Revista Matemática Iberoamericana
Vol. 16, N.º 2, 2000

Measure-preserving quality within mappings

Stephen Semmes

Abstract. In [6], Guy David introduced some methods for finding controlled behavior in Lipschitz mappings with substantial images (in terms of measure). Under suitable conditions, David produces subsets on which the given mapping is bilipschitz, with uniform bounds for the bilipschitz constant and the size of the subset. This has applications for boundedness of singular integral operators and "uniform rectifiability" of sets, as in [6], [7], [11], [13]. Some special cases of David's results, concerning projections of subsets of Euclidean spaces of codimension 1, or mappings defined on Euclidean spaces (rather than sets or metric spaces of less simple nature), have been given alternate and much simpler proofs, as in [8], [19], [10]. In general this has not occurred.

Here we shall present a variation on David's method which breaks down into simpler pieces. We shall also take advantage of some components of the work of Peter Jones [19]. Jones' approach uses some Littlewood-Paley theory, and one of the important features of David's method was to avoid this, operating in a more directly geometric way which could be applied more broadly. To some extent, the present organization gives a reconciliation between the two, and between David's stopping-time argument and techniques related to Carleson measures and Carleson's Corona construction.

1. Introduction.

Let $(M, d(x, y))$ and $(N, \rho(u, v))$ be metric spaces. Thus M is a nonempty set, $d(x, y)$ is a symmetric nonnegative function on $M \times M$ that vanishes exactly when $x = y$ and satisfies the triangle inequality, and similarly for $(N, \rho(u, v))$. If E is a subset of M and $f : E \longrightarrow N$ is a mapping, then f is said to be *Lipschitz* if there is a constant C such that

$$\rho(f(x), f(y)) \le C\, d(x, y)\,, \qquad \text{for all } x, y \in E\,. \tag{1.1}$$

Similarly, f is *bilipschitz* if there is a constant C so that

$$C^{-1} d(x, y) \le \rho(f(x), f(y)) \le C\, d(x, y)\,, \qquad \text{for all } x, y \in E\,. \tag{1.2}$$

Lipschitz and bilipschitz mappings provide basic ways of making comparisons between metric spaces. In particular, two metric spaces are practically the same for many purposes when they are bilipschitz equivalent (*i.e.*, when there is a bilipschitz mapping from one onto the other).

On the other hand, bilipschitz parameterizations can also be hard to come by, even in situations in which it might appear as though they ought to exist. This is discussed further in [22], [23], [24], [26].

In practice it is often much easier to find "pieces" of bilipschitz equivalence, rather than whole parameterizations. That is, one might be able to find bilipschitz mappings between sets of significant size in terms of measure, if not between sets which are open, or large enough to contain a given ball, etc. Some basic tools for doing this are given in [6], [19]. Here we shall describe a kind of reorganization of these tools, which work in the same contexts as David's method in [6], but do so in a more relaxed way, and with the same type of (slightly stronger) conclusions as in [19].

An important simplification that comes from Jones' method [19] is that, instead of looking directly for bilipschitz pieces, it is enough to find approximate bilipschitz behavior at many locations and scales. This permits one to have "gaps" in the information about bilipschitz behavior, gaps that are sorted out at the end through a coding argument (from [19]), and it enables one to concentrate on estimates which are more local (and much simpler) than the ones in [6]. This part of the story is reviewed in Section 8, in terms of "weak bilipschitz" conditions, as in [10]. (This is really just a reformulation of part of Jones' argument from [19].)

The basic starting point in both [6] and [19] is a Lipschitz mapping whose image is reasonably large in the sense of measure. The bilipschitz pieces are then extracted from the given mapping, at least under some conditions. In [19] the given Lipschitz mapping f is defined on a cube in a Euclidean space, for instance, and one is able to use Littlewood-Paley theory to say that f is well-approximated by affine mappings at "most" scales and locations, with uniform bounds. Here "most" can be made precise through the use of Carleson measures. Affine mappings (with bounded gradient) are very nice to work with, because they are always either bilipschitz in a controlled way or quite degenerate, and one can account for the degeneracies through the behavior of f in terms of measure (as in [19]).

David's approach is more complicated. A basic feature of his argument is to be able to focus on ranges of scales and locations where the mapping f is almost measure-preserving. This provides a degree of rigidity which is not as directly potent as affine approximations would be, but which is quite useful nonetheless, and more readily accessible in a broader range of situations. (When affine approximations are available, Jones' method can be extended easily, as in [10].)

Here we shall take the issues of approximate measure-preserving behavior and treat them separately from the rest, in sections 3-7. In the end we shall only need a kind of "weak" condition, in the spirit of the "weak geometric lemma" (WGL) and other "weak" conditions, as discussed in [11]. This is analogous to the kind of "weak" conditions of affine approximate used in Jones' argument. (See also [10].) Stronger conditions are also available, in the spirit of Carleson's Corona construction (just as for approximation of Lipschitz functions by affine functions, or for the approximation of uniformly-rectifiable sets by flat pieces).

The passage from near-preservation of measure to bilipschitz conditions does not work abstractly, but requires extra information. In this we shall follow [6], for an auxiliary condition which is sufficient. Note, however, that the extraction of pieces which are almost measure-preserving is extremely general. It came up in slightly different ways in [12], *e.g.*, in the extraction of measure-preserving weak tangents, as in [12, Proposition 12.42]. (To some extent, [12] is exactly about the kind of structure and rigidities one can get when bilipschitz behavior is not available. Part of the point of the present paper is to make a better fit between [6], [19] and other contexts, like the ones in [12].) One could also use the construction in [6] to extract measure-preserving behavior,

simply dropping the issue of significant bilipschitz pieces (and one of the steps in the stopping-time argument), as well as the extra condition needed to get the bilipschitz pieces. This would give a bit more information than we shall derive or use here. (See Remark 6.14.)

David's condition is reviewed and used in Section 9. The basic result, a kind of conglomeration of the theorems of David and Jones, is given in Section 10. Section 2 covers some background information and basic lemmas, and Section 11 mentions some modest refinements of the material in this paper which are useful in certain contexts.

We should emphasize that many of the steps here have clear counterparts in [6]. (Otherwise they are close to [19], or to standard ideas related to Carleson measures or Carleson's Corona construction.) In some cases we perhaps gain some advantage in needing only relatively "local" computations or arguments. At any rate, it is also pleasant to have a better reconciliation between the methods of [6] and [19].

2. Some background information.

2.1. Hausdorff measure.

Let us begin by recalling the definition of Hausdorff measure. Let $(M, d(x, y))$ be a metric space, and fix a positive number n. Note that n need not be an integer for this discussion, although integer values will be of particular importance in this paper. Given $\delta > 0$ and a subset E of M, define $H^n_\delta(E)$ by

$$H^n_\delta(E) = \inf \Big\{ \sum_j (\operatorname{diam} A_j)^n : \{A_j\} \text{ is a sequence of sets in } M \text{ which covers } E \text{ and satisfies } \operatorname{diam} A_j < \delta \text{ for all } j \Big\}. \tag{2.1}$$

It is easy to see that $H^n_\delta(E)$ can only become larger as δ gets smaller, so that the limit

$$H^n(E) = \lim_{\delta \to 0} H^n_\delta(E) \tag{2.2}$$

exists (but may be infinite). This is the n-dimensional Hausdorff measure of E. (See [15], [16], [20] for more information.)

A simple and basic fact about Hausdorff measure is that if $(M, d(x,y))$ and $(N, \rho(u,v))$ are metric spaces and $f : M \longrightarrow N$ is Lipschitz with constant K, then

$$H^n(f(E)) \le K^n H^n(E)\,. \tag{2.3}$$

This follows easily from the definitions. (Note that the H^n on the left side of (2.3) is defined using the metric on N, while the H^n on the right is defined on M.)

We shall often write $|E|$ for $H^n(E)$, for simplicity.

It is well-known that Borel sets and sets of H^n-measure 0 are "measurable" for H^n, so that one has the usual additivity properties for these sets (beyond the subadditivity which works for arbitrary sets). See [15], [16], [20], for instance. Let us mention one other technical fact related to measurability, which will sometimes be needed in this paper.

Lemma 2.4. *Let M and N be metric spaces, and assume that M is a countable union of compact sets with finite H^n-measure. Suppose that D is a subset of M which is H^n-measurable, and let f be a Lipschitz mapping from D into N. Then $f(D)$ is H^n-measurable.*

This is pretty standard, but we include a proof for the sake of completeness. One could weaken the hypotheses a bit, but we shall not bother with this. (For the main purposes of this paper, one might as well make the stronger assumption that M be "Ahlfors-regular", as defined in Section 2.2.)

To prove Lemma 2.4, it suffices to show the following.

Claim 2.5. *Under the assumptions of the lemma, there is a sequence of compact sets $\{K_j\}$ such that each K_j is contained in D and*

$$H^n\Big(D \backslash \bigcup_j K_j\Big) = 0\,. \tag{2.6}$$

If Claim 2.5 is true, then we can write $f(D)$ as

$$f(D) = \Big(\bigcup_j f(K_j)\Big) \cup f\Big(D \backslash \bigcup_j K_j\Big)\,, \tag{2.7}$$

with the K_j's as above. Each $f(K_j)$ is a compact subset of N, since the K_j's are compact subsets of M and f is continuous, and thus each $f(D_j)$

is H^n-measurable in N. The last piece, $f(D \setminus \bigcup_j K_j)$, has measure 0, because of (2.6) and (2.3). This implies that $f(D)$ is measurable, which is what we want for Lemma 2.4.

It remains to verify Claim 2.5. We may as well assume that M is compact and of finite H^n-measure, because if we can prove Claim 2.5 in this restricted situation, then the general case follows automatically. This uses the assumption from Lemma 2.4 that M be a countable union of compact sets of finite measure. (In other words, if M is the union of a countable family $\{J_\ell\}$ of compact sets of finite measure, and if we know (from the restricted version of Claim 2.5) that the intersection of D with each J_ℓ can be realized as the countable union of compact subsets together with a set of measure 0, then the same follows for D itself by taking the (countable) union over ℓ.)

Thus we assume now that M is compact and has finite measure. Set $E = M \setminus D$. We want to show that there is a subset G of M such that $G \supseteq E$, G is a countable intersection of open sets, and

$$H^n(G) = H^n(E)\,. \tag{2.8}$$

If we can do this, then we are finished, because

$$D = M \setminus E = (M \setminus G) \cup (G \setminus E)\,, \tag{2.9}$$

where $M \setminus G$ is a countable union of compact sets (since M is compact and G is a countable intersection of open sets), while $G \setminus E$ has measure 0, by (2.8). For this last we also use the fact that E is contained in G and is measurable (since D is, by the hypotheses of Lemma 2.4), and that the measure of G is finite (because M has finite measure).

To produce G, we use the following standard argument. (See [20] for more information and results along these lines.) Let $\delta > 0$ be arbitrary, and let $\{A_j\}$ be a sequence of subsets of M which covers E and satisfies $\operatorname{diam} A_j < \delta$ for all j and

$$\sum_j (\operatorname{diam} A_j)^n < H^n_\delta(E) + \delta\,, \tag{2.10}$$

as in (2.1). We may assume that the A_j's are all open sets, because we can expand each of them by a tiny amount to make this true, and without disturbing the inequalities above. Let $U(\delta)$ be the open set which is the union of the A_j's. Notice that $U(\delta) \supseteq E$, and that

$$H^n_\delta(U(\delta)) \le \sum_j (\operatorname{diam} A_j)^n < H^n_\delta(E) + \delta\,. \tag{2.11}$$

The first inequality comes from (2.1) and the fact that $U(\delta)$ is covered by the A_j's, and the second is just (2.10).

We can do this for every $\delta > 0$, and then put

$$G = \bigcap_{\ell=1}^{\infty} U\Big(\frac{1}{\ell}\Big)\,. \tag{2.12}$$

Thus G is a countable intersection of open sets by construction, and G contains E as a subset because each $U(\delta)$ does. It remains to check that (2.8) holds. Of course $H^n(E) \le H^n(G)$ automatically, since $G \supseteq E$. For each $\delta = 1/\ell$ we have that

$$H^n_\delta(G) \le H^n_\delta(U(\delta)) < H^n_\delta(E) + \delta\,, \tag{2.13}$$

by (2.11). This is enough to ensure that $H^n(G) \le H^n(E)$, by the definition (2.2) of H^n (and because the limit in δ in (2.2) always exists).

This completes the proof of Claim 2.5, and hence of Lemma 2.4 too.

2.2. Ahlfors regularity.

A metric space $(M, d(x,y))$ is said to be *Ahlfors regular of dimension* n, $0 < n < \infty$, if it is complete, and if there is a constant C such that

$$C^{-1}\, r^n \le H^n(\overline{B}(x,r)) \le C\, r^n\,, \tag{2.14}$$

for all $x \in M$ and $0 < r \le \operatorname{diam} M$. Here $\overline{B}(x,r)$ denotes the closed ball with center x and radius r in M; sometimes we may write $\overline{B}_M(x,r)$ to emphasize the metric space in question. We write $B(x,r)$ for the open ball with center x and radius r, and in general "ball" will be used to mean "open ball". Let us make the standing assumption that Ahlfors-regular metric spaces be nondegenerate in the sense of having positive diameter.

