \, dt = \int_x^y \tilde{g}_u \, dt,$$

i.e., $\tilde{g}_u$ is an upper gradient of u. Hence $g_{u,\mu} \leq \tilde{g}_u$ μ-a.e. in $\mathbb{R}$.

The fundamental theorem of calculus again, together with the Hölder inequality and the Fubini theorem, now yields (with $I = (a, b)$)

$$\int_I |u(x) - u(a)|^q \, d\mu(x) \leq |I|^{q-q/p} \int_I \left(\int_I |u'(t)|^p \, dt \right)^{q/p} d\mu(x)$$

$$\leq |I|^{q-q/p} \mu(I) \left(\int_I g_{u,\mu}^p \, dt \right)^{q/p}.$$

Since $dt \leq w^{-1} \, d\mu$, we obtain

$$\left(\fint_I |u(x) - u(a)|^q \, d\mu(x) \right)^{1/q} \leq |I|^{1-1/p} \left(\frac{\mu(I)}{\operatorname{ess\,inf}_I w} \right)^{1/p} \left(\fint_I g_{u,\mu}^p \, d\mu \right)^{1/p},$$

and the required inequality then follows by a standard argument in which the constant $u(a)$ is replaced by the mean value $u_{I,\mu}$; see e.g. Lemma 4.17 in Björn–Björn [6]. $\qquad\square$

We have seen that our theory can be applied directly to the measure $d\mu = dx + \alpha\, dx_1$ on $\mathbb{R}^2$, or even $d\mu = dx + w(x_1)\, dx_1$ for a suitable weight w, and we can thus study the minimizers of the corresponding energy. It might be of interest to see which equation they satisfy.

Let $\Omega \subset \mathbb{R}^2$ be a domain. In $\Omega \setminus \mathbb{R}$, a minimizer u with respect to μ is a minimizer with respect to the ordinary dx measure, and is hence, after redefinition on a set of capacity zero, a p-harmonic function and thus locally $C^{1,\alpha}$ in $\Omega \setminus \mathbb{R}$. As $u|_{\Omega \cap \mathbb{R}} \in N^{1,p}(\Omega \cap \mathbb{R}, dx_1)$, $u|_{\Omega \cap \mathbb{R}}$ must be absolutely continuous. Since all the points in $\Omega \cap \mathbb{R}$ are regular boundary points of $\{(x_1, x_2) \in \Omega \colon \pm x_2 > 0\}$ (for all $p > 1$), it follows that u is continuous across $\mathbb{R}$ and thus (after the redefinition above) u is continuous in Ω.

For simplicity we assume that $p = 2$. In this case u is harmonic, and thus real-analytic, in $\Omega \setminus \mathbb{R}$. It locally minimizes the energy

$$\int ((\partial_1 u)^2 + (\partial_2 u)^2)\, dx_1\, dx_2 + \int (\partial_1 u)^2\, w\, dx_1.$$

Therefore u must satisfy the corresponding Euler–Lagrange equation, which in weak form becomes

$$\int_\Omega \nabla u \cdot \nabla \varphi\, dx_1\, dx_2 + \int_{\Omega \cap \mathbb{R}} \partial_1 u\, \partial_1 \varphi\, w\, dx_1 = 0 \quad \text{for all } \varphi \in C_0^\infty(\Omega).$$

Consider a function $\varphi(x_1, x_2) = \varphi_1(x_1)\varphi_2(\tau x_2) \in C_0^\infty(\Omega)$, where $\tau \geq 1$ and $\varphi_2(0) = 1$. Inserting this into the Euler–Lagrange equation gives

(10.6)
$$\int_\mathbb{R} \left(\int_\mathbb{R} \partial_1 u(x_1, x_2)\partial_1 \varphi_1(x_1)\, dx_1 \right) \varphi_2(\tau x_2)\, dx_2$$
$$+ \int_\mathbb{R} \left(\int_\mathbb{R} \tau \partial_2 u(x_1, x_2)\partial_2 \varphi_2(\tau x_2)\, dx_2 \right) \varphi_1(x_1)\, dx_1 - \int_{\Omega \cap \mathbb{R}} \varphi_1 \partial_1(w\partial_1 u)\, dx_1 = 0.$$

