# QUASICONFORMAL MAPPINGS WITH SOBOLEV BOUNDARY VALUES

KARI ASTALA, MARIO BONK, AND JUHA HEINONEN

ABSTRACT. We consider quasiconformal mappings in the upper half space $\mathbb{R}^{n+1}_+$ of $\mathbb{R}^{n+1}$, $n \geq 2$, whose almost everywhere defined trace in $\mathbb{R}^n$ has distributional differential in $L^n(\mathbb{R}^n)$. We give both geometric and analytic characterizations for this possibility, resembling the situation in the classical Hardy space $H^1$. More generally, we consider certain positive functions defined on $\mathbb{R}^{n+1}_+$, called conformal densities. These densities mimic the averaged derivatives of quasiconformal mappings, and we prove analogous trace theorems for them. The abstract approach of general conformal densities sheds new light to the mapping case as well.

## CONTENTS



*Date*: February 20, 2002.

K.A. was supported by the Academy of Finland, Project 41933. M.B. was supported by a Heisenberg fellowship of the Deutsche Forschungsgemeinschaft. J.H. was supported by NSF grant DMS 9970427 and by a grant from the Väisälä Foundation. Part of this research was done when K.A. was visiting the University of Michigan, and M.B. and J.H. the University of Jyväskylä. The authors want to thank the respective institutes for their hospitality.

*Mathematics Subject Classification:* 30C65, 46E35.

## 1. Introduction

A quasiconformal mapping that is defined in the upper half space $\mathbb{R}^{n+1}_+$ of $\mathbb{R}^{n+1}$ has an almost everywhere defined trace on $\mathbb{R}^n$ via its radial (angular) limits. In this paper, we study what happens when this trace possesses some degree of regularity. More specifically, we consider the case when the distributional differential of the trace is in $L^n(\mathbb{R}^n)$ for $n \geq 2$. It turns out that this case admits interesting characterizations. For example, we shall show that the averaged derivative of a quasiconformal mapping satisfies a Hardy type $n$-summability condition if and only if its radial maximal function is $n$-summable in $\mathbb{R}^n$. The crucial observation here is that both of these conditions are equivalent to the $n$-summability of the differential of the trace. This result closely resembles the situation in the holomorphic Hardy space $H^1$.

As an application, we prove a compactness theorem for quasiconformal mappings of $\mathbb{R}^{n+1}_+$. This result can be applied in particular for a family of maps for which the Hausdorff $n$-measure of the boundary of the image domain has a fixed upper bound.

The characterizations that we present in this paper are not true for quasiconformal mappings in the plane. It is crucial for all the aforementioned results that the dimension of the boundary $\mathbb{R}^n = \partial\mathbb{R}^{n+1}_+$ is at least two. Ultimately, this dimensional restriction is due to the existence of singular boundary values for quasiconformal mappings of the upper half plane $\mathbb{R}^2_+$. Analytically, the failure is seen through the mapping properties of various maximal functions on $L^1$.

Basically our results are not really about quasiconformal mappings. In a recent paper [BKR], the authors developed a theory of "conformal densities" as a generalization of the theory of (quasi-)conformal mappings of the unit ball (alternatively, the upper half space) in Euclidean space. It turns out that our main result for quasiconformal mappings, Theorem 1.1, has a general formulation in terms of conformal densities. We present this generalization in Theorem 4.1, which should be regarded as the main result of this paper.

There are two reasons for doing such a generalization. First, the theory of conformal densities appears to be a natural tool in proving results also about quasiconformal mappings. This principle is demonstrated in the proof of Theorem 5.1 below, where trace results are established (compare Corollary 5.2). Second, we believe that the use of general densities, and its accompanying theory of Banach space-valued Sobolev functions as in [HKST], is of independent interest, and useful in other contexts as well.

The work here was partly motivated by our quest to try to solve the longstanding open problem of "inverse absolute continuity" for quasiconformal mappings on lower dimensional surfaces [G1], [G2], [G3], [V2], [BM]. This problem asks whether the restriction of a quasiconformal mapping to a smooth (hyper)surface transforms sets of positive surface measure to sets of positive surface measure in case the image of the hypersurface has locally finite Hausdorff measure of the correct dimension. A related question asks for a similar measure preservation property for quasisymmetric mappings from $\mathbb{R}^n$, $n \geq 2$, onto an arbitrary metric space of locally finite Hausdorff $n$-measure [HS, Question 15]. We hope that the techniques in the present paper will be useful in further attempts to conquer these open problems.

Before stating our results, we require some notation and terminology.

1.1. **Basic notation.** We denote by $\mathbb{R}^{n+1}_+$ the open upper half space in $\mathbb{R}^{n+1}$ and the points in $\mathbb{R}^{n+1}_+$ by $z = (x, t)$. Thus,

$$\mathbb{R}^{n+1}_+ = \{z = (x,t) : x \in \mathbb{R}^n \quad \text{and} \quad t > 0\}.$$

Although much of the preliminary discussion remains valid whenever $n \geq 1$, our main results demand that $n \geq 2$. The dimensional assumptions will be made precise as required.

Lebesgue $n$-measure is denoted by $m_n$. For $\alpha > 0$, the Hausdorff $\alpha$-dimensional measure in a metric space is written as $\mathcal{H}_\alpha$, while $\mathcal{H}^\infty_\alpha$ denotes the Hausdorff $\alpha$-content arising in the Carathéodory construction for Hausdorff measures. See [Fe, p. 170].

Mean values of integrable functions are denoted by barred integral signs, and the abbreviation

$$u_E = \fint_E u \tag{1.1}$$

is used for integrable functions $u$ and for measurable sets $E$ of positive measure. The measure in expressions like (1.1) is usually clear from the context.

Whenever $Q \subset \mathbb{R}^n$ is an $n$-cube, we denote its edge length by $\ell(Q)$, and then set

$$\widehat{Q} = Q \times [\ell(Q), 2\ell(Q)] \subset \mathbb{R}^{n+1}_+, \tag{1.2}$$

so that $\widehat{Q}$ is the upper half of the $(n+1)$-dimensional box $Q \times [0, 2\ell(Q)]$. For definiteness, we assume that cubes are closed. We denote by $z_Q$ the center of a cube, and by $Q_{x,t}$ the $n$-cube with sides parallel to the coordinate axes, with center $x \in \mathbb{R}^n$ and edge length $\ell(Q_{x,t}) = t$. Thus,

$$\widehat{Q_{x,t}} = Q_{x,t} \times [t, 2t]. \tag{1.3}$$

Note that the hyperbolic diameter of the cube $\widehat{Q}$ is a constant that is independent of $Q$, where the hyperbolic metric in $\mathbb{R}^{n+1}_+$ is determined by the length element $|dz|/t$.

We denote by $\mathcal{D}(\mathbb{R}^n)$ the countable collection of all dyadic cubes in $\mathbb{R}^n$. Thus, $Q \in \mathcal{D}(\mathbb{R}^n)$ if and only if the corners of $Q$ lie in $2^k\,\mathbb{Z}^n$ and $\ell(Q) = 2^k$ for some $k \in \mathbb{Z}$. Similarly, if $Q_0 \subset \mathbb{R}^n$ is a cube, we denote by $\mathcal{D}(Q_0)$ the dyadic subcubes of $Q_0$. Finally, set

$$\mathcal{W} = \{\widehat{Q} : Q \in \mathcal{D}(\mathbb{R}^n)\}. \tag{1.4}$$

The members of $\mathcal{W}$ will be referred to as (*dyadic*) *Whitney cubes* of $\mathbb{R}^{n+1}_+$.

We employ throughout the notation $a \simeq b$ (respectively, $a \lesssim b$) meaning that there is a positive constant $C > 0$, depending only on some obvious data at hand, such that $(1/C)a \le b \le Ca$ (respectively, $a \le Cb$). We also use the phrase $a \simeq b$ (resp. $a \lesssim b$) *with constants depending only on* $\alpha, \beta, \dots$ meaning that $C$ depends only on $\alpha, \beta, \dots$. Alternatively, we write $C = C(\alpha, \beta, \dots)$ in this case.

There will also be some self-explanatory notation, such as “diam” and “dist”.

1.2. **Quasiconformal mappings and averaged derivatives.** For the basic theory of quasiconformal mappings in Euclidean spaces we refer to [V1]. We make the convention that all quasiconformal mappings are sensepreserving. The notation $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$ means that $F$ is a mapping into $\mathbb{R}^{n+1}$. The image $D = F(\mathbb{R}^{n+1}_+)$ is explicitly mentioned only if it is relevant to the discussion at hand. The *maximal dilatation* of a quasiconformal mapping $F$ is denoted by $K(F)$ (see [V1, p. 42] for the definition).

If $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$ is a quasiconformal mapping, then the limit

$$\lim_{t\to 0} F(x,t) =: f(x) \tag{1.5}$$

exists for almost every $x \in \mathbb{R}^n$ with respect to Lebesgue $n$-measure. In fact, the limit in (1.5) exists for $x \in \mathbb{R}^n$ outside an exceptional set of $(n+1)$-capacity zero, in particular outside a set of Hausdorff dimension zero. (See [Zo], [Vu, Lemma 14.7 and Theorem 15.1], [BKR, Theorem 4.4 and Remark 4.5 (b)].) For now, we understand that $f$ is defined almost everywhere in $\mathbb{R}^n$ via expression (1.5). We return to the issue of quasieverywhere defined boundary values later in Section 5.

The *averaged derivative* $a_F$ of a quasiconformal mapping $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$ is defined by

$$a_F(x,t) = \left( \fint_{\widehat{Q_{x,t}}} J_F(z)\, dm_{n+1}(z) \right)^{1/(n+1)}, \tag{1.6}$$

where $J_F = \det(DF)$ designates the Jacobian determinant of $F$, which exists almost everywhere in $\mathbb{R}^{n+1}_+$. Recall that in this article quasiconformal mappings are assumed to be sensepreserving, so that $J_F$ is nonnegative.

Definition (1.6) (in an equivalent form) was first used in [AG], and it has turned out to be a natural quasiconformal analog for the absolute value of the derivative of a conformal mapping. We recall that quasiconformal mappings do not in general have pointwise defined derivatives.

The following three statements were proved in [AG]: For $\widehat{Q} \in \mathcal{W}$ and $z \in \widehat{Q}$ we have that

$$a_F(z) \simeq a_F(z_{\widehat{Q}}) =: a_F(\widehat{Q}) \tag{1.7}$$

and that

$$a_F(\widehat{Q}) \simeq \frac{\operatorname{diam}(F(\widehat{Q}))}{\operatorname{diam}(\widehat{Q})} \simeq \frac{\operatorname{dist}\big(F(\widehat{Q}), \partial D\big)}{\operatorname{diam}(\widehat{Q})}, \tag{1.8}$$

where $D = F(\mathbb{R}^{n+1}_+)$; moreover, if $\widehat{Q}$, $\widehat{Q}' \in \mathcal{W}$ are adjacent, we have that

$$a_F(\widehat{Q}) \simeq a_F(\widehat{Q}'), \tag{1.9}$$

with constants of comparability in all three cases depending only on $n$ and $K(F)$.

Properties (1.7) and (1.9) constitute a Harnack type inequality for $a_F$. It was observed in [BKR] that this Harnack inequality, together with another basic property of $a_F$ called a volume growth property, can be abstracted and a rich function theory emerges from a study of conformal densities satisfying these two properties. We shall discuss this in Subsection 2.1 below.

1.3. **The space** $\operatorname{Osc}^{n,\infty}$**.** Following Rochberg and Semmes [RS1], [RS2], we next introduce a Besov type space $\operatorname{Osc}^{n,\infty}$. Later in the paper, we shall need this space for functions with values in a Banach space and for this reason we take up the general definition here. For our first main theorem, Theorem 1.1, one can take $V = \mathbb{R}^{n+1}$ in what follows.

Let $V = (V, |\cdot|)$ be a Banach space. In the following, $Q_0$ will either be a cube in $\mathbb{R}^n$ or $Q_0 = \mathbb{R}^n$. We denote by $L^1(Q_0; V)$ the vector space of integrable and by $L^1_{\mathrm{loc}}(Q_0; V)$ the vector space of locally integrable measurable functions $f\colon Q_0 \to V$. Throughout, the notions involving vector-valued integration refer to the classical Bochner integral [DU, Chapter II], [BL, Section 5.1]. In particular, by definition functions in $L^1_{\mathrm{loc}}(Q_0;V)$ are assumed to be essentially separably valued; that is, given $f \in L^1_{\mathrm{loc}}(Q_0;V)$, there is a subset $Z$ of $Q_0$ of measure zero such that $f(Q_0 \setminus Z)$ is a separable subset of $V$.

For a cube $Q \subset Q_0$, the mean value

$$f_Q = \fint_Q f \, dm_n \in V \tag{1.10}$$

of a function $f \in L^1_{\mathrm{loc}}(Q_0;V)$ has the property that

$$\langle \Lambda, f_Q \rangle = \fint_Q \langle \Lambda, f \rangle \, dm_n \tag{1.11}$$

for all elements $\Lambda$ in the dual space $V^*$. Throughout, we shall use the standard pairing notation $\langle \Lambda, v \rangle = \Lambda(v)$ for $v \in V$ and $\Lambda \in V^*$.

We define the space $\mathrm{Osc}^{n,\infty}(\mathbb{R}^n;V)$ to consist of all functions $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n;V)$ for which the map

$$Q \mapsto A_f(Q) = \fint_Q |f - f_Q| \, dm_n \tag{1.12}$$

belongs to the weak Lebesgue space weak-$\ell^n$, uniformly over translated families of dyadic cubes. More precisely, $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n;V)$ belongs to $\mathrm{Osc}^{n,\infty}(\mathbb{R}^n;V)$ if and only if the "norm"

$$\|f\|_{\mathrm{Osc}^{n,\infty}} = \|f\|_{\mathrm{Osc}^{n,\infty}(\mathbb{R}^n;V)} = \sup_{w \in \mathbb{R}^n} \sup_{\lambda > 0} \lambda N(f,w,\lambda)^{1/n} \tag{1.13}$$

is finite, where

$$N(f,w,\lambda) := \#\{Q \in \mathcal{D}(\mathbb{R}^n) : A_f(Q - w) > \lambda\}. \tag{1.14}$$

Here # denotes cardinality, and

$$Q - w = \{y - w : y \in Q\}.$$

The expression $\|f\|_{\mathrm{Osc}^{n,\infty}}$ given in (1.13) can be defined alternatively as

$$\sup_{w \in \mathbb{R}^n} \sup_{\lambda > 0} \lambda \left( \#\{Q \in \mathcal{D}(\mathbb{R}^n) : A_{\tau_w(f)}(Q) > \lambda\} \right)^{1/n}, \tag{1.15}$$

where $\tau_w$ is the translation operator

$$\tau_w(f)(x) = f(x - w). \tag{1.16}$$

We also write

$$\tau_w(\widehat{Q}) = \widehat{Q - w}. \tag{1.17}$$

1.4. **The Riesz class.** We now present our first main theorem. In addition to the preceding discussion, recall notation from (1.6), (1.2), and (1.4).

**Theorem 1.1.** *Let* $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$, $n \geq 2$, *be a quasiconformal mapping with the radial boundary values* $f(x) = \lim_{t\to 0} F(x,t)$ *for a.e.* $x \in \mathbb{R}^n$. *Then the following five conditions are equivalent:*

(i) $a_F^*(\cdot) := \sup_{t>0} a_F(\cdot, t) \in L^n(\mathbb{R}^n)$,

(ii) $\|a_F\|_{QH^n} := \left(\sup_{t>0} \int_{\mathbb{R}^n} a_F(x,t)^n \, dm_n(x)\right)^{1/n} < \infty$,

(iii) $\sup_{\lambda>0} \lambda \left(\#\{Q \in \mathcal{D}(\mathbb{R}^n) : \operatorname{diam}\big(F(\widehat{Q})\big) > \lambda\}\right)^{1/n} < \infty$,

(iv) $f \in \operatorname{Osc}^{n,\infty}(\mathbb{R}^n; \mathbb{R}^{n+1})$,

(v) $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n; \mathbb{R}^{n+1})$ *and the distributional differential* $Df$ *belongs to* $L^n(\mathbb{R}^n)$.

*Moreover, the various Lebesgue type norms appearing in conditions* (i)–(v) *are equivalent with multiplicative constants only depending on the dimension* $n$ *and the dilatation* $K(F)$.

In Section 3, we introduce a Dirichlet-Sobolev space of functions with values in an arbitrary Banach space. We then formulate a more general version of Theorem 1.1 in terms of conformal densities in Section 4. See Theorem 4.1.

Observe that if $F$ is a conformal mapping in the upper half space $\mathbb{R}^2_+$, then $a_F(z) \simeq |F'(z)|$. Thus, for $n = 1$ and $F$ conformal, the equivalence that is analogous to the equivalence of (i) and (ii) in Theorem 1.1 is implied by the classical fact that the radial maximal function of a function in the Hardy space $H^1(\mathbb{R})$ belongs to $L^1(\mathbb{R})$ with bounds [Ga, p. 57]. We find it interesting that such an analytic fact about holomorphic functions related to the Hardy space theory has an analog in the nonlinear theory of quasiconformal mappings. Our proofs are necessarily different from the classical arguments.