If $(M, d(x,y))$ is the same as $\mathbb{R}^n$ with the standard metric, for instance, then H^n is a constant multiple of Lebesgue measure, and (2.14) holds because $H^n(\overline{B}(x,r))$ is simply the product of r^n with a positive constant that depends only on n (namely, the volume of the unit ball).

Ahlfors-regular metric spaces are automatically *doubling*, which means that there is a constant L so that every ball in the metric space can be covered by at most L balls of half the radius. This is well-known, and not very difficult to prove. As a consequence, closed and bounded subsets of an Ahlfors-regular metric space are always compact. Indeed, the doubling property implies that every bounded set is totally bounded – covered by a finite number of balls of arbitrarily small radius – and then compactness follows from this and completeness, by a standard characterization of compact sets in metric spaces.

In particular, if M is Ahlfors-regular of dimension n, then M is a countable union of compact sets of finite H^n-measure, and therefore satisfies the hypothesis of Lemma 2.4.

2.3. Cubes.

A nice feature of Euclidean spaces is the existence of standard partitions into dyadic cubes. Ahlfors-regular metric spaces admit similar partitions, of the following nature.

Fix $(M, d(x, y))$, an n-dimensional Ahlfors-regular metric space. Set $j_0 = \infty$ if M is unbounded, and otherwise choose j_0 to be the integer such that

$$2^{j_0} \le \operatorname{diam} M < 2^{j_0+1}\,. \tag{2.15}$$

Instead of ordinary dyadic cubes we shall be interested in having a family $\{\Delta_j\}_{j<j_0}$, $j \in \mathbb{Z}$, of measurable subsets of M, with the following properties:

$$\begin{aligned} &\text{each } \Delta_j \text{ is a partition of } M, \textit{ i.e.}, M = \bigcup_{Q\in\Delta_j} Q \text{ for any } j, \\ &\text{and } Q \cap Q' = \varnothing \text{ whenever } Q, Q' \in \Delta_j \text{ and } Q \neq Q'\,, \end{aligned} \tag{2.16}$$

$$\begin{aligned} &\text{if } Q \in \Delta_j \text{ and } Q' \in \Delta_k \text{ for some } k \ge j, \\ &\text{then either } Q \subseteq Q' \text{ or } Q \cap Q' = \varnothing\,, \end{aligned} \tag{2.17}$$

$$\begin{aligned} &C^{-1}\, 2^j \le \operatorname{diam} Q \le C\, 2^j \text{ and } C^{-1}\, 2^{jn} \le |Q| \le C\, 2^{jn}, \\ &\text{for all } j \text{ and all } Q \in \Delta_j\,, \end{aligned} \tag{2.18}$$

given any j, $Q \in \Delta_j$, and $\tau > 0$, we have that

$$|\{x \in Q : \ \mathrm{dist}\,(x, M \backslash Q) \le \tau\, 2^j\}| \le C\, \tau^{1/C}\, |Q|\,, \quad \text{and} \tag{2.19}$$

$$|\{x \in M \backslash Q : \ \mathrm{dist}\,(x, Q) \le \tau\, 2^j\}| \le C\, \tau^{1/C}\, |Q|\,.$$

This last condition says that the boundary of $Q \in \Delta_j$ is reasonably tame, just as for ordinary dyadic cubes in a Euclidean space. Note that all of (2.16)-(2.19) are valid for the standard partitions of $\mathbb{R}^n$ into dyadic cubes, *i.e.*, with Δ_j taken to be the collection of dyadic cubes of sidelength 2^j. (For ordinary cubes one can take the exponent of τ in (2.19) to be equal to 1.)

If M is actually a subset of some $\mathbb{R}^N$, equipped with the ordinary Euclidean distance, then the existence of such a family $\{\Delta_j\}_{j<j_0}$ has been established by David [6], [7]. For this the formulation in [7] is closer to the present discussion. Note that in [7] the set M was assumed to be unbounded. This was not a serious requirement, and one can reduce to that case anyway by adding to M an unbounded Ahlfors-regular set E such that $\mathrm{dist}\,(E, M)$ is approximately equal to the diameter of M. (When n is an integer, for instance, one can take E to be a n-plane, but all of this works when n is not an integer too.)

The existence of $\{\Delta_j\}_{j<j_0}$ for a general Ahlfors-regular metric space can be derived from the special case of subsets of Euclidean spaces, as follows. Given $s \in (0, 1)$, consider M with the metric $d(x, y)^s$. This is indeed a metric, satisfying the triangle inequality in particular, as is well-known (and not hard to verify). A result of Assouad [1], [2], [3] implies that $(M, d(x, y)^s)$ is bilipschitz equivalent to a subset of some $\mathbb{R}^N$ for each $s \in (0, 1)$, with the dimension N depending on s. More precisely, Assouad shows that such embeddings exist as soon as $(M, d(x, y))$ is doubling, and the doubling property holds automatically when M is Ahlfors regular, as mentioned earlier.

Since $(M, d(x, y)^s)$ is bilipschitz equivalent to a subset of some $\mathbb{R}^N$, the existence of a family $\{\Delta_j\}_{j<j_0}$ for it follows from the construction of David. In other words, the properties (2.16)-(2.19) are not disturbed by bilipschitz equivalence, except for changing the constant C. This also uses the fact that $(M, d(x, y)^s)$ is Ahlfors-regular in its own right, with dimension n/s, as one can check. (Note that n-dimensional Hausdorff measure for $(M, d(x, y))$ is exactly the same as n/s-dimensional Hausdorff measure for $(M, d(x, y)^s)$, by definitions. Thus the measures stay the same, and the subsets of M which are balls remain the same. It is only the radii of the balls which change, but this washes out for the Ahlfors-regularity property.)

Given such a family $\{\Delta_j\}_{j<j_0}$ for $(M, d(x,y)^s)$, it is not hard to make minor modifications to get a family that works for $(M, d(x,y))$ itself. One has to adjust the parameters slightly, to compensate for the different dimensions of Ahlfors regularity and different measurements of diameter, but this does not create significant difficulties. (Keep in mind that the measures for $(M, d(x,y)^s)$ and $(M, d(x,y))$ remain the same, even if they are given slightly different names.)

In short, an Ahlfors-regular metric space $(M, d(x,y))$ always admits a family of partitions $\{\Delta_j\}_{j<j_0}$ as above. The constant C in (2.18) and (2.19) can be chosen so that it depends only on the dimension n and the Ahlfors-regularity constant for M.

Normally, when we have an Ahlfors-regular metric space $(M, d(x,y))$ in hand, we shall fix a family of partitions $\{\Delta_j\}_{j<j_0}$ as above, and we shall refer to the elements of the Δ_j's as *cubes*. (One might also say *pseudocubes*, to avoid confusion with ordinary cubes in Euclidean spaces.) We shall also typically set

$$\Delta = \bigcup_{j<j_0} \Delta_j \ . \tag{2.20}$$

This is like looking at the totality of all dyadic cubes, rather than just ones of a fixed size.

Note that a single cube Q may lie in Δ_j for more than one choice of j, *i.e.*, there is nothing in the conditions above to prevent this. This could not happen for more than a bounded number of consecutive j's, because of (2.18). In any case, this will not cause any serious difficulties.

Given cubes Q and Q' in Δ, we shall refer to Q' as a *child* of Q if there is an integer $j < j_0$ such that $Q \in \Delta_j$, $Q' \in \Delta_{j-1}$, and $Q' \subseteq Q$. We shall also call Q the *parent* of Q'.

If $Q, Q' \in \Delta$, then either Q and Q' are disjoint from each other, or one contains the other. This follows from (2.17).

2.4. Stopping-time regions.

Let $(M, d(x,y))$ be a n-dimensional Ahlfors-regular metric space, and let $\{\Delta_j\}_{j<j_0}$, Δ be as in Section 2.3. A collection S of cubes in Δ will be called a *stopping-time region* if it is nonempty and satisfies the

following conditions:

(2.21) S contains a *top cube* $Q(S)$ that contains all other cubes in S as subsets,

(2.22) if Q, Q' are elements of Δ such that $Q \in S$ and $Q \subseteq Q' \subseteq Q(S)$, then $Q' \in S$ too.

In practice, stopping-time regions come about from stopping-time arguments, of roughly the following form. One has some property P of cubes that one is interested in, and one starts with a cube Q_0 which enjoys this property. To define S, one first takes Q_0 as the top cube $Q(S)$. For each of the children of Q_0, one asks whether it has the property P as well. If so, then one puts that child into S and repeats the process for each of its children. If not, then one stops, and does not worry about the progeny of that cube. In the end, one gets a collection S which satisfies (2.21) and (2.22), and every cube Q in S satisfies the property P under consideration. Also, if a cube Q is contained in Q_0, and if Q does not lie in S but the parent of Q does, then Q does not satisfy P, by construction.

Here is a simple example of a stopping-time region. Fix a cube $Q_0 \in \Delta$, and also a nonempty subset E of Q_0. Put

$$\Delta(Q_0, E) = \{Q \in \Delta : \ Q \subseteq Q_0 \text{ and } Q_0 \cap E \neq \varnothing\} . \tag{2.23}$$

This clearly satisfies (2.21) and (2.22). In general the cubes in a stopping-time region may not "go all the way to the bottom", down to individual points, as they do in (2.23).

Often one is concerned not just with individual stopping-time regions, but with disjoint families of them, obtained by repeating the same kind of construction as above. That is, after having to "stop" in the construction of one stopping-time region S, one starts all over again, perhaps with an adjustment in the property P under consideration. This will come up in sections 3 and 4.

Sometimes one derives decompositions of all of Δ, or nearly all of Δ, into stopping-time regions. A basic paradigm for this is provided by Carleson's Corona construction, as in [17, Chapter VIII]. An important point then is to know that one does not have to "stop" too often, and this is often made precise by the notions of Carleson measures and Carleson sets. We shall discuss the latter in the next section.

2.5. Packing conditions and Carleson sets.

Let $(M, d(x,y))$, $\{\Delta_j\}_{j<j_0}$, and Δ be as before. Fix a cube Q_0 in Δ, and imagine that we have a collection $\mathcal{E}$ of cubes in Δ. We shall be interested in packing conditions of the form

$$\sum_{\substack{R \in \mathcal{E} \\ R \subseteq Q_0}} |R| \le C\, |Q_0| \,. \tag{2.24}$$

Here C is some constant (over which one would hope to have some control), and $|R|$ denotes the measure of the cube R (with respect to n-dimensional Hausdorff measure).

Definition 2.25 (Carleson sets)**.** *If there is a constant C so that* (2.24) *holds for all $Q_0 \in \Delta$, then $\mathcal{E}$ is called a Carleson set.*

Note that finite unions of Carleson sets are automatically Carleson sets. For the record, let us mention also the following simple observation.

Lemma 2.26. *If the elements of $\mathcal{E}$ are pairwise disjoint, then $\mathcal{E}$ is a Carleson set with constant equal to* 1*.*

This is easy to verify.

It is helpful to reformulate (2.24) as follows. With Q_0 and $\mathcal{E}$ as above, let $N(x) = N_{Q_0}(x)$ be equal to the number of cubes $R \in \mathcal{E}$ such that $x \in R$ and $R \subseteq Q_0$. It is not hard to see that this function is measurable, and that

$$\sum_{\substack{R \in \mathcal{E} \\ R \subseteq Q_0}} |R| = \int_{Q_0} N(x)\, dx \,, \tag{2.27}$$

by Fubini's theorem. Here dx denotes n-dimensional Hausdorff measure on M.

With this identity, we see that (2.24) is equivalent to asking that the average of $N(x)$ over Q_0 be bounded by C. In particular, (2.24) is automatic when $N(x)$ is pointwise bounded by C on Q_0, and it implies that $N(x)$ is bounded by $2\,C$ on at least half of the points in Q_0. For Carleson sets one can take this further, as follows.

Lemma 2.28. *Let $\mathcal{E}$ be a family of cubes in Δ, and suppose that there are positive constants K, η such that*

$$|\{x \in Q_0 : \ N_{Q_0}(x) \le K\}| \ge \eta\, |Q_0|\,, \tag{2.29}$$

for every cube $Q_0 \in \Delta$. Then $\mathcal{E}$ is a Carleson set, with constants which depend only on K and η.

This is a version of well-known results of John, Nirenberg, and Strömberg in the context of BMO functions, as in [17]. The proof shows not only that the average of N_{Q_0} on Q_0 is bounded, but also gives exponential-integrability as well, as in the original John-Nirenberg theorem.

