

# CAPACITY, CARLESON MEASURES, BOUNDARY CONVERGENCE, AND EXCEPTIONAL SETS

N. ARCOZZI, R. ROCHBERG, AND E. SAWYER

## 1. INTRODUCTION AND SUMMARY

There is a fundamental relation between the capacity of a set and energy integrals of probability measures supported on that set. If the capacity is small the energy integral will be large; in particular sets of capacity zero cannot support probability measures of finite energy. Here we develop similar ideas relating capacity to Carleson measures. We show that if a set has small capacity then any probability measure supported on it must have large Carleson embedding constant. In particular sets of capacity zero are exactly the simultaneous null set for all nontrivial Carleson measures.

Functions having limited smoothness often exhibit attractive or convenient behavior at most points, the exceptional set being of capacity zero; that is, the good behavior holds *quasi-everywhere*, henceforth *q.e.*. Using the relationship between capacity and Carleson measures such a conclusion can be reformulated by saying the function exhibits the good behavior  $\mu - a.e.$  for every Carleson measure  $\mu$ . This can be useful because sometimes it is relatively easy, even tautological, to establish that a property holds  $\mu - a.e.$ . We will use this viewpoint to give a new approach to results related to boundary behavior of holomorphic and harmonic functions.

The dyadic Dirichlet space is a Hilbert space of functions on a dyadic tree. In many ways that space models the classical Dirichlet space, the space of holomorphic functions,  $f$ , on the unit disk  $\mathbb{D}$  for which  $\int_{\mathbb{D}} |f'|^2 < \infty$ . In Sections 3 and 4 we present background material on the dyadic Dirichlet space and the associated theory of Carleson measures. In Section 5 we present our new results on Carleson measures for the dyadic Dirichlet space. Those include the use of Carleson measures to measure capacity and a direct proof of the equivalence of the measure theoretic and capacity theoretic criteria for a Carleson embedding.

In Section 6 we show how those results can be used to study boundary behavior and exceptional sets. Roughly, the idea is to work with the tree geometry to construct a majorant of the variation of the function being studied. If this majorant is in a discrete function space,  $X$ , then the boundary set on which the majorant is infinite must be a null set for every  $X$ -Carleson measure. We then appeal to results of Section 4 to recast this as a statement about the capacity of the exceptional set.

The discrete case is a model case in which the proofs are relatively straightforward and the geometric issues we wish to highlight are particularly clear; and in this paper we will focus almost exclusively on that model case. However the

---

The first author partially supported by the COFIN project Analisi Armonica, funded by the Italian Minister for Research. The second author's work supported by the National Science Foundation under Grant No. 0070642. The third author's work was supported by the National Science and Engineering Council of Canada.

ideas we present are very general and we believe they can be extended to the study of smooth functions whenever there is control of local oscillation. The details of extending these techniques to classes of smooth functions, and to function spaces that are not Hilbert spaces, requires some further ideas and technical machinery. In Section 7 we briefly discuss the types of extensions we have in mind and our approach to them; however we postpone a detailed presentation of that work to a later paper [ARSp].

## 2. EMBEDDING MAPS

**2.1. Capacity Conditions.** Here we record some results which go back to classic work by Maz'ya. For proofs and references we refer to [AH] and [AE].

For a function  $f$  defined on  $\mathbb{R}^2$  we define its Riesz potential of order one by  $I_1 f = (-\Delta)^{-1/2} f$ . We define the associated ( $p = 2$ ) capacity of a compact  $E \subset \mathbb{R}^2$  by

$$\begin{aligned} \text{Cap}(E) &= \text{Cap}_{1,2}(E) \\ &= \inf \left\{ \int_{\mathbb{R}^2} g^2 dx : I_1 g \geq 1 \text{ on } E, g \geq 0, g \in L^2(\mathbb{R}^2) \right\} \\ &\approx \inf \left\{ \int_{\mathbb{R}^2} |\nabla u|^2 dx : u \geq 1 \text{ on } E, u \in C_0^\infty \right\}. \end{aligned}$$

(Below we will tacitly assume all sets discussed are capacitable.) The basic trace theorem in this context is:

**Theorem 1.** *Suppose  $\omega$  is a nonnegative Borel measure on  $\mathbb{R}^2$ . The following are equivalent:*

- (1) *There is a constant  $C$  so that  $\forall u \in C_0^\infty$*

$$\|u\|_{L^2(d\omega)} \leq C \|\nabla u\|_{L^2(dx)}.$$

- (2) *There is a constant  $C$  so that  $\forall f \in L^2(dx)$*

$$\|I_1 f\|_{L^2(d\omega)} \leq C \|f\|_{L^2(dx)}.$$

- (3) *There is a constant  $C$  so that for all compact  $E \subset \mathbb{R}^2$*

$$\omega(E) \leq C \text{Cap}(E).$$

There are similar results for functions defined in regions of  $\mathbb{R}^2$  and for spaces of holomorphic functions. In particular we recall the result of Stegenga [St] for  $\mathcal{D}$ , the Dirichlet space of holomorphic functions on the unit disk  $\mathbb{D}$ ;

$$\mathcal{D} = \left\{ f \in \text{Hol}(\mathbb{D}) : \|f\|_{\mathcal{D}}^2 = |f(0)|^2 + \int_{\mathbb{D}} |f'|^2 dV < \infty \right\}.$$

Recall that functions in  $\mathcal{D}$  have boundary values *a.e.* on  $\mathbb{T}$ .

For any  $z \in \mathbb{D}$  we define the associated interval  $I_z$  in  $\mathbb{T}$ , the boundary of  $\mathbb{D}$ , to be the interval with center  $z/|z|$  and length  $2(1 - |z|)$ . For any open set  $\mathcal{O}$  contained in  $\mathbb{T}$  we define the *tent over  $\mathcal{O}$* ,  $T(\mathcal{O})$  by

$$T(\mathcal{O}) = \{z \in \mathbb{D} : I_z \subset \mathcal{O}\}.$$

For such an  $\mathcal{O}$  we define its  $\mathcal{D}$ -capacity by

$$\text{Cap}_{\mathcal{D}}(\mathcal{O}) = \inf \left\{ \|f\|_{\mathcal{D}}^2 : \text{Re } f \geq 1 \text{ on } \mathcal{O} \right\}.$$

In particular, for an interval  $I \subset T$  we have the estimate

$$\text{Cap}_{\mathcal{D}}(I) \sim \frac{1}{|\log |I||}.$$

**Theorem 2.** [St] *The following are equivalent for a positive measure  $\omega$  supported on  $\mathbb{D}$ :*

(1) *There is a constant  $C$  so that  $\forall f \in \mathcal{D}$*

$$(2.1) \quad \|f\|_{L^2(d\omega)} \leq C \|f\|_{\mathcal{D}}.$$

(2) *There is a constant  $C$  so that for any open  $\mathcal{O} \subset \mathbb{T}$*

$$(2.2) \quad \omega(T(\mathcal{O})) \leq C \text{Cap}_{\mathcal{D}}(\mathcal{O}).$$

**2.2. Testing conditions.** One approach to results such as the previous two is to find a good collection of testing functions and evaluate the desired inequality, for instance (2.1), on those functions. The inequalities obtained are of course necessary in order to have (2.1) hold for all functions and, in the favorable cases, those conditions are also sufficient.

A famous example where this approach succeeds is the following. Let  $H^2$  be the Hardy space, the subspace of  $L^2(\mathbb{T}, d\theta)$  consisting of functions,  $f$ , whose negative Fourier coefficients vanish. Such a function,  $f$ , extends to a holomorphic function in  $\mathbb{D}$  which we also denote by  $f$ .

**Theorem 3.** (Carleson) *The following are equivalent for a positive measure  $\omega$  supported on  $\mathbb{D}$ :*

(1) *There is a constant  $C$  so that  $\forall f \in H^2(\mathbb{T})$*

$$(2.3) \quad \|f\|_{L^2(d\omega)} \leq C \|f\|_{L^2(\mathbb{T}, d\theta)}$$

(2) *There is a constant  $C$  so that for any open interval  $I \subset \mathbb{T}$*

$$(2.4) \quad \omega(T(I)) \leq C |I|.$$

$H^2$  is a Hilbert space with reproducing kernels; the functions  $k_z(e^{i\theta}) = (1 - \bar{z}e^{i\theta})^{-1}$  are the kernel functions. Testing (2.3) on  $k_z$  gives the estimate (2.4) for the interval  $I_z$ .