First results about Hardy type conditions and quasiconformal mappings were proved in [J], [JW], [Zin], [As]. These papers mostly dealt with Hardy type conditions for the mapping itself rather than its derivative. The question about the equivalence of (i) and (ii) in Theorem 1.1 for $n \geq 2$ has been around for some time, although we have not seen it in print. The equivalence is false for quasiconformal mappings when $n = 1$, as shown by Hanson [Ha, Theorem E]. For recent studies on

Hardy type classes, quasiconformal mappings and their derivatives, see [AK], [BK].

To facilitate the future language, we introduce the following definition.

**Definition 1.2.** A quasiconformal mapping $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$, $n \geq 2$, is said to belong to the *Riesz class* if $F$ satisfies any of the equivalent conditions (i)–(v) in Theorem 1.1.

In 1916, F. and M. Riesz [Ri] proved that the derivative of a conformal mapping $F$, defined in the upper half plane $\mathbb{R}^2_+$, belongs to the Hardy space $H^1$ if and only if the boundary of the image domain has finite length. Moreover, as a function in $H^1$, the boundary values of $F'$ cannot vanish in a set of positive length, and $F$ preserves boundary sets of zero length.

For quasiconformal mappings, we have the following interesting open problem:

**Problem 1.3.** Let $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$, $n \geq 2$, be a quasiconformal mapping in the Riesz class. Is it then true that

$$m_n\big(\{x \in \mathbb{R}^n : Df(x) = 0\}\big) = 0\,? \tag{1.18}$$

In fact, if $F$ is in the Riesz class, then one can show that the sets

$$\{x \in \mathbb{R}^n : Df(x) = 0\} \quad \text{and} \quad \{x \in \mathbb{R}^n : \limsup_{t\to 0} a_F(x,t) = 0\} \tag{1.19}$$

agree up to a set of measure zero (cf. Corollary 5.2) . Moreover, if (1.18) were true, one could infer from this and from [BK, Lemma 6.2] that $\mathcal{H}_n(E) > 0$ implies $\mathcal{H}_n\big(f(E)\big) > 0$ for each Borel set $E \subset \mathbb{R}^n$, if $F$ is in the Riesz class. On the other hand, if the Hausdorff $n$-measure of the boundary of the image $F(\mathbb{R}^{n+1}_+) = D$ is finite, i.e., if $\mathcal{H}_n(\partial D) < \infty$, then $F$ belongs to the Riesz class by [BK, Definition 7.5, Theorems 7.6 and 7.8]. Thus an affirmative answer to Problem 1.3 would solve the problem of inverse absolute continuity, mentioned earlier in the introduction.

In search of a full higher dimensional analog of the F. and M. Riesz theorem, one faces some geometric measure theoretic problems that do not arise in the plane. We review next what is known about mappings in the Riesz class.

As we mentioned, $F$ belongs to the Riesz class, if $\mathcal{H}_n(\partial D) < \infty$, where $D = F(\mathbb{R}^{n+1}_+)$. On the other hand, $F$ can be in the Riesz class even when $D$ is bounded with $\mathcal{H}_n(\partial D) = \infty$. The precise statement here is that $F$ is in the Riesz class if and only if the Hausdorff $n$-measure of a porous part of the boundary is finite; the degree of porosity is dependent on $n$ and the dilatation of $F$. This fact is a recent result of Koskela and the second author; see [BK, Definition 7.5, Theorems 7.6 and 7.8, and Example 7.9].

Earlier, the following important special case was proved by Hanson [Ha, Theorem B]: if $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$ is a restriction of a global quasiconformal self mapping of $\mathbb{R}^{n+1}$, then $F$ is in the Riesz class if and only if $\mathcal{H}_n(\partial D) < \infty$. Even in this case, it is not known whether (1.18) and its consequence, the inverse absolute continuity, hold.

In the case of a global mapping as above, one has absolute continuity in the sense that sets of $n$-measure zero on the boundary get mapped to sets of Hausdorff $n$-measure zero, as proved by Gehring [G1], [G2], [G3] long ago. On the other hand, the third author has shown [H2, Theorem 1.3] that it is not true that $\mathcal{H}_n(E) = 0$ implies $\mathcal{H}_n\big(f(E)\big) = 0$ for every $F$ in the Riesz class, even when $F$ is bounded and has homeomorphic extension to $\mathbb{R}^n \cup \{\infty\}$ with $\mathcal{H}_n(\partial D) < \infty$.

The fact that examples as in [BK] and [H2] exist is accounted for not so much by the difference between the quasiconformal and conformal mappings, but rather by the fact that having finite Hausdorff $n$-measure generally means less when $n \geq 2$. For more precise statements about the implication "$\mathcal{H}_n(E) = 0 \Rightarrow \mathcal{H}_n\big(f(E)\big) = 0$", and its history, see the works [H2], [H3], [Se], [V3].

Finally, we recall that quasiconformal mappings of the upper half plane $\mathbb{R}^2_+$ onto itself need not be absolutely continuous on the boundary [BA], so that for $n = 1$ there are no results along the above lines.

Next we say a few words about the proofs of Theorem 1.1. Although the theorem will be derived from the more general Theorem 4.1, where no quasiconformal mappings are explicitly present, the ensuing remarks are still valid.

We have two proofs for the passage from (ii) to (i) in Theorem 1.1. The first one goes through the geometric condition (iii) and its analytic counterparts (iv) and (v). In fact, we shall prove

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow \text{(iv)} \Rightarrow \text{(v)} \Rightarrow \text{(i)}$$

for the implications in Theorem 1.1. Only the first implication here is trivial. For the second implication, we use the recent work [BK], where submartingale type estimates were proved for the averaged derivative of a quasiconformal mapping, and the ideas around the so called *wall theorem* [BK], [Ha], [H2], [V3]. The implication (iii) $\Rightarrow$ (iv) requires some harmonic analysis together with properties of quasiconformal mappings, and the ideas here are similar to those in [RS2]. The implication (iv) $\Rightarrow$ (v) is a general fact that has nothing to do with quasiconformal mappings; it is due to Stephen Semmes in collaboration with Connes, Sullivan, and Teleman [CST]. We shall present a Banach space-valued analog of this implication in Section 4. Finally, we prove the implication (v) $\Rightarrow$ (i) again by using some basic quasiconformal and harmonic analysis.

Our second proof for the implication (ii) $\Rightarrow$ (i) in Theorem 1.1 is more direct and avoids the Besov space $\mathrm{Osc}^{n,\infty}$. It relies, on the other hand, on the theory of Banach valued Sobolev functions. See Section 5 for this approach.

As an application of Theorems 1.1 and 4.1, we present a compactness result for quasiconformal mappings and conformal densities in the Riesz class in Section 6.

It is in the proof of Theorem 4.1 in Section 4 that we use the abstract theories of conformal densities and Banach space-valued Sobolev mappings. Although a direct proof could be given, we feel that avoiding the language of conformal densities in this context would amount to hiding a definite source of ideas. Moreover, with some amount of the basic theory from [BKR] and [HKST], reviewed in Sections 2 and 3, the implication (ii) $\Rightarrow$ (i) in Theorem 4.1 becomes rather transparent. (Compare Remark 5.3.)

We have chosen to formulate the results of this paper for quasiconformal mappings defined in the upper half space $\mathbb{R}^{n+1}_+$. This practice deviates from related earlier works where mappings of the open unit ball $\mathbb{B}^{n+1}$ were considered. This choice is mostly technical; it is more pleasant to work with $\mathbb{R}^n$ as the boundary rather than the sphere $\partial\mathbb{B}^{n+1}$. Consequently, many results remain true for mappings $\mathbb{B}^{n+1} \to \mathbb{R}^{n+1}$, $n \geq 2$, with obvious modifications in statements and in proofs.

## 2. Conformal densities

In this section, we collect some auxiliary results on conformal densities as defined in [BKR].

2.1. **Basic properties.** Consider a positive continuous function $\varrho$ defined in $\mathbb{R}^{n+1}_+$. (One can take $n \geq 1$ here although later our concern will be in the case $n \geq 2$.) Each such $\varrho$ determines a metric measure space

$$(\mathbb{R}^{n+1}_+, d_\varrho, \mu_\varrho) \tag{2.1}$$

which is a conformal deformation of $\mathbb{R}^{n+1}_+$. Thus,

$$d_\varrho(a,b) = \inf \int_\gamma \varrho\, ds, \quad a,b \in \mathbb{R}^{n+1}_+, \tag{2.2}$$

where the infimum is taken over all rectifiable paths $\gamma$ joining $a$ and $b$ in $\mathbb{R}^{n+1}_+$, and

$$\mu_\varrho(A) = \int_A \varrho^{n+1}(z)\, dm_{n+1}(z) \tag{2.3}$$

for a Borel set $A \subset \mathbb{R}^{n+1}_+$.

We call $\varrho$ a *conformal density* if there exist positive constants $C_1$ and $C_2$ such that the following two conditions, a *Harnack inequality* and a *volume growth condition*, hold:

$$\frac{1}{C_1}\varrho(a) \le \varrho(b) \le C_1\, \varrho(a) \tag{2.4}$$

whenever $Q \subset \mathbb{R}^n$ is a cube and $a, b \in \widehat{Q}$, and

$$\mu_\varrho\big(B_\varrho(x,R)\big) \le C_2\, R^{n+1} \tag{2.5}$$

whenever $B_\varrho(x,R)$ is an open metric ball in $(\mathbb{R}^{n+1}_+, d_\varrho, \mu_\varrho)$. The constants $C_1$, $C_2$, and the dimension $n$ together form the *data* of $\varrho$.

If $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$ is a quasiconformal mapping, then $\varrho = a_F$ satisfies (2.4) and (2.5) with constants $C_1$, $C_2$ depending only on $n$ and $K(F)$ [BKR, 2.4]. On the other hand, at least in dimensions $n+1 \ge 3$ there are conformal densities that are not comparable to the averaged derivative of any quasiconformal mapping [BHR, Theorem 1.8].

In the aforementioned sources, conformal densities are studied in the unit ball rather than the upper half space. However, the results that are cited here remain true in $\mathbb{R}^{n+1}_+$ as well, with only routine changes in the arguments.

Given a conformal density $\varrho$ in $\mathbb{R}^{n+1}_+$ and a Whitney cube $\widehat{Q} \in \mathcal{W}$, we set

$$\varrho(\widehat{Q}) := \varrho(z_{\widehat{Q}}) \tag{2.6}$$

analogously to (1.7); recall that $z_{\widehat{Q}}$ denotes the center of $\widehat{Q} \in \mathcal{W}$. We employ the notation

$$r_\varrho(\widehat{Q}) := \varrho(\widehat{Q})\, \mathrm{diam}(\widehat{Q}), \quad \widehat{Q} \in \mathcal{W}. \tag{2.7}$$

Then it is easy to see that

$$r_\varrho(\widehat{Q}) \simeq \mathrm{diam}_\varrho(\widehat{Q}) \tag{2.8}$$

and that

$$r_\varrho(\widehat{Q})^{n+1} \simeq \mu_\varrho(\widehat{Q}) \tag{2.9}$$

whenever $\widehat{Q} \in \mathcal{W}$. Here $\operatorname{diam}_\varrho$ designates the diameter in the metric $d_\varrho$ and the constants of comparability depend only on the data of $\varrho$. In light of (1.8), $r_\varrho(\widehat{Q})$ should be thought of as an abstract version of $\operatorname{diam}(F(\widehat{Q}))$. Indeed, $r_\varrho(\widehat{Q})$ is comparable to the diameter of the image of $\widehat{Q}$ under the identity map

$$I_\varrho := \mathrm{id}\colon \mathbb{R}^{n+1}_+ \to (\mathbb{R}^{n+1}_+, d_\varrho). \tag{2.10}$$

We shall use the notation $I_\varrho$ for the identity map in (2.10) for the sake of clarity.

We next explain why it is advantageous to consider the metric space $(\mathbb{R}^{n+1}_+, d_\varrho)$ and the map $I_\varrho$ in (2.10) even when $\varrho = a_F$ for a quasiconformal mapping $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$. The Harnack inequality (2.4) guarantees that the map $I_\varrho$ in (2.10) has nice local properties. In particular, up to a scaling it behaves like a bi-Lipschitz mapping on each Whitney cube with a bi-Lipschitz constant depending only on the data of $\varrho$. More precisely, for each conformal density $\varrho$ we have that

$$C^{-1}\varrho(\widehat{Q})|a-b| \le d_\varrho\big(I_\varrho(a), I_\varrho(b)\big) \le C\varrho(\widehat{Q})|a-b| \tag{2.11}$$

whenever $Q \subset \mathbb{R}^n$ is a cube, and $a, b \in \widehat{Q}$. Here $C \ge 1$ depends only on the data of $\varrho$. By using [BKR, Proposition 6.2] and an argument similar to that in [GO, Theorem 3, pp. 62–63], it follows therefore that the map $I_\varrho$ is bi-Lipschitz in the quasihyperbolic metrics of $\mathbb{R}^{n+1}_+$ and $(\mathbb{R}^{n+1}_+, d_\varrho)$.

*Remark* 2.1. Recall that the *quasihyperbolic metric* in a proper subdomain $D$ of $\mathbb{R}^{n+1}$ is obtained by changing the Euclidean metric by the conformal factor $\operatorname{dist}(x, \partial D)^{-1}$ [GO], [Vu]. A similar definition can be used in any noncomplete locally compact and rectifiably connected metric space such as $(\mathbb{R}^{n+1}_+, d_\varrho)$ [BHK]. Quasiconformal mappings between Euclidean domains are uniformly continuous in the quasihyperbolic metrics, and bi-Lipschitz for large quasihyperbolic distances [GO, Theorem 3], [Vu, Corollary 12.19]. On the other hand, locally in a given Whitney cube quasiconformal mappings need not be Lipschitz. The map $I_\varrho$ in (2.10) can be thought of as a local smoothening of $F$ if $\varrho = a_F$ for a quasiconformal mapping. The large scale geometry of the metric space $(\mathbb{R}^{n+1}_+, d_\varrho)$ is similar to that of $D = F(\mathbb{R}^{n+1}_+)$ equipped with its inner metric, and the study of the boundary behavior of $F$ can often be transferred to the study of the boundary behavior of $I_\varrho$. In other words, using $I_\varrho$ with some caution, we may assume that $F$

has a nice local behavior in $\mathbb{R}^{n+1}_+$. The price we have to pay for this supposition is that the target space is something more abstract than a Euclidean domain with is Euclidean (inner) boundary. The usefulness of this point of view will become clear later in the paper.

We can think of the transition from the mapping $F$ to the locally nice density $\varrho$ as a simple analytic substitute for the deep results of Sullivan, Tukia and Väisälä [Su], [TV] which (in dimensions $n+1 \neq 4$) allow for local Lipschitz smoothening of quasiconformal mappings while keeping the boundary values fixed.

If $\varrho$ is a conformal density in $\mathbb{R}^{n+1}_+$, then the identity map $I_\varrho$ in (2.10) has an almost everywhere defined trace on $\mathbb{R}^n$, which we denote by $i_\varrho$. The trace can be defined because the $d_\varrho$-length of the open line segment from $(x,1) \in \mathbb{R}^{n+1}_+$ to $x \in \mathbb{R}^n$ is finite for $m_n$-almost every point $x \in \mathbb{R}^n$; in fact, it is finite for $(n+1)$-capacity almost every $x \in \mathbb{R}^n$ [BKR, Theorem 4.4]. By using this fact about trace, in Section 4 we formulate a general $\varrho$-version of Theorem 1.1. In that version, the separable metric space $(\mathbb{R}^{n+1}_+, d_\varrho)$ is isometrically embedded in the Banach space $\ell^\infty$. The values of $i_\varrho$ on $\mathbb{R}^n$ lie in the completion of the metric space $(\mathbb{R}^{n+1}_+, d_\varrho)$, hence in $\ell^\infty$ when the embedding is understood.

2.2. **The Gehring-Hayman theorem.** One of the most important and clarifying tools in the theory of general conformal densities has turned out to be a general form of a theorem of Gehring and Hayman [GH]. The following result was proved in [BKR, Theorem 3.1].

**Proposition 2.2.** *Let $\varrho$ be a conformal density on $\mathbb{R}^{n+1}_+$, $n \geq 1$. Let $\gamma_h$ be a hyperbolic geodesic in $\mathbb{R}^{n+1}_+$ and let $\gamma$ in $\mathbb{R}^{n+1}_+$ be any curve with same end points as $\gamma_h$, where the end points are allowed to lie in $\mathbb{R}^n = \partial\mathbb{R}^{n+1}_+$. Then*

$$\int_{\gamma_h} \varrho\, ds \lesssim \int_\gamma \varrho\, ds. \tag{2.12}$$

*In particular, if the end points of $\gamma_h$ lie in $\mathbb{R}^n$, then*

$$r_\varrho(\widehat{Q}_0) \lesssim \int_\gamma \varrho\, ds, \tag{2.13}$$

*where $\widehat{Q}_0 \in \mathcal{W}$ denotes a Whitney cube of largest diameter that meets $\gamma_h$. The constants in* (2.12) *and* (2.13) *depend only on the data of $\varrho$.*

The classical Gehring-Hayman theorem corresponds to the case in Proposition 2.2, where $n = 1$ and $\varrho = |f'|$ for a conformal map $f\colon \mathbb{R}^2_+ \to \mathbb{R}^2$.