Lemma 2.28 is quite standard, and given in [11, Part IV, Lemma 1.12]. Let us briefly review the main steps in the proof. Let Q_0 be any cube in Δ, and let E_1 denote the set of points $x \in Q_0$ such that $N_{Q_0}(x) \le K$. Thus $|Q_0 \backslash E_1| \le (1 - \eta)\, |Q_0|$, by (2.29). Using the definition of E_1, it is not hard to see that $Q_0 \backslash E_1$ is a union of subcubes of Q_0. (That is, if $N_{Q_0}(y) > K$ for some point y, then there is a cube which contains y so that $N_{Q_0} > K$ for every point in the cube.) One can realize $Q_0 \backslash E_1$ as a union of *maximal* cubes (maximal among ones contained in $Q_0 \backslash E_1$), and these are automatically disjoint (by (2.17)). If Q_1 is one of these maximal cubes, then we can repeat the argument and take $E(Q_1)$ to be the set of points $x \in Q_1$ such that $N_{Q_1}(x) \le K$. When $x \in E(Q_1)$, we also have that $N_{Q_0}(x) \le 2\, K$; indeed, there are at most K cubes Q in $\mathcal{E}$ which contain x and lie in Q_1, since $N_{Q_1}(x) \le K$, and there are at most K cubes Q in $\mathcal{E}$ which contain Q_1 as a proper subset, are contained in Q_0, and lie in $\mathcal{E}$. This follows from the *maximality* of Q_1 as a cube contained in $Q_0 \backslash E_1$, *i.e.*, the next larger cube containing Q_1 must also contain an element of E_1.

Let E_2 denote the union of E_1 with the sets $E(Q_1)$, where Q_1 ranges over the maximal cubes contained in $Q_0 \backslash E_1$. Then $N_{Q_0} \le 2\, K$ on E_2, and $|Q_0 \backslash E_2| \le (1 - \eta)^2\, |Q_0|$. This last comes from $|Q_0 \backslash E_1| \le (1 - \eta)\, |Q_0|$ and its analogue $|Q_1 \backslash E(Q_1)| \le (1 - \eta)\, |Q_1|$ for the maximal cubes Q_1.

Again $Q_0 \backslash E_2$ is a union of cubes, and a union of (maximal) subcubes of the maximal cubes Q_1, as before. This permits one to repeat the argument. In general one obtains for each positive integer j a subset E_j of Q_0 such that $Q_0 \backslash E_j$ is a union of cubes, $N_{Q_0} \le j K$ on E_j, each of the constituent cubes in $Q_0 \backslash E_j$ is properly contained in at most $j K$ cubes in $\mathcal{E}$ which are subcubes of Q_0, and $|Q_0 \backslash E_j| \le (1 - \eta)^j\, |Q_0|$. The passage from j to $j + 1$ is very much like the argument above.

Thus we have exponential decay for the distribution function of N_{Q_0}, and in particular the finiteness of the average of N_{Q_0} over Q_0, with uniform bounds. This gives Lemma 2.28.

Keep in mind that there are always infinitely many cubes $R \in \Delta$ which contain a given point x, about one for each "scale". The packing and Carleson conditions imply a bound on average for the number of scales involved in $\mathcal{E}$ above a given x, and hence a precise sense in which $\mathcal{E}$ is "small" as a subset of Δ. However, the particular choice of scales may vary from point to point, and with little control or pattern.

A basic scenario which comes up in this paper is to have a family $\mathcal{F}$ of stopping-time regions in Δ, and for $\mathcal{E}$ to be the collection of their top cubes. The packing and Carleson conditions then have the effect of saying that, on average, one did not have to stop more than a bounded number of times in the stopping-time argument which produced $\mathcal{F}$. This is exactly what happens in Carleson's Corona construction. (See also [11, Part I, Section 3.2], especially the notion of a "coronization".)

2.6. Some lemmas.

There is a simple "stability" property of the packing and Carleson conditions that we should record. Fix an (arbitrary) number $A > 1$, and let us say that two cubes $Q, Q' \in \Delta$ are *neighbors* (with constant A) if

$$\text{dist}\,(Q, Q') \le A\,(\text{diam}\,Q + \text{diam}\,Q') \tag{2.30}$$

and

$$A^{-1}\text{diam}\,Q \le \text{diam}\,Q' \le A\,\text{diam}\,Q\,. \tag{2.31}$$

If one thinks of Δ as providing a discrete model for the "upper half-space" $M \times (0, \text{diam}\,M)$, then (2.30) and (2.31) correspond to the idea that Q and Q' lay at bounded distance from each other in a (quasi-) hyperbolic distance.

Suppose that we are given $\mathcal{E}$, Q_0 as above, and let $\mathcal{E}_A$ denote the set of all cubes $Q \in \Delta$ such that Q is a neighbor (in the sense of (2.30) and (2.31)) of a cube $Q' \in \mathcal{E}$.

Lemma 2.32. *Notation and assumptions as above. If $\mathcal{E}$ satisfies the packing condition* (2.24) *with the constant C, and if every cube in $\mathcal{E}$ is a*

subcube of the fixed cube Q_0, then $\mathcal{E}_A$ satisfies (2.24) *with C increased by a factor that depends only on A, n, and the Ahlfors regularity constant for M.*

If $\mathcal{E}$ is a Carleson set (*and without restriction on where the cubes in $\mathcal{E}$ might lie*), *then $\mathcal{E}_A$ is too, and with the same type of bound for the Carleson constant.*

This is not hard to check, and we omit the details. (Remember that our partitions Δ_j are chosen so that the constants in (2.18) and (2.19) depend only on n and the Ahlfors regularity constant for M.)

Although the passage from $\mathcal{E}$ to $\mathcal{E}_A$ preserves the packing condition (2.24) to within a bounded factor, this is not always true for the stronger requirement that multiplicity function $N(x)$ be bounded. For instance, the cubes in $\mathcal{E}$ might be pairwise disjoint, so that $N(x)$ is uniformly bounded by 1, but it could also happen (at the same time) that $\mathcal{E}_A$ contains all cubes Q which contain some fixed point $z \in M$. This is not difficult to arrange, and it leads to logarithmic blow-up for the analogue of $N(x)$ for $\mathcal{E}_A$ around z.

In many situations, packing or Carleson conditions are used to say that some "bad" or inconvenient event does not occur too often. Lemma 2.32 permits one to automatically extend this, to say that even being remotely close to a bad event does not occur too often. This can be very convenient for making proofs, in that one is free to take A to be as large as one wants, with the "price" for doing this not coming until the very end. This is because one often does not care what the bad sets look like, as long as they are suitably controlled by packing or Carleson conditions. (These were recurring themes in [11].)

On the other hand, stopping-time regions in Δ are often used to represent ranges of cubes where something good happens. In practice, one may wish to avoid the boundaries of stopping-time regions, and our next task is to provide some lemmas which facilitate this.

We begin with the following simplified situation. Fix a cube $T \in \Delta$, and define $\mathcal{C}_A(T)$ by

$$\begin{aligned} \mathcal{C}_A(T) = \{Q \in \Delta : \ & \text{there is a cube } Q' \in \Delta \text{ such that} \\ & Q \text{ and } Q' \text{ are neighbors , and either} \\ & Q \subseteq T,\ Q' \not\subseteq T, \text{ or } Q' \subseteq T,\ Q \not\subseteq T\}\,. \end{aligned} \tag{2.33}$$

Lemma 2.34. *There is a constant D so that*

$$\sum_{Q \in \mathcal{C}_A(T)} |Q| \le D\,|T|\,, \tag{2.35}$$

where D depends only on n, A, and the Ahlfors-regularity constant for M.

To prove this, we shall use the small boundary property (2.19) of the cubes in Δ. Fix $j \in \mathbb{Z}$ so that $T \in \Delta_j$.

Notice first that

$$\operatorname{diam} Q \le A \operatorname{diam} T\,, \qquad \text{when } Q \in \mathcal{C}_A(T)\,. \tag{2.36}$$

This is not hard to derive, from the definitions. From this it follows that there are only boundedly many cubes $Q \in \mathcal{C}_A(T)$ such that $Q \in \Delta_\ell$ with $\ell > j$. The contribution of these cubes to (2.35) is bounded by a constant times $|T|$, because of (2.18), and so we may forget about them. Thus we set

$$\mathcal{C}_A^*(T) = \{Q \in \mathcal{C}_A(T) :\ Q \in \Delta_k \text{ for some } k \le j\}\,, \tag{2.37}$$

and it suffices to check (2.35) for $\mathcal{C}_A^*(T)$ instead of $\mathcal{C}_A(T)$.

Let Q_1 be another cube in Δ_j, and let us show that

$$\sum_{\substack{R \in \mathcal{C}_A^*(T) \\ R \subseteq Q_1}} |R| \le C\,|Q_1|\,, \tag{2.38}$$

for a suitable constant C. If we can do this, then (2.35) will follow. This is because any cube R in $\mathcal{C}_A^*(T)$ must lie either in T itself or a cube $Q_1 \in \Delta_j$ which is not too far from T (by (2.16), (2.17), and (2.37)), and there are only boundedly many such cubes Q_1. Notice also that $|Q_1|$ is bounded by a constant multiple of $|T|$, by (2.18).

Fix $Q_1 \in \Delta_j$. Because of (2.17), Q_1 must either be equal to T, or disjoint from T.

Given $k \le j$, let $\mathcal{T}_k$ denote the set of cubes $R \in \Delta_k$ such that $R \in \mathcal{C}_A^*(T)$ and $R \subseteq Q_1$. Thus

$$\bigcup_{k \le j} \mathcal{T}_k = \{R \in \mathcal{C}_A^*(T) :\ R \subseteq Q_1\}\,, \tag{2.39}$$

by (2.37). The elements of $\mathcal{T}_k$ are pairwise disjoint, because of (2.16), and so

$$\sum_{R \in \mathcal{T}_k} |R| = \Big| \bigcup_{R \in \mathcal{T}_k} R \Big| . \tag{2.40}$$

On the other hand, if $R \in \mathcal{T}_k$, then there is a cube $R' \in \Delta$ such that R and R' are neighbors and either R is contained in T and R' is not, or the other way around. (See (2.33).) In both cases we have that $R' \not\subseteq Q_1$, because $R \subseteq Q_1$, and because Q_1 is itself either equal to T or disjoint from T. (If $Q_1 = T$, then $R \subseteq T$ and $R' \not\subseteq T = Q_1$; if Q_1 is disjoint from T, then $R \not\subseteq T$, $R' \subseteq T$, and so $R' \not\subseteq Q_1$ because Q_1 and T are disjoint from each other.) Since $R' \not\subseteq Q_1$, we must either have that R' contains Q_1, or is disjoint from Q_1. This can be derived from (2.17). Each of these possibilities implies that the distance from R to the complement of Q_1 is bounded by a constant (depending on A) times the diameter of R, because R and R' are neighbors.

The diameter of R is bounded by a constant times 2^k, by (2.18), and therefore

$$\bigcup_{R \in \mathcal{T}_k} R \subseteq \{x \in Q_1 : \ \mathrm{dist}\,(x, M \backslash Q_1) \le C(A)\, 2^k\} . \tag{2.41}$$

This constant $C(A)$ may depend on n and the Ahlfors-regularity constant for M in addition to A, but not on anything else. Combining (2.41) and (2.40) with the small boundary condition (2.19) we obtain that

$$\sum_{R \in \mathcal{T}_k} |R| \le C\, 2^{\gamma (k-j)}\, |Q_1| , \tag{2.42}$$

where C and γ are positive constants which may depend on n, A, and the Ahlfors-regularity constant for M. We can sum in k to get

$$\sum_{k \le j} \sum_{R \in \mathcal{T}_k} |R| \le C'\, |Q_1| , \tag{2.43}$$

where C' depends only on n, A, and the Ahlfors-regularity constant for M. This implies (2.38), as desired, and Lemma 2.34 follows.

Lemma 2.44. *Let $\mathcal{X}$ be a collection of cubes in Δ which is a Carleson set, and define $\widehat{\mathcal{X}}_A$ by*

$$\widehat{\mathcal{X}}_A = \bigcup_{T \in \mathcal{X}} \mathcal{C}_A(T) . \tag{2.45}$$

Then $\widehat{\mathcal{X}}_A$ is also a Carleson set, with a constant that depends only on n, A, the Ahlfors-regularity constant for M, and the Carleson constant for $\mathcal{X}$.

This is fairly easy to prove, because Carleson conditions "compose" with each other in a nice way. To make this precise, fix a cube $Q_0 \in \Delta$, and let us verify the packing condition (2.24) with $\mathcal{E} = \widehat{\mathcal{X}}_A$.

We first would like to verify that

$$\{R \in \widehat{\mathcal{X}}_A : \, R \subseteq Q_0\} \subseteq \mathcal{C}_A(Q_0) \cup \Big(\bigcup_{\substack{T \in \mathcal{X} \\ T \subseteq Q_0}} \mathcal{C}_A(T) \Big) . \tag{2.46}$$

If R lies in the left side of (2.46), then $R \subseteq Q_0$, and there is a $T \in \mathcal{X}$ such that $R \in \mathcal{C}_A(T)$. We may as well assume that T is *not* a subset of Q_0, since otherwise R lies in the last part of (2.46) automatically. In this case, T must either contain Q_0, or be disjoint from it, because of (2.17). If $T \supseteq Q_0$, then the fact that $R \in \mathcal{C}_A(T)$ and $R \subseteq Q_0 \subseteq T$ implies that R has a neighbor R' which is not contained in T. This means that R' cannot be contained in Q_0 either, so that $R \in \mathcal{C}_A(Q_0)$. If T is disjoint from Q_0, then R is not contained in T, and so R has a neighbor R' which is contained in T. Thus R' is disjoint from Q_0, and not contained in Q_0 in particular, so that again $R \in \mathcal{C}_A(Q_0)$.