A digression on terminology: The issues in the previous theorems are certainly similar to each other in spirit. However the questions arose in different contexts and are discussed using different words. The result in Theorem 1 is referred to as a *trace theorem*. On the other hand the measures in Theorem 3 are referred to as *Carleson measures (for the Hardy space)* and the results in Theorems 2 and 3 are described as *characterizing the Carleson measures* for the respective spaces.

(Actually there is a lack of consensus on terminology here. Some call a measure a "Carleson measure" if it satisfies a measure theoretic condition in the style of (2.4); the measure of a region of simple shape is controlled by a function of the size of that region. Other people use the name to denote measures for which conclusions such as (2.1) and (2.3) hold; there is a continuous embedding of the space being studied into a Lebesgue space associated with the measure. We use the second style and call the measures of

Theorem 2 Carleson measures. Also we call the embeddings whose continuity is insured by (2.1) or (2.3) Carleson embeddings and refer to the norms of those embeddings as the embedding constants of the measures or as the Carleson measure norms of the measures.)

Although the testing philosophy is a guide to the correct conjecture in this case, that is not always so. Similar ideas for the space  $\mathcal{D}$  would lead to speculation that it was sufficient to have (2.1) for all intervals  $I$ , rather than all open sets. However that is not the correct answer and Stegenga provides the appropriate examples. Yet there is a slightly more sophisticated testing scheme which can be used effectively to study Carleson measures on the Dirichlet space. The idea, which goes back to work by Kerman and Sawyer [KS], is to test the embedding operator and also test its adjoint.

One main theme of this paper is the relationship between results such as Theorems 1 and 2 formulated in terms of capacity and the more measure theoretic results suggested by the testing scheme just described.

For more about capacity in general, and in particular the fact that set functionals which appear to be quite different can be used to measure capacity, we refer to [AH], [AE], [V], and [KV].

### 3. THE DYADIC DIRICHLET SPACE

We will work with functions on the dyadic tree  $T$  which we now define.  $T$  is a rooted, directed, loopless graph. It contains a root vertex,  $o$ , which is connected by two edges to vertices  $o_r$  and  $o_l$  (its right and left children). Every other element  $\alpha$  is connected to three vertices; one, its parent,  $\alpha^-$ , is on the path connecting  $\alpha$  to  $o$ , the other two are its children,  $\alpha_r$  and  $\alpha_l$ . For  $\alpha, \beta \in T$  we denote by  $[\alpha, \beta]$  the set of vertices on the geodesic path from  $\alpha$  to  $\beta$ . We let  $d(\alpha, \beta)$  denote the length of that path,  $d(\alpha, \beta) = |[\alpha, \beta]| - 1$ , and we will abbreviate  $d(o, \alpha)$  to  $d(\alpha)$ . For  $\alpha, \beta \in T$  we write  $\beta \prec \alpha$  if  $\beta \in [o, \alpha]$  and we define the successor set of  $\beta$  to be  $S(\beta) = \{\alpha : \alpha \succeq \beta\}$

We will work with several operators defined on functions on  $T$ . For  $f$  a function on  $T$  we define functions  $Df$ ,  $If$ , and  $I^*f$  by

$$\begin{aligned} Df(\alpha) &= \begin{cases} f(o) & \text{if } \alpha = o \\ f(\alpha) - f(\alpha^+) & \text{otherwise,} \end{cases} \\ If(\alpha) &= \sum_{\gamma \in [0, \alpha]} f(\gamma), \\ I^*f(\alpha) &= \sum_{\gamma \in S(\alpha)} f(\gamma). \end{aligned}$$

We have  $ID = DI = Id$ , the identity operator, and  $I^*$  is the adjoint of  $I$  with respect to the inner product on  $l^2(T)$ .

We envision  $T$  as a subset of  $\mathbb{D}$  with  $o$  at the origin and with the  $2^n$  vertices in  $\{\alpha : d(\alpha) = n\}$  located at equally spaced points on the circle centered at the origin and of radius  $1 - 2^{-n}$ . With this in mind we define the *dyadic Dirichlet space*,  $\mathcal{D}_d$ , to be the space of functions  $F$  on  $T$  for which

$$\|F\|_{\mathcal{D}_d}^2 = |F(o)|^2 + \sum_T |DF|^2 < \infty.$$

Informally we think of  $T$  as sitting in  $\mathbb{D}$  and of holomorphic functions  $F \in \mathcal{D}$  as being modeled by their restrictions to  $T$ . Oversimplifying a bit we have these restriction are in  $\mathcal{D}_d$  and there is an informal correspondence between the operators  $\int_0^z$  and  $(1 - |z|^2) \frac{d}{dz}$  acting on holomorphic functions and the operators  $I$  and  $D$  operating on functions on the tree. This point of view is developed in detail in [Ar], [AR], [ARS1], and [ARS2].

In this context we will associate to any measure,  $\mu$ , defined on  $\mathbb{D}$  a measure  $\mu_T$  which is defined on  $T$  by setting  $\mu_T(\alpha)$  equal to the  $\mu$  measure of the set of  $z \in \mathbb{D}$  which are hyperbolically closer to  $\alpha$  (thought of as sitting in  $D$ ) than to any other point of  $T$  (with any convenient choice in cases of ties).

We will also be interested in functions and measures defined on the (abstract) boundary of  $T$ ,  $\partial T$ , which we now introduce. A point of  $\partial T$  is an equivalence class of geodesics in  $T$  each infinite in one direction. The geodesics  $\Gamma$  and  $\Gamma'$  are equivalent exactly if their symmetric difference,  $(\Gamma \cup \Gamma') \setminus (\Gamma \cap \Gamma')$ , is finite. A convenient representative for a class is the unique geodesic in the class starting at the root. For each  $\alpha \in T$  let  $\partial S(\alpha)$  be the set of  $\tau$  in  $\partial T$  with the property that every representative geodesic for  $\tau$  intersects  $S(\alpha)$ . The topology of  $\partial T$  is defined by declaring the sets  $\partial S(\alpha)$  to be a basis for the open neighborhoods. We set  $\bar{T} = T \cup \partial T$ . Note that any representative geodesic  $\Gamma$  for a point of  $\partial T$ , when viewed as a subset of  $\mathbb{D}$ , is a sequence that converges to a point of  $\mathbb{T}$ . Hence when we think of  $T$  as a subset of  $\mathbb{D}$  it is also convenient to think of  $\partial T$  as identified with  $\mathbb{T}$  with its usual topology. This lets us identify open intervals of  $\mathbb{T}$  with open sets  $\partial T$ . Using this identification we will regard measures defined on  $\mathbb{T}$  as also defined on  $\partial T$  and vice versa. (The pairing of  $\partial T$  with  $\mathbb{T}$  has a countable set of ambiguous points, but these cause no difficulty in the discussion which follows.) Finally, if a function  $F$  on  $T$  has a limit along a geodesic  $\Gamma$  then we extend it by continuity to the corresponding point of  $\partial T$ . In particular if  $f$  has finite support then  $If$  extends to the entire boundary.

We define capacity for subsets  $E$  of  $\bar{T}$  by

$$(3.1) \quad \text{Cap}_T(E) = \inf\{\|\varphi\|_{\ell^2}^2 : I\varphi|_E \geq 1\}.$$

Alternatively, there is a straightforward variation which we will use later

$$(3.2) \quad \text{Cap}_T(E) = \inf\{\|I^*\mu\|_{\ell^2}^2 : \mu \text{ a positive measure on } E, II^*\mu|_E \geq 1\}.$$

The quantity  $\|I^*\mu\|_{\ell^2}^2$  is sometimes called the *energy integral of  $\mu$*  and denoted  $\mathcal{E}(\mu)$ .

#### 4. CARLESON MEASURES FOR $\mathcal{D}_d$ AND $\mathcal{D}$

Here is the characterization of Carleson measures for  $\mathcal{D}_d$  and  $\mathcal{D}$  from [ARS1] which complements Stegenga's description of Carleson measures for the Dirichlet space. (Actually the final statement of the theorem, a capacity condition in the style of Stegenga, is not in [ARS1] but can be included, for instance, by virtue of Theorem 7 below.)