2.3. **Submartingale properties.** We turn into some more technical facts about conformal densities. Recall that the *conformal modulus* of a family $\Gamma$ of paths in $\mathbb{R}^N$ is the number

$$\operatorname{mod}_N(\Gamma) = \inf \int_{\mathbb{R}^N} \sigma^N dm_N, \tag{2.14}$$

where the infimum is taken over all Borel functions $\sigma\colon \mathbb{R}^N \to [0,\infty]$ such that

$$\int_\gamma \sigma\, ds \geq 1 \tag{2.15}$$

for all locally rectifiable paths $\gamma \in \Gamma$. A function $\sigma$ as above is called an *admissible density* for the curve family $\Gamma$. See [V1, Chapter1] for the basic theory of conformal modulus. We shall use conformal modulus in this paper both on $\mathbb{R}^{n+1}_+$ and on its boundary $\mathbb{R}^n$, for $n \geq 2$.

We state an important sub-mean value property of conformal densities, which follows from [BK, Proposition 5.5]:

**Proposition 2.3.** *Let $\varrho$ be a conformal density in $\mathbb{R}^{n+1}_+$, $n \geq 2$. Then*

$$r_\varrho(\widehat{Q}_0)^n \operatorname{mod}_n(\Gamma) \lesssim \sum_{Q\in\mathcal{S}} r_\varrho(\widehat{Q})^n \tag{2.16}$$

*whenever $Q_0 \subset \mathbb{R}^n$ is a cube, $\Gamma$ is a family of paths in $Q_0$ joining two opposite faces of $Q_0$, and $\mathcal{S}$ is a finite collection of cubes contained in $Q_0$ whose union covers the paths in $\Gamma$. The constant in* (2.16) *depends only on the data of $\varrho$.*

Inequality (2.16) implies in particular that the power $r_\varrho(\widehat{Q}_0)^n$ of the "diameter function" $r_\varrho$ defined in (2.7) is controlled by any sum of the values $r_\varrho(\widehat{Q})^n$, provided the union of the "shadows" $Q$ of the cubes $\widehat{Q}$ is equal to the shadow $Q_0$ of $\widehat{Q}_0$.

One should think of (2.16) as a nonlinear analog of the sub-mean value property of the subharmonic function $|F'(z)|$ for $F$ conformal. Inequality (2.16) means that $Q \mapsto \varrho(\widehat{Q})^n$ defines a submartingale on dyadic cubes in $\mathbb{R}^n$, up to a multiplicative constant. The following lemma shows that the appearance of this constant is inconsequential.

**Lemma 2.4.** *Let $\lambda\colon \mathcal{D}(Q_0) \to (0,\infty)$ be a positive function defined on dyadic subcubes of a cube $Q_0 \subset \mathbb{R}^n$ and suppose there exists a positive constant $C \geq 1$ such that*

$$\lambda(Q) \leq C \sum_{R\in\mathcal{S}} \lambda(R) \tag{2.17}$$

*whenever* $Q \in \mathcal{D}(Q_0)$ *and* $\mathcal{S} \subset \mathcal{D}(Q_0)$ *is a finite collection satisfying*

$$\bigcup_{R \in \mathcal{S}} R = Q. \tag{2.18}$$

*Then there exists a positive function* $\lambda' \colon \mathcal{D}(Q_0) \to (0, \infty)$ *such that*

$$\lambda'(Q) \leq \lambda(Q) \leq C\, \lambda'(Q) \tag{2.19}$$

*for each* $Q \in \mathcal{D}(Q_0)$ *and that*

$$\lambda'(Q) \leq \sum_{R \in \mathcal{S}} \lambda'(R) \tag{2.20}$$

*whenever* $Q \in \mathcal{D}(Q_0)$ *and* $\mathcal{S} \subset \mathcal{D}(Q_0)$ *is a collection satisfying* (2.18).

*Proof.* Define, for $Q \in \mathcal{D}(Q_0)$,

$$\lambda'(Q) := \inf \sum_{R \in \mathcal{S}} \lambda(R),$$

where the infimum is taken over all collections $\mathcal{S} \subset \mathcal{D}(Q_0)$ satisfying (2.18). Then $\lambda'$ satisfies (2.19) by definition and by assumption (2.17). To prove (2.20), fix $Q \in \mathcal{D}(Q_0)$ and a collection $\mathcal{S} \subset \mathcal{D}(Q_0)$ as in (2.18). Then fix $\varepsilon > 0$ and choose for each $R \in \mathcal{S}$ a collection $\mathcal{S}(R)$ such that

$$\sum_{P \in \mathcal{S}(R)} \lambda(P) \leq \lambda'(R) + \varepsilon.$$

Then

$$\lambda'(Q) \leq \sum_{R \in \mathcal{S}} \sum_{P \in \mathcal{S}(R)} \lambda(P) \leq \sum_{R \in \mathcal{S}} \lambda'(R) + \varepsilon(\#\mathcal{S}),$$

and the claim follows by letting $\varepsilon \to 0$. The lemma is proved. $\square$

**Proposition 2.5.** *Let* $\varrho$ *be a conformal density in* $\mathbb{R}^{n+1}_+$, $n \geq 2$, *such that*

$$\sup_{t>0} \int_{\mathbb{R}^n} \varrho(x,t)^n \, dm_n(x) < \infty. \tag{2.21}$$

*Then there exists a constant* $C \geq 1$ *depending only on the data of* $\varrho$ *such that*

$$\limsup_{t \to 0} \varrho(x,t) \leq C \liminf_{t \to 0} \varrho(x,t) < \infty \tag{2.22}$$

*for almost every* $x \in \mathbb{R}^n$. *In particular, the two sets*

$$\left\{x \in \mathbb{R}^n : \liminf_{t \to 0} \varrho(x,t) = 0\right\} \text{ and } \left\{x \in \mathbb{R}^n : \limsup_{t \to 0} \varrho(x,t) = 0\right\} \tag{2.23}$$

*are equal up to a set of Lebesgue* $n$*-measure zero.*

*Proof.* By the Harnack inequality (2.4), and by (2.16) and Lemma 2.4, we find that there is a submartingale $Q \mapsto \varrho'(\widehat{Q})^n$ defined on dyadic cubes $Q$ of $\mathbb{R}^n$ such that $\varrho(\widehat{Q}) \simeq \varrho'(\widehat{Q})$. Assumption (2.21) implies that this submartingale is $L^1$-bounded. Assertion (2.22) then follows from the submartingale convergence theorem [D, p. 450]. $\square$

It is an interesting open problem to determine whether, for a conformal density $\varrho$, the sets appearing in (2.23) indeed have $n$-measure zero, provided (2.21) holds. If that were the case, one could show by a standard "sawtooth domain" argument that the set appearing on the right in (2.23) has $n$-measure zero for each conformal density, independently whether (2.21) holds or not. Such a result could be seen as a nonlinear analog of the classical theorem of Lusin-Privalov [Pr, p. 210 and 212], [Po, p. 126]. See Problem 4.2 below for a precise statement of this problem.

Finally, we invite the reader to compare the above discussion with that in [Ma], where the relationship between conformal maps and martingale theory is discussed.

*Remark* 2.6. Although $L^1$-bounded submartingales have limits almost everywhere, one cannot expect the same to be true for a given conformal density satisfying (2.21). On the other hand, if $\varrho$ is as in Proposition 2.5, then one can always replace $\varrho$ by a comparable (in general discontinuous) conformal density $\varrho'$ for which

$$\lim_{t \to 0} \varrho'(x,t) =: \bar{\varrho}'(x) \tag{2.24}$$

exists for almost every $x \in Q_0$. For example, one can define $\varrho'$ on the cubes in $\mathcal{W}$ according to Lemma 2.4. For many problems in practice, the switch between two comparable densities is of no consequence.

The fact that such a function $\varrho'$ is only piecewise continuous makes no difference for the theory developed in [BKR] and elsewhere; the crucial properties, the Harnack inequality (2.4) and the volume growth condition (2.5) remain intact, and this suffices. We shall not use (2.24) in this paper however.

## 3. Banach space-valued Sobolev functions

In this section, the basic theory of abstract Sobolev spaces is reviewed. We follow the approach based on upper gradients as in [HKST]. An alternative, and an essentially equivalent way would be to use duality as in [Am], [Re], or "energy densities" as in [KS]. We prefer the upper gradient approach mainly because it fits well to the theory of conformal densities; the latter are upper gradients in a natural way.

The upper gradient approach also provides more information about pointwise behavior of Sobolev functions: they are automatically quasicontinuous and quasieverywhere defined. On the other hand, we have to pay a price for these advantages; the fact that Sobolev functions are automatically quasieverywhere defined means that we cannot freely move within the Lebesgue equivalence class of a function. This inflexibility causes some subtle technical problems. Therefore, in the discussion to follow, we assume initially that all the functions are pointwise defined. Later the equivalence classes based on capacity will be discussed.

For simplicity, we shall only consider functions whose domain of definition $Q_0$ is either a cube in $\mathbb{R}^n$, or all of $\mathbb{R}^n$. We assume also that $n \geq 2$. For the target, we let $V = (V, |\cdot|)$ be an arbitrary Banach space.

Given an arbitrary function $f\colon Q_0 \to V$, a Borel function $\sigma\colon Q_0 \to [0, \infty]$ is said to be an *upper gradient* of $f$ if

$$|f(a) - f(b)| \leq \int_\gamma \sigma \, ds \tag{3.1}$$

whenever $a$ and $b$ are two points in $Q_0$ and $\gamma$ is a rectifiable path in $Q_0$ joining $a$ and $b$.

We define the *Dirichlet-Sobolev space* $L^{1,n}(Q_0; V)$ to consist of all measurable and locally integrable functions $f\colon Q_0 \to V$ such that there exists an upper gradient of $f$ in $L^n(Q_0)$. There is a natural seminorm in $L^{1,n}(Q_0; V)$ given by

$$\|f\|_{L^{1,n}} = \|f\|_{L^{1,n}(Q_0;V)} := \inf \|\sigma\|_{L^n(Q_0)}, \tag{3.2}$$

where the infimum is taken over all upper gradients $\sigma$ of $f$.

We say that $f \in L^{1,n}(Q_0; V)$ is *separably valued* if $f(Q_0)$ is a separable subset of $V$. Recall from the introduction that locally integrable functions are assumed to be essentially separably valued in any case. The following proposition follows from [HKST, Corollary 6.8]. (Note that in [HKST], one generally assumes $n$-summability both for the function and for its upper gradient, but this assumption has no bearing on the results used and quoted here.)

**Proposition 3.1.** *Let $Q_0 \subset \mathbb{R}^n$ be a cube or $Q_0 = \mathbb{R}^n$, let $V$ be a Banach space, and let $f \in L^{1,n}(Q_0; V)$. Then $f$ is $n$-quasicontinuous and upon changing its values in a set of $n$-capacity zero $f$ becomes separably valued.*

A function $f$ is said to be *$n$-quasicontinuous* if there are open sets of arbitrarily small $n$-capacity with $f$ continuous in the complement.

The *n-capacity* of a measurable set $A \subset \mathbb{R}^n$ is defined by

$$\mathrm{cap}_n(A) = \inf ||u||^n_{W^{1,n}(\mathbb{R}^n)}\,, \tag{3.3}$$

where the infimum is taken over all functions $u$ in the standard Sobolev space $W^{1,n}(\mathbb{R}^n)$ such that $u \geq 1$ a.e. in a neighborhood of $A$. An event is said to take place *n-quasieverywhere* if it takes place outside a set of $n$-capacity zero. See [HKM, Chapter 4] or [MZ, Section 2.1] for the discussion of the above concepts in the case $V = \mathbb{R}$. The definitions and basic facts are similar in the general case.

If $V = \mathbb{R}^m$ for some $m \geq 1$, then the above definition leads to the standard Dirichlet-Sobolev space $L^{1,n}(Q_0; \mathbb{R}^m)$ consisting of all locally integrable $m$-tuples of $n$-quasicontinuous functions with $n$-summable distributional gradients [Sh, Theorem 4.5]. We abbreviate $L^{1,n}(Q_0; \mathbb{R}) = L^{1,n}(Q_0)$. Recall that every Lebesgue equivalence class of a function with $n$-summable distributional gradient on a domain in Euclidean space has an $n$-quasicontinuous representative which is unique up to a set of capacity zero. The approach via upper gradients automatically picks up the good representative.

For the applications in this paper, it is important to know how to handle convergence and compactness for Banach space-valued Sobolev functions. It is rather easy to see that there cannot be a compactness theorem of the Rellich-Kondrachev type for Banach space-valued functions. The way to bypass this difficulty is to consider the "components" of the functions. To this end, it is necessary to have a weaker notion of upper gradient.

A Borel function $\sigma\colon Q_0 \to [0, \infty]$ is said to be an *n-weak upper gradient* of a function $f\colon Q_0 \to V$, if $\sigma$ satisfies the upper gradient inequality (3.1) for every curve outside a family $\Gamma$ of curves of $n$-modulus zero. It is easy to see that $\mathrm{mod}_n(\Gamma) = 0$ if and only if there exists a Borel function $\tau\colon Q_0 \to [0, \infty]$, $\tau \in L^n(Q_0)$, such that

$$\int_\gamma \tau\, ds = \infty \tag{3.4}$$

for every locally rectifiable path $\gamma \in \Gamma$. It therefore follows that the existence of an $n$-weak upper gradient suffices for the membership in $L^{1,n}(Q_0; V)$, and moreover that

$$\|f\|_{L^{1,n}(Q_0;V)} = \inf \|\sigma\|_{L^n(Q_0)}, \tag{3.5}$$

where the infimum is taken over all $n$-weak upper gradients of $f$.

By using the concept of a weak upper gradient, and [Sh, Lemma 3.6], we find that if two functions $f_1, f_2 \in L^{1,n}(Q_0; V)$ agree outside a set of $n$-capacity zero, then $||f_1 - f_2||_{L^{1,n}(Q_0;V)} = 0$. We shall, therefore, make the convention that *a function in* $L^{1,n}(Q_0; V)$ *is pointwise* *defined outside a set of $n$-capacity zero.* In particular, two functions in $L^{1,n}(Q_0;V)$ are identified if they agree outside a set of $n$-capacity zero.

See [KM] or [Sh] for a detailed theory of upper gradients based on the use of modulus. The Banach space-valued discussion can be found in [HKST].

**Proposition 3.2.** *Let $Q_0 \subset \mathbb{R}^n$ be a cube or $Q_0 = \mathbb{R}^n$, let $V$ be a Banach space, and let $f\colon Q_0 \to V$ be a separably valued function. Then upon changing the Lebesgue equivalence class of $f$ the following three conditions are equivalent:*

(i) *$f$ belongs to the Dirichlet-Sobolev space $L^{1,n}(Q_0;V)$;*

(ii) *for each $\Lambda \in V^*$ with $|\Lambda| \leq 1$ the function $\langle \Lambda, f\rangle$ belongs to the standard Dirichlet-Sobolev space $L^{1,n}(Q_0)$ and there exists a function $\sigma \in L^n(Q_0)$ that is an $n$-weak upper gradient of $\langle \Lambda, f\rangle$ for each such $\Lambda$;*

(iii) *for each $\Lambda \in V^*$ with $|\Lambda| \leq 1$ the function $\langle \Lambda, f\rangle$ has a distributional gradient in $L^n(Q_0)$ and there exists a function $\sigma \in L^n(Q_0)$ such that*

$$|\nabla \langle \Lambda, f\rangle(x)| \leq \sigma(x) \tag{3.6}$$

*for almost every $x \in Q_0$, for each such $\Lambda$.*

*Moreover, in both cases* (ii)–(iii) *we have that*

$$\|f\|_{L^{1,n}} = \inf \|\sigma\|_{L^n(Q_0)}, \tag{3.7}$$

*where the infimum is taken over all such common $\sigma$.*

Proposition 3.2 follows from [HKST, Theorem 3.17 and Proposition 5.4]. Notice the distinction between (ii) and (iii): according to our definitions, the former requires that $\langle \Lambda, f\rangle$ is quasicontinuous for all $\Lambda$, while the latter only implies that $\langle \Lambda, f\rangle$ has a quasicontinuous Lebesgue representative, a priori dependent on $\Lambda$. Also notice that the requirement that $f$ be separably valued is automatically satisfied for $f \in L^{1,n}(Q_0;V)$ by Proposition 3.1.

**Proposition 3.3.** *Let $Q_0 \subset \mathbb{R}^n$ be a cube or $Q_0 = \mathbb{R}^n$, let $V$ be a Banach space, and let $(f_\nu)$ be a sequence of mappings in $L^{1,n}(Q_0;V)$ with corresponding upper gradient sequence $(\sigma_\nu)$. Assume that*

$$\sup_\nu ||\sigma_\nu||_{L^n(Q_0)} < \infty \tag{3.8}$$

*and that the functions $f_\nu$ converge pointwise almost everywhere in $Q_0$ to a mapping $f \in L^1_{\mathrm{loc}}(Q_0;V)$ with separable image. Then upon changing the Lebesgue equivalence class of $f$ we have that $f \in L^{1,n}(Q_0;V)$. If $f_\nu \to f$ $n$-quasieverywhere, then automatically $f \in L^{1,n}(Q_0;V)$.*

*Finally, if* $\sigma \in L^n(Q_0)$ *is any weak limit in* $L^n(Q_0)$ *of a subsequence of* $(\sigma_\nu)$*, then* $\sigma$ *is an* $n$*-weak upper gradient of* $f$*. In particular,*

$$\|f\|_{L^{1,n}} \leq \liminf_{\nu\to\infty} \|f_\nu\|_{L^{1,n}}. \tag{3.9}$$

*Proof.* The proposition is essentially contained in the arguments in [Sh], [HKST], but in the lack of a precise reference, we provide the details. We use Proposition 3.2.