Thus we have (2.46). This implies that

$$\sum_{\substack{R \in \widehat{\mathcal{X}}_A \\ R \subseteq Q_0}} |R| \le \sum_{R \in \mathcal{C}_A(Q_0)} |R| + \sum_{T \in \mathcal{X}(Q_0)} \sum_{R \in \mathcal{C}_A(T)} |R| , \tag{2.47}$$

where $\mathcal{X}(Q_0)$ denotes the set of $T \in \mathcal{X}$ with $T \subseteq Q_0$. Using Lemma 2.34 we may convert this into

$$\sum_{\substack{R \in \widehat{\mathcal{X}}_A \\ R \subseteq Q_0}} |R| \le C \, |Q_0| + C \sum_{T \in \mathcal{X}(Q_0)} |T| , \tag{2.48}$$

for a suitable constant C. This reduces further to

$$\sum_{\substack{R \in \widehat{\mathcal{X}}_A \\ R \subseteq Q_0}} |R| \le C' \, |Q_0| , \tag{2.49}$$

because of the requirement that $\mathcal{X}$ be a Carleson set. This completes the proof of Lemma 2.44.

Here is another version of "composing" Carleson conditions.

Lemma 2.50. *Let $\mathcal{F}$ be a family of stopping-time regions which are pairwise disjoint (as subsets of Δ). Assume that the collection of top cubes*

$$\{Q(S) : \ S \in \mathcal{F}\} \tag{2.51}$$

(as in (2.21)) is a Carleson set, with constant C_1. Suppose that for each $S \in \mathcal{F}$ we have a collection of cubes $\mathcal{E}(S)$ which is contained in S and which is a Carleson set with constant C_2. Then the union

$$\mathcal{E}^* = \bigcup_{S \in \mathcal{F}} \mathcal{E}(S) \tag{2.52}$$

is a Carleson set, with constant $(C_1 + 1) \cdot C_2$.

To prove this, fix a cube Q_0, and let us estimate

$$\sum_{\substack{R \in \mathcal{E}^* \\ R \subseteq Q_0}} |R| \, . \tag{2.53}$$

Put

$$\mathcal{E}_1^* = \bigcup \{\mathcal{E}(S) : \ S \in \mathcal{F}, \ Q(S) \subseteq Q_0\} \, , \tag{2.54}$$

and

$$\mathcal{E}_2^* = \{R \in \mathcal{E}^* \backslash \mathcal{E}_1^* : \ R \subseteq Q_0\} \, . \tag{2.55}$$

For $\mathcal{E}_1^*$ we have that

$$\begin{aligned} \sum_{R \in \mathcal{E}_1^*} |R| &\leq \sum_{\substack{S \in \mathcal{F} \\ Q(S) \subseteq Q_0}} \sum_{R \in \mathcal{E}(S)} |R| \\ &\leq \sum_{\substack{S \in \mathcal{F} \\ Q(S) \subseteq Q_0}} C_2 \, |Q(S)| \\ &\leq C_1 \, C_2 \, |Q_0| \, . \end{aligned} \tag{2.56}$$

The second inequality uses the Carleson condition for the $\mathcal{E}(S)$'s together with the fact that $R \subseteq Q(S)$ when $R \in \mathcal{E}(S)$, since $\mathcal{E}(S)$ is a

subset of S by assumption. (This also uses the definition of the top cube $Q(S)$ in (2.21).) The third inequality uses the Carleson condition for the collection (2.51) of top cubes.

Now consider $\mathcal{E}_2^*$. Let R be a cube in $\mathcal{E}_2^*$. Then $R \subseteq Q_0$, and there is an $S \in \mathcal{F}$ such that $R \in \mathcal{E}(S)$ but $Q(S) \not\subseteq Q_0$. We also have that $R \subseteq Q(S)$, since $\mathcal{E}(S) \subseteq S$, as above. This implies that Q_0 and $Q(S)$ are not disjoint, because they both contain R.

Given any two cubes in Δ, either one contains the other, or they are disjoint, by (2.17). For Q_0 and $Q(S)$ we have that $Q_0 \subseteq Q(S)$, since the other two possibilities have already been excluded. From this we conclude that Q_0 is an element of S, by (2.22) and the fact that $R \subseteq Q_0$.

The stopping-time regions $S \in \mathcal{F}$ are pairwise disjoint, by hypothesis, and so there is at most one $S \in \mathcal{F}$ such that $Q_0 \in S$. This means that there is a single $S_0 \in \mathcal{F}$ such that $\mathcal{E}_2^* \subseteq \mathcal{E}(S_0)$. Thus

$$\sum_{R \in \mathcal{E}_2^*} |R| \le \sum_{\substack{R \in \mathcal{E}(S_0) \\ R \subseteq Q_0}} |R| \le C_2 \, |Q_0| \,, \tag{2.57}$$

by the Carleson condition for $\mathcal{E}(S_0)$.

Combining (2.56) and (2.57), we obtain that the sum in (2.53) is bounded by $(C_1 \, C_2 + C_2) \, |Q_0|$, which is what we wanted. This completes the proof of Lemma 2.50.

Lemma 2.58. *Let $\mathcal{F}$ be a family of pairwise-disjoint stopping-time regions in Δ. For each $S \in \mathcal{F}$, put*

$$S_A = \{Q \in S : Q' \in S \text{ whenever } Q \text{ and } Q' \text{ are neighbors}\} \,, \tag{2.59}$$

and set

$$\mathcal{B}_A = \bigcup_{S \in \mathcal{F}} (S \backslash S_A) \,. \tag{2.60}$$

(Thus $\mathcal{B}_A$ consists of the cubes in Δ which lie in S for some $S \in \mathcal{F}$, but which are not so far from cubes outside of S.) If the collection of top cubes $Q(S)$, $S \in \mathcal{F}$, is a Carleson set, then so is $\mathcal{B}_A$, with a bound for the Carleson constant for $\mathcal{B}_A$ which depends only on the Carleson constant for (2.51)*, n, A, and the Ahlfors-regularity constant for M.*

It suffices to show that $S\backslash S_A$ is a Carleson set for any fixed stopping time region S, with uniform bounds (which are independent of S in particular). If we can do this, then Lemma 2.58 will follow, by Lemma 2.50.

We want to give another description of $S\backslash S_A$ in terms of the "functor" $\mathcal{C}_A(\cdot)$, and then reduce to lemmas 2.34 and 2.44.

Fix a cube Q in $S\backslash S_A$. Thus there is a cube $Q' = \tau(Q)$ such that Q and $\tau(Q)$ are neighbors, $Q \in S$, and $\tau(Q) \notin S$. Let us say that Q is of type 1, 2, or 3, according to whether $\tau(Q) \supseteq Q(S)$, $\tau(Q) \cap Q(S) = \varnothing$, or $\tau(Q) \subseteq Q(S)$, respectively. These three alternatives exhaust all possibilities, because of (2.17). Put

$$(S\backslash S_A)_i = \{Q \in S\backslash S_A : \ Q \text{ is of type } i\}\,, \qquad i = 1, 2, 3\,. \tag{2.61}$$

It suffices to show that each $(S\backslash S_A)_i$ is a Carleson set, with bounds for the Carleson constants.

If $\tau(Q) \supseteq Q(S)$, then Q and $Q(S)$ must be neighbors. This is not hard to check, since $Q \subseteq Q(S) \subseteq \tau(Q)$. From this it follows that $(S\backslash S_A)_1$ has only a bounded number of elements, and hence is a Carleson set with a bounded constant.

If $\tau(Q)$ is disjoint from $Q(S)$, then Q lies in $\mathcal{C}_A(Q(S))$. That is, $Q \subseteq Q(S)$ automatically (since $Q \in S$), $\tau(Q)$ is not contained in $Q(S)$ (since it is disjoint from $Q(S)$), and Q, $\tau(Q)$ are neighbors, as above. Thus Q, $Q' = \tau(Q)$ meet the requirements of the definition (2.33) of $\mathcal{C}_A(Q(S))$, and we conclude that

$$(S\backslash S_A)_2 \subseteq \mathcal{C}_A(Q(S))\,. \tag{2.62}$$

The Carleson condition for $(S\backslash S_A)_2$ now follows from the one for $\mathcal{C}_A(Q(S))$. (The latter corresponds to Lemma 2.44, with $\mathcal{X}$ consisting of the single cube $Q(S)$.)

We are left with the case of type 3 cubes. Set

$$\begin{aligned} b(S) = \{Q \in \Delta : \ & Q \subseteq Q(S),\, Q \notin S, \text{ and} \\ & Q \text{ is maximal with these properties}\}\,. \end{aligned} \tag{2.63}$$

Thus $b(S)$ consists of cubes "at the bottom", just below S. Notice that

$$\text{the elements of } b(S) \text{ are pairwise disjoint}\,. \tag{2.64}$$

This follows from maximality, and (2.17). If Q is an element of $(S\backslash S_A)_3$, then $\tau(Q) \subseteq Q(S)$ but $\tau(Q) \notin S$, and hence $\tau(Q) \subseteq T$ for some $T \in$

$b(S)$. That is, one simply takes T to be the maximal subcube of $Q(S)$ which contains $\tau(Q)$ and does not lie in S. We also have that $Q \not\subseteq T$, because otherwise T would have to lie in S, since S is a stopping-time region. (See (2.21).) This shows that $Q \in \mathcal{C}_A(T)$, since Q and $\tau(Q)$ are neighbors. Thus we get that

$$(S \backslash S_A)_3 \subseteq \bigcup_{T \in b(S)} \mathcal{C}_A(T) \, . \tag{2.65}$$

On the other hand, $b(S)$ is a Carleson set, with constant equal to 1, since the elements of $b(S)$ are pairwise disjoint. Lemma 2.44 now applies to say that the right side of (2.65) is a Carleson set, as desired. This completes the proof of Lemma 2.58.

The lemmas in this subsection are similar to ones in [11], especially [11, Part I, Section 3.2]. We have gone through them in some detail for the sake of clarity and completeness.

3. Measure-preserving behavior.

Standing Assumptions 3.1. *Let $(M, d(x, y))$ be a n-dimensional Ahlfors-regular metric space, and let $\{\Delta_j\}_{j<j_0}$, Δ be as in Section* 2.3. *In particular, the constants in* (2.18) *and* (2.19) *depend only on the dimension n and the Ahlfors regularity constant for M. Fix $L > 0$, let Q_0 be a cube in Δ, and let h be a mapping from Q_0 into another metric space N. We require that*

$$h \text{ is Lipschitz with constant } L \, . \tag{3.2}$$

For Proposition 3.6 below, we also ask that

$$|h(Q_0)| \geq \delta \, |Q_0| \, , \tag{3.3}$$

where δ is some (fixed) positive number. Here we use $|A|$ to denote the n-dimensional Hausdorff measure of A, whether A lies in M or N. Note that

$$\delta \leq L^n \, , \tag{3.4}$$

because of (2.3), (3.3) and (3.2).

Remark 3.5. Although everything is being stated (for simplicity) in this paper in terms of Hausdorff measures, one can allow slightly more general measures, like "Ahlfors-regular measures" on M and "Ahlfors-sub-regular" measures for the image space N. The main points are to have Ahlfors-regularity inequalities as in (2.14) for M, and to know that Lipschitz mappings from M to N do not expand measures too much, as in (2.3). (There are also minor technical issues of something like Borel regularity.) For this section even the Ahlfors-regularity does not really matter, and we shall say more about this later. In Section 11 we shall also discuss weakening the Lipschitz condition on h.

Proposition 3.6 is concerned with finding substantial regions in Δ where h behaves approximately like a measure-preserving mapping, and in a nondegenerate way. Here and later we shall want to have estimates which do not depend on Q_0 or h, but which may depend on the constants above.

Proposition 3.6. *Notation and assumptions as above. Let $\tau > 0$ be given (normally small). There exist positive constants k and α, depending only on τ and the constants n, L, and δ above, so that the following is true. There is a family $\mathcal{F}$ of pairwise-disjoint stopping-time regions in Δ, and a measurable subset E of Q_0, with the following properties*:

a) $|E| \geq \alpha\, |Q_0|$,

b) *if $Q \in \Delta$ satisfies $Q \subseteq Q_0$ and $Q \cap E \neq \varnothing$, then Q lies in S for some stopping-time region $S \in \mathcal{F}$,*

c) *if $Q \in S$ and $S \in \mathcal{F}$, then $Q \subseteq Q_0$, and*

$$(1+\tau)^{-1} \frac{|h(Q(S))|}{|Q(S)|} \leq \frac{|h(Q)|}{|Q|} \leq (1+\tau) \frac{|h(Q(S))|}{|Q(S)|}\ , \tag{3.7}$$

d) $|h(Q(S))| \geq \delta\, |Q(S)|$ *for all $S \in \mathcal{F}$,*

e) *for each $x \in M$, there are at most k choices of $S \in \mathcal{F}$ such that $x \in Q(S)$.*

Recall that $Q(S)$ denotes the top cube in the stopping-time region S, as in (2.21).

