

BILIPSCHITZ EMBEDDINGS OF METRIC SPACES INTO EUCLIDEAN SPACES

S. SEMMES*

Abstract

When does a metric space admit a bilipschitz embedding into some finite-dimensional Euclidean space? There does not seem to be a simple answer to this question. Results of Assouad [A1], [A2], [A3] do provide a simple answer if one permits some small (“snowflake”) deformations of the metric, but unfortunately these deformations immediately disrupt some basic aspects of geometry and analysis, like rectifiability, differentiability, and curves of finite length. Here we discuss a (somewhat technical) criterion which permits more modest deformations, based on small powers of an A_1 weight. For many purposes this type of deformation is quite innocuous, as in standard results in harmonic analysis about A_p weights [J], [Ga], [St2]. In particular, it cooperates well with “uniform rectifiability” [DS2], [DS4].

1. Preliminaries

Let $(M, d(x, y))$ be a metric space. Thus M is a nonempty set, and $d(x, y)$ is a symmetric nonnegative function on $M \times M$ which vanishes exactly on the diagonal and satisfies the triangle inequality.

Question 1.1. *Under what conditions does $(M, d(x, y))$ admit a bilipschitz embedding into some $\mathbf{R}^m$?*

*Partially supported by the U.S. National Science Foundation. The author would also like to thank Kevin Rogovin for pointing out some oversights in a previous version of this paper. The author is grateful too to the referee for his or her comments and suggestions.

In other words, under what conditions does there exist a mapping $f : M \rightarrow \mathbf{R}^m$ (for some finite m) such that

$$(1.2) \quad C^{-1} d(x, y) \leq |f(x) - f(y)| \leq C d(x, y)$$

for some constant C and all $x, y \in M$? When this type of embedding exists, it roughly means that one might as well think of M as living in a finite-dimensional Euclidean space to begin with, which can be very useful.

There is a simple necessary condition for this to occur, which is that M be *doubling*. This means that there is a constant k so that every ball in M can be covered by k balls of half the radius. It is not hard to show that Euclidean spaces satisfy the doubling property, and that the doubling property is inherited by spaces which admit bilipschitz embeddings into other spaces that are doubling. A partial converse to this was given by Assouad [A1], [A2], [A3].

Theorem 1.3 (Assouad). *Let $(M, d(x, y))$ be a metric space which is doubling. Then for each $s \in (0, 1)$, the metric space $(M, d(x, y)^s)$ admits a bilipschitz embedding into some $\mathbf{R}^m$ (with m and the bilipschitz constant depending on s and the doubling constant for M).*

It is well-known (and not difficult to prove) that $(M, d(x, y)^s)$ is automatically a metric space when $(M, d(x, y))$ is, assuming that $0 < s < 1$. One can also check that $(M, d(x, y)^s)$ is doubling if and only if $(M, d(x, y))$ is doubling. The doubling condition seems to capture all of the “size” requirements needed in order to admit a bilipschitz embedding into some $\mathbf{R}^m$, and Assouad’s theorem helps to make that precise. (One could formulate Assouad’s theorem slightly differently, and say that $(M, d(x, y))$ is doubling if and only if $(M, d(x, y)^s)$ admits a bilipschitz embedding into some $\mathbf{R}^m$, where $0 < s < 1$.)

However, the doubling condition is *not* sufficient for the existence of a bilipschitz embedding. A basic family of examples, which was known to Assouad, is given by the Heisenberg groups with their natural (invariant) Carnot metrics. Like Euclidean spaces, the Heisenberg groups with their Carnot metrics admit transitive groups of isometries (given by left translations) and also one-parameter families of “dilations”, which have the effect of rescaling the metric. It turns out that Lipschitz mappings from Heisenberg groups into Euclidean spaces (or other Carnot groups) are “differentiable almost everywhere” in a suitable sense.

See [Pa]. This is analogous to more classical results for Lipschitz mappings on Euclidean spaces [Fe], [St1], and it permits one to show that bilipschitz mappings from the Heisenberg groups into Euclidean spaces do not exist. More precisely, if one did have a bilipschitz mapping from a Heisenberg group into a Euclidean space, then the differentials of the mapping (which exist almost everywhere) would be bilipschitz as well. However, the differentials are given almost everywhere by group homomorphisms, and this leads to a contradiction, because homomorphisms from Heisenberg groups into Euclidean spaces have to have nontrivial kernel, since the latter are commutative (as additive groups) while the former are not. The Heisenberg groups are also doubling, as one can see most easily by using the translation and dilation symmetries (to reduce the doubling property to the case of a single ball). See [Se2] for further discussion of this family of examples.

These examples seem to indicate that the existence of a bilipschitz embedding into some $\mathbf{R}^m$ is a rather delicate issue, and indeed no clear way exists at present to decide when a metric space which is doubling admits such an embedding. The result of Assouad provides a pretty good substitute, but the “snowflake transform” which replaces $(M, d(x, y))$ with $(M, d(x, y)^s)$, $0 < s < 1$, distorts the geometry of M too severely for some purposes. It changes the Hausdorff dimension, for instance, and one can check that $(M, d(x, y)^s)$ can never contain nonconstant rectifiable curves when $s < 1$ and $(M, d(x, y))$ is a metric space.

In analysis it is often natural to have not just a metric but also a *measure*. Given a metric space $(M, d(x, y))$, a nonnegative Borel measure μ on M is said to be *doubling* if there is a constant C such that

$$(1.4) \quad \mu(2B) \leq C\mu(B)$$

for all balls B in M , where $2B$ denotes the ball with the same center as B and twice the radius. (For the record, when we refer to a “ball” in M , we shall mean an open ball by default, although closed balls would work just as well in practice.) The existence of a (nonzero) doubling measure on M implies that $(M, d(x, y))$ is doubling as a metric space, as is well-known and not hard to show. If μ is doubling with respect to $d(x, y)$, then it is also doubling with respect to $d(x, y)^s$ for any s , and this is easy to verify too. Let us make the convention that “doubling measures” are always not identically zero. The doubling condition then implies that a doubling measure has a positive value on any nonempty open ball in the metric space.

Much of standard “order 0” harmonic analysis makes sense as soon as one has a metric space (or quasi-metric space) and a doubling measure. Of course L^p spaces make sense as soon as one has a measure, but with the extra structure of a metric and the doubling condition for μ one can make sense of Calderón-Zygmund operators, maximal operators, H^1 , BMO, A_p weights, etc., and with much the same results as usual. See [CW1], [CW2], [CM], [J], for instance. For this type of analysis the passage from $d(x, y)$ to $d(x, y)^s$ makes no real difference, and Assouad’s theorem implies that one might as well think of working with subsets of Euclidean spaces rather than abstract metric spaces. For analysis roughly like that of Hölder continuous functions on $\mathbf{R}^n$ of some order α strictly between 0 and 1, the “snowflake transform” $d(x, y) \mapsto d(x, y)^s$ is not so important either, except that the Hölder exponent α changes with s . However, for analysis which sees *integer* orders of smoothness, like Lipschitz functions or Sobolev spaces $W^{1,p}$, something like the snowflake transform is much more serious. Integer orders of smoothness on Euclidean spaces behave very differently from non-integer orders, and this reflects something quite substantial about the geometry of Euclidean spaces. Roughly speaking, Hölder conditions like

$$(1.5) \quad |g(x) - g(y)| \leq C|x - y|^\alpha$$

for real-valued functions g on $\mathbf{R}^n$ are much “flabbier” when $0 < \alpha < 1$ than for the $\alpha = 1$ case of Lipschitz functions. This is reflected in the differentiability almost everywhere of Lipschitz functions on Euclidean spaces, for instance. (See also [Se3] for more discussion of this theme.)

In this paper, our sympathies lie largely with contexts in which integer orders of smoothness, rectifiability of curves and surfaces, differentiability of functions, and so forth, are of concern. We would like to be able to put a metric space into some $\mathbf{R}^m$, and with only relatively modest distortions in geometry that do not cause too much trouble for considerations like these. We shall not really be able to do this outright, but we shall give a way to combine pieces of information at different scales and locations into a single and more coherent picture.

2. Definitions and the main result

Fix a metric space $(M, d(x, y))$, and a doubling (Borel) measure μ on M . This data will be used throughout this section. Note that the combination of $(M, d(x, y))$ and μ is sometimes called a “space of homogeneous type”, as in [CW2].

Definition 2.1 (The BPE condition). We say that $(M, d(x, y), \mu)$ satisfies the *BPE condition* (BPE for “big pieces of Euclidean spaces”) if there exist positive constants m , k , and θ , with m an integer, so that for each ball B in M there is measurable set $E \subseteq B$ and a mapping $h : E \rightarrow \mathbf{R}^m$ with the properties

$$(2.2) \quad \mu(E) \geq \theta \mu(B)$$

and

$$(2.3) \quad k^{-1} d(x, y) \leq |h(x) - h(y)| \leq k d(x, y) \quad \text{for all } x, y \in E$$

(i.e., h is k -bilipschitz on E).

With the BPE condition we bring aspects of measure into the problem of bilipschitz embeddings. Note that there are examples of spaces which satisfy the BPE condition but do not admit bilipschitz embeddings into any finite-dimensional Euclidean space. We shall say more about this in Section 4.

The following is a basic concept from harmonic analysis. (General references include [Ga], [J], [St2].)

Definition 2.4 (A_1 weights). Let $w(x)$ be a measurable real-valued function on M which is positive μ -almost everywhere. Then $w(x)$ is an A_1 weight if there is a constant C so that

$$(2.5) \quad \frac{1}{\mu(B)} \int_B w(x) d\mu(x) \leq C \operatorname{essinf}_B w$$

for all balls B in M .

Here “essinf” means the *essential infimum*, which is defined (as usual) by taking the largest possible value of the infimum that occurs when a set of measure 0 is removed from the set over which the essential infimum is being taken. (In other words, the essential infimum is not effected by small values of the function which are attained only on sets of measure 0.) Of course

$$(2.6) \quad \operatorname{essinf}_B w \leq \frac{1}{\mu(B)} \int_B w(x) d\mu(x)$$

automatically, and one can reformulate the A_1 condition as saying that the essential infimum of w over any ball is always comparable in size to the average of w over the same ball.

Observe that

$$(2.7) \quad \begin{array}{l} w(x) d\mu(x) \text{ is a doubling measure} \\ \text{on } M \text{ when } w(x) \text{ is an } A_1 \text{ weight.} \end{array}$$

This is not hard to see, and there is a stronger “approximate monotonicity” property that holds too. Namely, if B_1 and B_2 are balls which satisfy $B_1 \supseteq B_2$, then

$$(2.8) \quad \operatorname{essinf}_{B_2} w \geq \operatorname{essinf}_{B_1} w$$

automatically, and the A_1 condition yields

$$(2.9) \quad \frac{1}{\mu(B_2)} \int_{B_2} w(x) d\mu(x) \geq C^{-1} \frac{1}{\mu(B_1)} \int_{B_1} w(x) d\mu(x),$$

using also (2.6). This is much stronger than the doubling condition, because (2.9) applies no matter how small B_2 is compared to B_1 (i.e., with a uniform constant).

To illustrate the notion of an A_1 weight, let us consider functions of the form $|x|^\alpha$ on the real line. For this we use the standard metric and Lebesgue measure on $\mathbf{R}$, as our choice of background metric and measure. These functions are A_1 weights when $-1 < \alpha \leq 0$, as one can check through straightforward computation. The A_1 property fails when $\alpha > 0$, because of the vanishing at the origin, and also when $\alpha \leq -1$, since $|x|^\alpha$ is not even locally integrable at the origin in that case. Similarly, on $\mathbf{R}^n$, $|x|^\alpha$ is an A_1 weight exactly when $-n < \alpha \leq 0$. If x_1 denotes the first coordinate of a point $x \in \mathbf{R}^n$, then $|x_1|^\alpha$ is an A_1 weight on $\mathbf{R}^n$ if and only if $-1 < \alpha \leq 0$, and for practically the same reasons as when $n = 1$.

Notice that if $w(x)$ is just a positive constant, then w is an A_1 weight with constant equal to 1 in (2.5). More generally, if $w(x)$ is any A_1 weight, and c is a positive constant, then $cw(x)$ is an A_1 weight, and with the same A_1 -constant as $w(x)$. Thus, one should never take individual values of an A_1 weight too seriously, since constant multiplicative factors behave like free parameters.

Conversely, if $w(x)$ is an A_1 weight with constant equal to 1 in (2.5), then $w(x)$ must be equal to a fixed constant μ -almost everywhere on M . This is not hard to check, since one would have equality in (2.6).

If $w(x)$ is an A_1 weight, then $w(x)$ is locally bounded away from 0, by (2.5). However, it is possible for A_1 weights to tend to zero (at modest rates) at infinity, as in the preceding examples. The *sets* (of measure 0) on which an A_1 weight blows up can be practically arbitrary (as shown by a construction in [CR] —see also [J], [St2]), but the nature of the blowing up is regulated rather strongly by (2.5). In particular, one has approximate monotonicity as one shrinks down to specific locations, as in (2.9).

We shall use A_1 weights to make suitable deformations of metrics for bilipschitz embeddings into Euclidean spaces, as in the next assertion (which is the main result of this paper).

Proposition 2.10. *Let $(M, d(x, y))$ be a metric space, and let μ be a doubling measure on M . If $(M, d(x, y), \mu)$ satisfies the BPE condition, then there is an A_1 weight $w(v)$ on M with the following property.*

Given $\delta \in (0, 1]$, define $D_\delta(x, y)$ for $x, y \in M$ by

$$(2.11) \quad D_\delta(x, y) = d(x, y) \cdot \inf \left\{ \frac{1}{\mu(B)} \int_B w(v)^\delta d\mu(v) : B \text{ is a ball} \right. \\ \left. \text{in } M \text{ which contains } x \text{ and } y \text{ and satisfies } \text{diam } B < 3d(x, y) \right\}.$$

(If $x = y$, set $D_\delta(x, y) = 0$.) Then for each $\delta \in (0, 1]$ there is an integer ℓ and a mapping $f : M \rightarrow \mathbf{R}^\ell$ such that

$$(2.12) \quad C^{-1} D_\delta(x, y) \leq |f(x) - f(y)| \leq C D_\delta(x, y)$$

for some constant C and all $x, y \in M$. In other words, the deformation $(M, D_\delta(x, y))$ of $(M, d(x, y))$ admits a bilipschitz embedding into $\mathbf{R}^\ell$. These constants C and ℓ may be chosen so that they depend only on δ , the doubling constant for μ , and the constants associated to the BPE condition for $(M, d(x, y), \mu)$ from Definition 2.1.

One can think of (2.11) as saying that $D_\delta(x, y)$ is obtained by changing $d(x, y)$ at a given location and scale by an amount which is the average of $w(v)^\delta$ at the same approximate location and scale. (See also Remark 2.13 below.) This type of deformation is pretty mild, at least when δ is small, or the weight $w(v)$ is sufficiently moderate. We shall discuss this further in Section 3. When δ is not small, pathologies can occur for general weights. This will also be discussed in Section 3. This is not really a problem in the context of Proposition 2.10, though.

As we mentioned towards the end of Section 1, Proposition 2.10 is not really about producing bilipschitz embeddings directly, without any such knowledge in advance. Instead it provides a method for converting many separate pieces of partial information into something global. This can be quite convenient, e.g., for aspects of M which involve topology. A particular context in which Proposition 2.10 applies is that of “uniformly rectifiable” metric spaces (to which we shall return in Section 4). For these Proposition 2.10 shows that the setting of abstract metric spaces is not too different from that of subsets of Euclidean spaces. In particular, most of the existing literature about uniform rectifiability is formulated for subsets of Euclidean spaces, and Proposition 2.10 provides a way to reduce to that case. Similarly, if one is interested in working with differential forms, currents, and exterior differentiation on a uniformly rectifiable metric space, then one avenue would be to use Proposition 2.10 to reduce to the existing theory for subsets of Euclidean spaces (as in [Fe]).

Note that the BPE condition for $(M, d(x, y), \mu)$ can be recovered from the embedding of $(M, D_\delta(x, y))$ provided by Proposition 2.10 (and for any choice of $\delta > 0$), as we shall see in Section 3 (just after Lemma 3.40).

In harmonic analysis, one is accustomed to the idea of a weight as giving a deformation of the background measure, and to weighted-norm inequalities as saying that such deformations do not change basic properties of interest too much. See [Ga], [J], [St2], for instance. The present use of weights is similar in spirit. (See also [DS1], [Se3], and Section 3 below.)

The proof of Proposition 2.10 will be given in Section 6. In Sections 4 and 5 we shall discuss the embeddings in Proposition 2.10 in a couple of slightly more specialized situations.

Remark 2.13 (The infimum in (2.11)). Notice that there always exist balls B which satisfy the conditions in (2.11), i.e., which contain x and y and have diameter $< 3d(x, y)$. One can take $B(x, \frac{4}{3}d(x, y))$ and $B(y, \frac{4}{3}d(x, y))$, for instance. (The factors $\frac{4}{3}$ are included because “balls” are open, by default. If closed balls are used, this extra factor can be dropped, but this is not a serious matter in any event.) If B_1 and B_2 are any two balls which contain x and y and have diameter $< 3d(x, y)$, then the averages

$$(2.14) \quad \frac{1}{\mu(B_i)} \int_{B_i} w(v)^\delta d\mu(v), \quad i = 1, 2$$

are approximately the same, in the sense that the average over B_1 is bounded by a constant multiple of the average over B_2 , and vice-versa. This is an easy consequence of the doubling properties for μ and

$w(v)^\delta d\mu(v)$. The same would be true for any roughly similar class of balls B (living near x, y and having size around $d(x, y)$), for the same reason of the doubling conditions.

3. On the behavior of $D_\delta(x, y)$

Throughout this section, we assume that $M, d(x, y), \mu, \delta$, and $D_\delta(x, y)$ are as in Proposition 2.10 (except for the BPE hypothesis). We want to look at the behavior of $D_\delta(x, y)$ in comparison with $d(x, y)$ in some detail. This will not be used in the proof of Proposition 2.10, but it helps to make clear what the proposition really means. More precisely, in this section we shall look at the behavior of $D_\delta(x, y)$ when $w(v)$ is any A_1 weight, or is more general than that, rather than the special case of weights produced as in Proposition 2.10.

In particular, we shall be interested in the behavior of $D_\delta(x, y)$ when δ is small. For applying Proposition 2.10, note that one is free to take δ to be as small as one wants.

The material in this section is fairly standard in some circles (and not in others), and we include it largely for the sake of clarity and completeness. We begin with the following observation.

Lemma 3.1. *Suppose that $w(v)$ is an A_1 weight on M , with constant C_w in (2.5), and let $\delta \in (0, 1]$ be given. Then $w(v)^\delta$ is an A_1 weight with constant C_w^δ , and*

$$(3.2) \quad \begin{aligned} C_w^{-\delta} \left(\frac{1}{\mu(B)} \int_B w(v) d\mu(v) \right)^\delta &\leq \frac{1}{\mu(B)} \int_B w(v)^\delta d\mu(v) \\ &\leq \left(\frac{1}{\mu(B)} \int_B w(v) d\mu(v) \right)^\delta \end{aligned}$$

for all balls B in M .

Thus, if we set

$$(3.3) \quad D'_\delta(x, y) = d(x, y) \cdot \inf \left\{ \left(\frac{1}{\mu(B)} \int_B w(v) d\mu(v) \right)^\delta : B \text{ is a ball} \right. \\ \left. \text{in } M \text{ which contains } x \text{ and } y \text{ and satisfies } \text{diam } B < 3d(x, y) \right\}$$

(with the proviso that $D'_\delta(x, y) = 0$ when $x = y$), then

$$(3.4) \quad C_w^{-\delta} D'_\delta(x, y) \leq D_\delta(x, y) \leq D'_\delta(x, y),$$

by (3.2). Thus one might as well think in terms of $D'_\delta(x, y)$ instead of $D_\delta(x, y)$, if one wishes. This version is a bit more “stable” in some ways.

Lemma 3.1 is quite easy to prove. The second inequality in (3.2) follows directly from Jensen’s inequality, or one could use Hölder’s inequality instead. For the first part of (3.2), one uses the A_1 condition (2.5) and the trivial fact that

$$(3.5) \quad \left(\operatorname{ess\,inf}_B w \right)^\delta \leq \frac{1}{\mu(B)} \int_B w(v)^\delta d\mu(v)$$

for any $\delta > 0$.

To see that $w(v)^\delta$ is an A_1 weight with constant C_w^δ , one uses the second inequality in (3.2) together with the A_1 -condition (2.5) for $w(v)$ to get that

$$(3.6) \quad \frac{1}{\mu(B)} \int_B w(v)^\delta d\mu(v) \leq C_w^\delta \left(\operatorname{ess\,inf}_B w \right)^\delta,$$

which is exactly what we need. This proves Lemma 3.1.

One can think of the $D_\delta(x, y)$ ’s for different choices of $\delta \in (0, 1]$ as all being obtained from the same recipe, applied to the different A_1 weights $w(v)^\delta$. When $w(v)$ is an A_1 weight, the A_1 -constant for $w(v)^\delta$ tends to 1 as $\delta \rightarrow 0$, since it is $\leq C_w^\delta$, by Lemma 3.1. Remember from Section 2 that 1 is the smallest that an A_1 constant can be, and that an A_1 constant can be equal to 1 only when the weight in question is constant μ -almost everywhere. Constant weights correspond to constant multiples of $d(x, y)$ at the level of distance functions. In our case, where we use powers of a single weight to define $D_\delta(x, y)$, one can check directly that $D_\delta(x, y) \rightarrow d(x, y)$ as $\delta \rightarrow 0$, for any fixed x and y in M (and under much weaker conditions on $w(v)$).

In this section, we shall be interested in properties like this, but with more quantitative or uniform information. It will often be more convenient to work with $D'_\delta(x, y)$ instead of $D_\delta(x, y)$ (at least at first). For instance, it is a bit simpler to see that $D'_\delta(x, y) \rightarrow d(x, y)$ as $\delta \rightarrow 0$ for each $x, y \in M$ (than for $D_\delta(x, y)$), using just the fact that $a^\delta \rightarrow 1$ as $\delta \rightarrow 0$ when a is a positive number. Of course, some mild assumptions on μ and $w(v) d\mu(v)$ are needed here, such as positivity and finiteness of these measures on arbitrary open balls in M .

We begin with the observation in the next lemma. As indicated at the beginning of this section, we shall always assume here that μ is as in Proposition 2.10, i.e., a doubling measure on M . Let us also make the standing assumption that

$$(3.7) \quad w(v) d\mu(v) \text{ is a doubling measure on } M.$$

We shall be ready to assume more about $w(v)$, like the A_1 condition, but it will be nice to first look at what one can get from the doubling property.

Lemma 3.8. *If $\delta > 0$ is small enough, then there is a constant C_0 so that*

$$(3.9) \quad D'_\delta(x, y) \geq C_0^{-1} D'_\delta(x, z) \\ \text{whenever } x, y, z \in M \text{ and } d(x, y) \geq d(x, z).$$

More precisely, C_0 depends only on the doubling constants for μ and $w(v) d\mu(v)$, and (3.9) holds as soon as δ is less than or equal to a constant δ_0 that depends only on the doubling constant for μ .

First let us explain why Lemma 3.8 is true, and then say something about what it means and how it can be used.

To prove the lemma, let $x, y, z \in M$ be given, with $d(x, y) \geq d(x, z)$, as in (3.9). We may as well assume that $x \neq z$, and hence that $x \neq y$, since otherwise $D'_\delta(x, z) = 0$ automatically. Let $B_{x,y}$ be an open ball in M which satisfies the conditions in (3.3), i.e., $B_{x,y}$ contains x and y and has diameter less than $3d(x, y)$. Let $B_{x,z}$ be a ball in M which satisfies the analogous properties with respect to x and z . Thus

$$(3.10) \quad d(x, y) \left(\frac{1}{\mu(B_{x,y})} \int_{B_{x,y}} w(v) d\mu(v) \right)^\delta$$

and

$$(3.11) \quad d(x, z) \left(\frac{1}{\mu(B_{x,z})} \int_{B_{x,z}} w(v) d\mu(v) \right)^\delta$$

provide approximations to $D'_\delta(x, y)$ and $D'_\delta(x, z)$, as in (3.3). To establish (3.9), it suffices to show that

$$(3.12) \quad d(x, y) \left(\frac{1}{\mu(B_{x,y})} \int_{B_{x,y}} w(v) d\mu(v) \right)^\delta \\ \geq C_0^{-1} d(x, z) \left(\frac{1}{\mu(B_{x,z})} \int_{B_{x,z}} w(v) d\mu(v) \right)^\delta$$

for a suitable constant C_0 , and for all choices of $B_{x,y}$ and $B_{x,z}$ (that satisfy the conditions described above).