Thus, fix $\Lambda \in V^*$ with $|\Lambda| \leq 1$. By Proposition 3.2, the functions $\langle \Lambda, f_\nu \rangle$ belong to the standard Dirichlet-Sobolev space $L^{1,n}(Q_0)$ with uniformly bounded norms. Therefore, because $\langle \Lambda, f_\nu \rangle \to \langle \Lambda, f \rangle$ pointwise almost everywhere in $Q_0$, we have by standard reasoning that $\langle \Lambda, f \rangle$ has a distributional gradient in $L^n(Q_0)$. To be more precise, we may clearly assume that $Q_0$ is a finite cube. Then by the Sobolev embedding theorem [GT, Chapter 7] (compare (6.6) below) we have that $(\langle \Lambda, f_\nu \rangle)$ is a bounded sequence in the standard Sobolev space $W^{1,n}(Q_0)$. By the reflexivity of $W^{1,n}(Q_0)$, the sequence $(\langle \Lambda, f_\nu \rangle)$ subconverges weakly to a function $u$ in $W^{1,n}(Q_0)$. But because $\langle \Lambda, f_\nu \rangle \to \langle \Lambda, f \rangle$ pointwise almost everywhere, we must have that $u = \langle \Lambda, f \rangle$. This latter conclusion is a standard consequence either of Mazur's lemma [Y, Section V. 1], or of the compactness of the embedding of $W^{1,n}(Q_0)$ in $L^n(Q_0)$ [GT, Chapter 7].

Next we show that

$$|\nabla \langle \Lambda, f \rangle(x)| \leq \sigma(x) \tag{3.10}$$

for almost every $x \in Q_0$, whenever $\sigma \in L^n(Q_0)$ is any weak limit of a subsequence of $(\sigma_\nu)$. Indeed, it then follows from Proposition 3.2 that there is a Lebesgue representative of $f$ in $L^{1,n}(Q_0; V)$. The proof of (3.10) is an application of Mazur's lemma: by passing to subsequences and convex combinations, both for the functions $\sigma_\nu$ and for the corresponding functions $f_\nu$, we may assume that $\nabla\langle \Lambda, f_\nu \rangle \to \nabla\langle \Lambda, f \rangle$ and $\sigma_\nu \to \sigma$ in $L^n(Q_0)$ and pointwise almost everywhere; notice here that the upper gradient inequality (3.1) is insensitive for passing to convex combinations. This understood, we also have that $\sigma_\nu$ is an upper gradient of $\langle \Lambda, f_\nu \rangle$, and a standard Lebesgue point argument using (3.1) (compare [Sh, Proof of Theorem 4.5]) gives that

$$|\nabla \langle \Lambda, f_\nu \rangle(x)| \leq \sigma_\nu(x) \tag{3.11}$$

for almost every $x \in Q_0$. Thus (3.10) follows by passing to a limit in (3.11) as $\nu \to \infty$.

The fact that $\sigma$ is an $n$-weak upper gradient of $f$ follows from [HKST, Proof of (3') $\Rightarrow$ (1) in Theorem 3.17], while inequality (3.9) is, with all the proven facts, obvious.

Finally, if $f_\nu \to f$ $n$-quasieverywhere, then also $\langle \Lambda, f_\nu \rangle \to \langle \Lambda, f \rangle$ $n$-quasieverywhere for each given $\Lambda \in V^*$, and by strengthening the above argument by the nice trick of Fuglede [Fu, Theorem 3 (f)], we have that $\langle \Lambda, f \rangle \in L^{1,n}(Q_0)$ automatically. More precisely, after passing to convex combinations, we can extract from the sequence $(\sigma_\nu)$ a subsequence such that the line integral of $\sigma_\nu$ converges to the line integral of $\sigma$ outside a family of curves of $n$-modulus zero. By using the $n$-quasieverywhere convergence, and the fact that the $n$-modulus of all curves that pass through a fixed set of $n$-capacity zero has $n$-modulus zero, we hence have that the upper gradient inequality holds for the pair $\langle \Lambda, f \rangle$ and $\sigma$ for curves outside a family of $n$-modulus zero. It follows that $\sigma$ is an $n$-weak upper gradient of $\langle \Lambda, f \rangle$. (See [Sh, Section 3] or [KSh, Lemma 3.1] for more details here.) The second claim in the proposition now follows from [HKST, Theorem 3.17].

The proposition is proved. □

## 4. Statement and proof of the main result

The goal of this section is to state and prove Theorem 4.1, which is a generalization of Theorem 1.1 for general conformal densities.

Statements (i)–(iii) in Theorem 1.1 have obvious counterparts for general conformal densities. Condition (iv) can be formulated by the aid of the general definition for $\mathrm{Osc}^{n,\infty}(\mathbb{R}^n; V)$ given in (1.13), and condition (v) uses the Sobolev-Dirichlet space described in Section 3.

We make one more remark and convention before formulating Theorem 4.1. Given a conformal density $\varrho$ in $\mathbb{R}^{n+1}_+$, the metric space $(\mathbb{R}^{n+1}_+, d_\varrho)$ is separable and hence can be embedded isometrically in the Banach space $\ell^\infty$. An explicit embedding is given by

$$z \mapsto \big(d_\varrho(z, z_1) - d_\varrho(e_{n+1}, z_1), d_\varrho(z, z_2) - d_\varrho(e_{n+1}, z_2), \dots\big), \tag{4.1}$$

where $e_{n+1} = (0, 1) \in \mathbb{R}^{n+1}_+$ and $\{z_\nu : \nu \in \mathbb{N}\}$ is a fixed dense subset of $\mathbb{R}^{n+1}_+$. Embedding (4.1) entails some choices, but we tacitly assume from now on that the choices have been made independently of the conformal density $\varrho$. Thus, in particular, we can always think of $(\mathbb{R}^{n+1}_+, d_\varrho)$, as well as its metric completion, as being a subset of $\ell^\infty$.

As explained in Section 2, it was proved in [BKR, Theorem 4.4] that the identity map $I_\varrho$, defined in (2.10), has an $(n+1)$-capacity almost everywhere defined (radial) extension to $\mathbb{R}^n$,

$$i_\varrho \colon \mathbb{R}^n \to \ell^\infty, \qquad i_\varrho(x) = \lim_{t \to 0} I_\varrho(x, t). \tag{4.2}$$

The mapping $i_\varrho$ is measurable and essentially separably valued because it is almost everywhere a pointwise limit of continuous mappings

$$I_\varrho(\cdot,t)\colon \mathbb{R}^n \to \ell^\infty, \qquad x \mapsto I_\varrho(x,t), \qquad t>0, \tag{4.3}$$

whose images are separable.

**Theorem 4.1.** *Let $\varrho$ be a conformal density in $\mathbb{R}^{n+1}_+$, $n \geq 2$. Then the following five conditions are equivalent:*

(i) $\varrho^*(\cdot) := \sup_{t>0} \varrho(\cdot,t) \in L^n(\mathbb{R}^n)$,

(ii) $\|\varrho\|_{QH^n(\mathbb{R}^n)} := \Big(\sup_{t>0} \int_{\mathbb{R}^n} \varrho(x,t)^n \, dm_n(x)\Big)^{1/n} < \infty$,

(iii) $\sup_{\lambda>0} \lambda \Big(\#\{Q \in \mathcal{D}(\mathbb{R}^n) : r_\varrho(\widehat{Q}) > \lambda\}\Big)^{1/n} < \infty$,

(iv) $i_\varrho \in \mathrm{Osc}^{n,\infty}(\mathbb{R}^n;\ell^\infty)$,

(v) $i_\varrho \in L^{1,n}(\mathbb{R}^n;\ell^\infty)$.

*Moreover, the various Lebesgue type norms appearing in conditions* (i)–(v) *are equivalent with multiplicative constants only depending on the data of $\varrho$.*

Note that (v) has equivalent formulations given in Proposition 3.2. At this point, (v) should be understood so that the almost everywhere defined function $i_\varrho$ has an $n$-quasicontinuous representative in $L^{1,n}(\mathbb{R}^n;\ell^\infty)$. We shall later show (in Section 5) that the $(n+1)$-capacity everywhere defined trace in (4.2) is this representative.

**Problem 4.2.** Let $\varrho$ be a conformal density in $\mathbb{R}^{n+1}_+$, $n \geq 2$, satisfying any of the equivalent conditions (i)–(v) in Theorem 4.1. Is it then true that

$$m_n\big(\{x \in \mathbb{R}^n : \liminf_{t\to 0} \varrho(x,t) = 0\}\big) = 0 \ ? \tag{4.4}$$

As explained earlier, an affirmative answer to Problem 4.2 yields an affirmative answer to Problem 1.3.

*Remark* 4.3. In view of the discussion in Section 2, the equivalence of (i)–(iii) in Theorem 1.1 follows immediately from Theorem 4.1. Strictly speaking, this is not the case for the assertions (iv) and (v), for the target space for the functions in Theorem 4.1 is $\ell^\infty$ rather than $\mathbb{R}^{n+1}$. Moreover, the isometric embedding in $\ell^\infty$ uses the inner metric in the image domain rather than the Euclidean metric.

To handle this discrepancy, one should simply follow the ensuing proof of Theorem 4.1 and observe that the arguments go over wholesale in the case of a quasiconformal mapping with boundary values in $\mathbb{R}^{n+1}$. The details even simplify here because the classical Sobolev space theory applies. One can also derive the implication "(v) in Theorem 4.1 $\Rightarrow$ (v) in Theorem 1.1" more formally as follows: Let $\varrho = a_F$ for a quasiconformal mapping $F$ as in Theorem 1.1, and assume that $(\mathbb{R}^{n+1}_+, d_\varrho)$ has been isometrically embedded in $\ell^\infty$ as described in the beginning of this section. Then define a map

$$\pi\colon i_\varrho(\mathbb{R}^n) \to \mathbb{R}^{n+1} \tag{4.5}$$

by

$$y \mapsto \pi(y) := f(x)\,,$$

if $x \in \mathbb{R}^n$ is a point such that $y = i_\varrho(x)$. The map $\pi$ in (4.5) is indeed well defined, for if $x$ and $x'$ are two distinct points on $\mathbb{R}^n$, then the distance between $i_\varrho(x)$ and $i_\varrho(x')$ in $\ell^\infty$ is positive by the Gehring-Hayman Theorem 2.2—see (2.13). Moreover, the map $\pi$ is easily seen to be Lipschitz, with constant depending only on the data of the problem. Consequently,

$$f = \pi \circ i_\varrho$$

is a function in the Dirichlet-Sobolev space $L^{1,n}(\mathbb{R}^n;\mathbb{R}^{n+1})$, as follows directly from the definitions. Finally note that (iv) and (v) are equivalent by Proposition 4.7, irrespectively of the situation at hand.

We now turn to the proof of Theorem 4.1.

The implication (i) $\Rightarrow$ (ii) in Theorem 4.1 is of course trivial with the estimate

$$\|\varrho\|_{QH^n(\mathbb{R}^n)} \le \|\varrho^*\|_{L^n(\mathbb{R}^n)}.$$

Next, consider the implication (ii) $\Rightarrow$ (iii). In the proof, all constants in the relations $\lesssim$ and $\simeq$ depend only on the data of $\varrho$. To begin with, it follows from the sub-mean value property (2.16) in Proposition 2.3 that

$$\varrho(x,t) \lesssim \frac{\|\varrho\|_{QH^n(\mathbb{R}^n)}}{t} \tag{4.6}$$

whenever $t > 0$. If we define $r_\varrho$ as in (2.7), then (4.6) implies

$$\sup_{Q\in\mathcal{D}(\mathbb{R}^n)} r_\varrho(\widehat{Q}) \lesssim \|\varrho\|_{QH^n(\mathbb{R}^n)}. \tag{4.7}$$

To ease notation, we perform a scaling and assume that the supremum on the left hand side of (4.7) is bounded by one. We then claim that the following growth estimate for the number of Whitney cubes,

$$\#_k := \#\big\{Q \in \mathcal{D}(\mathbb{R}^n) : r_\varrho(\widehat{Q}) \in (2^{-k}, 2^{-k+1}]\big\} \lesssim 2^{kn} \tag{4.8}$$

for $k \in \mathbb{N}$, is enough to prove (iii). Indeed, it follows from (4.8) that

$$\#\{Q \in \mathcal{D}(\mathbb{R}^n) : r_\varrho(\widehat{Q}) > \lambda\} \le \sum_{k=1}^{m} \#_k \lesssim 2^{mn} \simeq \frac{1}{\lambda^n}, \tag{4.9}$$

where $m$ is such that $\lambda \in (2^{-m}, 2^{-m+1}]$. It therefore suffices to prove (4.8) under the normalization $\|\varrho\|_{QH^n} \simeq 1$.

To this end, fix $k \ge 1$, and let $\widehat{Q}_1, \dots, \widehat{Q}_N$ be Whitney cubes such that

$$r_\varrho(\widehat{Q}_i) \in (2^{-k}, 2^{-k+1}], \quad i = 1, \dots, N.$$

Then fix a positive number $t$ with

$$t \ll \min_{i=1,\dots,N} \{\operatorname{diam}(\widehat{Q}_i)\},$$

and set, for each $i = 1, \dots, N$,

$$S_t(\widehat{Q}_i) = \{(x,t) \in \mathbb{R}^{n+1}_+ : (x,s) \in \widehat{Q}_i \quad \text{for some } s > t\}.$$

Thus $S_t(\widehat{Q}_i)$ is the shadow of the cube $\widehat{Q}_i$ on $\mathbb{R}^n \times \{t\}$.

For the following lemma, fix a cube $\widehat{Q} \in \{\widehat{Q}_1, \dots, \widehat{Q}_N\}$.

**Lemma 4.4.** *For each $\varepsilon > 0$ there exists $\lambda \ge 1$, depending on the data of $\varrho$, and a set $T_t(\widehat{Q}) \subset S_t(\widehat{Q})$ such that*

$$\mathcal{H}^\infty_{n-1}\big(S_t(\widehat{Q}) \setminus T_t(\widehat{Q})\big) < \varepsilon(\operatorname{diam}(\widehat{Q}))^{n-1} \tag{4.10}$$

*and that*

$$T_t(\widehat{Q}) \subset \{y \in \mathbb{R}^{n+1}_+ : \operatorname{dist}_\varrho(y, \widehat{Q}) \le \lambda \operatorname{diam}_\varrho(\widehat{Q})\} =: N_\varrho(\widehat{Q}, \lambda). \tag{4.11}$$

Recall that $\mathcal{H}^\infty_\alpha$ denotes the Hausdorff $\alpha$-content, $\alpha > 0$. In order to prove Lemma 4.4, we require the following standard modulus estimate [KR, Proposition 4.3], [BK, Proposition 5.1]:

**Lemma 4.5.** *Let $\widehat{Q} \in \mathcal{W}$ be a cube, let $0 \le t < \operatorname{diam}(\widehat{Q})$, and let $S_t(\widehat{Q})$ be the shadow of $\widehat{Q}$ on $\mathbb{R}^n \times \{t\}$. If $E \subset S_t(\widehat{Q})$ and if $\Gamma$ is a family of curves joining $\widehat{Q}$ and $E$ in $\mathbb{R}^{n+1}_+$, then*

$$\frac{\mathcal{H}^\infty_\alpha(E)}{(\operatorname{diam}(\widehat{Q}))^\alpha} \le C \operatorname{mod}_{n+1}(\Gamma), \tag{4.12}$$

*where $C = C(n, \alpha) \ge 1$.*

*Proof of Lemma 4.4.* The proof is a standard application of estimate (4.12) with $\alpha = n-1$. Analogous results appear in several places in the literature, for example in [KR, Proposition 4.3] and in [BK, Section 5]. For convenience, we repeat the argument here.

For $\lambda > 1$ denote by $E_\lambda$ the set of all points in $S_t(\widehat{Q})$ that lie outside the neighborhood $N_\varrho(\widehat{Q}, \lambda)$ of $\widehat{Q}$ with respect to metric $d_\varrho$ as defined in (4.11). If $\Gamma = \Gamma(\widehat{Q}, t; \lambda)$ is the family of all rectifiable curves $\gamma$ joining $\widehat{Q}$ to $E_\lambda$ in $\mathbb{R}^{n+1}_+$, then every curve $\gamma \in \Gamma$ leaves $N_\varrho(\widehat{Q}, \lambda)$. Then by the basic modulus estimate [BKR, Lemma 3.2], we have that

$$\operatorname{mod}_{n+1}(\Gamma) \leq C(\log \lambda)^{-n} \tag{4.13}$$

for $\lambda \geq 4$. On the other hand, (4.12) shows that

$$\frac{\mathcal{H}^\infty_{n-1}(E_\lambda)}{(\operatorname{diam}(\widehat{Q}))^{n-1}} \leq C \operatorname{mod}_{n+1}(\Gamma). \tag{4.14}$$

The lemma follows from estimates (4.13) and (4.14). $\square$

Next, choose $\varepsilon = \frac{1}{2}$ in Lemma 4.4. Then the $n$-modulus in $S_t(\widehat{Q})$ of the family $\Gamma(\widehat{Q}, t)$ of all curves inside $T_t(\widehat{Q})$ joining two opposite faces of the cube $S_t(\widehat{Q})$ satisfies

$$\operatorname{mod}_n(\Gamma(\widehat{Q}, t)) \geq \frac{1}{2} > 0. \tag{4.15}$$

This is simply because at least half of the straight line segments joining the faces necessarily miss $S_t(\widehat{Q}) \setminus T_t(\widehat{Q})$, as follows from (4.10) and from the fact that Hausdorff content decreases in projections. The assumption $n \geq 2$ is important here.