To rephrase the conclusions of the proposition, a) says that E contains a definite proportion of Q_0, b) provides a precise sense in

which the stopping-time regions in $\mathcal{F}$ contain a substantial part of the cubes in Δ which are subsets of Q_0, c) says that h is approximately measure-preserving (up to a scale factor) on each stopping-time region in $\mathcal{F}$, d) gives a lower bound for the scale factors in c), and e) implies that there are not too many of the stopping-time regions in $\mathcal{F}$. In particular, the family of top cubes $Q(S)$, $S \in \mathcal{F}$, is a Carleson set, in the sense of Section 2.5. This follows from (2.27).

The proposition would not be very interesting without e), or something like a Carleson or packing condition, because then the stopping-time regions would be able to "stop" too often. If S did not contain any element besides the top cube $Q(S)$, for instance, c) would not contain any information at all. As it is, there have to be a lot of pretty big stopping-time regions S in $\mathcal{F}$, because of the properties a), b), and e) above.

An important point here is that we are free to take τ as small as we wish, so that, in effect, the behavior of h on the stopping-time regions $S \in \mathcal{F}$ is as nice as we want. The price for this comes in the constants k and α, but for the purposes of making proofs this is often very simple, and causes no trouble.

For this proposition the Ahlfors regularity of M will not really be important, nor the measure-theoretic properties of the cubes in Δ (beyond the fact that they have finite H^n-measure). In other words, this proposition really works at a "martingale" level. In the next sections we shall give refinements of it which do rely on the Ahlfors-regularity of M, and the properties (2.18), (2.19) of cubes in Δ (and not just (2.16), (2.17)).

The remainder of this section will be devoted to the proof of Proposition 3.6. Let $\tau > 0$ be given, as above. To find $\mathcal{F}$ we run the obvious stopping-time argument. We begin with Q_0 itself, and we consider the following two conditions for stopping at a cube $Q \subseteq Q_0$

$$\frac{|h(Q)|}{|Q|} < (1 + \tau)^{-1} \, \frac{|h(Q_0)|}{|Q_0|} \, , \tag{3.8}$$

$$\frac{|h(Q)|}{|Q|} > (1 + \tau) \, \frac{|h(Q_0)|}{|Q_0|} \, . \tag{3.9}$$

The first condition will be more serious for us, in that we shall really stop when we reach a cube which satisfies it. When we reach a cube which satisfies the second condition, we shall stop the given stopping-time region, but then start a new one.

More precisely, to define our first stopping-time region S_0, we begin by taking Q_0 to be the top cube of S_0. We then look at the children of Q_0, and keep (for S_0) the ones that satisfy both (3.8) and (3.9), discarding the others, at least for the moment. For the cubes that are kept we repeat the process over and over again. In the end we get a collection S_0 of cubes which is indeed a stopping-time region (in the sense of (2.21) and (2.22), with $Q(S_0) = Q_0$), and for which properties c) and d) in Proposition 3.6 hold automatically.

Let $b(S_0)$ be as in (2.63), near the end of Section 2.6. By construction, each cube Q in $b(S_0)$ satisfies one of (3.8) and (3.9), *i.e.*, otherwise we would not have stopped. Let $b_1(S_0)$ denote the set of cubes Q in $b(S_0)$ which satisfy (3.8), and let $b_2(S_0)$ be the remaining set of cubes in $b(S_0)$ which satisfy (3.9). Note that the cubes in $b(S_0)$ need not cover all of Q_0, *i.e.*, one may be able to go all the way down to individual points without ever having to stop.

Let us write $\mathcal{F}_0$ for the "family" which consists of S_0 alone. Next we want to define a family $\mathcal{F}_1$ as follows. Let Q_1 be an element of $b_2(S_0)$, if there are any. Using exactly the same procedure as before, we can get a stopping-time region S_1 with $Q(S_1) = Q_1$. That is, we start with Q_1, and proceed to its children, grandchildren, etc., stopping whenever we reach a cube Q which satisfies the analogue of one of (3.8) and (3.9), but with Q_1 in place of Q_0. We do this for every element Q_1 of $b_2(S_0)$, ignoring the elements of $b_1(S_0)$. We take for $\mathcal{F}_1$ the family of stopping-time regions produced in this way.

Similarly, we define $\mathcal{F}_2$ by applying the same procedure to elements of $b_2(S)$, $S \in \mathcal{F}_1$, where $b_2(S)$ is defined exactly as before (the set of cubes in $b(S)$ for which the analogue of (3.9) holds). We repeat the process, obtaining families $\mathcal{F}_3$, $\mathcal{F}_4$, etc., until we run out of cubes from which to start again. In the end we take $\mathcal{F}$ to be the union of all the $\mathcal{F}_j$'s, $j \geq 0$, which are produced in this manner.

If $S \in \mathcal{F}_j$, then

$$|h(Q(S))| \geq (1 + \tau)^j \, \delta \, |Q(S)| \, . \tag{3.10}$$

This follows from the construction; in proceeding from one generation to the next, the mass ratio always went up by at least a factor of $1 + \tau$, because we were careful to start new stopping-time regions only for cubes which satisfied (3.9) (and its counterparts in successive generations). When $j = 0$, we have that $Q(S) = Q_0$, and (3.10) reduces to (3.3).

From this we may conclude that there is a constant k, depending only on τ, δ, L, and n, so that

$$\mathcal{F}_j \text{ is empty when } j \ge k\,. \tag{3.11}$$

To see this, remember that h is Lipschitz with constant at most L, by assumption, so that

$$|h(Q)| \le L^n\,|Q|\,, \tag{3.12}$$

for all cubes Q on which h is defined. (See (2.3).) This upper bound is incompatible with (3.10) when j is sufficiently large.

Another basic feature of the $\mathcal{F}_j$'s is that

$$\text{the top cubes } Q(S), S \in \mathcal{F}_j, \text{ are pairwise disjoint}\,. \tag{3.13}$$

It is important that j be fixed (but arbitrary) in (3.13), since the top cubes certainly do intersect from one generation to the next. To prove (3.13), one argues by induction. When $j = 0$, there is only one stopping-time region in $\mathcal{F}_0$, and (3.13) is trivial. Suppose now that (3.13) is true for some value of j, and let us check it for $j + 1$. If $S \in \mathcal{F}_j$, then the "bottom" cubes in $b(S)$ are pairwise disjoint, as in (2.64). The totality of all cubes Q which arise in some $b(S)$, $S \in \mathcal{F}_j$, are therefore pairwise disjoint as well. This is because (3.13) holds for j, by assumption, and because the elements of $b(S)$ are all contained (as subsets) in $Q(S)$. This implies (3.13) for $j + 1$, since the top cubes of the stopping-time regions in $\mathcal{F}_{j+1}$ are always chosen among the elements of $b(S)$, $S \in \mathcal{F}_j$. Thus we have (3.13) for all j.

For future reference, let us record the fact that

$$\text{the totality of cubes } Q \in b(S),\ S \in \mathcal{F}_j\ , \text{ are pairwise disjoint}\,. \tag{3.14}$$

Again it is important that j be fixed here. This can be derived from the same argument as above, or viewed as a consequence of (3.13), using the pairwise disjointness of cubes in a fixed $b(S)$ (as in (2.64)), and the fact that $Q \subseteq Q(S)$ when $Q \in b(S)$.

Property e) in Proposition 3.6 follows from (3.13) and (3.11). Specifically, if x is any element of M, and if j is a nonnegative integer, then (3.13) implies that x can lie in $Q(S)$ for at most one choice of $S \in \mathcal{F}_j$. This gives e), since (3.11) ensures that there are at most k values of j to worry about anyway.

From (3.10) we get d) in Proposition 3.6. Property c) was incorporated directly into the construction of the stopping-time regions. It is also easy to see that the stopping-time regions in $\mathcal{F}$ are pairwise disjoint by construction (as subsets of Δ), as required in the statement of the proposition.

Let us now define $E \subseteq Q_0$ by

$$Q_0 \backslash E = \bigcup_{S \in \mathcal{F}} \bigcup_{Q \in b_1(S)} Q \,, \tag{3.15}$$

where $b_1(S)$ is as above (*i.e.*, the set of cubes Q in $b(S)$ such that the reason for "stopping" was (3.8) with Q_0 replaced by $Q(S)$). Property b) in Proposition 3.6 holds automatically, because of the definitions. This uses the fact that it was only for the cubes in $b_1(S)$ that we would stop permanently; for the elements of $b_2(S)$, we always started a new stopping-time region immediately.

It remains to establish a lower bound for the measure of E, as in a) in Proposition 3.6. For this we use the following.

Lemma 3.16. *Let Q be a cube in Δ, and suppose that $f : Q \longrightarrow N$ is Lipschitz (say). Let $\{Q_i\}_i$ be a family of subcubes of Q which are pairwise-disjoint and satisfy*

$$\frac{|f(Q_i)|}{|Q_i|} \leq (1+\tau)^{-1} \, \frac{|f(Q)|}{|Q|} \,, \tag{3.17}$$

for each i. Then

$$\left| f(Q \setminus \bigcup_i Q_i) \right| \geq \frac{\tau}{1+\tau} \, |f(Q)| \,. \tag{3.18}$$

In particular, if f is Lipschitz with constant L, then

$$\left| Q \setminus \bigcup_i Q_i \right| \geq L^{-n} \, \frac{\tau}{1+\tau} \, |f(Q)| \,. \tag{3.19}$$

This is quite straightforward. For simple reasons of subadditivity we have that

$$|f(Q)| \leq \left| f(Q \setminus \bigcup_i Q_i) \right| + \sum_i |f(Q_i)| \,. \tag{3.20}$$

Using (3.17) we have that

$$\begin{aligned} \sum_i |f(Q_i)| &\le (1+\tau)^{-1} \frac{|f(Q)|}{|Q|} \sum_i |Q_i| \\ &\le (1+\tau)^{-1} \frac{|f(Q)|}{|Q|} |Q| \\ &= (1+\tau)^{-1} |f(Q)| \, . \end{aligned} \tag{3.21}$$

Thus

$$|f(Q)| \le \Big| f(Q \setminus \bigcup_i Q_i) \Big| + (1+\tau)^{-1} |f(Q)| \, . \tag{3.22}$$

The requirement that f be Lipschitz ensures that $|f(Q)| < \infty$, so that we may subtract the last term on right from both sides of the inequality. This gives (3.18), and (3.19) follows from (3.18) and (2.3). This completes the proof of Lemma 3.16.

We want to apply this to the situation of h and E. We begin with the following basic step. Fix an $S \in \mathcal{F}$, and let $b_1(S)$ be as before, the elements of $b(S)$ for which the reason for stopping was (3.8), but with Q_0 replaced by $Q(S)$. Thus

$$\frac{|h(R)|}{|R|} \le (1+\tau)^{-1} \frac{|h(Q(S))|}{|Q(S)|} \, , \tag{3.23}$$

for all $R \in b_1(S)$. We also have that the R's in $b_1(S)$ are pairwise disjoint, as in (2.64). If we set

$$E_0(S) = Q(S) \setminus \bigcup_{R \in b_1(S)} R \, , \tag{3.24}$$

then we get that

$$|E_0(S)| \ge L^{-n} \frac{\tau}{1+\tau} |h(Q(S))| \, , \tag{3.25}$$

by Lemma 3.16. In particular

$$|E_0(S)| \ge L^{-n} \frac{\tau}{1+\tau} \, \delta \, |Q(S)| \, , \tag{3.26}$$

by d) in Proposition 3.6 (which was derived above, from (3.10)).

Given a nonnegative integer j, define $E_j \subseteq Q_0$ through the formula

$$Q_0 \backslash E_j = \bigcup_{i=0}^{j} \bigcup_{S \in \mathcal{F}_i} \bigcup_{R \in b_1(S)} R \,. \tag{3.27}$$

Remember that $\mathcal{F}_0$ consists of only the single stopping-time region S_0, and that $Q(S_0) = Q_0$; thus E_0 is the same as $E_0(S_0)$, with the latter defined as in (3.24). Thus

$$|E_0| \ge L^{-n} \, \frac{\tau}{1+\tau} \, \delta \, |Q_0| \,, \tag{3.28}$$

by (3.26). We also have that the set E from (3.15) is the same as E_{k-1}, because of (3.11).