Roughly speaking, $B_{x,y}$ is “larger” than $B_{x,z}$. This may not quite be true in terms of set-theoretic inclusion, but we do have that

$$(3.13) \quad B_{x,z} \subseteq 7B_{x,y}.$$

To check this, let c and r denote the center and radius of $B_{x,y}$, respectively. Our assumptions on $B_{x,y}$ require that x and y lie in $B_{x,y}$, so that

$$(3.14) \quad d(x, y) \leq d(x, c) + d(c, y) < 2r.$$

On the other hand, $B_{x,z}$ contains x and has diameter less than $3d(x, z)$, by the conditions on $B_{x,z}$. Since $d(x, z) \leq d(x, y)$, again by assumption, we have that $\text{diam } B_{x,z} < 6r$. Because x lies in $B_{x,y} = B(c, r)$, we conclude that $B_{x,z} \subseteq B(c, 7r)$. This is the same as (3.13).

We now apply (3.13) to get that

$$(3.15) \quad \int_{B_{x,z}} w(v) d\mu(v) \leq \int_{7B_{x,y}} w(v) d\mu(v) \leq C \int_{B_{x,y}} w(v) d\mu(v).$$

The last step uses the doubling condition for $w(v) d\mu(v)$.

This estimate (3.15) goes in the “right direction” for (3.12). One also has factors of $\mu(B_{x,y})^{-1}$ and $\mu(B_{x,z})^{-1}$ in (3.12), and for these the inclusion (3.13) goes in the “wrong direction”. To get (3.12), we need to have upper bounds for $\mu(B_{x,y})$ in terms of $\mu(B_{x,z})$.

Set $\rho = d(x, y)/d(x, z)$. Thus $\rho \geq 1$, and one can verify that

$$(3.16) \quad B_{x,y} \subseteq 7\rho B_{x,z}.$$

This is analogous to (3.13), and we omit the details.

Using the doubling condition for μ and (3.16) we obtain that

$$(3.17) \quad \mu(B_{x,y}) \leq C_1 \rho^\alpha \mu(B_{x,z}),$$

where C_1 and α depend only on the doubling constant for μ . That is, one first iterates the doubling condition (1.4) to get $\mu(2^k B_{x,z}) \leq C^k \mu(B_{x,z})$ for any positive integer k , and then one applies this with k chosen so that $7\rho \leq 2^k$ and k is as small as possible.

We are now ready to derive (3.12). For the condition that δ be sufficiently small, we ask that $\delta \leq \alpha^{-1}$. This ensures that $\rho^{\delta\alpha} \leq \rho = d(x, y)/d(x, z)$, since $\rho \geq 1$. Once we have this, (3.12) follows from (3.15) and (3.17) by straightforward computation. Lemma 3.8 is then an easy consequence of this and the various definitions involved.

Note that one can take the constant C_0 in Lemma 3.8 to be close to 1 when δ is small. More precisely, the constant that comes out is the δ th power of a constant that depends only on the doubling conditions for μ and $w d\mu$.

It is not hard to show that Lemma 3.8 would not work in general if one did not require that δ be small. This is true even when $w(v)$ is an A_1 weight. In fact, for this purpose it would be better if $w(v)$ were roughly the opposite of an A_1 weight, like a negative power of an A_1 weight. In this case the averages $\mu(B)^{-1} \int_B w d\mu$ would never decrease too much (by more than a bounded factor) as B increases. This is the “opposite” of the approximate monotonicity inequality (2.9). If this were the case, then we could get (3.9) more directly, and without needing δ to be small.

Instead of requiring that δ be small, one could look at conditions on $w(v)$, to the effect that its averages do not decrease too rapidly as one moves away from a given location. In this regard, one can normally try to get more than the inequality (3.15) in the context of deriving (3.12). If $d(x, y)$ is very large compared to $d(x, z)$, then one would expect that the integral of $w d\mu$ over $B_{x, y}$ to be large compared to the integral of $w d\mu$ over $B_{x, z}$, which would make it easier to get (3.15). In general, however, the amount that one might gain from this can be quite modest (as a function of $d(x, y)/d(x, z)$, say), and it can easily be much less than what one loses from having $\mu(B_{x, y})$ be larger than $\mu(B_{x, z})$.

For the purposes of this paper, one might as well think in terms of taking δ to be small. One is free to do this in applying Proposition 2.10, and by doing so one avoids the need for other assumptions on $w(v)$.

As an application of Lemma 3.8, let us observe that $D'_\delta(x, y)$ defines a *quasimetric* on M when δ is small enough (and μ and $w(v) d\mu(v)$ are doubling measures). This means that $D'_\delta(x, y)$ satisfies the same conditions as a metric, except that the triangle inequality is weakened to the requirement that

$$(3.18) \quad D'_\delta(x, z) \leq K(D'_\delta(x, y) + D'_\delta(y, z))$$

for some constant K and all $x, y, z \in M$. Indeed, the conditions besides (3.18) —nonnegativity, symmetry, and the requirement that $D'_\delta(x, y) = 0$ if and only if $x = y$ — follow easily from the definition of $D'_\delta(x, y)$ (and the doubling conditions on μ and $w(v) d\mu(v)$). For (3.18), it is helpful to distinguish between two cases, where $d(x, y) \leq d(x, z)$ and $d(x, y) \geq d(x, z)$. The first case can be handled directly using the doubling conditions for μ and $w(v) d\mu(v)$ and the definition of $D'_\delta(x, y)$. (Notice that either $d(x, y)$ or $d(y, z)$ is greater than or equal to $d(x, z)/2$, by the triangle inequality for $d(u, v)$. This sets one up for using the

doubling condition in deriving (3.18).) If instead $d(x, y) \geq d(x, z)$, then (3.18), follows from (3.9), at least if δ is small enough for the hypotheses of Lemma 3.8.

If δ is not small, then (3.18) can fail, and this works in the same way as for Lemma 3.8. Let us point out that, as a consequence of (3.4), $D_\delta(x, y)$ (rather than $D'_\delta(x, y)$) is also a quasimetric when δ is small enough, if we assume that $w(v)$ is an A_1 weight on M .

In general, $D'_\delta(x, y)$ and $D_\delta(x, y)$ are not actual metrics, i.e., they do not satisfy the triangle inequality (without an extra constant factor, as in (3.18)). We shall say more about this soon. For the purposes of defining open sets, convergence of sequences, continuity of mappings, and so on, quasimetrics are practically as good as metrics, though. This leads to the question of whether $D'_\delta(x, y)$ and $D_\delta(x, y)$ are equivalent to $d(x, y)$, in terms of defining the same topology on M . To address this point, we shall use the following variant of Lemma 3.8.

Lemma 3.19. *If $\delta > 0$ is small enough, then there is a constant $\tilde{C}_0$ so that*

$$(3.20) \quad \frac{D'_\delta(x, y)}{D'_\delta(x, z)} \geq \tilde{C}_0^{-1} \left(\frac{d(x, y)}{d(x, z)} \right)^{1/2}$$

whenever $x, y, z \in M$ and $d(x, y) \geq d(x, z) > 0$.

Here $\tilde{C}_0$ depends only on the doubling constants for μ and $w(v) d\mu(v)$, and how small δ needs to be depends only on the doubling constant for μ .

This can be established in exactly the same manner as Lemma 3.8. The only difference is that instead of asking that $\delta \leq \alpha^{-1}$, where α is as in (3.17), one asks that $\delta \leq (2\alpha)^{-1}$. Once one has this, one can use (3.15) and (3.17) in essentially the same manner as before, and obtain (3.20). This gives Lemma 3.19. (One can also take $\tilde{C}_0$ to be close to 1 when δ is small enough, as in Lemma 3.8.)

Using Lemma 3.19, it is not hard to show that $D'_\delta(x, y)$ and $d(x, y)$ are equivalent topologically on M when δ is small enough (and μ and $w(v) d\mu(v)$ are doubling measures). For this it is helpful to rewrite (3.20) as

$$(3.21) \quad \frac{D'_\delta(x, z)}{D'_\delta(x, y)} \leq \tilde{C}_0 \left(\frac{d(x, z)}{d(x, y)} \right)^{1/2}$$

whenever $x, y, z \in M$ and $d(x, y) \geq d(x, z) > 0$.

To put it informally, if one fixes a point x in M and asks “which points $z \in M$ are close to x ?”, then (3.21) shows that points which are close to x in terms of $d(\cdot, \cdot)$ distances are also close in terms of $D'_\delta(\cdot, \cdot)$ distances. For this one would also fix $y \in M$ different from x , as a kind of basepoint, and observe that both $d(x, y)$ and $D'_\delta(x, y)$ are positive numbers. (We may as well assume here that M has more than one element, since otherwise there is nothing to prove anyway.) To go in the other direction, imagine that we fix $x \in M$, and an auxiliary point $z \in M$, $z \neq x$. Let y be another point in M , and assume that $D'_\delta(x, y)$ is very small, and small compared to $D'_\delta(x, z)$ in particular. If $D'_\delta(x, y)$ is sufficiently small compared to $D'_\delta(x, z)$, then we may apply (3.9) (or (3.20)) to conclude that $d(x, y) \leq d(x, z)$. That is, if $d(x, y)$ is not less than or equal to $d(x, z)$, then (3.9) implies that $D'_\delta(x, y)$ is greater than or equal to $D'_\delta(x, z)$ times a constant, whereas we are assuming that $D'_\delta(x, y)$ is much smaller than $D'_\delta(x, z)$. Thus we get that $d(x, y) \leq d(x, z)$. Once one has this, it is not hard to check that $d(x, y)$ tends to 0 as $D'_\delta(x, y)$ tends to 0. In other words, if there are y 's in M such that $d(x, y)$ does not become small, but $d(x, y) \leq d(x, z)$ (so that $d(x, y)$ remains bounded in particular), then one can verify that $D'_\delta(x, y)$ cannot get too small either. More precisely, none of the factors that occur in the definition (3.3) of $D'_\delta(x, y)$ can become too small in these circumstances (using also the doubling conditions for μ and $w(v) d\mu(v)$).

This shows that $D'_\delta(x, y)$ and $d(x, y)$ determine the same topology on M when μ and $w(v) d\mu(v)$ are doubling measures, and δ is small enough. As usual, this does not work in general without the requirement that δ be small. Similarly, $D_\delta(x, y)$ (rather than $D'_\delta(x, y)$) determines the same topology on M as $d(x, y)$ when δ is small and $w(v)$ is an A_1 weight. This follows from (3.4) and the corresponding statement for $D'_\delta(x, y)$.

Instead of the quasimetric condition, there is another (and stronger) way to say that $D_\delta(x, y)$ or $D'_\delta(x, y)$ is approximately a metric. Given $\delta \in (0, 1]$, let us consider the possibility that there is a metric $\rho_\delta(x, y)$ on M (satisfying the ordinary triangle inequality, without an extra factor as in (3.18)) such that

$$(3.22) \quad C^{-1} D_\delta(x, y) \leq \rho_\delta(x, y) \leq C D_\delta(x, y)$$

for some constant C and all $x, y \in M$. This condition implies that $D_\delta(x, y)$ is a quasimetric, but there are plenty of quasimetrics which are not approximately metrics in this sense. (A basic example is given by $|x - y|^\gamma$ on $\mathbf{R}^n$ when $\gamma > 1$.)

If $w(v)$ is as in Proposition 2.10, then it is true that there is a metric $\rho_\delta(x, y)$ on M so that (3.22) holds, and this works for every $\delta \in (0, 1]$.

Indeed, if $f : M \rightarrow \mathbf{R}^\ell$ is as in Proposition 2.10, then one can take

$$(3.23) \quad \rho_\delta(x, y) = |f(x) - f(y)|,$$

and this will have the required features. (Note that f depends on δ .) Specifically, (3.22) is the same as (2.12) in this case, while the triangle inequality for $\rho_\delta(x, y)$ follows easily from its particular form (and from the triangle inequality for the Euclidean metric on $\mathbf{R}^\ell$).

The existence of a metric $\rho_\delta(x, y)$ satisfying (3.22) is actually much easier than the existence of an embedding $f : M \rightarrow \mathbf{R}^\ell$ as in Proposition 2.10, and it happens more generally. To analyze this further, let us define a candidate for $\rho_\delta(x, y)$ by

$$(3.24) \quad \rho_\delta(x, y) = \inf \sum_{i=1}^k D_\delta(z_i, z_{i-1}),$$

where the infimum is taken over all finite chains $z_0, z_1, \dots, z_k$ of points in M which begin at x and end at y ($z_0 = x$, $z_k = y$). This is a kind of universal construction, which makes sense independently of the particular form for $D_\delta(x, y)$ that we have here.

It is easy to see that $\rho_\delta(x, y)$, defined in this manner, is nonnegative and symmetric, using the corresponding properties for $D_\delta(x, y)$. The triangle inequality also holds automatically for $\rho_\delta(x, y)$, as one can check using the fact that any chain from x to y can be combined with any chain from y to w to give a chain from x to w . Furthermore,

$$(3.25) \quad \rho_\delta(x, y) \leq D_\delta(x, y) \quad \text{for all } x, y \in M,$$

as one can see by using the one-step chain $z_0 = x$, $z_1 = y$ in (3.24).

The first inequality in (3.22) does not follow automatically, even if one knows that $D_\delta(x, y)$ is a quasimetric. The particular form of $D_\delta(x, y)$ does not especially help here either, nor is a doubling condition for $w(v)^\delta d\mu(v)$ (or for $w(v) d\mu(v)$) sufficient. As a basic counterexample, imagine that M is $\mathbf{R}^2$, with the Euclidean metric, and that $w(v) = |v_1|^\beta$, where v_1 denotes the first component of $v \in \mathbf{R}^2$, and β is any positive number. Then $w(v)^\delta dv$ is a doubling measure on $\mathbf{R}^2$ for any $\delta > 0$ (and indeed $w(v)^\delta$ is always an “ A_∞ weight”), but one can check that $\rho_\delta(x, y)$ as defined in (3.24) actually vanishes for all pairs of points $x, y \in \mathbf{R}^2$ whose first coordinates are 0. (A variant of this will be discussed in Section 5.)

For that matter, there is no metric $\rho_\delta(x, y)$ on $\mathbf{R}^2$ which will satisfy (3.22) in this case, and not just the candidate in (3.24). It is not hard to verify that if there is ever a metric $\rho_\delta(x, y)$ that satisfies (3.22), then the construction in (3.24) will work as well. This works in general, without regard to the specific form of $D_\delta(x, y)$ that we have here.

However, if $w(v)$ is actually an A_1 weight on M , then we are in much better shape. The A_1 condition prevents any kind of local vanishing of $w(v)$, as occurs in the examples on $\mathbf{R}^2$ mentioned above.

More precisely, if $w(v)$ is an A_1 weight on M , and if $\delta > 0$ is small enough (as in Lemma 3.8), then the first inequality in (3.22) will hold (for some constant C), and with $\rho_\delta(x, y)$ defined as in (3.24). This is very similar to an observation in [Se1] (Example (a) in Section 4 in [Se1]). In more concrete terms, one wants to show that there is a constant C so that

$$(3.26) \quad D_\delta(x, y) \leq C \sum_{i=1}^k D_\delta(z_i, z_{i-1})$$

for any pair of points $x, y \in M$ and any chain $z_0, z_1, \dots, z_k$ of points in M which begins at x and ends at y . To see this, it is helpful to consider the following two cases separately. First, if each point z_i in the chain $z_0, z_1, \dots, z_k$ lies in the ball $B(x, 10d(x, y))$, say, then (3.26) is not hard to derive from the definition of $D_\delta(x, y)$, using the triangle inequality for $d(\cdot, \cdot)$ and the A_1 property for $w(v)$. For this it is helpful to employ Lemma 3.29 below, which puts the A_1 condition for $w(v)$ in a more convenient form (at the level of $D_\delta(\cdot, \cdot)$, from which the triangle inequality for $d(\cdot, \cdot)$ is easier to apply). Second, if the z_i 's do not all lie in $B(x, 10d(x, y))$, then one can reduce to this case by looking at the first of the z_i 's which does not lie in $B(x, 10d(x, y))$. This is similar to Lemma 3.3 in [Se1], and is not difficult to manage anyway. It is for this case that we need the assumption that δ be small, in order to apply Lemma 3.8. Specifically, we need to know that $D_\delta(x, z) \geq C^{-1}D_\delta(x, y)$ when $z \notin B(x, 10d(x, y))$. In applying Lemma 3.8 here, we are switching the roles of y and z , and also replacing $D'_\delta(x, y)$ with $D_\delta(x, y)$. The latter is allowed because of (3.4), and since we are assuming that $w(v)$ is an A_1 weight.

Now let us look more at comparisons between $D_\delta(x, y)$ and $d(x, y)$. We begin with a “quasisymmetry” condition. If $w(v)$ is an A_1 weight on M , and $\delta > 0$ is small enough, then there is a function $\eta : [0, \infty) \rightarrow [0, \infty)$ such that $\lim_{t \rightarrow 0} \eta(t) = 0$ and

$$(3.27) \quad D_\delta(x, z) \leq \eta(t) D_\delta(x, y) \quad \text{whenever} \quad d(x, z) \leq t d(x, y),$$

for all $x, y, z \in M$ and $t > 0$. In other words, this says that the identity mapping on M is quasisymmetric as a map from $(M, d(x, y))$ to $(M, D_\delta(x, y))$ (in the sense of [TV]). This property implies that *relative* distances are approximately the same for $d(x, y)$ and $D_\delta(x, y)$. This is true uniformly in δ , i.e., with a single choice of η , but one can make “better” choices of η as δ gets smaller. We shall say more about this in a moment.

One can also describe (3.27) as saying that $d(x, y)$ and $D_\delta(x, y)$ determine roughly the same class of “balls” in M , even if they might assign very different diameters to sets in M .

To establish the quasisymmetry property (3.27), it is helpful to first use Lemma 3.1 to reduce to the analogous assertion for $D'_\delta(x, y)$ instead of $D_\delta(x, y)$. Once this reduction is made, the A_1 property for $w(v)$ is not needed, but only the doubling conditions for μ and $w(v) d\mu(v)$. The quasisymmetry condition then holds for all small δ with a single choice of $\eta(t)$ of the form $C \max(t^{1/2}, t^2)$, for instance. For $t \leq 1$, this is essentially the same as Lemma 3.19. When $t \geq 1$, one may as well restrict one’s attention to $x, y, z \in M$ which satisfy

$$(3.28) \quad d(x, y) \leq d(x, z) \leq t d(x, y)$$

(rather than just the second inequality). This is because of the $t = 1$ case already established. Under the condition (3.28), it is not hard to show that $D'_\delta(x, z) \leq C t^a D'_\delta(x, y)$ for suitable constants C and a , and with C and a close to 1 when δ is small. This is analogous to the computations made in the proof of Lemma 3.8, but with some minor changes. (In particular, the roles of expressions like $\int_B w(v) d\mu(v)$ and $\mu(B)$ are now the opposite of what they were before.)

As δ tends to 0, the quasisymmetry conditions for $D'_\delta(\cdot, \cdot)$ and $D_\delta(\cdot, \cdot)$ improve, and one can choose $\eta : [0, \infty) \rightarrow [0, \infty)$ so that it becomes uniformly close to the identity on bounded subsets of $[0, \infty)$. This is not hard to show, using the same kinds of computations as above. One can also choose $\eta(t)$ so that its growth at infinity and decay near 0 are only slightly worse than linear (in terms of powers of t).

These quasisymmetry properties provide fairly good senses in which the $D_\delta(x, y)$ ’s represent only mild deformations of the geometry of $d(x, y)$, especially as $\delta \rightarrow 0$. One can do better than this, however. We begin with the following lower bounds.

Lemma 3.29. *Fix points $x, y \in M$ and $\delta \in (0, 1]$, and put*

$$(3.30) \quad \tau_\delta(x, y) = \frac{D_\delta(x, y)}{d(x, y)}.$$

Let $z, u \in M$ be points such that $z, u \in B(x, 10d(x, y))$ (where the ball $B(x, 10d(x, y))$ is defined in terms of $d(\cdot, \cdot)$, as usual). If $w(v)$ is an A_1 weight on M , then

$$(3.31) \quad D_\delta(z, u) \geq (1 + C\delta)^{-1} \tau_\delta(x, y) d(z, u),$$

where C depends only on the A_1 constant for w and the doubling constant for μ (and not on the specific choices of x, y, z , or u in particular).

This is a kind of “ A_1 property” at the level of distance functions, rather than measures. Roughly speaking, it says that $D_\delta(z, u)$ is always approximately “larger” than $d(z, u)$, except that we allow for a scale factor $\tau(x, y)$ (which does not depend on z or u). Of course the choice of the constant 10 in $B(x, 10d(x, y))$ does not matter, and could be replaced by any other number, with larger choices leading to larger values of C in (3.31).

To prove Lemma 3.29, it is enough to show

$$(3.32) \quad D'_\delta(z, u) \geq (1 + \tilde{C}\delta)^{-1} \tau'_\delta(x, y) d(z, u)$$

in place of (3.31), where $\tilde{C}$ is a constant, and $\tau'_\delta(x, y)$ is defined in the same way as (3.30), but with $D_\delta(\cdot, \cdot)$ replaced by $D'_\delta(\cdot, \cdot)$. This follows from (3.4), which ensures that the substitution of $D'_\delta(\cdot, \cdot)$ for $D_\delta(\cdot, \cdot)$ leads at worst to additional factors of the form C_w^δ . Note that C_w^δ is bounded by the sum of 1 and a constant multiple of δ when δ lies in $(0, 1]$.

Next, we only need to worry about $\delta = 1$ in order to prove the lemma. This is because

$$(3.33) \quad \frac{D'_\delta(\xi, \eta)}{d(\xi, \eta)} = \left(\frac{D'_1(\xi, \eta)}{d(\xi, \eta)} \right)^\delta$$

for all $\xi, \eta \in M$, by the definition of $D'_\delta(\cdot, \cdot)$ in (3.3). In other words, the fact that we can get $(1 + C\delta)^{-1}$ on the right side of (3.31), instead of a constant without such good control on the dependence on δ , follows directly from a bound for $\delta = 1$ with any reasonable constant and the presence of the exponent δ on the right side of (3.33).

Thus it suffices to show that

$$(3.34) \quad D_1(z, u) \geq C^{-1} \tau_1(x, y) d(z, u)$$

under the conditions of the lemma. We have gone back to D_1 and τ_1 here, instead of D'_1 and τ'_1 , because $D'_1(z, u) = D_1(z, u)$ and $\tau'_1(x, y) = \tau_1(x, y)$, by definitions. (See (2.11) and (3.3).)

Fix x, y, z , and u as in the statement of Lemma 3.29. Set

$$(3.35) \quad I(x, y) = \inf \left\{ \frac{1}{\mu(B)} \int_B w(q) d\mu(q) : B \ni x, y \right. \\ \left. \text{and } \text{diam } B < 3d(x, y) \right\}.$$

More precisely, the infimum is taken over B 's which are balls in M (with respect to $d(\cdot, \cdot)$). Define $I(z, u)$ in the same manner, but with z and u instead of x and y . The desired estimate (3.34) is then the same as

$$(3.36) \quad I(z, u) \geq C^{-1} I(x, y),$$

as one can see by unwinding the definitions (of $D_1(\cdot, \cdot)$ and $\tau_1(x, y)$).

To prove (3.36), let B be a ball in M which contains z and u and has diameter $< 3d(z, u)$. By assumption, $z, u \in B(x, 10d(x, y))$, and so $d(z, u) \leq 20d(x, y)$ and $\text{diam } B < 60d(x, y)$. This implies in turn that

$$(3.37) \quad B \subseteq B(x, 70d(x, y)),$$

since z and u lie in both B and $B(x, 10d(x, y))$.

Using (3.37) and the A_1 property for $w(v)$ we obtain that

$$(3.38) \quad \frac{1}{\mu(B)} \int_B w(v) d\mu(v) \\ \geq C^{-1} \frac{1}{\mu(B(x, 70d(x, y)))} \int_{B(x, 70d(x, y))} w(v) d\mu(v),$$

as in (2.9). By taking the infimum over B we get that

$$(3.39) \quad I(z, u) \geq C^{-1} \frac{1}{\mu(B(x, 70d(x, y)))} \int_{B(x, 70d(x, y))} w(v) d\mu(v),$$

by the definition of $I(z, u)$. From here it is easy to derive (3.36) (with a modestly different constant C), using the doubling condition for μ . This completes the proof of Lemma 3.29.

The next lemma records estimates in the direction opposite to that of Lemma 3.29, i.e., with *upper* bounds for $D_\delta(z, u)$ instead of lower bounds. We no longer have uniform control, but only for “most” points, in the sense of μ -measure.