We will also be interested in the radial variation of functions in  $\mathcal{D}$  and we now introduce a function which measures that. For  $f$  in  $\mathcal{D}$  we define

$$V(f)(re^{i\theta}) = \int_0^r |f'(te^{i\theta})| dt.$$

**Theorem 4.** *Let  $\mu$  be a positive Borel measure on  $\overline{\mathbb{D}}$ . Then, the following conditions are equivalent:*

- (1)  $\mu$  is a Carleson measure for  $\mathcal{D}$ ; i.e. (2.1) holds.
- (2)  $\mu$  satisfies a Carleson inequality for the radial variation. There is a constant  $C$  so that  $\forall f \in \mathcal{D}$

$$\int_{\overline{\mathbb{D}}} |V(f)(z)|^2 d\mu \leq C \|f\|_{\mathcal{D}}^2.$$

- (3)  $\mu_T$  is a Carleson measure for  $\mathcal{D}_d$ . There is a constant  $C$  so that  $\forall f \in \mathcal{D}_d$

$$(4.1) \quad \sum_{\alpha \in T} |F(\alpha)|^2 \mu_T(\alpha) \leq C \|F\|_{\mathcal{D}_d}^2.$$

- (4)  $\mu_T$  satisfies the tree condition. There is a constant  $C$  so that  $\forall \alpha \in T$

$$(4.2) \quad I^*(I^* \mu_T)^2(\alpha) \leq C I^* \mu_T(\alpha) < \infty.$$

- (5)  $\mu_T$  satisfies a tree capacity condition. There is a constant  $C$  so that for all sets  $E = \cup_j \partial S(x_j)$  in  $\partial T$ ,

$$(4.3) \quad \mu_T \left( \cup_j \overline{S(x_j)} \right) \leq C \text{Cap}_T(\cup_j \partial S(x_j)).$$

The condition (4.1) states that the "integration map"  $I$  is bounded from  $l^2(T)$  into  $l^2(T, \mu_T)$ . We could analyze that condition by testing it on simple functions as follows. Pick  $\alpha \in T$  and set  $f = d(\alpha)^{-1} \chi_{(o, \alpha]}$ . If (4.1) holds then we find

$$\begin{aligned} (I^* \mu_T)(\alpha) &= \mu(S(\alpha)) \leq \|If\|_{l^2(T, \mu)}^2 \\ &\leq C \|If\|_{\mathcal{D}_d}^2 = d(\alpha)^{-1}. \end{aligned}$$

The condition obtained,  $d(\alpha)\mu(S(\alpha)) \leq C$ , is equivalent to the condition (4.3) restricted to the case where  $\{x_j\}$  has only one element. As was true in Stegenga's theorem that condition is necessary for the boundedness of  $I$  but not sufficient. However if  $I$  is bounded then so is its adjoint,  $I^*$ , mapping  $l^2(T, \mu_T)$  to  $l^2(T)$ . Condition (4.2) tests the boundedness of  $I^*$  on the functions  $\chi_{S(\alpha)}$ ; (4.2) is a slightly weakened version of the statement that

$$\left\| I^* \chi_{S(\alpha)} \right\|_{l^2(T)} \leq C \left\| \chi_{S(\alpha)} \right\|_{l^2(T, \mu_T)}.$$

Let  $\mu$  be a positive, Borel measure on  $\overline{T}$ . We denote the best constant in the testing condition (4.2) by  $\|\mu\|_{CM(T)}$ ;

$$(4.4) \quad \|\mu\|_{CM(T)} = \sup_{\alpha \in T} \frac{I^*(I^* \mu)^2(\alpha)}{I^* \mu(\alpha)}.$$

By the previous theorem this quantity is comparable to the norm of the Carleson embedding on  $\mathcal{D}_d$ .

## 5. CAPACITY AND CARLESON MEASURES ON TREES.

**5.1. Computing Capacity Using Carleson Measures.** The ordinary definition of capacity for a subset  $E$  of  $T$  is that in (3.1). We now give an alternative way of computing capacity for subsets of  $\overline{T}$  using Carleson measures. To minimize the notational burden we give the proof below for  $E \subset \partial T$  which is the case of primary interest. The extension to the general case is straightforward.

**Theorem 5.** *Let  $E \subset \bar{T}$  be compact. Then,*

$$(5.1) \quad \text{Cap}_T(E) = \sup_{\text{Supp}(\mu) \subset E} \frac{\mu(E)}{\|\mu\|_{CM(T)}}.$$

Before proceeding with the proof we note that the quantities which appear in the definition  $\|\mu\|_{CM(T)}$  are in fact related to quantities which show up in the study of capacity. From the definitions we have that  $I^*\mu(\alpha) = \mu(\bar{S}(\alpha))$ . If we multiply out the square in  $I^*(I^*\mu)^2(\alpha)$  we find the comparable quantity

$$I^*(I^*\mu)^2(\alpha) \sim \sum_{\beta, \beta' \succ \alpha} d(\alpha, \beta \wedge \beta') \mu(\beta) \mu(\beta').$$

Thus, completely informally, the ratio in (4.2) is comparing the mass of the measure  $\mu\chi_{S(\alpha)}$  to a type of energy integral of the same measure.

*Proof.* We start with inequality  $[\leq]$ . For fixed  $\alpha \in T$  we denote by  $\bar{T}_\alpha = \bar{S}(\alpha)$  the subtree of  $\bar{T}$  having root  $\alpha$  and we add a subscript  $\alpha$  to the corresponding tree objects:  $Cap_\alpha$  is the capacity of subsets of  $\bar{T}_\alpha$ ,  $\mathcal{E}_\alpha$  is the energy in  $\bar{T}_\alpha$ ,  $\omega_\alpha$  is the extremal measure in the definition of capacity, and so on. Let  $E_\alpha = E \cap \bar{T}_\alpha$ . The extremal measure  $\omega_\alpha$  and the function  $\varphi_\alpha = I_\alpha^*\omega$  satisfy

$$(5.2) \quad Cap_\alpha(E_\alpha) = \omega_\alpha(E_\alpha) = \mathcal{E}_\alpha(\omega_\alpha) = \|\varphi_\alpha\|_{L^2(\bar{T}_\alpha)}^2.$$

We **claim** that  $\omega_\alpha$  is a rescaling of the extremal measure for  $E$  in  $\bar{T}$ ,  $\omega$ , restricted to  $E_\alpha$ :

$$(5.3) \quad \omega_\alpha = \frac{\omega|_{E_\alpha}}{1 - II^*\omega(\alpha^{-1})}.$$

Here,  $\alpha^{-1}$  is the predecessor of  $\alpha$  in  $T$ . In fact,  $\omega_\alpha$  minimizes  $\mathcal{E}_\alpha(\mu)$  over all measures  $\mu$  such that  $I_\alpha I_\alpha^*\mu(\xi) \geq 1$  on  $E_\alpha$  (with the possible exception of a set having null-capacity). On the other hand, we claim that  $\omega|_{E_\alpha}$  minimizes  $\mathcal{E}_\alpha(\nu)$  among all measures  $\nu$  on  $E_\alpha$  such that  $I_\alpha I_\alpha^*\nu(\xi) \geq 1 - II^*\omega(\alpha^{-1})$  *p.p.* on  $E_\alpha$ .

Suppose this is not the case. Then there exists a measure  $\nu$  on  $E_\alpha$  such that  $I_\alpha I_\alpha^*\nu(\xi) \geq 1 - II^*\omega(\alpha^{-1})$  for *p.p.*  $\xi \in E_\alpha$  and

$$\mathcal{E}_\alpha(\nu) = \sum_{T_\alpha} (I^*\nu)^2 < \sum_{T_\alpha} (I^*\omega)^2 = \mathcal{E}_\alpha(\omega|_{E_\alpha}).$$

Consider the functions  $I^*\omega$  in  $\bar{T}$  and  $I_\alpha^*\nu = I^*\nu$  in  $\bar{T}_\alpha$ . Define a new function  $\psi$  on  $T$ :

$$\psi(x) = \begin{cases} I^*\nu(x) & \text{if } x \in T_\alpha, \\ I^*\omega(x) & \text{if } x \in T \setminus T_\alpha. \end{cases}$$

We have

$$I\psi(\xi) \geq 1 \text{ q.e. on } E,$$

hence  $\|\psi\|_{L^2}^2 \geq Cap(E)$ . On the other hand,

$$\|\psi\|_{L^2}^2 = \mathcal{E}_\alpha(\nu) + [\mathcal{E}(\omega) - \mathcal{E}_\alpha(\omega)] < \mathcal{E}(\omega) = Cap(E),$$

and we have reached a contradiction.

The measure

$$\lambda = \frac{\omega|_{E_\alpha}}{1 - II^*\omega(\alpha^{-1})},$$

then, minimizes  $\mathcal{E}_\alpha(\mu)$  over the set of the measures  $\mu$  such that  $II^*\mu(\xi) \geq 1$  for  $p.p.$   $\xi$  in  $E_\alpha$ , hence  $\lambda = \omega_\alpha$ . The claim is proved.

