Wasserstein Distance and the Rectifiability of Doubling Measures: Part II

Jonas Azzam Guy David Tatiana Toro*

October 29, 2014

Abstract

We study the structure of the support of a doubling measure by analyzing its self-similarity properties, which we estimate using a variant of the L^1 Wasserstein distance. We show that a measure satisfying certain self-similarity conditions admits a unique (up to multiplication by a constant) flat tangent measure at almost every point. This allows us to decompose the support into rectifiable pieces of various dimensions.

Soit μ une mesure doublante dans $\mathbb{R}^n$. On introduit deux parties du support où μ a certaines propriétés d'autosimilarité, que l'on mesure à l'aide d'une variante de la L^1-distance de Wasserstein, et on montre qu'en chaque point de ces deux parties, toutes les mesures tangentes à μ sont des multiples d'une mesure plate (la mesure de Lebesgue sur un sous-espace vectoriel). On utilise ceci pour donner une décomposition de ces deux parties en ensembles rectifiables de dimensions diverses.

Key words/Mots clés. Rectifiability, Tangent measures, Doubling measures, Wasserstein distance

Contents

1 Introduction **2**

1.1 Statement of Results 2

1.2 Acknowledgements 7

2 Control of the averages $\alpha_{\mathscr{G}}$ **8**

3 Tangent measures are self-similar **9**

*The first author was partially supported by NSF RTG grant 0838212. The second author acknowledges the generous support of the Institut Universitaire de France, and of the ANR (programme blanc GEOMETRYA, ANR-12-BS01-0014). The third author was partially supported by an NSF grants DMS-0856687 and DMS-1361823, a grant from the Simons Foundation (# 228118) and the Robert R. & Elaine F. Phelps Professorship in Mathematics.

4 Self-similar measures are flat **14**

5 A smoother version of the Wasserstein $\mathbb{W}_1$ distance **19**

6 Uniqueness of the tangent measure at "good points" **24**

7 The decomposition of the "good set" in Σ_1 **27**

8 The decomposition of the "good set" in Σ_2 **29**

1 Introduction

1.1 Statement of Results

In this paper we are concerned with understanding the rectifiability properties of doubling measures. Our ultimate goal is to characterize rectifiable doubling measures. Recently Tolsa provided such a characterization for 1-rectifiable measures with upper density bounded below (see [10]). His conditions are expressed in terms of the properties of the density ratios. We are interested in whether self-similarity properties yield some sort of rectifiabilty. Roughly speaking we analyze how the distance between the dilations of a measure appropriately scaled yield information about the structure of its support. We provide a criteria which ensures that the support of a doubling measure can be decomposed as a union of rectifiable pieces of different dimensions. In a previous paper [1], similar decompositions were obtained by looking at conditions that were expressed in terms of the properties of the *local* distance between the measure and flat measures (that is multiples of Hausdorff measures restricted to affine subsets of Euclidean space). In both cases a minor variant of the L^1 Wasserstein distance is used to estimate the *good features* of a measure.

To present our results we need to define local distances between measures as well as several quantities which describe the self similar character of a measure. In this paper, μ denotes a Radon measure on $\mathbb{R}^n$ (i.e., a locally finite positive Borel measure), and $\Sigma = \Sigma_\mu$ denotes its support. That is,

$$\Sigma = \big\{x \in \mathbb{R}^n \,;\, \mu(B(x,r)) > 0 \text{ for } r > 0\big\}, \tag{1.1}$$

where $B(x,r)$ denotes the open ball centered at x and with radius r. We say that μ is *doubling* when there is a constant $C_\delta > 0$ for which

$$\mu(B(x,2r)) \leq C_\delta\, \mu(B(x,r)) \text{ for all } x \in \Sigma \text{ and } r > 0. \tag{1.2}$$

Let $\mathbb{B} = B(0,1)$ denote the unit ball in $\mathbb{R}^n$. For $M \geq 0$, denote by $\mathrm{Lip}_M(\mathbb{B})$ the set of functions $\psi : \mathbb{R}^n \to \mathbb{R}$ that are M-Lipschitz, i.e., such that

$$|\psi(x) - \psi(y)| \leq M|x-y| \;\text{ for } x, y \in \mathbb{R}^n, \tag{1.3}$$

and for which

$$\psi(x) = 0 \;\text{ for } x \in \mathbb{R}^n \setminus \mathbb{B}. \tag{1.4}$$

Definition 1.1. Let μ and ν be measures on $\mathbb{R}^n$, whose restrictions to $\mathbb{B} := B(0,1)$ are probability measures. We set

$$\mathbb{W}_1(\mu,\nu) := \sup_{\psi\in \mathrm{Lip}_1(\mathbb{B})} \left| \int \psi d\mu - \int \psi d\nu \right|. \tag{1.5}$$

Thus $\mathbb{W}_1(\mu,\nu)$ only measures some distance between the restrictions to $\mathbb{B}$ of μ and ν. This quantity is similar to the usual L^1-Wasserstein distance, which by the Kantorovich duality theorem has the same definition as $\mathbb{W}_1$ except that the infimum ranges over all 1-Lipschitz nonnegative functions in $\mathbb{B}$. Note that $\mathbb{W}_1$ has appeared before in the study of rectifiability of measures; see for example [7], [8], [9], and [1]. In Section 5, we replace $\mathbb{W}_1$ with a smoother version of local distance $\mathbb{W}_\varphi$ which is easier to manipulate. Lemmas 5.1 and 5.3 state that $\mathbb{W}_1$ and $\mathbb{W}_\varphi$ are essentially comparable. We refer to [11] for a detailed introduction to Wasserstein distances and their properties.

To estimate the self-similarity properties of μ we use several groups of affine transformations of $\mathbb{R}^n$. Denote by $\mathscr{R}$ the group of affine isometries of $\mathbb{R}^n$ (i.e., compositions of translations, rotations, and symmetries). Then let $\mathscr{G}$ denote the group of similar affine transformations, defined by

$$\mathscr{G} = \big\{\lambda R\,;\, \lambda > 0 \text{ and } R \in \mathscr{R}\big\}. \tag{1.6}$$

For $G \in \mathscr{G}$, we denote by $\lambda(G)$ the unique positive number such that $G = \lambda(G)R$ for some $R \in \mathscr{R}$. Denote by $\mathscr{D}$ the group of translations and dilations:

$$\mathscr{D} = \big\{\lambda I + a\,;\, \lambda > 0 \text{ and } a \in \mathbb{R}^n\big\}, \tag{1.7}$$

where I denotes the identity on $\mathbb{R}^n$.

The transformations that map a given $x \in \mathbb{R}^n$ to the origin, are denoted by

$$\mathscr{G}(x) = \big\{G \in \mathscr{G};\, G(x) = 0\big\} \text{ and } \mathscr{D}(x) = \big\{D \in \mathscr{D};\, D(x) = 0\big\}. \tag{1.8}$$

To each $G \in \mathscr{G}$, we associate the measure $\mu^G = G_\sharp \mu$, which is defined by

$$\mu^G(A) = \mu(G^{-1}(A)) \text{ for every Borel set } A \subset \mathbb{R}^n. \tag{1.9}$$

When $G \in \mathscr{G}(x)$ for some $x \in \Sigma$, be may normalize μ^G and set

$$\mu_0^G = \frac{\mu^G}{\mu^G(\mathbb{B})} = \frac{\mu^G}{\mu(G^{-1}(\mathbb{B}))} = \frac{\mu^G}{\mu(B(x,\lambda(G)^{-1}))} \tag{1.10}$$

because $\mu(B(x,\lambda(G)^{-1})) > 0$. This normalization is needed if we want to compute $\mathbb{W}_1$-distances.

A special case of this is when $G = T_{x,r}$, the element of $\mathscr{D}$ that maps $B(x,r)$ to $\mathbb{B}$; then μ^G and μ_0^G are denoted by $\mu^{x,r}$ and $\mu_0^{x,r}$ respectively. That is,

$$\mu^{x,r}(A) = \mu(x + rA) \text{ for } A \subset \mathbb{R}^n, \text{ and } \mu_0^{x,r} = \frac{\mu^{x,r}}{\mu(B(x,r))}. \tag{1.11}$$

To measure the self-similar nature of μ we introduce two quantities $\alpha_{\mathscr{G}}$ and $\alpha_{\mathscr{D}}$. We fix two parameters $1 < \lambda_1 < \lambda_2 < \infty$. Set

$$\mathscr{G}(x,r) = \big\{G \in \mathscr{G}(x)\,;\, \lambda_1 r \le \lambda(G)^{-1} \le \lambda_2 r\big\} \tag{1.12}$$

and then

$$\alpha_{\mathscr{G}}(x,r) = \inf\big\{\mathbb{W}_1(\mu_0^G, \mu_0^{x,r})\,;\, G \in \mathscr{G}(x,r)\big\}. \tag{1.13}$$

Thus, if $\alpha_{\mathscr{G}}(x,r)$ is small, this means that in $\mathbb{B}$, $\mu_0^{x,r}$ is close to some measure μ_0^G, obtained via a transformation G that contracts more than $T_{x,r}$ and possibly rotates as well. After composition with $T_{x,r}^{-1}$, the fact that $\alpha_{\mathscr{G}}(x,r)$ is small can be interpreted as saying that in $B(x,r)$, μ is quite close to the measure $a\mu^{G'}$, where $G' = T_{x,r}^{-1} \circ G$ is a contracting element of $\mathscr{G}$ that fixes x and $a > 0$ is a normalizing constant. It is important to note that even though we allow some flexibility in the choice of G and G', we demand that $G'(x) = x$. This is the reason why the usual fractal measures do not satisfy the conditions below.

We also use the analogue of $\alpha_{\mathscr{G}}(x,r)$ for the smaller group $\mathscr{D}$.That is, set

$$\mathscr{D}(x,r) = \big\{D \in \mathscr{D}(x)\,;\, \lambda_1 r \leq \lambda(G)^{-1} \leq \lambda_2 r\big\} \tag{1.14}$$

and

$$\alpha_{\mathscr{D}}(x,r) = \inf\big\{\mathbb{W}_1(\mu_0^G, \mu_0^{x,r})\,;\, G \in \mathscr{D}(x,r)\big\}. \tag{1.15}$$

For $\alpha_{\mathscr{D}}(x,r)$ we only compare μ with its image by some dilation centered at x. Obviously $\alpha_{\mathscr{D}}(x,r) \geq \alpha_{\mathscr{G}}(x,r)$. Thus conditions on $\alpha_{\mathscr{D}}(x,r)$ are more restrictive than those on $\alpha_{\mathscr{G}}(x,r)$.

Our goal is to get a control on the part of Σ (see (1.1)) where either $\alpha_{\mathscr{G}}(x,r)$, or $\alpha_{\mathscr{D}}(x,r)$, are sufficiently small. More precisely, we want to control the sets

$$\Sigma_1 = \big\{x \in \Sigma\,;\, \int_0^1 \alpha_{\mathscr{D}}(x,r)\frac{dr}{r} < \infty\big\} \tag{1.16}$$

and

$$\Sigma_2 = \big\{x \in \Sigma\,;\, \int_0^1 \alpha_{\mathscr{G}}(x,r)\frac{\log(1/r)dr}{r} < \infty\big\}. \tag{1.17}$$

Theorem 1.2. *Let μ be a doubling measure on $\mathbb{R}^n$, and denote by Σ its support. Let $1 < \lambda_1 < \lambda_2 < \infty$ be given, and define the sets Σ_1 and Σ_2 as above. Then there are sets $\mathscr{S}_0, ..., \mathscr{S}_n \subset \Sigma$, such that*

$$\mu\Big((\Sigma_1 \cup \Sigma_2) \setminus \Big(\bigcup_{d=0}^{n} \mathscr{S}_d\Big)\Big) = 0, \tag{1.18}$$

and moreover

- *$\mathscr{S}_0$ is the set of points where Σ has an atom; it is at most countable, and every point of $\mathscr{S}_0$ is an isolated point of Σ.*
- *For $1 \leq d \leq n$, if $x \in \mathscr{S}_d$, there exists a d-dimensional vector space V_x such that all the tangent measures to μ at x (defined below) are multiples of the Lebesgue measure on V_x.*
- *For $1 \leq d \leq n$, $\mathscr{S}_d$ is d-rectifiable, and it can be covered by a countable family of Lipschitz graphs of dimension d.*

We may see Theorem 1.2 as a structural decomposition of the good parts of Σ. Tangent measures will play an important role in proof and the definition of the $\mathscr{S}_d$. Recall that the set of *tangent measures* to μ at x, which will be denoted by $\mathrm{Tan}(\mu, x)$, is the set of non-zero Radon

measures σ for which there exist a sequence $\{r_k\}$, with $\lim_{k\to\infty} r_k = 0$, and a sequence $\{a_k\}$ of nonnegative numbers, such that

$$\sigma \text{ is the weak limit of the measures } a_k \mu^{x,r_k}, \tag{1.19}$$

where the μ^{x,r_k} are as in (1.11). That is, for every continuous function f with compact support,

$$\int f d\sigma = \lim_{k\to\infty} a_k \int f d\mu^{x,r_k}. \tag{1.20}$$

Note that since here μ is doubling, $\mathrm{Tan}(\mu, x)$ is not empty (see for instance the proof of Lemma 2.1 in [1]). Furthermore if μ satisfies (1.2) and $\sigma \in \mathrm{Tan}(\mu, x)$ then σ is also doubling with a constant at most C_δ^2. A priori $\mathrm{Tan}(\mu, x)$ may be large. Nevertheless Theorem 1.2 ensures that for $x \in \mathscr{S}_d$, $\mathrm{Tan}(\mu, x)$ is of dimension 1.

A Lipschitz graph of dimension d is a set Γ_A such that

$$\Gamma_A = \big\{x + A(x)\,;\, x \in V\big\},$$

where V is a vector space of dimension d, $A : V \to V^\perp$ is a Lipschitz map and $V^\perp$ denotes the $(n-d)$-dimensional vector space perpendicular to V. In the statement of Theorem 1.2, $\mathscr{S}_d$ can be covered by Lipschitz graphs where the corresponding function A has Lipschitz constant less than ε, where $\varepsilon > 0$ is any small number given in advance. Note that this yields that $\mathscr{S}_d$ is d-rectifiable while providing additional information in the sense that $\mathscr{S}_d$ is completely covered by Lipschitz graphs not simply up to a set of $\mathscr{H}^d$-measure zero. Let us make a few more remarks on Theorem 1.2 and its proof. The advantage of using the quantities $\alpha_{\mathscr{G}}$ and $\alpha_{\mathscr{D}}$ is that they yield information not only about the geometry of the support but also about how the measure is distributed on it. The decomposition of $\Sigma_1 \cup \Sigma_2$ into pieces of different dimensions is possible once we prove that for μ-almost every point $x \in \Sigma_1 \cap \Sigma_2$, $\mathrm{Tan}(\mu, x)$ is entirely composed of flat measures of a same dimension depending on x. Recall that *flat measures* are multiples of Lebesgue measures on vector subspaces of $\mathbb{R}^n$; that is, for each integer $d \in [0, n]$, set

$$\mathscr{F}_d = \big\{c\mathscr{H}^d \llcorner V\,;\, c \geq 0 \text{ and } V \in G(d, n)\big\}, \tag{1.21}$$

where $\mathscr{H}^d$ denotes the d-dimensional Hausdorff measure (see [6] or [4]) and $G(d, n)$ is the set of d-planes in $\mathbb{R}^n$. The set of flat measures is $\mathscr{F} = \bigcup_{0\leq d\leq n} \mathscr{F}_d$.