Lemma 3.40. *Assume that $w(v)$ is an A_1 weight on M . Fix $x, y \in M$ and $\delta \in (0, 1]$, and define $\tau_\delta(x, y)$ as before, in (3.30). For each $\lambda > 0$, there is a (relatively closed and hence measurable) subset F_λ of $B(x, 10d(x, y))$ with the properties described below. (Here the ball $B(x, 10d(x, y))$ is defined in terms of $d(\cdot, \cdot)$, as usual. Also, F_λ does not depend on δ .)*

Let $z, u \in M$ be points such that $z, u \in B(x, 10d(x, y))$ and at least one of z and u lies in F_λ . Then

$$(3.41) \quad D_\delta(z, u) \leq (C\lambda)^\delta \tau_\delta(x, y) d(z, u).$$

Also, F_λ is a subset of $B(x, 10d(x, y))$ of substantial size, in the sense that

$$(3.42) \quad \mu(B(x, 10d(x, y)) \setminus F_\lambda) \leq C\lambda^{-1} \mu(B(x, 10d(x, y))).$$

Here C depends only on the A_1 constant for w and the doubling constant for μ (and not on the specific choices of x, y, z , or u in particular).

Before we explain why this is true, let us say a few things about what it means.

If we take $\lambda = 2C$, with C as in Lemma 3.40, then (3.42) implies that F_λ contains at least half of the elements of $B(x, 10d(x, y))$, as measured by μ , and then (3.41) reduces to

$$(3.43) \quad D_\delta(z, u) \leq (2C^2)^\delta \tau_\delta(x, y) d(z, u).$$

This together with Lemma 3.29 shows that $D_\delta(z, u)$ is bounded from above and below by (bounded) constant multiples of $\tau_\delta(x, y) d(z, u)$ when z and u range among at least half of the points in $B(x, 10d(x, y))$. In particular, this shows that the existence of the embedding $f : M \rightarrow \mathbf{R}^\ell$ as in the conclusions of Proposition 2.10 implies that $(M, d(\cdot, \cdot), \mu)$ satisfies the BPE condition (Definition 2.1). In other words, the passage from the BPE condition to the conclusions of Proposition 2.10 does not lead to loss of information (modulo worse constants and embedding dimensions).

Notice also that Lemma 3.40 very much does not work (in general) if instead of $D_\delta(\cdot, \cdot)$ we used deformations of $d(\cdot, \cdot)$ of the form $d(\cdot, \cdot)^{1-\delta}$, as in Theorem 1.3. Everything else in this section so far—like Lemma 3.29 and the quasisymmetry properties of the deformations of $d(\cdot, \cdot)$ —would work just as well for these “snowflake” deformations $d(\cdot, \cdot)^{1-\delta}$. With Lemma 3.40 we make use of the measure theory in a significant way.

Remark 3.44. Lemma 3.40 would work just as well if $w(v)$ were an “ A_∞ weight” (see [Ga], [J], [St2]) instead of an A_1 weight. Lemma 3.29 does not work for A_∞ weights in general, but one can repair this by taking measure theory into account in the same manner as for Lemma 3.40, i.e., only asking for lower bounds when one of z or u lies in a set $\widetilde{F}_\lambda \subseteq B(x, 10d(x, y))$ of large size (in the sense of (3.42)). In this case, the lower bounds would depend on λ in the same manner as for the upper bounds in (3.41).

Note that for A_∞ weights, one does not in general have the approximation of $D_\delta(x, y)$ by actual metrics, as in (3.22). We shall return to this point in Section 5.

Let us now prove Lemma 3.40. We begin by using exactly the same reductions as in the proof of Lemma 3.29. That is, one can replace $D_\delta(\cdot, \cdot)$ with $D'_\delta(\cdot, \cdot)$ throughout, because of (3.4), and then we may reduce to the case where $\delta = 1$, because the powers of δ can be washed out, as in (3.33).

In the end, it suffices to find $F_\lambda \subseteq B(x, 10d(x, y))$ so that (3.42) holds, and so that

$$(3.45) \quad D_1(z, u) \leq C\lambda\tau_1(x, y)d(z, u)$$

when $z, u \in B(x, 10d(x, y))$ and at least one of z or u lies in F_λ . As in (3.36), we can reformulate (3.45) as

$$(3.46) \quad I(z, u) \leq C\lambda I(x, y),$$

where $I(\cdot, \cdot)$ is defined in (3.35). This follows from the definitions of $D_1(\cdot, \cdot)$ and $\tau_1(x, y)$.

Before we choose F_λ , let us record the following fact. Let z and w be arbitrary points in $B(x, 10d(x, y))$, and let B be a ball which contains z and w and has diameter which is $< 3d(z, u)$. For instance, B might simply be $B(z, \frac{4}{3}d(z, u))$. Just as in (3.37), we have that

$$(3.47) \quad B \subseteq B(x, 70d(x, y)).$$

Let f be the function which is equal to w on $B(x, 70d(x, y))$ and which vanishes on the complement of $B(x, 70d(x, y))$. Thus

$$(3.48) \quad \begin{aligned} \int_M f(p) d\mu(p) &= \int_{B(x, 70d(x, y))} w(p) d\mu(p) \\ &\leq C I(x, y) \mu(B(x, 10d(x, y))). \end{aligned}$$

The second step follows from the definition of $I(x, y)$ in (3.35) and the doubling conditions for μ and $w d\mu$. In particular, this constant C depends only on the doubling constants for μ and $w d\mu$.

Let f^* denote the (uncentered) Hardy-Littlewood maximal function of f , given by

$$(3.49) \quad f^*(p) = \sup_{B \ni p} \frac{1}{\mu(B)} \int_B f(q) d\mu(q),$$

where the supremum is taken over all balls B in M that contain p . Put

$$(3.50) \quad E_\lambda = \{p \in M : f^*(p) > \lambda I(x, y)\}.$$

It is not hard to check that E_λ is an *open* subset of M , since the balls B in (3.49) are open (by our standing convention that “ball” means “open ball”, by default). We also have that

$$(3.51) \quad \mu(E_\lambda) \leq C (\lambda I(x, y))^{-1} \int_M f(p) d\mu(p),$$

where C depends only on the doubling constant for μ . This is the usual “weak-type $(1, 1)$ inequality” for the maximal function, extended to the case of general doubling measures as in [CW1], [CW2], and proved through a Vitali-type covering lemma. Combining this with (3.48), we obtain that

$$(3.52) \quad \mu(E_\lambda) \leq C \lambda^{-1} \mu(B(x, 10 d(x, y))),$$

but with a modestly larger choice of C .

Now set $F_\lambda = B(x, 10 d(x, y)) \setminus E_\lambda$. Thus F_λ is a relatively closed subset of $B(x, 10 d(x, y))$, since E_λ is an open set, and (3.42) follows immediately from (3.52).

Let $z, u \in B(x, 10 d(x, y))$ be given, with one of z, u lying in F_λ . It remains to verify (3.46). For the sake of definiteness, let us assume that $z \in F_\lambda$, the other case being equivalent to this one. Fix any ball B in M such that B contains z and u and $\text{diam } B < 3 d(z, u)$. Thus

$$(3.53) \quad I(z, u) \leq \frac{1}{\mu(B)} \int_B w(q) d\mu(q)$$

by the definition (3.35) of $I(\cdot, \cdot)$. We can rewrite this as

$$(3.54) \quad I(z, u) \leq \frac{1}{\mu(B)} \int_B f(q) d\mu(q),$$

because of (3.47) and the definition of f . On the other hand,

$$(3.55) \quad \frac{1}{\mu(B)} \int_B f(q) d\mu(q) \leq f^*(z)$$

by the definition (3.49) of f^* and the fact that B contains z . Our assumption that $z \in F_\lambda$ yields

$$(3.56) \quad f^*(z) \leq \lambda I(x, y),$$

and thus we conclude that

$$(3.57) \quad I(z, u) \leq \lambda I(x, y).$$

This shows that (3.46) holds (with $C = 1$), and the proof of Lemma 3.40 is now complete.

To recapitulate once more, from Lemmas 3.29 and 3.40 we have that $D_\delta(\cdot, \cdot)$ and $d(\cdot, \cdot)$ are practically the same in any given ball in M , modulo bounded distortions, “bad” sets of small measure (as in (3.42)), and a positive scale factor.

4. Variations on the theme

In this section, we shall restrict our attention to metric spaces that are “Ahlfors regular”, in the sense of the next definition. For the record, $\overline{B}(x, t)$ denotes the closed ball with center x and radius t , in a given metric space.

Definition 4.1. A metric space $(M, d(x, y))$ is said to be *Ahlfors regular of dimension s* (where s is some positive number) if it is complete, and if there is a Borel measure μ on M such that

$$(4.2) \quad C^{-1} t^s \leq \mu(\overline{B}(x, t)) \leq C t^s$$

for some constant $C > 0$ and all $x \in M$, $0 < t \leq \text{diam } M$.

One sometimes also refers to the measure μ as being Ahlfors-regular of dimension s . Note that μ is automatically doubling in this case, with a constant that depends only on the dimension s and the constant C in (4.2).

If $(M, d(x, y))$ is Ahlfors-regular of dimension s , then (4.2) necessarily holds with μ replaced by s -dimensional Hausdorff measure H^s , and in fact μ and H^s are each bounded by a constant multiple of the other. This is not too hard to prove, but we shall not really need it here anyway.

Let us make the standing assumption for the time being that $(M, d(x, y))$ is Ahlfors-regular of dimension s , and that μ is as in Definition 4.1. In this case we have that

$$(4.3) \quad d(x, y) \approx (\mu(B(x, d(x, y))))^{\frac{1}{s}} \approx (\mu(B(y, d(x, y))))^{\frac{1}{s}},$$

where we write $A_1 \approx A_2$ to mean that A_1 and A_2 are bounded from above and below by constant multiples of each other. This is an immediate consequence of (4.2). If μ were merely doubling (i.e., without Ahlfors regularity), then expressions like $(\mu(B(x, d(x, y))))^{\frac{1}{s}}$ provide quasisymmetrically-equivalent ways to measure distance in M compared to the original distance $d(x, y)$. (This assertion uses the assumption that $(M, d(x, y))$ be Ahlfors regular in a mild way, to know that M is “uniformly perfect”. This last means that for any point $x \in M$ and any radius $r \leq \text{diam } M$ there is a point $y \in M$ with $d(x, y) \approx r$.) With Ahlfors-regularity the two measurements of distance are almost the same, as in (4.3).

Let $w(v)$ be an A_1 weight on M (with respect to $d(\cdot, \cdot)$ and μ), and let $D_\delta(x, y)$ be as defined in (2.11). The assumption of Ahlfors-regularity gives us another way to look at $D_\delta(x, y)$, as follows.

Lemma 4.4. *Let $w(v)$ be an A_1 weight on M , as above, and suppose that $\delta, s > 0$ satisfy $\delta, \delta \cdot s \leq 1$. Define $\tilde{D}_\delta(x, y)$ by*

$$(4.5) \quad \tilde{D}_\delta(x, y) = \inf \left\{ \left(\int_B w(v)^{\delta s} d\mu(v) \right)^{\frac{1}{s}} : B \text{ is a ball in } M \text{ which contains } x \text{ and } y \text{ and satisfies } \text{diam } B < 3d(x, y) \right\}.$$

Then

$$(4.6) \quad C^{-1} D_\delta(x, y) \leq \tilde{D}_\delta(x, y) \leq C D_\delta(x, y)$$

for some constant C and all $x, y \in M$.

In other words, one might as well think in terms of $\tilde{D}_\delta(x, y)$ instead of $D_\delta(x, y)$. As in Remark 2.13, one should not take the infimum in (4.5) at all seriously, as all competing balls B will give approximately the same answer (to within bounded factors). In the present situation one could also simply take the infimum over all balls B , since the integral of $w^{\delta s} d\mu$ over B is monotone with respect to set-theoretic inclusion. (This is a bit different from what happens in the definition (2.11) of $D_\delta(x, y)$.)

Lemma 4.4 is pretty straightforward, but we include a proof for the sake of completeness. Consider $\hat{D}_\delta(x, y)$ defined by

$$(4.7) \quad \hat{D}_\delta(x, y) = d(x, y) \cdot \inf \left\{ \left(\frac{1}{\mu(B)} \int_B w(v)^{\delta s} d\mu(v) \right)^{\frac{1}{s}} : B \text{ is a ball} \right. \\ \left. \text{in } M \text{ which contains } x \text{ and } y \text{ and satisfies } \text{diam } B < 3d(x, y) \right\}.$$

It is easy to see that $\tilde{D}_\delta(x, y)$ is bounded from above and below by constant multiples of $\hat{D}_\delta(x, y)$, using Ahlfors regularity to “cancel” the $d(x, y)$ and the $1/\mu(B)$ which occur in (4.7).

Thus it suffices to show that $\hat{D}_\delta(x, y)$ and $D_\delta(x, y)$ are each bounded by constant multiples of each other. Because of the definitions (2.11) and (4.7) of $D_\delta(x, y)$ and $\hat{D}_\delta(x, y)$, it is enough to show that

$$(4.8) \quad \frac{1}{\mu(B)} \int_B w(v)^\delta d\mu(v) \quad \text{and} \quad \left(\frac{1}{\mu(B)} \int_B w(v)^{\delta s} d\mu(v) \right)^{\frac{1}{s}}$$

are each bounded by constant multiples of the other, where B is any ball in M . The second average in (4.8) is always greater than or equal to the first one when $s \geq 1$, because of Jensen’s inequality or Hölder’s inequality, and the reverse is true when $s \leq 1$. To get the opposite estimates (in either case), one can use the A_1 property for $w(v)$. This is practically the same as (3.2) (which one could also simply invoke here). Specifically, both “averages” in (4.8) are greater than or equal to

$$(4.9) \quad \left(\text{essinf}_B w \right)^\delta$$

automatically, and both are less than or equal to

$$(4.10) \quad \left(\frac{1}{\mu(B)} \int_B w(v) d\mu(v) \right)^\delta,$$

by Jensen's or Holder's inequalities. This last relies on our hypothesis that δ and $\delta \cdot s$ each be less than or equal to 1. The A_1 property for $w(v)$ implies that (4.10) is bounded by a constant times (4.9), which implies the desired inequalities for the two averages in (4.8). This proves Lemma 4.4.

The restriction to δ 's with $\delta \cdot s \leq 1$ in Lemma 4.4 should not be taken too seriously. In the context of Proposition 2.10, it is the small δ 's which are preferred anyway, in that the corresponding $D_\delta(\cdot, \cdot)$'s are then "milder" as deformations of the original metric $d(\cdot, \cdot)$. One can also think of replacing $w(v)$ by a small power of itself, in which case the conclusions of Proposition 2.10 remain the same. (Similar points came up in Section 3.)

Part of the point of Lemma 4.4 is that it helps to bring the deformations of metrics that come up in Proposition 2.10 closer to standard themes in harmonic analysis, in which deformations of measures are normally given greater emphasis. Let us take this a bit further and look at expressions like $\tilde{D}_\delta(x, y)$ in a more general way. Given a doubling measure ν on M (with respect to $d(x, y)$), define $\tilde{D}_\nu(x, y)$ for $x, y \in M$ by

$$(4.11) \quad \tilde{D}_\nu(x, y) = \inf \left\{ \nu(B)^{\frac{1}{s}} : B \text{ is a ball} \right. \\ \left. \text{in } M \text{ which contains } x, y \text{ and satisfies } \text{diam } B < 3d(x, y) \right\}.$$

The doubling condition for ν ensures that $\tilde{D}_\nu(x, y)$ is a quasimetric (in the sense of (3.18)), and one can also show that the quasisymmetry property (3.27) is satisfied (with $\tilde{D}_\nu(x, y)$ in the role of $D_\delta(x, y)$). (As indicated just after (4.3), the latter uses the assumption that $(M, d(x, y))$ be Ahlfors regular to know that M is "uniformly perfect", and thus lacking isolated islands which might otherwise cause trouble for the quasisymmetry condition.)

A nice feature of $\tilde{D}_\nu$ is that one automatically has Ahlfors-regularity for it, and using ν for the measure (in place of μ , as before). In other words, the ν -measure of a $\tilde{D}_\nu$ -ball of radius r is bounded from above and below by constant multiples of r^s , at least if r is not greater than the $\tilde{D}_\nu$ -diameter of M . This is a fairly standard fact, which is reviewed in Section 4 of [Se3] (beginning around (4.12)).

From Lemma 4.4 and the preceding remarks we have that $(M, D_\delta(x, y))$ is Ahlfors-regular with the same dimension s , and with respect to the measure $w(v)^{s\delta} d\mu(v)$. (The fact that $D_\delta(x, y)$ may not be a metric, but only a quasimetric, does not really cause trouble here, even if it was not

included in the definition of Ahlfors regularity above.) In other words, the type of deformation of the metric that we make in Proposition 2.10 preserves Ahlfors regularity, and the dimension s in particular.

Let us now consider a slightly more specific class of Ahlfors-regular metric spaces.

Definition 4.12 (Uniform rectifiability). Let $(M, d(x, y))$ be a metric space, and let μ be a Borel measure on M . Assume that $(M, d(x, y))$ is Ahlfors-regular of dimension s , and with this choice of μ , as in Definition 4.1. Then we say that $(M, d(x, y))$ is *uniformly rectifiable* if it satisfies the same conditions as in Definition 2.1 (the definition of the BPE condition), except that the dimension m of the Euclidean space $\mathbf{R}^m$ into which embeddings are made is required to be equal to s .

This condition is a lot stronger than the BPE condition, because of the equality of dimensions. It implies that the structure of M must be very much like that of $\mathbf{R}^s$, while embeddings into Euclidean spaces of larger dimension would allow plenty of fractal behavior (like self-similar Cantor sets and snowflakes). Note that the Ahlfors-regularity dimension s has to be an integer here.

Because uniform rectifiability is stronger than the BPE condition, it means that we can automatically apply Proposition 2.10 to it, to get a bilipschitz embedding into some $\mathbf{R}^\ell$ under modest deformations of the metric. It is not too difficult to show that these deformations do not disturb uniform rectifiability, i.e., that $(M, D_\delta(x, y))$ is uniformly rectifiable if $(M, d(x, y))$ is. This follows from Lemmas 3.29 and 3.40, and it would work just as well for any A_∞ (let alone A_1) weight $w(v)$, as in Remark 3.44. (This also uses implicitly the Ahlfors regularity of $(M, D_\delta(x, y))$, as discussed above.)

Even for uniformly rectifiable spaces, it may not be that a bilipschitz embedding into some $\mathbf{R}^\ell$ exists. See [Se2] for counterexamples. Proposition 2.10 implies that one can get a bilipschitz embedding after making a small perturbation in the metric, perturbations which can be seen (thanks to Lemma 4.4) as resulting in a simple manner from deformations of Hausdorff measure by small powers of an A_1 weight.

This is a very convenient fact, and one of the motivations behind the present paper. It provides a general way to reduce questions about uniformly rectifiable metric spaces to the special case of subsets of Euclidean spaces (which is the setting of [DS2], [DS3], [DS4]). The distortion of the geometry through weights of the type that we have here does not cause much trouble, because so many basic properties are not disturbed by such deformations. This follows standard themes in harmonic analysis (as in [Ga], [J], [St2], [Se3]).

5. Another special case

Fix a positive integer n and a doubling measure ν on $\mathbf{R}^n$, and define $\tilde{D}_\nu(x, y)$ for $x, y \in \mathbf{R}^n$ as in (4.11), with $s = n$. If ν is simply Lebesgue measure on $\mathbf{R}^n$, then it is automatically doubling, and $\tilde{D}_\nu(x, y)$ is just a constant multiple of the usual Euclidean distance $|x - y|$. In general, $\tilde{D}_\nu(x, y)$ is automatically a quasimetric (when ν is doubling), as mentioned in the previous section, but it need not be a metric.

Definition 5.1 (Metric doubling measures). A doubling measure ν on $\mathbf{R}^n$ is said to be a *metric doubling measure* if there is a constant C and a metric $\rho_\nu(\cdot, \cdot)$ on $\mathbf{R}^n$ in the ordinary sense (i.e., satisfying the triangle inequality) such that

$$(5.2) \quad C^{-1} \rho_\nu(x, y) \leq \tilde{D}_\nu(x, y) \leq C \rho_\nu(x, y)$$

for all $x, y \in \mathbf{R}^n$.

This definition makes sense more generally than on Euclidean spaces, but let us restrict ourselves to that situation for simplicity.

When $n = 1$, we have that

$$(5.3) \quad \tilde{D}_\nu(x, y) = \nu([x, y])$$

when $x \leq y$ (otherwise $\tilde{D}_\nu(x, y) = \nu([y, x])$), and this satisfies the triangle inequality. Thus doubling measures on $\mathbf{R}$ are always metric doubling measures. (For that matter, $\tilde{D}_\nu(x, y)$ is automatically isometrically equivalent to the standard metric in that case, as one can check using (5.3).)

This is not true when $n \geq 2$. For instance, consider the measure ν defined by

$$(5.4) \quad d\nu(x) = |x_1|^a dx$$

on $\mathbf{R}^n$, where x_1 denotes the first coordinate of x and dx denotes n -dimensional Lebesgue measure on $\mathbf{R}^n$. If x, y are elements of $\mathbf{R}^n$ which satisfy $x_1 = y_1 = 0$, then

$$(5.5) \quad \tilde{D}_\nu(x, y) = c |x - y|^{1 + \frac{a}{n}},$$

where c is a positive constant that does not depend on x or y . This is not hard to check, directly from the definitions. Using this, one can verify that ν is not a metric doubling measure when $a > 0$ and $n \geq 2$.

Specifically, fix $x, y \in \mathbf{R}^n$ with $x_1 = y_1 = 0$ and $x \neq y$, let L be an arbitrarily large positive integer. Define z_j for $0 \leq j \leq L$ by

$$(5.6) \quad z_j = \left(1 - \frac{j}{M}\right)x + \frac{j}{L}y.$$

Thus $z_0 = x$, $z_M = y$, and

$$(5.7) \quad \sum_{j=1}^L \tilde{D}_\nu(z_{j-1}, z_j) = c L^{-\frac{a}{n}} |x - y|^{1+\frac{a}{n}}.$$

If $a > 0$, then the right-hand side of (5.7) goes to 0 as L tends to infinity. If ν were a metric doubling measure, however, the left side of (5.7) would have to be bounded away from 0, independently of L . This follows by making a comparison with a true metric $\rho_\nu(\cdot, \cdot)$ (as in Definition 5.1), and then using the triangle inequality for ρ_ν to bound the sum from below. (Note that it was necessary to have $n > 1$ in this argument in order to be able to choose x different from y .)

Metric doubling measures on $\mathbf{R}^n$ are always mutually absolutely continuous with Lebesgue measure when $n \geq 2$, and in fact the density is always an A_∞ weight in the sense of harmonic analysis (as in [Ga], [J], [St2], [Se3]). This was proved in [DS1] (in a slightly different formulation — see also [Se1]), essentially by observing that the well-known method of Gehring [Ge] carries over to this context. (In [DS1], [Se1], “metric doubling measures” were called “strong A_∞ weights”, motivated by this fact.)

If ν is a metric doubling measure on $\mathbf{R}^n$, then $(\mathbf{R}^n, D_\nu(x, y))$ is practically as good as a metric space, and it is also Ahlfors-regular of dimension n , as mentioned in Section 4. It is uniformly rectifiable as well: when $n = 1$, $(\mathbf{R}^n, D_\nu(x, y))$ is isometrically equivalent to $\mathbf{R}$ with the usual metric, as mentioned above, while for $n \geq 2$ this follows from the A_∞ property of the density, as in the extensions of Lemmas 3.29 and 3.40 to the case of A_∞ weights indicated in Remark 3.44.

One might hope that, when $n \geq 2$, $(\mathbf{R}^n, D_\nu(x, y))$ is always bilipschitz equivalent to $\mathbf{R}^n$ with the standard metric when ν is a metric doubling measure. This question was one of the motivations behind [DS1], and it is equivalent to asking whether every metric doubling measure ν on $\mathbf{R}^n$ arises as the Jacobian of a quasiconformal mapping on $\mathbf{R}^n$ times Lebesgue measure, to within a factor which is bounded and bounded away from 0. It is not hard to see that the metric doubling property is *necessary* for a doubling measure to arise from the Jacobian of a quasiconformal mapping in this way, using standard distortion theorems. The absolute continuity and A_∞ conditions are also necessary when $n \geq 2$, but they hold automatically, as mentioned above.

The answer to this question turns out to be “no” (in general), by examples in [Se2] (and straightforward extensions of them). More precisely, there are examples where $(\mathbf{R}^n, D_\nu(x, y))$ is bilipschitz equivalent to a subset of a larger-dimension Euclidean space, and not to $\mathbf{R}^n$ itself, and there are examples where it is not even bilipschitz equivalent to a subset of any $\mathbf{R}^\ell$.