After the change of variables $y = \tau x_2$, the inner integral in the second term becomes

$$\int_\mathbb{R} \partial_2 u(x_1, y/\tau)\partial_2 \varphi_2(y)\, dy$$

which tends to

$$\partial_2^- u(x_1, 0) \int_{-\infty}^0 \partial_2 \varphi_2\, dy + \partial_2^+ u(x_1, 0) \int_0^\infty \partial_2 \varphi_2\, dy = \partial_2^- u(x_1, 0) - \partial_2^+ u(x_1, 0),$$

as $\tau \to \infty$, where $\partial_2^\pm u(x_1, 0) = \lim_{x_2 \to 0\pm} \partial_2 u(x_1, x_2)$. Since the first term in (10.6) tends to 0, as $\tau \to \infty$, we obtain that

$$\int_\mathbb{R} (\partial_2^- u(x_1, 0) - \partial_2^+ u(x_1, 0))\varphi_1(x_1)\, dx_1 - \int_{\Omega \cap \mathbb{R}} \varphi_1 \partial_1(w\partial_1 u)\, dx_1 = 0.$$

Thus u needs to fulfill

$$\partial_2^- u(x_1, 0) - \partial_2^+ u(x_1, 0) = \partial_1(w\partial_1 u)(x_1, 0) \quad \text{for } x_1 \in \Omega \cap \mathbb{R}$$

(in a weak sense), in addition to being harmonic in $\Omega \setminus \mathbb{R}$. For this derivation we have assumed that u is sufficiently smooth.

A. Consequences of the Fuglede and Mazur lemmas

In this appendix we prove two convergence results, which were used in the earlier sections. They are generalizations to Dirichlet spaces of results from Björn–Björn–Parviainen [8] (which can also be found in Björn–Björn [6]). Note that these results hold on arbitrary metric spaces without any additional assumptions.

Proposition A.1. *Let $f_j \in D^p(X)$ have a p-weak upper gradient $g_j \in L^p(X)$, $j = 1, 2, \ldots$ Assume that $f_j - f \to 0$ and $g_j \to g$ in $L^p(X)$, as $j \to \infty$, and that g is nonnegative. Then there is a function $\tilde{f} = f$ a.e. such that g is a p-weak upper gradient of $\tilde{f}$, and thus $\tilde{f} \in D^p(X)$. There is also a subsequence $\{f_{j_k}\}_{k=1}^{\infty}$ such that $f_{j_k} \to \tilde{f}$ q.e., as $k \to \infty$.*

When we say that $f_j - f \to 0$ in $L^p(X)$ we implicitly require that $f_j - f \in L^p(X)$, which in particular requires that f_j and f are real-valued almost everywhere. Note that we do *not* require $f_j \in L^p(X)$ and can therefore *not* use Proposition 3.1 in [8] (nor Proposition 2.3 in [6]).

Proof. By passing to a subsequence if necessary we may assume that $f_j \to f$ a.e., and (by the Fuglede lemma; see Shanmugalingam [36], Lemma 3.4 and Remark 3.5, or Björn–Björn [6], Lemma 2.1), that $\int_\gamma g_j \, ds \to \int_\gamma g \, ds \in \mathbb{R}$, as $j \to \infty$, for all curves $\gamma \notin \Gamma$, where $\mathrm{Mod}_p(\Gamma) = 0$. Let $\tilde{f} = \limsup_{j \to \infty} f_j$, and observe that $\tilde{f}$ is defined at every point of X and $\tilde{f} = f$ a.e. in X. Let $A = \{x \in X : |\tilde{f}(x)| = \infty\}$.

By definition, p-almost every curve γ is such that (2.1) holds for all f_j and g_j, $j = 1, 2, \ldots$, on γ and all its subcurves, and neither γ nor any of its subcurves belongs to Γ. Consider such a curve $\gamma \colon [0, l_\gamma] \to X$. We see that either $\gamma(0), \gamma(l_\gamma) \in A$ or

$$|\tilde{f}(\gamma(l_\gamma)) - \tilde{f}(\gamma(0))| \leq \limsup_{j \to \infty} |f_j(\gamma(l_\gamma)) - f_j(\gamma(0))| \leq \limsup_{j \to \infty} \int_\gamma g_j \, ds = \int_\gamma g \, ds.$$

As $\mu(A) = 0$, Proposition 2.5 in Björn–Björn–Parviainen [8] (or Corollary 1.51 in Björn–Björn [6]) shows that g is indeed a p-weak upper gradient of $\tilde{f}$, and thus $\tilde{f} \in D^p(X)$.