By observing (4.15), Proposition 2.3 implies that

$$r_\varrho(\widehat{Q})^n \lesssim \sum_{R \in \mathcal{T}_t(Q)} r_\varrho(\widehat{R})^n, \tag{4.16}$$

whenever $\mathcal{T}_t(Q)$ is any finite collection of cubes $R \subset Q$ in $\mathbb{R}^n$ whose union covers the projection of $T_t(\widehat{Q})$ in $\mathbb{R}^n$; that is, we require in (4.16) that

$$\big\{x \in \mathbb{R}^n : (x,t) \in T_t(\widehat{Q})\big\} \subset \bigcup_{R \in \mathcal{T}_t(Q)} R, \tag{4.17}$$

where the union is assumed finite. Recalling that $r_\varrho(\widehat{Q}) \in (2^{-k}, 2^{-k+1}]$, we now obtain from (4.16) that

$$2^{-kn} \simeq r_\varrho(\widehat{Q})^n \lesssim \sum_{R \in \mathcal{T}_t(Q)} r_\varrho(\widehat{R})^n. \tag{4.18}$$

Next we note that there is only a fixed amount of overlap among the neighborhoods $N_\varrho(\widehat{Q}_1, \lambda), \ldots, N_\varrho(\widehat{Q}_N, \lambda)$, where the notation is as in Lemma 4.4 and $\lambda \geq 1$ corresponds to the value $\varepsilon = \frac{1}{2}$. Indeed, because

$$\operatorname{diam}_\varrho(\widehat{Q}_i) \simeq r_\varrho(\widehat{Q}_i) \simeq 2^{-k}$$

for each $i = 1, \dots, N$, no point in $\mathbb{R}^{n+1}_+$ can belong to more than $C$ of the sets $N_\varrho(\widehat{Q}_i, \lambda)$ by a simple packing argument based on (2.4) and (2.5), where $C > 0$ only depends on the data. From this and from the way the sets $T_t(\widehat{Q}_i)$ were constructed in the proof of Lemma 4.4, we see that no point $z = (x, t)$ can belong to more than $C$ of the sets $T_t(\widehat{Q}_i)$, where again $C > 0$ only depends on the data. Because the sets $T_t(\widehat{Q}_i)$ are closed and there are only finitely many of them, there exists $\delta > 0$ such that no set of diameter less than $\delta$ can meet more than $C$ of the sets $T_t(\widehat{Q}_i)$. Now choose a finite covering $\mathcal{T}_t(Q_i)$ of the projection to $\mathbb{R}^n$ of the set $T_t(\widehat{Q}_i)$ by cubes of a fixed diameter $t' < \min\{t, \delta\}$. Then no point $x \in \mathbb{R}^n$ belongs to more than a fixed number of cubes from all the coverings $\mathcal{T}_t(Q_i)$, $i = 1, \dots, N$. Clearly, this fixed number depends only on the data of $\varrho$. It therefore follows from (4.18) that

$$N\, 2^{-kn} \lesssim \sum_{i=1}^{N} \sum_{R \in \mathcal{T}_t(Q_i)} r_\varrho(\widehat{R})^n \lesssim \int_{\mathbb{R}^n} \varrho(x, t')^n \, dm_n(x) \lesssim 1,$$

where our normalization $\|\varrho\|_{QH^n} \simeq 1$ was also used.

This proves the growth estimate (4.8), and thereby the implication (ii) $\Rightarrow$ (iii) in Theorem 1.1. The proof shows that the quantity in (iii) is bounded by a constant (depending only on the data of $\varrho$) times the quantity in (ii).

Next we prove the implication (iii) $\Rightarrow$ (iv) in Theorem 1.1. First notice that (iii) gives

$$\sup_{w \in \mathbb{R}^n} \sup_{\lambda > 0} \lambda \left( \#\{Q \in \mathcal{D}(\mathbb{R}^n) : r_\varrho(\widehat{Q - w}) > \lambda\} \right)^{1/n} < \infty,$$

as is easily seen by using the Harnack inequality (2.4). We shall show that the norm $\|i_\varrho\|_{\mathrm{Osc}^{n,\infty}}$ is bounded by a constant $C(n, K)$ times this quantity. Here we follow the ideas from [RS1], [RS2].

It suffices to consider a fixed translation of the dyadic cubes, and without loss of generality we may consider the dyadic cubes in $\mathcal{D}(\mathbb{R}^n)$ themselves. Thus, fix a cube $Q \in \mathcal{D}(\mathbb{R}^n)$ and let $z_Q = (x_Q, t_Q)$ be the center of the cube $\widehat{Q} \in \mathbb{R}^{n+1}_+$. Then

(4.19)
$$\begin{aligned}
&\fint_Q |i_\varrho(x) - (i_\varrho)_Q| \, dm_n(x) \lesssim \fint_Q |i_\varrho(x) - I_\varrho(z_Q)| \, dm_n(x) \\
&\le \fint_Q |i_\varrho(x) - I_\varrho(x, t_Q)| \, dm_n(x) + \fint_Q |I_\varrho(x, t_Q) - I_\varrho(z_Q)| \, dm_n(x) \\
&\lesssim \fint_Q \int_0^{t_Q} \varrho(x,s) \, ds \, dm_n(x) + r_\varrho(\widehat{Q}).
\end{aligned}$$

In order to estimate the last integral in the preceding expression we introduce a "maximal function" $\mathcal{M}(\phi)\colon \mathcal{D}(\mathbb{R}^n) \to \mathbb{R}_+$ associated with a function $\phi\colon \mathcal{D}(\mathbb{R}^n) \to \mathbb{R}_+$ as follows

$$\mathcal{M}(\phi)(Q) := \frac{1}{m_n(Q)} \sum_{\substack{R \subset Q \\ R \in \mathcal{D}(\mathbb{R}^n)}} \phi(R) \, m_n(R). \tag{4.20}$$

It then follows from (4.19) that the mean oscillation function $Q \mapsto A_{i_\varrho}(Q)$ is bounded from above by a constant $C(n, K)$ times the maximal function defined in (4.20) for the function $Q \mapsto r_\varrho(\widehat{Q})$:

$$A_{i_\varrho}(Q) \lesssim \mathcal{M}(r_\varrho)(Q) \quad \text{for} \quad Q \in \mathcal{D}(\mathbb{R}^n).$$

Note that instead of $\mathcal{M}(r_\varrho(\widehat{\cdot}))$ we here write $\mathcal{M}(r_\varrho)$ for simplicity.

It follows from Lemma 4.6 below that $\mathcal{M}$ is a bounded operator weak-$\ell^n$ $\to$ weak-$\ell^n$. Hence the mapping $Q \mapsto A_\varrho(Q)$ belongs to weak-$\ell^n$ with norm bounded by a constant times that of $r_\varrho$.

This completes the proof of the implication (iii) $\Rightarrow$ (iv) in Theorem 1.1.

The following lemma was needed in the above proof.

**Lemma 4.6.** *The operator*

$$\mathcal{M}\colon \text{weak-}\ell^n \to \text{weak-}\ell^n$$

*defined in* (4.20) *is bounded with norm depending only on* $n$.

*Proof.* The lemma is a generalized Schur lemma, and appears in [RS2, p. 269], for example. For the convenience of the reader, we indicate the proof. The following lemma appears in [W, p. 87]:

Let $(X, \mu)$ be a measure space, let $\mathcal{K}\colon X \times X \to [0, \infty)$ be measurable, and let $1 < p < \infty$. Suppose that there exists a measurable function $g\colon X \to (0, \infty)$ together with a number $M > 0$ such that

$$\int_X \mathcal{K}(x,y) g(y)^{p/(p-1)} \, d\mu(y) \le M g(x)^{p/(p-1)}$$

for all $x \in X$ and

$$\int_X \mathcal{K}(x,y) g(x)^p \, d\mu(x) \leq M g(y)^p$$

for all $y \in X$. Then the operator

$$Tf(x) = \int_X \mathcal{K}(x,y) f(y) \, d\mu(y)$$

is bounded $T\colon L^p(X,\mu) \to L^p(X,\mu)$ with $||T|| \leq M^2$.

By applying this lemma to the spaces $\ell^p$, $1 < p < \infty$, we find that $\mathcal{M}\colon \ell^p \to \ell^p$ is bounded. Indeed, let $X = \mathcal{D}(\mathbb{R}^n)$ and let $\mu$ be the counting measure on $X$. Define $\mathcal{K}(Q,Q') = |Q'|/|Q|$ if $Q' \subset Q$, and $\mathcal{K}(Q,Q') = 0$ otherwise, and choose $g(Q) = |Q|^\alpha$, where $0 < \alpha < 1/p$. The boundedness $\mathcal{M}\colon$ weak-$\ell^n \to$ weak-$\ell^n$ follows by interpolation [SW, V.3]. This proves Lemma 4.6. □

The implication (iv) ⇒ (v) follows from the left inequality in (4.21) in the following proposition.

**Proposition 4.7.** *Let $V$ be a Banach space and let $f\colon \mathbb{R}^n \to V$ be a function in $L^1_{\mathrm{loc}}(\mathbb{R}^n; V)$. Then*

$$\frac{1}{C} \|f\|_{L^{1,n}(\mathbb{R}^n;V)} \leq \|f\|_{\mathrm{Osc}^{n,\infty}(\mathbb{R}^n;V)} \leq C\, \|f\|_{L^{1,n}(\mathbb{R}^n;V)}, \tag{4.21}$$

*where the constant $C \geq 1$ depends only on $n$.*

One should interpret Proposition 4.7 to mean that if a function $f$ belongs to $L^{1,n}(\mathbb{R}^n;V)$, then it belongs to $\mathrm{Osc}^{n,\infty}(\mathbb{R}^n;V)$ with the norm inequality on the right in (4.21), and if a function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n;V)$ belongs to $\mathrm{Osc}^{n,\infty}(\mathbb{R}^n;V)$, then it has a Lebesgue representative in $L^{1,n}(\mathbb{R}^n;V)$. Moreover, $f$ is assumed to be a priori separably valued, which is essentially automatic if $f \in L^{1,n}(\mathbb{R}^n;V)$ by Proposition 3.1.

The right inequality in (4.21) for scalar-valued functions was proved by Rochberg and Semmes [RS1, p. 228]. Their proof extends to the general case with some modifications, but as we do not need the right inequality here, we omit the proof. It is the left inequality that is required in this paper. As mentioned earlier, the left inequality for scalar-valued functions appears, in a slightly different formulation, in [CST, Appendix]. Since the argument in [CST] is somewhat sketchy, and since a few words need to be added in order to reach the general Banach space-valued statement, for the convenience of the reader we provide a detailed proof of the left inequality.

To that end, fix a nonnegative, radially symmetric bump-function $\eta \in C_0^\infty(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} \eta \, dm_n = 1$ whose support is contained in the unit ball of $\mathbb{R}^n$. For each function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n; V)$ and $\varepsilon > 0$, the convolutions

$$f_\varepsilon(x) = (f \star \eta_\varepsilon)(x) = \int_{\mathbb{R}^n} f(x-w)\,\eta_\varepsilon(w)\,dm(w), \tag{4.22}$$

are smooth functions from $\mathbb{R}^n$ into $V$. Here $\eta_\varepsilon(w) = \varepsilon^{-n}\,\eta(w/\varepsilon)$.

**Lemma 4.8.** *For each $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n; V)$ and $\varepsilon > 0$, the convolution $f_\varepsilon$ satisfies*

$$A_{f_\varepsilon}(Q) \le \int_{\mathbb{R}^n} A_f(Q-w)\,\eta_\varepsilon(w)\,dm_n(w) \tag{4.23}$$

*for each cube $Q \subset \mathbb{R}^n$.*

*Proof.* Recall the definition for $A_f(Q)$ from (1.12). Estimate (4.23) follows from a straightforward computation:

$$\begin{aligned}
A_{f_\varepsilon}(Q) &= \fint_Q |f_\varepsilon(x) - (f_\varepsilon)_Q|\,dm_n(x)\\
&= \fint_Q \left| \fint_Q \left( \int_{\mathbb{R}^n} \big(f(x-w)\; - f(y-w)\big)\eta_\varepsilon(w)\,dm_n(w)\right) dm_n(y)\right| dm_n(x)\\
&\le \int_{\mathbb{R}^n} \fint_Q \left| f(x-w) - \fint_Q f(y-w)\,dm_n(y)\right| dm_n(x)\,\eta_\varepsilon(w)\,dm_n(w)\\
&= \int_{\mathbb{R}^n} A_f(Q-w)\,\eta_\varepsilon(w)\,dm_n(w),
\end{aligned}$$

whence the lemma follows. $\square$

We wish to reduce the left inequality in (4.21) to the case when $f$ is a smooth function. To this end, we shall invoke some general interpolation theorems as in [SW, V.3]. For small $\varepsilon > 0$ consider the product measure $\# \times \nu_\varepsilon$ on $\mathcal{D}(\mathbb{R}^n) \times \mathbb{R}^n$, where

$$\begin{aligned}
\# &= \text{counting measure},\\
d\nu_\varepsilon &= \eta_\varepsilon\,dm_n.
\end{aligned}$$

Then define an operator $T$ from functions on $\mathcal{D}(\mathbb{R}^n) \times \mathbb{R}^n$ to functions on $\mathcal{D}(\mathbb{R}^n)$ by

$$Ta(Q) = \int_{\mathbb{R}^n} a(Q,w)\,d\nu_\varepsilon(w). \tag{4.24}$$

We claim that

$$T\colon L^p(\# \times \nu_\varepsilon) \to L^p(\#) = \ell^p$$

is bounded for all $1 < p < \infty$ with operator norm $\|T\| \le 1$. Indeed, because $\nu_\varepsilon(\mathbb{R}^n) = 1$, we have that

$$\begin{aligned} \|Ta\|_{\ell^p} &= \Big( \sum_{Q \in \mathcal{D}(\mathbb{R}^n)} |Ta(Q)|^p \Big)^{1/p} \\ &\le \Big( \int_{\mathbb{R}^n} \sum_{Q \in \mathcal{D}(\mathbb{R}^n)} |a(Q,w)|^p \, d\nu_\varepsilon(w) \Big)^{1/p} = \|a\|_{L^p(\# \times \nu_\varepsilon)}. \end{aligned}$$

Consequently, by interpolation [SW, Theorem 3.15, p. 197], we conclude that

$$T \colon \text{weak-}L^p(\# \times \nu_\varepsilon) \to \text{weak-}\ell^p$$

is bounded for each $1 < p < \infty$. Indeed, we obtain the inequality

$$\|Ta\|_{\text{weak-}\ell^n} \le B \, \|a\|_{\text{weak-}L^n(\# \times \nu_\varepsilon)},$$

where $B = B(n) \ge 1$ is a constant depending only on $n$ (see [SW, p. 200]).

We apply the last inequality for the function $a$ defined by

$$a(Q,w) := A_f(Q - w).$$

Then

$$\begin{aligned} \|a\|_{\text{weak-}L^n(\# \times \nu_\varepsilon)} &= \sup_{\lambda > 0} \lambda \big( (\# \times \nu_\varepsilon)(\{(Q,w) : A_f(Q-w) > \lambda\}) \big)^{1/n} \\ &\le \|f\|_{\mathrm{Osc}^{n,\infty}}. \end{aligned}$$

Since $A_{f_\varepsilon}(Q) \le Ta(Q)$ by Lemma 4.8, and because translations commute with convolutions, we obtain from the preceding discussion that

$$\|f_\varepsilon\|_{\mathrm{Osc}^{n,\infty}} \le B(n) \, \|f\|_{\mathrm{Osc}^{n,\infty}} \tag{4.25}$$

whenever $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n; V)$ and $\varepsilon > 0$.

**Lemma 4.9.** *There exists a constant $C = C(n) \ge 1$ such that*

$$\| \, |Dg| \, \|_{L^n(\mathbb{R}^n)} \le C \, \|g\|_{\mathrm{Osc}^{n,\infty}(\mathbb{R}^n;V)} \tag{4.26}$$

*whenever $g \colon \mathbb{R}^n \to V$ is a smooth function.*

The (operator) norm of the differential $|Dg(x)|$ of a smooth function $g \colon \mathbb{R}^n \to V$ is a continuous upper gradient of $g$. All these are well known facts for scalar-valued functions $g$, and for the general Banach space-valued function, see [Fe, pp. 209, 211, 347]. (Note that Federer assumes separability for the target Banach space, but all the claims here are valid in general, because smooth functions are separably valued. The proofs in [Fe] apply in this case.)

Before we prove Lemma 4.9, we point out how the left inequality in Proposition 4.7 follows. If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n; V)$, then we have seen that

(4.25) is valid for $\varepsilon > 0$. Moreover, since $f_\varepsilon$ is smooth, we can apply Lemma 4.9 and conclude that

$$\| \, |Df_\varepsilon| \, \|_{L^n(\mathbb{R}^n)} \leq C(n) \|f\|_{\mathrm{Osc}^{n,\infty}(\mathbb{R}^n;V)} \tag{4.27}$$

for $\varepsilon > 0$. Now $|Df_\varepsilon|$ is an upper gradient for $f_\varepsilon$, $\varepsilon > 0$, and $f_\varepsilon$ converges pointwise almost everywhere to $f$ if $\varepsilon \to 0$. Therefore, by (4.27) we can apply Proposition 3.3 to the our functions $f$ and $f_\varepsilon$. This shows that $f$ has an $n$-weak upper gradient $\sigma$ such that

$$\|\sigma\|_{L^n(\mathbb{R}^n)} \leq C(n) \|f\|_{\mathrm{Osc}^{n,\infty}(\mathbb{R}^n;V)} \tag{4.28}$$

and so the left inequality in Proposition 4.7 follows.