Thus, even for the special case of spaces that arise from metric doubling measures on $\mathbf{R}^n$, bilipschitz embeddings into Euclidean spaces need not exist. This case was analyzed in some detail in [Se1], where sufficient conditions for the existence of a bilipschitz embedding were given. For instance, if a doubling measure ν is of the form $w(u) du$, where $w(u)$ is an A_1 weight on $\mathbf{R}^n$, then a bilipschitz embedding into some $\mathbf{R}^m$ always exists. (It is not known whether $(\mathbf{R}^n, D_\nu(x, y))$ should have to be bilipschitz equivalent to $\mathbf{R}^n$ with the standard metric in this case, however.) Roughly speaking, the difficulties come when the density $w(u)$ of ν is small, and this does not really happen when $w(u)$ is an A_1 weight. This is born out by the examples in [Se2], where the geometric complexity came in the way that the density $w(u)$ vanished on certain sets.

The main result in [Se1] says that if one has a metric doubling measure $d\nu(u) = w(u) du$, and if ν does not lie on the “edge” of the set of all metric doubling measures in a certain precise sense, then a bilipschitz embedding into some $\mathbf{R}^m$. As a special case of this, one can start with any metric doubling measure and make it slightly “larger”, by multiplying by an arbitrarily small (positive) power of a certain A_1 weight, to get a new metric doubling measure which lies in the range for which bilipschitz embeddings are guaranteed to exist.

This last assertion is essentially the same as what one gets from Proposition 2.10 in this setting. The advantage of Proposition 2.10 is that it makes sense for arbitrary metric spaces, and not just ones that arise from deformations of $\mathbf{R}^n$ by metric doubling measures.

6. Proof of Proposition 2.10

The argument will be similar to the one in [Se1].

Standing Assumptions 6.1. $(M, d(x, y))$ is a metric space, μ is a Borel measure on M which is doubling (1.4), and $(M, d(x, y), \mu)$ satisfies the BPE condition (Definition 2.1).

Of course these are the same as the hypotheses of Proposition 2.10.

Let us fix constants m , k , and θ as in Definition 2.1. Thus, inside of each ball B in M there is a measurable subset $E \subseteq B$ and a mapping $g : E \rightarrow \mathbf{R}^m$ such that $\mu(E) \geq \theta \mu(B)$ and g is bilipschitz with constant k . These assertions will be in force throughout the present section, and the various constants which come up below will (mostly) depend only on these parameters and the doubling constant for μ .

Remember that the existence of the doubling measure μ on M implies that M is doubling as a metric space (i.e., every ball in M can be covered by a bounded number of balls of half the radius). This was mentioned in Section 1. Also, the doubling constant for M as a metric space is controlled by the doubling constant for μ as a measure, as an easy consequence of the argument.

To prove Proposition 2.10, we want to combine local mappings from the BPE condition into larger ones. Let us recall the following simple fact, concerning extensions of Lipschitz functions.

Proposition 6.2. *Let A be a (nonempty) subset of M , and suppose that $f : A \rightarrow \mathbf{R}$ is Lipschitz with constant L . Then there exists an extension of f to a real-valued function $F : M \rightarrow \mathbf{R}$ which is also Lipschitz with constant L .*

This is well-known, and one of the standard proofs is to take

$$(6.3) \quad F(x) = \inf\{f(y) + L d(x, y) : y \in A\}.$$

It is not hard to check that this is Lipschitz with constant L (since $x \mapsto d(x, u)$ is Lipschitz with constant 1 for any $u \in M$) and that it is equal to f on A . (Note that here, for the Lipschitz condition for $x \mapsto d(x, u)$, it is important that $d(x, y)$ be an actual metric—satisfying the ordinary triangle inequality—rather than the weaker quasimetric condition (3.18).)

The next proposition gives a substitute for Whitney decompositions for our metric space M . (See [St1] for the usual Whitney decomposition. This type of extension is well known in some quarters, as in [CW1].)

Proposition 6.4 (Generalized Whitney decompositions). *Let Ω be a open subset of M , with $M \setminus \Omega \neq \emptyset$. Then there is a subset A of Ω with the following properties:*

- (i) *if we set $B_a = \overline{B}(a, 10^{-1} \text{dist}(a, M \setminus \Omega))$, then*

$$(6.5) \quad \Omega = \bigcup_{a \in A} B_a;$$

- (ii) *there is a constant C_0 (depending only on the doubling constant for M) so that for each $a \in A$ there are at most C_0 choices of $b \in A$ such that $5B_b$ intersects $5B_a$;*
- (iii) *if a_1, a_2 are distinct elements of A , then*

$$(6.6) \quad \frac{1}{3}B_{a_1} \cap \frac{1}{3}B_{a_2} = \emptyset.$$

Let us indicate how Proposition 6.4 can be proved, both for the sake of completeness, and for the details of its formulation. We begin with the following observation. Suppose that $a, b \in \Omega$ satisfy

$$(6.7) \quad 5B_a \cap 5B_b \neq \emptyset.$$

Then

$$(6.8) \quad d(a, b) \leq \frac{1}{2} \operatorname{dist}(a, M \setminus \Omega) + \frac{1}{2} \operatorname{dist}(b, M \setminus \Omega),$$

by the definition of B_a, B_b . For any $y, z \in M$ we have that

$$(6.9) \quad \operatorname{dist}(y, M \setminus \Omega) \leq \operatorname{dist}(z, M \setminus \Omega) + d(y, z),$$

because of the triangle inequality, and therefore

$$(6.10) \quad \begin{aligned} \operatorname{dist}(b, M \setminus \Omega) &\leq \operatorname{dist}(a, M \setminus \Omega) + d(a, b) \\ &\leq \frac{3}{2} \operatorname{dist}(a, M \setminus \Omega) + \frac{1}{2} \operatorname{dist}(b, M \setminus \Omega), \end{aligned}$$

using (6.8). This reduces to

$$(6.11) \quad \operatorname{dist}(b, M \setminus \Omega) \leq 3 \operatorname{dist}(a, M \setminus \Omega)$$

(by moving the $\operatorname{dist}(b, M \setminus \Omega)$ term from the right side of (6.10) to the left, and then simplifying the constant factors). Switching the roles of a and b we also have that

$$(6.12) \quad \operatorname{dist}(a, M \setminus \Omega) \leq 3 \operatorname{dist}(b, M \setminus \Omega).$$

In short, $\operatorname{dist}(a, M \setminus \Omega)$ and $\operatorname{dist}(b, M \setminus \Omega)$ are nearly the same, by (6.11) and (6.12).

Going back to (6.8), we have that

$$(6.13) \quad d(a, b) \leq \frac{1}{2}(1 + 3) \operatorname{dist}(a, M \setminus \Omega) \leq 2 \operatorname{dist}(a, M \setminus \Omega),$$

and similarly with b instead of a on the right-hand side.

Imagine now that we have chosen $A \subseteq \Omega$ so that

$$(6.14) \quad d(a_1, a_2) \geq \min_{i=1,2} \frac{1}{10} \operatorname{dist}(a_i, M \setminus \Omega) \quad \text{when } a_1, a_2 \in A, a_1 \neq a_2.$$

Let us verify that part (ii) of Proposition 6.4 holds automatically if (6.14) is true. Fix $a \in A$, and let A_a denote the set of $b \in A$ such that (6.7) holds. Thus we have (6.13) when $b \in A_a$, by the discussion above. Also,

$$(6.15) \quad d(b_1, b_2) \geq \frac{1}{30} \operatorname{dist}(a, M \setminus \Omega) \quad \text{when } b_1, b_2 \in A_a \text{ and } b_1 \neq b_2,$$

because of (6.14) and (6.12). These two conditions imply that the number of elements in A_b is bounded by a constant that depends only on the doubling constant for M , as required in (ii) of Proposition 6.4. Specifically, the doubling condition guarantees that

$$(6.16) \quad \overline{B}(a, 2 \operatorname{dist}(a, M \setminus \Omega))$$

can be covered by a bounded number of open balls of radius

$$(6.17) \quad \frac{1}{60} \operatorname{dist}(a, M \setminus \Omega).$$

Each $b \in A_a$ lies in (6.16), since it satisfies (6.13), but no two distinct b 's in A_a can lie in the smaller (open) balls of radius (6.17), because of the lower bound in (6.15). This shows that the number of b 's in A_a is bounded by the number of balls in our covering of (6.16), which does the job.

Now let us check that part (iii) of Proposition 6.4 holds automatically if A satisfies (6.14). Suppose to the contrary that we have distinct points $a_1, a_2 \in A$ such that

$$(6.18) \quad \frac{1}{3} B_{a_1} \cap \frac{1}{3} B_{a_2} \neq \emptyset.$$

This implies that

$$(6.19) \quad d(a_1, a_2) \leq \frac{1}{30} \operatorname{dist}(a_1, M \setminus \Omega) + \frac{1}{30} \operatorname{dist}(a_2, M \setminus \Omega),$$

by the definition of B_{a_1}, B_{a_2} . We may as well assume that

$$(6.20) \quad \operatorname{dist}(a_1, M \setminus \Omega) \leq \operatorname{dist}(a_2, M \setminus \Omega),$$

since we can always reverse the indices to ensure that this is true. Thus

$$(6.21) \quad d(a_1, a_2) \geq \frac{1}{10} \operatorname{dist}(a_1, M \setminus \Omega),$$

by (6.14). As in (6.9) we have that

$$(6.22) \quad \operatorname{dist}(a_2, M \setminus \Omega) \leq \operatorname{dist}(a_1, M \setminus \Omega) + d(a_1, a_2),$$

and we can put this back into (6.19) to obtain that

$$(6.23) \quad d(a_1, a_2) \leq \frac{2}{30} \operatorname{dist}(a_1, M \setminus \Omega) + \frac{1}{30} d(a_1, a_2).$$

This yields

$$(6.24) \quad d(a_1, a_2) \leq \frac{2}{29} \operatorname{dist}(a_1, M \setminus \Omega),$$

by moving the $d(a_1, a_2)$ term on the right side of (6.23) to the left side, and then simplifying the fractions. However, (6.24) is incompatible with (6.21). This shows that the intersection of balls in (6.18) was impossible, which is exactly what we wanted (to derive part (iii) of Proposition 6.4 from (6.14)).

To obtain part (i) of Proposition 6.4, it suffices to choose $A \subseteq \Omega$ so that (6.14) holds and A is *maximal* with respect to these properties. More precisely, to say that A is maximal means that if $z \in \Omega \setminus A$, then $A \cup \{z\}$ does not satisfy (6.14). This says exactly that there is a point $a \in A$ so that

$$(6.25) \quad \begin{aligned} d(z, a) &\leq \min \left(10^{-1} \operatorname{dist}(z, M \setminus \Omega), 10^{-1} \operatorname{dist}(a, M \setminus \Omega) \right) \\ &\leq 10^{-1} \operatorname{dist}(a, M \setminus \Omega) \end{aligned}$$

holds. Thus $z \in B_a$ for this choice of $a \in A$, and we conclude that part (i) of Proposition 6.4 holds when A is maximal.

In order to have an $A \subseteq \Omega$ which satisfies (6.14) and is maximal, one does not need anything too abstract like Zorn's lemma. Let $\{U_j\}_{j=1}^\infty$ be a sequence of subsets of Ω such that U_j is bounded, $\operatorname{dist}(U_j, M \setminus \Omega) > 0$, and $U_j \subseteq U_{j+1}$ for all j , and so that

$$(6.26) \quad \Omega = \bigcup_{j=1}^{\infty} U_j.$$

For instance, one can fix $p \in M$ and take

$$(6.27) \quad U_j = \{x \in B(p, j) : \text{dist}(x, M \setminus \Omega) \geq j^{-1}\}.$$

If A is any subset of Ω that satisfies (6.14), then the number of elements of $A \cap U_j$ is bounded by a constant that depends on U_j but not on A , because of the doubling property of M . This is like the argument above, for deriving part (ii) of Proposition 6.4 from (6.14). (That is, no two elements of $U_j \cap A$ can lie in a single ball of sufficiently small radius (depending on $\text{dist}(U_j, M \setminus \Omega)$), because of (6.14), and hence the number of elements in $U_j \cap A$ is controlled by the (finite) number of balls of this radius needed to cover U_j .) Thus one can first choose $A_1 \subseteq U_1$ so that it satisfies (6.14) and is maximal with respect to these properties (by choosing A_1 so that it contains as many elements as possible), then add points to get a subset A_2 of U_2 which satisfies (6.14) and is maximal, and so on. In the end, the set A which is the union of the A_j 's, $1 \leq j < \infty$, will satisfy (6.14) and be maximal in Ω . Indeed, if A were not maximal in Ω , so that there is some point $w \in \Omega \setminus A$ such that $A \cup \{w\}$ satisfies (6.14), then $A_j \cup \{w\}$ also satisfies (6.14) for all j , and this contradicts the maximality of A_j in U_j when j is large enough so that $w \in U_j$ (which happens eventually, by (6.26)).

This completes the proof of Proposition 6.4 (i.e., a maximal set $A \subseteq \Omega$ which satisfies (6.14) exists, and parts (i), (ii), and (iii) of Proposition 6.4 are true for any A with these features, as explained above).

To prove Proposition 2.10, we shall construct a sequence $\{\Omega_j\}_{j=0}^\infty$ of open sets in M such that the Ω_j 's are decreasing in j , and so that we understand fairly well how to make bilipschitz embeddings into Euclidean spaces on $M \setminus \Omega_j$. In other words, we shall operate by "layers", but where the scales involved in level j vary from place to place in M . (Specifically, the relevant "scales" at level j will be given essentially by the distance to $M \setminus \Omega_j$.) In the end, we shall have to combine all these layers in a slightly careful way, both for constructing an embedding of M into some $\mathbf{R}^\ell$ and for the choice of the A_1 weight $w(x)$ on M (as in Proposition 2.10). The latter will be chosen so that it increases geometrically on the Ω_j 's.

We define $\{\Omega_j\}_{j=0}^\infty$ recursively as follows. Fix a basepoint $p_0 \in M$, and set $\Omega_0 = M \setminus \{p_0\}$. We assume now that Ω_j has already been chosen for some $j \geq 0$, and we want to define Ω_{j+1} . We also assume that Ω_j is open, and that $M \setminus \Omega_j$ is nonempty. We shall choose Ω_{j+1} so that it is an open subset of Ω_j , which will permit the recursion to continue afterwards.

Let $A_j = A(\Omega_j)$ be the subset of Ω_j which is provided by Proposition 6.4. Given $a \in M$, put

$$(6.28) \quad B_{a,j} = \overline{B}(a, 10^{-1} \operatorname{dist}(a, M \setminus \Omega_j)),$$

as in Proposition 6.4. For each $a \in A_j$, let $E_{a,j}$ be a subset of $B_{a,j}$ such that

$$(6.29) \quad \mu(E_{a,j}) \geq \theta_1 \mu(B_{a,j})$$

and

$$(6.30) \quad \text{there is a } k\text{-bilipschitz mapping } g_{a,j} : E_{a,j} \rightarrow \mathbf{R}^m.$$

It will be convenient to require also that

$$(6.31) \quad E_{a,j} \text{ is closed, and } E_{a,j} \subseteq \frac{1}{4} B_{a,j}.$$

The existence of such a set $E_{a,j}$ follows from the BPE condition, as reviewed at the beginning of the section, except for a couple of minor changes. Namely, the constant θ_1 in (6.29) should be taken a bit smaller than the original parameter θ (with $\theta_1 \geq C^{-1}\theta$ for a suitable constant C), to accommodate (6.31) and the fact that $B_{a,j}$ is a closed ball (while the BPE condition was stated for open balls). Also, to get $E_{a,j}$ to be closed (as in (6.31)), one can simply take the closure of the (possibly non-closed) set provided by the BPE condition. The bilipschitz embedding $g_{a,j}$ into $\mathbf{R}^m$ automatically extends to the closure, by standard reasoning, and the bilipschitz constant is not changed by this extension.

Set

$$(6.32) \quad E_j = \bigcup_{a \in A_j} E_{a,j}$$

and

$$(6.33) \quad \Omega_{j+1} = \Omega_j \setminus E_j.$$

Thus $\Omega_{j+1} \subseteq \Omega_j$ automatically. Note that the complement of Ω_{j+1} in M is given by $E_j \cup (M \setminus \Omega_j)$.

If Ω_j happens to be empty, then A_j , E_j , and Ω_{j+1} are empty too. This is not very interesting, but we do not mind it.

Lemma 6.34. $E_j \cup (M \setminus \Omega_j)$ is a closed subset of M (and hence Ω_{j+1} is open in M).

To see this, let $\{z_s\}_s$ be any sequence of points in $E_j \cup (M \setminus \Omega_j)$ which converges to some point $z \in M$. We want to show that z is contained in $E_j \cup (M \setminus \Omega_j)$. If $z \in M \setminus \Omega_j$, then there is nothing to do, and so we assume that $z \in \Omega_j$. This means that z lies in $B_{a,j}$ for some $a \in A_j$, by part (i) of Proposition 6.4. Because $\lim_{s \rightarrow \infty} z_s = z$, we have that $z_s \in 2B_{a,j}$ for all but finitely many s 's. In particular, $z_s \in \Omega_j$ for all but finitely many s 's. Since the z_s 's lie in $E_j \cup (M \setminus \Omega_j)$, we have that for all but finitely many s 's there is an $b(s) \in A_j$ such that $z_s \in E_{b(s),j}$. On the other hand, part (ii) of Proposition 6.4 implies that there are only finitely many b 's in A_j such that $B_{b,j}$ intersects $2B_{a,j}$. Because $E_{b(s),j} \subseteq B_{b(s),j}$, we conclude that the $b(s)$'s represent only finitely many distinct elements of A_j . In particular, there is a single $b_0 \in A_j$ such that $b(s) = b_0$ for infinitely many s 's. Thus $z_s \in E_{b_0,j}$ for infinitely many values of s . This implies that $z \in E_{b_0,j}$, since each individual $E_{b,j}$ is closed, by construction, and since $\lim_{s \rightarrow \infty} z_s = z$. Thus z lies in E_j , and Lemma 6.34 follows.

Remember that the openness of Ω_j and nonemptiness of $M \setminus \Omega_j$ were the “induction hypotheses” from which we started. Lemma 6.34 shows that Ω_{j+1} satisfies the same conditions, so that the process may be repeated indefinitely. In the end we get a decreasing sequence $\{\Omega_j\}_{j=0}^\infty$ of open subsets of M , together with the auxiliary sets A_j , $B_{a,j}$, $E_{a,j}$, and E_j .

Eventually we shall want to combine the $g_{a,j}$'s associated to the $E_{a,j}$'s to obtain our embedding f of M into some $\mathbf{R}^\ell$ (as in Proposition 2.10), but first we want to give some quantitative estimates about the way that the Ω_j 's decrease with j . We begin with the following simple observation.

Lemma 6.35. *If $x \in B_{a,j}$, then*

$$(6.36) \quad \text{dist}(x, M \setminus \Omega_j) \leq \frac{11}{10} \text{dist}(a, M \setminus \Omega_j)$$

and

$$(6.37) \quad \text{dist}(a, M \setminus \Omega_j) \leq \frac{10}{9} \text{dist}(x, M \setminus \Omega_j).$$

Indeed,

$$(6.38) \quad \text{dist}(x, M \setminus \Omega_j) \leq \text{dist}(a, M \setminus \Omega_j) + d(x, a),$$

as in (6.9), and so

$$(6.39) \quad \text{dist}(x, M \setminus \Omega_j) \leq (1 + 10^{-1}) \text{dist}(a, M \setminus \Omega_j),$$

since $x \in B_{a,j}$ (and $B_{a,j}$ is as in (6.28)). This proves (6.36). Similarly,

$$(6.40) \quad \begin{aligned} \text{dist}(a, M \setminus \Omega_j) &\leq \text{dist}(x, M \setminus \Omega_j) + d(a, x) \\ &\leq \text{dist}(x, M \setminus \Omega_j) + 10^{-1} \text{dist}(a, M \setminus \Omega_j), \end{aligned}$$

again because $x \in B_{a,j}$. It is easy to derive (6.37) from (6.40), by subtracting the last term on the right and simplifying the constant factors. This proves Lemma 6.35.

Lemma 6.41. *For each $j \geq 0$ and $x \in M$ we have that*

$$(6.42) \quad \text{dist}(x, M \setminus \Omega_{j+1}) \leq \frac{1}{7} \text{dist}(x, M \setminus \Omega_j).$$

Thus Ω_{j+1} is fairly dense in Ω_j , at the scale of the largest balls in Ω_j .

To prove Lemma 6.41, let $x \in \Omega_j$ be given, and choose $a \in A_j$ so that $x \in B_{a,j}$ (as in part (i) of Proposition 6.4). From (6.28), (6.29), and (6.31) we have that

$$(6.43) \quad E_{a,j} \cap \frac{1}{4} B_{a,j} = E_{a,j} \cap B\left(a, \frac{1}{40} \text{dist}(a, M \setminus \Omega_j)\right) \neq \emptyset$$

(i.e., the left-hand side has positive μ -measure, and hence is nonempty). Thus there is a point $z \in E_{a,j}$ such that

$$(6.44) \quad \begin{aligned} d(x, z) &\leq d(x, a) + d(a, z) \\ &\leq \frac{1}{10} \text{dist}(a, M \setminus \Omega_j) + \frac{1}{40} \text{dist}(a, M \setminus \Omega_j) \\ &= \frac{1}{8} \text{dist}(a, M \setminus \Omega_j). \end{aligned}$$

Combining this with (6.37) yields

$$(6.45) \quad d(x, z) \leq \frac{10}{9 \cdot 8} \text{dist}(x, M \setminus \Omega_j) < \frac{1}{7} \text{dist}(x, M \setminus \Omega_j).$$

On the other hand, $z \in M \setminus \Omega_{j+1}$ automatically, by the definition (6.33), (6.32) of Ω_{j+1} . This gives (6.42) when $x \in \Omega_j$, and Lemma 6.41 follows (since the case where $x \in M \setminus \Omega_j$ is trivial).

The next observation controls the rate at which the Ω_j 's decrease in terms of measure.

Lemma 6.46. *There is a constant $\theta_2 > 0$ such that for every $x \in M$, $r > 0$, and $j \geq 0$ we have that*

$$(6.47) \quad \mu(B(x, r) \cap \Omega_{j+1}) \leq (1 - \theta_2) \mu(B(x, \rho) \cap \Omega_j),$$

where

$$(6.48) \quad \rho = r + \frac{2}{9} \sup_{y \in B(x, r)} \text{dist}(y, M \setminus \Omega_j).$$

The estimate (6.47) would be a bit simpler and more standard if ρ were equal to r —so that exponential decay of $\mu(B(x, r) \cap \Omega_j)$ in j would follow immediately by iteration— but the small correction to ρ in (6.48) will not cause much trouble. Also, the constant θ_2 in Lemma 6.46 can be taken to be a geometric constant times θ_1 , where θ_1 is as in (6.29).

To prove Lemma 6.46, let x , r , and j be given as above. Let $\mathcal{A}$ denote the set of $a \in A_j$ such that $B_{a,j}$ intersects $B(x, r)$. Thus

$$(6.49) \quad B(x, r) \cap \Omega_j \subseteq \bigcup_{a \in \mathcal{A}} B_{a,j},$$

by part (i) of Proposition 6.4 (with $\Omega = \Omega_j$).

Notice that

$$(6.50) \quad \mu \left(\left(\bigcup_{a \in \mathcal{A}} B_{a,j} \right) \cap E_j \right) \geq C^{-1} \sum_{a \in \mathcal{A}} \mu(B_{a,j} \cap E_j),$$

because of the bounded overlap property of the $B_{a,j}$'s provided by part (ii) of Proposition 6.4. On the other hand,

$$(6.51) \quad B_{a,j} \cap E_j \supseteq E_{a,j},$$

since $E_{a,j} \subseteq B_{a,j}$ by construction (as mentioned just after (6.28)), and since E_j is the union of the $E_{a,j}$'s, $a \in A_j$, by definition. (See (6.32).) This permits us to rewrite (6.50) as

$$(6.52) \quad \mu \left(\left(\bigcup_{a \in \mathcal{A}} B_{a,j} \right) \cap E_j \right) \geq C^{-1} \sum_{a \in \mathcal{A}} \mu(E_{j,a}).$$

Combining this with (6.29) we obtain that

$$(6.53) \quad \begin{aligned} \mu \left(\left(\bigcup_{a \in \mathcal{A}} B_{a,j} \right) \cap E_j \right) &\geq C^{-1} \sum_{a \in \mathcal{A}} \theta_1 \mu(B_{a,j}) \\ &\geq C^{-1} \theta_1 \mu \left(\bigcup_{a \in \mathcal{A}} B_{a,j} \right). \end{aligned}$$

Next we observe that

$$(6.54) \quad \left(\bigcup_{a \in \mathcal{A}} B_{a,j} \right) \cap E_j = \left(\bigcup_{a \in \mathcal{A}} B_{a,j} \right) \setminus \Omega_{j+1}.$$

This follows from the definition (6.33) of Ω_{j+1} and the fact that the $B_{a,j}$'s, $a \in A_j$, are contained in Ω_j , by construction. (Remember also that $\mathcal{A} \subseteq A_j$.) Thus (6.53) becomes

$$(6.55) \quad \mu \left(\left(\bigcup_{a \in \mathcal{A}} B_{a,j} \right) \setminus \Omega_{j+1} \right) \geq C^{-1} \theta_1 \mu \left(\bigcup_{a \in \mathcal{A}} B_{a,j} \right),$$

and this yields

$$(6.56) \quad \mu \left(\left(\bigcup_{a \in \mathcal{A}} B_{a,j} \right) \cap \Omega_{j+1} \right) \leq (1 - C^{-1} \theta_1) \mu \left(\bigcup_{a \in \mathcal{A}} B_{a,j} \right).$$

From (6.49) and (6.56) we obtain that

$$(6.57) \quad \mu(B(x, r) \cap \Omega_{j+1}) \leq (1 - C^{-1} \theta_1) \mu \left(\bigcup_{a \in \mathcal{A}} B_{a,j} \right).$$

To finish the proof of Lemma 6.46, it remains to show that

$$(6.58) \quad \bigcup_{a \in \mathcal{A}} B_{a,j} \subseteq B(x, \rho) \cap \Omega_j,$$

where ρ is as in (6.48). In other words, (6.47) will follow from (6.57) once we have (6.58).