Now let $\hat{f} = \liminf_{j \to \infty} f_j$. Arguing exactly as above we see that g is also a p-weak upper gradient of $\hat{f} \in D^p(X)$ and that $\hat{f} = f = \tilde{f}$ a.e. Hence $\hat{f} = \tilde{f}$ q.e., and thus $f_j \to \tilde{f}$ q.e., as $j \to \infty$. □

Lemma A.2. *Let $1 < p < \infty$ and $f \in D^p(X)$. Assume that g_j is a p-weak upper gradient of u_j, $j = 1, 2, \ldots$, and that both sequences $\{u_j - f\}_{j=1}^{\infty}$ and $\{g_j\}_{j=1}^{\infty}$ are bounded in $L^p(X)$. Then there are functions u and g and convex combinations $v_j = \sum_{i=j}^{N_j} a_{j,i} u_i$ with p-weak upper gradients $\bar{g}_j = \sum_{i=j}^{N_j} a_{j,i} g_i$, such that*

(a) *$u - f \in N^{1,p}(X)$ and $g \in L^p(X)$;*

(b) *both $v_j - u \to 0$ and $\bar{g}_j \to g$ in $L^p(X)$, as $j \to \infty$;*

(c) *$v_j \to u$ q.e., as $j \to \infty$;*

(d) *g is a p-weak upper gradient of u.*

Proof. Let $w_j = u_j - f$, $j = 1, 2, \ldots$ Then $g_{w_j} \leq g_j + g_f$ and $\{w_j\}_{j=1}^\infty$ is bounded in $N^{1,p}(X)$. Since $L^p(X)$ is reflexive, its unit ball is weakly compact (by the Banach–Alaoglu theorem) and thus there is a subsequence of $\{w_j\}_{j=1}^\infty$ which converges weakly in $L^p(X)$. Taking a subsequence of this subsequence and again using the Banach–Alaoglu theorem we obtain a subsequence (again denoted $\{w_j\}_{j=1}^\infty$) such that both $\{w_j\}_{j=1}^\infty$ and $\{g_j\}_{j=1}^\infty$ converge weakly in $L^p(X)$ say to w and g. As g_j, $j = 1, 2, \ldots$, are nonnegative we may choose g nonnegative.

Applying the Mazur lemma (see, e.g., Yosida [38], pp. 120–121) repeatedly to the sequences $\{w_i\}_{i=j}^\infty$, $j \geq 1$, we find convex combinations $w'_j = \sum_{i=j}^{N'_j} a'_{i,j} w_i$ such that $\|w'_j - w\|_{L^p(X)} < 1/j$. Let $v'_j = w'_j + f = \sum_{i=j}^{N'_j} a'_{i,j} u_i$. Then $g'_j := \sum_{i=j}^{N'_j} a'_{i,j} g_i$ is a p-weak upper gradient of v'_j. Since moreover $g'_j \to g$ weakly in $L^p(X)$, as $j \to \infty$, we can again apply the Mazur lemma (repeatedly) to obtain convex combinations $v_j = \sum_{i=j}^{N_j} a_{i,j} u_i$ with p-weak upper gradients $\bar{g}_j = \sum_{i=j}^{N_j} a_{i,j} g_i$ such that $v_j - v \to 0$ and $\bar{g}_j \to g$ in $L^p(X)$, as $j \to \infty$. By Proposition A.1, there is a function $u = v$ a.e. satisfying (b)–(d).

As $g + g_f \in L^p(X)$ is a p-weak upper gradient of $u - f \in L^p(X)$, we see that $u - f \in N^{1,p}(X)$. $\qquad\square$

B. The variational capacity cap_p on nonopen sets

In this appendix we define the variational capacity with respect to nonopen sets, which was used to prove the Adams criterion in Section 6. We also state those properties of the variational capacity that we have needed in this paper. For proofs of Lemma B.2 and Theorem B.3, and a considerably more extensive discussion, we refer to Björn–Björn [7].

Let $E \subset X$ be a nonempty bounded set.

Definition B.1. For an arbitrary set $A \subset E$ we define the *variational capacity*

$$\mathrm{cap}_p(A, E) = \inf \int_X g_u^p \, d\mu,$$

where the infimum is taken over all $u \in N_0^{1,p}(E)$ (extended by 0 outside E) such that $u \geq 1$ on A.