By the homogeneity of the energy,

$$\begin{aligned}\mathcal{E}_\alpha(\omega|_{E_\alpha}) &= (1 - II^*\omega(\alpha^{-1}))^2 \mathcal{E}_\alpha(\omega_\alpha) \\ &= (1 - II^*\omega(\alpha^{-1}))^2 \omega_\alpha(E_\alpha) \\ &= (1 - II^*\omega(\alpha^{-1})) \omega(E_\alpha).\end{aligned}$$

As a consequence,

$$(5.4) \quad \frac{\sum_{x \geq \alpha} (I^*\omega)^{p'} \rho^{1-p'}}{I^*\omega} = \frac{\mathcal{E}_\alpha(\omega|_{E_\alpha})}{\omega(E_\alpha)} = (1 - V(\omega)(\alpha^{-1}))^{p'-1} \leq 1,$$

with equality if and only if  $\alpha = o$  (we set the default value  $II^*\omega(o^{-1}) = 0$ ). Hence,  $\|\omega\|_{CM} = 1$  and

$$Cap(E) = \omega(E) = \frac{\omega(E)}{\|\omega\|_{CM}}.$$

We now prove  $[\geq]$ . By definition of  $\|\cdot\|_{CM}$ ,  $\mathcal{E}(\mu) \leq \|\mu\|_{CM} \mu(E)$  for all measures  $\mu$ . Then,

$$\frac{\mu(E)}{\|\mu\|_{CM}} \leq \frac{\mu(E)}{\left(\frac{\mathcal{E}(\mu)}{\mu(E)}\right)} = \frac{\mu(E)^2}{\mathcal{E}(\mu)} \leq Cap(E),$$

as wished.  $\square$

**Corollary 1.**  *$E$  has capacity zero if and only if it is a null set for all Carleson measures.*

**5.2. On a Question of Maz'ya.** Conditions 4 and 5 of Theorem 4 are, by that theorem, equivalent to each other. Some time ago Prof. Maz'ya asked if we could give a direct proof of that equivalence, one not relying on the fact that both conditions characterize the same class of Carleson measures. We do that now. In our proof we will use the fact that if a measure satisfies condition 4 then so does any smaller positive measure. That fact follows from knowing that condition 4 characterizes a class of Carleson measures. However, in the spirit of this section, we also give a proof of that fact which is relatively elementary and which does not involve the theory of Carleson measures. We begin with the monotonicity.

For a measure  $\mu$  on  $\overline{T}$ , let  $\sigma_\mu = (I^*\mu)^2$ .

**Theorem 6.** *Let  $\mu$  be a measure on  $\overline{T}$  and let  $\lambda$  be a measurable function on  $\overline{T}$ ,  $0 \leq \lambda \leq 1$ . If  $I^*\sigma_\mu \leq I^*\mu$  on  $T$ , then  $I^*\sigma_{\lambda\mu} \leq 2 \cdot I^*(\lambda\mu)$ .*

**Corollary 2.** *If  $\nu \leq \mu$  and  $\mu, \nu$  are measures on  $\overline{T}$ , then  $\|\nu\|_{CM(T)} \leq 2\|\mu\|_{CM(T)}$ .*

*Proof.* By rescaling, it suffices to verify the conclusion at the root. We use a simple argument based on distribution functions. Let

$$\mathcal{M}_\mu \lambda(\alpha) = \max_{o \leq y \leq \alpha} \frac{I^*(\lambda\mu)(y)}{I^*\mu(y)}$$



be the discrete maximal function we used in [ARS3]. If necessary, we can extend the definition to  $\alpha \in \partial T$  in the obvious way. Then,

$$\begin{aligned} I^* \sigma_{\lambda\mu}(o) &= \sum_{\alpha \in T} \left[ \frac{I^*(\lambda\mu)(\alpha)}{I^*\mu(\alpha)} \right]^2 (I^*\mu(\alpha))^2 \\ &\leq \sum_{\alpha \in T} [\mathcal{M}_\mu \lambda(\alpha)]^2 \sigma_\mu(\alpha) \\ &= 2 \int_0^1 t \sigma_\mu(\zeta \in \bar{T} : \mathcal{M}_\mu \lambda(\zeta) > t) dt. \end{aligned}$$

Now,  $\{\zeta \in \bar{T} : \mathcal{M}_\mu \lambda(\zeta) > t\} = \cup S(\alpha_j)$  is the disjoint union in  $\bar{T}$  (by the definition of the maximal function, we do not need to consider the closure of  $S(\alpha_j)$  in  $\bar{T}$ ). Then,

$$\begin{aligned} t \sigma_\mu(\zeta \in \bar{T} : \mathcal{M}_\mu \lambda(\zeta) > t) &= \sum_j t \sigma_\mu(S(\alpha_j)) \\ &\leq \sum_j t I^*\mu(\alpha_j) \\ &\leq \sum_j I^*(\lambda\mu)(\alpha_j) \\ &\leq I^*(\lambda\mu)(o). \end{aligned}$$

Inserting this estimate in the previous one, we have

$$I^* \sigma_{\lambda\mu}(o) \leq 2 \cdot I^* \sigma_\mu(o).$$

□

We now give a direct proof that the tree condition is equivalent to the capacity condition, Theorem 7 below. That the capacity condition implies the tree condition was noted in [KS] and earlier as Theorem 4 in [Ad]. We proceed to the opposite implication. We need an estimate for measures supported in  $\bar{T}$ .

**Lemma 1.** *We have*

$$\text{Cap}_T(S(E)) \leq 4 \text{Cap}_T(E),$$

for  $E = \cup \partial S(\alpha_j)$ , where  $S(E) = \cup \overline{S(\alpha_j)}$ .

Note that  $\text{Cap}_T(S(E)) = \text{Cap}_T(\{\alpha_j\})$ . Obviously,  $\text{Cap}_T(S(E)) \geq \text{Cap}_T(E)$ .

*Proof.* Let  $\varphi$  be the extremal function for  $\text{Cap}_T(E)$ :

$$I\varphi(\zeta) = 1 \text{ on } E, \quad \text{Cap}_T(E) = \sum \varphi^2.$$

We show that it is near extremal for  $\overline{S(E)}$ .

The function  $\varphi$  can be recovered from the equilibrium measure,  $\varphi = I^*\mu$ , and  $\mu$  is clearly constant on each  $\partial S(\alpha_j)$ : there is  $\Gamma_j > 0$  such that

$$\varphi = \Gamma_j 2^{-d(\beta)} \quad \forall \beta \in S(\alpha_j).$$

Now, for all  $\zeta \in \partial S(\alpha_j)$ ,

$$\begin{aligned} 1 - I\varphi(\alpha_j) &= I\varphi(\zeta) - I\varphi(\alpha) \\ &= \sum_{\beta \in [\alpha_j, \zeta]} \varphi(\beta) \\ &= \Gamma_j \sum_{[\alpha_j, \zeta]} 2^{-d(\beta)} \\ &= \Gamma_j 2^{-d(\alpha_j)} \varphi(\alpha_j). \end{aligned}$$

Hence,  $1 - \varphi(\alpha_j) = I\varphi(\alpha_j)$ . Note that  $\varphi$ , the extremal function, is monotone increasing with respect to the partial ordering in  $T$ , thus  $\varphi(\alpha_j) \geq 1/d(\alpha_j)$ . Hence,

$$I\varphi(\alpha_j) \geq 1 - 1/d(\alpha_j) \geq 1/2.$$

This means that  $2\varphi$  is an admissible function for  $E$ :

$$\text{Cap}_T(S(E)) \leq 4 \text{Cap}_T(E).$$

□

**Theorem 7.** *Let  $\mu$  be a positive, Borel measure on  $\bar{T}$ . Then  $\mu$  satisfies*

$$(5.5) \quad \sup_{\alpha \in T} \frac{I^*[I^*\mu]^2(\alpha)}{I^*\mu(\alpha)} \leq C_1(\mu).$$

*if and only if  $\mu$  satisfies, for all sets  $E = \cup_j \partial S(\alpha_j)$  in  $\partial T$ ,*

$$(5.6) \quad \mu\left(\overline{\cup_j S(\alpha_j)}\right) \leq C_2(\mu) \text{Cap}_T(\cup_j \partial S(\alpha_j)).$$

*Moreover,  $C_2(\mu) \sim C_1(\mu)$ .*

*Proof.* Suppose that  $\mu$  satisfies (5.5) with  $C_1(\mu) = 1$ . Recall that  $\overline{S(E)} = \overline{\cup S(\alpha_j)}$ . Then,  $\mu_E = \mu|_{\overline{S(E)}} \leq \mu$  satisfies  $\|\mu_E\|_{CM(T)} \leq 2$  by Theorem 6. Hence,

$$\begin{aligned} \text{Cap}_T(\overline{S(E)}) &= \sup_{\sup(\nu) \subseteq \overline{S(E)}} \frac{\nu(\overline{S(E)})}{\|\nu\|_{CM(T)}} \\ &\geq \frac{\mu(\overline{S(E)})}{\|\mu_E\|_{CM(T)}} \\ &\geq \frac{\mu(\overline{S(E)})}{2}. \end{aligned}$$

i.e.,  $\mu(\overline{S(E)}) \leq 2 \text{Cap}_T(\overline{S(E)}) \leq 8 \text{Cap}_T(E)$ , by Lemma 1. As mentioned, the opposite inequality is already known. □

## 6. BOUNDARY BEHAVIOR AND EXCEPTIONAL SETS

In this section we give a number of results about boundary behavior and exceptional sets for the dyadic Dirichlet space. In several cases we show that certain behavior occurs with an exceptional set that is a null set for a class of Carleson measures. Then by Corollary 1, or a variation of that corollary, we conclude that the possible exceptional set has capacity zero.