The $B_{a,j}$'s are all contained in Ω_j by construction (when $a \in \mathcal{A} \subseteq A_j$), and so it suffices to show that

$$(6.59) \quad B_{a,j} \subseteq B(x, \rho) \quad \text{when } a \in \mathcal{A}$$

in order to establish (6.58). Let $a \in \mathcal{A}$ be given, and arbitrary. Thus $a \in A_j$ and $B_{a,j}$ intersects $B(x, r)$, by the definition of $\mathcal{A}$. Also, the radius of $B_{a,j}$ is $10^{-1} \text{dist}(a, M \setminus \Omega_j)$, as in (6.28). Let y be any element of $B_{a,j} \cap B(x, r)$. From Lemma 6.35 we have that

$$(6.60) \quad \text{dist}(a, M \setminus \Omega_j) \leq \frac{10}{9} \text{dist}(y, M \setminus \Omega_j),$$

and so the radius of $B_{a,j}$ is at most

$$(6.61) \quad \frac{1}{9} \operatorname{dist}(y, M \setminus \Omega_j).$$

If z is any element of $B_{a,j}$, then we have that

$$(6.62) \quad d(x, z) \leq d(x, y) + d(y, z) < r + \frac{2}{9} \operatorname{dist}(y, M \setminus \Omega_j).$$

That is, $d(x, y) < r$ because $y \in B(x, r)$, while

$$d(y, z) \leq \frac{2}{9} \operatorname{dist}(y, M \setminus \Omega_j)$$

because $y, z \in B_{a,j}$ and the radius of $B_{a,j}$ is bounded by (6.61). This implies that $d(x, z) < \rho$, where ρ is as in (6.48). Thus (6.59) holds, since this works for all $z \in B_{a,j}$, and Lemma 6.46 follows.

Next we want to define a weight $w(x)$ (as promised in Proposition 2.10) and establish some of its basic properties. We shall use the notation $\mathbf{1}_V(x)$ for the indicator (or characteristic) function of a set V , i.e., $\mathbf{1}_V(x) = 1$ when $x \in V$ and $\mathbf{1}_V(x) = 0$ when $x \notin V$.

Set

$$(6.63) \quad w(x) = 1 + \sum_{j=1}^{\infty} \lambda^j \mathbf{1}_{\Omega_j}(x),$$

where λ is a positive number. We shall always require that

$$(6.64) \quad 1 < \lambda \leq 2;$$

by choosing λ sufficiently close to 1, as in the next lemma, we can guarantee that $w(x)$ is an A_1 weight. We shall “officially” fix the choice of λ later on, but for now let us point out that different choices of λ amount to different positive powers of the same basic function, to within bounded factors. In other words, different choices of λ do not really lead to radically different $w(x)$ ’s.

One could easily adjust the definition of $w(x)$ so that different choices of λ literally lead to different powers of the same function. The expression in (6.63) was chosen somewhat for notational convenience (i.e., for writing down formulas more easily). Notice also that

$$(6.65) \quad w(x) = \sum_{j=0}^{\infty} \lambda^j \mathbf{1}_{\Omega_j}(x) \quad \text{when } x \neq p_0,$$

where p_0 is the basepoint chosen at the beginning of the construction (shortly before (6.28)).

Lemma 6.66. *If $\lambda > 1$ is sufficiently close to 1 so that*

$$(6.67) \quad \lambda(1 - \theta_2) < 1,$$

where θ_2 is as in Lemma 6.46, then $w(x)$ is an A_1 weight on M (with bounds that depend only on $\lambda(1 - \theta_2)$ and geometric constants for $(M, d(x, y), \mu)$).

To prove this, let $x \in M$ and $r > 0$ be given. We want to show that

$$(6.68) \quad \frac{1}{\mu(B(x, r))} \int_{B(x, r)} w(y) d\mu(y) \leq C \inf_{z \in B(x, r)} w(z)$$

for a suitable constant C , assuming that (6.67) holds.

Let k denote the largest nonnegative integer such that

$$(6.69) \quad B(x, r) \subseteq \Omega_k,$$

if such an integer exists. Note that if there is an integer k so that (6.69) holds, then there is a largest such integer, i.e., (6.69) cannot hold for arbitrarily large k , because of Lemma 6.41. If there is no nonnegative integer k for which (6.69) is satisfied, then it means that $p_0 \in B(x, r)$ (since $\Omega_0 = M \setminus \{p_0\}$ by definition, as indicated just above (6.28)). In this event we set $k = 0$.

In either situation we have that

$$(6.70) \quad B(x, r) \not\subseteq \Omega_{k+1}.$$

Using this and (6.69) we get that

$$(6.71) \quad \inf_{z \in B(x, r)} w(z) = \sum_{j=0}^k \lambda^j,$$

by the definition (6.63) of $w(x)$. (Note that this works in the exceptional case where $p_0 \in B(x, r)$ too.) We also have that

$$(6.72) \quad \sup_{y \in B(x, r)} \text{dist}(y, M \setminus \Omega_{k+1}) \leq 2r,$$

because of (6.70).

In order to establish (6.68), we shall iterate Lemma 6.46. Let m be an integer, $m \geq 2$, and let us estimate

$$(6.73) \quad \mu(B(x, r) \cap \Omega_{k+m}).$$

Define radii $r_m, r_{m-1}, \dots, r_1$ by the rules $r_m = r$,

$$(6.74) \quad r_{i-1} = r_i + \frac{2}{9} \sup_{x \in B(x, r_i)} \text{dist}(y, M \setminus \Omega_{k+i-1})$$

for $i = m, m-1, \dots, 2$. These rules ensure that

$$(6.75) \quad \mu(B(x, r_i) \cap \Omega_{k+i}) \leq (1 - \theta_2) \mu(B(x, r_{i-1}) \cap \Omega_{k+i-1})$$

when $i = m, m-1, \dots, 2$, by Lemma 6.46 (applied with $j = k+i-1$). By combining these inequalities we obtain that

$$(6.76) \quad \begin{aligned} \mu(B(x, r) \cap \Omega_{k+m}) &= \mu(B(x, r_m) \cap \Omega_{k+m}) \\ &\leq (1 - \theta_2)^{m-1} \mu(B(x, r_1) \cap \Omega_{k+1}). \end{aligned}$$

To make use of this we need to control the size of r_1 .

Fix $i \geq 2$ for the moment. Applying Lemma 6.41 repeatedly ($i-2$ times) we obtain that

$$(6.77) \quad \sup_{y \in B(x, r_i)} \text{dist}(y, M \setminus \Omega_{k+i-1}) \leq 7^{-(i-2)} \sup_{y \in B(x, r_i)} \text{dist}(y, M \setminus \Omega_{k+1}).$$

On the other hand, we can use (6.70) to get that $\text{dist}(x, M \setminus \Omega_{k+1}) < r$, and hence that

$$(6.78) \quad \sup_{y \in B(x, r_i)} \text{dist}(y, M \setminus \Omega_{k+1}) \leq r_i + r.$$

We can reduce this to

$$(6.79) \quad \sup_{y \in B(x, r_i)} \text{dist}(y, M \setminus \Omega_{k+1}) \leq 2r_i,$$

since $r \leq r_i$ for all i , $1 \leq i \leq m$, by the definition of the r_i 's. (That is, $r_m = r$, and the r_i 's increase as i decreases, by the recursive formula (6.74).) Combining (6.77) and (6.79) gives

$$(6.80) \quad \sup_{y \in B(x, r_i)} \text{dist}(y, M \setminus \Omega_{k+i-1}) \leq 2 \cdot 7^{-(i-2)} r_i.$$

Going back to (6.74) we obtain that

$$(6.81) \quad r_{i-1} \leq r_i + \frac{2}{9} (2 \cdot 7^{-(i-2)} r_i) = (1 + 4 \cdot 7^{-(i-1)}) r_i$$

when $i \geq 2$. Iterating this inequality we have that

$$(6.82) \quad r_1 \leq \left(\prod_{i=2}^m (1 + 4 \cdot 7^{-(i-1)}) \right) r,$$

since $r_m = r$ by definition. The infinite product

$$(6.83) \quad \prod_{i=2}^{\infty} (1 + 4 \cdot 7^{-(i-1)})$$

is finite (and not very large), by standard reasoning, and so we may conclude that

$$(6.84) \quad r_1 \leq P r,$$

where $P \geq 1$ is an absolute constant (namely, the product (6.83)). (One has $P \leq 4$, for instance.)

Substituting this into (6.76) yields

$$(6.85) \quad \mu(B(x, r) \cap \Omega_{k+m}) \leq (1 - \theta_2)^{m-1} \mu(B(x, P r) \cap \Omega_{k+1}).$$

We can throw away the Ω_{k+1} on the right-hand side to obtain

$$(6.86) \quad \mu(B(x, r) \cap \Omega_{k+m}) \leq (1 - \theta_2)^{m-1} \mu(B(x, P r)).$$

Although we have restricted ourselves to $m \geq 2$ in the preceding discussion (as mentioned just before (6.73)), (6.86) holds trivially for $m = 1$ as well.

We are now ready to prove (6.68). Define $w_1(y)$ and $w_2(y)$ by

$$(6.87) \quad w_1(y) = 1 + \sum_{j=1}^k \lambda^j \mathbf{1}_{\Omega_j}(y), \quad w_2(y) = \sum_{j=k+1}^{\infty} \lambda^j \mathbf{1}_{\Omega_j}(y).$$

Thus

$$(6.88) \quad w(y) = w_1(y) + w_2(y)$$

automatically, by the definition (6.63) of $w(y)$. We also have that

$$(6.89) \quad w_1(y) = \sum_{j=0}^k \lambda^j \quad \text{when } y \in B(x, r),$$

because of (6.69). More precisely, (6.69) holds except when $p_0 \in B(x, r)$, and when it holds we also have that $B(x, r) \subseteq \Omega_j$ for all $j \leq k$, since the Ω_j 's are decreasing, by construction. (See (6.33).) The inclusions $B(x, r) \subseteq \Omega_j$ for $j \leq k$ imply (6.89), by the definition of $w_1(y)$. In the exceptional case where $p_0 \in B(x, r)$, we have taken k to be 0 (as mentioned a few lines below (6.69)), and in this case (6.89) is still true, with the sum in the definition of $w_1(y)$ in (6.87) interpreted as being 0 when $k = 0$.

From (6.89) we obtain that

$$(6.90) \quad \frac{1}{\mu(B(x, r))} \int_{B(x, r)} w_1(y) d\mu(y) \leq \inf_{z \in B(x, r)} w(z).$$

This also uses

$$(6.91) \quad \sum_{j=0}^k \lambda^j \leq \inf_{z \in B(x, r)} w(z),$$

which one can get from (6.89). (In fact, we have equality in (6.90) and (6.91), as in (6.71).) In order to prove (6.68), it suffices to show that

$$(6.92) \quad \frac{1}{\mu(B(x, r))} \int_{B(x, r)} w_2(y) d\mu(y) \leq C \sum_{j=0}^k \lambda^j$$

for a suitable constant C .

From the definition (6.87) of $w_2(y)$ we have that

$$(6.93) \quad \int_{B(x, r)} w_2(y) d\mu(y) = \sum_{j=k+1}^{\infty} \lambda^j \mu(B(x, r) \cap \Omega_j).$$

Combining this with (6.86) yields

$$(6.94) \quad \int_{B(x, r)} w_2(y) d\mu(y) \leq \sum_{j=k+1}^{\infty} \lambda^j (1 - \theta_2)^{j-k-1} \mu(B(x, Pr)).$$

Let us rewrite this as

$$(6.95) \quad \int_{B(x, r)} w_2(y) d\mu(y) \leq \lambda^{k+1} \left(\sum_{i=0}^{\infty} \lambda^i (1 - \theta_2)^i \right) \mu(B(x, Pr)).$$

Our assumption (6.67) ensures that the sum

$$(6.96) \quad \sum_{i=0}^{\infty} \lambda^i (1 - \theta_2)^i$$

converges, and so we get that

$$(6.97) \quad \int_{B(x,r)} w_2(y) d\mu(y) \leq C_0 \lambda^{k+1} \mu(B(x, Pr)),$$

where C_0 depends on $\lambda(1 - \theta_2)$ (and, in particular, how close this quantity is to 1). Once we have (6.97), (6.92) follows easily, using also the doubling condition for μ .

To summarize, we now have (6.68), from (6.90), (6.91), (6.92), and (6.88). This completes the proof of Lemma 6.66.

Standing Assumptions 6.98. From now on we assume that λ is chosen (and fixed) in such a way that $1 < \lambda \leq 2$ and $\lambda(1 - \theta_2) < 1$ (as in (6.67)). We also take $w(x)$ to be defined as in (6.63), so that $w(x)$ is a (fixed) A_1 weight on M , as in Lemma 6.66.

The specific choice of λ can be made so that it depends only on the original constants for $(M, d(x, y), \mu)$, i.e., the doubling and BPE constants. Thus, in the computations that follow, we shall not be too concerned about the dependence of our bounds on λ , or on how close λ or $\lambda(1 - \theta_2)$ are to 1.

In the next lemma, we record separately some estimates that come from the proof of Lemma 6.66.

Lemma 6.99. *Let B be an (open) ball in M , and let k be*

$$(6.100) \quad \text{the largest nonnegative integer such that } B \subseteq \Omega_k.$$

(If no such k exists, then $p_0 \in B$, since $\Omega_0 = M \setminus \{p_0\}$ (as prescribed shortly before (6.28)), and we take $k = 0$.) For each $\delta \in (0, 1]$ we have that

$$(6.101) \quad C(\delta)^{-1} \lambda^{\delta k} \leq \frac{1}{\mu(B)} \int_B w(y)^\delta d\mu(y) \leq C(\delta) \lambda^{\delta k}.$$

Here $C(\delta)$ depends on λ , θ_2 , and geometric constants for $(M, d(u, v), \mu)$, in addition to δ . (In particular, $C(\delta)$ depends on how close $\lambda(1 - \theta_2)$ is to 1, as usual.)

When $\delta = 1$ this is practically the same as what we got in Lemma 6.66. Specifically, we write B as $B(x, r)$, as in the notation above, and then (6.101) follows easily from (6.87), (6.88), (6.89), and (6.97). One can also think in terms of going directly to the conclusions of the earlier argument, and applying the A_1 condition (6.68) together with (6.71). When $\delta < 1$, one can derive (6.101) from its counterpart for $\delta = 1$, using Lemma 3.1. Alternatively, one can also think of (6.101) when $\delta < 1$ as being a special case of the argument for Lemma 6.66, since

$$(6.102) \quad w(y)^\delta \approx 1 + \sum_{j=1}^{\infty} \lambda^{\delta j} \mathbf{1}_{\Omega_j}(y).$$

That is, each side of (6.102) is bounded by a constant multiple of the other, so that we can think of $w(y)^\delta$ as being like replacing λ with λ^δ in the preceding arguments. This change is compatible with the requirement (6.67), since $\lambda > 1$ implies that $1 < \lambda^\delta \leq \lambda$.

Thus we have Lemma 6.99. Next, we want to derive some estimates which will be useful for checking Lipschitz and bilipschitz conditions with respect to $D_\delta(u, v)$ on M , where $D_\delta(u, v)$ is as in (2.11) in the statement of Proposition 2.10. We begin with the following auxiliary definition.

Definition 6.103 (Admissible families of mappings). Let p be a nonnegative integer, and let $\{\psi_{a,j}\}_{a,j}$ be a family of mappings from M into $\mathbf{R}^d$, where j runs through all nonnegative integers and a runs through the set A_j for each individual j . (Remember that $A_j = A(\Omega_j)$ is the set associated to Ω_j in Proposition 6.4, as mentioned just above (6.28).) This family is said to be *admissible* if the following two conditions hold:

$$(6.104) \quad \text{supp } \psi_{a,j} \subseteq 5B_{a,j},$$

where $B_{a,j}$ is as in (6.28), and $\text{supp } \psi_{a,j}$ denotes the closure of the set of x 's such that $\psi_{a,j}(x) \neq 0$; and

$$(6.105) \quad \begin{array}{l} \text{the } \psi_{a,j} \text{'s are Lipschitz (with respect to } d(\cdot, \cdot) \text{)} \\ \text{on } M, \text{ and with uniformly bounded constant.} \end{array}$$

In general, when we refer to a function as being “Lipschitz on M ”, we mean Lipschitz with respect to our original distance $d(\cdot, \cdot)$ (rather than something like $D_\delta(\cdot, \cdot)$).

As a concrete example of an admissible family of mappings, consider $\{h_{a,j}\}_{a,j}$, with each $h_{a,j} : M \rightarrow \mathbf{R}$ defined by

$$(6.106) \quad h_{a,j}(x) = \max(2 \text{ radius } B_{a,j} - d(x, a), 0).$$

Here “radius $B_{a,j}$ ” means $10^{-1} \text{dist}(a, M \setminus \Omega_j)$, as in (6.28). (One should consider this as an assignment of a number to radius $B_{a,j}$, rather than a general definition. This is because $B_{a,j}$ might be representable as a ball in M with center a and a different radius, i.e., M might have some kind of gaps. Notice also that $M \setminus \Omega_j$ is nonempty for all j , because $\Omega_j \subseteq \Omega_0 = M \setminus \{p_0\}$, by construction.) These mappings $h_{a,j}$ are all Lipschitz with constant 1 (since functions of the form $x \mapsto \text{dist}(x, p)$ for any $p \in M$ are Lipschitz with constant 1, by the triangle inequality (as in (6.9)), and because this Lipschitz condition is not disturbed by the other operations in (6.106)). The localization condition (6.104) also holds automatically for each $h_{a,j}$, with the 5 in (6.104) replaced with a 2. Thus $\{h_{a,j}\}_{a,j}$ is indeed an admissible family. This type of family will be useful later on, for making some localizations.

In Lemma 6.142, we shall consider another admissible family of mappings, based on the $g_{a,j}$ ’s from (6.30). Before we get to that, we shall give some general estimates for admissible families of mappings.

Lemma 6.107. *Let $\{\psi_{a,j}\}_{a,j}$ be an admissible family of mappings into some $\mathbf{R}^p$. For each $\delta \in (0, 1]$ there is a constant $C(\delta)$ such that*

$$(6.108) \quad \sum_{j=0}^{\infty} \sum_{a \in A_j} \lambda^{\delta j} |\psi_{a,j}(x) - \psi_{a,j}(y)| \leq C(\delta) D_{\delta}(x, y)$$

for all $x, y \in M$. This constant $C(\delta)$ also depends on the Lipschitz bound for the $\psi_{a,j}$ ’s, as in (6.105), and the other usual constants, like λ and geometric constants for $(M, d(u, v), \mu)$ (but not on x or y).

Let $\{\psi_{a,j}\}_{a,j}$ and δ be given as in the statement of the lemma. We shall first derive some preliminary bounds, and then deal with Lemma 6.107 afterwards.

Sublemma 6.109. *For each nonnegative integer j_1 and each $x \in M$ we have that*

$$(6.110) \quad \sum_{j=j_1}^{\infty} \sum_{a \in A_j} \lambda^{\delta j} |\psi_{a,j}(x)| \leq C \lambda^{\delta j_1} \text{dist}(x, M \setminus \Omega_{j_1}).$$

Here C depends on the Lipschitz bound for the $\psi_{a,j}$ ’s and the other usual parameters, but not on x or j_1 . (Actually, C does not depend on δ in this case.)

To see this, notice that the property of bounded overlap in part (ii) of Proposition 6.4 and the condition (6.104) on the supports of the $\psi_{a,j}$'s ensure that

$$(6.111) \quad \psi_{a,j}(x) \neq 0 \text{ for at most a bounded number of } a\text{'s in } A_j$$

for each fixed x and j . Thus there are only a bounded number of terms in the inner sum on the left side of (6.110) that really matter.

Let us check that

$$(6.112) \quad |\psi_{a,j}(x)| \leq C \operatorname{dist}(x, M \setminus \Omega_j)$$

for a suitable constant C and all $x \in M$, $j \geq 0$, and $a \in A_j$. The support condition (6.104) and the Lipschitz bound (6.105) imply that $|\psi_{a,j}(x)|$ is bounded by a constant multiple of the radius of $B_{a,j}$. The latter is equal to $10^{-1} \operatorname{dist}(a, M \setminus \Omega_j)$, by (6.28). If $\psi_{a,j}(x) \neq 0$ (and these are the only x 's which we need to consider for (6.112)), then $x \in 5B_{a,j}$, by (6.104). This implies that

$$(6.113) \quad \operatorname{dist}(a, M \setminus \Omega_j) \leq 2 \operatorname{dist}(x, M \setminus \Omega_j),$$

by the same kind of argument as used to derive (6.37) in Lemma 6.35. Combining these pieces of information gives (6.112), as desired.

From (6.111) and (6.112) we obtain that

$$(6.114) \quad \sum_{j=j_1}^{\infty} \sum_{a \in A_j} \lambda^{\delta j} |\psi_{a,j}(x)| \leq \sum_{j=j_1}^{\infty} C' \lambda^{\delta j} \operatorname{dist}(x, M \setminus \Omega_j)$$

for a suitable constant C' . Because of Lemma 6.41, this reduces further to

$$(6.115) \quad \sum_{j=j_1}^{\infty} \sum_{a \in A_j} \lambda^{\delta j} |\psi_{a,j}(x)| \leq \sum_{j=j_1}^{\infty} C'' \lambda^{\delta j} 7^{j_1-j} \operatorname{dist}(x, M \setminus \Omega_{j_1}).$$

We can rewrite this as

$$(6.116) \quad \sum_{j=j_1}^{\infty} \sum_{a \in A_j} \lambda^{\delta j} |\psi_{a,j}(x)| \leq C'' \lambda^{\delta j_1} \left(\sum_{i=0}^{\infty} \lambda^{\delta i} 7^{-i} \right) \operatorname{dist}(x, M \setminus \Omega_{j_1}).$$

The infinite series on the right-hand side converges, because $\lambda \leq 2$ (as in Standing Assumptions 6.98) and $\delta \leq 1$. Thus the sum is really just a finite constant, and Sublemma 6.109 is an immediate consequence of (6.116).

Sublemma 6.117. *For each nonnegative integer j_1 and every pair of points $x, y \in M$ we have that*

$$(6.118) \quad \sum_{j=0}^{j_1} \sum_{a \in A_j} \lambda^{\delta j} |\psi_{a,j}(x) - \psi_{a,j}(y)| \leq C(\delta) \lambda^{\delta j_1} d(x, y).$$

This constant $C(\delta)$ depends on the usual parameters (in addition to δ), but it does not depend on j_1 , x , or y .

For each fixed j we have that

$$(6.119) \quad \sum_{a \in A_j} |\psi_{a,j}(x) - \psi_{a,j}(y)| \leq C d(x, y)$$

for some constant C and all $x, y \in M$. This uses (6.111) to say that only boundedly many terms in the sum are nonzero, and then the Lipschitz condition (6.105) to bound the individual terms by a constant multiple of $d(x, y)$. Thus

$$(6.120) \quad \sum_{j=0}^{j_1} \sum_{a \in A_j} \lambda^{\delta j} |\psi_{a,j}(x) - \psi_{a,j}(y)| \leq C \sum_{j=0}^{j_1} \lambda^{\delta j} d(x, y),$$

and Sublemma 6.117 follows by summing the geometric series. This uses the assumption that $\lambda > 1$, as in Standing Assumptions 6.98, and in particular the constant gets large when λ approaches 1.

