Analysis & PDE

Volume 12   No. 4   2019

Tuomas Orponen

ABSOLUTE CONTINUITY AND α-NUMBERS ON THE REAL LINE

msp

ABSOLUTE CONTINUITY AND α-NUMBERS ON THE REAL LINE

TUOMAS ORPONEN

Let μ, ν be Radon measures on $\mathbb{R}$, with μ nonatomic and ν doubling, and write $\mu = \mu_a + \mu_s$ for the Lebesgue decomposition of μ relative to ν. For an interval $I \subset \mathbb{R}$, define $\alpha_{\mu,\nu}(I) := \mathbb{W}_1(\mu_I, \nu_I)$, the Wasserstein distance of normalised blow-ups of μ and ν restricted to I. Let $\mathcal{S}_\nu$ be the square function

$$\mathcal{S}_\nu^2(\mu) = \sum_{I \in \mathcal{D}} \alpha_{\mu,\nu}^2(I)\chi_I,$$

where $\mathcal{D}$ is the family of dyadic intervals of side-length at most 1. I prove that $\mathcal{S}_\nu(\mu)$ is finite μ_a almost everywhere and infinite μ_s almost everywhere. I also prove a version of the result for a nondyadic variant of the square function $\mathcal{S}_\nu(\mu)$. The results answer the simplest "$n = d = 1$" case of a problem of J. Azzam, G. David and T. Toro.

1. Introduction

Wasserstein distance and α-numbers. In this paper, μ and ν are nonzero Radon measures on $\mathbb{R}$. The measure ν is generally assumed to be either *dyadically doubling* or *globally doubling*. Dyadically doubling means that

$$\nu(\hat{I}) \leq C\nu(I), \quad I \in \mathcal{D}, \tag{1.1}$$

where $\mathcal{D}$ is the standard family of dyadic intervals, and $\hat{I}$ is the *parent* of I, that is, the smallest interval in $\mathcal{D}$ strictly containing I. Globally doubling means that $\nu(B(x, 2r)) \leq C\nu(B(x, r))$ for $x \in \mathbb{R}$ and $r > 0$; in particular, this implies spt $\nu = \mathbb{R}$. The main example for ν is the Lebesgue measure $\mathcal{L}$, and the proofs in this particular case would differ little from the ones presented below. No a priori homogeneity is assumed of μ.

Definition 1.2 (Wasserstein distance). I will use the following definition of the (first) Wasserstein distance: given two Radon measures ν_1, ν_2 on [0, 1], set

$$\mathbb{W}_1(\nu_1, \nu_2) := \sup_{\psi} \left| \int \psi \, d\nu_1 - \int \psi \, d\nu_2 \right|,$$

where the sup is taken over all 1-Lipschitz functions $\psi \colon \mathbb{R} \to \mathbb{R}$ which are supported on [0, 1]. Such functions will be called *test functions*. A slightly different — and also quite common — definition would allow the sup to run over all 1-Lipschitz functions $\psi \colon [0, 1] \to \mathbb{R}$. To illustrate the difference, let $\nu_1 = \delta_0$ and $\nu_2 = \delta_1$. Then $\mathbb{W}_1(\nu_1, \nu_2) = 0$, but the alternative definition, say $\widetilde{\mathbb{W}}_1$, would give $\widetilde{\mathbb{W}}_1(\nu_1, \nu_2) = 1$. The main reason for using $\mathbb{W}_1$ instead of $\widetilde{\mathbb{W}}_1$ in this paper is to comply with the definitions in [Azzam et al. 2016; 2017].

As in the paper [Azzam et al. 2016] of J. Azzam, G. David and T. Toro, I make the following definition:

Definition 1.3 (α-numbers). Assume that $I \subset \mathbb{R}$ is an interval. Define

$$\alpha_{\mu,\nu}(I) := \mathbb{W}_1(\mu_I, \nu_I),$$

where μ_I and ν_I are normalised blow-ups of μ and ν restricted to I. More precisely, let T_I be the increasing affine mapping taking $\bar{I}$ to [0, 1], and define

$$\mu_I := \frac{T_{I\sharp}(\mu|_I)}{\mu(I)} \quad \text{and} \quad \nu_I := \frac{T_{I\sharp}(\nu|_I)}{\nu(I)}.$$

If $\mu(I) = 0$ (or $\nu(I) = 0$), define $\mu_I \equiv 0$ (or $\nu_I \equiv 0$).

The quantity defined above is somewhat awkward to work with, as it lacks (see Example 5.2) the following desirable stability property: if $I, J \subset \mathbb{R}$ are intervals of comparable length, and $I \subset J$, then $\alpha_{\mu,\nu}(I) \lesssim \alpha_{\mu,\nu}(J)$. Chiefly for this reason, I also need to consider the following "smooth" α-numbers; the definition below is essentially the same as the one given in [Azzam et al. 2017, Section 5]:

Definition 1.4 (smooth α-numbers). Let $\varphi := \text{dist}(\,\cdot\,, \mathbb{R} \setminus (0, 1))$. For an interval $I \subset \mathbb{R}$, define $\alpha_{s,\mu,\nu}(I) := \mathbb{W}_1(\mu_{\varphi,I}, \nu_{\varphi,I})$, where

$$\mu_{\varphi,I} := \frac{T_{I\sharp}(\mu|_I)}{\mu(\varphi_I)} \quad \text{and} \quad \nu_{\varphi,I} := \frac{T_{I\sharp}(\nu|_I)}{\nu(\varphi_I)}.$$

Here T_I is the map from Definition 1.3, $\varphi_I = \varphi \circ T_I$, and $\mu(\varphi_I) = \int \varphi_I \, d\mu$. If $\mu(\varphi_I) = 0$ (or $\nu(\varphi_I) = 0$), set $\mu_{\varphi,I} \equiv 0$ (or $\nu_{\varphi,I} \equiv 0$).

The only difference between the numbers $\alpha_{\mu,\nu}(I)$ and $\alpha_{s,\mu,\nu}(I)$ is in the normalisation of the measures μ_I, φ_I and $\mu_{\varphi,I}, \nu_{\varphi,I}$: if I is closed, the measures μ_I, ν_I are probability measures on [0, 1], while $\mu_{\varphi,I}([0, 1]) = \mu(I)/\mu(\varphi_I)$. The numbers $\alpha_{s,\mu,\nu}(I)$ enjoy the stability property alluded to above. Moreover, if either μ or ν is a doubling, one has $\alpha_{s,\mu,\nu}(I) \lesssim \alpha_{\mu,\nu}(I)$. These facts are contained in Proposition 5.4 (or see [Azzam et al. 2017, Section 5]).

Remark 1.5. The α-numbers were first introduced by X. Tolsa [2009], where he used them to characterise the uniform rectifiability of Ahlfors–David regular measures in $\mathbb{R}^d$. Tolsa's original definition of the

α-numbers has a different, asymmetric, normalisation compared to either $\alpha_{\mu,v}$ or $\alpha_{s,\mu,v}$ above; see [Tolsa 2009, p. 394].

Main results. Before explaining the results in [Azzam et al. 2016], and their connection to the current manuscript, I emphasise that that paper treats "n-dimensional" measures in $\mathbb{R}^d$ for any $1 \leq n \leq d$. For the current paper, only the case $n = d = 1$ is relevant. So, to avoid digressing too much, I need to state the results of [Azzam et al. 2016] in far smaller generality than they deserve.

With this proviso in mind, the main results of [Azzam et al. 2016] imply the following: if μ is a doubling measure on $\mathbb{R}$, and the numbers $\alpha_{\mu,\mathcal{L}}$ satisfy a Carleson condition of the form

$$\int_{B(x,2r)} \int_0^{2r} \alpha_{\mu,\mathcal{L}}(B(y,t)) \frac{dt \, d\mu y}{t} \leq C\mu(B(x,r)), \tag{1.6}$$

then μ, or at least a large part of μ, is absolutely continuous with respect to $\mathcal{L}$, with quantitative upper and lower bounds on the density. As the authors of [Azzam et al. 2016] point out, the main shortcoming of their result is that condition (1.6) imposes a hypothesis on the first powers of the numbers $\alpha_{\mu,\mathcal{L}}$, whereas evidence suggests (see in that paper Remark 1.6.1, the discussion after Theorem 1.7, and Example 4.6) that the correct power should be 2. More support for this belief comes from the following "converse" result [Tolsa 2015, Lemma 2.2]: if μ is a finite Borel measure on $\mathbb{R}$ then

$$\int_0^\infty \tilde{\alpha}_{\mu,\mathcal{L}}^2(x,r) \frac{dr}{r} < \infty \quad \text{for } \mathcal{L} \text{ a.e. } x \in \mathbb{R}. \tag{1.7}$$

In particular, if $\mu \ll \mathcal{L}$, then (1.7) holds for μ almost every $x \in \mathbb{R}$. I should again mention that this is only the easiest $n = d = 1$ case of Tolsa's result. Here $\tilde{\alpha}_{\mu,\mathcal{L}}$ is a variant of the α-number (in fact the one discussed in Remark 1.5).

The purpose of this paper is to address the problem of Azzam, David and Toro in the simplest case $n = d = 1$. I show that control for the second powers of the $\alpha_{\mu,\mathcal{L}}$-numbers does guarantee absolute continuity with respect to Lebesgue measure. In fact, the doubling assumption on μ can be dropped, the Carleson condition (1.6) can be relaxed considerably, and the results remain valid, if $\mathcal{L}$ is replaced by any doubling measure v. The results below also contain the "converse" statement, analogous to (1.7).

I prove two variants of the main result: one dyadic, and the other nondyadic. Here is the dyadic version:

Theorem 1.8. *Let $\mathcal{D}$ be the family of dyadic subintervals of $[0, 1)$, and let μ, v be Borel probability measures on $[0, 1)$. Assume that μ does not charge the boundaries of intervals in $\mathcal{D}$, and v is dyadically doubling. Write $\mu = \mu_a + \mu_s$ for the Lebesgue decomposition of μ relative to v, where $\mu_a \ll v$ and $\mu_s \perp v$. Finally, let $\mathcal{S}_{\mathcal{D},v}(\mu)$ be the square function*

$$\mathcal{S}_{\mathcal{D},v}^2(\mu) = \sum_{I \in \mathcal{D}} \alpha_{\mu,v}^2(I) \chi_I.$$

Then,

(a) *$\mathcal{S}_v(\mu)$ is finite μ_a almost surely, and*

(b) *$\mathcal{S}_v(\mu)$ is infinite μ_s almost surely.*

In particular,

$$\sum_{I \in \mathcal{D}} \alpha_{\mu,\nu}^2(I)\mu(I) < \infty \quad \Longrightarrow \quad \mu \ll \nu.$$

Heuristically, this corresponds to assuming (1.6) at the scale $r = 1$, but I could not find a way to *reduce* the continuous problem to the dyadic one; on the other hand, a reduction in the other direction does not appear straightforward either, so perhaps one needs to treat the cases separately. A caveat of the dyadic set-up is the "nonatomicity" hypothesis on μ. It cannot be dispensed with: for instance, if $\mu = \delta_x$ for any $x \in [0, 1)$, which only belongs to the interiors of finitely many dyadic intervals, then $\mathcal{S}_{\mathcal{D},\mathcal{L}}(\mu)$ is uniformly bounded (for instance $\mathcal{S}_{\mathcal{D},\mathcal{L}}(\delta_0) \equiv 0$), but $\mu \perp \mathcal{L}$.

Here is the nondyadic version of the main theorem:

Theorem 1.9. *Assume that μ, ν are Radon measures, and ν is globally doubling. Write $\mu = \mu_a + \mu_s$, as in Theorem 1.8. Let $\mathcal{S}_\nu$ be the square function*

$$\mathcal{S}_\nu^2(\mu)(x) = \int_0^1 \alpha_{s,\mu,\nu}^2(B(x,r))\, \frac{dr}{r}, \quad x \in \mathbb{R},$$

defined via the smooth α-numbers $\alpha_{s,\mu,\nu}$. Then,

(a) $\mathcal{S}_\nu(\mu)$ *is finite μ_a almost surely, and*

(b) $\mathcal{S}_\nu(\mu)$ *is infinite μ_s almost surely.*

Recall that $\alpha_{s,\mu,\nu}(B(x,r)) \lesssim \alpha_{\mu,\nu}(B(x,r))$ whenever ν is doubling, such as $\nu = \mathcal{L}$; see Proposition 5.4. So, Theorem 1.9 has the following corollary:

Corollary 1.10. *Assume that μ, ν are Radon measures, and ν is globally doubling. If*

$$\int_0^1 \alpha_{\mu,\nu}^2(B(x,t))\, \frac{dt}{t} < \infty \tag{1.11}$$

for μ almost every $x \in \mathbb{R}$, then $\mu \ll \nu$.

The following question remains open:

Question 1. *In the setting of Theorem 1.9 and Corollary 1.10, is the square function in (1.11) finite μ_a almost everywhere?*

The difficulties arise from the nonstability of the numbers $\alpha_{\mu,\nu}$. See [Azzam et al. 2017, Section 5], and in particular Lemma 5.3 of that paper, for related discussion.

Assuming the full Carleson condition (1.6), and that μ is globally doubling, the authors of [Azzam et al. 2016] prove something more quantitative than $\mu \ll \mathcal{L}$; see in particular Theorem 1.9 of that paper. The same ought to be true for the second powers of the α-numbers, and indeed the following result can be easily deduced with the method of the current paper:

Theorem 1.12. *Assume that μ, ν are Borel probability measures on $[0, 1)$, both dyadically doubling, and assume that the Carleson condition*

$$\sum_{I \subset J} \alpha_{\mu,\nu}^2(I)\mu(I) \le C\mu(J), \quad J \in \mathcal{D}, \tag{1.13}$$

holds for some $C \geq 1$. Then μ belongs to $A_\infty^{\mathcal{D}}(\nu)$, the dyadic A_∞ class relative to ν. Similarly, if μ, ν are Radon measures on $\mathbb{R}$, both globally doubling, and the Carleson condition (1.6) holds for the second powers $\alpha_{\mu,\nu}^2(B(y, t))$, then $\mu \in A_\infty(\nu)$.