It is natural to use the self-similarity properties of μ to get information on the structure of Σ, as in Theorem 1.2. In particular the numbers $\alpha_{\mathscr{G}}(x, r)$ provide an intrinsic way to measure the regularity of μ. We contrast this approach with the one taken in [1] where we were interested on the local approximation of the measure by flat measures.

As we shall see in Section 8, the additional logarithm in (1.17) is used to sum a series which allows us to control the density of μ on most of Σ_2. It may well be an artifact of the proof.

To prove Theorem 1.2 we need to find a set Σ_0 which covers almost all $\Sigma_1 \cup \Sigma_2$ and such that all tangents to μ at points in Σ_0 are flat. To accomplish this we define an average analogue of the numbers $\alpha_{\mathscr{G}}(x, r)$ by

$$\alpha^*_{\mathscr{G}}(x, r) = \fint_{B(x,r)} \fint_r^{2r} \alpha_{\mathscr{G}}(y, t) d\mu(y) dt \tag{1.22}$$

(where $\fint$ is our notation for an average) and consider the set

$$\Sigma_0 = \big\{x \in \Sigma\,;\, \lim_{r\to 0} \alpha^*_{\mathscr{G}}(x,r) = 0\big\}. \tag{1.23}$$

Theorem 1.3. *Let μ be a doubling measure on $\mathbb{R}^n$, and denote by Σ its support. Let $1 < \lambda_1 < \lambda_2 < \infty$ be given, and define the functions $\alpha_{\mathscr{G}}(x,r)$ and $\alpha^*_{\mathscr{G}}(x,r)$ and the set Σ_0 as above (see* (1.13)*,* (1.22)*, and* (1.23)*). Then*

$$\textit{Tan}(\mu,x) \subset \mathscr{F} \ \textit{ for every } x \in \Sigma_0. \tag{1.24}$$

A consequence of (1.24) and the fact that elements of different $\mathscr{F}_d$ are far away from each other is that for each $x \in \Sigma_0$, there is an integer $d \in [0,n]$ such that

$$\mathrm{Tan}(\mu,x) \subset \mathscr{F}_d. \tag{1.25}$$

This is not too hard to prove. In the case of Theorem 1.2, we obtain more than (1.25) directly, thus we omit the proof of this fact.

To deduce Theorem 1.2 from Theorem 1.3, we shall first check that

$$\mu\big((\Sigma_1 \cup \Sigma_2) \setminus \Sigma_0\big) = 0. \tag{1.26}$$

This uses standard techniques from measure theory including the Lebesgue density theorem. Then we show that for each $x \in \Sigma_0 \cap (\Sigma_1 \cup \Sigma_2)$,

$$\mathrm{Tan}(\mu,x) = \big\{c\sigma\,;\, c > 0\big\} \ \text{ for some } \sigma \in \mathscr{F}. \tag{1.27}$$

Let us now say a few words about the definition of the $\mathscr{S}_d$. Set

$$\mathscr{S}_d = \big\{x \in \Sigma_0 \cap (\Sigma_1 \cup \Sigma_2)\,;\, \mathrm{Tan}(\mu,x) \subset \mathscr{F}_d\big\}; \tag{1.28}$$

these sets are disjoint, and by (1.27) or (1.25)

$$\Sigma_0 \cap (\Sigma_1 \cup \Sigma_2) = \bigcup_{d=0}^{n} \mathscr{S}_d. \tag{1.29}$$

The special set $\mathscr{S}_0$ is easily dealt with at the beginning of Section 7, and Theorem 1.2 follows as soon as we prove that for $d \geq 1$,

$$\begin{aligned} &\mathscr{S}_d \text{ can be covered by a countable collection} \\ &\qquad\qquad\qquad\qquad \text{of Lipschitz graphs of dimension } d. \end{aligned} \tag{1.30}$$

The fact that information on the tangent measures may imply rectifiability properties for the measure is much better understood since D. Preiss [7] showed that if μ is a Radon measure, not necessarily doubling, such that for μ-almost every $x \in \mathbb{R}^n$, the d-density $\lim_{r\to 0} \frac{\mu(B(x,r))}{r^d}$ exists and is positive and finite, then $\mathbb{R}^n$ may be covered, up to a set of μ-measure zero, by a countable collection of d-dimensional Lipschitz graphs. He deduced this from the hypothesis

on the d-density and the fact that at μ-almost every $x \in \Sigma$, $\mathrm{Tan}(\mu, x) \subset \mathscr{F}$. In our case, we are unable to use [7] because we are not given any information on the density of μ.

We shall use the fact that since μ is doubling, (1.27) implies the existence of a tangent d-plane to Σ at x, and then (1.30) for the set where (1.27) holds. We include a proof of these simple observations in Section 7.

To prove (1.27), we shall use the numbers $\alpha_{\mathscr{D}}(x, r)$ and $\alpha_{\mathscr{G}}(x, r)$ to control the variations of the measures $\mu^{x,r}$ on Σ_1 and Σ_2. Eventually we compare them to the tangent measures.

For points of Σ_1 we use the triangle inequality and the summability of the $\alpha_{\mathscr{D}}(x, r)$, to show that the distance between two different tangent measures at x is controlled by integrals that tend to 0. To deal with some of the technical complications that arise with the distance $\mathbb{W}_1$, we shall introduce in Section 5 a smoother variant $\mathbb{W}_\varphi$ of the Wasserstein distance, study it briefly, and then use it in Section 6 to prove (1.27) on $\Sigma_1 \cap \Sigma_0$.

For points of Σ_2, we'll use the bounds on the numbers $\alpha_{\mathscr{G}}(x, r)$ to compute the $\mathbb{W}_\varphi$-distance between the $\mu^{x,r}$ and the tangent measures. This time we can only work modulo rotations, but this is enough to control the $\mathbb{W}_1$-distance from the $\mu_0^{x,r}$ to flat measures, and apply Theorem 1.5 in [1]. This yields additional information on $\Sigma_2 \cap \Sigma_0$. In particular, it guarantees that on the sets $\Sigma_2 \cap \mathscr{S}_d$, μ is absolutely continuous with respect to the Hausdorff measure $\mathscr{H}^d$, with a density that can be computed from the measure of balls, and that some local mutual absolute continuity of μ and $\mathscr{H}^d_{|\Sigma_2 \cap \mathscr{S}_d}$ holds. See near (8.17) for a statement, and the rest of Section 8 for the proof.

There is a significant difference between (1.24) (or even (1.25)) and the stronger (1.27). For instance, let Σ be an asymptotically flat snowflake in $\mathbb{R}^2$, constructed in the usual way but with angles that tend slowly to 0. Put on Σ the natural measure μ, coming from the parameterization of Σ (see [3]). In this case for μ-almost every $x \in \Sigma$, $\mathrm{Tan}(\mu, x) = \mathscr{F}_1$. Of course Σ is not rectifiable, and Theorem 1.2 says that $\Sigma_1 \cup \Sigma_2$ is μ-negligible.

The definitions (1.22) and (1.23) ensure that for $x \in \Sigma_0$, the local self-similarity character of μ improves as the balls get smaller and smaller, which yields self-similar tangent measures at all point of Σ_0. That is, we show that if $x \in \Sigma_0$, $\sigma \in \mathrm{Tan}(\mu, x)$ and y lies in the support of σ, there is a transformation $G \in \mathscr{G}$ such that $\lambda_1 \leq \lambda(H)^{-1} \leq \lambda_2$, $H(y) = y$, and $H_\sharp \sigma = c\sigma$ for some $c > 0$. See Lemma 3.1, in Section 3. Once we prove this, showing that σ is flat is mostly a matter of playing with the invariance properties of the support and the measure; see Section 4.

1.2 Acknowledgements

The authors are grateful to Alessio Figalli and Xavier Tolsa for helpful discussions. Parts of this work were done while the first author was visiting IPAM and while the second author was visiting the University of Washington.

2 Control of the averages $\alpha_{\mathscr{G}}$

The main goal of this section is to prove (1.26). To this effect define the set

$$\Sigma_3 = \big\{x \in \Sigma\,;\, \int_0^1 \alpha_{\mathscr{G}}(x,r)\frac{dr}{r} < \infty\big\}. \tag{2.1}$$

Notice that $\Sigma_1 \cup \Sigma_2 \subset \Sigma_3$, by (1.16), (1.17), and because $\alpha_{\mathscr{G}}(x,r) \le \alpha_{\mathscr{D}}(x,r)$. Thus (1.26) follows once we prove that

$$\mu(\Sigma_3 \setminus \Sigma_0) = 0. \tag{2.2}$$

For $N > 0$ large and $k \ge 2$, let

$$\Sigma_3(N) = \big\{x \in \Sigma_3 \cap B(0,N)\,;\, \int_0^1 \alpha_{\mathscr{G}}(x,r)\frac{dr}{r} \le N\big\} \tag{2.3}$$

and

$$\varepsilon_k = \int_{\Sigma_3(N)} \int_{2^{-k}}^{2^{-k+2}} \alpha_{\mathscr{G}}(y,r)\frac{d\mu(y)dr}{r}. \tag{2.4}$$

Then

$$\sum_{k\ge 2} \varepsilon_k \le 2\int_{\Sigma_3(N)} \int_0^1 \alpha_{\mathscr{G}}(x,r)\frac{d\mu(x)dr}{r} \le 2N\mu(\Sigma_3(N)) < \infty. \tag{2.5}$$

Choose a decreasing sequence $\{\gamma_k\}$ such that

$$\lim_{k\to\infty} \gamma_k = 0 \ \text{ and } \ \sum_{k\ge 2} \gamma_k^{-1}\varepsilon_k < \infty. \tag{2.6}$$

For $x \in \Sigma_3(N)$, define auxiliary functions α_k by

$$\alpha_k(x) = \int_{\Sigma_3(N)\cap B(x,2^{-k+1})} \int_{2^{-k}}^{2^{-k+2}} \alpha_{\mathscr{G}}(y,r)\frac{d\mu(y)dr}{r}. \tag{2.7}$$

Consider the bad sets

$$Z_k = \big\{x \in \Sigma_3(N)\,;\, \alpha_k(x) \ge \gamma_k \mu(B(x,2^{-k+1}))\big\}. \tag{2.8}$$

Our goal is to show that Z_k is small. Let $X \subset Z_k$ be a maximal subset whose points lie at distance at least 2^{-k+2} from each other. Thus the balls $\overline{B}(x,2^{-k+2})$, $x \in X$, cover Z_k, so by (1.2) and (2.8)

$$\begin{aligned} \mu(Z_k) &\le \sum_{x\in X} \mu(\overline{B}(x,2^{-k+2})) \le C_\delta^2 \sum_{x\in X} \mu(B(x,2^{-k+1})) \\ &\le C_\delta^2 \gamma_k^{-1} \sum_{x\in X} \alpha_k(x). \end{aligned} \tag{2.9}$$

Since the balls $B(x,2^{-k+1})$, $x \in X$, are disjoint,

$$\sum_{x\in X} \alpha_k(x) \le \varepsilon_k$$

(compare (2.4) and (2.7)); thus $\mu(Z_k) \leq C_\delta^2 \gamma_k^{-1} \varepsilon_k$. We are not interested in the precise bound, but merely the fact that $\sum_k \mu(Z_k) < \infty$, from which we deduce that if we set $Z_l^* = \bigcup_{k \geq l} Z_k$ for $l \geq 2$, then $\lim_{l \to \infty} \mu(Z_l^*) = 0$. Thus for μ-almost every $x \in \Sigma_3(N)$ there is $k_x \in \mathbb{N}$ such that

$$x \in \Sigma_3(N) \setminus Z_k \text{ for } k \geq k_x. \tag{2.10}$$

By the Lebesgue differentiation theorem applied to the doubling measure μ, we have that for μ-almost every $x \in \Sigma_3(N)$

$$\lim_{r \to 0} \frac{\mu(B(x,r) \cap \Sigma \setminus \Sigma_3(N))}{\mu(B(x,r) \cap \Sigma)} = 0; \tag{2.11}$$

see for instance Corollary 2.14 in [6].

Let $x \in \Sigma_3(N)$ satisfy (2.10) and (2.11); we want to estimate $\alpha_{\mathscr{G}}^*(x,r)$ for r small. Choose k such that $2^{-k} \leq r \leq 2^{-k+1}$; then $k \geq k_x$ for r small. Recall from (1.22) that

$$\alpha_{\mathscr{G}}^*(x,r) = \mu(B(x,r))^{-1} \int_{y \in B(x,r)} \fint_r^{2r} \alpha_{\mathscr{G}}(y,t) d\mu(y) dt. \tag{2.12}$$

We decompose the domain of integration above into two parts and estimate each one separately. By (1.2), (2.7), (2.10), and (2.8),

$$\begin{aligned} &\mu(B(x,r))^{-1} \int_{\Sigma_3(N) \cap B(x,r)} \fint_r^{2r} \alpha_{\mathscr{G}}(y,t) d\mu(y) dt \\ &\quad \leq \; 4\mu(B(x,2^{-k}))^{-1} \int_{\Sigma_3(N) \cap B(x,r)} \int_{2^{-k}}^{2^{-k+2}} \alpha_{\mathscr{G}}(y,t) \frac{d\mu(y) dt}{t} \\ &\quad \leq \; 4C_\delta \mu(B(x,2^{-k+1}))^{-1} \alpha_k(x) \leq 4C_\delta \gamma_k. \end{aligned} \tag{2.13}$$

This term tends to 0 when r tends to 0, by (2.6).

For the second part, we notice that $\alpha_{\mathscr{G}}(y,t) \leq 2$ by definition (see (1.13) and (1.5)), so

$$\begin{aligned} &\mu(B(x,r))^{-1} \int_{\Sigma \cap B(x,r) \setminus \Sigma_3(N)} \fint_r^{2r} \alpha_{\mathscr{G}}(y,t) d\mu(y) dt \\ &\qquad \leq \; 2\mu(B(x,r))^{-1} \mu(\Sigma \cap B(x,r) \setminus \Sigma_3(N)), \end{aligned} \tag{2.14}$$

which tends to 0 by (2.11). Combining (2.13) and (2.14) we get that

$$\lim_{r \to 0} \alpha_{\mathscr{G}}^*(x,r) = 0 \text{ for } \mu\text{-almost every } x \in \Sigma_3(N).$$

In other words, $\mu(\Sigma_3(N) \setminus \Sigma_0) = 0$ (see (1.23)); (2.2) follows easily, and so does (1.26).

3 Tangent measures are self-similar

In this section we start the proof of Theorem 1.3 and prove the basic self-similarity estimate for tangent measures.