*Proof of Lemma 4.9.* The proof here is essentially that given in [CST, pp. 679–680].

Fix a cube $Q_0 \in \mathcal{D}(\mathbb{R}^n)$. It suffices to show that

$$\int_{M_s} |Dg|^n \, dm_n \leq C(n) \, \|g\|^n_{\mathrm{Osc}^{n,\infty}(\mathbb{R}^n;V)} \tag{4.29}$$

for each $s > 0$, where $M_s := \{x \in Q_0 : |Dg(x)| > s\}$. Thus, fix $s > 0$, and let

$$\mathcal{Q}_s := \big\{Q \in \mathcal{D}(\mathbb{R}^n) : Q \subset Q_0, \, \ell(Q) = 2^m, \text{ and } Q \cap M_s \neq \emptyset\big\},$$

where $m \in \mathbb{Z}$ is chosen so small that

$$|Dg(x) - Dg(y)| << s \tag{4.30}$$

whenever $x, y \in Q$ and $Q \in \mathcal{Q}_s$. In particular we have

$$\frac{1}{2} \, |Dg(x)| \leq |Dg(y)| \leq 2 \, |Dg(x)| \tag{4.31}$$

whenever $x, y \in Q$ and $Q \in \mathcal{Q}_s$. Note that

$$M_s \subset \bigcup_{Q \in \mathcal{Q}_s} Q \subset Q_0.$$

If $R \subset Q_0$ is a dyadic cube we write

$$|Dg(R)| = \max_{x \in R} |Dg(x)|.$$

Then a routine computation based on (4.30) shows that there is constant $b = b(n) \geq 1$ such that

$$\frac{1}{b} \, \ell(R) \, |Dg(R)| \leq A_g(R) \leq b \, \ell(R) \, |Dg(R)| \tag{4.32}$$

if $R \subset Q \in \mathcal{Q}_s$. Now fix $0 < \lambda < s$ such that

$$\lambda < A_g(Q) \tag{4.33}$$

for all $Q \in \mathcal{Q}_s$. Then for $Q \in \mathcal{Q}_s$ we have that

$$\lambda < b\,\ell(Q)\,|Dg(Q)|,$$

and hence there is a unique integer $k_Q \geq 1$ such that

$$b\,2^{-k_Q}\,\ell(Q)\,|Dg(Q)| \leq \lambda < b\,2^{-k_Q+1}\,\ell(Q)\,|Dg(Q)|. \tag{4.34}$$

Thus (4.31) and (4.34) give that

$$\begin{aligned}\int_{M_s} |Dg|^n\,dm_n &\leq \sum_{Q\in\mathcal{Q}_s} \sum_{\substack{R\subset Q\\ \ell(R)=2^{-k_Q}\ell(Q)}} \int_R |Dg|^n\,dm_n \\ &\simeq \sum_{Q\in\mathcal{Q}_s} \sum_{\substack{R\subset Q\\ \ell(R)=2^{-k_Q}\ell(Q)}} \ell(R)^n\,|Dg(R)|^n \\ &\simeq \sum_{Q\in\mathcal{Q}_s} 2^{k_Q n}\,2^{-k_Q n}\,\ell(Q)^n\,|Dg(Q)|^n \simeq \sum_{Q\in\mathcal{Q}_s} 2^{k_Q n}\,\lambda^n.\end{aligned}$$

In the above sums it is understood that $R$ is a dyadic cube.

On the other hand, $A_g(R) \gtrsim \lambda$ for each cube $R \subset Q$ with $\ell(R) = 2^{-k_Q}\,\ell(Q)$, if $Q \in \mathcal{Q}_s$. It follows that the last sum above is bounded by a dimensional constant times $\|g\|^n_{\mathrm{Osc}^{n,\infty}}$. Hence we obtain (4.29), and the proof of the lemma is complete. □

We have now completed the proof of the implication (iv) $\Rightarrow$ (v) in Theorem 4.1.

It remains to prove the implication (v) $\Rightarrow$ (i). We first require a lemma. The lemma is valid for arbitrary conformal densities $\varrho$.

**Lemma 4.10.** *Suppose $\varrho : \mathbb{R}^{n+1}_+ \to (0,\infty)$ is a conformal density. For each cube $Q \subset \mathbb{R}^n$ there exists a set $E \subset Q$ such that $m_n(E) \geq \frac{1}{2}\,m_n(Q)$ and that*

$$|i_\varrho(x) - (i_\varrho)_Q| \geq \delta r_\varrho(\widehat{Q}) \tag{4.35}$$

*for $x \in E$, where $\delta > 0$ depends only on the data of $\varrho$.*

Recall that $r_\varrho$ is defined in (2.7). Here and below $|\cdot|$ denotes the norm in $\ell^\infty$.

*Proof.* Note that $r_\varrho(\widehat{Q}) \simeq \operatorname{diam}_\varrho(\widehat{Q})$. Then fix $\varepsilon > 0$ and let $G_\varepsilon \subset Q$ consist of those points $x$ for which $|i_\varrho(x) - (i_\varrho)_Q| \leq \varepsilon \operatorname{diam}_\varrho(\widehat{Q})$. Let $\Gamma_\varepsilon$ denote the family of straight line segments between the points $x$ and $(x, \ell(Q))$ for $x \in G_\varepsilon$. Then

$$\operatorname{mod}_{n+1}(\Gamma_\varepsilon) = \frac{m_n(G_\varepsilon)}{m_n(Q)} \tag{4.36}$$

by [V1, 7.2]. On the other hand,

$$
\begin{aligned}
\operatorname{mod}_{n+1}(\Gamma_\varepsilon) &\le C_1\left(\log\left(\frac{\operatorname{dist}_\varrho(\widehat{Q}, G_\varepsilon)}{\varepsilon \operatorname{diam}_\varrho(\widehat{Q})}\right)\right)^{-n} \\
&\le C_1\big(\log(C_2/\varepsilon)\big)^{-n}
\end{aligned}
\tag{4.37}
$$

by the modulus estimate [BKR, Lemma 3.2] and by (2.4). Here the constants $C_1, C_2 > 0$ depend only on the data of $\varrho$. By combining (4.36) and (4.37), choosing $\varepsilon > 0$ judiciously, and letting $E = Q \setminus G_\varepsilon$, we obtain the required set. The lemma follows. □

Note that the vector $(i_\varrho)_Q$ can be replaced by any other vector in $\ell^\infty$ in the statement and proof of the preceding lemma.

To finish the proof of the implication (v) ⇒ (i), fix a cube $Q \subset \mathbb{R}^n$, and let $E \subset Q$ be as in Lemma 4.10. By [HKST, Theorem 6.2], the following Poincaré inequality holds for $\ell^\infty$-valued Sobolev function:

$$
⨍_Q |i_\varrho(x) - (i_\varrho)_Q| \, dm_n(x) \le C \operatorname{diam}(Q) ⨍_{\tau Q} \sigma(x)\, dm_n(x) \tag{4.38}
$$

for each upper gradient $\sigma$ of $i_\varrho$, where $C = C(n) \ge 1$ and $\tau = \tau(n) \ge 1$, and $\tau Q$ is the cube with same center as $Q$ but $\ell(\tau Q) = \tau \ell(Q)$. From this and (4.35) we obtain

$$
\begin{aligned}
r_\varrho(\widehat{Q}) &\lesssim ⨍_E |i_\varrho(x) - (i_\varrho)_Q| \, dm_n(x) \lesssim ⨍_Q |i_\varrho(x) - (i_\varrho)_Q| \, dm_n(x) \\
&\lesssim \operatorname{diam}(Q) ⨍_{\tau Q} \sigma(x)\, dm_n(x),
\end{aligned}
\tag{4.39}
$$

whence

$$
\varrho(\widehat{Q}) \lesssim ⨍_{\tau Q} \sigma(x)\, dm_n(x). \tag{4.40}
$$

This implies that

$$
\varrho^*(x) \lesssim (M\sigma)(x), \quad x \in \mathbb{R}^n,
$$

where $M$ denotes the Hardy-Littlewood maximal operator [St, Chapter 1]. Because $M$ maps $L^n(\mathbb{R}^n)$ boundedly to $L^n(\mathbb{R}^n)$ (recall that $n > 1$), the implication (v) ⇒ (i) follows. Moreover, the implication is quantitative in that

$$
\|\varrho^*\|_{L^n(\mathbb{R}^n)} \le C(n)\, \|\sigma\|_{L^n(\mathbb{R}^n)}.
$$

We have thereby completely proved Theorem 4.1.

## 5. Quasieverywhere defined trace

As already mentioned in the introduction, a quasiconformal map $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$ has radially defined trace $f$ as in (1.5) outside an exceptional set $E_F \subset \mathbb{R}^n$ of zero $(n+1)$-capacity in $\mathbb{R}^{n+1}$. It turns out that a slightly better result is true for mappings in the Riesz class (cf. discussion after Corollary 5.2).

**Theorem 5.1.** *Let $\varrho$ be a conformal density in $\mathbb{R}^{n+1}_+$, $n \geq 2$, such that*

$$\sup_{t>0} \int_{\mathbb{R}^n} \varrho(x,t)^n \, dm_n(x) < \infty. \tag{5.1}$$

*Then there is a Borel set $E_\varrho \subset \mathbb{R}^n$ of zero $n$-capacity in $\mathbb{R}^n$ such that*

$$\int_0^1 \varrho(x,s)\, ds < \infty \quad \text{for all} \quad x \in \mathbb{R}^n \backslash E_\varrho. \tag{5.2}$$

*Moreover, if*

$$\bar{\varrho}(x) = \limsup_{t\to 0} \varrho(x,t) \tag{5.3}$$

*for $x \in \mathbb{R}^n$, then $\bar{\varrho}$ is an $n$-weak upper gradient of the boundary mapping $i_\varrho\colon \mathbb{R}^n \to \ell^\infty$, and $\|\bar{\varrho}\|_{L^n(\mathbb{R}^n)} \simeq \|i_\varrho\|_{\mathrm{Osc}^{n,\infty}(\mathbb{R}^n;\ell^\infty)}$ with constants only depending on the data of $\varrho$.*

Recall that we assume that an isometric embedding $(\mathbb{R}^{n+1}_+, d_\varrho) \to \ell^\infty$ has been chosen; see (4.1).

*Proof.* Fix an arbitrary cube $Q_1 \subset \mathbb{R}^n$, and let

$$E_0 := \left\{ x \in Q_1 : \int_0^1 \varrho(x,s)\, ds = \infty \right\}.$$

It is easy to see that $E_0$ is a Borel set. We claim that $\mathrm{cap}_n(E_0) = 0$, where $\mathrm{cap}_n$ denotes $n$-capacity in $\mathbb{R}^n$ as defined in (3.3). This will be enough to show the first part of the theorem.

We may assume that $Q_1$ is the unit cube in $\mathbb{R}^n$. Fix $k \in \mathbb{N}$. If $\mathrm{cap}_n(E_0) > 0$, then there exists $\delta > 0$, independent of $k$ (see [HKM, Theorem 2.2 (v)]), such that for some $t_k > 0$ we have that $\mathrm{cap}_n(E_k) > \delta$, where

$$E_k := \left\{ x \in Q_1 : \int_{t_k}^1 \varrho(x,s)\, ds \geq k \right\}.$$

Fix a cube $Q_2$ in $\mathbb{R}^n$ with $\ell(Q_2) = \ell(Q_1) = 1$ and $\mathrm{dist}(Q_1, Q_2) = 1$, and consider the family $\Gamma_k$ of curves in $\mathbb{R}^n$ connecting $E_k \subset Q_1$ and $Q_2$. Then $\mathrm{cap}_n(E_t) > \delta$ implies

$$\mathrm{mod}_n(\Gamma_k) > \delta' > 0, \tag{5.4}$$

where $\delta'$ is independent of $k$ (see e.g. [Zie]).

On the other hand, consider an arbitrary locally rectifiable curve $\gamma \in \Gamma_k$, and suppose $x_1 \in E_k \subset Q_1$ and $x_2 \in Q_2$ are the end points of $\gamma$. By the Gehring-Hayman theorem (cf. Proposition 2.2), up to a fixed constant depending only on the data of $\varrho$, the hyperbolic geodesic $\alpha$ joining $(x_1, t_k)$ and $(x_2, t_k)$ has the smallest $\varrho$-length $\ell_\varrho(\alpha)$ among all curves in $\mathbb{R}^{n+1}_+$ with the same end points. By choice of $Q_1$ and $Q_2$, the hyperbolic geodesic $\alpha$ has bounded hyperbolic distance to $e_{n+1}$. Therefore, by definition of $E_k$ we have that $\ell_\varrho(\alpha) \gtrsim k$. It follows that

$$\int_\gamma \varrho(\gamma(x), t_k)\,|dx| \geq Ck$$

with a constant $C > 0$ independent of $\gamma$ and $k$. This shows that the density $\sigma(x) = \frac{1}{Ck}\varrho(x, t_k)$ on $\mathbb{R}^n$ is admissible for the curve family $\Gamma_k$. Hence

$$\operatorname{mod}_n(\Gamma_k) \lesssim (1/k^n) \int_{\mathbb{R}^n} \varrho(x, t_k)^n\, dm_n(x) \lesssim 1/k^n \to 0 \quad \text{as} \quad k \to \infty,$$

where we used the assumption (5.1). This contradicts (5.4).

To prove the second part of the theorem consider the continuous mappings

$$I_t\colon \mathbb{R}^n \to \ell^\infty, \qquad x \mapsto I_t(x) = I_\varrho(x, t), \tag{5.5}$$

and the functions

$$\varrho_t\colon \mathbb{R}^n \to (0, \infty), \qquad x \mapsto \varrho_t(x) = \varrho(x, t), \tag{5.6}$$

for $0 < t < 1$. Let $\gamma\colon [0, \ell(\gamma)] \to \mathbb{R}^n$ be a rectifiable curve parametrized by the arc length, with end points $a$ and $b$. Then $\gamma_t(s) = \big(\gamma(s), t\big)$, $s \in [0, \ell(\gamma)]$, is a curve in $\mathbb{R}^n \times \{t\} \subset \mathbb{R}^{n+1}_+$ with end points $(a, t)$ and $(b, t)$. Moreover,

$$\begin{aligned}|I_t(a) - I_t(b)| = d_\varrho\big((a,t),(b,t)\big) &\leq \int_0^{\ell(\gamma)} \varrho\big(\gamma(s), t\big)\, ds \\ &= \int_0^{\ell(\gamma)} \varrho_t\big(\gamma(s)\big)\, ds,\end{aligned}$$

which implies that $\varrho_t$ is an upper gradient of $I_t$. Inequality (5.2) implies that the equality $\lim_{t\to 0} I_t(\cdot) = i_\varrho(\cdot)$ is valid $n$-quasieverywhere on $\mathbb{R}^n$, and by our hypothesis (5.1) the $L^n$-norm of the functions $\varrho_t$ is uniformly bounded independent of $\varrho$. So from Proposition 3.3 we obtain that $i_\varrho \in L^{1,n}(\mathbb{R}^n; \ell^\infty)$ (without change of the Lebesgue class of $i_\varrho$; see the remark following Theorem 4.1). In addition, any weak limit $\sigma\colon \mathbb{R}^n \to [0, \infty]$ of $\varrho_t$ as $t \to 0$ is an $n$-weak upper gradient of $i_\varrho$. By Mazur's lemma such a weak limit $\sigma$ will be the pointwise limit almost everywhere of convex combinations of the functions $\varrho_t$. This shows that

$$\sigma(x) \le \bar{\varrho}(x)$$

for almost every $x \in \mathbb{R}^n$. It is easy to see that this implies that $\int_\gamma \sigma\, ds \le \int_\gamma \bar{\varrho}\, ds$ for all curves $\gamma$ in $\mathbb{R}^n$ outside a family $\Gamma$ of curves of $n$-modulus zero (cf. [HKST, Lemma 3.23]). Hence $\bar{\varrho}$ is an $n$-weak upper gradient of $i_\varrho$.