The a priori doubling assumptions cannot be omitted (that is, they are not implied by the Carleson condition): Just consider $\mu = 2\chi_{[0,1/2)} d\mathcal{L}$. It is clear that the Carleson condition (1.13) holds for the numbers $\alpha_{\mu,\mathcal{L}}^2(I)$, but nevertheless $\mu \notin A_\infty^{\mathcal{D}}(\mathcal{L}|_{[0,1]})$.

Outline of the paper, and the main steps of the proofs. The main substance of the paper is proving the dyadic result, Theorem 1.8, and in particular part (b). This work takes up Sections 2-4. The proof of part (a) is simpler, and closely follows a previous argument of Tolsa — namely the one used to prove (1.7). The details (both in the dyadic and continuous settings) are given in Section 6. Modifications required to prove part (b) of the "continuous" Theorem 1.9 are outlined in Section 5.

The proof of Theorem 1.8(b) has three main steps. First, the numbers $\alpha_{\mu,\nu}(I)$ are used to control something analyst-friendlier, namely the dyadic variants

$$\Delta_{\mu,\nu}(I) = \left| \frac{\mu(I_-)}{\mu(I)} - \frac{\nu(I_-)}{\nu(I)} \right|. \tag{1.14}$$

Here I_- stands for the left half of I. This would be simple, if $\chi_{[0,1/2)}$ happened to be one of the admissible test functions ψ in the definition of $\mathbb{W}_1$. It is not, however, and in fact there seems to be no direct (and sufficiently efficient) way to control $\Delta_{\mu,\nu}(I)$ by $\alpha_{\mu,\nu}(I)$, or even $\alpha_{\mu,\nu}(3I)$. However, it turns out that the numbers are equivalent at the level of certain Carleson sums over trees; proving this statement is the main content of Section 2.

The numbers $\Delta_{\mu,\nu}(I)$ are well-known quantities: they are the (absolute values of the) coefficients in an orthogonal representation of μ in terms of ν-adapted Haar functions, and it is known that they can be used to characterise A_∞. The following theorem is due to S. Buckley [1993]:

Theorem 1.15 [Buckley 1993, Theorem 2.2(iii)]. *Let μ, ν be dyadically doubling Borel probability measures on $[0, 1]$. Then $\mu \in A_\infty^{\mathcal{D}}(\nu)$ if and only if*

$$\sum_{I \subset J} \Delta_{\mu,\nu}^2(I)\mu(I) \leq C\mu(J), \quad J \in \mathcal{D}. \tag{1.16}$$

The result in [Buckley 1993] is only stated for $\nu = \mathcal{L}|_{[0,1]}$, but the proof works in the greater generality. Note the similarity between the Carleson conditions (1.16) and (1.13): The dyadic part of Theorem 1.12 is, in fact, nothing but a corollary of Buckley's result, assuming that one knows how to control the numbers $\Delta_{\mu,\nu}(I)$ by the numbers $\alpha_{\mu,\nu}(I)$ at the level of Carleson sums; consequently, the short proof of this half of Theorem 1.12 can be found in Section 2. The continuous version is discussed briefly in Remark 5.19.

Buckley's result is not applicable for Theorem 1.8: the measure μ is not dyadically doubling, and the information available is much weaker than the Carleson condition (1.13). Handling these issues constitutes the remaining two steps in the proof: all dyadic intervals are split into trees, where μ is "tree-doubling" (Section 4), and the absolute continuity of μ with respect to ν is studied in each tree separately (Section 3).

2. Comparison of α-numbers and Δ-numbers

In this section, μ and ν are Borel probability measures on $[0, 1)$, μ does not charge the boundaries of dyadic intervals, and ν is dyadically doubling inside $[0, 1)$:

$$\nu(\hat{I}) \leq D_\nu \nu(I), \quad I \in \mathcal{D} \setminus \{[0, 1)\}.$$

This implies, in particular, that $\nu(I) > 0$ for all $I \in \mathcal{D}$ with $I \subset [0, 1)$. The main task of the section is to bound the numbers $\Delta_{\mu,\nu}(I)$ by the numbers $\alpha_{\mu,\nu}(I)$, where $\Delta_{\mu,\nu}(I)$ is the quantity

$$\Delta_{\mu,\nu}(I) = \left| \frac{\mu(I_-)}{\mu(I)} - \frac{\nu(I_-)}{\nu(I)} \right| = \left| \int \chi_{(0,1/2)} \, d\mu_I - \int \chi_{(0,1/2)} \, d\nu_I \right|.$$

The task would be trivial if $\chi_{(0,1/2)}$ were a 1-Lipschitz function vanishing at the boundary of $[0, 1]$. It is not: in fact, the difference between $\Delta_{\nu_1,\nu_2}(I)$ and $\alpha_{\nu_1,\nu_2}(I)$ can be rather large for a given interval I.

Example 2.1. If $\nu_1 = \delta_{1/2-1/n}$ and $\nu_2 = \delta_{1/2+1/n}$, then $\Delta_{\nu_1,\nu_2}([0, 1)) = 1$, but $\alpha_{\nu_1,\nu_2}([0, 1)) \lesssim 1/n$. These measures do not satisfy the assumptions of the section, so consider also the following example. Let $\mu = f \, d\mathcal{L}$, where f takes the value 1 everywhere, except in the 2^{-n}-neighbourhood of $\frac{1}{2}$. Let $f \equiv \frac{1}{2}$ on the interval $\left[\frac{1}{2} - 2^{-n}, \frac{1}{2} \right]$, and $f \equiv \frac{3}{2}$ on the interval $\left(\frac{1}{2}, \frac{1}{2} + 2^{-n} \right]$. Then μ is a dyadically 4-doubling probability measure on $[0, 1]$, $\Delta_{\mu,\mathcal{L}}([0, 1)) \sim 2^{-n}$, and $\alpha_{\mu,\mathcal{L}}([0, 1)) \sim 2^{-2n}$.

Fortunately, "pointwise" estimates between $\Delta_{\mu,\nu}(I)$ and $\alpha_{\mu,\nu}(I)$ are not really needed in this paper, and it turns out that certain sums of these numbers are comparable, up to a manageable error. To state such results, I need to introduce some terminology. A family $\mathcal{C} \subset \mathcal{D}$ of dyadic intervals is called *coherent* if the implication

$$Q, R \in \mathcal{C} \text{ and } Q \subset P \subset R \quad \Longrightarrow \quad P \in \mathcal{C}$$

holds for all $Q, P, R \in \mathcal{D}$.

Definition 2.2 (trees, leaves, boundary). A *tree* $\mathcal{T} \subset \mathcal{D}$ is any coherent family of dyadic intervals with a unique largest interval, $\mathrm{Top}(\mathcal{T}) \in \mathcal{T}$, and with the property that

$$\mathrm{card}(\mathrm{ch}(I) \cap \mathcal{T}) \in \{0, 2\}, \quad I \in \mathcal{T}.$$

For the tree $\mathcal{T}$, define the set family $\mathrm{Leaves}(\mathcal{T})$ to consist of the minimal intervals in $\mathcal{T}$, in other words those $I \in \mathcal{T}$ with $\mathrm{card}(\mathrm{ch}(I) \cap \mathcal{T}) = 0$. Abusing notation, I often write $\mathrm{Leaves}(\mathcal{T})$ also for the set $\bigcup \{I : I \in \mathrm{Leaves}(\mathcal{T})\}$. Finally, define the *boundary* of the tree $\partial\mathcal{T}$ by

$$\partial\mathcal{T} := \mathrm{Top}(\mathcal{T}) \setminus \mathrm{Leaves}(\mathcal{T}).$$

Then $x \in \partial\mathcal{T}$ if and only if $x \in \mathrm{Top}(\mathcal{T})$ and all intervals $I \in \mathcal{D}$ with $x \in I \subset \mathrm{Top}(\mathcal{T})$ are contained $\mathcal{T}$.

Definition 2.3 (($\mathcal{T}, D$)-doubling measures). A Borel probability measure μ on $[0, 1]$ is called ($\mathcal{T}, D$)-doubling if

$$\mu(\hat{I}) \leq D\mu(I), \quad I \in \mathcal{T} \setminus \mathrm{Top}(\mathcal{T}).$$

Here is the main result of this section:

Proposition 2.4. *Let μ, ν be measures satisfying the assumptions of the section, and let $\mathcal{T} \subset \mathcal{D}$ be a tree. Moreover, assume that μ is $(\mathcal{T}, D)$-doubling for some constant $D \geq 1$. Then*

$$\sum_{I \in \mathcal{T}} \Delta_{\mu,\nu}^2(I)\mu(I) \lesssim_{D_\nu, D} \sum_{I \in \mathcal{T} \backslash \text{Leaves}(\mathcal{T})} \alpha_{\mu,\nu}^2(I)\mu(I) + \mu(\text{Top}(\mathcal{T})).$$

The "dyadic part" of Theorem 1.12 is an immediate corollary:

Proof of Theorem 1.12, dyadic part. By hypothesis, both measures μ and ν are $(\mathcal{D}, C)$-doubling. Hence, by the Carleson condition (1.13), and Proposition 2.4 applied to the trees $\mathcal{T}_J := \{I \in \mathcal{D} : I \subset J\}$, one has

$$\sum_{I \subset J} \Delta_{\mu,\nu}^2(I)\mu(I) \lesssim_C \sum_{I \subset J} \alpha_{\mu,\nu}^2(I)\mu(I) + \mu(J) \lesssim \mu(J).$$

This is precisely the condition in Buckley's result, Theorem 1.15, so $\mu \in A_\infty^{\mathcal{D}}(\nu)$. $\qquad\square$

I then begin the proof of Proposition 2.4. It would, in fact, suffice to assume that ν is also just $(\mathcal{T}, D_\nu)$-doubling, but checking this would result in some unnecessary book-keeping below. The proof is based on the observation that $\chi_{(0,1/2)}$ can be written as a series of Lipschitz functions, each supported on subintervals of $[0, 1]$. This motivates the following considerations.

Assume that

$$\Psi := \Psi_0 := \sum_{j \geq 0} \psi_j$$

is a bounded function such that each $\psi_j \colon \mathbb{R} \to [0, \infty)$ is an L_j-Lipschitz function supported on some interval $I_j \in \mathcal{D}_j$. Assume moreover that the intervals I_j are nested: $[0, 1) \supset I_1 \supset I_2 \supset \cdots$. Then, as a first step in proving Proposition 2.4, I claim that

$$\left| \int \Psi \, d\mu - \int \Psi \, d\nu \right|$$

$$\leq \sum_{k=0}^{N} \frac{L_k}{2^k} \alpha_{\mu,\nu}(I_k)\mu(I_k) + \sum_{k=0}^{N} \left(\frac{1}{\nu(I_{k+1})} \int \Psi_{k+1} \, d\nu \right) \Delta_{\mu,\nu}(I_k)\mu(I_k) + 2\|\Psi\|_\infty \mu(I_{N+1}) \quad (2.5)$$

for any $N \in \{0, 1, \ldots, \infty\}$, where

$$\Psi_k := \sum_{j \geq k} \psi_j, \quad m \geq 0.$$

For $N = \infty$, the symbol "I_{N+1}" should be interpreted as the intersection of all the intervals I_j. I will first verify that, for any $m \geq 0$,

$$\left| \frac{1}{\mu(I_m)} \int \Psi_m \, d\mu - \frac{1}{\nu(I_m)} \int \Psi_m \, d\nu \right| \leq \frac{L_m}{2^m} \alpha_{\mu,\nu}(I_m) + \left(\frac{1}{\nu(I_{m+1})} \int \Psi_{m+1} \, d\nu \right) \Delta_{\mu,\nu}(I_m)$$

$$+ \frac{\mu(I_{m+1})}{\mu(I_m)} \left| \frac{1}{\mu(I_{m+1})} \int \Psi_{m+1} \, d\mu - \frac{1}{\nu(I_{m+1})} \int \Psi_{m+1} \, d\nu \right|, \quad (2.6)$$

from which it will be easy to derive (2.5). If $\mu(I_m) = 0$, the corresponding term should be interpreted as "0" (recall that $\nu(I_m)$ is never zero by the doubling assumption). The proof of (2.6) is straightforward.

First, note that since $\psi_m \colon \mathbb{R} \to \mathbb{R}$ is an L_m-Lipschitz function supported on I_m, and $|I_m| = 2^{-m}$, one has

$$\left| \frac{1}{\mu(I_m)} \int \psi_m \, d\mu - \frac{1}{\nu(I_m)} \int \psi_m \, d\nu \right| = \left| \int \psi_m \circ T_{I_m}^{-1} \, d\mu_{I_m} - \int \psi_m \circ T_{I_m}^{-1} \, d\nu_{I_m} \right| \leq \frac{L_m}{2^m} \alpha_{\mu,\nu}(I_m).$$

(The mappings T_I are familiar from Definition 1.3). This gives rise to the first term in (2.6). What remains is bounded by

$$\left| \frac{1}{\mu(I_m)} \int \Psi_{m+1} \, d\mu - \frac{1}{\nu(I_m)} \int \Psi_{m+1} \, d\nu \right| \leq \frac{\mu(I_{m+1})}{\mu(I_m)} \left| \frac{1}{\mu(I_{m+1})} \int \Psi_{m+1} \, d\mu - \frac{1}{\nu(I_{m+1})} \int \Psi_{m+1} \, d\nu \right|$$
$$+ \left(\frac{1}{\nu(I_{m+1})} \int \Psi_{m+1} \, d\nu \right) \left| \frac{\mu(I_{m+1})}{\mu(I_m)} - \frac{\nu(I_{m+1})}{\nu(I_m)} \right|.$$

This is (2.6), observing that

$$\Delta_{\mu,\nu}(I_m) = \left| \frac{\mu(I_{m+1})}{\mu(I_m)} - \frac{\nu(I_{m+1})}{\nu(I_m)} \right|,$$

since either $I_{m+1} = (I_m)_+$ or $I_{m+1} = (I_m)_-$, and both possibilities give the same number $\Delta_{\mu,\nu}(I_m)$. Finally, (2.5) is obtained by repeated application of (2.6). By induction, one can check that N iterations of (2.6) (starting from $m = 0$, and recalling that μ, ν are probability measures on $[0, 1)$) leads to

$$\left| \int \Psi \, d\mu - \int \Psi \, d\nu \right| \leq \sum_{k=0}^{N} \frac{L_k}{2^k} \alpha_{\mu,\nu}(I_k) \mu(I_k) + \sum_{k=0}^{N} \left(\frac{1}{\nu(I_{k+1})} \int \Psi_{k+1} \, d\nu \right) \Delta_{\mu,\nu}(I_k) \mu(I_k)$$
$$+ \mu(I_{N+1}) \left| \frac{1}{\mu(I_{N+1})} \int \Psi_{N+1} \, d\mu - \frac{1}{\nu(I_{N+1})} \int \Psi_{N+1} \, d\nu \right|. \quad (2.7)$$

This gives (2.5) immediately, observing that $\|\Psi_{N+1}\|_\infty \leq \|\Psi\|_\infty$.