Let us now use Sublemmas 6.109 and 6.117 to prove Lemma 6.107. Fix $x, y \in M$ and $\delta \in (0, 1]$, and set

$$(6.121) \quad B = B(x, \tfrac{4}{3} d(x, y)),$$

say. (We may as well assume that $x \neq y$, so that $d(x, y) > 0$, since otherwise Lemma 6.107 is trivial.) We have that

$$(6.122) \quad D_\delta(x, y) \approx d(x, y) \left(\frac{1}{\mu(B)} \int_B w(z)^\delta d\mu(z) \right),$$

i.e., each side of (6.122) is bounded by a constant multiple of the other. This follows from (2.11) and the doubling properties of μ and $w(z)^\delta d\mu(z)$, as in Remark 2.13.

Let k be as in Lemma 6.99, for this choice of B . From (6.101) in Lemma 6.99 we get that

$$(6.123) \quad C(\delta)^{-1} \lambda^{\delta k} d(x, y) \leq D_\delta(x, y) \leq C(\delta) \lambda^{\delta k} d(x, y)$$

for some constant $C(\delta)$ (which is not quite the same as the one in (6.101), in that it incorporates the one from (6.122) too).

To prove Lemma 6.107, it is enough to show that

$$(6.124) \quad \sum_{j=0}^{\infty} \sum_{a \in A_j} \lambda^{\delta j} |\psi_{a,j}(x) - \psi_{a,j}(y)| \leq C(\delta) \lambda^{\delta k} d(x, y)$$

for some constant $C(\delta)$, i.e., (6.124) in place of (6.108), because of (6.123). Let S denote the left side of (6.124), and let S_1 and S_2 denote the pieces of S that correspond to the sums over $j \leq k$ and $j > k$, respectively. Thus

$$(6.125) \quad S = S_1 + S_2,$$

by definitions. For S_1 we have that

$$(6.126) \quad S_1 \leq C(\delta) \lambda^{\delta k} d(x, y)$$

for some constant $C(\delta)$, because of Sublemma 6.117. For S_2 we get

$$(6.127) \quad S_2 \leq C \lambda^{\delta(k+1)} (\text{dist}(x, M \setminus \Omega_{k+1}) + \text{dist}(y, M \setminus \Omega_{k+1}))$$

using Sublemma 6.109. We also have that

$$(6.128) \quad B \not\subseteq \Omega_{k+1},$$

because of the choice of k (as in Lemma 6.99). Thus B intersects $M \setminus \Omega_{k+1}$, from which we may conclude that

$$(6.129) \quad \text{dist}(x, M \setminus \Omega_{k+1}) + \text{dist}(y, M \setminus \Omega_{k+1}) \leq 4 d(x, y),$$

using also the definition of B in (6.121). Combining (6.127) with (6.129) we obtain

$$(6.130) \quad S_2 \leq C' \lambda^{\delta(k+1)} d(x, y).$$

This together with (6.126) and (6.125) yields (6.124), which is what we needed to finish the proof of Lemma 6.107.

The following is a mild strengthening of Lemma 6.107 which will also be useful.

Lemma 6.131. *Let $\{\psi_{a,j}\}_{a,j}$ be an admissible family of mappings into some $\mathbf{R}^q$. Fix a positive integer p and points $x, y \in M$, $x \neq y$. Let B be the ball given in (6.121) (and depending on x, y), and let k be the integer associated to B as in Lemma 6.99. For each $\delta \in (0, 1]$ there is a constant $C(\delta)$ such that*

$$(6.132) \quad \sum_{j: |j-k| \geq p} \sum_{a \in A_j} \lambda^{\delta j} |\psi_{a,j}(x) - \psi_{a,j}(y)| \leq C(\delta) \lambda^{-\delta p} D_\delta(x, y).$$

Here the sum on the left is implicitly restricted to j 's which are nonnegative integers. The constant $C(\delta)$ does not depend on p, x, y , or k (but does depend on the Lipschitz bound for the $\psi_{a,j}$'s, and the other usual constants, like λ and geometric constants for $(M, d(u, v), \mu)$).

In other words, the sum in (6.108) becomes small (compared to $D_\delta(x, y)$) if we restrict ourselves to j 's which are far from k .

To prove this, we proceed in exactly the same manner as for Lemma 6.107. Specifically, let H denote the double-sum on the left side of (6.132), and let H_1, H_2 denote the portions of the sum which correspond to $j \leq k - p$ and to $j \geq k + p$, respectively, so that

$$(6.133) \quad H = H_1 + H_2$$

by definitions. Thus

$$(6.134) \quad H_1 \leq C(\delta) \lambda^{\delta(k-p)} d(x, y),$$

by Sublemma 6.117. (More precisely, we may apply Sublemma 6.117 when $k - p \geq 0$; if $k < p$, then H_1 is 0, and there is nothing to do.) From Sublemma 6.109 we have that

$$(6.135) \quad H_2 \leq C \lambda^{\delta(k+p)} (\text{dist}(x, M \setminus \Omega_{k+p}) + \text{dist}(y, M \setminus \Omega_{k+p})).$$

Using Lemma 6.41 we can reduce this to

$$(6.136) \quad H_2 \leq C \lambda^{\delta(k+p)} 7^{1-p} (\text{dist}(x, M \setminus \Omega_{k+1}) + \text{dist}(y, M \setminus \Omega_{k+1})).$$

On the other hand, (6.129) holds for the same reason as before (i.e., (6.128)), and so we may replace (6.136) with

$$(6.137) \quad H_2 \leq 4C \lambda^{\delta(k+p)} 7^{1-p} d(x, y).$$

Combining (6.134) and (6.137) we have that

$$(6.138) \quad H = H_1 + H_2 \leq C' (\lambda^{-\delta p} + \lambda^{\delta p} 7^{1-p}) \lambda^{\delta k} d(x, y).$$

We now apply (6.123) (which we may do, because the ball B and the integer k have been chosen here in exactly the same manner as before) to convert (6.138) into

$$(6.139) \quad H \leq C(\delta) (\lambda^{-\delta p} + \lambda^{\delta p} 7^{1-p}) D_\delta(x, y).$$

From here we get

$$(6.140) \quad H \leq C(\delta) \lambda^{-\delta p} D_\delta(x, y),$$

(with a slightly larger choice of $C(\delta)$). This uses the assumptions $\lambda \leq 2$ (as in Standing Assumptions 6.98) and $\delta \leq 1$ (from the statement of the lemma) to ensure that

$$(6.141) \quad \lambda^{\delta p} 7^{-p} \leq \lambda^{-\delta p},$$

which is exactly what we need to go from (6.139) to (6.140). The proof of Lemma 6.131 is now complete, because (6.140) is the same as (6.132) (by definition of H).

Now that we have these basic estimates, we want to adjust the mappings $g_{a,j}$ from (6.30) to get an admissible family. Recall that the $g_{a,j}$'s map into $\mathbf{R}^m$, as in (6.30).

Lemma 6.142. *For each integer $j \geq 0$ and element a of A_j there is a mapping $G_{a,j} : M \rightarrow \mathbf{R}^m$ such that*

$$(6.143) \quad \text{supp } G_{a,j} \subseteq \frac{1}{3} B_{a,j};$$

$$(6.144) \quad G_{a,j} \text{ is Lipschitz on } M \text{ with constant } \leq C_0;$$

$$(6.145) \quad G_{a,j} \text{ is bilipschitz on } E_{a,j} \text{ with constant } \leq C_0.$$

Here $C_0 \geq 1$ depends on the doubling and BPE constants for $(M, d(x, y), \mu)$, but not on a or j .

To prove Lemma 6.142, we start with the $g_{a,j}$'s from (6.30). Fix $j \geq 0$ and $a \in A_j$. Without loss of generality, we may assume that $g_{a,j}$ takes the value 0 somewhere on $E_{a,j}$; otherwise, we could make this be true by composing $g_{a,j}$ with a translation on $\mathbf{R}^m$, which would not affect the bilipschitz condition (6.30) (nor the $E_{a,j}$'s).

Using this normalization, we get that

$$(6.146) \quad |g_{a,j}(z)| \leq k \text{ diam } E_{a,j}$$

for all $z \in E_{a,j}$, since $g_{a,j}$ is Lipschitz on $E_{a,j}$ with constant k by (6.30). Let us write radius $B_{a,j}$ once again for the radius of the ball $B_{a,j}$, whose value is $10^{-1} \text{dist}(a, M \setminus \Omega_j)$, as in (6.28). (The actual value of the radius will not matter for the computation that we are about to perform, i.e., it will wash out in the end.) Because $E_{a,j} \subseteq \frac{1}{4}B_{a,j}$, by (6.31), we have that

$$(6.147) \quad \text{diam } E_{a,j} \leq 2 \text{ radius } \frac{1}{4}B_{a,j} = \frac{1}{2} \text{ radius } B_{a,j}.$$

Thus

$$(6.148) \quad |g_{a,j}(z)| \leq \frac{k}{2} \text{ radius } B_{a,j}$$

for all $z \in E_{a,j}$, by (6.146).

Let us extend $g_{a,j}$ to $M \setminus \frac{1}{3}B_{a,j}$ by setting it to be 0 there. This does not affect the original choice of $g_{a,j}$ on $E_{a,j}$, since the latter is contained in $\frac{1}{4}B_{a,j}$. Let us check that

$$(6.149) \quad g_{a,j} \text{ is Lipschitz with constant } 6k \text{ on } E_{a,j} \cup \left(M \setminus \frac{1}{3}B_{a,j} \right),$$

i.e.,

$$(6.150) \quad |g_{a,j}(x) - g_{a,j}(y)| \leq 6k d(x, y) \\ \text{whenever } x, y \in E_{a,j} \cup \left(M \setminus \frac{1}{3}B_{a,j} \right).$$

If x and y both lie in $M \setminus \frac{1}{3}B_{a,j}$, then (6.150) is trivial, since $g_{a,j}$ vanishes on $M \setminus \frac{1}{3}B_{a,j}$. If $x, y \in E_{a,j}$, then (6.150) follows from (6.30). The remaining possibility is that one of x, y lies in $M \setminus \frac{1}{3}B_{a,j}$ and the other in $E_{a,j} \subseteq \frac{1}{4}B_{a,j}$. In this case we have that

$$(6.151) \quad d(x, y) \geq \frac{1}{12} \text{ radius } B_{a,j},$$

and (6.150) follows from this and the fact that one of x and y satisfies (6.148), while $g_{a,j}$ vanishes at the other one. This proves (6.149).

Now choose $G_{a,j}$ to be a Lipschitz extension of $g_{a,j}$ from $E_{a,j} \cup (M \setminus \frac{1}{3}B_{a,j})$ to all of M (with values still in $\mathbf{R}^m$). We can do this while maintaining a bound for the Lipschitz constant, because of Proposition 6.2. Thus $G_{a,j}$ satisfies (6.144). We already set $g_{a,j}$ to be 0 on $M \setminus \frac{1}{3}B_{a,j}$, so that (6.143) holds, and (6.145) follows from the bilipschitz condition (6.30) for $g_{a,j}$ and the fact that $G_{a,j}$ is the same as $g_{a,j}$ on $E_{a,j}$, by construction. This completes the proof of Lemma 6.142.

Instead of the $h_{a,j}$'s from (6.106), it will be helpful to make localizations in a slightly more precise manner, as in the next lemma.

Lemma 6.152. *Let an integer $j \geq 0$ and an element a of A_j be given. Define $H_{a,j} : M \rightarrow \mathbf{R}$ by*

$$(6.153) \quad H_{a,j}(x) = \max(1.5 \text{ radius } B_{a,j} - \text{dist}(x, E_{a,j}), 0).$$

(As usual, $\text{radius } B_{a,j} = 10^{-1} \text{dist}(a, M \setminus \Omega_j)$, as in (6.28).) Then $H_{a,j}$ satisfies the following properties:

$$(6.154) \quad H_{a,j} \text{ is Lipschitz with constant } \leq 1;$$

$$(6.155) \quad \text{supp } H_{a,j} \subseteq 2B_{a,j};$$

$$(6.156) \quad H_{a,j}(x) \geq \frac{1}{4} \text{ radius } B_{a,j} \quad \text{when } x \in B_{a,j}.$$

In particular, $\{H_{a,j}\}_{a,j}$ is an admissible family (in the sense of Definition 6.103).

This is quite straightforward from the definitions, but let us be a bit careful. Notice first that each $E_{a,j}$ is nonempty, because of (6.29), so that the distance function on the right side of (6.153) makes sense. Any function of the form $x \mapsto \text{dist}(x, Z)$, where Z is a nonempty subset of M , is a 1-Lipschitz function of x , by a standard application of the triangle inequality. (This is practically the same as (6.9), for instance.) From this it follows that $H_{a,j}$ is 1-Lipschitz too, because the Lipschitz condition is not disturbed by the addition of a constant, or by taking the maximum (or minimum) with a constant (or any other 1-Lipschitz function). This gives (6.154).

In order to check (6.155) and (6.156), let us observe that

$$(6.157) \quad \begin{aligned} d(x, a) - \frac{1}{4} \text{ radius } B_{a,j} &\leq \text{dist}(x, E_{a,j}) \\ &\leq d(x, a) + \frac{1}{4} \text{ radius } B_{a,j}. \end{aligned}$$

These inequalities come from the fact that $E_{a,j} \subseteq \frac{1}{4}B_{a,j}$, as in (6.31). (More precisely, we are also using the nonemptiness of $E_{a,j}$, as mentioned above, and the knowledge that $B_{a,j}$ is centered at a , by its definition (6.28).) Once we have (6.157), the desired properties (6.155) and (6.156) follow from the definition (6.153) of $H_{a,j}$ and easy calculation. This proves Lemma 6.152.

The embedding f in Proposition 2.10 will be obtained by combining the $G_{a,j}$'s from Lemma 6.142 and the $H_{a,j}$'s from Lemma 6.152. To do this, we shall need a bit of coding, and this is our next task.

Let L_0 and L_1 be large positive constants, to be chosen later. The only penalty for making L_0, L_1 large will be in the size of the dimension of the target space $\mathbf{R}^\ell$ for the embedding f in Proposition 2.10. Put

$$(6.158) \quad \mathcal{C} = \{(a, j) : j \text{ is a nonnegative integer, and } a \in A_j\}.$$

Given j , let us say that $a, b \in A_j$ are L_1 -neighbors if

$$(6.159) \quad L_1^{-1} \text{dist}(b, M \setminus \Omega_j) \leq \text{dist}(a, M \setminus \Omega_j) \leq L_1 \text{dist}(b, M \setminus \Omega_j)$$

and

$$(6.160) \quad d(a, b) \leq L_1 \min(\text{dist}(a, M \setminus \Omega_j), \text{dist}(b, M \setminus \Omega_j)).$$

Lemma 6.161 (Coding lemma). *If L_2 is a positive integer which is sufficiently large (depending only on L_0, L_1 , and the doubling constant for $(M, d(x, y))$), then there is a mapping $\Gamma : \mathcal{C} \rightarrow \{1, 2, \dots, L_2\}$ (which is the “coding”) which satisfies the following property.*

Suppose that $(a, j), (a', j')$ are elements of $\mathcal{C}$ such that $\Gamma(a, j) = \Gamma(a', j')$. Then

$$(6.162) \quad j = j' \bmod L_0$$

and

$$(6.163) \quad a = a' \text{ if also } j = j' \text{ and } a, a' \text{ are } L_1\text{-neighbors.}$$

To prove this we shall use the following.

Sublemma 6.164. *For each $j \geq 0$ and $a \in A_j$ there are at most $C(L_1)$ choices of $b \in A_j$ such that a and b are L_1 -neighbors (where $C(L_1)$ depends on L_1 and the doubling constant for $(M, d(x, y))$, but not on j or a).*

We can derive this from the doubling condition for $(M, d(x, y))$. Let $N_j(a)$ denote the set of $b \in A_j$ which are L_1 -neighbors of a . Then

$$(6.165) \quad d(a, b) \leq L_1^2 \text{dist}(a, M \setminus \Omega_j)$$

for all $b \in N_j(a)$, by (6.159) and (6.160). Also,

$$(6.166) \quad B(b, (30L_1)^{-1} \text{dist}(a, M \setminus \Omega_j)) \subseteq \frac{1}{3} B_{b,j}$$

for all $b \in N_j(b)$, by (6.159) and the definition (6.28) of $B_{b,j}$. Thus the balls

$$(6.167) \quad B(b, (30L_1)^{-1} \text{dist}(a, M \setminus \Omega_j)), \quad b \in N_j(a),$$

are pairwise disjoint, because of part (iii) of Proposition 6.4. However, the doubling property for $(M, d(x, y))$ ensures that the ball

$$(6.168) \quad B(a, L_1^2 \text{dist}(a, M \setminus \Omega_j))$$

can be covered by at most $C(L_1)$ open balls of radius

$$(6.169) \quad \frac{1}{60L_1} \text{dist}(a, M \setminus \Omega_j).$$

In particular, every element b of $N_j(a)$ should be contained in at least one of these smaller balls, by (6.165). On the other hand, the disjointness of the balls in (6.167) guarantees that no two elements of $N_j(a)$ can lie in a single open ball of radius (6.169) (otherwise these elements would lie at distance $< (30L_1)^{-1} \text{dist}(a, M \setminus \Omega_j)$ from each other). Thus $N_j(a)$ can have at most $C(L_1)$ elements, which is exactly what we wanted. This proves Sublemma 6.164.

Next, let us observe that each A_j is at most countable. One can see this as a consequence of Sublemma 6.164, by taking $L_1 \rightarrow \infty$. That is, any pair of elements of A_j are L_1 -neighbors when L_1 is large enough. (One can also go back to the proof of Proposition 6.4.)

From this we have that $\mathcal{C}$ is at most countable as well, so that we can think of the elements of $\mathcal{C}$ as being enumerated by a single sequence. To define a mapping $\Gamma : \mathcal{C} \rightarrow \{1, 2, \dots, L_2\}$, one can proceed recursively, by choosing the value of Γ at one element of $\mathcal{C}$ after another, with respect to the enumeration that we have. We simply need to make certain that at each step of the process there is a choice available in $\{1, 2, \dots, L_2\}$ for the value of Γ , and one that is compatible with the requirements of Lemma 6.161 and the choices made at earlier steps.

The condition (6.162) is easy to arrange; it is enough to require that $L_0 \leq L_2$ and

$$(6.170) \quad \Gamma(a, j) = j \bmod L_0$$

for every $(a, j) \in \mathcal{C}$, for instance. As for (6.163), imagine that the values of Γ have already been chosen for the first q elements of $\mathcal{C}$ in our enumeration, and that one must now choose the value of Γ at the next term in the sequence. Let (a, j) denote this next term in the sequence.

We want $\Gamma(a, j)$ to satisfy (6.170), and this leaves about L_2/L_0 possible choices for $\Gamma(a, j)$ among the elements of $\{1, 2, \dots, L_2\}$.

In order to respect (6.163), we require that

$$(6.171) \quad \Gamma(a, j) \neq \Gamma(a', j)$$

if (a', j) lies among the q elements of $\mathcal{C}$ at which Γ has already been defined, and a, a' are L_1 -neighbors. There are at most $C(L_1)$ such points $a' \in A_j$, by Sublemma 6.164, and so it is possible to choose $\Gamma(a, j) \in \{1, 2, \dots, L_2\}$ in this manner as long as L_2 is sufficiently large, depending on L_0 and L_1 .

Thus one can choose the values of Γ recursively, and in such a way that the conditions of Lemma 6.161 are satisfied at each finite stage of the process. In the end we get a function Γ with values in $\{1, 2, \dots, L_2\}$ which is defined on all of $\mathcal{C}$, and which satisfies the required properties. This proves Lemma 6.161.

Let $\delta \in (0, 1]$ be arbitrary but fixed now, for the rest of the proof. Also let L_2 be as in Lemma 6.161, depending on choices of L_0 and L_1 that will be made later, near the end. For each integer i , $1 \leq i \leq L_2$, define $G_i : M \rightarrow \mathbf{R}^m$ (with m as in (6.30) and Lemma 6.142) and $H_i : M \rightarrow \mathbf{R}$ by

$$(6.172) \quad G_i(x) = \sum_{\substack{(a,j) \in \mathcal{C} \\ \Gamma(a,j)=i}} \lambda^{\delta j} G_{a,j}(x), \quad H_i(x) = \sum_{\substack{(a,j) \in \mathcal{C} \\ \Gamma(a,j)=i}} \lambda^{\delta j} H_{a,j}(x),$$

where the $G_{a,j}$'s are as in Lemma 6.142 and the $H_{a,j}$'s are as in (6.153). Thus we are using the "coding" $\Gamma : \mathcal{C} \rightarrow \{1, 2, \dots, L_2\}$ to separate the $G_{a,j}$'s and $H_{a,j}$'s into a bounded number of groups, each of which is then converted into a single mapping by summing, as in (6.172). The individual $G_{a,j}$'s and $H_{a,j}$'s in a single group are rather independent of each other, because of the way that the coding was selected.

We now define ℓ and $f : M \rightarrow \mathbf{R}^\ell$ by setting

$$(6.173) \quad \ell = (m+1)L_2$$

and

$$(6.174) \quad f = (G_1, G_2, \dots, G_{L_2}, H_1, H_2, \dots, H_{L_2}).$$

In other words, f is obtained by combining the G_i 's and H_i 's into a single mapping through a Cartesian product. This uses the obvious identification between $\mathbf{R}^\ell$ and the Cartesian product of L_2 copies of $\mathbf{R}^m$ and L_2 copies of $\mathbf{R}$.

To finish the proof of Proposition 2.10, it remains to show that f is bilipschitz as a map from $(M, D_\delta(x, y))$ into $\mathbf{R}^\ell$ with the standard Euclidean metric (and with suitable bounds for the bilipschitz constant).

Lemma 6.175. *There is a constant C so that*

$$(6.176) \quad |f(x) - f(y)| \leq C D_\delta(x, y)$$

for all $x, y \in M$, where C depends on δ , the BPE and doubling constants for $(M, d(u, v), \mu)$, and the choice of λ (from Standing Assumptions 6.98), but not on x or y (or on L_0, L_1 , for that matter).

This is an easy consequence of Lemma 6.107 and the admissibility of the families $\{G_{a,j}\}$, $\{H_{a,j}\}$. The admissibility of these families comes from Lemmas 6.142 and 6.152.

Lemma 6.177. *If L_0, L_1 are large enough (depending on δ , the BPE and doubling constants for $(M, d(u, v), \mu)$, and the choice of λ (from Standing Assumptions 6.98)), then*

$$(6.178) \quad |f(x) - f(y)| \geq C^{-1} D_\delta(x, y)$$

for some constant C and all $x, y \in M$. Here C also depends on δ , the BPE and doubling constants for $(M, d(u, v), \mu)$, and the choice of λ (from Standing Assumptions 6.98), but not on x or y .

Once we have this, the proof of Proposition 2.10 will be complete.

There are two main points behind the proof of Lemma 6.177 that is given below. The first is that for each pair of points $x, y \in M$ we shall be able to find $(a, j) \in \mathcal{C}$ such that

$$(6.179) \quad \text{either } |G_{a,j}(x) - G_{a,j}(y)| \text{ or } |H_{a,j}(x) - H_{a,j}(y)|$$

is not too small, compared to $d(x, y)$. The second is that the coding can be used to convert this into a lower bound for

$$(6.180) \quad |G_i(x) - G_i(y)| \quad \text{or} \quad |H_i(x) - H_i(y)|, \quad i = \Gamma(a, j).$$

The coding will help to ensure that the total contribution of the other terms involved in G_i or H_i is small enough so as not to disturb the lower bounds for (6.179). It is here that we shall need to take L_0, L_1 large enough, so that the other terms in G_i or H_i are sufficiently “far” from the main $G_{a,j}, H_{a,j}$ terms, in location or scale.

We begin with the following observation.

Sublemma 6.181. *Let $j \geq 0$ and $a, b \in A_j$ be given, and suppose that*

$$(6.182) \quad 5B_{a,j} \cap 5B_{b,j} \neq \emptyset.$$

If $L_1 \geq 3$, then a and b are L_1 -neighbors in A_j (in the sense of (6.159) and (6.160)).

In effect, this was verified in the first part of the proof of Proposition 6.4. Specifically, the assumption (6.182) corresponds to (6.7) in the proof of Proposition 6.4, and then (6.159) (with $L_1 = 3$) is the same as (6.11) and (6.12) before. Similarly, (6.160) follows (with $L_1 = 2$, in fact) from (6.13) and its variant with the roles of a and b reversed. This proves Sublemma 6.181.

From now on we shall assume that L_1 is at least 3, so that the conclusions of Sublemma 6.181 hold. (We shall impose more stringent conditions on L_1 later.)

Let us now begin the proof of Lemma 6.177 in earnest. Let $x, y \in M$ be given. We may as well assume that $x \neq y$, since the estimate in Lemma 6.177 is trivial when $x = y$.