If we define

$$\underline{\varrho}(x) := \liminf_{t \to 0} \varrho(x,t)$$

for $x \in \mathbb{R}^n$, then the same argument shows that

$$\underline{\varrho}(x) \le \sigma(x)$$

for almost every $x \in \mathbb{R}^n$, whenever $\sigma$ is a $n$-weak upper gradient of $i_\varrho$. By Proposition 2.3 there is a constant $C > 0$ depending only on the data of $\varrho$ sucht that $\bar{\varrho}(x) \le C\underline{\varrho}(x)$ for almost every $x \in \mathbb{R}^n$. Hence

$$\|i_\varrho\|_{L^{1,n}(\mathbb{R}^n;\ell^\infty)} \le \|\bar{\varrho}\|_{L^n(\mathbb{R}^n)} \lesssim \inf \|\sigma\|_{L^n(\mathbb{R}^n)} = \|i_\varrho\|_{L^{1,n}(\mathbb{R}^n;\ell^\infty)},$$

and the proof of the theorem is complete. □

**Corollary 5.2.** *Let $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$, $n \ge 2$, be a quasiconformal mapping in the Riesz class. Then there is a Borel set $E_F \subset \mathbb{R}^n$ of zero $n$-capacity in $\mathbb{R}^n$ such that the limit*

$$\lim_{t \to 0} F(x,t) = f(x) \tag{5.7}$$

*exists for each $x \in \mathbb{R}^n \setminus E_F$. The function $f$ as defined radially in (5.7) for $x \in \mathbb{R}^n \setminus E_f$ is an $n$-quasicontinuous representative of the trace $f$ in the Dirichlet-Sobolev space $L^{1,n}(\mathbb{R}^n;\mathbb{R}^{n+1})$. Moreover, the sets*

$$\{x \in \mathbb{R}^n : Df(x) = 0\} \text{ and } \{x \in \mathbb{R}^n : \limsup_{t \to 0} a_F(x,t) = 0\} \tag{5.8}$$

*are equal up to a set of Lebesgue $n$-measure zero.*

Note that the class of sets in $\mathbb{R}^n$ with vanishing $n$-capacity in $\mathbb{R}^n$ is contained in the class of sets of vanishing $(n+1)$-capacity in $\mathbb{R}^{n+1}$ (where $\mathbb{R}^n$ is considered as an $n$-dimensional hyperplane in $\mathbb{R}^{n+1}$), but the former class is strictly smaller than the latter. See e.g. [Mz, Chapter 7.2.3]. If $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$ is an arbitrary quasiconformal map, then the limit in (5.7) exists only outside a set $E_F \subset \mathbb{R}^n$ which has $(n+1)$-capacity zero in $\mathbb{R}^{n+1}$. Therefore, according to Corollary 5.2 the exceptional sets $E_F$ for quasiconformal maps $F$ in the Riesz class are *a priori* smaller than for general quasiconformal maps. On the other hand, it is not clear which exceptional sets can actually arise at all. In the plane the situation is better understood according to a result by Kaplan [K]: Every compact set $E \subset \mathbb{R}$ of vanishing logarithmic capacity (which is the same as vanishing 2-capacity in $\mathbb{R}^2$) can arise as the exceptional set of a conformal map $F\colon \mathbb{R}^2_+ \to \mathbb{R}^2$ in the sense that the limit in (5.7) exists except for $x \in E$, where it is infinite.

*Proof of Corollary 5.2.* We apply the previous theorem to the density $\varrho = a_F$. Since $\int_0^1 \varrho(x,t)\,dt < \infty$ implies the existence of the limit in (5.7) (cf. [BK, Lemma 7.4]), a set $E_F$ as required exists.

Recall the Lipschitz map $\pi$ from Remark 4.3. Then $f = \pi \circ i_\varrho$, and it follows that $f$ as defined in (5.7) is an $n$-quasicontinuous representative of the boundary trace of $F$, since $i_\varrho$ is such a representative for the boundary trace of $I_\varrho$.

To see that the sets in (5.8) agree up to a set of measure zero, note that up to a dimensional constant $|Df|$ is a minimal weak upper gradient of $f$ (cf. [Sh, Proof of Theorem 4.5]). So if $\sigma$ is any $n$-weak upper gradient of $f$, then

$$|Df(x)| \le C(n)\sigma(x)$$

for almost every $x \in \mathbb{R}^n$. By Theorem 5.1 we know that

$$\sigma = \bar{a}_F := \limsup_{t\to 0} a_F(\cdot, t)$$

is an $n$-weak upper gradient of $f$. Hence up to a set of measure zero, the set on the right side of (5.8) is contained in the set on the left side.

For the other direction we use (4.40) from the proof of the implication (v) $\Rightarrow$ (i) in Theorem 4.1. We apply this to $\varrho = a_F$, the $n$-weak upper gradient $\sigma = |Df|$ of $f$, and let the cube $Q$ shrink to a given point $x$. Then the Lebesgue differentiation implies that

$$\bar{a}_F(x) \le |Df(x)|$$

for almost every $x \in \mathbb{R}^n$. This shows that up to a set of measure zero, the set on the left side of (5.8) is contained in the set on the right side. The corollary follows. $\square$

*Remark* 5.3. If $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$, $n \ge 2$, is a quasiconformal mapping satisfying (ii) in Theorem 1.1, it is tempting to try to adapt the argument in Theorem 5.1 and prove that the boundary trace $f$ belongs to the Dirichlet-Sobolev space $L^{1,n}(\mathbb{R}^n;\mathbb{R}^{n+1})$. This approach does not work, however, for the simple reason that the maps $f_t(x) = f(x,t)$, $t > 0$, need not belong to $L^{1,n}(\mathbb{R}^n;\mathbb{R}^{n+1})$. Theorem 5.1 is another instance where the approach via general densities and Sobolev spaces comes in handy.

## 6. Two compactness results

In this section, as an application of the techniques introduced earlier in this paper, we state and prove two compactness results.

Our second compactness result (Theorem 6.2) is about general conformal densities, and it contains the first result (Theorem 6.1), which is about quasiconformal mappings, as a special case. We find it instructive to present and prove the quasiconformal case first, and then indicate the changes that are required in the general case.

**Theorem 6.1.** *Let $(F_\nu)$ be a sequence of $K$-quasiconformal mappings from $\mathbb{R}^{n+1}_+$ into $\mathbb{R}^{n+1}$, $n \geq 2$. Suppose that*

$$H := \sup_\nu \|a_{F_\nu}\|_{QH^n(\mathbb{R}^n)} < \infty, \tag{6.1}$$

*and that the sequence $(F_\nu)$ converges uniformly on compacta to a nonconstant mapping $F\colon \mathbb{R}^{n+1}_+ \to \mathbb{R}^{n+1}$.*

*Then $F$ is $K$-quasiconformal, and we have that*

$$\|a_F\|_{QH^n(\mathbb{R}^n)} \lesssim H < \infty, \tag{6.2}$$

*where the constant implicit in (6.2) depends only $n$ and $K$.*

*Moreover, for every cube $Q_0 \subset \mathbb{R}^n$, the sequence of boundary mappings $(f_\nu|Q_0)$ converges to the boundary map $f|Q_0$ weakly in the Sobolev space $W^{1,n}(Q_0;\mathbb{R}^{n+1})$.*

Recall the notational convention: $f_\nu$ and $f$ denote the traces of the maps $F_\nu$ and $F$, respectively, as defined in (1.5). Note that the limiting map $F$ is quasiconformal [V1, Section 21]. The assumption that $F$ is nonconstant can be replaced by the normalization

$$a_{F_\nu}(0,1) = 1$$

for the averaged derivative.

By using the properties of the averaged derivative, (1.8), we infer that (6.2) follows from (6.1) and from the locally uniform convergence. We thus have, by Theorem 1.1, that the traces $f_\nu$ and $f$ belong to the Dirichlet-Sobolev space $L^{1,n}(\mathbb{R}^n;\mathbb{R}^{n+1})$ with uniformly bounded norm. We shall see that the traces restricted to a fixed cube $Q_0 \subset \mathbb{R}^n$ belong to the standard Sobolev space $W^{1,n}(Q_0;\mathbb{R}^{n+1})$ with uniformly bounded norm. By the reflexivity of the Sobolev space, the sequence $(f_\nu|Q_0)$ has weak limits, and Theorem 6.1 further asserts that there is precisely one weak limit namely $f|Q_0$. This, of course, is the main point of Theorem 6.1.

Note that if $(F_\nu)$ is a locally uniformly convergent sequence of conformal maps in the upper half plane $\mathbb{R}^2_+$ with derivatives uniformly bounded in the Hardy space $H^1$, then an analogous result to Theorem 6.1 can be proved by using the fact that $H^1$ is a dual space (of VMO).

Finally, we remark that the hypothesis (6.1) is satisfied if

$$\sup_\nu \mathcal{H}_n(\partial F_\nu(\mathbb{R}^{n+1}_+)) < \infty$$

by [BK, Theorem 1.4]).

*Proof of Theorem 6.1.* After the preceding discussion it remains to prove the statement about the convergence of the boundary traces.

Fix a cube $Q_0 \subset \mathbb{R}^n$. For simplicity, we denote the traces $f_\nu|Q_0$ and $f|Q_0$ by $f_\nu$ and $f$, respectively. As already remarked above, these traces exist. We first claim that $(f_\nu)$ is a uniformly bounded sequence in the Sobolev space $W^{1,n}(Q_0)$. Indeed, a uniform bound

$$\sup_\nu \int_{Q_0} |Df_\nu|^n \, dm_n \leq C < \infty \tag{6.3}$$

follows from the hypothesis (6.1) and from Theorem 1.1, and so it remains to show that

$$\int_{Q_0} |f_\nu|^n \leq C < \infty. \tag{6.4}$$

Here and in the rest of the proof $C$ will denote various positive constants independent of $\nu$. The bound (6.4) can be obtained from (6.3) and from a Poincaré type inequality as follows. First note that by Proposition 2.3 and by (6.1) we have that

$$a_{F_\nu}(z_0) \leq C < \infty,$$

where $z_0$ is the center of $\widehat{Q}_0$. Now modulus estimates as, for example, in Lemma 4.4 (see [HeK, Lemma 6.6]) and the locally uniform convergence $F_\nu \to F$ imply that there exists, for each $\nu \in \mathbb{N}$, a set $E_\nu \subset Q_0$ such that $m_n(E_\nu) \geq \frac{1}{2} m_n(Q_0)$ and that

$$\sup_\nu \sup_{x \in E_\nu} \operatorname{dist}\big(f_\nu(x), F(z_0)\big) \leq C < \infty. \tag{6.5}$$

Therefore

$$\begin{aligned}\int_{Q_0} |f_\nu|^n \, dm_n &\lesssim \int_{Q_0} |f_\nu - (f_\nu)_{E_\nu}|^n \, dm_n + m_n(Q_0)|(f_\nu)_{E_\nu}|^n \\ &\leq C \int_{Q_0} |Df_\nu|^n + C,\end{aligned}$$

where the second inequality follows from the Poincaré inequality

$$\int_{Q_0} |u - u_E|^n \, dm_n \leq C(n) \, \frac{m_n(Q_0)^2}{m_n(E)} \int_{Q_0} |Du|^n \, dm_n \tag{6.6}$$

valid for all functions $u$ with $n$-summable distributional gradient in $L^n(Q_0)$ and for all measurable sets $E \subset Q_0$ [GT, Section 7.8]. Thus (6.4) follows.

Now let $g$ be any weak limit of a subsequence of $(f_\nu)$ in $W^{1,n}(Q_0;\mathbb{R}^{n+1})$. We continue to denote the subsequence by $(f_\nu)$, and our task is to show that $g = f$.

To this end, let $Q \subset Q_0$ be a dyadic cube. As before, $z_Q = (x_Q, t_Q)$ denotes the center of the cube $\widehat{Q}$ as in (1.2). We estimate

$$\begin{aligned}
\fint_Q |f(x) - g(x)|\, dm_n(x) \leq & \fint_Q |f(x) - F(z_Q)|\, dm_n(x) \\
& + \fint_Q |F(z_Q) - F_\nu(z_Q)|\, dm_n(x) \\
& + \fint_Q |F_\nu(z_Q) - f_\nu(x)|\, dm_n(x) \\
& + \fint_Q |f_\nu(x) - g(x)|\, dm_n(x).
\end{aligned} \tag{6.7}$$

Recall the definition for $r_{a_F}$ from (2.7), let $r := r_{a_F}$ and $r_\nu = r_{a_{F_\nu}}$ for $\nu \in \mathbb{N}$, and recall the maximal function $\mathcal{M}$ from (4.20). The first and the third term in the preceding sum can be bounded from above by a constant $C(n, K)$ times the maximal functions $\mathcal{M}(r)(Q)$ and $\mathcal{M}(r_\nu)(Q)$, respectively. For this estimation, see the proof of the implication (iii) $\Rightarrow$ (iv) of Theorem 4.1 in Section 4.

Next, by the Rellich-Kondrachev compactness theorem [GT, Section 7.10], we may pass to a subsequence, still denoted by $(f_\nu)$, which converges to $g$ in $L^1(Q_0)$. It follows that the Hardy-Littlewood maximal functions

$$M(f_\nu - g)(x) = \sup_{x \in Q} \fint_Q |f_\nu(x) - g(x)|\, dm_n(x), \qquad Q \subset Q_0,$$

converge to zero in measure as $\nu \to \infty$, and so by passing to yet another subsequence, we may assume that $M(f_\nu - g)(x) \to 0$ for almost every $x \in Q_0$ as $\nu \to \infty$.

Assume now that the set $\{x \in Q_0 : f(x) \neq g(x)\}$ has positive measure. Then there are $x_0 \in Q_0$ and $\delta > 0$ such that

$$\delta < \liminf_{Q \downarrow x_0} \fint_Q |f(x) - g(x)|\, dm_n(x), \tag{6.8}$$

where the notation $Q \downarrow x_0$ means that $Q$ runs through all dyadic cubes in $Q_0$ that contain $x_0$ and shrink down to $x_0$. Moreover, by the preceding discussion, we may assume that

$$\lim_{\nu\to\infty} M(f_\nu - g)(x_0) = 0. \tag{6.9}$$

The preceding understood, we conclude from (6.7) and (6.9) that there is a positive integer $N_0$ such that

$$\delta < \mathcal{M}(r)(Q) + |F(z_Q) - F_\nu(z_Q)| + \mathcal{M}(r_\nu)(Q) \tag{6.10}$$

for all dyadic cubes $Q$ in $Q_0$ that contain $x_0$ and satisfy $\ell(Q) \le 2^{-N_0}\,\ell(Q_0)$, and for all large enough $\nu$. If we recall that $r(\widehat{Q}) \simeq \operatorname{diam}(F(\widehat{Q}))$, we obtain from Theorem 4.1(iii), and from (6.2), that the function $Q \mapsto r(\widehat{Q})$ is in weak-$\ell^n$ (on the cubes $\mathcal{D}(\mathbb{R}^n)$). The same holds true for $r_\nu$, and by the hypotheses (6.1), the weak-$\ell^n$ norms are uniformly bounded (that is, independent of $\nu$).

Next recall that the maximal function $\mathcal{M}\colon$ weak-$\ell^n \to$ weak-$\ell^n$ is bounded (see the discussion after (4.20)). We can therefore pick an integer $L > 0$, independent of $\nu$, such that

$$\begin{aligned} &\#\big\{Q : Q \subset Q_0 \text{ dyadic and } \mathcal{M}(r)(Q) > \delta/3\big\} \\ &\quad + \#\big\{Q : Q \subset Q_0 \text{ dyadic and } \mathcal{M}(r_\nu)(Q) > \delta/3\big\} < L. \end{aligned} \tag{6.11}$$

We may obviously assume that the integer $N_0$, chosen above before (6.10), is large enough so that $\mathcal{M}(r)(Q) \le \delta/3$ for all dyadic $Q \subset Q_0$ with $\ell(Q) \le 2^{-N_0}\,\ell(Q_0)$. Inequality (6.10) then implies

$$\frac{2\delta}{3} < |F(z_0) - F_\nu(z_0)| + \mathcal{M}(r_\nu)(Q) \tag{6.12}$$

whenever $\ell(Q) \le 2^{-N_0}\,\ell(Q_0)$. On the other hand, for each $\nu$ there exists, by (6.11), a dyadic cube $Q \subset Q_0$ that contains $x_0$ and satisfies both

$$2^{-L-1}\,2^{-N_0}\,\ell(Q_0) \le \ell(Q) \le 2^{-N_0}\,\ell(Q_0) \tag{6.13}$$

and

$$\frac{\delta}{3} < |F(z_Q) - F_\nu(z_Q)|. \tag{6.14}$$

Now there are only finitely many cubes satisfying condition (6.13), and hence by passing to a subsequence, we may assume that (6.14) holds for a fixed cube $Q$ and for all $\nu$. But this is an obvious contradiction as $F_\nu \to F$ locally uniformly in $\mathbb{R}^{n+1}_+$.

We have thus shown that $f_\nu$ converges weakly to $f$ in $W^{1,n}(Q_0;\mathbb{R}^{n+1})$, and the proof of Theorem 6.1 is thereby complete. $\square$

Next we formulate and prove an abstract version of Theorem 6.1.

**Theorem 6.2.** *Let $(\varrho_\nu)$ be a sequence of conformal densities in $\mathbb{R}^{n+1}_+$, $n \geq 2$, with uniformly bounded data. Suppose that*

$$H := \sup_\nu \|\varrho_\nu\|_{QH^n(\mathbb{R}^n)} < \infty, \tag{6.15}$$

*and that the sequence $(\varrho_\nu)$ converges uniformly on compacta to a density $\varrho_0 \not\equiv 0$ on $\mathbb{R}^{n+1}_+$. Then $\varrho_0$ is a conformal density with*

$$\|\varrho_0\|_{QH^n(\mathbb{R}^n)} \leq H. \tag{6.16}$$

*Moreover, for each cube $Q_0 \subset \mathbb{R}^n$ and each $\Lambda \in (\ell^\infty)^*$, the functions $\langle \Lambda, i_{\varrho_\nu}\rangle | Q_0$ converge weakly in the Sobolev space $W^{1,n}(Q_0)$ to the function $\langle \Lambda, i_{\varrho_0}\rangle | Q_0$.*

Note that estimate (6.16) is a direct consequence of the locally uniform convergence $\varrho_\nu \to \varrho_0$. It is also clear that $\varrho_0$ is a conformal density, that is, it satisfies (2.4) and (2.5).