Now, it is time to specify the functions ψ_j. I first define a hands-on Whitney decomposition for $(0, \frac{1}{2})$. Pick a small parameter $\tau > 0$, to be specified later, and let $U_0 := [\tau, \frac{1}{2} - \tau)$. Then, set $U_{-k} := [\tau 2^{-k}, \tau 2^{-k+1})$ and $U_k := \frac{1}{2} - U_{-k}$ for $k \geq 1$. Let $\{\psi_k\}_{k \in \mathbb{Z}}$ be a partition of unity subordinate to slightly enlarged versions of the sets U_k, $k \in \mathbb{Z}$. By this, I first mean that each ψ_k is nonnegative and L_k-Lipschitz with

$$L_k \leq \frac{C 2^{|k|}}{\tau}. \quad (2.8)$$

Second, the supports of the functions ψ_k should satisfy $\psi_0 \subset [\frac{1}{2}\tau, \frac{1}{2} - \frac{1}{2}\tau)$,

$$\operatorname{spt} \psi_{-k} \subset [(\tfrac{1}{2}\tau) 2^{-k}, 2\tau 2^{-k+1}) \subset (0, 2\tau 2^{-k+1}) \quad \text{and} \quad \psi_k \subset \tfrac{1}{2} - (0, 2\tau 2^{-k+1})$$

for $k \geq 1$. Third,

$$\sum_{k \in \mathbb{Z}} \psi_k = \chi_{(0,1/2)}.$$

Let $\Psi^- := \sum_{k>0} \psi_{-k} + \frac{1}{2}\psi_0$ and $\Psi^+ := \sum_{k>0} \psi_k + \frac{1}{2}\psi_0$. Then

$$\Delta_{\mu,\nu}([0,1)) \leq \left| \int \Psi^- \, d\mu - \int \Psi^- \, d\nu \right| + \left| \int \Psi^+ \, d\mu - \int \Psi^+ \, d\nu \right|. \quad (2.9)$$

This is the only place in the paper where the assumption of μ not charging the boundaries of dyadic intervals is used (however, the estimate (2.9) will eventually be applied to all the measures μ_I, $I \in \mathcal{D}$, so the full strength of the hypothesis is needed). The function Ψ^- is precisely of the form treated above with $I_j := [0, 2^{-j})$, since clearly spt $\psi_{-k} \subset I_k$. Applying the inequality (2.5) with any $N_1 \in \{0, 1, \ldots, \infty\}$ yields

$$\left| \int \Psi^- \, d\mu - \int \Psi^- \, d\nu \right|$$

$$\leq \sum_{k=0}^{N_1} \frac{L_{-k}}{2^k} \alpha_{\mu,\nu}(I_k) \mu(I_k) + \sum_{k=0}^{N_1} \left(\frac{1}{\nu(I_{k+1})} \int \Psi_{k+1}^- \, d\nu \right) \Delta_{\mu,\nu}(I_k) \mu(I_k) + 2\mu(I_{N_1+1}). \quad (2.10)$$

Next, observe that each function Ψ_{k+1}^-, $k \geq 0$, is bounded by 1 and vanishes outside

$$\bigcup_{j=k+1}^{\infty} \text{spt } \psi_{-k} \subset (0, 2\tau 2^{-k}).$$

It follows that

$$\frac{1}{\nu(I_{k+1})} \int \Psi_{k+1}^- \, d\nu \leq \frac{\nu((0, 2\tau 2^{-k}))}{\nu(I_{k+1})} = o_{D_\nu}(\tau),$$

where the implicit constants only depend on the dyadic doubling constant D_ν of ν. In the sequel, I assume that τ is so small that $o_{D_\nu}(\tau) \leq \kappa$, where $\kappa > 0$ is another small constant, which will eventually depend on the $(\mathcal{T}, D)$-doubling constant D for μ. Recalling also (2.8), the estimate (2.10) then becomes

$$\left| \int \Psi^- \, d\mu - \int \Psi^- \, d\nu \right| \leq \frac{C}{\tau} \sum_{k=0}^{N_1} \alpha_{\mu,\nu}(I_k) \mu(I_k) + \kappa \sum_{k=0}^{N_1} \Delta_{\mu,\nu}(I_k) \mu(I_k) + 2\mu(I_{N_1+1}). \quad (2.11)$$

The last term simply vanishes if $N_1 = \infty$, because $\mu(\{0\}) = 0$. A heuristic point to observe is that the left-hand side is roughly $\Delta_{\mu,\nu}([0, 1])$; the right-hand side also contains the same term, but multiplied by a small constant $\kappa > 0$. This gain is "paid for" by the large constant C/τ.

Next, the estimate is replicated for Ψ^+. This time, the inequality (2.5) is applied to the sequence $\tilde{I}_0 = [0, 1)$, $\tilde{I}_1 = \left[0, \frac{1}{2}\right)$, $\tilde{I}_2 = (\tilde{I}_1)_+$, and in general $\tilde{I}_{k+1} = (\tilde{I}_k)_+$ for $k \geq 1$ (here J_+ is the right half of J). Then, if τ is small enough, it is again clear that spt $\psi_k \subset \tilde{I}_k$. Thus, by inequality (2.5),

$$\left| \int \Psi^+ \, d\mu - \int \Psi^+ \, d\nu \right|$$

$$\leq \sum_{k=0}^{N_2} \frac{L_k}{2^k} \alpha_{\mu,\nu}(\tilde{I}_k) \mu(\tilde{I}_k) + \sum_{k=0}^{N_2} \left(\frac{1}{\nu(\tilde{I}_{k+1})} \int \Psi_{k+1}^+ \, d\nu \right) \Delta_{\mu,\nu}(\tilde{I}_k) \mu(\tilde{I}_k) + 2\mu(\tilde{I}_{N_2+1}) \quad (2.12)$$

for any $N_2 \geq 0$. As before, the term $\mu(\tilde{I}_{N_2})$ vanishes for $N_2 = \infty$ $\left(\text{because } \mu\left(\left\{\frac{1}{2}\right\}\right) = 0\right)$, and one can ensure

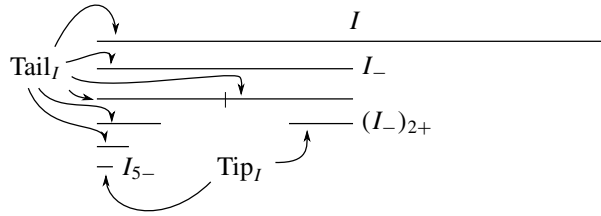$$\frac{1}{\nu(\tilde{I}_{k+1})} \int \Psi_{k+1}^+ \, d\nu \leq \kappa$$

Figure 1. An example of $\mathrm{Tail}_I(4, 1)$ and Tip_I.

by choosing $\tau = \tau(D_\nu) > 0$ small enough. Consequently (recalling (2.9)), (2.11) and (2.12) together imply

$$\Delta_{\mu,\nu}([0, 1)) \leq \frac{C}{\tau} \sum_{I \in \mathrm{Tail}} \alpha_{\mu,\nu}(I)\mu(I) + \kappa \sum_{I \in \mathrm{Tail}} \Delta_{\mu,\nu}(I)\mu(I) + 2\mu(I_{N_1+1}) + 2\mu(\tilde{I}_{N_2+1}). \qquad (2.13)$$

Here Tail is the collection of all the intervals $I_0, \ldots, I_{N_1}$ and $\tilde{I}_0, \ldots, \tilde{I}_{N_2}$. The intervals $[0, 1)$ and $\left[0, \frac{1}{2}\right)$ arise a total of two times from (2.11) and (2.12), but this has no visible impact on the end result, (2.13). The estimate (2.13) generalises in a simple way to other intervals $I \in \mathcal{D}$, besides $I = [0, 1)$, but requires an additional piece of notation. Let $I \in \mathcal{D}$, and write $I_{0-} := I =: I_{0+}$. For $k \geq 1$, define $I_{k-} := (I_{(k-1)-})_-$ and $I_{k+} := (I_{(k-1)+})_+$. Now, for a fixed dyadic interval $I \subset [0, 1)$, and $N_1, N_2 \geq 0$, let $\mathrm{Tail}_I = \mathrm{Tail}_I(N_1, N_2)$ be the collection of subintervals of I, which includes I_{k-} for all $0 \leq k \leq N_1$ and $(I_-)_{k+}$ for all $0 \leq k \leq N_2$; see Figure 1. Then, the generalisation of (2.13) reads

$$\Delta_{\mu,\nu}(I)\mu(I) \leq \frac{C}{\tau} \sum_{J \in \mathrm{Tail}_I} \alpha_{\mu,\nu}(J)\mu(J) + \kappa \sum_{J \in \mathrm{Tail}_I} \Delta_{\mu,\nu}(J)\mu(J) + 2\mu(\mathrm{Tip}_I), \qquad (2.14)$$

where $\mathrm{Tip}_I = I_{(N_1+1)-} \cup (I_-)_{(N_2+1)+}$. If $N_1 < \infty$ and $N_2 = \infty$, for instance, then $\mathrm{Tip}_I = I_{(N_1+1)-}$. The proof is nothing but an application of (2.13) to the measures μ_I and ν_I. For minor technical reasons, I also wish to allow the choice $N_1 = 0$ and $N_2 = -1$: by definition, this choice means that $\mathrm{Tail}_I = \{I\}$ and $\mathrm{Tip}_I := I_-$. It is easy to see that (2.14) remains valid in this case, with "2" replaced by "4" (for $I = [0, 1)$, this follows by applying (2.11) and (2.12) with the choices $N_1 = 0 = N_2$).

Now, the table is set to prove Proposition 2.4, which I recall here:

Proposition 2.4. *Let μ, ν be measures satisfying the assumptions of the section, and let $\mathcal{T} \subset \mathcal{D}$ be a tree. Moreover, assume that μ is $(\mathcal{T}, D)$-doubling for some constant $D \geq 1$. Then*

$$\sum_{I \in \mathcal{T}} \Delta_{\mu,\nu}^2(I)\mu(I) \lesssim_{D_\nu, D} \sum_{I \in \mathcal{T} \setminus \mathrm{Leaves}(\mathcal{T})} \alpha_{\mu,\nu}^2(I)\mu(I) + \mu(\mathrm{Top}(\mathcal{T})).$$

Proof. The sum over $I \in \mathrm{Leaves}(\mathcal{T})$ is evidently bounded by $4\mu(\mathrm{Top}(\mathcal{T}))$, so it suffices to consider

$$I \in \mathcal{T} \setminus \mathrm{Leaves}(\mathcal{T}) =: \mathcal{T}^-. \qquad (2.15)$$

Let $I \in \mathcal{T}$, and define the number $N_1 = N_1(I) \geq 0$ as the smallest index such that $I_{(N_1+1)-} \in \mathrm{Leaves}(\mathcal{T})$. If no such index exists, set $N_1 = \infty$. If $I_- \in \mathrm{Leaves}(\mathcal{T})$, then $N_1 = 0$, and I define $N_2 = -1$: then $\mathrm{Tail}_I := \{I\}$,

and $\mathrm{Tip}_I := I_-$. Otherwise, if $I_- \in \mathcal{T}^-$, let $N_2 \geq 0$ be the smallest index such that $(I_-)_{(N_2+1)+} \in \mathrm{Leaves}(\mathcal{T})$. If no such index exists, let $N_2 = \infty$. Now $\mathrm{Tail}_I \subset \mathcal{T}^-$ and $\mathrm{Tip}_I \subset \mathrm{Leaves}(\mathcal{T})$ are defined as after (2.14). Start by the following combination of (2.14) and Cauchy–Schwarz:

$$\Delta^2_{\mu,\nu}(I)\mu(I)^2 \lesssim \frac{1}{\tau^2}\left(\sum_{J\in\mathrm{Tail}_I}\alpha^2_{\mu,\nu}(J)\mu(J)^{3/2}\right)\left(\sum_{J\in\mathrm{Tail}_I}\mu(J)^{1/2}\right)$$
$$+\kappa^2\left(\sum_{J\in\mathrm{Tail}_I}\Delta^2_{\mu,\nu}(J)\mu(J)^{3/2}\right)\left(\sum_{J\in\mathrm{Tail}_I}\mu(J)^{1/2}\right)+\mu(\mathrm{Tip}_I)^2. \quad (2.16)$$

The factors $\sum_{J\in\mathrm{Tail}_I}\mu(J)^{1/2}$ are under control, thanks to the $(\mathcal{T}, D)$-doubling hypothesis on μ, and the fact that $\mathrm{Tail}_I \subset \mathcal{T}$. Since Tail_I consists of two "branches" of nested intervals inside I, and the $(\mathcal{T}, D)$-doubling hypothesis implies that the μ-measures of intervals decay geometrically along these branches, one arrives at

$$\sum_{J\in\mathrm{Tail}_I}\mu(J)^{1/2}\lesssim_D \mu(I)^{1/2}.$$

Thus, by (2.16),

$$\Delta^2_{\mu,\nu}(I)\mu(I)\lesssim_D \frac{1}{\tau^2}\sum_{J\in\mathrm{Tail}_I}\alpha^2_{\mu,\nu}(J)\frac{\mu(J)^{3/2}}{\mu(I)^{1/2}}+\kappa^2\sum_{J\in\mathrm{Tail}_I}\Delta^2_{\mu,\nu}(J)\frac{\mu(J)^{3/2}}{\mu(I)^{1/2}}+\frac{\mu(\mathrm{Tip}_I)^2}{\mu(I)}. \quad (2.17)$$