Set

$$(6.183) \quad B_x = B(x, \tfrac{4}{3}d(x, y)), \quad B_y = B(y, \tfrac{4}{3}d(x, y)).$$

Let k_x, k_y be the nonnegative integers associated to these balls as in Lemma 6.99. We may assume that

$$(6.184) \quad k_y \leq k_x,$$

since otherwise we could interchange x and y . If it happens that $k_x = k_y$, then we require that

$$(6.185) \quad \text{dist}(y, M \setminus \Omega_{k_y}) \leq \text{dist}(x, M \setminus \Omega_{k_x}),$$

which, again, can be arranged by switching the roles of x and y . With these provisos, we now set

$$(6.186) \quad B = B_x, \quad k = k_x.$$

As in (6.123), we have that

$$(6.187) \quad C(\delta)^{-1} \lambda^{\delta k} d(x, y) \leq D_\delta(x, y) \leq C(\delta) \lambda^{\delta k} d(x, y).$$

That is, B and k are chosen here in exactly the same manner as in (6.123), and so the same argument applies as before.

Let us check that

$$(6.188) \quad \text{dist}(y, M \setminus \Omega_k) \leq \text{dist}(x, M \setminus \Omega_k),$$

as a consequence of our normalizing assumptions (6.184) and (6.185) on the roles of x and y . If $k_x = k_y$, then (6.188) is the same as (6.185), and there is nothing to do. Thus we assume that $k = k_x > k_y$, since the opposite inequality is ruled out by (6.184). In particular, we have that $k > 0$, since k_x and k_y are always nonnegative integers. The definition of $k = k_x$ and k_y implies that

$$(6.189) \quad B_x \subseteq \Omega_k \quad \text{and} \quad B_y \not\subseteq \Omega_k$$

(as in (6.100)). Using the definitions of B_x , B_y in (6.183), we obtain that

$$(6.190) \quad \text{dist}(x, M \setminus \Omega_k) \geq \frac{4}{3} d(x, y) \quad \text{and} \quad \text{dist}(y, M \setminus \Omega_k) < \frac{4}{3} d(x, y),$$

i.e., (6.190) is the same as (6.189). In this case we again have (6.188), by combining the two inequalities in (6.190).

We should also check that

$$(6.191) \quad x \in \Omega_k \quad \text{and} \quad d(x, y) \leq 2 \text{dist}(x, M \setminus \Omega_k).$$

There is a small subtlety here, in that k could be 0. That is, “normally” the choice of $k = k_x$ ensures that $B_x \subseteq \Omega_k$, as in (6.100), and this would give (6.191) immediately (by the definition (6.183) of B_x). However, if $k = 0$, then we have $\Omega_0 = M \setminus \{p_0\}$ (as mentioned shortly before (6.28)), and the conventions established in Lemma 6.99 allow for the possibility that $p_0 \in B_x$. However, if $k_x = k = 0$, then $k_y = 0$ as well, because of (6.184), and (6.185) becomes

$$(6.192) \quad d(y, p_0) \leq d(x, p_0).$$

This leads to

$$(6.193) \quad d(x, y) \leq 2d(x, p_0),$$

by adding $d(x, p_0)$ to both sides of (6.192) and applying the triangle inequality. This gives the inequality $d(x, y) \leq 2 \text{dist}(x, M \setminus \Omega_k)$ in (6.191) in this case, while $x \in \Omega_0$ follows from this inequality and the fact that

$d(x, y) > 0$ (since $x \neq y$). Thus (6.191) is always true, whether or not $k = 0$.

Because x lies in Ω_k , there is a $b \in A_k$ such that

$$(6.194) \quad x \in B_{b,k}.$$

This follows from part (i) of Proposition 6.4 (and the fact that $A_k = A(\Omega_k)$, by definition (just before (6.28))). Let us now fix a choice of $b \in A_k$ so that (6.194) holds, and also so that

$$(6.195) \quad \text{dist}(x, E_{b,k}) \text{ is minimal,}$$

i.e., minimal among the (finitely many) b 's in A_k for which (6.194) holds. This extra condition will be useful for guaranteeing that $\text{dist}(x, E_{b,k})$ is small compared to $\text{dist}(x, M \setminus \Omega_k)$ when this is possible, which is one of the basic situations of concern here.

Set

$$(6.196) \quad i = \Gamma(b, k).$$

These choices of b and i will be in force for the rest of the section.

To establish the inequality (6.178), and thereby prove Lemma 6.177, we shall distinguish between four cases. There is some flexibility in the details of the division of cases, but the main point is to either be in a position to use the local bilipschitz conditions in Lemma 6.142, or otherwise to be able to use the localizations provided by the $H_{a,j}$'s to separate points adequately.

Case I: $y \notin 2B_{b,k}$. For this we begin with the lower bound

$$(6.197) \quad |f(x) - f(y)| \geq |H_i(x) - H_i(y)|,$$

which follows immediately from the definition (6.174) of f . From the definition of H_i in (6.172) we can write

$$(6.198) \quad |H_i(x) - H_i(y)| \geq \beta_1 - \epsilon_1,$$

where the main term β_1 is equal to

$$(6.199) \quad \beta_1 = \left| \sum_{\substack{a \in A_k \\ \Gamma(a,k)=i}} \lambda^{\delta_k} (H_{a,k}(x) - H_{a,k}(y)) \right|,$$

and the error term ϵ_1 is given by

$$(6.200) \quad \epsilon_1 = \sum_{\substack{(a,j) \in \mathcal{C}: j \neq k \\ \Gamma(a,k)=i}} \lambda^{\delta j} |H_{a,j}(x) - H_{a,j}(y)|.$$

More precisely, (6.198) is derived from (6.172) by separating the terms in the definition of H_i according to whether $j = k$ or not, and then applying the triangle inequality to get

$$(6.201) \quad \beta_1 \leq |H_i(x) - H_i(y)| + \epsilon_1$$

(which is the same as (6.198)).

We want to analyze β_1 and ϵ_1 .

Sublemma 6.202. *Notations and assumptions as above. If $L_1 \geq 40$, then*

$$(6.203) \quad \beta_1 \geq \frac{1}{88} \lambda^{\delta k} \text{dist}(x, M \setminus \Omega_k) \geq \frac{1}{176} \lambda^{\delta k} d(x, y).$$

The second inequality in (6.203) follows automatically from (6.191), and so we only need to consider the first one.

We should first check which terms in the sum in (6.199) are nonzero. Suppose that $a \in A_k$ satisfies $\Gamma(a, k) = i$ and $H_{a,k}(x) \neq 0$. The latter implies that $x \in 2B_{a,k}$, by (6.155). We also have that $x \in B_{b,j}$, as in (6.194), so that $2B_{a,k}$ and $B_{b,j}$ have nonempty intersection. This means that a and b are L_1 -neighbors in A_k , by Sublemma 6.181. Therefore $a = b$, because of Lemma 6.161 and the fact that $\Gamma(a, k) = i = \Gamma(b, k)$.

Conversely, $H_{b,k}(x)$ is indeed nonzero, and

$$(6.204) \quad H_{b,k}(x) \geq \frac{1}{44} \lambda^{\delta k} \text{dist}(x, M \setminus \Omega_k).$$

To see this, notice first that

$$(6.205) \quad H_{b,k}(x) \geq \frac{1}{4} \text{radius } B_{b,k} = \frac{1}{40} \text{dist}(b, M \setminus \Omega_k).$$

The first step in (6.205) is the same as (6.156), while the second step is just the definition of radius $B_{b,k}$. To go from (6.205) to (6.204), we apply (6.36) in Lemma 6.35 (with b instead of a), again using the fact that $x \in B_{b,k}$.

Now suppose that $H_{a,k}(y) \neq 0$ for some $a \in A_k$ with $\Gamma(a, k) = i$. Then $y \in 2B_{a,k}$, by (6.155). If a' is another element of A_k with $\Gamma(a', k) = i$ and $H_{a',k}(y) \neq 0$, then we also have that $y \in 2B_{a',k}$, so that $2B_{a,k}$ and $2B_{a',k}$ intersect. As before, this means that a and a' are L_1 -neighbors in A_k , by Sublemma 6.181, and hence that $a = a'$, because of the properties of the coding in Lemma 6.161. In short, there is at most one possible choice for $a \in A_k$ with $H_{a,k}(y) \neq 0$ and $\Gamma(a, k) = i$, and there may simply be none.

To summarize, either

$$(6.206) \quad \beta_1 = \lambda^{\delta k} H_{b,k}(x)$$

or

$$(6.207) \quad \beta_1 = \lambda^{\delta k} |H_{b,k}(x) - H_{a,k}(y)|,$$

where $a \in A_k$, $H_{a,k}(y) \neq 0$, and $\Gamma(a, k) = i$. All of the other terms in (6.199) must vanish, by the arguments above.

If (6.206) holds, then the first inequality in (6.203) follows from (6.204), and we are finished. Thus we suppose that (6.207) is true instead. If

$$(6.208) \quad H_{a,k}(y) \leq \frac{1}{80} \text{dist}(b, M \setminus \Omega_k),$$

then

$$(6.209) \quad \begin{aligned} H_{b,k}(x) - H_{a,k}(y) &\geq \frac{1}{40} \text{dist}(b, M \setminus \Omega_k) - \frac{1}{80} \text{dist}(b, M \setminus \Omega_k) \\ &\geq \frac{1}{80} \text{dist}(b, M \setminus \Omega_k) \\ &\geq \frac{1}{88} \text{dist}(x, M \setminus \Omega_k), \end{aligned}$$

where the first step comes from (6.205), and the last step uses Lemma 6.35 again, just as in the transition from (6.205) to (6.204). This and (6.207) imply (6.203).

Finally, we consider the possibility that (6.207) holds, with $a \in A_k$, $H_{a,k}(y) \neq 0$, and $\Gamma(a, k) = i$, and that

$$(6.210) \quad H_{a,k}(y) > \frac{1}{80} \text{dist}(b, M \setminus \Omega_k)$$

too (i.e., the opposite of (6.208)). By the definition (6.153) of $H_{a,k}$, the maximal value of $H_{a,k}$ is

$$(6.211) \quad \begin{aligned} 1.5 \cdot \text{radius } B_{a,k} &= 1.5 \cdot 10^{-1} \text{dist}(a, M \setminus \Omega_k) \\ &\leq \frac{1}{5} \text{dist}(a, M \setminus \Omega_k), \end{aligned}$$

and so (6.210) yields

$$(6.212) \quad \text{dist}(a, M \setminus \Omega_k) > \frac{5}{80} \text{dist}(b, M \setminus \Omega_k) = \frac{1}{16} \text{dist}(b, M \setminus \Omega_k).$$

We want to use this to show that

$$(6.213) \quad a \text{ and } b \text{ are } L_1\text{-neighbors in } A_k,$$

at least if L_1 is large enough.

Let us first check that

$$(6.214) \quad \text{dist}(a, M \setminus \Omega_k) \leq 2 \text{dist}(b, M \setminus \Omega_k).$$

Since $H_{a,k}(y) \neq 0$, we have that

$$(6.215) \quad d(y, a) \leq 2 \text{radius } B_{a,k} = \frac{1}{5} \text{dist}(a, M \setminus \Omega_k),$$

by (6.155) and the fact that $\text{radius } B_{a,k} = 10^{-1} \text{dist}(a, M \setminus \Omega_k)$. On the other hand,

$$(6.216) \quad \text{dist}(a, M \setminus \Omega_k) \leq d(y, a) + \text{dist}(y, M \setminus \Omega_k),$$

as in (6.9), so that

$$(6.217) \quad \frac{4}{5} \text{dist}(a, M \setminus \Omega_k) \leq \text{dist}(y, M \setminus \Omega_k),$$

using (6.215). Combining this with (6.188) yields

$$(6.218) \quad \frac{4}{5} \text{dist}(a, M \setminus \Omega_k) \leq \text{dist}(x, M \setminus \Omega_k).$$

We also have that

$$(6.219) \quad \text{dist}(x, M \setminus \Omega_k) \leq \frac{11}{10} \text{dist}(b, M \setminus \Omega_k),$$

by (6.36) in Lemma 6.35 (with b instead of a) and the fact that $x \in B_{b,k}$, and (6.214) follows immediately from this and (6.218).

From (6.212) and (6.214) we have that a and b satisfy the first requirement (6.159) for being L_1 -neighbors in A_k as soon as $L_1 \geq 16$. We want to verify that the second requirement (6.160) also holds when L_1 is large enough. From the triangle inequality we have that

$$(6.220) \quad d(a, b) \leq d(a, y) + d(y, x) + d(x, b).$$

This leads to

$$(6.221) \quad d(a, b) \leq \frac{1}{5} \operatorname{dist}(a, M \setminus \Omega_k) + 2 \operatorname{dist}(x, M \setminus \Omega_k) + \frac{1}{10} \operatorname{dist}(b, M \setminus \Omega_k),$$

because of (6.215), (6.191), and the fact that $x \in B_{b,k}$. Combining this with (6.219) we obtain that

$$(6.222) \quad d(a, b) \leq \frac{1}{5} \operatorname{dist}(a, M \setminus \Omega_k) + \frac{23}{10} \operatorname{dist}(b, M \setminus \Omega_k).$$

If we use (6.212), then we can convert (6.222) into

$$(6.223) \quad d(a, b) \leq \frac{2 + 368}{10} \operatorname{dist}(a, M \setminus \Omega_k),$$

while (6.214) gives

$$(6.224) \quad d(a, b) \leq \frac{4 + 23}{10} \operatorname{dist}(b, M \setminus \Omega_k).$$

Thus

$$(6.225) \quad d(a, b) \leq 40 \min(\operatorname{dist}(a, M \setminus \Omega_k), \operatorname{dist}(b, M \setminus \Omega_k)).$$

This shows that (6.160) holds when $L_1 \geq 40$.

To recapitulate, for this last part of the proof of Sublemma 6.202, we have assumed that (6.210) and (6.207) hold, with $a \in A_k$, $H_{a,k}(y) \neq 0$, and $\Gamma(a, k) = i$, and we have shown that (6.213) is true (i.e., a and b are L_1 -neighbors), at least when $L_1 \geq 40$. From this we may conclude that $a = b$, because of (6.163) in Lemma 6.161 (and the fact that $\Gamma(a, k) = i = \Gamma(b, k)$). On the other hand, the information that $H_{a,k}(y) \neq 0$ implies that $y \in 2B_{a,k}$, by (6.155). This is the same as saying that $y \in 2B_{b,k}$, since $a = b$, which is exactly what is ruled out by the hypothesis of Case I, under which we are currently working.

Thus the last situation for Sublemma 6.202 (described in (6.210) and the lines immediately before it) simply does not occur in Case I. We have already seen that the conclusion (6.203) of Sublemma 6.202 holds in the other possible circumstances (i.e., when (6.206) is true, or (6.207) and (6.208) are satisfied), and so Sublemma 6.202 is completely proved.

Now that we have a good lower bound for β_1 , we want to bound ϵ_1 from above, and show that it is small compared to β_1 .

Let us begin with the observation that

$$(6.226) \quad \epsilon_1 \leq \sum_{\substack{(a,j) \in \mathcal{C} \\ |j-k| \geq L_0}} \lambda^{\delta j} |H_{a,j}(x) - H_{a,k}(y)|.$$

The right-hand side here in (6.226) differs from the definition (6.200) of ϵ_1 only in the collection of pairs (a, j) over which the sum extends. The collection of pairs (a, j) in (6.226) contains the corresponding collection for (6.200) (and hence the sum in (6.226) is larger). This follows from the properties of our coding, as formulated in Lemma 6.161. Namely, if $(a, j) \in \mathcal{C}$ satisfies $\Gamma(a, j) = i$ and $j \neq k$, as in (6.200), then we must have $|j - k| \geq L_0$, because of (6.162) in Lemma 6.161 and the fact that $\Gamma(b, k) = i$ by construction (as in (6.196)). This gives (6.226) from (6.200).

We may now apply Lemma 6.131 (with $\psi_{a,j} = H_{a,j}$ and $p = L_0$) to obtain that

$$(6.227) \quad \epsilon_1 \leq C \lambda^{-\delta L_0} D_\delta(x, y),$$

as a consequence of (6.226). This constant C does not depend on L_0 , x , or y , but only on the usual parameters. Let us rewrite (6.203) in Sublemma 6.202 as

$$(6.228) \quad \beta_1 \geq C^{-1} D_\delta(x, y),$$

using (6.187). From (6.227) and (6.228) we conclude that

$$(6.229) \quad \beta_1 - \epsilon_1 \geq C^{-1} D_\delta(x, y)$$

when L_0 is large enough, depending on the usual parameters (and not on x or y in particular). The constant C in (6.229) also depends only on the usual parameters.

Combining (6.229) with (6.197) and (6.198), we obtain that

$$(6.230) \quad |f(x) - f(y)| \geq C^{-1} D_\delta(x, y)$$

when L_0 is large enough, with the same conditions on L_0 and the same constant C as in (6.229). This gives (6.178) in Lemma 6.177 under the circumstances of Case I, which is what we wanted.

Case II: $y \in 2B_{b,k}$, and

$$(6.231) \quad |H_{b,k}(x) - H_{b,k}(y)| \geq \frac{1}{40} C_0^{-2} d(x, y),$$

where $C_0 \geq 1$ is the constant from Lemma 6.142.

This case can be handled in practically the same manner as Case I. We use the same two initial steps (6.197), (6.198) as before, with the same definitions of β_1 and ϵ_1 . The analogue of Sublemma 6.202 is much simpler now, because we have

$$(6.232) \quad \beta_1 = \lambda^{\delta k} |H_{b,k}(x) - H_{b,k}(y)|$$

in place of (6.206), (6.207). This can be checked through the same kind of arguments as in the first part of the proof of Sublemma 6.202. The main point is that the $a \in A_k$ which occurs in (6.207) has to be equal to b under the circumstances of the present case, since $y \in 2B_{b,k}$, and because of the coding property (6.163) in Lemma 6.161.

Combining (6.232) and (6.231) we get that

$$(6.233) \quad \beta_1 \geq \frac{1}{40} C_0^{-2} \lambda^{\delta k} d(x, y).$$

This is exactly analogous to the conclusion (6.203) of Sublemma 6.202. From here one can proceed as in Case I, to conclude that (6.178) holds in Case II as well.

Case III: $y \in 2B_{b,k}$, and we also have that

$$(6.234) \quad |H_{b,k}(x) - H_{b,k}(y)| < \frac{1}{40} C_0^{-2} d(x, y),$$

and

$$(6.235) \quad \text{dist}(x, E_k) < \frac{1}{40} C_0^{-2} d(x, y),$$

where $C_0 \geq 1$ is again the constant from Lemma 6.142.

Let us begin with some preliminary observations pertaining to the hypotheses (6.234) and (6.235) above.

Sublemma 6.236. $\text{dist}(x, E_k) = \text{dist}(x, E_{b,k})$.

Indeed, suppose to the contrary that there is a point $z \in E_k \setminus E_{b,k}$ such that

$$(6.237) \quad d(x, z) < \text{dist}(x, E_{b,k}).$$

Because of (6.235), such a point z satisfies

$$(6.238) \quad d(x, z) < \frac{1}{40} C_0^{-2} d(x, y).$$

This gives

$$(6.239) \quad d(x, z) < \frac{1}{20} C_0^{-2} \operatorname{dist}(x, M \setminus \Omega_k),$$

by (6.191).

By definition (see (6.32)), E_k is the union of the sets $E_{a,k}$, $a \in A_k$. Thus there is an $a \in A_k$ such that $z \in E_{a,k}$. Since we are assuming that $z \notin E_{b,k}$, we should have that $a \neq b$.

We would like to show that $x \in B_{a,k}$. If we can do that, then (6.237) would contradict the minimality condition (6.195) in the choice of b (since $\operatorname{dist}(x, E_{a,k}) \leq d(x, z)$ automatically), and we would be finished with the proof of Sublemma 6.236.

Remember from (6.31) that $E_{a,k} \subseteq \frac{1}{4} B_{a,k}$. To show that $x \in B_{a,k}$, it therefore suffices to show that

$$(6.240) \quad d(x, z) \leq \frac{3}{4} \operatorname{radius} B_{a,k} = \frac{3}{4} 10^{-1} \operatorname{dist}(a, M \setminus \Omega_k),$$

by the definition (6.28) of $B_{a,k}$. Because of (6.36) in Lemma 6.35 (applied to $z \in B_{a,k}$ instead of x), it is enough to show that

$$(6.241) \quad d(x, z) \leq \frac{3}{4} 11^{-1} \operatorname{dist}(z, M \setminus \Omega_k)$$

(i.e., this would imply (6.240)).

As a general fact, we have that

$$(6.242) \quad \operatorname{dist}(x, M \setminus \Omega_k) \leq d(x, z) + \operatorname{dist}(z, M \setminus \Omega_k),$$

as in (6.9). Combining this with (6.239) (and the inequality $C_0 \geq 1$) yields

$$(6.243) \quad \operatorname{dist}(x, M \setminus \Omega_k) \leq \frac{1}{20} \operatorname{dist}(x, M \setminus \Omega_k) + \operatorname{dist}(z, M \setminus \Omega_k).$$

Hence

$$(6.244) \quad \operatorname{dist}(x, M \setminus \Omega_k) \leq \frac{20}{19} \operatorname{dist}(z, M \setminus \Omega_k).$$

Putting this back into (6.239) we get that

$$(6.245) \quad d(x, z) \leq \frac{1}{19} \operatorname{dist}(z, M \setminus \Omega_k).$$

This implies (6.241). Sublemma 6.236 follows from here, as mentioned before.

Sublemma 6.246. $H_{b,k}(x) = 1.5 \text{ radius } B_{b,k} - \text{dist}(x, E_{b,k})$.

Indeed, $H_{b,k}(x) = \max(1.5 \text{ radius } B_{b,k} - \text{dist}(x, E_{b,k}), 0)$, by definition. (See (6.153).) Because $x \in B_{b,k}$, as in (6.194), we have that

$$(6.247) \quad \text{dist}(x, E_{b,k}) \leq 1.5 \text{ radius } B_{b,k},$$

by the second inequality in (6.157). This proves Sublemma 6.246.

Sublemma 6.248. $\text{dist}(y, E_{b,k}) < \frac{1}{20} C_0^{-2} d(x, y)$.

To see this, we begin with the observation that

$$(6.249) \quad H_{b,k}(y) \geq H_{b,k}(x) - |H_{b,k}(x) - H_{b,k}(y)|,$$

by the triangle inequality. Using Sublemma 6.246 we can convert this into

$$(6.250) \quad H_{b,k}(y) \geq 1.5 \text{ radius } B_{b,k} - \text{dist}(x, E_{b,k}) - |H_{b,k}(x) - H_{b,k}(y)|.$$

Substituting (6.234) and (6.235) into (6.250), we obtain

$$(6.251) \quad \begin{aligned} H_{b,k}(y) &> 1.5 \text{ radius } B_{b,k} - \frac{1}{40} C_0^{-2} d(x, y) - \frac{1}{40} C_0^{-2} d(x, y) \\ &= 1.5 \text{ radius } B_{b,k} - \frac{1}{20} C_0^{-2} d(x, y). \end{aligned}$$

If we can show that

$$(6.252) \quad H_{b,k}(y) > 0,$$

then Sublemma 6.248 will follow, because

$$(6.253) \quad H_{b,k}(y) = 1.5 \text{ radius } B_{b,k} - \text{dist}(y, E_{b,k}) \quad \text{when } H_{b,k}(y) > 0,$$

by the definition (6.153) of $H_{b,k}(y)$.

It is enough to show that

$$(6.254) \quad \frac{1}{20} d(x, y) < 1.5 \text{ radius } B_{b,k}$$

instead of (6.252), because of (6.251) and the fact that $C_0 \geq 1$. By definition,

$$(6.255) \quad \text{radius } B_{b,k} = \frac{1}{10} \text{dist}(b, M \setminus \Omega_k),$$

while

$$(6.256) \quad \text{dist}(x, M \setminus \Omega_k) \leq \frac{11}{10} \text{dist}(b, M \setminus \Omega_k),$$

as in (6.36) in Lemma 6.35 (with b in place of a). (Remember that $x \in B_{b,k}$, by our choice of b . See (6.194).) We also have that

$$(6.257) \quad d(x, y) \leq 2 \text{dist}(x, M \setminus \Omega_k),$$

by (6.191). Thus

$$(6.258) \quad \frac{1}{20} d(x, y) \leq \frac{1}{10} \text{dist}(x, M \setminus \Omega_k) \leq \frac{11}{100} \text{dist}(b, M \setminus \Omega_k),$$

using (6.256) for the second step. The last expression is strictly less than 1.5 radius $B_{b,k}$, by (6.255). This gives (6.254), and Sublemma 6.248 follows.