Lemma 3.1. *Let μ be a doubling measure, let Σ denote its support, let $\Sigma_0 \subset \Sigma$ be as in* (1.23)*, and $\sigma \in$ Tan(μ, x) be a tangent measure of μ at a point $x \in \Sigma_0$. For each $y \in \Xi$, the support of σ, there exist $H \in \mathscr{G}$ such that $H(y) = y$,*

$$\lambda_1 \leq \lambda(H)^{-1} \leq \lambda_2 \tag{3.1}$$

and

$$H_\sharp \sigma = c\sigma \ \textit{ for some } c > 0. \tag{3.2}$$

The numbers λ_1 and λ_2 are the same as in (1.12), the dilation number $\lambda(H)$ is defined below (1.6), and $H_\sharp \sigma$, the push forward image of σ by H, is defined as in (1.9).

Proof. We may assume, without loss of generality, that $x = 0$. Since $\sigma \in$ Tan(μ, x) there are coefficients $a_k \geq 0$ and radii $r_k > 0$, such that $\lim_{k\to\infty} r_k = 0$, and σ is the weak limit of the measures $\{\sigma_k\}$, where

$$\sigma_k = a_k \mu^{0,r_k} = a_k \mu^{R_k} \text{ with } R_k(u) = r_k^{-1} u \ \text{ for } u \in \mathbb{R}^n. \tag{3.3}$$

Note that R_k maps $B(x, r_k) = B(0, r_k)$ to $\mathbb{B}$. See (1.11) for the definition of $\mu^{x,r}$.

Let

$$\alpha_k = \sup \left\{ \alpha^*_{\mathscr{G}}(0, r) \, ; \, 0 < r < \sqrt{r_k} \right\}; \tag{3.4}$$

then since $x \in \Sigma_0$ (see (1.23))

$$\lim_{k\to\infty} \alpha_k = 0. \tag{3.5}$$

By (3.4), if for k large

$$r_k < \rho_k < \sqrt{r_k} \tag{3.6}$$

then $\alpha^*_{\mathscr{G}}(0, \rho_k) \leq \alpha_k$ for these k. Since σ is the weak limit of the σ_k, for each $y \in \Xi := \operatorname{supp} \sigma$ we can find points $y_k \in \operatorname{supp}(\sigma_k) = r_k^{-1}\Sigma$ such that

$$\lim_{k\to\infty} |y_k - y| = 0. \tag{3.7}$$

Let $\{\eta_k\}_{k\geq 1}$ and $\{\rho_k\}_{k\geq 1}$ be sequences such that (3.6) holds for k large, and also

$$\lim_{k\to\infty} \frac{\rho_k}{r_k} = \infty \ \text{ and } \ \lim_{k\to\infty} \eta_k = 0. \tag{3.8}$$

Consider

$$A_k = \fint_{B(r_k y_k, \eta_k r_k)} \fint_{\rho_k}^{2\rho_k} \alpha_{\mathscr{G}}(z, t) d\mu(z) dt. \tag{3.9}$$

For k large, the domain of integration $B(r_k y_k, \eta_k r_k)$ is contained in $B(0, \rho_k)$ (because y_k tends to y, η_k tends to 0, and $r_k^{-1}\rho_k$ tends to $+\infty$; see (3.7) and (3.8)). Recall that $B(0, \rho_k) \times [\rho_k, 2\rho_k]$ is the domain of integration in the definition of $\alpha^*_{\mathscr{G}}(0, \rho_k)$ (see (1.22)); hence for k large

$$A_k \leq \frac{\mu(B(0, \rho_k))}{\mu(B(r_k y_k, \eta_k r_k))} \alpha^*_{\mathscr{G}}(0, \rho_k) \leq C_\delta^{2+\log_2(\rho_k/(\eta_k r_k))} \alpha_k \tag{3.10}$$

by (3.9), (1.22), (3.6), (3.4), and the doubling property (1.2). Later on, we will choose ρ_k and η_k, depending on α_k, so that A_k is still small enough.

By Chebyshev's inequality there exist

$$z_k \in \Sigma \cap B(r_k y_k, \eta_k r_k) \text{ and } t_k \in [\rho_k, 2\rho_k] \tag{3.11}$$

such that

$$\alpha_{\mathscr{G}}(z_k, t_k) \leq A_k. \tag{3.12}$$

By the definition of $\alpha_{\mathscr{G}}$ (see (1.13)) there exists $G_k \in \mathscr{G}(z_k, t_k)$ such that

$$\mathbb{W}_1(\mu_0^{G_k}, \mu_0^{z_k,t_k}) \leq 2A_k, \tag{3.13}$$

which by (1.5) means that

$$\left| \int \psi d\mu_0^{G_k} - \int \psi \mu_0^{z_k,t_k} \right| \leq 2A_k \text{ for any } \psi \in \mathrm{Lip}_1(\mathbb{B}). \tag{3.14}$$

Our next goal is to interpret (3.14) in terms of σ_k. Let $T_k \in \mathscr{D}$ be such that for $u \in \mathbb{R}^n$

$$T_k(u) = \frac{u - z_k}{t_k} \text{ and so } T_k(B(z_k, t_k)) = \mathbb{B}. \tag{3.15}$$

The definitions (1.10) and (1.11) yield

$$\mu_0^{G_k} = e_k \mu^{G_k} \text{ and } \mu_0^{z_k,t_k} = e'_k \mu^{z_k,t_k} = e'_k \mu^{T_k}, \tag{3.16}$$

where e_k and e'_k come from the normalization, and are given by

$$e_k = \mu^{G_k}(\mathbb{B})^{-1} \text{ and } e'_k = \mu^{T_k}(\mathbb{B})^{-1}. \tag{3.17}$$

Let ψ be any Lipschitz function supported on $\mathbb{B}$ and set $I_k = e_k \int_{G_k(\Sigma)} \psi d\mu^{G_k}$; by (3.16) and (1.9),

$$I_k = e_k \int_{G_k(\Sigma)} \psi d\mu^{G_k} = e_k \int_\Sigma \psi(G_k(\xi)) d\mu(\xi) = e_k \int \psi \circ G_k \, d\mu. \tag{3.18}$$

By (3.3), $\sigma_k = a_k \mu^{R_k} = a_k (R_k)_\sharp \mu$, hence $\mu = a_k^{-1} (R_k^{-1})_\sharp \sigma_k$. Thus a similar computation to the one in (3.18) yields

$$I_k = e_k a_k^{-1} \int \psi \circ G_k \circ R_k^{-1} d\sigma_k. \tag{3.19}$$

A similar computation, with G_k replaced by T_k, yields

$$I'_k := \int \psi \mu_0^{z_k,t_k} = e'_k a_k^{-1} \int \psi \circ T_k \circ R_k^{-1} d\sigma_k. \tag{3.20}$$

We want to apply (3.19) and (3.20) to special functions ψ. Let φ be a compactly supported 1-Lipschitz function. Let

$$\psi = \varphi \circ R_k \circ T_k^{-1}. \tag{3.21}$$

Note that ψ is a Lipschitz function with constant less or equal to $r_k^{-1} t_k \leq 2 r_k^{-1} \rho_k$ (see (3.11)). If φ is supported in $B(0, R)$, for k large enough the support of ψ is contained in

$$\begin{aligned} T_k \circ R_k^{-1}(B(0,R)) &= T_k(B(0, r_k R)) = B\big(\frac{-z_k}{t_k}, t_k^{-1} r_k R\big) \\ &\subset B(0, r_k t_k^{-1}(R + \eta_k + |y_k|)) \subset B(0,1), \end{aligned}$$

where we have used (3.11), (3.8), and (3.7). Because of (3.10), $\widetilde{\psi} = (2r_k^{-1}\rho_k)^{-1}\psi$ is 1-Lipschitz; then (3.14) applies to $\widetilde{\psi}$, and (3.10) yields

$$|I'_k - I_k| \leq 4A_k r_k^{-1}\rho_k \leq 4r_k^{-1}\rho_k C_\delta^{2+\log_2(\rho_k/(\eta_k r_k))}\alpha_k =: \widetilde{\alpha}_k \tag{3.22}$$

where I_k and I'_k are as in (3.18) and (3.20) with ψ coming from (3.21). The final identity is the definition of $\widetilde{\alpha}_k$. Notice that even though φ does not depend on k, ψ does, but this is not an issue.

Note that by (3.21) and (3.20), we have

$$I'_k = e'_k a_k^{-1} \int \varphi d\sigma_k. \tag{3.23}$$

Similarly, by (3.19) and (3.21) we have

$$I_k = e_k a_k^{-1} \int \varphi \circ R_k \circ T_k^{-1} \circ G_k \circ R_k^{-1} d\sigma_k. \tag{3.24}$$

Set

$$H_k = R_k \circ T_k^{-1} \circ G_k \circ R_k^{-1}. \tag{3.25}$$

Then $H_k \in \mathscr{G}$, and by (3.15), its dilation factor $\lambda(H_k)$ (which is also the n-th root of its Jacobian determinant) is such that

$$\lambda(H_k)^{-1} = \lambda(T_k)\lambda(G_k)^{-1} = t_k^{-1}\lambda(G_k)^{-1} \in [\lambda_1, \lambda_2], \tag{3.26}$$

because $G_k \in \mathscr{G}(z_k, t_k)$, and by the definition (1.12). Note that by (3.3), the fact that $G_k \in \mathscr{G}(z_k, t_k)$ (see (1.12) and (1.8)), and (3.15), we have

$$\begin{aligned} H_k(r_k^{-1}z_k) &= R_k \circ T_k^{-1} \circ G_k \circ R_k^{-1}(r_k^{-1}z_k) = R_k \circ T_k^{-1} \circ G_k(z_k) \\ &= R_k \circ T_k^{-1}(0) = R_k(z_k) = r_k^{-1}z_k. \end{aligned} \tag{3.27}$$

Notice also that by (3.11)

$$|r_k^{-1}z_k - y| \leq |r_k^{-1}z_k - y_k| + |y_k - y| = r_k^{-1}|z_k - r_k y_k| + |y_k - y| \leq \eta_k + |y_k - y|.$$

Thus $|r_k^{-1}z_k - y|$ tends to 0 by (3.7) and (3.8), so

$$\lim_{k\to\infty} r_k^{-1}z_k = y. \tag{3.28}$$

Combining (3.26), (3.27), and (3.28), we deduce that H_k lies in a compact subset of $\mathscr{G}$. Thus we can replace $\{r_k\}$ by a subsequence for which the H_k converge to a limit H. In addition, (3.26), (3.27) and (3.28) imply that

$$\lambda(H)^{-1} \in [\lambda_1, \lambda_2] \text{ and } H(y) = y. \tag{3.29}$$

Combining (3.22), (3.23), (3.24), and (3.25) we see that if φ is a compactly supported 1-Lipschitz function, then for k large

$$\left| e_k a_k^{-1} \int \varphi \circ H_k \, d\sigma_k - e'_k a_k^{-1} \int \varphi d\sigma_k \right| \leq \widetilde{\alpha}_k. \tag{3.30}$$

By (3.17), (1.9), and (3.15),

$$e'_k = \mu^{T_k}(\mathbb{B})^{-1} = \mu(T_k^{-1}(\mathbb{B}))^{-1} = \mu(B(z_k, t_k))^{-1}. \tag{3.31}$$

Similarly notice that $G_k(z_k) = 0$ because $G_k \in \mathscr{G}(z_k, t_k) \subset \mathscr{G}(z_k)$ (see (1.12) and (1.8)); then (3.17) and (1.9) yield

$$e_k = \mu^{G_k}(\mathbb{B})^{-1} = \mu(G_k^{-1}(\mathbb{B}))^{-1} = \mu(B(z_k, \lambda(G_k)^{-1}))^{-1}. \tag{3.32}$$

In addition,

$$B(z_k, \lambda_1 t_k) \subset G_k^{-1}(\mathbb{B}) = B(z_k, \lambda(G_k)^{-1}) \subset B(z_k, \lambda_2 t_k) \tag{3.33}$$

because $G_k(z_k) = 0$ and by (1.12). Then (1.2), (3.31), (3.32), (3.33), and the fact that $\lambda_1 > 1$ yield

$$C^{-1} e'_k \leq e_k \leq e'_k \tag{3.34}$$

for some constant C that depends on C_δ and λ_2.

To estimate a_k, consider a test function f such that $\mathbb{1}_{\mathbb{B}} \leq f \leq \mathbb{1}_{2\mathbb{B}}$. By definition of σ, $\int f d\sigma = \lim_{k\to\infty} \int f d\sigma_k$. By (1.2), (3.3) and the definition above (1.11), we have

$$a_k \mu(B(0, r_k)) = \sigma_k(\mathbb{B}) \leq \int f d\sigma_k \leq \sigma_k(2\mathbb{B}) = a_k \mu(B(0, 2r_k)) \leq C_\delta a_k \mu(B(0, r_k)). \tag{3.35}$$

Moreover, since σ is also doubling (see the remark below (1.20)), we have that

$$\sigma(\mathbb{B}) \leq \int f d\sigma \leq \sigma(2\mathbb{B}) \leq C_\delta^2 \sigma(\mathbb{B}). \tag{3.36}$$

Thus by (3.35), (3.36) and the definition of σ there exists $C > 1$ such that for k large,

$$C^{-1} a_k \mu(B(0, r_k)) \leq \sigma(\mathbb{B}) \leq C C_\delta^2 a_k \mu(B(0, r_k)). \tag{3.37}$$

Recall that $\rho_k \leq t_k \leq 2\rho_k$ by (3.11), and that $r_k^{-1} z_k$ is bounded by (3.28); since $r_k^{-1}\rho_k$ tends to $+\infty$ by (3.8) we get that for k large, $|z_k| < C r_k < \rho_k \leq t_k$, and so $B(z_k, t_k) \subset B(0, 2t_k) \subset B(0, 4\rho_k)$. By (3.31), since $0 \in \Sigma$, and by (1.2),

$$(e'_k)^{-1} = \mu(B(z_k, t_k)) \leq \mu(B(0, 4\rho_k)) \leq C_\delta^{3+\log_2(\rho_k/r_k)} \mu(B(0, r_k)). \tag{3.38}$$

Combining (3.37) and (3.38) we obtain

$$a_k \leq C C_\delta^{5+\log_2(\rho_k/r_k)} e'_k \cdot \sigma(\mathbb{B}). \tag{3.39}$$

Return to (3.30), set $b_k = e_k / e'_k$, and observe that by (3.39)

$$\left| b_k \int \varphi \circ H_k \, d\sigma_k - \int \varphi d\sigma_k \right| \leq \frac{a_k \widetilde{\alpha}_k}{e'_k} \leq C C_\delta^{5+\log_2(\rho_k/r_k)} \widetilde{\alpha}_k \sigma(\mathbb{B}). \tag{3.40}$$