The constant $\kappa > 0$ will have to be chosen so small, eventually, that its product with the implicit constants above is notably less than 1. From now on, the precise restriction $J \in \mathrm{Tail}_I$ can be replaced by the conditions $J \in \mathcal{T}^-$ and $J \subset I$. With this in mind, observe first that

$$\sum_{\substack{I\in\mathcal{T}^- \\ }}\sum_{\substack{J\in\mathcal{T}^- \\ J\subset I}}\alpha^2_{\mu,\nu}(J)\frac{\mu(J)^{3/2}}{\mu(I)^{1/2}}=\sum_{J\in\mathcal{T}^-}\alpha^2_{\mu,\nu}(J)\mu(J)^{3/2}\sum_{\substack{I\in\mathcal{T}^- \\ I\supset J}}\frac{1}{\mu(I)^{1/2}}\lesssim_D\sum_{J\in\mathcal{T}^-}\alpha^2_{\mu,\nu}(J)\mu(J).$$

The final inequality uses, again, the geometric decay of μ-measures of intervals in $\mathcal{T}$. A similar estimate can be performed for the second term in (2.17). As for the third term,

$$\sum_{I\in\mathcal{T}^-}\frac{\mu(\mathrm{Tip}_I)^2}{\mu(I)}\lesssim\sum_{I\in\mathcal{T}^-}\frac{\mu(I_{(N_1+1)-})^2+\mu((I_-)_{(N_2+1)+})^2}{\mu(I)}$$
$$\lesssim\sum_{J\in\mathrm{Leaves}(\mathcal{T})}\mu(J)^2\sum_{\substack{I\in\mathcal{T}^- \\ I\supset J}}\frac{1}{\mu(I)}\lesssim_D\mu(\mathrm{Leaves}(\mathcal{T})),$$

relying once more on the geometric decay of μ in $\mathcal{T}$. Combining all the estimates gives

$$\sum_{I\in\mathcal{T}^-}\Delta^2_{\mu,\nu}(I)\mu(I)\lesssim_D\frac{1}{\tau^2}\sum_{I\in\mathcal{T}^-}\alpha^2_{\mu,\nu}(I)\mu(I)+\kappa^2\sum_{I\in\mathcal{T}^-}\Delta^2_{\mu,\nu}(I)\mu(I)+\mu(\mathrm{Leaves}(\mathcal{T})). \quad (2.18)$$

If the left-hand side is a priori finite, the proof of Proposition 2.4 is now completed by choosing κ small enough, depending on D. If not, consider any finite subtree $\mathcal{T}_j \subset \mathcal{T}$ with $\mathrm{Top}(\mathcal{T}_j) = \mathrm{Top}(\mathcal{T})$. Then, the

proof above gives (2.18) with $\mathcal{T}_j$ in place of $\mathcal{T}$. Hence

$$\sum_{I \in \mathcal{T}_j^-} \Delta_{\mu,\nu}^2(I)\mu(I) \lesssim_D \sum_{I \in \mathcal{T}_j^-} \alpha_{\mu,\nu}^2(I)\mu(I) + \mu(\text{Top}(\mathcal{T})),$$

where the constants do not depend on the choice of $\mathcal{T}_j$. Now the proposition follows by letting $\mathcal{T}_j \nearrow \mathcal{T}$. $\square$

3. Absolute continuity of tree-adapted measures

Recall the concepts of tree, leaves and boundaries from Definition 2.2, and the notion of $(\mathcal{T}, D)$-doubling measures from Definition 2.3. In the present section, I assume that $\mathcal{T} \subset \mathcal{D}$ is a tree, and μ, ν are two finite Borel measures, which satisfy the following two assumptions:

(A) $\min\{\mu(\text{Top}(\mathcal{T})), \nu(\text{Top}(\mathcal{T}))\} > 0$.

(B) μ, ν are $(\mathcal{T}, D)$-doubling for some constant $D \geq 1$.

In particular, the assumptions imply

$$\mu(I) > 0 \quad \text{and} \quad \nu(I) > 0, \quad I \in \mathcal{T}.$$

For reasons to become apparent soon, I define the $(\mathcal{T}, \mu)$-*adaptation of* ν,

$$\nu_{\mathcal{T}} := \nu|_{\partial\mathcal{T}} + \sum_{I \in \text{Leaves}(\mathcal{T})} \frac{\nu}{\mu}(I) \cdot \mu|_I,$$

where $(\nu/\mu)(I) := \nu(I)/\mu(I)$. Note that

$$\nu_{\mathcal{T}}(I) = \nu(I), \quad I \in \mathcal{T}, \tag{3.1}$$

because $\partial\mathcal{T}$ is disjoint from the leaves, which are also pairwise disjoint. In particular, $\nu_{\mathcal{T}}(\text{Top}(\mathcal{T})) = \nu(\text{Top}(\mathcal{T}))$. The main result of the section is the following:

Proposition 3.2. *Assume* (A) *and* (B), *and that*

$$\sum_{I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})} \Delta_{\mu,\nu}^2(I)\mu(I) < \infty.$$

Then $\mu|_{\text{Top}(\mathcal{T})} \ll \nu_{\mathcal{T}}$. *In particular* $\mu|_{\partial\mathcal{T}} \ll \nu$.

Remark 3.3. By the definition of $\nu_{\mathcal{T}}$, it is obvious that $\mu|_{\text{Leaves}(\mathcal{T})} \ll \nu_{\mathcal{T}}$. So, the main point of Proposition 3.2 is to show that $\mu|_{\partial\mathcal{T}} \ll (\nu_{\mathcal{T}})|_{\partial\mathcal{T}} = \nu|_{\partial\mathcal{T}}$.

Since $\mu(\text{Top}(\mathcal{T})) > 0$ and $\nu(\text{Top}(\mathcal{T})) > 0$, one may assume without loss of generality that

$$\mu(\text{Top}(\mathcal{T})) = 1 = \nu(\text{Top}(\mathcal{T})).$$

The proof of Proposition 3.2 is based on a "product representation" for $\nu_{\mathcal{T}}$, relative to μ, in the spirit of [Fefferman et al. 1991, Theorem 3.22] of Fefferman, Kenig and Pipher. Recall that every interval $I \in \mathcal{D}$ has exactly two children: I_- and I_+. Define the μ-adapted Haar functions

$$h_I^\mu := c_I^+ \chi_{I_+} - c_I^- \chi_{I_-}, \quad I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T}),$$

where

$$c_I^+ := \frac{\mu(I)}{\mu(I_+)} \quad \text{and} \quad c_I^- := \frac{\mu(I)}{\mu(I_-)}.$$

This ensures that $\int h_I^\mu \, d\mu = 0$ for $I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$. Note that $\mu(I_+), \mu(I_-) > 0$, because $I_+, I_- \in \mathcal{T}$. Now, the plan is to define coefficients $a_J \in \mathbb{R}$, for $J \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$, so that the following requirement is met:

$$\prod_{\substack{J \supsetneq I \\ J \in \mathcal{T}}} (1 + a_J h_J^\mu)(x) = \frac{\nu}{\mu}(I), \quad x \in I \in \mathcal{T}. \tag{3.4}$$

The left-hand side of (3.4) is certainly constant on I, so the equation has some hope; if $I = \text{Top}(\mathcal{T})$, then the product is empty, and the right-hand side of (3.4) equals 1 by the assumption $\mu(\text{Top}(\mathcal{T})) = \nu(\text{Top}(\mathcal{T})) = 1$. Now, assume that (3.4) holds for some interval $I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$. Then $I_-, I_+ \in \mathcal{T}$, so if (3.4) is supposed to hold for I_-, one has

$$\frac{\nu}{\mu}(I_-) = \prod_{\substack{J \supseteq I_- \\ J \in \mathcal{T}}} (1 + a_J h_J^\mu) = (1 - c_I^- a_I) \prod_{\substack{J \supseteq I \\ J \in \mathcal{T}}} (1 + a_J h_J^\mu) = (1 - c_I^- a_I) \frac{\nu}{\mu}(I), \tag{3.5}$$

and similarly

$$\frac{\nu}{\mu}(I_+) = (1 + c_I^+ a_I) \frac{\nu}{\mu}(I). \tag{3.6}$$

From (3.5) one solves

$$a_I = \frac{(\nu/\mu)(I) - (\nu/\mu)(I_-)}{(\nu/\mu)(I) c_I^-} = \frac{\mu(I_-)}{\mu(I)} - \frac{\nu(I_-)}{\nu(I)}, \tag{3.7}$$

and (3.6) gives

$$a_I = \frac{(\nu/\mu)(I_+) - (\nu/\mu)(I)}{(\nu/\mu)(I) c_I^+} = \frac{\nu(I_+)}{\nu(I)} - \frac{\mu(I_+)}{\mu(I)}. \tag{3.8}$$

Using that $\mu(I_-)/\mu(I) = 1 - \mu(I_+)/\mu(I)$ (and three other similar formulae), it is easy to see that the numbers on the right-hand sides of (3.7) and (3.8) agree. So, a_I can be defined consistently, and (3.4) holds for $I_+, I_- \in \mathcal{T}$. Moreover, the formulae for a_I look quite familiar:

Observation 1. $|a_I| = \Delta_{\mu,\nu}(I)$ for $I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$.

Now that the coefficients a_I have been successfully defined for $I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$, let g be the (at the moment) formal series

$$g(x) := \sum_{I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})} a_I h_I^\mu(x).$$

Since the Haar functions h_I^μ are orthogonal in $L^2(\mu)$, and satisfy

$$\int (h_I^\mu)^2 \, d\mu \leq \max\{c_I^+, c_I^-\}^2 \mu(I) \leq D^2 \mu(I), \quad I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T}),$$

one arrives at

$$\|g\|_{L^2(\mu)}^2 = \sum_{I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})} \Delta_{\mu,\nu}^2(I) \|h_I\|_{L^2(\mu)}^2 \leq D^2 \sum_{I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})} \Delta_{\mu,\nu}^2(I) \mu(I) < \infty,$$

by the assumption in Proposition 3.2. This means that the sequence

$$g_N := \sum_{\substack{I \in \mathcal{T} \backslash \text{Leaves}(\mathcal{T}) \\ |I| > 2^{-N}}} a_I h_I^\mu$$

converges in $L^2(\mu)$. In particular, one can pick a subsequence $(g_{N_j})_{j \in \mathbb{N}}$ which converges pointwise μ almost everywhere (in fact, the entire sequence converges by basic martingale theory, but this is not needed). Now, recall that the goal was to prove that $\mu|_{\text{Top}(\mathcal{T})} \ll \nu_{\mathcal{T}}$. To this end, one has to verify that

$$\liminf_{I \to x} \frac{\mu}{\nu_{\mathcal{T}}}(I) < \infty \tag{3.9}$$

at μ almost every $x \in \text{Top}(\mathcal{T})$. This is clear for $x \in \text{Leaves}(\mathcal{T})$, since the ratios $\mu(I)/\nu_{\mathcal{T}}(I)$, $I \ni x$, are eventually constant. So, it suffices to prove (3.9) at μ almost every point $x \in \partial\mathcal{T}$. Fix a point $x \in \partial\mathcal{T}$ with the properties that sequence $(g_{N_j}(x))_{j \in \mathbb{N}}$ converges, and also

$$\sum_{x \in J \in \mathcal{T}} a_J^2 = \sum_{x \in J \in \mathcal{T}} \Delta_{\mu,\nu}^2(I) < \infty. \tag{3.10}$$

These properties hold at μ almost every $x \in \partial\mathcal{T}$. Let $I \in \mathcal{D}$ be so small that $x \in I \in \mathcal{T}$, and note that

$$\log \frac{\nu_{\mathcal{T}}}{\mu}(I) = \log \frac{\nu}{\mu}(I) = \log \prod_{\substack{J \supsetneq I \\ J \in \mathcal{T}}} (1 + a_J h_J^\mu(x)) = \sum_{\substack{J \supsetneq I \\ J \in \mathcal{T}}} \log(1 + a_J h_J^\mu(x)).$$

Now, the plan is to use the estimate $\log(1+t) \geq t - C_\delta t^2$, which is valid as long as $t \geq \delta - 1$ for some $\delta > 0$. Observe that $a_J h_J^\mu(x) \in \{-c_J^- a_J, c_J^+ a_J\}$, where

$$-a_J c_J^- = \frac{(\nu/\mu)(J_-)}{(\nu/\mu)(J)} - 1 \geq \frac{1}{C} - 1 \quad \text{and} \quad a_J c_J^+ = \frac{(\nu/\mu)(J_+)}{(\nu/\mu)(J)} - 1 \geq \frac{1}{C} - 1. \tag{3.11}$$

Consequently, for $x \in I \in \mathcal{T}$ with $|I| = 2^{-N_j}$, one has

$$\log \frac{\nu_{\mathcal{T}}}{\mu}(I) \geq \sum_{\substack{J \supsetneq I \\ J \in \mathcal{T}}} a_J h_J^\mu(x) - C' \sum_{\substack{J \supsetneq I \\ J \in \mathcal{T}}} (a_J h_J^\mu(x))^2 \geq g_{N_j}(x) - C'D^2 \sum_{x \in J \in \mathcal{T}} a_J^2, \tag{3.12}$$

where $C' \lesssim_D 1$ only depends on the constant C in (3.11). Since the sequence $(g_{N_j}(x))_{j \in \mathbb{N}}$ converges and (3.10) holds, the right-hand side of (3.12) has a uniform lower bound $-M(x) > -\infty$. This implies

$$\limsup_{I \to x} \frac{\nu_{\mathcal{T}}}{\mu}(I) \geq \exp(-M(x)) > 0,$$

which gives (3.9) at x. The proof of Proposition 3.2 is complete.

4. Proof of Theorem 1.8(b)

In this section, Theorem 1.8(b) is proved via a simple tree construction, coupled with Propositions 2.4 and 3.2. Recall the statement of Theorem 1.8(b):

Theorem 4.1. *Assume that μ, ν are Borel probability measures on $[0, 1)$, μ does not charge the boundaries of dyadic intervals, and ν is dyadically doubling. Write $\mu = \mu_a + \mu_s$ for the Lebesgue decomposition of μ relative to ν, and let $\mathcal{S}_{\mathcal{D},\nu}(\mu)$ be the square function*

$$\mathcal{S}^2_{\mathcal{D},\nu}(\mu) = \sum_{I \in \mathcal{D}} \alpha^2_{\mu,\nu}(I) \chi_I.$$

Then, $\mathcal{S}_{\mathcal{D},\nu}(\mu)$ is infinite μ_s almost surely.