The results we present now are discrete analogs of established results about boundary convergence of smooth functions and about the associated exceptional sets. The literature on those problems is extensive. We offer [A], [AC], [Ca], [DB],

[GP], [Ki], [M], [NRS], [NS], [Tw], and [W] as recent representatives as well as places where the reader can get more information. In particular the versions of our next few results for harmonic functions are in [Tw]. A main theme here is to show that when appropriate oscillation estimates are available then there is a unified approach to such results. In particular this approach highlights the basic geometry of the tree model, or, what is roughly the same thing, the geometry of a Whitney covering of the domain.

**6.1. Boundary Values.** It is not clear at first glance that functions in  $\mathcal{D}_d$  must have boundary values on a large subset of  $\partial T$ . We now establish that with an argument whose basic form is at the core of the later discussions. For each positive integer  $n$  let  $\chi_n$  be the characteristic function of the set  $\{\alpha \in T : d(\alpha) \leq n\}$ . Given  $F \in \mathcal{D}_d$ , set, for each  $n$

$$F_n^* = I(|DF| \cdot \chi_n).$$

It is immediate that the sequence of functions  $\{F_n^*\}$  is increasing in  $n$ , that each function extends by continuity to all of  $\partial T$ , and that each function is in  $\mathcal{D}_d$  and has norm at most  $\|F\|$ . By monotonicity the extended limit  $\lim F_n^* = F^*$  is defined everywhere. It then follows using Fatou's lemma that  $|F^*|^2$  has finite integral with respect to any Carleson measure  $\mu$  for  $\mathcal{D}_d$ . Hence  $F^*$  is finite  $\mu - a.e.$ . With  $F^*$  in hand as a majorant for the variation of  $F$  along each geodesic it is easy to see that  $F$  also has boundary limits  $\mu - a.e.$  Hence, by Corollary 1,  $F$  has finite boundary limits *q.e.*

**6.2. Beurling's Theorem.** The prototypical result in this area is Beurling's theorem that any  $f \in \mathcal{D}$  has radial boundary values *q.e.* He did this by showing that the boundary values of the radial variation,  $V(f)(e^{it})$  were finite *q.e.* and noting that at such points  $f$  must have radial boundary values. In fact the argument in the previous paragraph is essentially a complete proof of a discrete analog of Beurling's theorem. For  $F \in \mathcal{D}_d$  we measure its radial variation by  $V_T(F)(\alpha) = I(|DF|)(\alpha)$ .

**Theorem 8** (Discrete Beurling's Theorem). *For  $F \in \mathcal{D}_d$ ,*

$$\text{Cap}_T(\{\Gamma \in \partial T : V_T(F)(\Gamma) = +\infty\}) = 0.$$

*Hence*

$$\text{Cap}_T(\{\Gamma \in \partial T : \lim_{\alpha \in \Gamma} F(\alpha) \text{ does not exist}\}) = 0.$$

*Proof.* Start with  $F \in \mathcal{D}_d$ . Hence  $DF \in l^2(T)$ , and hence  $|DF| \in l^2(T)$ , and thus  $V_T(F) \in \mathcal{D}_d$ . This insures that the set on which  $V_T(F) = \infty$  is a null set for any  $\mathcal{D}_d$  Carleson measure and hence, by Corollary 1, has vanishing capacity. Finally, one easily checks that if  $V_T(F)(\Gamma) < \infty$  then  $\lim_{\alpha \in \Gamma} F(\alpha)$  exists.  $\square$

**6.3. Algebraic Approach Regions.** We continue to focus on  $\mathcal{D}_d$  and now consider boundary limits through more general approach regions. For any subset  $S \subset T$  we define the boundary limit through  $S$  of a function  $F$  defined on  $T$  by

$$\lim_S F(\beta) = \lim_{\substack{\beta \in S \\ d(\beta) \rightarrow \infty}} F(\beta).$$

At the level of metaphor, convergence along the geodesics  $\Gamma \subset T$  is similar to both radial convergence in the disk and to non-tangential convergence through a narrow

wedge with apex at the boundary point corresponding to  $\Gamma$ . The analog of non-tangential convergence with wider wedges is obtained by also including points that are at most a fixed distance,  $k$ , from  $\Gamma$ . For  $\alpha \in T, n \in \mathbb{N}$  let

$$\{\alpha + n\} = \{(\beta \in T : \beta \succ \alpha, d(\alpha, \beta) = n)\}$$

and we define

$$\Gamma_1(k) = \bigcup_{\alpha \in \Gamma} \{\alpha + k\}.$$

**Theorem 9.** *Fix  $k \geq 1$ . For  $F \in \mathcal{D}_d$ ,  $F$  has  $\Gamma_1(k)$  limits q.e.; that is*

$$\text{Cap}_T(\{\Gamma \in \partial T : \lim_{\Gamma_1(k)} F(\alpha) \text{ does not exist}\}) = 0.$$

*Proof.* For  $F \in \mathcal{D}_d$  we define  $DF^{(1)}$  by

$$DF^{(1)}(\alpha) = |DF(\alpha)| + \sum_{\beta \in \{\alpha+k\}} |DF(\beta)|.$$

Because  $DF \in l^2(T)$  we also have  $DF^{(1)} \in l^2(T)$ , the reason is that each value  $|DF(\delta)|$  shows up as part of  $DF^{(1)}(\delta)$  and as part of at most one other  $DF^{(1)}(\delta')$ . Hence  $G = I(DF^{(1)}) \in \mathcal{D}_d$  and thus  $G$  has boundary values q.e. Suppose now that  $\Gamma$  is a geodesic representing a point in the boundary at which  $G$  has boundary limit  $G(\Gamma)$ . We claim that it follows that  $F$  has a boundary limit through the approach region  $\Gamma_1(k)$ . We are assuming

$$(6.1) \quad \lim_{\Gamma} G(\alpha) - G(\Gamma) = 0,$$

Now pick  $\alpha$  in  $\Gamma$  and let  $\beta, \beta'$  be two points of  $\Gamma_1(k)$  which are further from the root than  $\alpha$ . Let  $\alpha^{-k}$  be the direct ancestor of  $\alpha$  of order  $k$ . Now we have that

$$(6.2) \quad \begin{aligned} |F(\beta) - F(\beta')| &\leq |F(\beta) - F(\alpha)| + |F(\alpha) - F(\beta')| \\ &\leq \sum_{\gamma \in [\alpha, \beta]} |DF(\gamma)| + \sum_{\gamma \in [\alpha, \beta']} |DF(\gamma')| \\ &\leq 4(G(\Gamma) - G(\alpha^{-k})). \end{aligned}$$

To see this last inequality we concentrate on the first sum. The geodesic segment  $[\alpha, \beta]$  has two regions, the first consisting of  $\gamma$ 's in  $\Gamma$ , the second part consisting of the  $\gamma'$ 's not in  $\Gamma$ . In the first case  $|DF(\gamma)| \leq DG(\gamma)$ . In the second case set  $\beta(\gamma') = \gamma'^{-k}$ , the ancestor of  $\gamma'$  of order  $k$ . We have  $\beta(\gamma') \in [\alpha^{-k}, \alpha]$  and thus  $|DF(\gamma')| \leq DG(\beta(\gamma'))$ . Thus the sum of the  $|DF|$  along  $[\alpha, \beta]$  is dominated by twice the sum of the  $DG$  along  $[\alpha^{-k}, \alpha]$ . Finally, because  $G$  is increasing we get (6.2). Furthermore (6.1) insures that the right hand side of (6.2) will tend to zero as  $\alpha$  tends to the boundary; and hence  $F$  has the desired limit through  $\Gamma_1(k)$ .  $\square$