Now we choose ρ_k and η_k. Denote by

$$\beta_k = C C_\delta^{5+\log_2(\rho_k/r_k)} \widetilde{\alpha}_k = C C_\delta^{5+\log_2(\rho_k/r_k)} \cdot 4 r_k^{-1} \rho_k C_\delta^{2+\log_2(\rho_k/(\eta_k r_k))} \alpha_k$$

the right-hand side of (3.40) (see (3.22)). Since α_k tends to 0 by (3.5), we can choose ρ_k and η_k so that the constraints (3.8) and (3.6) hold, but the convergence in (3.8) is slow enough so that

$$\lim_{k\to\infty} \beta_k = 0. \tag{3.41}$$

Recall from (3.34) that $C^{-1} \leq b_k \leq 1$, hence modulo passing to a subsequence (which we relabel) we can guarantee that $\lim_{k\to\infty} b_k = b > 0$. Letting k tend to infinity in (3.40) we obtain

$$\lim_{k\to\infty} \int \varphi d\sigma_k = \int \varphi d\sigma. \tag{3.42}$$

Since φ is Lipschitz and compactly supported, so is $\varphi \circ H$ and

$$\lim_{k\to\infty} \int \varphi \circ H \, d\sigma_k = \int \varphi \circ H \, d\sigma \tag{3.43}$$

because σ is the weak limit of the σ_k. Note that there is also a ball B such that for k large $\varphi \circ H(x) = \varphi \circ H_k(x) = 0$ for $x \in \mathbb{R}^n \setminus B$; then

$$\begin{aligned} \int |\varphi \circ H - \varphi \circ H_k| \, d\sigma_k &\leq ||\varphi||_{lip} \int_B |H - H_k| \, d\sigma_k \\ &\leq ||\varphi||_{lip} ||H - H_k||_{L^\infty(B)} \sigma_k(B) \\ &\leq 2||\varphi||_{lip} ||H - H_k||_{L^\infty(B)} \sigma(2B). \end{aligned} \tag{3.44}$$

Thus

$$\lim_{k\to\infty} \left| \int \varphi \circ H_k \, d\sigma_k - \int \varphi \circ H \, d\sigma_k \right| = 0. \tag{3.45}$$

Combining (3.40), (3.41), (3.42), (3.43), (3.44) and (3.45) we obtain that for any 1-Lipschitz function φ with compact support,

$$b \int \varphi \circ H \, d\sigma = \int \varphi \, d\sigma. \tag{3.46}$$

Since the Radon measure σ is regular, (3.46) also holds for characteristic functions of Borel sets. Hence $bH_\sharp \sigma = \sigma$.

Recall that $\lambda(H)^{-1} \in [\lambda_1, \lambda_2]$ and $H(y) = y$ by (3.29); thus the conclusion of Lemma 3.1 hold, with $c = b^{-1}$. □

4 Self-similar measures are flat

In this section we complete the proof of Theorem 1.3. Using the notation in Section 3, our goal is to show that if $\sigma \in \text{Tan}(\mu, x_0)$, where $x_0 \in \Sigma_0$ (see (1.23)) then σ is a flat measure. Lemma 3.1 guarantees that for each $y \in \Xi$ (the support of σ), there is a transformation $H(y) \in \mathscr{G}$ and a constant $c(y) > 0$ such that

$$\lambda(y) := \lambda(H(y)) \in [\lambda_2^{-1}, \lambda_1^{-1}] \tag{4.1}$$

and

$$H(y)_\sharp \sigma = c(y)\sigma. \tag{4.2}$$

By definition of $\mathscr{G}$ (see (1.6)), the linear part of $H(y)$ is of the form $\lambda(y)R(y)$, where $R(y)$ is a linear isometry. Since $H(y)$ fixes y, this means that

$$H(y)(u) = y + \lambda(y)R(y)(u-y) \text{ for } u \in \mathbb{R}^n. \tag{4.3}$$

The next lemma allows us to replace $H(y)$ with one of its large powers, chosen so that its isometric part is very close to the identity.

Lemma 4.1. *For each choice of $\varepsilon > 0$, there is an integer m_0, that depends only on ε and n, such that for each $y \in \Xi$ and each integer $\ell \geq 1$, there is an integer $m(y) \in [1, m_0]$ such that*

$$||R(y)^{m(y)\ell} - I|| \leq \varepsilon. \tag{4.4}$$

Proof. Here I is the identity mapping. We use the compactness of the group of linear isometries of $\mathbb{R}^n$ to choose m_0 large enough so that if $R_1, \ldots, R_{m_0}$ are linear isometries, we can find integers m_1, m_2 such that $1 \leq m_1 < m_2 < m_0$ and $||R_{m_2} - R_{m_1}|| \leq \varepsilon$. We apply this with $R_m = R(y)^{m\ell}$ to find m_1 and m_2 such that $||R(y)^{m_2\ell} - R(y)^{m_1\ell}|| \leq \varepsilon$. Then $||R(y)^{(m_2-m_1)\ell} - I|| \leq \varepsilon$, as needed. □

Next we study elementary properties of Ξ. Notice that by (4.2), $H(y)(\Xi) = \Xi$, and iterations also yield

$$H(y)^m(\Xi) = \Xi \text{ for } m \geq 1. \tag{4.5}$$

Lemma 4.2. *The set Ξ is convex.*

Proof. Let $x, y \in \Xi$ be given. Our goal is to show that the segment $[x, y]$ is contained in Ξ. For each $\varepsilon > 0$ and $\ell \geq 1$, we construct a sequence $\{y_k\}$ in Ξ (depending on ε and ℓ) which will allow us to estimate how far $[x, y]$ is from Ξ. We start with $y_0 = y$. If $k \geq 0$ and $y_k \in \Xi$ has been defined, we define y_{k+1} as follows. Set

$$H_k = H(y_k)^{m(y_k)\ell} \text{ and } y_{k+1} = H_k(x). \tag{4.6}$$

By (4.5) and since $x \in \Xi$, $y_{k+1} \in \Xi$. Let us show that for ℓ large and ε small, the y_k stay close to the segment $[x, y]$ and converge (slowly) to x. First observe that by iterations of (4.3), H_k is given by

$$H_k(u) = y_k + \lambda'_k R_k(u - y_k) \text{ for } u \in \mathbb{R}^n, \tag{4.7}$$

with

$$\lambda'_k = \lambda(H(y_k))^{m(y_k)\ell} \in [\lambda_2^{-m_0\ell}, \lambda_1^{-\ell}] \text{ and hence } \lambda'_k < 1. \tag{4.8}$$

In addition, $R_k = R(y_k)^{m(y_k)\ell}$ and therefore, by (4.4), $||R_k - I|| \leq \varepsilon$.

Set $r_k = |x - y_k|$ and $y^*_{k+1} = y_k + \lambda'_k(x - y_k)$. Notice that by (4.7) and (4.4),

$$\begin{aligned} |y_{k+1} - y^*_{k+1}| &= \big|[y_k + \lambda'_k R_k(x - y_k)] - [y_k + \lambda'_k(x - y_k)]\big| \\ &= \big|\lambda'_k[R_k - I](x - y_k)\big| \leq \varepsilon\lambda'_k r_k. \end{aligned} \tag{4.9}$$

Since

$$|y^*_{k+1} - x| = (1 - \lambda'_k)|y_k - x| = (1 - \lambda'_k)r_k, \tag{4.10}$$

we also get that if $\varepsilon < 1/2$,

$$\begin{aligned} r_{k+1} &= |y_{k+1} - x| \le |y^*_{k+1} - x| + |y_{k+1} - y^*_{k+1}| \\ &\le (1 - \lambda'_k)r_k + \varepsilon\lambda'_k r_k \le (1 - \lambda'_k/2)r_k. \end{aligned} \tag{4.11}$$

Then since $y^*_{k+1} \in [x, y_k]$, and by (4.9) and (4.11) we have

$$\mathrm{dist}(y_{k+1}, [x, y_k]) \le |y_{k+1} - y^*_{k+1}| \le \varepsilon\lambda'_k r_k \le 2\varepsilon(r_k - r_{k+1}). \tag{4.12}$$

By elementary geometry,

$$\mathrm{dist}(y_{k+1}, [x, y]) \le \mathrm{dist}(y_{k+1}, [x, y_k]) + \mathrm{dist}(y_k, [x, y]). \tag{4.13}$$

An iteration of (4.12) combined with (4.13) yields

$$\mathrm{dist}(y_{k+1}, [x, y]) \le 2\varepsilon \sum_{0 \le j \le k} (r_j - r_{j+1}) \le 2\varepsilon r_0 = 2\varepsilon|x - y|. \tag{4.14}$$

Notice that by (4.11) and (4.8) r_k tends to 0. Thus we have constructed a sequence $\{y_k\}$ in Ξ, which goes from $y = y_0$ to $x = \lim_{k\to\infty} y_k$. The points y_k lie within $2\varepsilon|x - y|$ from $[x, y]$. Using the definition of y^*_{k+1}, (4.9), and (4.8) we can estimate their successive distances

$$\begin{aligned} |y_{k+1} - y_k| &\le |y_{k+1} - y^*_{k+1}| + |y^*_{k+1} - y_k| \le \varepsilon\lambda'_k r_k + \lambda'_k r_k \\ &\le 2\lambda'_k r_k \le 2\lambda_1^{-\ell}|x - y|. \end{aligned} \tag{4.15}$$

Let z_k be the orthogonal projection of y_k into the line containing x and y. Note that by (4.14) $z_k \in \left[x - 2\varepsilon|x - y|\frac{y-x}{|x-y|}, y + 2\varepsilon|x - y|\frac{y-x}{|x-y|}\right]$. By (4.15), $|z_{k+1} - z_k| \le 2\lambda_1^{-\ell}|x - y|$ and by the definition of z_k, $|z_k - y_k| \le 2\varepsilon|x - y|$. Therefore every point of $[x, y]$ lies within $2(\varepsilon + \lambda_1^{-\ell})|x - y|$ of some y_k, that is each point of $[x, y]$ is at most $2(\varepsilon + \lambda_1^{-\ell})|x - y|$ away from Ξ. Choosing ε arbitrarly small and ℓ arbitrarly large, we get that $[x, y] \subset \Xi$. Lemma 4.2 follows. □

Lemma 4.3. *The set Ξ is a vector subspace of $\mathbb{R}^n$.*

Proof. Let V be the smallest affine subspace of $\mathbb{R}^n$ that contains Ξ, and let d denote its dimension. Choose $d + 1$ affinely independent points $y_0, \dots y_d$ in V (this means that the vectors $y_j - y_0$, $j \ge 1$, are linearly independent), and set $y = \frac{1}{d+1}\sum_{j=0}^{d} y_j$. By Lemma 4.2, Ξ contains the convex hull of the y_j, so there is a small radius $r > 0$ such that $V \cap B(y, r) \subset \Xi$.

Recall from (4.5) that $H(y)^m(\Xi) = \Xi$ for $m \ge 1$. Applying the bijection $H(y)^{-m}$ to both sides, we see that $H(y)^{-m}(\Xi) = \Xi$ for $m \ge 1$, so

$$H(y)^{-m}(V \cap B(y, r)) \subset \Xi. \tag{4.16}$$

We know that $H(y)^{-m}(V \cap B(y,r))$ is a nontrivial open subset of the affine space $H(y)^{-m}(V)$, and since $\Xi \subset V$ we get that $H(y)^{-m}(V) \subset V$. Then by a dimension count $H(y)^{-m}(V) = V$, and by the description (4.3) of $H(y)$, we see that

$$H(y)^{-m}(V \cap B(y,r)) = V \cap B(y, \lambda(H(y))^{-m} r). \tag{4.17}$$

Recall that $\lambda(H(y)) < 1$; then $\lambda(H(y))^{-m} r$ is as large as we want by picking m as large as we need. Hence (4.16) guarantees that Ξ contains V, and this means that $\Xi = V$.

Note that Remark 14.4 (2) in [6] ensures that $0 \in \Xi$, as Ξ is the support of σ and $\sigma \in \mathrm{Tan}(\mu, x_0)$ where μ is a doubling measure and $x_0 \in \Sigma$. Hence $\Xi = V$ is a vector space. This completes our proof of Lemma 4.3. □

Now we study the distribution of σ on Ξ. If Ξ is reduced to the origin, then σ is a Dirac mass, and Dirac masses lie in $\mathscr{F}_0$. Thus in this case σ is trivially flat. We may now assume that Ξ is a vector space of dimension $d > 0$.

Lemma 4.4. *There is a dimension $D \geq 0$ such that σ is Ahlfors-regular of dimension D, which means that*

$$C^{-1}\rho^D \leq \sigma(B(y,\rho)) \leq C\rho^D \ \text{ for } y \in \Xi \text{ and } \rho > 0 \tag{4.18}$$

for some constant $C \geq 1$.

Proof. Set $\lambda(y) = \lambda(H(y))$, and recall from (4.1) that $1 < \lambda_1 \leq \lambda(y)^{-1} \leq \lambda_2$. Notice that by (4.2), for $y \in \Xi$ and $r > 0$

$$\begin{aligned} \sigma(B(y, \lambda(y)^{-1} r)) &= \sigma(H(y)^{-1}(B(y,r))) = H(y)_\sharp \sigma(B(y,r)) \\ &= c(y)\sigma(B(y,r)). \end{aligned} \tag{4.19}$$

Iterating we obtain for $m \geq 0$

$$\sigma(B(y, \lambda(y)^{-m} r)) = c(y)^m \sigma(B(y,r)). \tag{4.20}$$

Applying (4.20) to $\lambda(y)^m r$ instead of r we have that (4.20) also holds for $m \leq 0$. Observe that since $\lambda(y) < 1$ and $\sigma(B(y,r)) > 0$ when $y \in \Xi$, (4.20) yields $c(y) \geq 1$. If $c(y) = 1$, then $\sigma(B(y, \lambda(y)^{-m} r)) = \sigma(B(y,r))$ for all $m \in \mathbb{Z}$, and σ is a Dirac mass. This case was excluded before the statement of the lemma, so $c(y) > 1$.