We want to show that the measures of the E_j's do not decrease too fast as j increases. Let us begin by converting the definition (3.27) into

$$E_{j+1} = E_j \setminus \bigcup_{S \in \mathcal{F}_{j+1}} \bigcup_{R \in b_1(S)} R \,. \tag{3.29}$$

Next we want to check that

$$E_j \supseteq \bigcup_{S \in \mathcal{F}_j} \bigcup_{Q \in b_2(S)} Q \,, \tag{3.30}$$

for each $j \ge 0$. We argue by induction. When $j = 0$, there is exactly one stopping-time region S_0 in $\mathcal{F}_0$, and

$$E_0 = Q(S) \setminus \bigcup_{R \in b_1(S_0)} R \,. \tag{3.31}$$

This implies that E_0 contains every $Q \in b_2(S_0)$, as in (3.30), because $b_1(S)$ and $b_2(S)$ are always disjoint as subsets of $b(S)$, and because the individual elements of $b(S)$ are always disjoint from each other, as subsets of M, by (2.64).

Now suppose that (3.30) holds for some j, and let us check it for $j+1$. If $S \in \mathcal{F}_{j+1}$, then $Q(S)$ lies in $b_2(S')$ from some $S' \in \mathcal{F}_j$, by the definition of the $\mathcal{F}_i$'s. In particular, $Q(S) \subseteq E_j$, by (3.30). We want to

show that the cubes Q in $b_2(S)$ are subsets of E_{j+1}. From (3.29) and the fact that $Q(S) \subseteq E_j$ we have that

$$E_{j+1} \supseteq Q(S) \setminus \bigcup_{\widehat{S} \in \mathcal{F}_{j+1}} \bigcup_{R \in b_1(\widehat{S})} R \,. \tag{3.32}$$

Fix $Q \in b_2(S)$. Thus $Q \subseteq Q(S)$ (automatically), and Q is disjoint from every $R \in b_1(\widehat{S})$, $\widehat{S} \in \mathcal{F}_{j+1}$, because of (3.14) (with j replaced by $j+1$). Using (3.32) we get that $E_{j+1} \supseteq Q$ when $Q \in b_2(S)$. This implies (3.30) with j replaced by $j+1$, since $S \in \mathcal{F}_{j+1}$ and $Q \in b_2(S)$ are arbitrary.

This finishes the proof that (3.30) holds for every $j \geq 0$. We can rewrite (3.30) as

$$E_j \supseteq \bigcup_{S \in \mathcal{F}_{j+1}} Q(S) \,. \tag{3.33}$$

Indeed, the collection of cubes $Q(S)$, $S \in \mathcal{F}_{j+1}$, is identical to the collection of cubes Q such that $Q \in b_2(S')$ for some $S' \in \mathcal{F}_j$, by construction.

We want to show that

$$|E_{j+1}| \geq L^{-n} \, \frac{\tau}{1+\tau} \, \delta \, |E_j| \,, \tag{3.34}$$

for each $j \geq 0$ (although one could improve on this a bit). Notice first that

$$E_{j+1} \setminus \bigcup_{S \in \mathcal{F}_{j+1}} Q(S) = E_j \setminus \bigcup_{S \in \mathcal{F}_{j+1}} Q(S) \,, \tag{3.35}$$

i.e., E_{j+1} and E_j are the same outside the cubes $Q(S)$, $S \in \mathcal{F}_{j+1}$. This follows from (3.29), since

$$\bigcup_{S \in \mathcal{F}_{j+1}} \bigcup_{R \in b_1(S)} R \subseteq \bigcup_{S \in \mathcal{F}_{j+1}} Q(S) \tag{3.36}$$

automatically, so that the parts which are removed from E_j to make E_{j+1} (as in (3.29)) are contained inside the union of the $Q(S)$'s, $S \in \mathcal{F}_{j+1}$. (More precisely,

$$\bigcup_{R \in b_1(S)} R \subseteq Q(S) \,, \tag{3.37}$$

for each S, *i.e.*, the elements of $b_1(S) \subseteq b(S)$ are subcubes of S by definition, as in (2.63).)

Because of (3.35), in order to establish (3.34), it suffices to show that

$$\left| E_{j+1} \cap \bigcup_{S \in \mathcal{F}_{j+1}} Q(S) \right| \geq L^{-n} \, \frac{\tau}{1+\tau} \, \delta \left| E_j \cap \bigcup_{S \in \mathcal{F}_{j+1}} Q(S) \right| . \tag{3.38}$$

This is the same as

$$\begin{aligned} &\left| \bigcup_{S \in \mathcal{F}_{j+1}} Q(S) \setminus \bigcup_{\widehat{S} \in \mathcal{F}_{j+1}} \bigcup_{R \in b_1(\widehat{S})} R \right| \\ &\qquad \geq L^{-n} \, \frac{\tau}{1+\tau} \, \delta \left| \bigcup_{S \in \mathcal{F}_{j+1}} Q(S) \right| , \end{aligned} \tag{3.39}$$

by (3.33) and (3.29). We can simplify this a bit further, through the following remarks. The $Q(S)$'s with $S \in \mathcal{F}_{j+1}$ are pairwise disjoint, as in (3.13). Thus (3.39) is equivalent to

$$\sum_{S \in \mathcal{F}_{j+1}} \left| Q(S) \setminus \bigcup_{\widehat{S} \in \mathcal{F}_{j+1}} \bigcup_{R \in b_1(\widehat{S})} R \right| \geq L^{-n} \, \frac{\tau}{1+\tau} \, \delta \sum_{S \in \mathcal{F}_{j+1}} |Q(S)| , \tag{3.40}$$

and so it suffices to show that

$$\left| Q(S) \setminus \bigcup_{\widehat{S} \in \mathcal{F}_{j+1}} \bigcup_{R \in b_1(\widehat{S})} R \right| \geq L^{-n} \, \frac{\tau}{1+\tau} \, \delta \, |Q(S)| , \tag{3.41}$$

for every $S \in \mathcal{F}_{j+1}$. Now, if $S, \widehat{S} \in \mathcal{F}_{j+1}$, and if $S \neq \widehat{S}$, then

$$Q(S) \cap Q(\widehat{S}) = \varnothing , \tag{3.42}$$

by (3.13). This implies that $Q(S) \cap R = \varnothing$ for all $R \in b_1(\widehat{S})$ (and in fact for all $R \in b(\widehat{S})$, since R is then a subset of $Q(\widehat{S})$, by the definition (2.63) of $b(\widehat{S})$). Thus (3.41) is equivalent to

$$\left| Q(S) \setminus \bigcup_{R \in b_1(S)} R \right| \geq L^{-n} \, \frac{\tau}{1+\tau} \, \delta \, |Q(S)| . \tag{3.43}$$

We have already shown that this inequality holds for all $S \in \mathcal{F}$, as in (3.26). (Remember that $E_0(S)$ is defined in (3.24), which exactly matches with the left side of (3.43).)

This completes the proof of (3.34). Combining (3.34) with (3.28) we obtain that

$$|E_\ell| \ge \left(L^{-n} \frac{\tau}{1+\tau}\, \delta\right)^{\ell+1} |Q_0|\,, \tag{3.44}$$

for all $\ell \ge 0$. We also mentioned before (just after (3.28)) that our set E (defined in (3.15)) is equal to E_{k-1}. Therefore

$$|E| \ge \left(L^{-n} \frac{\tau}{1+\tau}\, \delta\right)^{k} |Q_0|\,. \tag{3.45}$$

Thus we have a lower bound for the mass of E_0, as required in a) of Proposition 3.6. The proof of Proposition 3.6 is now completely finished.

Remark 3.46. Let $\mathcal{F}$ be as above, and set

$$G = \bigcup_{S \in \mathcal{F}} S\,. \tag{3.47}$$

Thus G is the collection of all of the cubes Q which occur in some stopping-time region S. From the construction we have that $Q \subseteq Q_0$ for every $Q \in G$, and that Q_0 lies in G. In fact, G is itself a stopping-time region: if Q and Q' are cubes in Δ such that $Q \subseteq Q' \subseteq Q_0$ and $Q \in G$, then $Q' \in G$. This is not hard to check from the construction either. The point is that when we were choosing our stopping-time regions, we started at Q_0, and each time we "stopped" for one stopping-time region S, we either stopped for good (as for cubes in $b_1(S)$), or we started a new stopping-time region again immediately (for cubes in $b_2(S)$), with no gaps between the end of one stopping-time region and the beginning of another.

To put this another way, if Q lies in G, then all of the successive parents of Q lie in G as well, until we get to Q_0. Of course, in each stopping-time region $S \in \mathcal{F}$ successive parents of a cube in S also lie in S until one gets to the top cube $Q(S)$. If S is not the first stopping-time region S_0, then one can keep going (upwards), because the parent of $Q(S)$ lies in some $S' \in \mathcal{F}$, by construction. By repeating this, one obtains that all of the ancestors of Q which are contained in Q_0 also lie in G, as desired. This observation will be useful in Section 4.

4. First improvement – more stopping-time regions.

We continue to use the notation and assumptions from Standing Assumptions 3.1. (Note that (3.3) was *not* part of Standing Assumptions 3.1.) Given $Q \in \Delta$, set

$$\Delta(Q) = \{R \in \Delta : \ R \subseteq Q\} . \tag{4.1}$$

In Proposition 3.6 we have good behavior of h, in terms of preservation of measure, on a substantial subset of $\Delta(Q_0)$. Now we want to expand this to a larger part of $\Delta(Q_0)$, so that we do not "stop" without a good reason.

Proposition 4.2. *Let Q_0, h, etc., be as in Standing Assumptions* 3.1, *and fix $\delta, \tau > 0$. There exists a constant k_1, depending only on n, L, δ, and τ, as well as a family $\mathcal{F}_1$ of stopping-time regions in Δ and a collection $\{Q_i\}_{i \in I}$ of cubes in M, so that the following are true*:

a) *the Q_i's are pairwise disjoint subcubes of Q_0, and the stopping-time regions in $\mathcal{F}_1$ are pairwise-disjoint subsets of $\Delta(Q_0)$,*

b) *if $R \in \Delta(Q_0)$, then either $R \subseteq Q_i$ for some $i \in I$, or $R \in S$ for some $S \in \mathcal{F}_1$ (and not both),*

c) *if $Q \in S$ and $S \in \mathcal{F}_1$, then*

$$(1+\tau)^{-1} \frac{|h(Q(S))|}{|Q(S)|} \leq \frac{|h(Q)|}{|Q|} \leq (1+\tau) \frac{|h(Q(S))|}{|Q(S)|} , \tag{4.3}$$

d) $|h(Q(S))| \geq \delta \, |Q(S)|$ *for all $S \in \mathcal{F}_1$,*

e) *the family of cubes*

$$\{Q(S) : \ S \in \mathcal{F}_1\} \tag{4.4}$$

is a Carleson set with constant k_1,

f) $|h(Q_i)| < \delta |Q_i|$ *for all $i \in I$.*

This is very similar to Proposition 3.6, especially in the properties c) and d) of the stopping-time regions. The chief difference is that now we account for every cube R contained in Q_0, through b), rather than "many" such cubes, as before. Notice that f) implies that

$$\Big| h\Big(\bigcup_{i \in I} Q_i \Big) \Big| < \delta \, \Big| \bigcup_{i \in I} Q_i \Big| , \tag{4.5}$$

at least if there are any Q_i's, so that the union is not empty. In particular,

$$\left| Q_0 \setminus \bigcup_{i \in I} Q_i \right| \tag{4.6}$$

is of definite size if $|h(Q_0)| \geq 2 \, \delta \, |Q_0|$, say. (In this case (4.6) would have measure at least $L^{-n} \, \delta \, |Q_0|$, because of (2.3).)

Note that if Q_0 itself satisfies $|h(Q_0)| < \delta \, |Q_0|$, then Proposition 4.2 is trivial. We can simply take $\{Q_i\}_{i \in I}$ to consist only of Q_0, and $\mathcal{F}_1$ to be empty.

The Carleson condition in e) provides a way to say that there are not too many stopping-time regions in $\mathcal{F}_1$. As such it is similar to e) in Proposition 3.6, but a bit weaker. This reflects the fact that we now cover more of $\Delta(Q_0)$ with our stopping-time regions than we did before.

In order to prove Proposition 4.2, we basically just have to iterate Proposition 3.6. We may as well assume that

$$|h(Q_0)| \geq \delta \, |Q_0| \, , \tag{4.7}$$

since otherwise there is not much to do, as mentioned above. This permits us to apply Proposition 3.6, to get a family $\mathcal{F}$ of stopping-time regions in $\Delta(Q_0)$ with the properties described there.

Let G denote the union of the stopping-time regions $S \in \mathcal{F}$, as in (3.47). In Remark 3.46, we saw that G is itself a stopping-time region, with top cube Q_0. Let $b(G)$ be as in (2.63), *i.e.*, the collection of maximal subcubes of $Q_0 = Q(G)$ which do not lie in G. Thus

$$\Delta(Q_0) = G \cup \bigcup_{R \in b(G)} \Delta(R) \, , \tag{4.8}$$

and the union is a disjoint one, *i.e.*, G is disjoint from $\Delta(R)$ for every $R \in b(G)$, and the $\Delta(R_1)$, $\Delta(R_2)$ are disjoint from each other when R_1, R_2 are distinct elements of $b(G)$. These assertions are easy to derive from the fact that G is a stopping-time region in $\Delta(Q_0)$, with Q_0 for its top cube, and from the definition of $b(G)$. In particular, the elements of $b(G)$ are pairwise disjoint as subcubes of Q_0, by maximality (as in (2.64).)