An equivalent statement is that the restriction of μ to the set

$$G := \{x \in [0, 1) : \mathcal{S}_{\mathcal{D},\nu}(\mu)(x) < \infty\}$$

is absolutely continuous with respect to ν; this is the formulation proven below. For the rest of the section, fix the measures μ, ν as in the statement above, and let D be the doubling constant of ν. I record a simple lemma, which says that the doubling of ν implies the doubling of μ on intervals, where the α-number is small enough.

Lemma 4.2. *There are constants $\epsilon > 0$ and $C \geq 1$, depending only on D, such that the following holds. For every interval $I \in \mathcal{D}$, if $\alpha_{\mu,\nu}(I) < \epsilon$, then*

$$\mu(I) \leq C \min\{\mu(I_-), \mu(I_+)\}. \tag{4.3}$$

Proof. Let $I_{--} \subset I_-$ and $I_{++} \subset I_+$ be intervals which lie at distance $\geq \frac{1}{8}|I|$ from the boundaries of I_- and I_+, respectively, and have length $\frac{1}{8}|I|$. Let ψ_- and $\psi_+ : \mathbb{R} \to [0, 1]$ be $(C'/|I|)$-Lipschitz functions which equal 1 on I_{--} and I_{++}, respectively, and are supported on I_- and I_+. Then

$$\frac{\mu(I_-)}{\mu(I)} \geq \frac{1}{\mu(I)} \int \psi_- \, d\mu \geq \frac{1}{\nu(I)} \int \psi_- \, d\nu - C' \alpha_{\mu,\nu}(I) \geq \frac{\nu(I_{--})}{\nu(I)} - C' \alpha_{\mu,\nu}(I),$$

and the analogous inequality holds for $\mu(I_+)/\mu(I)$. The ratio $\nu(I_{--})/\nu(I)$ is at least $1/D^3$, so if $\alpha_{\mu,\nu}(I) < 1/(2C'D^3) =: \epsilon$, then both $\mu(I_-) \geq [1/(2D^3)]\mu(I)$ and $\mu(I_+) \geq [1/(2D^3)]\mu(I)$. This gives (4.3) with $C = 2D^3$. $\square$

In particular, if $\mathcal{T}$ is a tree, and $\alpha_{\mu,\nu}(I) < \epsilon$ for all $I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$, then μ is $(\mathcal{T}, C)$-doubling. I will now describe how such trees $\mathcal{T}_j \subset \mathcal{D}$ are constructed, starting with $\mathcal{T}_0$. Let $[0, 1) = \text{Top}(\mathcal{T}_0)$, and assume that some interval I is in $\mathcal{T}_0$. If

$$\sum_{I \subset J \subset [0,1)} \alpha^2_{\mu,\nu}(J) \geq \epsilon^2, \tag{4.4}$$

add I to $\text{Leaves}(\mathcal{T}_0)$. The children I_- and I_+ become the tops of new trees. If (4.4) fails, add I_- and I_+ to $\mathcal{T}_0$. The construction of $\mathcal{T}_0$ is now complete. If a new top T_j was created in the process of constructing $\mathcal{T}_0$, and $\mu(T_j) > 0$, construct a new tree $\mathcal{T}_j$ with $\text{Top}(\mathcal{T}_j) = T_j$ by repeating the algorithm above, only replacing $[0, 1)$ by T_j in the stopping criterion (4.4). Continue this process until all intervals in $\mathcal{D}$ belong to some tree, or all remaining tops T_j satisfy $\mu(T_j) = 0$. For all tops T_j with $\mu(T_j) = 0$, simply define $\mathcal{T}_j := \{I \in \mathcal{D} : I \subset T_j\}$, so there is no further stopping inside $\mathcal{T}_j$.

Remark 4.5. Let $\mathcal{T}$ be one of the trees constructed above, with $\mu(\text{Top}(\mathcal{T})) > 0$. Then μ is $(\mathcal{T}, C)$-doubling by Lemma 4.2, since it is clear that $\alpha_{\mu,\nu}(I) < \epsilon$ for all $I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$. In particular $\mu(I) > 0$ for all $I \in \mathcal{T}$.

The following observation is now rather immediate from the definitions:

Lemma 4.6. *Assume that* $\mathcal{T}_0, \ldots, \mathcal{T}_{N-1}$ *are distinct trees such that* $x \in \text{Leaves}(\mathcal{T}_j)$ *for all* $0 \leq j \leq N-1$. *Then*

$$\mathcal{S}^2_{\mathcal{D},\nu}(\mu)(x) \geq \epsilon^2 N.$$

Proof. For $0 \leq j \leq N-1$, let $I_j \in \text{Leaves}(\mathcal{T}_j)$ with $x \in I_j$. Then

$$\mathcal{S}^2_{\mathcal{D},\nu}(\mu)(x) \geq \sum_{j=0}^{N-1} \sum_{I_j \subset J \subset \text{Top}(\mathcal{T}_j)} \alpha^2_{\mu,\nu}(J) \geq \epsilon^2 N,$$

as claimed. $\square$

It follows that μ almost every point in $G = \{x \in [0,1) : \mathcal{S}_\nu(\mu)(x) < \infty\}$ belongs to $\text{Leaves}(\mathcal{T}_j)$ for only finitely many trees $\mathcal{T}_j$. This is equivalent to saying that μ almost every point in G belongs to $\partial \mathcal{T}$ for some tree $\mathcal{T}$. The converse is also true: if x belongs to $\partial \mathcal{T}$ for some tree $\mathcal{T}$, then clearly $\mathcal{S}_\nu(\mu)(x) < \infty$. Consequently

$$\mu|_G = \sum_{\text{trees } \mathcal{T}} \mu|_{\partial \mathcal{T}}.$$

To prove Theorem 4.1, it now suffices to show that $\mu|_{\partial \mathcal{T}} \ll \nu$ for every tree $\mathcal{T}$. This is clear, if $\mu(\text{Top}(\mathcal{T})) = 0$, so I exclude the trivial case to begin with. In the opposite case, note that

$$\sum_{I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})} \alpha^2_{\mu,\nu}(I) \mu(I) = \int \sum_{I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})} \alpha^2_{\mu,\nu}(I) \chi_I(x) \, d\mu x \leq \epsilon^2 \cdot \mu(\text{Top}(\mathcal{T})). \tag{4.7}$$

It then follows from Proposition 2.4 that

$$\sum_{I \in \mathcal{T}} \Delta^2_{\mu,\nu}(I) \mu(I) \lesssim \mu(\text{Top}(\mathcal{T})) < \infty,$$

and the claim $\mu|_{\partial \mathcal{T}} \ll \nu$ is finally a consequence of Proposition 3.2. The proof of Theorem 1.8(b) is complete.

5. The nondyadic square function

This section contains the proof of Theorem 1.9(b). The argument naturally contains many similarities to the one given above. The main novelty is that one needs to work with the smooth α-numbers, introduced in Definition 1.4 (or [Azzam et al. 2016, Section 5]).

Smooth α-numbers, and their properties. I recall the definition of the smooth α-numbers:

Definition 5.1 (smooth α-numbers). Write $\varphi(x) = \mathrm{dist}(x, \mathbb{R} \setminus (0, 1))$. For an interval $I \subset \mathbb{R}$, define $\alpha_{s,\mu,\nu}(I) := \mathbb{W}_1(\mu_{\varphi,I}, \nu_{\varphi,I})$, where

$$\mu_{\varphi,I} := \frac{T_{I\sharp}(\mu|_I)}{\mu(\varphi_I)} \quad \text{and} \quad \nu_{\varphi,I} := \frac{T_{I\sharp}(\nu|_I)}{\nu(\varphi_I)}.$$

Here $\varphi_I = \varphi \circ T_I$, and $\mu(\varphi_I) = \int \varphi_I \, d\mu$. If $\mu(\varphi_I) = 0$ set $\mu_{\varphi,I} \equiv 0$, and if $\nu(\varphi_I) = 0$, set $\nu_{\varphi,I} \equiv 0$. Unwrapping the definition, if $\mu(\varphi_I), \nu(\varphi_I) > 0$, then

$$\alpha_{s,\mu,\nu}(I) = \sup_{\psi} \left| \frac{1}{\mu(\varphi_I)} \int \psi \circ T_I \, d\mu - \frac{1}{\nu(\varphi_I)} \int \psi \circ T_I \, d\nu \right| = \sup_{\psi} \left| \frac{\mu(\psi_I)}{\mu(\varphi_I)} - \frac{\nu(\psi_I)}{\nu(\varphi_I)} \right|,$$

where the sup is taken over test functions ψ.

Recall that the main reason to prefer the smooth α-numbers over the ones from Definition 1.3 is the following stability property: if $I \subset J$ are intervals of comparable length, then $\alpha_{s,\mu,\nu}(I) \lesssim \alpha_{s,\mu,\nu}(J)$, whenever either μ or ν is doubling. This fact is essentially [Azzam et al. 2017, Lemma 5.2], but I include a proof in Proposition 5.4 for completeness. Similar stability is not true for the numbers $\alpha_{\mu,\nu}(I)$ and $\alpha_{\mu,\nu}(J)$, even for very nice measures μ and ν, as the following example demonstrates:

Example 5.2. Fix $n \in \mathbb{N}$, and let $I_-^n := \left[\frac{1}{2} - 2^{-n}, \frac{1}{2}\right]$ and $I_+^n := \left(\frac{1}{2}, \frac{1}{2} + 2^{-n}\right]$. Let μ be the same measure as in Example 2.1:

$$\mu = \chi_{\mathbb{R} \setminus (I_-^n \cup I_+^n)} + \tfrac{1}{2}\chi_{I_-^n} + \tfrac{3}{2}\chi_{I_+^n}.$$

Let $\nu = \mathcal{L}$. It is clear that both μ and ν are doubling, with constants independent of n. It is also easy to check that $\alpha_{\mu,\nu}(I) \lesssim 2^{-2n}$ for any interval I with length $|I| \sim 1$ such that $I_-^n \cup I_+^n \subset I$ (this implies $\mu(I) = \nu(I)$). However, $\alpha_{\mu,\nu}\left(\left[0, \frac{1}{2}\right]\right) \sim 2^{-n}$, because $\nu_{[0,1/2)} = \chi_{[0,1]}$, while

$$\mu_{[0,1/2]} = \left(1 + \frac{2^{-n}}{1 - 2^{-n}}\right)\chi_{[0, 1 - 2^{1-n})} + \frac{1}{2}\left(1 + \frac{2^{-n}}{1 - 2^{-n}}\right)\chi_{[1 - 2^{1-n}, 1]}.$$

So, for instance, it is clear that no inequality of the form $\alpha_{\mu,\nu}\left(\left[0, \frac{1}{2}\right]\right) \lesssim \alpha_{\mu,\nu}([-1, 1])$ can hold.

Without any doubling assumptions, even the smooth α-numbers can behave badly:

Example 5.3. Let $\mu = \delta_{1/2}$, and $\nu = (1-\epsilon) \cdot \delta_{1/2+\epsilon} + \epsilon \cdot \delta_{1/4}$. Then $\alpha_{s,\mu,\nu}([-1, 1]) \sim \epsilon$, but $\alpha_{s,\mu,\nu}\left(\left[0, \frac{1}{2}\right]\right) \sim 1$.

Proposition 5.4 (basic properties of the smooth α-numbers). *Let μ, ν be two Radon measures on $\mathbb{R}$, and let $I \subset \mathbb{R}$ be an interval. Then*

$$\alpha_{s,\mu,\nu}(I) \leq 2 \quad \text{and} \quad \alpha_{s,\mu,\nu}(I) \leq \frac{2\alpha_{\mu,\nu}(I)}{\nu_I(\varphi)}.$$

Moreover, if ν is doubling with constant D, the following holds: if $I \subset J \subset \mathbb{R}$ are intervals with $|I| \geq \theta|J|$ for some $\theta > 0$, then

$$\alpha_{s,\mu,\nu}(I) \lesssim_{D,\theta} \alpha_{s,\mu,\nu}(J). \tag{5.5}$$

Proof. For the duration of the proof, fix an interval $I \subset \mathbb{R}$ with $\mu(\varphi_I), \nu(\varphi_I) > 0$. The cases where $\mu(\varphi_I) = 0$ or $\nu(\varphi_I) = 0$ always require a little case chase, which I omit. Recall that $\varphi = \chi_{[0,1]} \operatorname{dist}(\,\cdot\,, \{0, 1\})$. Note that any 1-Lipschitz function $\psi \colon \mathbb{R} \to \mathbb{R}$ supported on $[0, 1]$ must satisfy $|\psi| \leq \varphi$. Consequently $|\psi_I| \leq \varphi_I$ for any interval I, and so

$$\alpha_{s,\mu,\nu}(I) \leq \sup_\psi \left[\frac{\mu(|\psi_I|)}{\mu(\varphi_I)} + \frac{\nu(|\psi_I|)}{\nu(\varphi_I)} \right] \leq 2.$$

This proves the first inequality. For the second inequality, one may assume that $\alpha_{\mu,\nu}(I) > 0$, since otherwise $\mu|_{\operatorname{int} I} = c\nu|_{\operatorname{int} I}$ for some constant $c > 0$, and this also gives $\alpha_{s,\mu,\nu}(I) = 0$. After this observation, it is easy to reduce to the case $\mu(\varphi_I) > 0$ and $\nu(\varphi_I) > 0$. Fix a test function ψ. Using that $\mu_I(|\psi|) = \mu(|\psi_I|)/\mu(I) \leq \mu(\varphi_I)/\mu(I) = \mu_I(\varphi)$, one obtains

$$\left| \frac{\mu(\psi_I)}{\mu(\varphi_I)} - \frac{\nu(\psi_I)}{\nu(\varphi_I)} \right| = \left| \frac{\mu_I(\psi)}{\mu_I(\varphi)} - \frac{\nu_I(\psi)}{\nu_I(\varphi)} \right| = \left| \frac{\mu_I(\psi)\nu_I(\varphi) - \nu_I(\psi)\mu_I(\varphi)}{\mu_I(\varphi)\nu_I(\varphi)} \right|$$
$$\leq \frac{\mu_I(|\psi|)}{\mu_I(\varphi)\nu_I(\varphi)} |\mu_I(\varphi) - \nu_I(\varphi)| + \frac{\mu_I(\varphi)}{\mu_I(\varphi)\nu_I(\varphi)} |\mu_I(\psi) - \nu_I(\psi)| \leq \frac{2\alpha_{\mu,\nu}(I)}{\nu_I(\varphi)}.$$