Now let $\rho > 0$ be given, and choose m such that

$$\lambda(y)^{-m} \leq \rho \leq \lambda(y)^{-m-1}. \tag{4.21}$$

By (4.20) applied to $r = 1$,

$$c(y)^m \sigma(B(y,1)) \leq \sigma(B(y,\rho)) \leq c(y)^{m+1} \sigma(B(y,1)), \tag{4.22}$$

hence, letting $\ell = \log(\sigma(B(y,1)))$, (4.22) yields

$$m \log(c(y)) + \ell \leq \log(\sigma(B(y,\rho))) \leq (m+1)\log(c(y)) + \ell. \tag{4.23}$$

By (4.21)

$$m\log(\lambda(y)^{-1}) \leq \log\rho \leq (m+1)\log(\lambda(y)^{-1}). \tag{4.24}$$

Hence combining (4.23) and (4.24) we have

$$\lim_{\rho\to+\infty} \frac{\log(\sigma(B(y,\rho)))}{\log\rho} = \frac{\log(c(y))}{\log(\lambda(y)^{-1})} =: D(y). \tag{4.25}$$

We claim that $D(y)$ does not depend on y. Indeed, if $z \in \Xi$, observe that $B(y,\rho) \subset B(z, \rho + |z-y|)$, hence

$$\begin{aligned} D(y) &= \lim_{\rho\to+\infty} \frac{\log(\sigma(B(y,\rho)))}{\log\rho} \leq \liminf_{\rho\to+\infty} \frac{\log(\sigma(B(z,\rho+|z-y|)))}{\log\rho} \\ &= \liminf_{\rho\to+\infty} \frac{\log(\sigma(B(z,\rho+|z-y|)))}{\log(\rho+|z-y|)} = D(z); \end{aligned} \tag{4.26}$$

the opposite inequality also holds exchanging the roles of y and z. Let D be the common value of the $D(y)$ for $y \in \Xi$. By definition (see (4.25)), $\lambda(y)^{-D} = c(y)$, and using (4.21) we can rewrite (4.22) as

$$\lambda(y)^{-mD}\sigma(B(y,1)) \leq \sigma(B(y,\rho)) \leq \lambda(y)^{-D}\lambda(y)^{-mD}\sigma(B(y,1)). \tag{4.27}$$

Thus, by (4.21),

$$\lambda(y)^{D}\rho^{D}\sigma(B(y,1)) \leq \sigma(B(y,\rho)) \leq \lambda(y)^{-D}\rho^{D}\sigma(B(y,1)), \tag{4.28}$$

which yields (4.18) with $C = \lambda_2^D \max(\sigma(B(y,1)), \sigma(B(y,1))^{-1})$ (because (4.1) guarantees that $\lambda(y)^{-1} \leq \lambda_2$). □

Lemma 4.5. *Let d be the dimension of the vector space Ξ, and denote by $\nu = \mathscr{H}^d \llcorner \Xi$ the restriction of $\mathscr{H}^d$ to Ξ. Let D be as in* (4.18)*. Then $d = D$, and there exists a constant $c_0 > 0$ such that $\sigma = c_0\nu$.*

Proof. Since σ is Ahlfors regular of dimension D (by Lemma 4.4), a standard covering argument (see for instance Lemma 18.11 in [2]) guarantees that there exists a constant $C > 0$ such that

$$C^{-1}\mathscr{H}^D \llcorner \Xi \leq \sigma \leq C\mathscr{H}^D \llcorner \Xi, \tag{4.29}$$

and Ξ is a D-dimensional Ahlfors regular set. Hence $D = d$.

By (4.29), σ is absolutely continuous with respect to ν, and the Radon-Nikodym derivative of σ with respect to ν is bounded. Thus there is a bounded function f on Ξ such that $\sigma = f\nu$.

We now show that f is constant. First observe that since $d = D = \frac{\log(c(y))}{\log(\lambda(y)^{-1})}$ (by (4.25)), (4.20) guarantees that for $y \in \Xi$, $r > 0$, and $m \in \mathbb{Z}$

$$\sigma(B(y, \lambda(y)^{-m}r)) = c(y)^m\sigma(B(y,r)) = \lambda(y)^{-md}\sigma(B(y,r)). \tag{4.30}$$

Since $\nu(B(y,\lambda(y)^{-m}r)) = \lambda(y)^{-md}\nu(B(y,r))$, we may rewrite (4.30) as

$$\frac{\sigma(B(y,\lambda(y)^{-m}r))}{\nu(B(y,\lambda(y)^{-m}r))} = \frac{\sigma(B(y,r))}{\nu(B(y,r))}. \tag{4.31}$$

The Lebesgue differentiation theorem says that for ν-almost every $y \in \Xi$,

$$f(y) = \lim_{\rho\to 0} \frac{\sigma(B(y,\rho))}{\nu(B(y,\rho))}. \tag{4.32}$$

For such an y and every $r > 0$, by (4.32) and (4.31) for $-m$

$$f(y) = \lim_{m\to\infty} \frac{\sigma(B(y,\lambda(y)^{m}r))}{\nu(B(y,\lambda(y)^{m}r))} = \frac{\sigma(B(y,r))}{\nu(B(y,r))}. \tag{4.33}$$

That is,

$$\sigma(B(y,r)) = f(y)\nu(B(y,r)) \text{ for } r > 0. \tag{4.34}$$

If $z \in \Xi$ is another Lebesgue point of f, since $B(z,r) \subset B(y, r+|y-z|)$ we have

$$\begin{aligned} f(z) &= \lim_{r\to\infty} \frac{\sigma(B(z,r))}{\nu(B(z,r))} \le \liminf_{r\to\infty} \frac{\sigma(B(y,r+|y-z|))}{\nu(B(z,r))} \\ &= \liminf_{r\to\infty} \frac{\sigma(B(y,r+|y-z|))}{\nu(B(y,r+|y-z|))} = f(y). \end{aligned} \tag{4.35}$$

Similarly $f(y) \le f(z)$, and f is constant. □

This completes the proof of Theorem 1.3. In fact we have proved that if $\sigma \in \mathrm{Tan}(\mu, x_0)$ with $x_0 \in \Sigma_0$ (see (1.23)) then $\sigma = c_0\mathscr{H}^d \llcorner \Xi$ for some vector space Ξ.

5 A smoother version of the Wasserstein $\mathbb{W}_1$ distance

So far we managed to work with the distance $\mathbb{W}_1$ defined by (1.5), but for the proof of Theorem 1.2, it is more convenient to use a slightly smoother variant, which attenuates the possible discontinuities in $r > 0$ of the normalizing factors $\mu^{x,r}(B(0,1))^{-1} = \mu(B(x,r))^{-1}$.

Let us choose a smooth radial function φ such that

$$\mathbb{1}_{B(0,1/2)} \le \varphi \le \mathbb{1}_{B(0,1)}; \tag{5.1}$$

If μ and ν are two Radon measures such that

$$\mu(B(0,1/2)) > 0 \text{ and } \nu(B(0,1/2)) > 0, \tag{5.2}$$

we define a new distance $\mathbb{W}_\varphi(\mu,\nu)$ by

$$\mathbb{W}_\varphi(\mu,\nu) = \sup_{\psi\in \mathrm{Lip}_1(\mathbb{B})} \left| \frac{\int \psi\varphi d\mu}{\int \varphi d\mu} - \frac{\int \psi\varphi d\nu}{\int \varphi d\nu} \right|. \tag{5.3}$$

Recall that $\mathrm{Lip}_1(\mathbb{B})$ is defined near (1.3). The distance $\mathbb{W}_1$ above essentially corresponds to $\varphi = \mathbb{1}_{\mathbb{B}}$ here.

We required (5.2) (and (5.1)) to make sure that we do not divide by 0. But notice that even when $\mu(B(0,1/2))$ or $\nu(B(0,1/2))$ is very small, we always get that

$$\mathbb{W}_\varphi(\mu,\nu) \le 2, \tag{5.4}$$

because $|\int \psi\varphi d\mu| \le \int \varphi d\mu$ and similarly for ν. Note that

$$\mathbb{W}_\varphi(a\mu, b\nu) = \mathbb{W}_\varphi(\mu,\nu) \text{ for } a,b>0, \tag{5.5}$$

so we do not need to normalize μ and ν in advance. Finally observe that $\mathbb{W}_\varphi$ satisfies the triangle inequality. That is, if σ is a third Radon measure such that $\sigma(B(0,1/2)) > 0$, if follows at once from the definition that

$$\mathbb{W}_\varphi(\mu,\sigma) \le \mathbb{W}_\varphi(\mu,\nu) + \mathbb{W}_\varphi(\nu,\sigma). \tag{5.6}$$

Let us check that if we restrict to measures that are not to small on $B(0,1/2)$, then $\mathbb{W}_1$ controls $\mathbb{W}_\varphi$.

Lemma 5.1. *Let μ and ν be Radon measures such that* (5.2) *holds and*

$$\mu(\mathbb{B}) = \nu(\mathbb{B}) = 1. \tag{5.7}$$

Then

$$\mathbb{W}_\varphi(\mu,\nu) \le \frac{1+2||\varphi||_{lip}}{\mu(B(0,1/2))}\,\mathbb{W}_1(\mu,\nu), \tag{5.8}$$

where $||\varphi||_{lip}$ denotes the Lipschitz norm of φ.

The fact that the estimate is not symmetric is not an issue. In particular we shall apply (5.8) to doubling measures μ and ν; in this case $\mu(B(0,1/2)) \sim \mu(\mathbb{B}) = 1 = \nu(\mathbb{B}) \sim \nu(B(0,1/2))$.

Proof. Let $\psi \in \mathrm{Lip}_1(\mathbb{B})$ be given. The definition (1.5), applied to $\psi\varphi$, yields

$$\Big|\int \psi\varphi d\mu - \int \psi\varphi d\nu\Big| \le ||\psi\varphi||_{lip}\mathbb{W}_1(\mu,\nu) \le (1+||\varphi||_{lip})\mathbb{W}_1(\mu,\nu). \tag{5.9}$$

The same definition, applied to φ itself, yields

$$\Big|\int \varphi d\mu - \int \varphi d\nu\Big| \le ||\varphi||_{lip}\mathbb{W}_1(\mu,\nu). \tag{5.10}$$

Set

$$\Delta = \left|\frac{\int \psi\varphi d\mu}{\int \varphi d\mu} - \frac{\int \psi\varphi d\nu}{\int \varphi d\nu}\right| \tag{5.11}$$

and write

$$\Delta = \Big|\frac{a}{b} - \frac{c}{d}\Big| = \left|\frac{ad-bc}{bd}\right| = \frac{|d(a-c)+c(d-b)|}{bd} \tag{5.12}$$

with $a = \int \psi\varphi d\mu$, $b = \int \varphi d\mu$, $c = \int \psi\varphi d\nu$, and $d = \int \varphi d\nu$. Notice that $|c| \le d$ because $\psi \in \mathrm{Lip}_1(\mathbb{B})$ and $\varphi \ge 0$. Also, $b \ge \mu(B(0,1/2))$ by (5.1). Hence by (5.9) and (5.10); we have

$$\Delta \le \frac{|a-c|+|d-b|}{b} \le \frac{1+2||\varphi||_{lip}}{\mu(B(0,1/2))}\,\mathbb{W}_1(\mu,\nu). \tag{5.13}$$

Taking the supremum over $\psi \in \mathrm{Lip}_1(\mathbb{B})$ in (5.13) (recall (5.11)) yields (5.8). □

The following lemma specifies the sense in which $\mathbb{W}_\varphi$ is more stable that $\mathbb{W}_1$.

Lemma 5.2. *Let μ and ν be Radon measures and let $\theta \in (0, 1/2]$ be such that*

$$\mu(B(0,\theta/2)) > 0 \text{ and } \nu(B(0,\theta/2)) > 0, \tag{5.14}$$

and define new measures μ_1 and ν_1 by

$$\mu_1(A) = \mu(\theta A) \text{ and } \nu_1(A) = \nu(\theta A) \text{ for } A \subset \mathbb{R}^n. \tag{5.15}$$

Then

$$\mathbb{W}_\varphi(\mu_1, \nu_1) \leq \theta^{-1}(1 + 4||\varphi||_{lip}) \frac{\mu(B(0,1))}{\mu(B(0,\theta/2))} \mathbb{W}_\varphi(\mu, \nu). \tag{5.16}$$

As in (5.8) the estimate is not symmetric in μ and ν, but is nonetheless valid. We require that $\mu(B(0,\theta/2)) \neq 0$ and $\nu(B(0,\theta/2)) \neq 0$ to make sure that $\mathbb{W}_\varphi(\mu_1, \nu_1)$ is easily defined. Often μ is the restriction to $\mathbb{B}$ of a doubling measure and its support contains the origin; then $\frac{\mu(B(0,1))}{\mu(B(0,\theta/2))} \geq C^{-1}$, for some C that depends on θ and the doubling constant C_δ.

Proof. Let $\psi \in \mathrm{Lip}_1(\mathbb{B})$ be given; we want to control the quantity

$$\Delta = \left| \frac{\int \psi\varphi d\mu_1}{\int \varphi d\mu_1} - \frac{\int \psi\varphi d\nu_1}{\int \varphi d\nu_1} \right| =: \left| \frac{a}{b} - \frac{c}{d} \right| \tag{5.17}$$

(as above, but with integrals relative to μ_1 and ν_1). Notice that by (5.15),

$$a = \int \psi\varphi d\mu_1 = \int \psi(\theta^{-1}x)\varphi(\theta^{-1}x)d\mu(x) = \int \psi(\theta^{-1}x)\varphi(\theta^{-1}x)\varphi(x)^2 d\mu(x), \tag{5.18}$$

where we just use the fact that $\varphi(\theta^{-1}x) = 0$ when $x \in \mathbb{R}^n \setminus B(0, 1/2)$, and the special shape of φ in (5.1), to add an extra $\varphi^2(x)$. Similarly,

$$c = \int \psi\varphi d\nu_1 = \int \psi(\theta^{-1}x)\varphi(\theta^{-1}x)\varphi^2(x)d\nu(x). \tag{5.19}$$

The same computations without ψ yield

$$b = \int \varphi d\mu_1 = \int \varphi(\theta^{-1}x)\varphi(x)^2 d\mu(x), \tag{5.20}$$

$$d = \int \varphi d\nu_1 = \int \varphi(\theta^{-1}x)\varphi(x)^2 d\nu(x). \tag{5.21}$$

It is also useful to introduce

$$e = \int \varphi d\mu \text{ and } f = \int \varphi d\nu. \tag{5.22}$$

Let us first estimate $\delta_1 = \frac{a}{e} - \frac{c}{f}$. We want to apply the definition of $\mathbb{W}_\varphi(\mu, \nu)$ to the function Ψ defined by $\Psi(x) = \psi(\theta^{-1}x)\varphi(\theta^{-1}x)\varphi(x)$. Notice that Ψ is supported in $\mathbb{B}$ (this is why we

added $\varphi(x)$), and its Lipschitz norm is at most $\theta^{-1}(1+||\varphi||_{lip})+||\varphi||_{lip} \le \theta^{-1}(1+2||\varphi||_{lip})$. Thus (5.3) yields

$$|\delta_1| = \left|\frac{a}{e} - \frac{c}{f}\right| \le \theta^{-1}||\Psi||_{lip}\mathbb{W}_\varphi(\mu,\nu) \le \theta^{-1}(1+2||\varphi||_{lip})\mathbb{W}_\varphi(\mu,\nu). \tag{5.23}$$

We can also apply the definition of $\mathbb{W}_\varphi(\mu,\nu)$ to $\varphi(\theta^{-1}x)\varphi(x)$, whose Lipschitz norm is at most $2\theta^{-1}||\varphi||_{lip}$, and we get that

$$|\delta_2| := \left|\frac{b}{e} - \frac{d}{f}\right| \le 2\theta^{-1}||\varphi||_{lip}\mathbb{W}_\varphi(\mu,\nu). \tag{5.24}$$

Thus

$$\frac{a}{b} = \frac{e}{b}\frac{a}{e} = \frac{e}{b}\Big(\frac{c}{f}+\delta_1\Big) = \frac{e}{b}\frac{c}{d}\frac{d}{f} + \frac{e\delta_1}{b} = \frac{e}{b}\frac{c}{d}\Big(\frac{b}{e}+\delta_2\Big) + \frac{e\delta_1}{b} = \frac{c}{d} + \frac{ec\delta_2}{bd} + \frac{e\delta_1}{b},$$

where δ_1 and δ_2 are as in(5.23) and (5.24). Thus

$$\Delta = \Big|\frac{a}{b} - \frac{c}{d}\Big| \le \frac{|ec\delta_2|}{bd} + \frac{|e\delta_1|}{b}. \tag{5.25}$$

Now $|c| \le d$ because $\varphi \ge 0$ and $|\psi| \le 1$, $e = \int \varphi d\mu \le \mu(B(0,1))$, and $b = \int \varphi(\theta^{-1}x)\varphi(x)^2 d\mu(x) \ge \mu(B(0,\theta/2))$ because of (5.1). Thus we have

$$\Delta \le (\delta_1+\delta_2)\frac{e}{b} \le \theta^{-1}(1+4||\varphi||_{lip})\mathbb{W}_\varphi(\mu,\nu)\,\frac{\mu(B(0,1))}{\mu(B(0,\theta/2))}. \tag{5.26}$$

Noting (5.17) and taking the supremum over $\psi \in \mathrm{Lip}_1(\mathbb{B})$, we obtain (5.16). □

The next lemma is used in Section 8. It shows that the distance function $\mathbb{W}_\varphi$ also controls $\mathbb{W}_1$ in some averaged way. Thus $\mathbb{W}_1$ and $\mathbb{W}_\varphi$ are basically interchangeable.