If $Q \in b(G)$ and

$$|h(Q)| < \delta \, |Q| \, , \tag{4.9}$$

then we stop there and place Q among the Q_i's. If not, then we can repeat the whole construction with Q instead of Q_0. This leads to a family of stopping-time regions $\mathcal{F}(Q)$ with the properties given in Proposition 3.6, and a collection $G(Q) \subseteq \Delta(Q)$ as above. We do this for all of the Q's in $b(G)$ with

$$|h(Q)| \geq \delta \, |Q| \, . \tag{4.10}$$

For each of the Q's in this second class we can again look at the cubes Q' in $b(G(Q))$, and separate them according to whether (4.9) or (4.10) hold (with Q' instead of Q). When the analogue of (4.9) holds, we add the given cube Q' to the collection of Q_i's. When the analogue of (4.10) holds, we repeat the process for Q', *i.e.*, applying Proposition 3.6 to obtain a family of stopping-time regions $\mathcal{F}(Q')$ with the usual properties.

We do this indefinitely, going on forever or until no more new cubes come about (as when a collection $G(Q)$ contains all subcubes of Q, so that $b(G(Q))$ is empty). In the end we take $\{Q_i\}_{i \in I}$ to be exactly the set of cubes which satisfied (4.9) in the process above, and we take $\mathcal{F}_1$ to be the union of all the families $\mathcal{F}(Q)$ that were produced above, including $\mathcal{F} = \mathcal{F}(Q_0)$.

Clearly the Q_i's are all subcubes of Q_0, and the stopping-time regions $S \in \mathcal{F}_1$ are all contained in $\Delta(Q_0)$, by construction. It is not hard to see that

$$\Delta(Q_0) = \Big(\bigcup_{S \in \mathcal{F}_1} S \Big) \cup \Big(\bigcup_{i \in I} \Delta(Q_i) \Big) \, . \tag{4.11}$$

The first step for this was given by (4.8). In (4.8), one can think of separating the $\Delta(R)$'s, $R \in b(G)$, into two groups, according to (4.9) and (4.10). The R's that satisfy (4.9) are included among the Q_i's, and thus the corresponding $\Delta(R)$'s are taken into account in the right side of (4.11). For the R's that satisfy (4.10), one feeds into a natural recursion, in which $\Delta(R)$ is similarly decomposed as in (4.8). Repeating the process indefinitely one gets (4.11), as desired.

The same type of argument also permits one to show that the union in (4.11) is a disjoint one, *i.e.*, the S's in $\mathcal{F}_1$ are pairwise disjoint, the $\Delta(Q_i)$'s are pairwise disjoint, and the S's are pairwise disjoint from the $\Delta(Q_i)$'s. One does not really have to worry about individual S's here, because the S's in a single $\mathcal{F}(Q)$ are pairwise-disjoint, by the properties of $\mathcal{F}$ in Proposition 3.6. Thus one may as well think in terms of the

$G(Q)$'s instead of individual S's. The first step in the proof of the disjointness of the unions in (4.11) is to realize that the unions in (4.8), are disjoint ones. Indeed, the $\Delta(R)$'s, $R \in b(G)$, are disjoint from G because of the definition (2.63) of $b(G)$, and the fact that G is a stopping-time region. The disjointness of the $\Delta(R)$'s, $R \in b(G)$, from each other follows from the fact that distinct cubes in $b(G)$ are pairwise disjoint as subsets of Q_0, as in (2.64). Once one has disjointness of the various pieces in (4.8), one can get similar disjointness at later steps for the same reasons. More precisely, in later steps one takes $\Delta(Q)$ for certain cubes Q and decomposes it according to

$$\begin{aligned} \Delta(Q) &= G(Q) \cup \bigcup_{R \in b(G(Q))} \Delta(R) \\ &= \Big(\bigcup_{S \in \mathcal{F}(Q)} S \Big) \cup \Big(\bigcup_{R \in b(G(Q))} \Delta(R) \Big) . \end{aligned} \tag{4.12}$$

These unions are disjoint, for exactly the same reasons as before. This permits one to preserve disjointness at each stage of the construction, with disjointness for (4.11) in the end.

We also need to know that the Q_i's, $i \in I$, are pairwise disjoint as subsets of Q_0. This can be shown in much the same way as above, with the pairwise-disjointness of the cubes in $b(G)$ or in any $b(G(Q))$ (as in (2.64)) providing the information needed at each individual step in the construction. (Actually, one should also keep track at each stage of the disjointness of the cubes in $b(G)$ or the $b(G(Q))$'s from the cubes which have been placed among the Q_i's by that point.) One can also derive this from the disjointness in (4.11), *i.e.*, two cubes Q_1, Q_2 are disjoint if and only if the corresponding $\Delta(Q_1)$, $\Delta(Q_2)$ are disjoint as subsets of Δ.

From (4.11) and these various considerations of disjointness we obtain properties a) and b) in Proposition 4.2. We do not need to do anything for c) and d), because they are inherited directly from Proposition 3.6. This also uses the fact that we applied Proposition 3.6 only to cubes which satisfy (4.10). Property f) also follows directly from our construction, *i.e.*, the cubes that we set aside for $\{Q_i\}_{i \in I}$ were the ones for which (4.9) was true.

It remains to show that e) holds, *i.e.*, that the top cubes of the stopping-time regions $S \in \mathcal{F}_1$ form a Carleson set, with a suitably bounded constant. Let us first set some notation. Let $\mathcal{G}$ denote the set of cubes Q to which Proposition 3.6 was applied in the construction

above. Thus Q_0 is the first cube that we put into $\mathcal{G}$; then we added all of the Q's in $b(G)$ for which (4.10) was true; for each of these cubes Q, we took the cubes $R \in b(G(Q))$ for which the analogue of (4.10) for R was true, and put them into $\mathcal{G}$ too; and so on. By construction (and as discussed above), the $G(Q)$'s, $Q \in \mathcal{G}$, are disjoint stopping-time regions contained in $\Delta(Q_0)$, and $\mathcal{F}_1$ is exactly the union of the families $\mathcal{F}(Q)$'s, $Q \in \mathcal{G}$ (which came from the application of Proposition 3.6 to Q).

We would like to show that $\mathcal{G}$ is a Carleson set. We shall do this using Lemma 2.28, and the existence of the set E in Proposition 3.6 (to verify the hypotheses of Lemma 2.28). We begin by reformulating the latter for the present context.

For each cube $Q \in \mathcal{G}$, Proposition 3.6 provides a measurable set $E(Q) \subseteq Q$ such that

$$|E(Q)| \geq \alpha \, |Q| \, , \tag{4.13}$$

(where α is as in Proposition 3.6, and depends only on τ, δ, n, and L), and so that

$$Q' \in G(Q) \text{ for every } Q' \in \Delta(Q) \text{ such that } Q' \cap E(Q) \neq \varnothing \, . \tag{4.14}$$

These properties correspond exactly to a) and b) in Proposition 3.6. We can reformulate (4.14) as saying that

$$\begin{aligned} &\text{for each } x \in E(Q), \ Q \text{ is the only cube} \\ &\text{in } \Delta(Q) \text{ which contains } x \text{ and lies in } \mathcal{G} \, . \end{aligned} \tag{4.15}$$

This is because Q is the only cube which is an element of both $G(Q)$ and $\mathcal{G}$, by construction. (This last can also be seen as part of the earlier matter of disjointness.)

Let us extend (4.15), as in the next claim.

Claim 4.16. *For each cube $R \in \Delta$, there is a measurable subset $F(R)$ of R such that $|F(R)| \geq \alpha \, |R|$, and so that for each $y \in F(R)$ there is at most one $Q \in \mathcal{G}$ which contains y and is contained in R.*

If R lies in $\mathcal{G}$ itself, then Claim 4.16 follows by taking $F(R)$ to be $E(R)$, as above. Otherwise, let $\{T_j\}$ denote the collection of maximal cubes contained in both $\Delta(R)$ and $\mathcal{G}$. Set

$$F(R) = \Big(R \setminus \bigcup_j T_j \Big) \cup \Big(\bigcup_j E(T_j) \Big) \, . \tag{4.17}$$

Note that the T_j's are disjoint, by maximality (and (2.17)). It is not hard to check that $|F(R)| \geq \alpha\,|R|$, because of the analogous property for the $E(T_j)$'s. Now suppose that y lies in $F(R)$, and let us check that there is at most one $Q \in \mathcal{G}$ which contains y and is contained in R. The main point is that if $Q \in \mathcal{G}$ and $Q \subseteq R$, then $Q \subseteq T_j$ for some j; one simply has to take T_j to be the maximal element of $\mathcal{G}$ which contains Q and is contained in R. Therefore, if y does not lie in any T_j, then there can be no such Q, while if y lies in some $E(T_j)$, then there is exactly one such Q, by (4.15) (with T_j instead of Q). This proves the claim.

From Lemma 2.28 it now follows that $\mathcal{G}$ is a Carleson set with a constant that depends only on α, and hence only on τ, δ, n, and L (in terms of our original constants). We want to go from this to a Carleson condition for the top cubes $Q(S)$ of the stopping-time regions S in our family $\mathcal{F}_1$.

There are a couple of ways to do this, using property e) from Proposition 3.6. In the present notation, this last asserts that for each $Q \in \mathcal{G}$ and each $z \in M$,

$$\begin{aligned} &\text{there are at most } k \text{ stopping-time regions } S \text{ in} \\ &\mathcal{F}(Q) \text{ such that } z \text{ lies in the top cube } Q(S)\,, \end{aligned} \tag{4.18}$$

where k depends only on τ, δ, and L. Keep in mind that our present family $\mathcal{F}_1$ is nothing but the union of the $\mathcal{F}(Q)$'s, $Q \in \mathcal{G}$.

One way to get the Carleson condition for the top cubes $Q(S)$, $S \in \mathcal{F}_1$, is to use Lemma 2.28, in much the same manner as above. Now one would employ (4.18) in place of (4.15), and the analogue of Claim 4.16 would be slightly more complicated, but not in a serious way.

Alternatively, one can argue as follows. If $Q \in \mathcal{G}$, then

$$\{Q(S) :\ S \in \mathcal{F}_1\} \cap G(Q) = \{Q(S) :\ S \in \mathcal{F}(Q)\}\,. \tag{4.19}$$

This follows from the construction and definitions, *i.e.*, $\mathcal{F}_1$ is the union of the $\mathcal{F}(Q)$'s, $Q \in \mathcal{G}$, while $G(Q)$ is exactly the union of the S's in $\mathcal{F}(Q)$, and the different $G(Q)$'s are pairwise disjoint. On the other hand, (4.18) implies that

$$\{Q(S) :\ S \in \mathcal{F}(Q)\} \tag{4.20}$$

is a Carleson set for each $Q \in \mathcal{G}$, with constant at most k. (This uses (2.27) and the remarks that followed it.) Remember that the $G(Q)$'s

are themselves stopping-time regions, as in Remark 3.46 at the end of Section 3. These pieces of information, together with the disjointness of the $G(Q)$'s, the fact that

$$\{Q(S) : \ S \in \mathcal{F}_1\} \subseteq \bigcup_{Q \in \mathcal{G}} G(Q)\,, \tag{4.21}$$

and the Carleson condition for $\mathcal{G}$, permit us to apply Lemma 2.50 from Section 2.6 to conclude that the total collection of top cubes $Q(S)$, $S \in \mathcal{F}_1$, is a Carleson set, with bounded constant.

This completes the proof of Proposition 4.2.

Remark 4.22. If one happens to know that $|h(Q)| \geq \delta\,|Q|$ for all cubes $Q \in \Delta(Q)$, then the collection of Q_i's in Proposition 4.2 is empty, and one has a decomposition of all of $\Delta(Q)$ into stopping-time regions on which h is approximately measure-preserving. In particular, this condition holds (for some $\delta > 0$) when h is a "regular" mapping in the sense of [5].

Alternatively, the given mapping h may degenerate on some cubes, but one might be in circumstances so that for each cube $Q \in \Delta$ there is a L-Lipschitz mapping $h_Q : Q \longrightarrow N$ such that

$$|h_Q(Q)| \geq \delta\,|Q|\,. \tag{4.23}$$

This happens for the condition of "big projections" (as in [10], [11]), for instance, and something like this happens with the looking-down relation in [12] (between spaces which have "big pieces of themselves", which is a self-similarity property). In such a situation one can get rid of the Q_i's by permitting oneself to start over with a new mapping, rather than always using the same h. (This should be compared with the various kinds of Corona conditions, as in [11].)

5. Second improvement – good stopping-time regions.

It will be useful later on to modify slightly the stopping-time regions from Proposition 4.2, as in the following notion.

Definition 5.1. *A stopping-time region $S \subseteq \Delta$ will be called good if for each cube Q in S it is either true that all of the children of Q also lie in S, or that none of them do.*

There is a general procedure for subdividing an arbitrary stopping-time regions into ones which are "good", and we shall discuss that in a moment. First, let us give an indication of how this property can be used.