To prove the final claim, start with the following estimate for a test function ψ:

$$\left| \frac{\mu(\psi_I)}{\mu(\varphi_I)} - \frac{\nu(\psi_I)}{\nu(\varphi_I)} \right| \leq \frac{\nu(\varphi_J)}{\nu(\varphi_I)} \left| \frac{\mu(\psi_I)}{\mu(\varphi_J)} - \frac{\nu(\psi_I)}{\nu(\varphi_J)} \right| + \frac{\mu(|\psi_I|)}{\mu(\varphi_I)} \frac{\nu(\varphi_J)}{\nu(\varphi_I)} \left| \frac{\mu(\varphi_I)}{\mu(\varphi_J)} - \frac{\nu(\varphi_I)}{\nu(\varphi_J)} \right|.$$

Then, recall that $\mu(|\psi_I|) \leq \mu(\varphi_I)$. Further, it follows from the doubling of ν that $\nu(\varphi_J) \lesssim_{D,\theta} \nu(\varphi_I)$. Finally, notice that $\psi_I = (\psi_I \circ T_J^{-1}) \circ T_J$ and $\varphi_I = (\varphi_I \circ T_J^{-1}) \circ T_J$, where both

$$\psi_I \circ T_J^{-1} \quad \text{and} \quad \varphi_I \circ T_J^{-1}$$

are $(|J|/|I|)$-Lipschitz functions supported on $T_J(I) \subset [0, 1]$. Consequently,

$$\max\left\{ \left| \frac{\mu(\psi_I)}{\mu(\varphi_J)} - \frac{\nu(\psi_I)}{\nu(\varphi_J)} \right|, \left| \frac{\mu(\varphi_I)}{\mu(\varphi_J)} - \frac{\nu(\varphi_I)}{\nu(\varphi_J)} \right| \right\} \leq \frac{\alpha_{s,\mu,\nu}(J)}{\theta},$$

and the estimate (5.5) follows. $\square$

***Proof of Theorem 1.9**(b).* In this section, ν is a globally doubling measure with constant $D \geq 1$, say. As in Section 4, it suffices to show that $\mu|_G \ll \nu$, where

$$G := \{x : \mathcal{S}_\nu(\mu)(x) < \infty\}.$$

Write

$$\alpha_{s,\mu,\nu}(J) =: \alpha(J), \quad J \subset \mathbb{R}.$$

Assume without loss of generality (or translate both measures μ and ν slightly) that $\mu(\partial I) = 0$ for all $I \in \mathcal{D}$. Also without loss of generality, one may assume that $\operatorname{spt} \mu \subset (0, 1)$: the reason is that the finiteness $\mathcal{S}_\nu(\mu)(x)$ is equivalent to the finiteness of $\mathcal{S}_\nu(\mu|_U)(x)$ for all $x \in U$, whenever $U \subset \mathbb{R}$ is open. So, it suffices to prove $\mu|_{U \cap G} \ll \nu$ for any bounded open set U. Whenever I write $\mathcal{D}$ in the sequel, I only mean the family $\{I \in \mathcal{D} : I \subset [0, 1)\}$.

I start with some standard discretisation arguments. For each $I \in \mathcal{D}$, associate a somewhat larger interval $B_I \supset I$ as follows. First, for $x \in \operatorname{spt} \mu$ and $k \in \mathbb{N}$, choose a radius $r_{x,k} > 0$ such that

$$\alpha(B(x, r_{x,k})) \leq 2 \inf\{\alpha(B(x, r)) : 1.1 \cdot 2^{-k-1} \leq r \leq 0.9 \cdot 2^{-k}\}. \tag{5.6}$$

Then

$$\alpha^2(B(x, r_{x,k})) \leq \left(\frac{1}{\ln[2 \cdot (0.9/1.1)]} \int_{1.1 \cdot 2^{-k-1}}^{0.9 \cdot 2^{-k}} 2\alpha(x, r) \frac{dr}{r} \right)^2 \lesssim \int_{2^{-k-1}}^{2^{-k}} \alpha^2(x, r) \frac{dr}{r}.$$

For $I \in \mathcal{D}$ with $|I| = 2^{-k}$ and $I \cap \operatorname{spt} \mu \neq \varnothing$, let B_I be some open interval of the form $B(x, r_{k-10})$, $x \in I$, such that

$$\alpha(B_I) \leq 2 \inf\{\alpha(B(y, r_{y,k-10})) : y \in I \cap \operatorname{spt} \mu\}.$$

The number "-10" simply ensures that $I \subset B_I$ with $\operatorname{dist}(I, \partial B_I) \sim |I|$, and

$$I \subset J \implies B_I \subset B_J \quad \text{for } I, J \in \mathcal{D}.$$

This implication also uses the slight separation between the scales, provided by the factors "1.1" and "0.9" in (5.6). For $I \in \mathcal{D}$ with $I \cap \operatorname{spt} \mu = \varnothing$, define $B_I := I$ (although this definition will never be really used). Now, a tree decomposition of $\mathcal{D}$ can be performed as in the previous section, replacing the stopping condition (4.4) by declaring $\operatorname{Leaves}(\mathcal{T})$ to consist of the maximal intervals $I \subset \operatorname{Top}(\mathcal{T})$ with

$$\sum_{I \subset J \subset \operatorname{Top}(\mathcal{T})} \alpha^2(B_I) \geq \epsilon^2,$$

where $\epsilon = \epsilon_D > 0$ is a suitable small number; in particular, $\epsilon > 0$ is chosen so small that $\alpha(B_I) \leq \epsilon$ implies $\mu(B_I) \lesssim \mu(I)$ (which is possible by a small modification of Lemma 4.2). If now $x \in \operatorname{Leaves}(\mathcal{T})$ for infinitely many different trees $\mathcal{T}$, then

$$\infty = \sum_{x \in I \in \mathcal{D}} \alpha^2(B_I) \leq 2 \sum_{k \in \mathbb{N}} \alpha^2(B(x, r_{x,k-10})) \lesssim \int_0^{2^{10}} \alpha^2(B(x, r)) \frac{dr}{r},$$

which implies that $x \notin G$. Repeating the argument from Section 4, this gives

$$\mu|_G \leq \sum_{\text{trees } \mathcal{T}} \mu|_{\partial \mathcal{T}}.$$

The converse inequality could also be deduced from the stability of the smooth α-numbers (Proposition 5.4), but it is not needed: the inequality already shows that it suffices to prove

$$\mu|_{\partial \mathcal{T}} \ll \nu \tag{5.7}$$

for any given tree $\mathcal{T}$. So, fix a tree $\mathcal{T}$. If $\epsilon > 0$ was chosen small enough (again depending on D), then μ is $(\mathcal{T}, C)$-doubling for some $C = C_D \geq 1$ in the usual sense:

$$\mu(\hat{I}) \leq C\mu(I), \quad I \in \mathcal{T} \setminus \operatorname{Top}(\mathcal{T}).$$

So, if one knew that

$$\sum_{I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})} \Delta_{\mu,\nu}^2(I)\mu(I) < \infty, \tag{5.8}$$

then the familiar Proposition 3.2 would imply (5.7), completing the entire proof.

The proof of (5.8) is based on the inequality

$$\sum_{I \in \mathcal{T}} \Delta_{\mu,\nu}^2(I)\mu(I) \lesssim \sum_{I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})} \alpha^2(B_I)\mu(B_I) + \mu(\text{Top}(\mathcal{T})). \tag{5.9}$$

The right-hand side is finite by the same estimate as in (4.7) (start with $\mu(B_I) \lesssim \mu(I)$, using $\alpha(B_I) \leq \epsilon$ for $I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$). So, (5.9) implies (5.8). I start the proof of (5.9) by noting that if $I \in \mathcal{D}$, then

$$\begin{aligned}
\Delta_{\mu,\nu}(I) &= \left| \frac{\nu(I_-)}{\nu(I)} - \frac{\mu(I_-)}{\mu(I)} \right| \\
&\leq \frac{\nu(\varphi_{B_I})}{\nu(I)} \left| \frac{\nu(I_-)}{\nu(\varphi_{B_I})} - \frac{\mu(I_-)}{\mu(\varphi_{B_I})} \right| + \frac{\mu(I_-)}{\mu(I)} \frac{\nu(\varphi_{B_I})}{\nu(I)} \left| \frac{\mu(I)}{\mu(\varphi_{B_I})} - \frac{\nu(I)}{\nu(\varphi_{B_I})} \right|.
\end{aligned} \tag{5.10}$$

Noting that $\nu(\varphi_{B_I})/\nu(I) \lesssim_D 1$, to prove (5.9), it suffices to control

$$\sum_{I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})} \left[\left| \frac{\nu(I_-)}{\nu(\varphi_{B_I})} - \frac{\mu(I_-)}{\mu(\varphi_{B_I})} \right|^2 + \left| \frac{\mu(I)}{\mu(\varphi_{B_I})} - \frac{\nu(I)}{\nu(\varphi_{B_I})} \right|^2 \right] \mu(I) \tag{5.11}$$

by the right-hand side of (5.9). The main task is to find a suitable replacement for the "Tail-Tip" inequality (2.14), which I replicate here for comparison:

$$\Delta_{\mu,\nu}(I)\mu(I) \leq \frac{C}{\tau} \sum_{J \in \text{Tail}_I} \alpha_{\mu,\nu}(J)\mu(J) + \kappa \sum_{J \in \text{Tail}_I} \Delta_{\mu,\nu}(J)\mu(J) + 2\mu(\text{Tip}_I). \tag{5.12}$$

Glancing at (5.11), one sees that an analogue for the inequality above is actually needed for both the terms

$$\tilde{\Delta}_{B_I}(I_-) = \left| \frac{\nu(I_-)}{\nu(\varphi_{B_I})} - \frac{\mu(I_-)}{\mu(\varphi_{B_I})} \right| \quad \text{and} \quad \tilde{\Delta}_{B_I}(I) = \left| \frac{\mu(I)}{\mu(\varphi_{B_I})} - \frac{\nu(I)}{\nu(\varphi_{B_I})} \right|.$$

If $I_- \in \text{Leaves}(\mathcal{T})$, then the trivial estimate $\tilde{\Delta}_{B_I}(I_-) \lesssim 1$ will suffice, so in the sequel I assume that

$$I, I_- \notin \text{Leaves}(\mathcal{T}). \tag{5.13}$$

The goal is inequality (5.18) below. Fix B_I and $J \in \{I, I_-\}$. Assume for notational convenience that $|B_I| = 1$, and hence, also $|J| \sim 1$. In a familiar manner, start by writing

$$\chi_J = \sum_{k \in \mathbb{Z}} \psi_k, \tag{5.14}$$

where ψ_k is a nonnegative $C2^{|k|}$-Lipschitz function supported on either $J \subset B_I$ (for $k = 0$) or $J_{|k|-}$ (for negative k) or J_{k+} (for positive k). As in the proof of the original Tail-Tip inequality, it suffices to first estimate

$$\left| \frac{1}{\mu(\varphi_{B_I})} \int \Psi_0^+ \, d\mu - \frac{1}{\nu(\varphi_{B_I})} \int \Psi_0^+ \, d\nu \right|, \tag{5.15}$$

where $\Psi_0^+ = \sum_{k \geq 1} \psi_k + \frac{1}{2}\psi_0$, and more generally $\Psi_j^+ = \sum_{k \geq j} \psi_j$ for $j \geq 1$; eventually one can just replicate the argument for the function $\Psi_0^- = \sum_{k \leq -1} \psi_k + \frac{1}{2}\psi_0$, and summing the bounds gives control for $\tilde{\Delta}_{B_I}(J)$. Start with the following estimate, which only uses the triangle inequality, and the fact that $\frac{1}{2}\psi_0$ is a C-Lipschitz function supported on B_I:

$$\left| \frac{1}{\mu(\varphi_{B_I})} \int \Psi_0^+ \, d\mu - \frac{1}{\nu(\varphi_{B_I})} \int \Psi_0^+ \, d\nu \right| \leq C\alpha(B_I) + \frac{\mu(\varphi_{B_{J_+}})}{\mu(\varphi_{B_I})} \left| \frac{1}{\mu(\varphi_{B_{J_+}})} \int \Psi_1^+ \, d\mu - \frac{1}{\nu(\varphi_{B_{J_+}})} \int \Psi_1^+ \, d\nu \right|$$

$$+ \left(\frac{1}{\nu(\varphi_{B_{J_+}})} \int \Psi_1^+ \, d\nu \right) \left| \frac{\mu(\varphi_{B_{J_+}})}{\mu(\varphi_{B_I})} - \frac{\nu(\varphi_{B_{J_+}})}{\nu(\varphi_{B_I})} \right|. \quad (5.16)$$

Here

$$\frac{1}{\nu(\varphi_{B_{J_+}})} \int \Psi_1^+ \, d\nu \lesssim 1,$$

since ν is doubling and Ψ_1^+ vanishes outside $J_+ \subset B_{J_+}$, and

$$\left| \frac{\mu(\varphi_{B_{J_+}})}{\mu(\varphi_{B_I})} - \frac{\nu(\varphi_{B_{J_+}})}{\nu(\varphi_{B_I})} \right| \leq \frac{|B_I|}{|B_{J_+}|} \cdot \alpha(B_I) \lesssim \alpha(B_I),$$

since $\varphi_{B_{J_+}} = (\varphi_{B_{J_+}} \circ T_{B_I}^{-1}) \circ T_{B_I}$, where $\varphi_{B_{J_+}} \circ T_{B_I}^{-1}$ is a $(|B_I|/|B_{J_+}|)$-Lipschitz function supported on $[0, 1]$. Consequently,

$$\left| \frac{1}{\mu(\varphi_{B_I})} \int \Psi_0^+ \, d\mu - \frac{1}{\nu(\varphi_{B_I})} \int \Psi_0^+ \, d\nu \right| \mu(\varphi_{B_I}) \leq C\alpha(B_I)\mu(\varphi_{B_I})$$

$$+ \left| \frac{1}{\mu(\varphi_{B_{J_+}})} \int \Psi_1^+ \, d\mu - \frac{1}{\nu(\varphi_{B_{J_+}})} \int \Psi_1^+ \, d\nu \right| \mu(\varphi_{B_{J_+}}).$$