Lemma 5.3. *Let μ and ν are Radon measures such that $\mu(B(0,1/4)) > 0$ and $\nu(B(0,1/4)) > 0$. Define μ_t and ν_t, $1/4 \le t \le 1/2$, by*

$$\mu_t(A) = \frac{\mu(tA)}{\mu(B(0,t))} \;\; \textit{and} \;\; \nu_t(A) = \frac{\mu(tA)}{\nu(B(0,t))} \;\; \textit{for } A \subset \mathbb{R}^n. \tag{5.27}$$

Then

$$\fint_{1/4}^{1/2} \mathbb{W}_1(\mu_t,\nu_t)dt \le \frac{(8+||\varphi||_{lip})\mu(\mathbb{B})}{\mu(B(0,1/4))}\,\mathbb{W}_\varphi(\mu,\nu). \tag{5.28}$$

Proof. First notice that the statement does not change when we multiply μ and ν by positive constants. So we may assume that

$$\int \varphi d\mu = \int \varphi d\nu = 1. \tag{5.29}$$

Next fix $t \in [1/4, 1/2]$ and $\psi \in \mathrm{Lip}_1(\mathbb{B})$. We want to estimate

$$\int \psi d\mu_t - \int \psi d\nu_t = \frac{\int \psi(t^{-1}x)d\mu(x)}{\mu(B(0,t))} - \frac{\int \psi(t^{-1}x)d\nu(x)}{\nu(B(0,t))} =: \frac{a}{b} - \frac{c}{d}. \tag{5.30}$$

As before,

$$\Delta = \frac{|ad - bc|}{bd} = \frac{|d(a-c) - c(b-d)|}{bd} \leq \frac{|a-c| + |b-d|}{b} \tag{5.31}$$

because $|c| = |\int \psi(t^{-1}x)d\nu(x)| \leq ||\psi||_\infty \nu(B(0,t)) \leq d$. Notice that $\varphi(x) = 1$ when $\psi(t^{-1}x) \neq 0$ since this implies that $|t^{-1}x| \leq 1$ and hence $|x| \leq t \leq 1/2$. Thus

$$\begin{aligned} |a-c| &= \left| \int \psi(t^{-1}x)d\mu(x) - \int \psi(t^{-1}x)d\nu(x) \right| \\ &= \left| \int \psi(t^{-1}x)\varphi^2(x)d\mu(x) - \int \psi(t^{-1}x)\varphi^2(x)d\nu(x) \right|. \end{aligned} \tag{5.32}$$

We apply the definition (5.3) of $\mathbb{W}_\varphi$ with the function $x \to \psi(t^{-1}x)\varphi(x)$, which is supported in $\mathbb{B}$ and $(t^{-1} + ||\varphi||_{lip})$-Lipschitz. We obtain that

$$|a-c| \leq (t^{-1} + ||\varphi||_{lip})\mathbb{W}_\varphi(\mu, \nu) \tag{5.33}$$

Notice also that

$$\frac{1}{b} = \frac{\int \varphi d\mu}{\mu(B(0,t))} \leq \frac{\mu(\mathbb{B})}{\mu(B(0,t))} \leq \frac{\mu(\mathbb{B})}{\mu(B(0,1/4))}. \tag{5.34}$$

Thus

$$\begin{aligned} \Delta &\leq \frac{|a-c| + |b-d|}{b} \leq \frac{\mu(\mathbb{B})}{\mu(B(0,1/4))} \left[(t^{-1} + ||\varphi||_{lip})\mathbb{W}_\varphi(\mu,\nu) + |b-d| \right] \\ &= \frac{\mu(\mathbb{B})}{\mu(B(0,1/4))} \left[(t^{-1} + ||\varphi||_{lip})\mathbb{W}_\varphi(\mu,\nu) + |\mu(B(0,t)) - \nu(B(0,t))| \right]. \end{aligned} \tag{5.35}$$

We take the supremum over $\psi \in \mathrm{Lip}_1(\mathbb{B})$ and get that

$$\mathbb{W}_1(\mu_t, \nu_t) \leq \frac{\mu(\mathbb{B})}{\mu(B(0,1/4))} \left[(t^{-1} + ||\varphi||_{lip})\mathbb{W}_\varphi(\mu,\nu) + |\mu(B(0,t)) - \nu(B(0,t))| \right]. \tag{5.36}$$

Since $t^{-1} \in [2, 4]$, (5.28) will follow as soon as we prove that

$$\int_{[1/4,1/2]} |\mu(B(0,t)) - \nu(B(0,t))| dt \leq \mathbb{W}_\varphi(\mu, \nu). \tag{5.37}$$

Let h be a bounded measurable function, defined on $[1/4, 1/2]$; we want to evaluate

$$I_h = \int_{[1/4,1/2]} h(t)[\mu(B(0,t)) - \nu(B(0,t))]dt. \tag{5.38}$$

Observe that by Fubini

$$\int_{[1/4,1/2]} h(t)\mu(B(0,t))dt = \int_{x \in B(0,1/2)} \Big\{ \int \mathbb{1}_{t \in [1/4,1/2]} \mathbb{1}_{t > |x|} h(t)dt \Big\} d\mu(x),$$

and similarly for ν. Set $\psi_h(x) = \int \mathbb{1}_{t\in[1/4,1/2]}\mathbb{1}_{t>|x|}h(t)dt$. This is a $||h||_\infty$-Lipschitz function of $|x|$, which vanishes when $|x| \geq 1/2$, so by (5.1), (5.3) and the normalization in (5.29) we have

$$\begin{aligned} |I_h| &= \left| \int_{B(0,1/2)} \psi_h d\mu - \int_{B(0,1/2)} \psi_h d\nu \right| \\ &= \left| \int \psi_h \varphi d\mu - \int \psi_h \varphi d\nu \right| \leq ||h||_\infty \mathbb{W}_\varphi(\mu,\nu). \end{aligned} \tag{5.39}$$

Thus (5.39) holds for all bounded measurable functions h defined on $[1/4, 1/2]$, and (5.37) follows by duality. We saw earlier that (5.37) implies (5.28). Lemma 5.3 follows. □

We conclude this section with an easy observation concerning the behavior of $\mathbb{W}_\varphi(\mu,\nu)$ when taking weak limits.

Lemma 5.4. *Let μ and ν satisfy* (5.2)*, suppose that μ is the weak limit of some sequence $\{\mu_k\}$, and that ν is the weak limit of some sequence $\{\nu_k\}$. Then*

$$\mathbb{W}_\varphi(\mu,\nu) \leq \liminf_{k\to\infty} \mathbb{W}_\varphi(\mu_k,\nu_k). \tag{5.40}$$

Proof. Set $L_k = \mathbb{W}_\varphi(\mu_k,\nu_k)$ and $L = \liminf_{k\to\infty} L_k$. Notice that $\int \varphi d\mu = \lim_{k\to\infty} \int \varphi d\mu_k$ by weak convergence; by (5.2), this implies that $\int \varphi d\mu_k > 0$ for k large. The same argument applied to a continuous function $f \leq \mathbb{1}_{B(0,1/2)}$ such that $\int f d\mu > 0$ shows that $\mu_k(B(0,1/2)) > 0$ for k large. Similar observations hold for ν and ν_k. For each $\psi \in \mathrm{Lip}_1(\mathbb{B})$, the weak convergence yields $\int \psi\varphi d\mu = \lim_{k\to\infty} \int \psi\varphi d\mu_k$. For k large,

$$\left| \frac{\int \psi\varphi d\mu_k}{\int \varphi d\mu_k} - \frac{\int \psi\varphi d\nu_k}{\int \varphi d\nu_k} \right| \leq L_k.$$

Since each term has a limit and the denominators are bounded away from 0, taking a $\liminf$ we have that

$$\left| \frac{\int \psi\varphi d\mu}{\int \varphi d\mu} - \frac{\int \psi\varphi d\nu}{\int \varphi d\nu} \right| \leq L.$$

Taking the supremum over $\psi \in \mathrm{Lip}_1(\mathbb{B})$ we conclude that (5.40) holds. □

6 Uniqueness of the tangent measure at "good points"

In this section show that for $x \in \Sigma_0 \cap \Sigma_1$, $\mathrm{Tan}(\mu, x)$ is a one-dimensional set of flat measures. Recall that Σ_1 and Σ_0 were defined in (1.16) and (1.23) respectively.

Lemma 6.1. *Let μ be a doubling measure, and let $x \in \Sigma_0 \cap \Sigma_1$. Then there is a nonzero flat measure σ such that Tan$(\mu, x) = \big\{c\sigma\,;\, c > 0\big\}$.*

Proof. Fix μ and $x \in \Sigma_0 \cap \Sigma_1$; without loss of generality, we may assume that $x = 0$. By Theorem 1.3, we know that

$$\mathrm{Tan}(\mu, 0) \subset \mathscr{F} \tag{6.1}$$

where $\mathscr{F}$ denotes the set of flat measures (see (1.21)). By definition of Σ_1,

$$\int_0^1 \alpha_{\mathscr{D}}(0, r) \frac{dr}{r} < \infty. \tag{6.2}$$

Thus it only remains to show that $\mathrm{Tan}(\mu, 0)$ is one-dimensional. Our initial goal is to bound the $\mathbb{W}_\varphi$ distance for two different scaled dilations of μ by $\alpha_{\mathscr{D}}(0, \cdot)$ at the right scale. For each $r \in (0, 1/4)$ we use Chebyshev's inequality to find $r_+ \in [2r, 4r]$ such that

$$\alpha_{\mathscr{D}}(0, r_+) \leq (2r)^{-1} \int_{2r}^{4r} \alpha_{\mathscr{D}}(0, t) dt \leq 2 \int_{2r}^{4r} \alpha_{\mathscr{D}}(0, t) \frac{dt}{t}. \tag{6.3}$$

By the definition of $\alpha_{\mathscr{D}}(0, r_+)$ (see (1.15)): there is a transformation $G \in \mathscr{D}(0, r_+)$, such that

$$\mathbb{W}_1(\mu_0^G, \mu_0^{0,r_+}) \leq 2\alpha_{\mathscr{D}}(0, r_+). \tag{6.4}$$

By the definition (1.14) of $\mathscr{D}(0, r_+)$, G is simply given by $G(u) = \lambda u$, with $\lambda^{-1} = \lambda(G)^{-1} \in [\lambda_1 r_+, \lambda_2 r_+]$. Set

$$r^* = \lambda^{-1} \in [\lambda_1 r_+, \lambda_2 r_+]; \tag{6.5}$$

notice that G is the homotety that sends $B(0, r^*)$ to $\mathbb{B}$, so $\mu_0^G = \mu_0^{0,r^*}$ (see near (1.11)) and now (6.4) says that

$$\mathbb{W}_1(\mu_0^{0,r^*}, \mu_0^{0,r_+}) \leq 2\alpha_{\mathscr{D}}(0, r_+). \tag{6.6}$$

First apply Lemma 5.1 to the measures μ_0^{0,r^*} and μ_0^{0,r_+};(6.6) yields that

$$\mathbb{W}_\varphi(\mu_0^{0,r^*}, \mu_0^{0,r_+}) \leq C\alpha_{\mathscr{D}}(0, r_+) \tag{6.7}$$

(where we do not record any more the dependence on φ or C_δ). Then we apply Lemma 5.2 to the same measures, with $\theta = r/r^*$. Notice that $\theta \leq 1/2$ because $r^* \geq r_+ \geq 2r$.

Recall from (5.15) that μ_1 is defined by

$$\begin{aligned} \mu_1(A) &= \mu_0^{0,r^*}(\theta A) = \mu_0^{0,r^*}(rA/r^*) = \frac{\mu(r^* rA/r^*)}{\mu(B(0, r^*))} \\ &= \frac{\mu(rA)}{\mu(B(0, r^*))} = \frac{\mu(B(0, r))}{\mu(B(0, r^*))} \mu_0^{0,r}(A) \end{aligned} \tag{6.8}$$

by (1.11). Similarly, ν_1 is defined by

$$\begin{aligned} \nu_1(A) &= \mu_0^{0,r_+}(\theta A) = \mu_0^{0,r_+}(rA/r^*) = \frac{\mu(r_+ rA/r^*)}{\mu(B(0, r_+))} \\ &= \frac{\mu(\rho(r)A)}{\mu(B(0, r_+))} = \frac{\mu(B(0, \rho(r)))}{\mu(B(0, r_+))} \mu_0^{0,\rho(r)}(A) \end{aligned} \tag{6.9}$$

where

$$\rho(r) = \frac{rr_+}{r^*} \in [\lambda_2^{-1} r, \lambda_1^{-1} r]. \tag{6.10}$$

By (5.5), (6.8), (6.9), Lemma 5.2, the fact that μ_0^{0,r^*} is also doubling with the same constant as μ (which controls mass ratio in (5.16)), and (6.7),

$$\mathbb{W}_\varphi(\mu_0^{0,r}, \mu_0^{0,\rho(r)}) = \mathbb{W}_\varphi(\mu_1, \nu_1) \leq C\mathbb{W}_\varphi(\mu_0^{0,r^*}, \mu_0^{0,r_+}) \leq C\alpha_{\mathscr{D}}(0, r_+). \tag{6.11}$$