We now consider larger regions. Fix an integer  $k$  and set

$$\Gamma_2(k) = \bigcup_{\alpha \in \Gamma} \{\alpha + kd(\alpha)\}.$$

We will consider limits over the sets  $\Gamma_2(k)$ . To do this we start with  $F$  and construct a majorant for its variation. Define

$$DF^{(2)}(\alpha) = |DF(\alpha)| + \sum_{j=0}^{k-1} \max\{|DF(\beta)| : \beta \in \{\alpha + kd(\alpha) + j\}\}$$

If  $F \in \mathcal{D}_d$  then  $DF \in l^2(T)$ . We now claim that  $DF^{(2)} \in l^2(T)$ . Each  $DF^{(2)}(\alpha)$  is a sum of  $k+1$  terms taken from the square summable sequence  $\{DF(\alpha)\}$  so we only need show that there is an upper bound on how many times an individual element of that latter sequence is used in this construction. However, in fact, no term is used more than twice. The term  $|DF(\beta)|$  does appear in  $DF^{(2)}(\beta)$ . The only other time that term can be used as a summand is in  $DF^{(2)}(\alpha)$  for the unique  $\alpha$  which satisfies the two conditions  $\alpha \in [o, \beta]$  and  $(k+1)d(\alpha) \leq d(\beta) < (k+2)d(\alpha)$ . This observation together with an application of the Cauchy-Schwarz inequality to the sum used in defining  $DF^{(2)}(\alpha)$  shows that  $DF^{(2)} \in l^2(T)$ .

A minor modification of the argument we used before now shows that  $F$  converges along  $\Gamma_2(k)$ . The only required change is where before we backed up from  $\alpha$  to an  $\alpha^{-k}$  now we have to back up from  $\alpha$  to  $\alpha' \in [o, \alpha]$  with  $(k+1)d(\alpha') > d(\alpha)$ .

**Theorem 10.** *Fix  $k \geq 1$ . For  $F \in \mathcal{D}_d$ ,  $F$  has  $\Gamma_2(k)$  limits q.e.; that is*

$$\text{Cap}_T(\{\Gamma \in \partial T : \lim_{\Gamma_2(k)} F(\alpha) \text{ does not exist}\}) = 0.$$

Again at the level of metaphor we can describe the geometry of the regions which correspond to these types of approach. We give the description in the upper half plane with the positive imaginary axis as the geodesic of interest. The geodesic convergence in the tree corresponds to convergence in the Stoltz region  $y > |x|$ . The thickened geodesics  $\Gamma_1(k)$  correspond to the wider, but still non-tangential, approach regions  $y > 2^{-k}|x|$ . The regions  $\Gamma_2(k)$  correspond to regions which are tangent to the boundary with the tangency being of finite order; roughly,  $\Gamma_2(k)$  corresponds to the region  $y > |x|^{k+1}$ . Results on tangential convergence and the size of the associated exceptional sets go back to Kinney [Ki] and more general versions are in [Tw].

**6.4. Beyond Algebraic Approach Regions.** The previous result can be extended to regions which are tangent of infinite order to the boundary but at a cost; the convergence will be quasi-everywhere but now quasi-everywhere with respect to a *different capacity*. The capacities will be those associated with the spaces  $\mathcal{D}_{d,\varepsilon}$ , the Hilbert space of  $F$  functions on  $T$  for which

$$(6.3) \quad \|F\|_{\mathcal{D}_{d,\varepsilon}}^2 = |F(o)|^2 + \sum_T |DF(\alpha)|^2 2^{-\varepsilon d(\alpha)} < \infty.$$

The approach regions of interest are these. For  $0 < \varepsilon \leq 1$  and  $\Gamma$  a geodesic in  $T$  which defines an element of  $\partial T$  we set

$$\Gamma_3(\varepsilon) = \bigcup_{\alpha \in \Gamma} \left\{ \alpha + \left[ 2^{\varepsilon d(\alpha)} \right] \right\}.$$

(Hereafter we will regard the nearest integer brackets as implicit and will not write them.)

We begin by a straightforward modification of the argument which gave Theorem 10. We start with  $F \in \mathcal{D}_d$  and construct a majorant for its variation. Define

$$(6.4) \quad DF^{(3)}(\alpha) = |DF(\alpha)| + \sum_{j=2^{\varepsilon d(\alpha)}}^{j=2 \cdot 2^{\varepsilon d(\alpha)}} \max \{ |DF(\beta)| : \beta \in \{\alpha + j\} \}$$

If  $F \in \mathcal{D}_d$  then  $DF \in l^2(T)$ , but now it need not hold that  $DF^{(3)} \in l^2(T)$ . For each  $\alpha$  let  $\{\beta(\alpha)_j\}$  be the vertices of  $T$  which appear on the right hand side of (6.4);

that is,  $\alpha$  and the selected elements where  $\max |DF(\beta)|$  is attained. Thus

$$DF^{(3)}(\alpha) = \sum_j |DF(\beta(\alpha)_j)|.$$

Hence

$$\begin{aligned} \left| DF^{(3)}(\alpha) \right|^2 &\leq \left( \sum_j |DF(\beta(\alpha)_j)|^2 \right) \cdot (\text{number of } j\text{'s}) \\ &\leq C 2^{\varepsilon d(\alpha)} \left( \sum_j |DF(\beta(\alpha)_j)|^2 \right). \end{aligned}$$

Hence the sequence of numbers  $\left\{ 2^{-\varepsilon d(\alpha)} |DF^{(3)}(\alpha)|^2 \right\}$  is summable because, again, no vertex shows up as a  $\beta(\alpha)_j$  more than a few times. Thus  $DF^{(3)} \in l^2(T, 2^{-\varepsilon d(\alpha)})$ . We now use the same arguments as before. Set  $G = I(DF^{(3)})$ ;  $G$  will have finite radial limits along every geodesic  $\Gamma$  with the possible exception of a set which is a null set for every Carleson measure for the space  $\mathcal{D}_{d,\varepsilon}$ . Also as in the previous proof, any boundary point  $\Gamma$  at which  $I(DF^{(3)})(\Gamma) < \infty$  will be a boundary point where we have good convergence of  $F$ ; in this case the good convergence meaning convergence over  $\Gamma_3(\varepsilon)$ .

The description of the Carleson measures for these spaces is given in [AR]. Here is the description for  $0 < \varepsilon < 1$ .

**Theorem 11.** *Suppose  $0 < \varepsilon \leq 1$ . Let  $\mu$  be a positive Borel measure on  $\bar{T}$ . Then, the following are equivalent:*

- (1)  $\mu$  is a Carleson measure for  $\mathcal{D}_{d,\varepsilon}$ ; i.e. there is a constant  $C$  so that  $\forall f \in \mathcal{D}_{d,\varepsilon}$

$$(6.5) \quad \sum_{\alpha \in \bar{T}} |F(\alpha)|^2 \mu(\alpha) \leq C \|F\|_{\mathcal{D}_{d,\varepsilon}}^2.$$

- (2)  $\mu$  satisfies the  $\varepsilon$ -tree condition. There is a constant  $C$  so that  $\forall \alpha \in T$

$$(6.6) \quad I^* \left[ (I^* \mu)^2(\cdot) 2^{\varepsilon d(\cdot)} \right] (\alpha) \leq C I^* \mu(\alpha).$$

Hence we have

**Theorem 12.** *Fix  $\varepsilon$ ,  $0 < \varepsilon < 1$ . For  $F \in \mathcal{D}_d$ ,  $F$  has  $\Gamma_3(\varepsilon)$  limits for all  $\Gamma \in \partial T$  with the possible exception of a set which is a null set for every measure  $\mu$  which satisfies the condition (6.6).*

In this case the approach regions have infinite order tangency, in fact subexponential contact. The Euclidean analogs of these regions shaped like the part of the upper halfplane where  $y > \exp(-|x|^{-\varepsilon})$ .

Finally if a variant of Theorem 5 is available in this context the result can be reformulated as *q.e.* convergence with respect to the appropriate capacity. In fact such a theorem does hold; its statement and proof are similar to the  $\varepsilon = 0$  case considered earlier; details will be in [ARSp]

**6.5. The Result of Nagel, Rudin, and Shapiro.** It is a result of Nagel, Rudin, and Shapiro [NRS] that, with a possible exceptional set of Lebesgue measure zero, functions in the Dirichlet space approach their radial boundary values through approach regions of full exponential contact; that is, with the shape of the set

$y > \exp(-|x|^{-1})$ . Further work in that direction is in [NS], [DB], and [Tw]. It would be interesting to know if an analogous result holds for  $\mathcal{D}_d$  and we leave that as a question.