Sublemmas 6.236 and 6.248 and the assumption (6.235) tell us that x and y lie close to $E_{b,k}$. We want to use this to bound $|f(x) - f(y)|$ from below, by reducing to the bilipschitz condition on $G_{b,k}$ from Lemma 6.142. We start with the inequality

$$(6.259) \quad |f(x) - f(y)| \geq |G_i(x) - G_i(y)|,$$

which follows automatically from (6.174). This time we write

$$(6.260) \quad |G_i(x) - G_i(y)| \geq \beta_2 - \epsilon_2,$$

where the main term β_2 is given by

$$(6.261) \quad \beta_2 = \left| \sum_{\substack{a \in A_k \\ \Gamma(a,k)=i}} \lambda^{\delta k} (G_{a,k}(x) - G_{a,k}(y)) \right|,$$

and the error term is

$$(6.262) \quad \epsilon_2 = \sum_{\substack{(a,j) \in \mathcal{C}: j \neq k \\ \Gamma(a,k)=i}} \lambda^{\delta j} |G_{a,j}(x) - G_{a,j}(y)|.$$

The derivation of (6.260) from the definition (6.172) of G_i is completely analogous to the derivation of (6.198) from the definition of H_i before, in (6.172); that is, we separate terms in the sum in (6.172) that defines G_i according to whether $j = k$ or not, and then use the triangle inequality to get (6.260).

Under the present conditions, the formula for β_2 reduces to

$$(6.263) \quad \beta_2 = \lambda^{\delta k} |G_{b,k}(x) - G_{b,k}(y)|.$$

In other words,

$$(6.264) \quad G_{a,k}(x) = G_{a,k}(y) = 0 \quad \text{when } a \in A_k, \Gamma(a, k) = i, a \neq b.$$

This follows from considerations of coding and L_1 -neighbors in A_k similar to ones in the treatment of the previous two cases. Specifically, we have that $x \in B_{b,k}$ by the choice of b , as in (6.194), while $y \in 2B_{b,k}$ by the hypothesis of this case. If $G_{a,k}(x)$ or $G_{a,k}(y)$ is not zero, it means that x or y lies in $\frac{1}{3}B_{a,k}$, by (6.143) in Lemma 6.142. This leads to

$$(6.265) \quad 2B_{b,k} \cap B_{a,k} \neq \emptyset,$$

and hence that a and b are L_1 -neighbors, as in Sublemma 6.181. Thus $a = b$ when $\Gamma(a, k) = i$ ($= \Gamma(b, k)$), because of the property (6.163) in Lemma 6.161. This gives (6.264), and (6.263) then follows from (6.261).

We want to use the bilipschitz condition (6.145) for $G_{b,k}$ in Lemma 6.142 to obtain a lower bound for β_2 . This bilipschitz condition applies only to elements of $E_{b,k}$, and so we cannot apply it directly to x and y . Thus we first choose $x', y' \in E_{b,k}$ such that

$$(6.266) \quad d(x, x') < \frac{1}{40}C_0^{-2}d(x, y) \quad \text{and} \quad d(y, y') < \frac{1}{20}C_0^{-2}d(x, y).$$

We can do this, because (6.235) and Sublemmas 6.236 and 6.248 ensure that $\text{dist}(x, E_{b,k}) < \frac{1}{40}C_0^{-2}d(x, y)$ and $\text{dist}(y, E_{b,k}) < \frac{1}{20}C_0^{-2}d(x, y)$.

The bilipschitz condition (6.145) for $G_{b,k}$ yields

$$(6.267) \quad |G_{b,k}(x') - G_{b,k}(y')| \geq C_0^{-1}d(x', y').$$

We want to convert this into information about x and y . The Lipschitz condition (6.144) for $G_{b,k}$ and (6.266) imply that

$$(6.268) \quad |G_{b,k}(x) - G_{b,k}(x')| + |G_{b,k}(y) - G_{b,k}(y')| \leq C_0(d(x, x') + d(y, y')) < \frac{3}{40}C_0^{-1}d(x, y).$$

Thus

$$\begin{aligned}
 |G_{b,k}(x) - G_{b,k}(y)| &\geq |G_{b,k}(x') - G_{b,k}(y')| \\
 &\quad - |G_{b,k}(x) - G_{b,k}(x')| - |G_{b,k}(y) - G_{b,k}(y')| \\
 &> C_0^{-1} d(x', y') - \frac{3}{40} C_0^{-1} d(x, y).
 \end{aligned}
 \tag{6.269}$$

On the other hand,

$$\begin{aligned}
 d(x', y') &\geq d(x, y) - d(x, x') - d(y, y') \\
 &> d(x, y) - \frac{3}{40} d(x, y),
 \end{aligned}
 \tag{6.270}$$

by the triangle inequality, (6.266), and the fact that $C_0 \geq 1$. This permits us to convert (6.269) into

$$\begin{aligned}
 |G_{b,k}(x) - G_{b,k}(y)| &> C_0^{-1} d(x, y) - \frac{6}{40} C_0^{-1} d(x, y) \\
 &\geq (2C_0)^{-1} d(x, y).
 \end{aligned}
 \tag{6.271}$$

Rewriting this in terms of β_2 , we have that

$$\beta_2 \geq (2C_0)^{-1} \lambda^{\delta k} d(x, y) \geq C^{-1} D_\delta(x, y)
 \tag{6.272}$$

for a suitable constant C , by (6.263) and (6.187). At this stage we are in practically the same situation as in Case I, once Sublemma 6.202 was established. Just as in (6.226), we have that

$$\epsilon_2 \leq \sum_{\substack{(a,j) \in \mathcal{C} \\ |j-k| \geq L_0}} \lambda^{\delta j} |G_{a,j}(x) - G_{a,k}(y)|,
 \tag{6.273}$$

because of the properties of our coding function Γ in Lemma 6.161. (I.e., $(a, j) \in \mathcal{C}$, $\Gamma(a, j) = i = \Gamma(b, k)$, and $j \neq k$ imply that $|j - k| \geq L_0$, by (6.162).) From here we can apply Lemma 6.131 (with $\psi_{a,j} = G_{a,j}$ and $p = L_0$) to get that

$$\epsilon_2 \leq C \lambda^{-\delta L_0} D_\delta(x, y).
 \tag{6.274}$$

(Recall that the admissibility of the family $\{G_{a,j}\}$, required for Lemma 6.131, comes from Lemma 6.142.) If L_0 is large enough (depending on the usual parameters), then we obtain that

$$\beta_2 - \epsilon_2 \geq C^{-1} D_\delta(x, y)
 \tag{6.275}$$

for a suitable constant C , because of (6.272) and (6.274). Combining this with (6.259) and (6.260), we conclude that

$$(6.276) \quad |f(x) - f(y)| \geq C^{-1} D_\delta(x, y).$$

Thus (6.178) holds under the conditions of Case III, as desired.

Case IV: We have that

$$(6.277) \quad \text{dist}(x, E_k) \geq \frac{1}{40} C_0^{-2} d(x, y),$$

with $C_0 \geq 1$ as in Lemma 6.142. (This is the opposite of the third assumption (6.235) in Case III.)

This case will be practically the same as Case I, but with $k+1$ in the role that k had before. The next sublemma will put (6.277) in a slightly more convenient form.

Sublemma 6.278. *If $z \in \Omega_k$, then*

$$(6.279) \quad \text{dist}(z, M \setminus \Omega_{k+1}) = \text{dist}(z, E_k).$$

To see this, remember that $\Omega_{k+1} = \Omega_k \setminus E_k$ by definition. (See (6.33).) If we can show that

$$(6.280) \quad \text{dist}(z, E_k) \leq \text{dist}(z, M \setminus \Omega_k),$$

then (6.279) will follow automatically.

Since $z \in \Omega_k$, there is an $a = a(z) \in A_k$ such that $z \in B_{a,k}$, as in part (i) of Proposition 6.4. We also have that $E_{a,k} \subseteq E_k$, as in (6.32), so that

$$(6.281) \quad \text{dist}(z, E_k) \leq \text{dist}(z, E_{a,k}).$$

This leads to

$$(6.282) \quad \text{dist}(z, E_k) \leq \frac{5}{4} \text{radius } B_{a,k} = \frac{5}{40} \text{dist}(a, M \setminus \Omega_k),$$

by (6.31), (6.29) (which ensures that $E_{a,k} \neq \emptyset$), (6.28) (for the value of radius $B_{a,k}$), and the fact that $z \in B_{a,k}$. From (6.37) in Lemma 6.35 (with x replaced with z and j replaced with k) we have that

$$(6.283) \quad \text{dist}(a, M \setminus \Omega_k) \leq \frac{10}{9} \text{dist}(z, M \setminus \Omega_k),$$

and hence

$$(6.284) \quad \text{dist}(z, E_k) \leq \frac{5}{36} \text{dist}(z, M \setminus \Omega_k).$$

Thus (6.280) holds, and Sublemma 6.278 follows.

Because of (6.279) (with z taken to be x), we may rewrite our assumption (6.277) as

$$(6.285) \quad \text{dist}(x, M \setminus \Omega_{k+1}) \geq \frac{1}{40} C_0^{-2} d(x, y).$$

In particular, $x \in \Omega_{k+1}$ (since $d(x, y) > 0$). By part (i) of Proposition 6.4 we can choose a point $q \in A_{k+1}$ such that

$$(6.286) \quad x \in B_{q, k+1}.$$

This choice of q should be considered as fixed for the rest of the argument.

The next assertion makes the similarity between Case I and the present circumstances more precise.

Sublemma 6.287. $y \notin 2B_{q, k+1}$.

Suppose to the contrary that y did lie in $2B_{q, k+1}$, so that

$$(6.288) \quad d(y, x) \leq 2 \text{radius } B_{q, k+1} = \frac{1}{5} \text{dist}(q, M \setminus \Omega_{k+1})$$

(using (6.28) in the last step). From Lemma 6.35 (with $B_{q, k+1}$ instead of $B_{a, j}$) we have that

$$(6.289) \quad \frac{9}{10} \text{dist}(q, M \setminus \Omega_{k+1}) \leq \text{dist}(x, M \setminus \Omega_{k+1}) \leq \frac{11}{10} \text{dist}(q, M \setminus \Omega_{k+1}),$$

so that

$$(6.290) \quad d(x, y) \leq \frac{2}{9} \text{dist}(x, M \setminus \Omega_{k+1}).$$

In particular,

$$(6.291) \quad \frac{4}{3} d(x, y) < \text{dist}(x, M \setminus \Omega_{k+1}),$$

and hence

$$(6.292) \quad B_x = B(x, \frac{4}{3} d(x, y)) \subseteq \Omega_{k+1}.$$

This contradicts the way that k was originally chosen, from (6.186) and the line just after (6.183). (In the end the choice of k goes back to Lemma 6.99. The main point is that $k = k_x$ was exactly chosen to be as small as possible so that (6.292) would *not* happen.) Thus our original assumption that $y \in 2B_{q,k+1}$ is not correct, and Sublemma 6.287 follows.

Set

$$(6.293) \quad t = \Gamma(q, k+1).$$

Analogous to the argument in Case I, we shall use the lower bound

$$(6.294) \quad |f(x) - f(y)| \geq |H_t(x) - H_t(y)|,$$

which follows immediately from the definition (6.174) of f . As before, we have that

$$(6.295) \quad |H_t(x) - H_t(y)| \geq \beta_3 - \epsilon_3,$$

where

$$(6.296) \quad \beta_3 = \left| \sum_{\substack{a \in A_{k+1} \\ \Gamma(a, k+1)=t}} \lambda^{\delta(k+1)} (H_{a, k+1}(x) - H_{a, k+1}(y)) \right|,$$

and

$$(6.297) \quad \epsilon_3 = \sum_{\substack{(a, j) \in \mathcal{C}: j \neq k+1 \\ \Gamma(a, k+1)=t}} \lambda^{\delta j} |H_{a, j}(x) - H_{a, j}(y)|.$$

Sublemma 6.298. *Notations and assumptions as above. If L_1 is sufficiently large, depending only on C_0 , then*

$$(6.299) \quad \beta_3 \geq \frac{1}{88} \lambda^{\delta(k+1)} \text{dist}(x, M \setminus \Omega_{k+1}).$$

This is similar to Sublemma 6.202. The first main point is that we must either have

$$(6.300) \quad \beta_3 = \lambda^{\delta(k+1)} H_{q, k+1}(x)$$

or

$$(6.301) \quad \beta_3 = \lambda^{\delta(k+1)} |H_{q, k+1}(x) - H_{a, k+1}(y)|$$

for some $a \in A_{k+1}$ such that $H_{a,k+1}(y) \neq 0$ and $\Gamma(a, k+1) = t$. This corresponds to (6.206) and (6.207) in the earlier situation, and the same considerations apply. (Specifically, one uses the property (6.163) in Lemma 6.161 of the coding function Γ to show that all of the other terms in the sum defining β_3 vanish. This also uses the localization for $H_{q,k+1}$ provided by (6.155) in Lemma 6.152, and the criterion for being neighbors in Sublemma 6.181.)

For $H_{q,k+1}(x)$ we have the lower bound

$$(6.302) \quad H_{q,k+1}(x) \geq \frac{1}{44} \lambda^{\delta(k+1)} \text{dist}(x, M \setminus \Omega_{k+1}).$$

This can be proved in exactly the same manner as (6.204). That is, one first observes that

$$(6.303) \quad H_{q,k+1}(x) \geq \frac{1}{4} \text{radius } B_{q,k+1} = \frac{1}{40} \text{dist}(q, M \setminus \Omega_{k+1}),$$

by (6.286), (6.156), and the definition (6.28) of $B_{q,k+1}$. One can then derive (6.302) from (6.303) and (6.289).

If the first alternative (6.300) for the value of β_3 occurs, then (6.299) is an immediate consequence of (6.302). Thus we assume instead that the second possibility (6.301) holds.

Assume for the moment that

$$(6.304) \quad H_{a,k+1}(y) \leq \frac{1}{80} \text{dist}(q, M \setminus \Omega_{k+1})$$

also holds, so that

$$\begin{aligned} (6.305) \quad H_{q,k+1}(x) - H_{a,k+1}(y) &\geq \frac{1}{40} \text{dist}(q, M \setminus \Omega_{k+1}) - \frac{1}{80} \text{dist}(q, M \setminus \Omega_{k+1}) \\ &= \frac{1}{80} \text{dist}(q, M \setminus \Omega_{k+1}) \\ &\geq \frac{1}{88} \text{dist}(x, M \setminus \Omega_{k+1}), \end{aligned}$$

by (6.303) and (6.289). Again we have that (6.299) holds, and so we assume instead that

$$(6.306) \quad H_{a,k+1}(y) > \frac{1}{80} \text{dist}(q, M \setminus \Omega_{k+1}).$$

As in (6.211), the maximal value of $H_{a,k+1}$ is given by

$$(6.307) \quad \begin{aligned} 1.5 \cdot \text{radius } B_{a,k+1} &= 1.5 \cdot 10^{-1} \text{dist}(a, M \setminus \Omega_{k+1}) \\ &\leq \frac{1}{5} \text{dist}(a, M \setminus \Omega_{k+1}), \end{aligned}$$

by the definition (6.153) of $H_{a,k+1}$. Thus (6.306) implies that

$$(6.308) \quad \text{dist}(a, M \setminus \Omega_{k+1}) > \frac{5}{80} \text{dist}(q, M \setminus \Omega_{k+1}) = \frac{1}{16} \text{dist}(q, M \setminus \Omega_{k+1}).$$

We want to use this to show that a and q must be L_1 -neighbors in A_{k+1} , at least if L_1 is large enough. This is very similar to what we did for Sublemma 6.202, but there are some small differences.

Let us get an upper bound for $\text{dist}(a, M \setminus \Omega_{k+1})$ in terms of $\text{dist}(q, M \setminus \Omega_{k+1})$. Because $H_{a,k+1}(y) \neq 0$ (by assumption), we have that

$$(6.309) \quad y \in 2B_{a,k+1},$$

by (6.155). Hence

$$(6.310) \quad \frac{4}{5} \text{dist}(a, M \setminus \Omega_{k+1}) \leq \text{dist}(y, M \setminus \Omega_{k+1}),$$

exactly as in (6.215)–(6.218). (This is also analogous to Lemma 6.35.) On the other hand (and this is slightly different from what we did before),

$$(6.311) \quad \begin{aligned} \text{dist}(y, M \setminus \Omega_{k+1}) &\leq d(x, y) + \text{dist}(x, M \setminus \Omega_{k+1}) \\ &\leq 41 C_0^2 \text{dist}(x, M \setminus \Omega_{k+1}), \end{aligned}$$

where the first step is completely general (as in (6.9)) while the second follows from (6.285). Thus

$$(6.312) \quad \begin{aligned} \text{dist}(a, M \setminus \Omega_{k+1}) &\leq 52 C_0^2 \text{dist}(x, M \setminus \Omega_{k+1}) \\ &\leq 58 C_0^2 \text{dist}(q, M \setminus \Omega_{k+1}), \end{aligned}$$

by (6.310), (6.311), and (6.289).

From (6.308) and (6.312) we get that a and q satisfy the first condition (6.159) for being L_1 -neighbors in A_{k+1} , at least if $L_1 \geq 58 C_0^2$. For the second condition (6.160) one uses the triangle inequality in practically the same manner as before. That is,

$$(6.313) \quad d(a, q) \leq d(a, y) + d(y, x) + d(x, q),$$

while $d(a, y) \leq \frac{1}{5} \text{dist}(a, M \setminus \Omega_{k+1})$ by (6.309), $d(x, y)$ is bounded in terms of $\text{dist}(q, M \setminus \Omega_{k+1})$ by (6.285) and (6.289), and $d(x, q) \leq 10^{-1} \text{dist}(q, M \setminus \Omega_{k+1})$ by (6.286). Using this and the previous estimates one can get the second requirement (6.160) for a and q to be L_1 -neighbors in A_{k+1} when L_1 is large enough, i.e., $d(a, q)$ is bounded by the minimum of $L_1 \text{dist}(a, M \setminus \Omega_{k+1})$ and $L_1 \text{dist}(q, M \setminus \Omega_{k+1})$. The size requirement for L_1 can be given explicitly in terms of the constant C_0 (from Lemma 6.142) and ordinary numbers (like the 58 above).

Thus we have shown that if (6.301) and (6.306) hold, and if L_1 is large enough, then a and q are L_1 -neighbors in A_{k+1} . Since we also have that $\Gamma(a, k+1) = t = \Gamma(q, k+1)$ (as mentioned just after (6.301)), the property (6.163) in Lemma 6.161 implies that $a = q$. This is incompatible with (6.309) and Sublemma 6.287, and so we conclude that this last scenario (in which (6.301) and (6.306) hold) is not possible (at least if L_1 is large enough).

The proof of Sublemma 6.298 is now complete, since we have shown that (6.299) is true under either of the conditions (6.300) or (6.301) together with (6.304), and that the remaining possibility of (6.301) and (6.306) simply does not occur (when L_1 is large enough).

From now on let us assume that L_1 is large enough for the purposes of Sublemma 6.298. Let us reformulate the conclusion (6.299) of Sublemma 6.298 as saying that

$$(6.314) \quad \beta_3 \geq \frac{1}{3520} C_0^{-2} \lambda^{\delta(k+1)} d(x, y).$$

This we can do because of (6.285). We conclude that

$$(6.315) \quad \beta_3 \geq C^{-1} D_\delta(x, y),$$

by (6.187). (As usual, this constant C depends only on suitable parameters, and not on x , y , or k in particular.)

It remains to estimate ϵ_3 , and this we can do in practically the same manner as in the earlier cases. Specifically, we first use the property (6.162) of our coding to convert (6.297) into

$$(6.316) \quad \epsilon_3 \leq \sum_{(a,j) \in \mathcal{C}: |j-(k+1)| \geq L_0} \lambda^{\delta j} |H_{a,j}(x) - H_{a,j}(y)|.$$

We then apply Lemma 6.131 to obtain

$$(6.317) \quad \epsilon_3 \leq C \lambda^{-\delta L_0} D_\delta(x, y).$$

This uses also the admissibility of the family $\{H_{a,j}\}_{a,j}$ provided by Lemma 6.152 for the applicability of Lemma 6.131.

If L_0 is large enough, depending on the usual parameters, then we have that

$$(6.318) \quad \beta_3 - \epsilon_3 \geq C^{-1} D_\delta(x, y),$$

where this constant C is 2 times the one in (6.315). This follows immediately from (6.315) and (6.317). From here we conclude that

$$(6.319) \quad |f(x) - f(y)| \geq C^{-1} D_\delta(x, y)$$

(with the same constant C as in (6.318)), because of (6.294) and (6.295). This gives (6.178) again in this case, which is what we wanted.

To summarize, we have shown that (6.178) holds in each of Cases I, II, III, and IV. These four cases cover all situations, as one can easily verify. Thus Lemma 6.177 is now completely proved, and Proposition 2.10 follows as well.

References

- [A1] P. ASSOUD, Espaces Métriques, Plongements, Facteurs, Thèse de Doctorat (January, 1977), Université de Paris XI, 91405 Orsay, France.
- [A2] P. ASSOUD, Étude d'une dimension métrique liée à la possibilité de plongement dans $\mathbf{R}^n$, *C. R. Acad. Sci. Paris* **288** (1979), 731–734.
- [A3] P. ASSOUD, Plongements Lipschitziens dans $\mathbf{R}^n$, *Bull. Soc. Math. France* **111** (1983), 429–448.
- [CM] R. COIFMAN AND Y. MEYER, Au-delà des opérateurs pseudo-différentiels, *Astérisque* **57**, Société Mathématique de France (1978).
- [CR] R. COIFMAN AND R. ROCHBERG, Another characterization of BMO, *Proc. Amer. Math. Soc.* **79** (1980), 249–254.
- [CW1] R. COIFMAN AND G. WEISS, “*Analyse Harmonique Non-commutative sur Certains Espaces Homogènes*,” Lecture Notes in Math. **242**, Springer-Verlag, 1971.
- [CW2] R. COIFMAN AND G. WEISS, Extensions of Hardy spaces and their use in analysis, *Bull. Amer. Math. Soc.* **83** (1977), 569–645.

- [DS1] G. DAVID AND S. SEMMES, Strong A_∞ -weights, Sobolev inequalities, and quasiconformal mappings, in “*Analysis and Partial Differential Equations*,” edited by C. Sadosky, Lecture Notes in Pure and Applied Mathematics **122**, Marcel Dekker, 1990, pp. 101–111.
- [DS2] G. DAVID AND S. SEMMES, Singular Integrals and Rectifiable Sets in $\mathbf{R}^n$: au-delà des graphes lipschitziens, *Astérisque* **193**, Société Mathématique de France (1991).
- [DS3] G. DAVID AND S. SEMMES, Quantitative rectifiability and Lipschitz mappings, *Trans. Amer. Math. Soc.* **337** (1993), 855–889.
- [DS4] G. DAVID AND S. SEMMES, “*Analysis of and on Uniformly Rectifiable Sets*,” Mathematical Surveys and Monographs **38**, 1993, American Mathematical Society.
- [DS5] G. DAVID AND S. SEMMES, “*Fractured Fractals and Broken Dreams: Self-Similar Geometry through Metric and Measure*,” Oxford Lecture Series in Mathematics and its Applications **7**, Oxford University Press, 1997.
- [Fe] H. FEDERER, “*Geometric Measure Theory*,” Springer-Verlag, 1969.
- [Ga] J. GARNETT, “*Bounded Analytic Functions*,” Academic Press, 1981.
- [Ge] F. GEHRING, The L^p integrability of the partial derivatives of a quasiconformal mapping, *Acta Math.* **130** (1973), 265–277.
- [J] J. L. JOURNÉ, “*Calderón-Zygmund Operators, Pseudodifferential Operators, and the Cauchy Integral of Calderón*,” Lecture Notes in Math. **994**, Springer-Verlag, 1983.
- [Pa] P. PANSU, Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un, *Ann. of Math.* **129** (1989), 1–60.
- [Se1] S. SEMMES, Bilipschitz mappings and strong A_∞ weights, *Ann. Acad. Sci. Fenn. Ser. A I Math.* **18** (1993), 211–248.
- [Se2] S. SEMMES, On the nonexistence of bilipschitz parameterizations and geometric problems about A_∞ weights, *Rev. Mat. Iberoamericana* **12** (1996), 337–410.
- [Se3] S. SEMMES, Metric Spaces and Mappings Seen at Many Scales, appendix, in “*Metric Structures in Riemannian and non-Riemannian Spaces*,” by M. Gromov et al, Birkhäuser.
- [St1] E. M. STEIN, “*Singular Integrals and Differentiability Properties of Functions*,” Princeton University Press, 1970.
- [St2] E. M. STEIN, “*Harmonic Analysis : Real-Variable Methods, Orthogonality, and Oscillatory Integrals*,” Princeton University Press, 1993.

- [TV] P. TUKIA AND J. VÄISÄLÄ, Quasisymmetric embeddings of metric spaces, *Ann. Acad. Sci. Fenn. Ser. A I Math.* **5** (1980), 97–114.

Department of Mathematics
Rice University
Box 1892
Houston TX 77251-1892
U.S.A.

e-mail: semmes@math.rice.edu

Primera versió rebuda el 2 de setembre de 1998,
darrera versió rebuda el 15 de març de 1999