The infimum can equivalently be taken over all nonnegative $u \in N_0^{1,p}(E)$ such that $u = 1$ on A. If E is measurable we can also equivalently integrate over E instead of X. The variational capacity $\mathrm{cap}_p(A, E)$ can also be regarded as the condenser capacity $\mathrm{cap}_p(X \setminus E, A, X)$, as in Definition 5.12.

Note that as $N_0^{1,p}(E) \subset N^{1,p}(X)$, it is natural to consider the minimal p-weak upper gradient g_u with respect to X. On the other hand, by Proposition 3.10, $g_u = g_{u,E}$ in this case (if E is measurable).

Lemma B.2. *Assume that X supports the p-Friedrichs inequality* (2.3) *and that* $C_p(X \setminus E) > 0$. *Let* $A \subset E$. *Then* $C_p(A) = 0$ *if and only if* $\mathrm{cap}_p(A, E) = 0$.

Theorem B.3.

   (i) *If $A_1 \subset A_2 \subset E$, then* $\operatorname{cap}_p(A_1, E) \leq \operatorname{cap}_p(A_2, E)$;

  (ii) cap_p *is countably subadditive, i.e., if $A_1, A_2, \ldots \subset E$, then*

$$\operatorname{cap}_p\Big(\bigcup_{i=1}^{\infty} A_i, E \Big) \leq \sum_{i=1}^{\infty} \operatorname{cap}_p(A_i, E);$$

 (iii) *if $1 < p < \infty$ and $A_1 \subset A_2 \subset \cdots \subset E$, then*

$$\operatorname{cap}_p\Big(\bigcup_{i=1}^{\infty} A_i, E \Big) = \lim_{i \to \infty} \operatorname{cap}_p(A_i, E).$$

Observe the following more or less direct consequence of Theorem 7.3. We leave the proof to the reader.

Proposition B.4. *Assume that X is complete and supports a $(1, p)$-Poincaré inequality, that μ is doubling and that $p > 1$. Let $E \subset X$ be bounded and let $A \subset E$. Then*

$$\operatorname{cap}_p(A, E) = \begin{cases} \operatorname{cap}_p(A \cap \text{fine-int } E, \text{fine-int } E), & \text{if } C_p(A \setminus \text{fine-int } E) = 0, \\ \infty, & \text{if } C_p(A \setminus \text{fine-int } E) > 0. \end{cases}$$

References

[1] Adamowicz, T., Björn, A., Björn, J. and Shanmugalingam, N.: Prime ends for domains in metric spaces. *Adv. Math.* **238** (2013), 459–505.

[2] Adams, D. R.: Capacity and obstacle problems. *Appl. Math. Optim.* **8** (1981), 39–57.

[3] Björn, A.: Cluster sets for Sobolev functions and quasiminimizers. *J. Anal. Math.* **112** (2010), 49–77.

[4] Björn, A. and Björn, J.: Boundary regularity for p-harmonic functions and solutions of the obstacle problem on metric spaces. Preprint, LiTH-MAT-R-2004-09, Linköping, 2004.

[5] Björn, A. and Björn, J.: Boundary regularity for p-harmonic functions and solutions of the obstacle problem on metric spaces. *J. Math. Soc. Japan* **58** (2006), 1211–1232.

[6] Björn, A. and Björn, J.: *Nonlinear potential theory on metric spaces.* EMS Tracts in Mathematics 17, European Mathematical Society, Zürich, 2011.

[7] Björn, A. and Björn, J.: The variational capacity with respect to nonopen sets in metric spaces. *Potential Anal.* **40** (2014), 57–80.

[8] Björn, A., Björn, J. and Parviainen, M.: Lebesgue points and the fundamental convergence theorem for superharmonic functions on metric spaces. *Rev. Mat. Iberoam.* **26** (2010), 147–174.

[9] Björn, A., Björn, J. and Shanmugalingam, N.: Quasicontinuity of Newton–Sobolev functions and density of Lipschitz functions on metric spaces. *Houston J. Math.* **34** (2008), 1197–1211.

[10] Björn, A. and Marola, N.: Moser iteration for (quasi)minimizers on metric spaces. *Manuscripta Math.* **121** (2006), 339–366.