The proof just given shows that any  $F \in \mathcal{D}_d$  has limits along regions  $\Gamma_3(1)$  with an exceptional set that is a null set for all the measures which satisfy (6.6) with  $\varepsilon = 1$ . However the full boundary is, in fact, such a set. The quickest way to see that is to note  $G = Ig$  with  $g(a) = d(\alpha)^{-1}$  has boundary values identically  $+\infty$ . We could consider the subspace of  $\mathcal{D}_{d,1}$  consisting of martingales; the Carleson measures for that subspace have null sets which are exactly the sets of Lebesgue, for that see [AR]. However it is not clear how to use that result in this context.

One difference between the two cases is that the proofs for harmonic functions make systematic use of the reconstruction of the interior values of functions from the boundary values; in contrast the values of an  $F \in \mathcal{D}_d$  on  $T$  are not determined by the boundary values.

**6.6. Boundary Convergence for a  $BMO$ -type Space.** The Carleson measures for  $\mathcal{D}$  are the positive measures on the disk which satisfy the equivalent conditions of Theorem 4. There is an interesting subspace  $\mathcal{X}$  of  $\mathcal{D}$  consisting of those  $f \in \mathcal{D}$  which generate Carleson measures in the following sense:

$$\mathcal{X} = \left\{ f \in \mathcal{D} : |f'(z)|^2 dx dy \text{ is a Carleson measure for } \mathcal{D} \right\}.$$

The discrete analog is

$$\mathcal{X}_d = \left\{ F \in \mathcal{D}_d : \mu_F = |DF|^2 \text{ is a Carleson measure for } \mathcal{D}_d \right\}.$$

As discussed briefly in [AR], the space  $\mathcal{X}$  has a relation to  $\mathcal{D}$  similar to the relation the space  $BMO$  has to the Hardy space  $H^2$ . In fact one of the characterizations of functions in  $BMO$  is that  $f \in BMO$  if and only if  $|f'(z)|^2(1 - |z|^2) dx dy$  is a Carleson measure for the Hardy space. The functions in  $BMO$  are both smaller than and smoother than generic functions in  $H^2$ . Similarly functions  $\mathcal{X}$  are smaller than generic functions in  $\mathcal{D}$ . Specifically if  $f \in \mathcal{D}$  then for some small  $\varepsilon_f$  it holds that  $\exp(\varepsilon_f |f|^2)$  has integrable boundary values; however  $f \in \mathcal{X}$  insures that the boundary values of  $\exp(\exp(\varepsilon_f |f|))$  are integrable. Similar results also hold for the model spaces on trees, for all this see [LL]. Here we obtain a different result, but one in the same spirit, the functions in  $\mathcal{X}_d$  have nicer properties than the generic elements of  $\mathcal{D}_d$ .

For comparison recall that Theorem 8 states that

$$(6.7) \quad \forall F \in \mathcal{D}_d \quad \lim_{\alpha \in \Gamma} F(\alpha) = \sum_{\alpha \in \Gamma} DF(\alpha) \text{ exists for quasi-every } \Gamma.$$

Suppose now that  $F \in \mathcal{X}_d$  is fixed  $\mu(\alpha) = \mu_F(\alpha) = |DF(\alpha)|^2$ . Recall that the tree condition for  $\mu$  is that there is a  $C$  so that for all  $\alpha$

$$I^*(I^*\mu)^2(\alpha) \leq CI^*\mu(\alpha).$$

**Theorem 13.** *Suppose  $\mu$  is a Carleson measure  $\mathcal{D}_d$  then*

$$\sum_{\alpha \in \Gamma} \mu(S(\alpha)) < \infty \text{ for quasi-every } \Gamma.$$

Equivalently

$$\int \log^+ \frac{1}{|z - e^{i\theta}|} d\mu(z) < \infty \text{ for quasi-every } \theta.$$

In particular, if  $F \in \mathcal{X}_d$  then

$$\sum_{\alpha \in \Gamma} (I^*(DF)^2)(\alpha) < \infty \text{ for quasi-every } \Gamma.$$

*Proof.* The argument proving Theorem 8 applies to  $Ih$  for any square summable function  $h$  defined on  $T$ . The tree condition evaluated at the origin insures that  $h(\alpha) = I^*\mu(\alpha) = \mu(S(\alpha))$  is such a sequence. That gives the first statement. The second follows from the first by estimating how often each value  $\mu(\beta)$  occurs in the sum. By writing out all the terms in  $(I^*(DF)^2)(\alpha)$  and then discarding those corresponding to vertices not on  $\Gamma$  we obtain a weaker, but more transparent, corollary.  $\square$

**Corollary 3.** *If  $F \in \mathcal{X}_d$  then*

$$\sum_{\alpha \in \Gamma} d(\alpha) |DF|^2(\alpha) < \infty \text{ for quasi-every } \Gamma.$$

These results describe radial convergence and, as with the results for functions in  $\mathcal{D}$ , they can be extended to larger convergence regions for both

$$\lim_{\beta} II^*(|DF|^2)(\beta) \text{ and}$$

$$\lim_w \int \log^+ \frac{1}{|1 - \bar{w}z|} d\mu(z).$$

## 7. POSSIBLE EXTENSIONS

Here we briefly and very informally discuss how some of these ideas will be taken further in [ARSp].

**7.1. Other Function Spaces on Trees.** Various function spaces on  $T$  have been studied both on their own and as discrete models for spaces of smooth functions such as Besov spaces. This view is developed among other places in [Ar], [AR], [ARS1], [ARS2] where in addition to  $l^2(T)$  study is also made of various weighted  $l^p(T)$  spaces. The arguments of the previous sections adapt directly to show that such functions converge to boundary values through various approach regions with exceptional sets that are null sets for classes of Carleson measures. One way to get further insight is to develop geometric characterizations of the relevant classes of Carleson measures. For the spaces mentioned that is done in the earlier work by the authors. To go to results involving capacity we need a result such as Corollary 1. For the function space described in (6.3) the proof we gave in the case  $\alpha = 0$  continues to work with straightforward changes. However for  $p \neq 2$ , for instance for the dyadic Besov spaces of [AR], one needs to work with the nonlinear potential theory appropriate for  $l^p$  spaces and the arguments are more complicated. That work will be presented in [ARSp]



**7.2. Holomorphic, Harmonic, and other Smooth Functions.** Results such as those we described for model spaces such as  $\mathcal{D}_d$  can be used to obtain results for spaces of smooth functions. Suppose for instance that we want to derive a version of Beurling's theorem [Beu].

**Theorem 14** (Beurling, 1940). *For all  $f \in \mathcal{D}$*

$$\text{Cap}_{\mathcal{D}}(\{e^{it} : \lim_r f(re^{i\theta}) \text{ fails to exist}\}) = 0.$$

First select and fix  $f \in \mathcal{D}$ . Also select  $R$  so large that the hyperbolic disks of radius  $R$  centered at points of the tree,  $\{D(\alpha, R) : \alpha \in T\}$  is a cover for  $\mathbb{D}$  and so that for all  $\alpha \in T$  we have  $\{\alpha^-, \alpha_r, \alpha_l\} \subset D(\alpha, R)$ . For each  $\alpha \in T$  we measure the local oscillation of  $f$  by

$$\text{Osc}(\alpha) = \text{Osc}(\alpha, f) = \sup \{|f(z) - f(z')| : z, z' \in D(\alpha, R)\}.$$

Straightforward considerations of the geometry of the placement of  $T$  in  $\mathbb{D}$  show that the disks  $\{D(\alpha, 5R) : \alpha \in T\}$ , have the property that there is an  $M$  so that no point is in more than  $M$  disks. This insures that  $\{\text{Osc}(\alpha)\} \in l^2(T)$  because function theoretic estimates yield

$$\text{Osc}(\alpha)^2 \leq C \int_{D(\alpha, 5R)} |f'|^2 dx dy.$$

The finite overlap of the disks and the definition  $\mathcal{D}$  insures that the integrals on the right can be summed.

Theorem 8 then insures that  $\lim_{\Gamma} I^* \text{Osc}(\alpha)$  is finite except for a set of  $\Gamma$  of  $\mathcal{D}_T$  capacity zero. This insures that  $f$  has a limit along the path connecting the vertices in  $\Gamma$  and that in turn is enough to insure that  $f$  also has a limit along the radius which terminates at the point of the circle corresponding to the boundary element determined by  $\Gamma$ .

This outline gives convergence off of an exceptional set is of  $\mathcal{D}_T$  capacity zero rather than  $\mathcal{D}$  capacity zero. We hope to return to the general question of the relationship between null sets for discrete capacities and for continuous capacities. In this particular case however, the capacities associated with  $\mathcal{D}_T$  and with  $\mathcal{D}$ , the two collections of null sets are known to agree, as is shown by Benjamini and Peres [BP].