Here Ψ_1^+ vanishes outside $J_+ \subset B_{J_+}$, so the estimate can be iterated. After $N \geq 0$ repetitions (the case $N = 0$ was seen above), one ends up with

$$\left| \frac{1}{\mu(\varphi_{B_I})} \int \Psi_0^+ \, d\mu - \frac{1}{\nu(\varphi_{B_I})} \int \Psi_0^+ \, d\nu \right| \mu(\varphi_{B_I})$$

$$\leq C \sum_{k=0}^{N} \alpha(B_{J_{k+}})\mu(\varphi_{B_{J_{k+}}})$$

$$+ \mu(\varphi_{B_{J_{(N+1)+}}}) \left| \frac{1}{\mu(\varphi_{B_{(N+1)+}})} \int \Psi_{N+1}^+ \, d\mu - \frac{1}{\nu(B_{(N+1)+})} \int \Psi_{N+1}^+ \, d\nu \right|, \quad (5.17)$$

where one needs to interpret $J_{0+} = I$ (which is different from J in the case $J = I_-$). What is a good choice for N? Let $N_1 \geq 0$ be the smallest number such that $J_{(N_1+1)+} \in \text{Leaves}(\mathcal{T})$. If there is no such number, let $N_1 = \infty$. In the case $N_1 = \infty$, the term in the last line of (5.17) vanishes, since $\mu(B_{J_{N+}})$ decays rapidly as long as $N \in \mathcal{T}$ (using the doubling of ν, and the fact that $\alpha(B_I) \leq \epsilon$ for $I \in \mathcal{T}$). If $N_1 < \infty$, the term in the last line of (5.17) is clearly bounded by $\leq 2\mu(B_{J_{(N_1+1)+}})$, since $\Psi_{N_1+1}^+$ vanishes outside $J_{(N_1+1)+}$, which is well inside $B_{(N_1+1)+}$. Observing that also $\mu(I) \lesssim \mu(\varphi_{B_I})$, it follows that

$$\left| \frac{1}{\mu(\varphi_{B_I})} \int \Psi_0^+ \, d\mu - \frac{1}{\nu(\varphi_{B_I})} \int \Psi_0^+ \, d\nu \right| \mu(I) \lesssim \sum_{k=0}^{N_1} \alpha(B_{J_{k+}})\mu(B_{J_{k+}}) + \mu(B_{J_{(N_1+1)+}}).$$

Finally, by symmetry, the same argument can be carried out for the series $\Psi_0^- = \sum_{k<0} \psi_k + \frac{1}{2}\psi_0$. If $N_2 \geq 0$ is the smallest number such that $J_{(N_2+1)-} \in \text{Leaves}(\mathcal{T})$, this leads to the following analogue of the Tail-Tip inequality:

$$\tilde{\Delta}_{B_I}(J)\mu(I) \lesssim \sum_{P \in \text{Tail}_J} \alpha(B_P)\mu(B_P) + \mu(\text{Tip}_J), \quad J \in \{I, I_-\}, \ I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T}). \tag{5.18}$$

Here Tail_J is the collection of dyadic intervals $\text{Tail}_J = \{J_{N_2-}, \ldots, J, \ldots, J_{N_1+}\} \subset \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$, and $\text{Tip}_J = B_{J_{(N_2+1)-}} \cup B_{J_{(N_1+1)+}}$. Finally, in the excluded special case, where $J = I_- \in \text{Leaves}(\mathcal{T})$ (recall (5.13)), the same estimate holds if one defines $\text{Tail}_J = \varnothing$ and $\text{Tip}_J := J$ (noting that $I \in \mathcal{T}$, so $\mu(I) \lesssim \mu(J)$).

Armed with the Tail-Tip inequality (5.18), the proof of the main estimate (5.9) is a replica of the argument in the dyadic case, namely the proof of Proposition 2.4. I only sketch the details. For $I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$ and $J \in \{I, I_-\}$, start with

$$\tilde{\Delta}_{B_I}^2(J)\mu(I) \lesssim \sum_{P \in \text{Tail}_J} \alpha^2(B_P)\frac{\mu(B_P)^{3/2}}{\mu(I)^{1/2}} + \frac{\mu(\text{Tip}_J)^2}{\mu(I)}$$

$$\leq \sum_{\substack{P \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T}) \\ P \subset I}} \alpha^2(B_P)\frac{\mu(B_P)^{3/2}}{\mu(I)^{1/2}} + \frac{\mu(\text{Tip}_J)^2}{\mu(I)}.$$

The second inequality is trivial, and the first is proved with the same Cauchy–Schwarz argument as (2.17), using the fact that $\sum_{P \in \text{Tail}_J} \mu(B_P)^{1/2} \lesssim \mu(I)^{1/2}$, which follows from $\text{Tail}_J \subset \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$, and in particular the geometric decay of the measures $\mu(B_P)$ for $P \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$. Now, the inequality above can be summed for $I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$ precisely as in the proof of (2.18). In particular, one should first use the estimate

$$\mu(\text{Tip}_J) \leq \mu(B_{J_{(N_2+1)-}}) + \mu(B_{J_{(N_1+1)+}}) \lesssim \mu(J_{(N_2+1)-}) + \mu(J_{(N_1+1)+}),$$

which follows from $\alpha(B_{J_{N_1+}}), \alpha(B_{J_{N_2-}}) < \epsilon$, if ϵ is small enough, depending on the doubling constant of ν. The conclusion is

$$\sum_{I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})} \tilde{\Delta}_{B_I}^2(J)\mu(I) \lesssim \sum_{P \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})} \alpha^2(B_P)\mu(B_P) + \mu(\text{Leaves}(\mathcal{T}))$$

for $J \in \{I, I_-\}$. As observed in and around (5.11), this implies (5.9).

Remark 5.19. In the proof of (5.9), the uniform bound $\alpha(B_I) < \epsilon$, $I \in \mathcal{T} \setminus \text{Leaves}(\mathcal{T})$, was only used to guarantee that μ is sufficiently doubling along, and inside, the balls B_I. If such properties are assumed a priori in some given tree $\mathcal{T}$, then (5.9) continues to hold for $\mathcal{T}$. In particular, if μ is doubling on the whole real line, and the Carleson condition

$$\int_{B(x,2r)} \int_0^{2r} \alpha_{\mu,\nu}^2(B(y,t)) \frac{dt\, d\mu y}{t} \leq C\mu(B(x,r))$$

holds, then the dyadic Carleson condition of Theorem 1.12 holds for any dyadic system $\mathcal{D}$ (a family of half-open intervals covering $\mathbb{R}$, where every interval has length of the form 2^{-k} for some $k \in \mathbb{Z}$, and

every interval is the union of two further intervals in the family; the proof of Theorem 1.12 seen in Section 2 works for any such system). It follows from this that $\mu \in A_\infty^{\mathcal{D}}(\nu)$ for every dyadic system $\mathcal{D}$, and consequently $\mu \in A_\infty(\nu)$. (To see this, pick a finite collection $\mathcal{D}_1, \ldots, \mathcal{D}_N$ of dyadic systems such that the max of the corresponding dyadic maximal functions $M_\nu^{\mathcal{D}_i}$,

$$M_\nu^{\mathcal{D}_i} f(x) = \sup_{x \in I \in \mathcal{D}_i} \frac{1}{\nu(I)} \int_I |f| \, d\nu,$$

bounds the usual Hardy–Littlewood maximal function M_ν, up to a constant depending only on the doubling of ν. The construction of such systems is well known, and in $\mathbb{R}$ as few as two systems do the trick; for a reference, see for instance Section 5 in [Muscalu et al. 2002]. Then, for every $1 \leq i \leq N$, there exists $p_i < \infty$ such that $\mu \in A_{p_i}^{\mathcal{D}_i}(\nu)$; see [Grafakos 2014, Theorem 9.33(f)]. In particular $\mu \in A_p^{\mathcal{D}_i}(\nu)$ for $p := \max p_i$, and hence $\|M_\nu^{\mathcal{D}_i}\|_{L^p(\mu) \to L^p(\mu)} < \infty$ for $1 \leq i \leq N$. It follows that $\|M_\nu\|_{L^p(\mu) \to L^p(\mu)} < \infty$, which is one possible definition for $\mu \in A_\infty(\nu)$. For much more information, see [Grafakos 2014, Section 9.11].) This proves the "continuous" part of Theorem 1.12.

6. Parts (a) of the main theorems

Parts (a) of Theorems 1.8 and 1.9 are proved in this section: $\mathcal{S}_{\mathcal{D}, \nu}(\mu)$ and $\mathcal{S}_\nu(\mu)$ are finite μ_a almost everywhere, where μ_a is the absolutely continuous part of μ relative to ν. The strategy is to prove the statement first for the dyadic square function $\mathcal{S}_{\mathcal{D}, \nu}(\mu)$, but allow $\mathcal{D}$ to be a slightly generalised system: a family $\mathcal{D} = \bigcup \mathcal{D}_k$, $k \geq 0$, of half-open intervals of length at most 1 such that

(D1) each $\mathcal{D}_k$ is a partition of $\mathbb{R}$,

(D2) each interval in $\mathcal{D}_k$ has length 2^{-k}, and

(D3) each interval $I \in \mathcal{D}_k$ has two children in $\mathcal{D}_{k+1}$, denoted by ch(I).

The added generality makes no difference in the proof, which closely follows previous arguments of Tolsa [2009; 2015]. The benefit is that the nondyadic square function $\mathcal{S}_\nu(\mu)$ can, eventually, be bounded by a finite sum of dyadic square functions $\mathcal{S}_{\mathcal{D}_1, \nu}(\mu), \ldots, \mathcal{S}_{\mathcal{D}_N, \nu}(\mu)$, so the nondyadic problem easily reduces to the dyadic one.

With the strategy in mind, fix a dyadic system $\mathcal{D}$ satisfying (D1)–(D3), and let $\mathcal{S}_{\mathcal{D}, \nu}(\mu)$ be the associated square function.

Lemma 6.1. *Assume that* μ, ν *are Radon measures on* $\mathbb{R}$, *with* μ *finite, and* ν *dyadically doubling* (*relative to* $\mathcal{D}$). *Then* $\mathcal{S}_\nu(\mu)$ *is finite* μ_a *almost surely.*

The proof of Lemma 6.1 is a combination of two arguments of Tolsa: the proofs of [Tolsa 2009, Theorem 1.1] and [Tolsa 2015, Lemma 2.2]. I start with an analogue of the first:

Lemma 6.2. *Assume that* $\mu \in L^2(\nu)$. *Then*

$$\sum_{\substack{I \in \mathcal{D} \\ \nu(I) > 0}} \alpha_{\mu, \nu}^2(I) \frac{\mu(I)^2}{\nu(I)} \lesssim \|\mu\|_{L^2(\nu)}^2.$$

Proof. It suffices to sum over the intervals $I \subset \mathcal{D}$ with $\mu(I) > 0$ and $\nu(I) > 0$; fix one of these I, and a 1-Lipschitz function $\psi \colon \mathbb{R} \to \mathbb{R}$, supported on $[0, 1]$. Then, write

$$\left| \int \psi \, d\mu_I - \int \psi \, d\nu_I \right| = \left| \frac{1}{\mu(I)} \int_I (\psi \circ T_I) g \, d\nu - \frac{1}{\nu(I)} \int_I (\psi \circ T_I) \, d\nu \right|, \qquad (6.3)$$

where g is the Radon–Nikodym derivative $d\mu/d\nu \in L^2(\nu)$. Express $g\chi_I$ in terms of standard (ν-adapted) martingale differences:

$$g\chi_I = \langle g \rangle_I^\nu \chi_I + \sum_{J \in \mathcal{D}(I)} \Delta_J^\nu g, \qquad (6.4)$$

where $\mathcal{D}(I) := \{J \in \mathcal{D} : J \subset I\}$, the sum converges in $L^2(\nu)$, and

$$\langle g \rangle_I^\nu = \frac{1}{\nu(I)} \int g \, d\nu = \frac{\mu(I)}{\nu(I)} \quad \text{and} \quad \Delta_J^\nu g = -\langle g \rangle_J^\nu \chi_J + \sum_{J' \in \mathrm{ch}(J)} \langle g \rangle_{J'}^\nu \chi_{J'}.$$

Note that $\Delta_J^\nu g$ is supported on J and has ν-mean zero. By (6.4),

$$\frac{1}{\mu(I)} \int_J (\psi \circ T_I) g \, d\nu = \frac{1}{\nu(I)} \int_I (\psi \circ T_I) \, d\nu + \sum_{J \in \mathcal{D}(I)} \frac{1}{\mu(I)} \int_J (\psi \circ T_I) \Delta_J^\nu g \, d\nu. \qquad (6.5)$$

Since the first term on the right-hand side of (6.5) cancels out the last term in (6.3), one can continue as follows:

$$(6.3) \leq \sum_{J \in \mathcal{D}(I)} \frac{1}{\mu(I)} \left| \int_J (\psi \circ T_I) \Delta_J^\nu g \, d\nu \right|$$

$$= \sum_{J \in \mathcal{D}(I)} \frac{1}{\mu(I)} \left| \int_J [(\psi \circ T_I) - (\psi \circ T_I(x_J))] \Delta_J^\nu g \, d\nu \right|.$$