[11] Björn, J.: Boundary continuity for quasiminimizers on metric spaces. *Illinois J. Math.* **46** (2002), 383–403.

[12] Björn, J.: Fine continuity on metric spaces. *Manuscripta Math.* **125** (2008), 369–381.

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

[14] Choquet, G.: Theory of capacities. *Ann. Inst. Fourier Grenoble* **5** (1953–1954), 131–295.

[15] Farnana, Z.: The double obstacle problem on metric spaces. *Ann. Acad. Sci. Fenn. Math.* **34** (2009), 261–277.

[16] Friedrichs, K.: Die Randwert- und Eigenwertprobleme aus der Theorie der elastischen Platten. *Math. Ann.* **98** (1928), 205–247.

[17] Fuglede, B.: Extremal length and functional completion. *Acta Math.* **98** (1957), 171–219.

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

[19] Hajłasz, P.: Sobolev spaces on metric-measure spaces. In *Heat kernels and analysis on manifolds, graphs and metric spaces* (*Paris, 2002*), 173–218. Contemp. Math. 338, Amer. Math. Soc., Providence, RI, 2003.

[20] Hajłasz, P. and Koskela, P.: Sobolev met Poincaré. *Mem. Amer. Math. Soc.* **145** (2000), no. 688.

[21] Heinonen, J., Kilpeläinen, T. and Martio, O.: *Nonlinear potential theory of degenerate elliptic equations.* Second edition. Dover Publications, Mineola, NY, 2006.

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

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

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

[25] Kinnunen, J. and Latvala, V.: Fine regularity of superharmonic functions on metric spaces. In *Future trends in geometric function theory,* 157–167. Rep. Univ. Jyväskylä Dep. Math. Stat. 92, Univ. Jyväskylä, Jyväskylä, 2003.

[26] Kinnunen, J. and Martio, O.: Nonlinear potential theory on metric spaces. *Illinois Math. J.* **46** (2002), 857–883.

[27] Korte, R.: A Caccioppoli estimate and fine continuity for superminimizers on metric spaces. *Ann. Acad. Sci. Fenn. Math.* **33** (2008), 597–604.

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

[29] Latvala, V.: Finely superharmonic functions of degenerate elliptic equations. *Ann. Acad. Sci. Fenn. Ser. A I Math. Dissertationes* **96** (1994).

[30] Lukeš, J. and Malý, J.:, Thinness, Lebesgue density and fine topologies (an interplay between real analysis and potential theory). In *Summer school in potential theory* (*Joensuu, 1990*)*,* 35–70. Joensuun Yliop. Luonnont. Julk. 26, Univ. Joensuu, Joensuu, 1992.

[31] Malý, J. and Ziemer, W. P.: *Fine regularity of solutions of elliptic partial differential equations.* Mathematical Surveys and Monographs 51, American Mathematical Society, Providence, RI, 1997.

[32] Maz′ya, V. G.: *Sobolev spaces.* Springer Series in Soviet Mathematics, Springer-Verlag, Berlin, 1985.

[33] Maz′ya, V. G.: Conductor and capacitary inequalities for functions on topological spaces and their applications to Sobolev-type imbeddings. *J. Funct. Anal.* **224** (2005), 408–430.

[34] Rajala, T.: Local Poincaré inequalitites from stable curvature conditions on metric spaces. *Calc. Var. Partial Differential Equations* **44** (2012), 477–494.

[35] Royden, H. L.: *Real analysis.* Third edition. Macmillan Publishing Company, New York, 1988.

[36] Shanmugalingam, N.: Newtonian spaces: An extension of Sobolev spaces to metric measure spaces. *Rev. Mat. Iberoamericana* **16** (2000), 243–279.

[37] Shanmugalingam, N.: Harmonic functions on metric spaces. *Illinois J. Math.* **45** (2001), 1021–1050.

[38] Yosida, K.: *Functional analysis.* Sixth edition. Fundamental Principles of Mathematical Sciences 123, Springer-Verlag, Berlin-New York, 1980.

Anders Björn: Department of Mathematics, Linköpings universitet, SE-581 83 Linköping, Sweden.
E-mail: anders.bjorn@liu.se

Jana Björn: Department of Mathematics, Linköpings universitet, SE-581 83 Linköping, Sweden.
E-mail: jana.bjorn@liu.se