In order to show that $\mathrm{Tan}(\mu, 0)$ is one-dimensional, we define a specific sequence of measures μ^{0,r_j}, which will be used to approximate all the tangent measures of $\mathrm{Tan}(\mu, 0)$ up to a multiplicative constant. We start with $r_0 = 1/4$, and define r_j by induction, taking $r_{j+1} = \rho(r_j)$ for $j \geq 0$. Note that for all choice of integers $0 \leq k \leq l$, by (5.6), (6.11), and (6.3) we have

$$\begin{aligned} \mathbb{W}_\varphi(\mu_0^{0,r_k}, \mu_0^{0,r_{l+1}}) &\leq \sum_{k\leq j\leq l} \mathbb{W}_\varphi(\mu_0^{0,r_j}, \mu_0^{0,r_{j+1}}) \leq C \sum_{k\leq j<l} \alpha_{\mathscr{D}}(0, (r_j)_+) \\ &\leq 2C \sum_{k\leq j<l} \int_{2r_j}^{4r_j} \alpha_{\mathscr{D}}(0, t)\frac{dt}{t}. \end{aligned} \tag{6.12}$$

Recall that $r_{j+1} = \rho(r_j) \leq \lambda_1^{-1} r_j$ (see (6.10)), thus the r_j's decay at a definite rate. Therefore the intervals $[2r_j, 4r_j]$ have bounded overlap, and since they are all contained in $(0, 4r_k]$, we obtain

$$\mathbb{W}_\varphi(\mu_0^{0,r_k}, \mu_0^{0,r_{l+1}}) \leq C \int_0^{4r_k} \alpha_{\mathscr{D}}(0, t)\frac{dt}{t}. \tag{6.13}$$

Let $\sigma \in \mathrm{Tan}(\mu, 0)$ be given. There exist sequences $\{\rho_k\}$ and $\{a_k\}$ such that $\rho_k \in (0, 1/4]$, $\lim_{k\to\infty} \rho_k = 0$, $a_k > 0$, and

$$\sigma_k = a_k \mu^{0,\rho_k} \text{ converges weakly to } \sigma. \tag{6.14}$$

Let $j = j(k)$ denote the largest integer such that $r_j \geq \rho_k$. Thus $j \geq 0$ (because $r_0 = 1/4$), and $r_{j+1} < \rho_k$; since $r_{j+1} = \rho(r_j) \in [\lambda_2^{-1} r_j, \lambda_1^{-1} r_j]$ (by (6.10)), we get that

$$\lambda_2^{-1} r_{j(k)} < \rho_k \leq r_{j(k)}. \tag{6.15}$$

Set $\theta_k = \rho_k / r_{j(k)} \in [\lambda_2^{-1}, 1]$; we may replace $\{\rho_k\}$ by a subsequence such that the θ_k converge to a limit θ. Consider the dilation D_k defined by $D_k(u) = \theta_k u$, and set $D(u) = \theta u$. Notice that by (6.14) and (1.11)

$$[D_k]_\sharp \sigma_k = a_k [D_k]_\sharp \mu^{0,\rho_k} = a_k \mu^{0,\rho_k/\theta_k} = a_k \mu^{0,r_{j(k)}}. \tag{6.16}$$

Also, the $[D_k]_\sharp \sigma_k$ converge weakly to $D_\sharp \sigma$. In fact for f continuous and compactly supported,

$$\begin{aligned} \lim_{k\to\infty} \int f\, d[D_k]_\sharp \sigma_k &= \lim_{k\to\infty} \int f(\theta_k^{-1} x) d\sigma_k(x) = \lim_{k\to\infty} \int f(\theta^{-1} x) d\sigma_k(x) \\ &= \int f(\theta^{-1} x) d\sigma(x) = \int f\, d[D]_\sharp \sigma \end{aligned} \tag{6.17}$$

(use the uniform continuity of f and local uniform bounds on the σ_k). By (6.1), σ is a flat measure; then $D_\sharp \sigma = \sigma$ and (6.17) shows that

$$\{a_k \mu^{r_{j(k)}}\} \text{ converges weakly to } \sigma. \tag{6.18}$$

If σ' is another nonzero element of $\mathrm{Tan}(\mu,0)$, we can find other sequences $\{j'(k)\}$ and $\{a'_k\}$, with $\lim_{k\to\infty} j'(k)=\infty$ (by the analogue of (6.15)), such that

$$\{a'_k\mu^{0,r_{j'(k)}}\} \text{ converges weakly to } \sigma'. \tag{6.19}$$

By Lemma 5.4, then (5.5), and then (6.13) and (6.2),

$$\begin{aligned}\mathbb{W}_\varphi(\sigma,\sigma') &\leq \liminf_{k\to\infty}\mathbb{W}_\varphi(a_k\mu^{0,r_{j(k)}},a'_k\mu^{0,r_{j'(k)}}) = \liminf_{k\to\infty}\mathbb{W}_\varphi(\mu^{0,r_{j(k)}},\mu^{0,r_{j'(k)}})\\ &\leq C\liminf_{k\to\infty}\int_0^{4\max(r_{j(k)},r_{j'(k)})}\alpha_{\mathscr{D}}(0,t)\frac{dt}{t}=0.\end{aligned} \tag{6.20}$$

Then $\sigma=\sigma'$, and this completes our proof of Lemma 6.1. □

7 The decomposition of the "good set" in Σ_1

In (1.28)-(1.29) we announced a decomposition of $\Sigma_0\cap(\Sigma_1\cup\Sigma_2)$ into pieces $\mathscr{S}_d$ $(0\leq d\leq n)$, which satisfy the property that for each $x\in\mathscr{S}_d$, $\mathrm{Tan}(\mu,x)\subset\mathscr{F}_d$. In this section we check that the pieces $\Sigma_1\cap\mathscr{S}_d$ satisfy the requirements of Theorem 1.2. The remaining sets $\Sigma_2\cap\mathscr{S}_d$ will be treated in Section 8.

We start with $d=0$. Set

$$\mathscr{S}_0=\big\{x\in\Sigma\,;\,\mathrm{Tan}(\mu,x)\subset\mathscr{F}_0\big\}. \tag{7.1}$$

We claim (as in the statement of Theorem 1.2) that $\mathscr{S}_0$ is the set of points where μ has an atom, and that every point of $\mathscr{S}_0$ is an isolated point of Σ.

Suppose that μ has an atom at x. Then since μ is doubling, x is an isolated point of Σ (Lemma 2.3 in [1]). We can check by hand that $\mathrm{Tan}(\mu,x)$ is the set $\mathscr{F}_0$ of multiples of the Dirac measure at the origin, and that $x\in\Sigma_0\cap\Sigma_1\cap\Sigma_2$ (because $\alpha_{\mathscr{D}}(x,r)=0$ for r small).

Conversely, suppose that $\mathrm{Tan}(\mu,x)\subset\mathscr{F}_0$, and let us check that x is an isolated point of Σ. Suppose instead that we can find a sequence $\{x_k\}$ in $\Sigma\setminus\{x\}$ that converges to x. Set $r_k=2|x-x_k|$. Since μ is doubling, there is a subsequence of $\{\mu_0^{x,r_k}\}$ which converges weakly to a measure σ. Since $\sigma\in\mathrm{Tan}(\mu,x)$, σ is a Dirac mass. Let ζ be smooth function such that $\mathbb{1}_{\mathbb{B}\setminus B(0,1/4)}\leq\zeta\leq\mathbb{1}_{(B(0,2)\setminus B(0,1/4)}$. Then $\int\zeta d\sigma=0$, so $\lim_{k\to\infty}\int\zeta\mu_0^{x,r_k}=0$. On the other hand, by (1.11) and (1.2)

$$\begin{aligned}\int\zeta d\mu_0^{x,r_k} &= \mu(B(x,r_k))^{-1}\int\zeta d\mu^{x,r_k}\\ &= \mu(B(x,r_k))^{-1}\int\zeta(r_k^{-1}(y-x))d\mu(y)\\ &\geq \mu(B(x,r_k))^{-1}\mu(B(x_k,r_k/4))\geq C_\delta^{-3}\end{aligned} \tag{7.2}$$

because $\zeta(r_k^{-1}(y-x))=1$ for $y\in B(x_k,r_k/4)$. This contradiction shows that if $x\in\mathscr{S}_0$, then x is an isolated point in Σ, and then μ has a Dirac mass at x. This gives the desired description of $\mathscr{S}_0$, the fact that $\mathscr{S}_0$ is at most countable is easy to see.

We may now concentrate on exponents $d\in[1,n]$. Set

$$\mathscr{S}'_d=\big\{x\in\Sigma_0\cap\Sigma_1\,;\,\mathrm{Tan}(\mu,x)\subset\mathscr{F}_d\big\}=\mathscr{S}_d\cap\Sigma_1, \tag{7.3}$$

where the last equality comes from (1.28). Together with $\mathscr{S}_0$, these sets are disjoint and cover $\Sigma_1 \cap \Sigma_0$ (by (1.29)), hence also μ-almost all of Σ_1 (by (1.26)). By Lemma 6.1, the only part of Theorem 1.2 concerning Σ_1 that remains to be proven is the fact that $\mathscr{S}'_d$ is rectifiable for $1 \leq d \leq n$, and more precisely

$$\mathscr{S}'_d \text{ can be covered by a countable family of } d\text{-dimensional Lipschitz graphs.} \tag{7.4}$$

(This is slightly more precise because we don not need to add a $\mathscr{H}^d$-negligible set.) This follows from the following lemma, which is essentially known, but which we prove for the reader's convenience.

Lemma 7.1. *Suppose that μ is a doubling measure, Σ is its support, $d \in \{1, ..., n\}$, and $E \subset \Sigma$ is such that for all $x \in E$, there is a vector space V_x of dimension d such that Tan$(x, \mu) = \big\{c\mathscr{H}^d|_{V_x}\,;\, c > 0\big\}$. Then E can be covered by a countable family of d-dimensional Lipschitz graphs.*

The fact that $E = \mathscr{S}'_d$ satisfies the assumption of the lemma comes from Lemma 6.1.

Proof. If $d = n$, $\mathbb{R}^n$ is a d-dimensional Lipschitz graph that covers E, thus we assume that $d < n$. We claim that for $x \in E$

$$x + V_x \text{ is a tangent plane to } \Sigma \text{ at } x. \tag{7.5}$$

If not there is a sequence $\{y_k\}$ in $\Sigma \setminus \{x\}$, that tends to x, and such that

$$\operatorname{dist}(y_k, x + V_x) \geq c|y_k - x| \tag{7.6}$$

for some $c > 0$. Set $r_k = 2|y_k - x|$, and replace $\{y_k\}$ with a subsequence for which the $\{\mu_0^{x,r_k}\}$ converges weakly to a measure $\sigma \in$ Tan(μ, x). Let ζ be a smooth compactly supported non-negative function such that $\zeta(0) = 0$ on V_x, but

$$\zeta(u) = 1 \ \text{ for } u \in \mathbb{B} \text{ such that } \operatorname{dist}(u, V_x) \geq c/2. \tag{7.7}$$

By assumption, σ is supported on V_x and so $\int \zeta d\sigma = 0$. Thus $\lim_{k\to\infty} \int \zeta \mu_0^{x,r_k} = 0$. On the other hand, (7.6) says that for $y \in B(y_k, cr_k/4)$,

$$\operatorname{dist}(r_k^{-1}(y - x), V_x) = r_k^{-1} \operatorname{dist}(y, x + V_x) \geq r_k^{-1}[\operatorname{dist}(y_k, x + V_x) - \frac{cr_k}{4}] \geq c/4,$$

hence $\zeta(r_k^{-1}(y - x)) = 1$ by (7.7), and (1.11) and (1.2) imply that

$$\begin{aligned} \int \zeta d\mu_0^{x,r_k} &= \mu(B(x, r_k))^{-1} \int \zeta d\mu^{x,r_k} \\ &= \mu(B(x, r_k))^{-1} \int \zeta(r_k^{-1}(y - x)) d\mu(y) \\ &\geq \mu(B(x, r_k))^{-1} \mu(B(x_k, cr_k/4)) \geq C^{-1}. \end{aligned} \tag{7.8}$$

This contradiction proves (7.5).

For $\varepsilon > 0$ small and $x \in E$, choose an integer $j = j(x) \geq 0$ such that

$$\operatorname{dist}(y, x + V_x) \leq \varepsilon |y - x| \text{ for } y \in \Sigma \cap B(x, 2^{-j(x)}). \tag{7.9}$$

On the Grassmann manifold $G(d,n)$ of vector spaces of dimension d in $\mathbb{R}^n$, let us for instance use the distance defined by $\operatorname{dist}(V,W) = ||\pi_V - \pi_W||$, where π_V and π_W denote the orthogonal projections on V and W. With this distance, $G(d,n)$ is compact. Choose a finite family $\mathscr{V}$ in $G(d,n)$ such that $\operatorname{dist}(V, \mathscr{V}) \leq \varepsilon$ for $V \in G(d,n)$. Set

$$E(V,j) = \big\{x \in E\,;\, j(x) = j \text{ and } \operatorname{dist}(V_x, V) \leq \varepsilon\big\} \tag{7.10}$$

for $V \in \mathscr{V}$ and $j \geq 0$. We now cover each $E(V,j)$ with a countable collection of d-dimensional Lipschitz graphs. We claim that for each ball B of radius 2^{-j-1},

$$E(V,j) \cap B \text{ is contained in a Lipschitz graph over } V. \tag{7.11}$$

Lemma 7.1 follows from this claim because the $E(V,j)$ cover E. To prove the claim, let $x, y \in E(V,j) \cap B$ be given. Observe that $|x - y| < 2^{-j}$ and $y \in E \subset \Sigma$, so (7.9) guarantees that $\operatorname{dist}(y, x + V_x) \leq \varepsilon |y - x|$. Then

$$|\pi_V(y) - \pi_V(x)| \leq |\pi_{V_x}(y) - \pi_{V_x}(x)| + ||\pi_V - \pi_{V_x}|||x - y| \leq 2\varepsilon |x - y|,$$

which yields (7.11). This completes our proof of Lemma 7.1.

□

8 The decomposition of the "good set" in Σ_2

Our goal in the section is to apply Theorem 1.5 in [1] to the set $\Sigma_0 \cap \Sigma_2$, to obtain the desired decomposition. For the reader's convenience we include the necessary background below.

Theorem 8.1 (Theorem 1.5, [1]). *Let μ be a doubling measure in $\mathbb{R}^n$, denote by Σ its support, and set*

$$\Sigma^0 = \big\{x \in \Sigma\,;\, \int_0^1 \alpha(x,r) \frac{dr}{r} < \infty\big\},$$

where

$$\alpha(x,r) = \min_{d=0,1,\ldots,n} \alpha_d(x,r),$$

and

$$\alpha_d(x,r) = \inf\big\{\mathbb{W}_1(\mu_0^{x,r}, \nu_V)\,;\, V \in A'(d,n)\big\}.$$

Here $A'(d,n)$ denotes the set of n dimensional affine spaces which intersect $B(0,1/2)$ and $\nu_V = c_V \mathscr{H}^d \llcorner V = c_V \mathbb{1}_V \mathscr{H}^d$, with $c_V = \mathscr{H}^d(V \cap \mathbb{B})^{-1}$. Then there are disjoint Borel sets $\Sigma^0(d) \subset \Sigma$, $0 \leq d \leq n$, such that

$$\Sigma^0 = \bigcup_{d=0}^{n} \Sigma^0(d),$$

with the following properties.