Above, x_J is the midpoint of J, and the zero-mean property of $\Delta_J^\nu g$ was used. Finally, recalling that ψ is 1-Lipschitz, one obtains

$$(6.3) \leq \sum_{J \in \mathcal{D}(I)} \frac{\ell(T_I(J))}{\mu(I)} \|\Delta_J^\nu g\|_{L^1(\nu)} \leq \sum_{J \in \mathcal{D}(I)} \frac{\ell(J)\nu(J)^{1/2}}{\mu(I)\ell(I)} \|\Delta_J^\nu g\|_{L^2(\nu)}.$$

Taking a sup over admissible functions $\psi \colon \mathbb{R} \to \mathbb{R}$ gives

$$\alpha_{\mu,\nu}(I) \leq \sum_{J \in \mathcal{D}(I)} \frac{\ell(J)\nu(J)^{1/2}}{\mu(I)\ell(I)} \|\Delta_J^\nu g\|_{L^2(\nu)}. \qquad (6.6)$$

Now, using (6.6) and Cauchy–Schwarz, we may sum over $I \in \mathcal{D}$ as follows (we suppress the requirement $\nu(I) > 0$ from the notation):

$$\sum_{I \in \mathcal{D}} \alpha_{\mu,\nu}^2(I) \frac{\mu(I)^2}{\nu(I)} \leq \sum_{J \in \mathcal{D}} \left(\sum_{J \in \mathcal{D}(I)} \frac{\ell(J)\nu(J)^{1/2}}{\ell(I)} \|\Delta_J^\nu g\|_{L^2(\nu)} \right)^2 \frac{1}{\nu(I)}$$

$$\leq \sum_{I \in \mathcal{D}} \left(\sum_{J \in \mathcal{D}(I)} \frac{\ell(J)}{\ell(I)} \|\Delta_J^\nu g\|_{L^2(\nu)}^2 \right) \sum_{J \in \mathcal{D}(I)} \frac{\ell(J)\nu(J)}{\ell(I)\nu(I)}.$$

Clearly,

$$\sum_{J \in \mathcal{D}(I)} \frac{\ell(J)\nu(J)}{\ell(I)\nu(I)} \lesssim 1,$$

so

$$\sum_{J \in \mathcal{D}} \alpha_{\mu,\nu}(I)^2 \frac{\mu(I)^2}{\nu(I)} \lesssim \sum_{J \in \mathcal{D}} \|\Delta_J^\nu g\|_{L^2(\nu)}^2 \sum_{I \supset J} \frac{\ell(J)}{\ell(I)} \lesssim \sum_{J \in \mathcal{D}} \|\Delta_J^\nu g\|_{L^2(\nu)}^2 \leq \|g\|_{L^2(\nu)}^2,$$

as claimed. $\qquad \square$

Corollary 6.7. *If $\mu \in L^2(\nu)$, then $\mathcal{S}_{\mathcal{D},\nu}(\mu)$ is finite μ almost everywhere.*

Proof. By Lemma 6.2, and the Lebesgue differentiation theorem, the following conditions hold μ almost everywhere:

$$\sum_{x \in I \in \mathcal{D}} \alpha_{\mu,\nu}(I)^2 \frac{\mu(I)}{\nu(I)} < \infty \quad \text{and there exists} \quad \lim_{I \to x} \frac{\mu(I)}{\nu(I)} = \mu(x) > 0.$$

Clearly $\mathcal{S}_{\mathcal{D},\nu}(\mu)(x) < \infty$ for such $x \in [0, 1)$. $\qquad \square$

Now, we can prove Lemma 6.1 by an argument similar to [Tolsa 2015, Lemma 2.2]:

Proof of Lemma 6.1. Perform a Calderón–Zygmund decomposition of μ with respect to ν, at some level $\lambda \geq 1$. More precisely, let $\mathcal{B}$ be the family of maximal intervals $I \in \mathcal{D}$ with $\mu(I) > \lambda\nu(I)$, and set $\mu = g + b$, where

$$g = \mu|_G + \sum_{I \in \mathcal{B}} \frac{\mu(I)}{\nu(I)} \nu|_I, \quad G := [0, 1) \setminus \bigcup_{I \in \mathcal{B}} I,$$

and

$$b = \sum_{I \in \mathcal{B}} \left[\mu|_I - \frac{\mu(I)}{\nu(I)} \nu|_I \right] =: \sum_{I \in \mathcal{B}} b_I.$$

Then $\|g\|_{L^\infty(\nu)} \lesssim \lambda$ (the implicit constants depend on the doubling of ν), and

$$\nu([0, 1) \setminus G) = \sum_{I \in \mathcal{B}} \nu(I) < \frac{1}{\lambda} \sum_{I \in \mathcal{B}} \mu(I) \leq \frac{1}{\lambda}.$$

Since $\mu_a \in L^1(\nu)$ (recall that μ is a finite measure), it follows that $\mu_a([0, 1) \setminus G) \to 0$ as $\lambda \to \infty$. Hence, it suffices to show that

$$\mathcal{S}_{\mathcal{D},\nu}(\mu)(x) < \infty \quad \text{for } \mu \text{ almost every } x \in G \cap \text{spt}_{\mathcal{D}} \mu,$$

where $\text{spt}_{\mathcal{D}} \mu = \{x \in \mathbb{R} : \mu(I) > 0 \text{ for all } x \in I \in \mathcal{D}\}$. Let $\mathcal{G} \subset \mathcal{D}$ be the intervals, which are not contained in any interval in $\mathcal{B}$. Fix $x \in G \cap \text{spt}_{\mathcal{D}} \mu$, and note that if $x \in I \in \mathcal{D}$, then $I \in \mathcal{G}$. Observe that $\mu(I) = g(I)$ for $I \in \mathcal{G}$, and consequently

$$\left| \int \psi \, d\mu_I - \int \psi \, d\nu_I \right| \leq \left| \int \psi \, d\mu_I - \int \psi \, dg_I \right| + \alpha_{g,\nu}(I)$$

$$= \frac{1}{\mu(I)} \left| \int_I (\psi \circ T_I) \, db \right| + \alpha_{g,\nu}(I), \quad I \ni x,$$

for any 1-Lipschitz function $\psi \colon \mathbb{R} \to \mathbb{R}$ supported on $[0, 1]$. Using the zero-mean property of the measures b_J, one can estimate further as follows:

$$\left| \int (\psi \circ T_I) \, db \right| \leq \sum_{J \in \mathcal{B}(I)} \left| \int (\psi \circ T_I) \, db_J \right| = \sum_{J \in \mathcal{B}(I)} \left| \int [(\psi \circ T_I) - (\psi \circ T_I(x_J))] \, db_J \right|,$$

where $\mathcal{B}(I) := \{J \in \mathcal{B} : J \subset I\}$, and x_J is the midpoint of J. Using the fact that ψ is 1-Lipschitz, one has

$$\frac{1}{\mu(I)} \left| \int_I (\psi \circ T_I) \, db \right| \leq \frac{1}{\mu(I)} \left| \int [(\psi \circ T_I) - (\psi \circ T_I(x_J))] \, db_J \right| \leq \frac{\ell(T_I(J))}{\mu(I)} \|b_J\| \lesssim \frac{\ell(J)\mu(J)}{\ell(I)\mu(I)},$$

and finally

$$\mathcal{S}^2_{\mathcal{D},\nu}(\mu)(x) \lesssim \mathcal{S}_{\mathcal{D},\nu}(g)^2(x) + \sum_{x \in I \in \mathcal{G}} \left(\sum_{J \in \mathcal{B}(I)} \frac{\ell(J)\mu(J)}{\ell(I)\mu(I)} \right)^2 =: \mathcal{S}_{\mathcal{D},\nu}(g)^2(x) + S^2(x).$$

Since $\mathcal{S}_{\mathcal{D},\nu}(g)$ is finite g almost everywhere by [Corollary 6.7](), and in particular $\mathcal{S}_{\mathcal{D},\nu}(g)(x) < \infty$ for μ almost every $x \in G$, it remains to prove that $S(x) < \infty$ for μ almost every $x \in \mathbb{R}$. First, note that

$$\sum_{J \in \mathcal{B}(I)} \frac{\ell(J)\mu(J)}{\ell(I)\mu(I)} \leq \frac{1}{\mu(I)} \sum_{J \in \mathcal{B}(I)} \mu(J) \leq 1,$$

as the intervals in $\mathcal{B}(I)$ are disjoint. Consequently,

$$\int S^2 \, d\mu \leq \int \sum_{x \in I \in \mathcal{G}} \sum_{J \in \mathcal{B}(I)} \frac{\ell(J)\mu(J)}{\ell(I)\mu(I)} \, d\mu(x) = \sum_{I \in \mathcal{G}} \sum_{J \in \mathcal{B}(I)} \frac{\ell(J)\mu(J)}{\ell(I)}$$

$$= \sum_{J \in \mathcal{B}} \mu(J) \sum_{J \subset I \in \mathcal{G}} \frac{\ell(J)}{\ell(I)} \lesssim \sum_{J \in \mathcal{B}} \mu(J) \leq \|\mu\| < \infty.$$

It follows that $S^2(x) < \infty$ for μ almost every $x \in \mathbb{R}$. This completes the proof of [Lemma 6.1](), and [Theorem 1.8]()(a).   □

Bounding the nondyadic square function. It remains to prove [Theorem 1.9]()(a). Assume that μ, ν are Radon measures on $\mathbb{R}$, with ν doubling, and recall that $\mathcal{S}_\nu(\mu)$ is the square function

$$\mathcal{S}^2_\nu(\mu)(x) = \int_0^1 \alpha^2_{s,\mu,\nu}(B(x,r)) \, \frac{dr}{r}, \quad x \in \mathbb{R}.$$

The claim is that $\mathcal{S}_\nu(\mu)$ is finite μ_a almost everywhere; since this is a local problem, one may assume that μ is a finite measure. Now, as in [Remark 5.19]() (or see [Muscalu et al. 2002, Section 5]), pick a finite number of dyadic systems $\mathcal{D}_1, \ldots, \mathcal{D}_N$ with the following property: for any interval $I \subset \mathbb{R}$, there exists $j \in \{1, \ldots, N\}$, depending on I, and an interval $J \in \mathcal{D}_j$ such that $I \subset J_i$ and $|J_i| \sim |I|$. As a little technical point, we actually need to restrict $\mathcal{D}_j$ to intervals of length at most 1, so also the defining property above only holds for intervals $I \subset \mathbb{R}$ of length $|I| \leq r_0$, say.

Then, apply Lemma 6.1 to each of the corresponding square functions $\mathcal{S}_{\mathcal{D}_j,\nu}(\mu)$ to infer

$$\mathcal{S}_{\mathcal{D},\nu}(\mu)(x) := \sum_{j=1}^{N} \mathcal{S}_{\mathcal{D}_j,\nu}(\mu)(x) < \infty$$

for μ_a almost every $x \in \mathbb{R}$ (note that ν is dyadically doubling relative to every $\mathcal{D}_j$). So, it suffices to argue that $\mathcal{S}_{\mathcal{D},\nu}(\mu)$ dominates $\mathcal{S}_\nu(\mu)$. Using the stability of the smooth α-numbers, and the fact that they are dominated by the regular α-numbers whenever ν is doubling (see Proposition 5.4), one has

$$\alpha_{s,\mu,\nu}^2(B(x,r)) \lesssim \alpha_{\mu,\nu}^2(I_{x,r}^j), \quad x \in \mathbb{R}, \ 0 < r < r_0,$$

where $j \in \{1, \ldots, N\}$, and $I_{x,r}^j \in \mathcal{D}_j$ is a dyadic interval of length at most 1, satisfying $x \in B(x,r) \subset I_{x,r}$ and $|I_{x,r}| \sim r$. The existence follows from the construction of the systems $\mathcal{D}_j$. It is now clear that $\mathcal{S}_\nu(\mu) \lesssim \mathcal{S}_{\mathcal{D},\nu}(\mu)$, and the proof of Theorem 1.9(a) is complete.

Remark 6.8. Lemma 5.4 in [Azzam et al. 2017] implies

$$\int_{1/4}^{1/2} \alpha_{\mu,\nu}(B(0,t)) \, dt \lesssim \alpha_{s,\mu,\nu}(B(0,1)),$$

whenever ν is doubling, and $\nu\left(B\left(0, \frac{1}{4}\right)\right) > 0$, $\mu\left(B\left(0, \frac{1}{4}\right)\right) > 0$. So, at the level of L^1-averages over scales, the smooth and regular α-numbers are comparable. One would need a similar comparison at the level of L^2-averages to answer Question 1.

Acknowledgements

References

[Azzam et al. 2016] J. Azzam, G. David, and T. Toro, "Wasserstein distance and the rectifiability of doubling measures: part I", *Math. Ann.* **364**:1-2 (2016), 151–224. MR Zbl

[Azzam et al. 2017] J. Azzam, G. David, and T. Toro, "Wasserstein distance and the rectifiability of doubling measures: part II", *Math. Z.* **286**:3-4 (2017), 861–891. MR Zbl

[Buckley 1993] S. M. Buckley, "Summation conditions on weights", *Michigan Math. J.* **40**:1 (1993), 153–170. MR Zbl

[Fefferman et al. 1991] R. A. Fefferman, C. E. Kenig, and J. Pipher, "The theory of weights and the Dirichlet problem for elliptic equations", *Ann. of Math.* (2) **134**:1 (1991), 65–124. MR Zbl

[Grafakos 2014] L. Grafakos, *Modern Fourier analysis*, 3rd ed., Graduate Texts in Mathematics **250**, Springer, 2014. MR Zbl

[Muscalu et al. 2002] C. Muscalu, T. Tao, and C. Thiele, "Multi-linear operators given by singular multipliers", *J. Amer. Math. Soc.* **15**:2 (2002), 469–496. MR Zbl

[Tolsa 2009] X. Tolsa, "Uniform rectifiability, Calderón–Zygmund operators with odd kernel, and quasiorthogonality", *Proc. Lond. Math. Soc.* (3) **98**:2 (2009), 393–426. MR Zbl

[Tolsa 2015] X. Tolsa, "Characterization of n-rectifiability in terms of Jones' square function: part I", *Calc. Var. Partial Differential Equations* **54**:4 (2015), 3643–3665. MR Zbl

TUOMAS ORPONEN: tuomas.orponen@helsinki.fi
Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

ANALYSIS & PDE

Volume 12   No. 4   2019