Next, consider the mappings

$$I_\nu = I_{\varrho_\nu} \colon \mathbb{R}^{n+1}_+ \to (\mathbb{R}^{n+1}_+, d_{\varrho_\nu})$$

as in (2.10), and similarly for $I_0 = I_{\varrho_0}$. We adopt the conventions made around (4.1) and assume that the mappings $I_\nu$, $I_0$ take values in the Banach space $\ell^\infty$. The discussion now parallels to that after the statement of Theorem 6.1. Conditions (6.15) and (6.16) imply by Theorem 4.1 that the traces

$$i_\nu = i_{\varrho_\nu} \colon \mathbb{R}^n \to \ell^\infty$$

belong to the Dirichlet-Sobolev space $L^{1,n}(\mathbb{R}^n; \ell^\infty)$ with uniform norm bound for $\nu \in \mathbb{N}$. From the Banach space-valued Sobolev embedding theorem [HKST, Theorem 6.2] we conclude that

$$\sup_\nu \int_{Q_0} |i_\nu(x) - (i_\nu)_{Q_0}|^n \, dm_n(x) \leq C < \infty,$$

where $(i_\nu)_{Q_0}$ stands for the mean value of $i_\nu$ over $Q_0$ as defined in (1.10) and (1.11). The modulus estimates alluded to in connection with (6.5) apply equally well in the setting of conformal densities and we conclude that there exist sets $E_\nu \subset Q_0$ such that $m_n(E_\nu) \geq \frac{1}{2}\, m_n(Q_0)$ and that

$$\sup_\nu \sup_{x \in E_\nu} \operatorname{dist}\big(i_\nu(x), I_0(z_0)\big) \leq C < \infty, \tag{6.17}$$

where $z_0$ is the center of $\widehat{Q}_0$ and the distance is taken in $\ell^\infty$. As in the proof of Theorem 6.1, therefore, we obtain a uniform bound for the traces $i_\nu$ in $L^n(Q_0; \ell^\infty)$, provided one can show a Poincaré inequality of the type (6.6) for Banach space-valued Sobolev mappings. Such an inequality follows from the results in [HaK], as is presented in Lemma 6.3 below.

The preceding discussion understood, we conclude that the sequence of mappings $i_\nu \colon Q_0 \to \ell^\infty$ is uniformly bounded in the Sobolev norm,

$$\|i_\nu\|_{W^{1,n}(Q_0;\ell^\infty)} := \|i_\nu\|_{L^n(Q_0;\ell^\infty)} + \inf \|\sigma_\nu\|_{L^n(Q_0)} \le C < \infty, \tag{6.18}$$

where the infimum is taken over all upper gradients $\sigma_\nu$ of $i_\nu$, and where $C$ is independent of $\nu \in \mathbb{N}$. On the other hand, there is no Rellich's theorem for Banach space-valued Sobolev functions, so that we have to be satisfied with the formulation of Theorem 6.2 via duality.

Before turning to the proof of Theorem 6.2, we require the following lemma.

**Lemma 6.3.** *Let $V$ be a Banach space and let $u \colon Q_0 \to V$ be a function in $L^1(Q_0;V) \cap L^{1,n}(Q_0;V)$, where $Q_0 \subset \mathbb{R}^n$ is a cube and $n \ge 2$. The inequality*

$$\int_{Q_0} |u - u_E|^n \, dm_n \le C(n) \frac{m_n(Q_0)^2}{m_n(E)} \int_{Q_0} \sigma^n \, dm_n \tag{6.19}$$

*holds for every upper gradient $\sigma$ of $u$ and for every measurable subset $E \subset Q_0$.*

*Proof.* It follows from [HaK, Proof of Theorem 5.2, p. 25, and Theorem 5.3] (see also [HKST, Section 6]) that for almost every $x$, $y$ in $Q_0$ the inequality

$$|u(x) - u(y)| \lesssim J\sigma(x) + J\sigma(y) \tag{6.20}$$

is valid, where $J$ is a generalized Riesz potential satisfying

$$\|J\sigma\|_{L^n(Q_0)} \lesssim \operatorname{diam}(Q_0) \, \|\sigma\|_{L^n(Q_0)}.$$

Thus, by integrating (6.20) first over $E$ and then over $Q_0$, we obtain

$$\begin{aligned} &\int_{Q_0} \int_E |u(x) - u(y)|^n \, dm_n(y) \, dm_n(x) \\ &\lesssim m_n(E) \int_{Q_0} J\sigma(x)^n \, dm_n(x) + m_n(Q_0) \int_E J\sigma(y)^n \, dm_n(y), \end{aligned}$$

and therefore

$$\begin{aligned} \fint_{Q_0} |u(x) - u_E|^n \, dm_n(x) &\lesssim \frac{m_n(Q_0)}{m_n(E)} \fint_{Q_0} J\sigma(x)^n \, dm_n(x) \\ &\lesssim \frac{m_n(Q_0)^2}{m_n(E)} \fint_{Q_0} \sigma^n(x) \, dm_n(x), \end{aligned}$$

as desired. Finally, the constants above only depend on $n$. The lemma follows. $\square$

*Proof of Theorem 6.2.* Let $\varrho_\nu$, $\varrho_0$ be as in the hypotheses, and let $Q_0 \subset \mathbb{R}^n$ be a cube. As we already pointed out in the preceding discussion, estimate (6.16) is an immediate consequence of the locally uniform convergence $\varrho_\nu \to \varrho_0$. Then fix $\Lambda \in (\ell^\infty)^*$. Without loss of generality, we assume that $|\Lambda| \leq 1$. By (6.17) and by (6.19) we have that (6.18) holds, as explained above. Because every upper gradient of $i_\nu$ is an upper gradient of $\langle \Lambda, i_\nu \rangle$ (see [HKST, Proof of Theorem 3.17]), and because the Sobolev space defined via upper gradients agrees with the standard Sobolev space in Euclidean domains [Sh, Theorem 4.5], we have from (6.18) that the sequence $(\langle \Lambda, i_\nu \rangle)$ is bounded in $W^{1,n}(Q_0)$. Let $g$ be any weak limit of a subsequence of $(\langle \Lambda, i_\nu \rangle)$ in $W^{1,n}(Q_0)$; we continue to denote the subsequence by $(\langle \Lambda, i_\nu \rangle)$. Analogously to the proof of Theorem 6.1, our task is to show that $g = \langle \Lambda, i_0 \rangle$. For this, we argue much as in the proof of Theorem 6.1.

Thus, let $Q \subset Q_0$ be a dyadic cube. We abbreviate $r_\nu := r_{\varrho_\nu}$ (see (2.7)) and estimate as in (6.7), and in the subsequent discussion,

$$
\begin{aligned}
&\fint_Q |\langle \Lambda, i_0(x) \rangle - g(x)| \, dm_n(x) \\
&\leq \fint_Q |\langle \Lambda, i_0(x) - I_0(z_Q) \rangle| \, dm_n(x) + |\langle \Lambda, I_0(z_Q) - I_\nu(z_Q) \rangle| \\
&\quad + \fint_Q |\langle \Lambda, I_\nu(z_Q) - i_\nu(x) \rangle| \, dm_n(x) \\
&\quad + \fint_Q |\langle \Lambda, i_\nu(x) \rangle - g(x)| \, dm_n(x) \\
&\lesssim \mathcal{M}(r_0)(Q) + |I_0(z_Q) - I_\nu(z_Q)| + \mathcal{M}(r_\nu)(Q) \\
&\quad + \fint_Q |\langle \Lambda, i_\nu(x) \rangle - g(x)| \, dm_n(x).
\end{aligned}
$$

Now the assertion $g = \langle \Lambda, i_0 \rangle$ follows precisely as in the proof of Theorem 6.1, by using Theorem 4.1.

This completes the proof of Theorem 6.2. □

## References


[Am] L. Ambrosio, *Metric space valued functions of bounded variation*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 17 **3** (1990), 439–478.

[As] K. Astala, *Quasiharmonic analysis and BMO*, University of Joensuu, Publications in Sciences, **14** (1989), 9–19.

[AG] K. Astala and F.W. Gehring, *Quasiconformal analogues of theorems of Koebe and Hardy-Littlewood*, Mich. Math. J. **32** (1985), 99–107.

[AK] K. Astala and P. Koskela, *$H^p$-theory for quasiconformal mappings*, in preparation.

[BL] Y. Benyamini and J. Lindenstrauss, *Geometric Nonlinear Functional Analysis, Volume I*, Vol. 48 of Colloquium Publications, American Mathematical Society, Providence, R.I., 2000.

[BM] A. Baernstein II and J.J. Manfredi, *Topics in quasiconformal mapping*, Topics in Modern Harmonic Analysis, Proc. Seminar held in Torino and Milano II (1982), 819–862.

[BA] A. Beurling and L.V. Ahlfors, *The boundary correspondence under quasiconformal mappings*, Acta Math. **96** (1956), 124–142.

[BK] M. Bonk and P. Koskela, *Conformal metrics and size of the boundary*, Preprint 240, Jyväskylä, 2001.

[BHK] M. Bonk, J. Heinonen and P. Koskela, *Uniformizing Gromov hyperbolic spaces*, Astérisque **270**, 2001.

[BHR] M. Bonk, J. Heinonen and S. Rohde, *Doubling conformal densities*, J. reine angew. Math. **541** (2001), 117–141.

[BKR] M. Bonk, P. Koskela and S. Rohde, *Conformal metrics on the unit ball in Euclidean space*, Proc. London Math. Soc. **77** (1998), 635–664.

[CST] A. Connes, D. Sullivan and N. Teleman, *Quasiconformal mappings, operators on Hilbert space, and local formulae for characteristic classes*, Topology **33** (1994), 663–681.

[DU] J. Diestel and J.J. Uhl, *Vector measures*, American Mathematical Society, Providence, R.I., 1977.

[D] J.L. Doob, *Classical Potential Theory and its Probabilistic Counterpart*, Springer-Verlag, New York, 1984.

[Fe] H. Federer, *Geometric Measure Theory*, Springer-Verlag, Berlin-Heidelberg-New York, 1969.

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

[Ga] J. Garnett, *Bounded Analytic Functions*, Academic Press, New York, 1981.

[G1] F.W. Gehring, *The Hausdorff measure of sets which link in euclidean space*, Contributions to analysis: A collection of papers dedicated to Lipman Bers, Academic Press, New York, 1974.

[G2] F.W. Gehring, *Lower dimensional absolute continuity properties of quasiconformal mappings*, Math. Proc. Cambridge Philos. Soc. **78** (1975), 81–93.

[G3] F.W. Gehring, *Absolute continuity properties of quasiconformal mappings*, Symposia Math. **XVIII** (1976), 551–559.

[GH] F.W. Gehring and W.K. Hayman, *An inequality in the theory of conformal mapping*, J. Math. Pure Appl. **41** (1962), 353–361.

[GO] F.W. Gehring and B.G. Osgood, *Uniform domains and the quasihyperbolic metric*, J. Anal. Math. **36** (1979), 50–74.

[GT] D. Gilbarg and N.S. Trudinger, *Elliptic Partial Differential Operators of Second Order*, 2nd Edition, Springer-Verlag, Berlin-Heidelberg-New York, 1983.

[HaK] P. Hajłasz and P. Koskela, *Sobolev met Poincaré*, Memoirs Amer. Math. Soc. **145**, 2000.

[Ha] B. Hanson, *Quasiconformal analogues of a theorem of Smirnov*, Math. Scand. **75** (1994), 133–149.

[H1] J. Heinonen, *The boundary absolute continuity of quasiconformal mappings*, Amer. J. Math. **116** (1994), 1545–1567.

[H2] J. Heinonen, *The boundary absolute continuity of quasiconformal mappings II*, Rev. Mat. Iberoamericana **12** (1996), 697–725.

[H3] J. Heinonen, *A theorem of Semmes and the boundary absolute continuity in all dimensions*, Rev. Mat. Iberoamericana **12** (1996), 783–789.

[HKM] J. Heinonen, T. Kilpeläinen and O. Martio, *Nonlinear Potential Theory of Degenerate Elliptic Equations*, Oxford University Press, 1993.

[HeK] J. Heinonen and P. Koskela, *The boundary distortion of a quasiconformal mapping*, Pacific J. Math. **165** (1994), 93–114.

[HKST] J. Heinonen, P. Koskela, N. Shanmugalingam and J.T. Tyson, *Sobolev classes of Banach space-valued functions and quasiconformal mappings*, J. Anal. Math. **85** (2001), 87–139.

[HS] J. Heinonen and S. Semmes, *Thirty-three YES or NO questions about mappings, measures, and metrics*, Conformal Geom. Dyn. **1** (1997), 1–12.

[JW] D. Jerison and A. Weitsman, *On the means of quasiregular and quasiconformal mappings*, Proc. Amer. Math. Soc. **83** (1981), 304–306.

[J] P.W. Jones, *Extension theorems for BMO*, Indiana Univ. Math. J. **29** (1980), 41–66.

[K] W. Kaplan, *On Gross's star theorem, schlicht functions, logarithmic potentials, and Fourier series*, Ann. Acad. Sci. Fenn. Ser. A I Math.-Phys. **86** (1951), 1–23.

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

[KS] N.J. Korevaar and R.M. Schoen, *Sobolev spaces and harmonic maps for metric space targets*, Comm. Anal. Geom. **1** (1993), 561–659.

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

[KR] P. Koskela and S. Rohde, *Hausdorff dimension and mean porosity*, Math. Ann. **309** (1997), 593–609.

[Ma] N.G. Makarov, *Probability methods in the theory of conformal mappings*, (in Russian) Algebra and Analysis **1** (1989), 1–59.

[Mz] V.G. Maz'ja, *Sobolev spaces*, Springer-Verlag, Berlin, 1985.

[MZ] J. Malý and W.P. Ziemer, *Fine regularity of solutions of elliptic partial differential equations*, AMS Mathematical Surveys and Monographs, **51**, Providence, R.I., 1997.

[Po] Ch. Pommerenke, *Boundary Behaviour of Conformal Maps*, Springer-Verlag, Berlin-Heidelberg-New York, 1992.

[Pr] I.I. Priwalow, *Randeigenschaften analytischer Funktionen*, Deutscher Verlag Wiss., Berlin, 1956.

[Re] Yu.G. Reshetnyak, *Sobolev classes of functions with values in a metric space*, Sibirsk. Mat. Zh. **38** (1997), 657–675.

[Ri] F. and M. Riesz, *Über die Randwerte einer analytischen Funktion*, 4. Cong. Scand. Math., Stockholm (1916), 27–44.

[RS1] R. Rochberg and S. Semmes, *End point results for estimates of singular values of singular integral operators*, Contributions to operator theory and its applications (Mosa, AZ, 1987), pp. 217–231, in Oper. Theory Adv. Appl. **35**, Birkhäuser, Basel, 1988.

[RS2] R. Rochberg and S. Semmes, *Nearly weakly orthonormal sequences, singular value estimates, and Calderón-Zygmund operators*, J. Funct. Anal. **86** (1989), 237–306.

[Se] S. Semmes, *Quasisymmetry, measure, and a question of Heinonen*, Rev. Mat. Iberoamericana **12** (1996), 727–781.

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

[St] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton Univ. Press, Princeton, N.J., 1970.

[SW] E. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton Univ. Press, Princeton, N.J., 1971.

[Su] D. Sullivan, *Hyperbolic geometry and homeomorphisms*, in Geometric Topology (Proc. Georgia Topology Conf. Athens, Ga, 1977), pp. 543–555, Academic Press, New York, 1979.

[TV] P. Tukia and J. Väisälä, *Lipschitz and quasiconformal approximation and extension*, Ann. Acad. Sci. Fenn. Ser. A I Math. **6** (1981), 303–342.

[V1] J. Väisälä, *Lectures on n-dimensional Quasiconformal Mappings*, Lecture Notes in Math. 229, Springer-Verlag, Berlin, 1971.

[V2] J. Väisälä, *Quasi-symmetric embeddings in Euclidean spaces*, Trans. Amer. Math. Soc. **264** (1981), 191–204.

[V3] J. Väisälä, *The wall conjecture on domains in Euclidean spaces*, Manuscripta Math. **93** (1997), 515–534.

[Vu] M. Vuorinen, *Conformal Geometry and Quasiregular Mappings*, Lecture Notes in Math. 1319, Springer-Verlag, Berlin, 1988.

[W] P. Wojtaszczyk, *Banach spaces for analysts*, Cambridge Studies in adv. math. 25, Cambridge University Press, 1991.

[Y] K. Yosida, *Functional Analysis*, Springer-Verlag, Berlin-Heidelberg-New York, 1965.

[Zie] W.P. Ziemer, *Extremal length and p-capacity*, Mich. Math. J. **16** (1969), 43–51.

[Zin] M. Zinsmeister, *A distortion theorem for quasiconformal mappings*, Bull. Soc. Math. France **114** (1986), 123–133.

[Zo] V.A. Zorich, *Angular boundary values of quasiconformal mappings of a ball*, Dokl. Akad. Nauk. SSSR **179** (1967), 1492–1494.

(K.A.) Department of Mathematics and Statistics, University of Jyväskylä, P.O. Box 35 (MaD), FIN-40351 Jyväskylä, Finland
*E-mail address*: astala@math.jyu.fi

(M.B.) Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA
*E-mail address*: mbonk@math.lsa.umich.edu

(J.H.) Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA
*E-mail address*: juha@math.lsa.umich.edu