1. *First, $\Sigma^0(0)$ is the set of points of Σ where μ has an atom; it is at most countable and each of its point is an isolated point of Σ.*
2. *For $1 \leq d \leq n$ and $x \in \Sigma^0(d)$, the limit*

$$\theta_d(x) := \lim_{r \to 0} r^{-d} \mu(B(x,r))$$

exists, and $0 < \theta_d(x) < \infty$.

3. *For $1 \leq d \leq n$ and $x \in \Sigma^0(d)$, Σ has a tangent d-plane at x, W, and set $W^* = W - x$. Then Tan$(x,\mu) = \{c\mathscr{H}^d \llcorner W^* \,;\, c > 0\}$. In addition, the measures $\mu_0^{x,r}$ converge weakly to $\mathscr{H}^d \llcorner W^*$.*
4. *Further decompose $\Sigma^0(d)$, $1 \leq d \leq n$, into the sets*

$$\Sigma^0(d,k) = \left\{x \in \Sigma^0(d) \,;\, 2^k \leq \theta_d(x) < 2^{k+1}\right\},\ k \in \mathbb{Z};$$

then each $\Sigma^0(d,k)$ is a rectifiable set of dimension d, with $\mathscr{H}^d(\Sigma^0(d,k) \cap B(0,R)) < \infty$ for every $R > 0$, μ and $\mathscr{H}^d$ are mutually absolutely continuous on $\Sigma^0(d,k)$, and $\mu = \theta_d \mathscr{H}^d$ there.

We want to apply Theorem 8.1, so we need to show that for each $x \in \Sigma_0 \cap \Sigma_2$,

$$\int_0^1 \alpha(x,t) \frac{dt}{t} < \infty. \tag{8.1}$$

Let $x \in \Sigma_0 \cap \Sigma_2$ be given. By Theorem 1.3 every tangent measure $\sigma \in \text{Tan}(\mu, x)$ is flat. To estimate to the distance from $\mu_0^{x,r}$ to σ we proceed as in Section 6 except that we work with the whole group $\mathscr{G}$ rather than $\mathscr{D}$. We now follow that argument, without some of the details but we do emphasize the differences.

Without loss of generality, we assume that $x = 0$. We use the definition of $\alpha_{\mathscr{G}}$ and Chebyshev's inequality to associate to each $r \in (0, 1/4]$ a radius $r_+ \in [2r, 4r]$ such that

$$\alpha_{\mathscr{G}}(0, r_+) \leq 2 \int_{2r}^{4r} \alpha_{\mathscr{G}}(0,t) \frac{dt}{t} \tag{8.2}$$

(see (6.3)). By the definition (1.13), there exists $G \in \mathscr{G}(0, r_+)$ such that

$$\mathbb{W}_1(\mu_0^G, \mu_0^{0,r_+}) \leq 2\alpha_{\mathscr{G}}(0, r_+) \tag{8.3}$$

(see (6.4)). By (1.6) $G = \lambda R$ for some isometry, and since $G \in \mathscr{G}(x)$, (1.8) guarantee that $G(0) = 0$ and hence $R(0) = 0$. That is, R is a linear isometry. We still have that $\lambda^{-1} = \lambda(G)^{-1} \in [\lambda_1 r_+, \lambda_2 r_+]$, and if we set

$$r^* = \lambda^{-1} \in [\lambda_1 r_+, \lambda_2 r_+]; \tag{8.4}$$

as in (6.5), we have that

$$G(B(0, r^*)) = \mathbb{B} \tag{8.5}$$

and μ_0^G is the image of μ_0^{0,r^*} by a linear isometry. That is, $\mu_0^G = R_\sharp \mu_0^{0,r^*}$ and (8.3) only yields

$$\mathbb{W}_1(R_\sharp \mu_0^{0,r^*}, \mu_0^{0,r_+}) \leq 2\alpha_{\mathscr{G}}(0, r_+) \tag{8.6}$$

instead of (6.6). We still multiply the radii by r/r^*, set

$$\rho(r) = \frac{rr_+}{r^*} \in [\lambda_2^{-1} r, \lambda_1^{-1} r] \tag{8.7}$$

as in (6.10), and deduce from (8.6) that

$$\mathbb{W}_\varphi(R_\sharp \mu_0^{0,r}, \mu_0^{0,\rho(r)}) \leq C\alpha_{\mathscr{G}}(0, r_+), \tag{8.8}$$

using the same proof which involves Lemma 5.1 and Lemma 5.2 (the extra rotation does not affect the argument). Inequality (8.8) is the analogue of (6.11). Let us write this slightly differently. Set $R^r = R^{-1}$; then by (8.8)

$$\mathbb{W}_\varphi(\mu_0^{0,r}, R^r_\sharp \mu_0^{0,\rho(r)}) \leq C\alpha_{\mathscr{G}}(0, r_+), \tag{8.9}$$

since the $\mathbb{W}_\varphi$-distance is invariant under isometry.

Given $r_0 \leq 1/4$, we can construct a decreasing sequence $\{r_j\}$ as we did before, defined by $r_{j+1} = \rho(r_j)$. Let us keep track of the rotations: set $S^0 = I$ and $S^{j+1} = S^j R^{r_j}$. For $k \geq 0$ we want to estimate the numbers

$$\delta_k = \mathbb{W}_\varphi(\mu_0^{0,r_0}, S^{k+1}_\sharp \mu_0^{0,r_{k+1}}). \tag{8.10}$$

Let us check by induction that

$$\delta_k \leq C \sum_{0 \leq j \leq k} \alpha_{\mathscr{G}}(0, (r_j)_+). \tag{8.11}$$

When $k = 0$, this is (8.9) for r_0. If $k \geq 1$ and (8.11) holds for $k-1$, the triangle inequality (5.6) yields

$$\begin{aligned} \delta_k &\leq \delta_{k-1} + \mathbb{W}_\varphi(S^k_\sharp \mu_0^{0,r_k}, S^{k+1}_\sharp \mu_0^{0,r_{k+1}}) \\ &= \delta_{k-1} + \mathbb{W}_\varphi(S^k_\sharp \mu_0^{0,r_k}, [S^k R^{r_k}]_\sharp \mu_0^{0,r_{k+1}}) \\ &= \delta_{k-1} + \mathbb{W}_\varphi(\mu_0^{0,r_k}, R^{r_k}_\sharp \mu_0^{0,\rho(r_k)}) \leq \delta_{k-1} + C\alpha_{\mathscr{G}}(0, (r_k)_+) \end{aligned} \tag{8.12}$$

by definition of S^{k+1}, the invariance of $\mathbb{W}_\varphi$ under linear isometries, and (8.9). This proves (8.11). Then (8.2) and the same argument as in (6.12)-(6.13) yield

$$\begin{aligned} \delta_k &\leq C \sum_{0 \leq j \leq k} \alpha_{\mathscr{G}}(0, (r_j)_+) \leq C \sum_{0 \leq j \leq k} \int_{2r_j}^{4r_j} \alpha_{\mathscr{G}}(0, t) \frac{dt}{t} \\ &\leq C \int_0^{4r_0} \alpha_{\mathscr{G}}(0, t) \frac{dt}{t}. \end{aligned} \tag{8.13}$$

The final integral is finite because $0 \in \Sigma_2$ (see the definition (1.17)).

The measures μ_0^{0,r_k} are suitably normalized, so there is a subsequence which converges weakly to some measure σ (again see Lemma 2.1 in [1] for a little more detail). There is

also a further subsequence for which the S^k converge to an isometry S, and then the $S^k_\sharp \mu_0^{0,r_k}$ converge to $S_\sharp\sigma$ (proceed as for (6.17)). By Lemma 5.4, (8.10), and (8.13),

$$\begin{aligned}\mathbb{W}_\varphi(\mu_0^{0,r_0}, S_\sharp\sigma) &\leq \liminf_{k\to\infty} \mathbb{W}_\varphi(\mu_0^{0,r_0}, S^k_\sharp \mu_0^{0,r_k}) = \liminf_{k\to\infty} \delta_k \\ &\leq C\int_0^{4r_0} \alpha_{\mathscr{G}}(0,t)\frac{dt}{t}. \end{aligned} \tag{8.14}$$

We now use Lemma 5.3 to translate estimate (8.14) into an upper bound for the $\int_0^1 \alpha(x,r)\frac{dr}{r}$. For $t \in [1/4, 1/2]$, the measure μ_t that is defined by (5.27) with μ replaced by μ_0^{0,r_0} is just μ_0^{0,tr_0}. Since σ is a flat measure, so is $S_\sharp\sigma$. Hence the measure ν_t built from $\nu = S_\sharp\sigma$ as in (5.27) is also a flat measure supported on a d-plane V passing through the origin. We use ν_t to estimate $\alpha(x,r)$. By (5.28), the fact that μ is doubling, and (8.14), we have

$$\begin{aligned}\int_{1/4}^{1/2} \alpha(x,tr_0)dt &\leq \int_{1/4}^{1/2} \mathbb{W}_1(\mu_0^{0,tr_0}, \nu_t)dt = \int_{1/4}^{1/2} \mathbb{W}_1(\mu_t,\nu_t)dt \\ &\leq (8+||\varphi||_{lip})C_\delta^2 \mathbb{W}_\varphi(\mu_0^{0,r_0}, S_\sharp\sigma) \leq C\int_0^{4r_0} \alpha_{\mathscr{G}}(0,t)\frac{dt}{t}. \end{aligned} \tag{8.15}$$

Note that (8.15) holds for $r_0 \leq 1/4$.

We are now ready to prove that for $x \in \Sigma_0 \cap \Sigma_2$, $\int_0^1 \alpha(x,r)\frac{dr}{r} < \infty$. Recall we are assuming $x = 0$. By (8.15) and the definition (1.17) of Σ_2 we have

$$\begin{aligned}\int_0^{1/8} \alpha(x,s)\frac{ds}{s} &= \sum_{k\geq 2}\int_{2^{-k-2}}^{2^{-k-1}} \alpha(0,s)\frac{ds}{s} = \sum_{k\geq 2}\int_{1/4}^{1/2} \alpha(0,t2^{-k})\frac{dt}{t} \\ &\leq 4\sum_{k\geq 2}\int_{1/4}^{1/2} \alpha(0,t2^{-k})dt \\ &\leq 4C\sum_{k\geq 2}\int_0^{2^{-k+2}} \alpha_{\mathscr{G}}(0,t)\frac{dt}{t} \\ &= 4C\int_0^1 \alpha_{\mathscr{G}}(0,t)\Big\{\sum_{k\geq 2}\mathbb{1}_{\{k:2^{-k+2}>t\}}(k)\Big\}\frac{dt}{t} \\ &\leq C\int_0^1 \alpha_{\mathscr{G}}(0,t)\frac{\log(2/t)dt}{t} < \infty. \end{aligned} \tag{8.16}$$

A brutal estimate shows that $\alpha(x,r) \leq 2$ for $r > 0$, hence $\int_{1/8}^1 \alpha(x,s)\frac{ds}{s} < \infty$. Hence the hypothesis of Theorem 8.1 hold. We obtain a decomposition of the set $\Sigma_0 \cap \Sigma_2$ into subsets $\mathscr{S}''_d$, $0 \leq d \leq n$, which satisfy all the requirements for Theorem 1.2. In fact we get some additional information which we record here. First, for every point $x \in \mathscr{S}''_d$, $d \geq 1$, $\mathrm{Tan}(\mu,x)$ is the vector space of dimension 1 spanned by some flat measure of dimension d (see (1.27)) which implies

$$\mathscr{S}''_d = \big\{x \in \Sigma_0 \cap \Sigma_2\,;\, \mathrm{Tan}(\mu,x) \subset \mathscr{F}_d\big\} = \Sigma_2 \cap \mathscr{S}_d, \tag{8.17}$$

where $\mathscr{S}_d$ is as in (1.28). Moreover once we know that $\mathrm{Tan}(\mu,x)$ is the space of dimension 1 spanned by some flat measure $\sigma \in \mathscr{F}_d$, we have that Σ has a tangent d-plane at x whose

direction is given by the support of σ; see (7.5). Then $\mathscr{S}_d''$ is rectifiable, and even satisfies (7.4); see the proof above, and also[1]. This completes the proof of Theorem 1.2.

As mentioned in Section 1 we have additional control on the size of μ on Σ_2 and the behavior of μ on $\mathscr{S}_d''$. For $1 \leq d \leq n$ and every point $x \in \mathscr{S}_d'' = \Sigma_2 \cap \mathscr{S}_d$, the density of μ exists, that is

$$\theta_d(x) = \lim_{r \to 0} r^{-d} \mu(B(x,r)) \in (0, \infty) \tag{8.18}$$

(see [1]). Morever, we have the further decomposition of $\mathscr{S}_d''$ into sets $\mathscr{S}_d'' \cap \Sigma^0(d,k)$ where μ and $\mathscr{H}^d$ are mutually absolutely continuous, as in 4. in Theorem 8.1.

References

[1] J. Azzam, G. David, and T. Toro, *Wasserstein distance and rectifiability of doubling measures: Part I*, preprint, 2014.

[2] G. David, *Singular sets of minimizers for the Mumford-Shah functional,* Progress in Mathematics, 233. Birkhäuser Verlag, Basel, 2005. xiv+581 pp.

[3] G. David, C. Kenig and T. Toro, *Asympotically Optimally Doubling Measures and Reifenberg Flat Sets with Vanishing Constant*, Comm. Pure Appl. Math, **54**, 2001, 385-449.

[4] H. Federer, *Geometric measure theory*, Grundlehren der Mathematishen Wissenschaften 153, Springer Verlag 1969.

[5] C. Kenig and T. Toro, *Free Boundary Regularity for Harmonic Measures and Poisson Kernels*, Annals of Math. **150**, (1999), 369-454.

[6] P. Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge Studies in Advanced Mathematics, vol. 44, Cambridge University Press, Cambridge, 1995, Fractals and rectifiability. MR MR1333890 (96h:28006)

[7] D. Preiss, *Geometry of measures in* $\mathbf{R}^n$*: distribution, rectifiability, and densities*, Ann. of Math. (2) **125** (1987), no. 3, 537–643. MR 890162 (88d:28008)

[8] X. Tolsa, *Uniform rectifiability, Calderón-Zygmund operators with odd kernel, and quasiorthogonality*, Proc. Lond. Math. Soc. (3) **98** (2009), no. 2, 393–426. MR 2481953 (2010f:28007)

[9] _____, *Mass transport and uniform rectifiability*, Geom. Funct. Anal. **22** (2012), no. 2, 478–527. MR 2929071

[10] _____ *Rectifiable measures, square functions involving densities, and Cauchy transform*, preprint.

[11] C. Villani, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009, Old and new. MR 2459454 (2010f:49001)

Jonas Azzam: Departament de matemàtiques, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona) Email: jazzam@mat.uab.cat

Guy David: Université Paris-Sud, Laboratoire de Mathématiques, UMR 8658 Orsay, F-91405 CNRS, Orsay, F-91405. Email: guy.david@math.u-psud.fr

Tatiana Toro: University of Washington, Department of Mathematics, Seattle, WA 98195-4350. Email: toro@math.washington.edu