A limitation of the proof we outlined is that it established the existence of boundary limits rather than finiteness of the variation functional  $V(f)(e^{i\theta})$ . However that was just for convenience of presentation. A slightly more elaborate definition of oscillation together with a similar argument, but using the fact that functions in  $\mathcal{D}_T$  have nontangential limits, would establish the variation result.

We would like to emphasize that large parts of the previous argument do not involve holomorphy at all. If one has local oscillation estimates and knows that the oscillation numbers live in a space  $X$ , for instance  $X$  could be a weighted  $l^p(T)$ , then the argument shows that limits exist along  $\Gamma$  with an exceptional set of  $\Gamma$  that is a null set for all the Carleson measures for  $X$ . So, for instance, these arguments can certainly be used with harmonic functions or holomorphic functions of several variables. Also, there are other, rather different types of function spaces such as  $\mathcal{A}$ -harmonic functions and monotone Sobolev functions where such oscillation estimates are available; see for instance [KMV], [MV]. These types of variations have not been explored

**7.3. Final Questions.** One of the themes in the study of boundary value results for, say, harmonic functions is consideration of whether the description of the exceptional sets is sharp. That is also a natural question in this context but we haven't considered it. We conclude by mentioning two areas where we do not know if the approach we have been describing can be used but the possibility is intriguing.

First, the study of radial variation for bounded holomorphic (or harmonic,  $p$ -harmonic, etc.) functions on the unit disk (or ball, tree, etc.) is a very active research area. The indications so far are that the results there are deeper than and different from the results for, for instance, general Hardy or Besov spaces. The paper [CFPR] includes some general discussion of the area and references.

Second, although we considered various types of approach regions for boundary convergence, they were all of the same sort, a geodesic  $\Gamma$  with a symmetrical enveloping shell which, in the Euclidean sense, narrowed as the region approached the boundary. These are all versions of having boundary limits along a collection of paths, the geometry of the envelope controlling the type of paths. However one can also consider convergence to boundary values through a collection of sets which is (in some appropriate sense approximately) translation invariant and contains no paths. This theme has a long history, recently it shows up in the alternative approach to the results of Nagel, Rudin, and Shapiro given by Nagel and Stein [NS]. In this more general context it also makes sense to look for descriptions of approach regions that are optimal in various senses. These topics are treated fully by DiBiasi in [DB]. Particularly interesting to us is that a substantial part of the work there proceeds through analysis on model spaces defined on dyadic trees.

#### REFERENCES

- [Ad] D. Adams, On the existence of capacity strong type estimates in  $R^n$ . *Ark. Mat.* 14 (1976), no. 1, 125–140.
- [AH] D. Adams, L. Hedberg, *Function spaces and potential theory*. Grundlehren der Mathematischen Wissenschaften, 314. Springer-Verlag, Berlin, 1996.
- [AC] P. Ahern, W. Cohn, Weighted maximal functions and derivatives of invariant Poisson integrals of potentials. *Pacific J. Math.* 163 (1994), no. 1, 1–16.
- [A] H. Aikawa, Capacity and Hausdorff content of certain enlarged sets. *Mem. Fac. Sci. Eng. Shimane Univ. Ser. B Math. Sci.* 30 (1997), 1–21.
- [AE] H. Aikawa, M. Essén, *Potential theory—selected topics*. Lecture Notes in Mathematics, 1633. Springer-Verlag, Berlin, 1996.
- [Ar] N. Arcozzi, Carleson measures for the analytic Besov spaces: the upper triangle case, *J. Inequal. Pure Appl. Math.* 6 (2005), no. 1, Article 60,
- [AR] N. Arcozzi, R. Rochberg, Topics in dyadic Dirichlet spaces, *New York J. Math.* 10 (2004), 45–67
- [ARS1] N. Arcozzi, R. Rochberg, E. Sawyer, Carleson measures for analytic Besov spaces, *Revista Math. Iberoamericana*, 18, (2002), 443–510.
- [ARS2] N. Arcozzi, R. Rochberg, E. Sawyer, Carleson measures and interpolating sequences for Besov spaces on complex balls. *Mem. Amer. Math. Soc.* 182 (2006), no. 859, vi+163 pp.
- [ARS3] N. Arcozzi, R. Rochberg, E. Sawyer, The Characterization of Carleson measures for analytic Besov spaces: a simple proof, *Complex and Harmonic Analysis*, A. Carbery, P. Duren, D. Khavinson, A. Siskakis, Eds., Destech Publ. 2007, 167-178.
- [ARSp] N. Arcozzi, R. Rochberg, E. Sawyer, Carleson Measures, Capacity, and Exceptional Sets, in preparation.
- [BP] I. Benjamini, Itai; Y. Peres, Random walks on a tree and capacity in the interval. *Ann. Inst. H. Poincaré Probab. Statist.* 28 (1992), no. 4, 557–592.
- [Beu] A. Beurling, Ensembles exceptionnels. (French) *Acta Math.* 72, (1940). 1–13.
- [CFPR] A. Cantón, J. L. Fernández, D. Pestana, J. M. Rodríguez, On harmonic functions on trees. *Potential Anal.* 15 (2001), no. 3, 199-244.

- [Ca] L. Carleson, Selected problems on exceptional sets. Van Nostrand Mathematical Studies, No. 13 D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London 1967
- [DB] F. Di Biase, Fatou type theorems. Maximal functions and approach regions. Progress in Mathematics, 147. Birkhäuser Boston, Inc., Boston, MA, 1998.
- [GP] D. Girela, J. A. Peláez, Boundary behaviour of analytic functions in spaces of Dirichlet type. J. Inequal. Appl. 2006, Art. ID 92795, 12 pp.
- [KV] N. Kalton, I. Verbitsky. Nonlinear equations and weighted norm inequalities. Trans. Amer. Math. Soc. 351 (1999), no. 9, 3441–3497.
- [KS] R. Kerman, E. Sawyer, The trace inequality and eigenvalue estimates for Schrödinger operators. Ann. Inst. Fourier (Grenoble) 36 (1986), no. 4, 207–228.
- [Ki] J. Kinney, Tangential limits of functions of the class  $S_\alpha$ . Proc. Amer. Math. Soc. 14 1963 68–70.
- [KMV] P. Koskela, J. Manfredi, E. Villamor, Regularity theory and traces of  $\mathcal{A}$ -harmonic functions. Trans. Amer. Math. Soc. 348 (1996), no. 2, 755–766.
- [LL] Y. Lin, Thesis, Washington University, in preparation, 2008.
- [MV] J. Manfredi, E. Villamor, Traces of monotone Sobolev functions. J. Geom. Anal. 6 (1996), no. 3, 433–444 (1997).
- [M] Y. Mizuta, Existence of tangential limits for  $\mathcal{A}$ -harmonic functions on half spaces. Potential Anal. 25 (2006), no. 1, 29–36.
- [NRS] A. Nagel, W. Rudin, J. Shapiro, Tangential boundary behavior of functions in Dirichlet-type spaces. Ann. of Math. (2) 116 (1982), no. 2, 331–360.
- [NS] A. Nagel, E. Stein, On certain maximal functions and approach regions. Adv. in Math. 54 (1984), no. 1, 83–106.
- [St] D. Stegenga, Multipliers of the Dirichlet space. Illinois J. Math. 24 (1980), no. 1, 113–139.
- [Tw] J. Twomey, Tangential boundary behaviour of harmonic and holomorphic functions. J. London Math. Soc. (2) 65 (2002), no. 1, 68–84.
- [V] I. Verbitsky, Nonlinear potentials and trace inequalities. The Maz'ya anniversary collection, Vol. 2 (Rostock, 1998), 323–343, Oper. Theory Adv. Appl., 110, Birkhäuser, Basel, 1999.
- [W] D. Walsh, Radial variation of functions in Besov spaces. Publ. Mat. 50 (2006), no. 2, 371–399.

(Arcozzi) DIPARTIMENTO DO MATEMATICA, UNIVERSITA DI BOLOGNA, 40127 BOLOGNA, ITALY  
*E-mail address:* arcozzi@dm.unibo.it

(Rochberg) DEPARTMENT OF MATHEMATICS, WASHINGTON UNIVERSITY, ST. LOUIS, MO 63130, U.S.A.

*E-mail address:* rr@math.wustl.edu

(Sawyer) DEPARTMENT OF MATHEMATICS & STATISTICS, MCMASTER UNIVERSITY, HAMILTON, ONTARIO, L8S 4K1, CANANDA

*E-mail address:* Saw6453CDN@aol.com