Lemma 5.2. *Suppose that $S \subseteq \Delta$ is a good stopping-time region. Let Q be a cube in S, and let $\{T_i\}$ be a finite family of pairwise-disjoint cubes in Δ such that $T_i \in S$ and $T_i \subseteq Q$ for all i. Then there is a finite family $\{W_\ell\}$ of pairwise-disjoint cubes in Δ such that $W_\ell \in S$ and $W_\ell \subseteq Q$ for all ℓ, the W_ℓ's are all disjoint from the T_i's, and*

$$Q = \Big(\bigcup_i T_i\Big) \cup \Big(\bigcup_\ell W_\ell\Big). \tag{5.3}$$

In other words, we have a lot of freedom in decomposing Q into disjoint cubes from S. To prove this, let S, Q, and $\{T_i\}$ be given as above, and let $\{W_\ell\}$ be the family of maximal subcubes of Q which are disjoint from all of the T_i's. Thus the W_ℓ's are disjoint from the T_i's by definition, and they are disjoint from each other by maximality (and (2.17)). The equality (5.3) follows from the assumption that there are only finitely many T_i's. That is, if $x \in Q$ does not lie in any T_i, then there is an entire cube in Δ which contains x and remains disjoint from the T_i's (as one can readily check). This ensures that x is contained in a maximal such cube, which is then among the W_ℓ's.

It is not hard to see that there are only finitely many W_ℓ's. Indeed, if j_1 is an integer such that each T_i lies in Δ_j for some $j \geq j_1$ (which exists, since there are only finitely many T_i's), then each W_ℓ is contained in a Δ_j with $j \geq j_1$ too. (In fact, one can initially realize $Q \backslash \bigcup_i T_i$ as a (finite) union of cubes in Δ_{j_1}, and then the W_ℓ's are all supersets of these "elementary" complementary cubes in Δ_j.)

The remaining point is that the W_ℓ's are all contained in S. To see this, fix ℓ, and let $\widehat{W}_\ell$ denote the parent of W_ℓ. Except in the trivial case where $\{T_i\}$ is empty, we must have that $\widehat{W}_\ell$ is contained in Q (since otherwise $W_\ell = Q$). Also, $\widehat{W}_\ell \cap T_i \neq \varnothing$ for at least one i, since W_ℓ is supposed to be maximal. By general properties of cubes (namely, (2.17)), either $\widehat{W}_\ell \subseteq T_i$, or $T_i \subseteq \widehat{W}_\ell$. This first alternative is not possible because W_ℓ is disjoint from T_i, by construction. Thus T_i is contained in $\widehat{W}_\ell$. Let R denote the child of $\widehat{W}_\ell$ which contains T_i. Then R and $\widehat{W}_\ell$ lie in S, because

$$T_i \subseteq R \subseteq \widehat{W}_\ell \subseteq Q \tag{5.4}$$

and because both Q and T_i lie in S by hypothesis. This uses the property (2.22) of stopping-time regions. From here we obtain that W_ℓ also lies in S, since S is a good stopping-time region, and since $\widehat{W_\ell}$ is the parent of both R and W_ℓ. This proves Lemma 5.2.

Let us now give a modified version of Proposition 4.2, in which the stopping-time regions are all "good".

Proposition 5.5. *Let Q_0, h, etc., be as in Standing Assumptions* 3.1, *and fix $\delta, \tau > 0$. There exists a constant k_2, depending only on L, δ, and τ, as well as a family $\mathcal{F}_2$ of stopping-time regions in Δ and a collection $\{Q_i\}_{i\in I}$ of cubes in M, so that the following are true*:

a) *the Q_i's are pairwise disjoint subcubes of Q_0, and the stopping-time regions in $\mathcal{F}_2$ are pairwise-disjoint subsets of $\Delta(Q_0)$,*

b) *if $R \in \Delta(Q_0)$, then either $R \subseteq Q_i$ for some $i \in I$, or $R \in S$ for some $S \in \mathcal{F}_2$ (and not both),*

c) *if $Q, \widetilde{Q} \in S$ and $S \in \mathcal{F}_2$, then*

$$(1+\tau)^{-2}\,\frac{|h(Q)|}{|Q|} \le \frac{|h(\widetilde{Q})|}{|\widetilde{Q}|} \le (1+\tau)^2\,\frac{|h(Q)|}{|Q|}\ , \tag{5.6}$$

d) $|h(Q)| \ge (1+\tau)^{-1}\,\delta\,|Q|$ *when $Q \in S$, $S \in \mathcal{F}_2$,*

e) *the family of cubes*

$$\{Q(S):\ S \in \mathcal{F}_2\} \tag{5.7}$$

is a Carleson set with constant k_2,

f) $|h(Q_i)| < \delta |Q_i|$ *for all $i \in I$,*

g) *each $S \in \mathcal{F}_2$ is a good stopping-time region, in the sense of Definition* 5.1.

To prove this we shall use exactly the same family $\{Q_i\}_{i\in I}$ as in Proposition 4.2, and the stopping-time regions in $\mathcal{F}_2$ will be obtained by decomposing the ones in $\mathcal{F}_1$ in a certain way. In particular, every stopping-time region in $\mathcal{F}_2$ will be a subset of one in $\mathcal{F}_1$, and every stopping-time region in $\mathcal{F}_1$ will be the disjoint union of the stopping-time regions in $\mathcal{F}_2$ that it contains. With this information alone we

have that the properties a), b), c), d), and f) above follow from their counterparts in Proposition 4.2. (In the case of c) and d) above, we have to use the old version of c) an extra time for each. The inequalities have been adjusted slightly to reflect the fact that the top cubes of the stopping-time regions in $\mathcal{F}_2$ do not have quite the same special status as the top cubes of the stopping-time regions in $\mathcal{F}_1$ did.)

It remains to decompose the stopping-time regions in $\mathcal{F}_1$ into good subsets, and to do this in such a way that we still have a Carleson condition, as in e). Fix an $S \in \mathcal{F}_1$. To make the decomposition we shall use a very simple stopping-time argument. We first define a subregion S_1 of S as follows. We automatically include $Q(S)$ in S_1, which is then the top cube of S_1 too. If all of the children of $Q(S)$ in Δ lie in S, then we put all of them into S_1 as well. Otherwise, we stop, so that S_1 consists only of $Q(S)$. If we do not stop, then we do the same thing for each of the children of $Q(S)$. That is, for each child Q of $Q(S)$, we ask if the children of Q all lie in S, and if they do, we include them all in S_1, and when they do not, we stop, and do not proceed further below Q.

We repeat the process for as long as possible, perhaps indefinitely in some directions (down from $Q(S)$ towards cubes of smaller and smaller diameter). This defines S_1. If it was necessary to stop at some time inside S, so that S_1 is a proper subset of S, then we do the same thing all over again at the places where we had to stop before. That is, if Q is some cube which lies in S_1, and if Q' is a child of Q which lies in S but not in S_1, then we begin a new stopping-time region at Q' (*i.e.*, with top cube Q'), with the same rules for stopping as before.

One does this as often as necessary, perhaps infinitely often, until all of S is exhausted in this manner. In the end we obtain a realization of S as the disjoint union of stopping-time subregions of S, and these subregions are all good in the sense of Definition 5.1 by construction.

Let $\mathcal{F}_2$ denote the collection of all stopping-time regions derived from the ones in $\mathcal{F}_1$ in this manner. This is consistent with the principles mentioned at the beginning of the proof. The only remaining issue is to show that the top cubes of the elements of $\mathcal{F}_1$ satisfy a Carleson packing condition.

When does a cube Q in Δ arise as the top cube of a stopping-time region in $\mathcal{F}_2$? There are two conditions under which this happens, namely when

$$Q \text{ is the top cube of some } S \in \mathcal{F}_1 \,, \tag{5.8}$$

and when

$$\begin{aligned} &Q \text{ and its parent belong to some } S \in \mathcal{F}_1, \\ &\text{but one of the other children of the parent} \\ &\text{(siblings of } Q\text{) does not belong to } S\,. \end{aligned} \tag{5.9}$$

This is not hard to check, and in fact one could use this to give an alternative description of the stopping-time regions in $\mathcal{F}_2$ starting from the ones in $\mathcal{F}_1$ (*i.e.*, in a sense it is easier to produce the stopping-time regions in $\mathcal{F}_2$ when one knows in advance a complete list of where the regions should start. In the present case this also gives a complete list in advance of where the stopping-time regions in $\mathcal{F}_2$ should end.)

We already know from part e) of Proposition 4.2 that the cubes which satisfy (5.8) satisfy a Carleson condition, and so we only need to worry about (5.9). If (5.9) holds, then either

$$\text{one of the siblings of } Q \text{ is the top cube of some } \widetilde{S} \in \mathcal{F}_1\,, \tag{5.10}$$

or

$$\text{one of the siblings of } Q \text{ is a cube in the family } \{Q_i\}_{i \in I}\,. \tag{5.11}$$

Indeed, if S is as in (5.9), then Q has to have a sibling Q' which does not lie in S. According to b) in Proposition 4.2, we must either have $Q' \in \widetilde{S}$ for some $\widetilde{S} \in \mathcal{F}_1$, $\widetilde{S} \neq S$, or $Q' \subseteq Q_i$ for some $i \in I$. In the first case we must have that Q' is the top cube of $\widetilde{S}$, because the common parent of Q' and Q is contained in $S \neq \widetilde{S}$ by assumption. (Remember that the stopping-time regions in $\mathcal{F}_1$ are disjoint subsets of Δ, as in part a) of Proposition 4.2.) Similarly, if $Q' \subseteq Q_i$ for some $i \in I$, then $Q' = Q_i$, because otherwise Q_i would contain the common parent of Q' and Q as a subcube, and this is ruled out by part b) in Proposition 4.2 and the fact that the parent already lies in $S \in \mathcal{F}_1$.

Thus (5.9) implies that one of (5.10) and (5.11) holds. The remaining point is that each of (5.10) and (5.11) describes a collection of cubes which is a Carleson set. To see this, remember that the set of top cubes of elements of $\mathcal{F}_1$ form a Carleson set, by e) in Proposition 4.2, and that the family $\{Q_i\}_{i \in I}$ forms a Carleson set, since the Q_i's are pairwise disjoint (as in a) in Proposition 4.2). Thus we are reduced to the assertion that if $\mathcal{E}$ is a subset of Δ which is a Carleson set, then so is the collection $\mathcal{E}^s$ of siblings of $\mathcal{E}$. This is not hard to verify, and one can also see it as a special case of Lemma 2.32 in Section 2.6.

This completes the proof of Proposition 5.5.

6. Third improvement – better preservation of measure.

We continue with the notation and assumptions in Standing Assumptions 3.1, and with the notation $\Delta(Q)$ from (4.1).

Suppose that we have a cube Q on which our mapping h is defined, and that h is approximately measure-preserving on Q, in the sense that the mass ratios

$$\frac{|h(T)|}{|T|} \tag{6.1}$$

are all nearly the same when $T = Q$ as when T is a child of Q. If this is the case, then the images under h of the children of Q cannot overlap very much. To see this, note that

$$|h(Q)| \le \sum_{R \in c(Q)} |h(R)|\,, \tag{6.2}$$

where $c(Q)$ denotes the set of children of Q, and that

$$|Q| = \sum_{R \in c(Q)} |R|\,. \tag{6.3}$$

If the mass ratios are all nearly the same, then both sides of (6.2) will be approximately the same multiple of $|Q|$, because of (6.3). In particular, the inequality in (6.2) will be very close to being an equality. However, if there is significant overlap in the images of the children of Q, then (6.2) will not be too close to being an equality. For instance, given any two distinct children R_1, R_2 of Q, we can strengthen (6.2) to get

$$|h(Q)| \le \sum_{R \in c(Q)} |h(R)| - |h(R_1) \cap h(R_2)|\,, \tag{6.4}$$

so that the deviation from being an equality is at least $|h(R_1) \cap h(R_2)|$. (This strengthening uses the general fact that

$$|E_1 \cup E_2| = |E_1| + |E_2| - |E_1 \cap E_2|\,, \tag{6.5}$$

whenever E_1 and E_2 are measurable. In our case we apply this with $E_i = h(R_i)$ to get that

$$\begin{aligned} |h(R_1 \cup R_2)| &= |h(R_1) \cup h(R_2)| \\ &= |h(R_1)| + |h(R_2)| - |h(R_1) \cap h(R_2)| \,. \end{aligned} \tag{6.6}$$

Note that $h(R)$ is indeed measurable when h is Lipschitz and R is a cube, as in Lemma 2.4 in Section 2.1.)

This is all well and good, but in the context of propositions 3.6, 4.2, and 5.5, we can do this in general only for cubes and their children, rather than cubes which are close to each other but do not happen to have the same parent (or grandparent, etc.). That is what we want to correct in the present section.

Let us begin by setting some notation. Let Q be a cube in Δ, and let $j = j(Q)$ be the largest integer such that $Q \in \Delta_j$. (For ordinary dyadic cubes in $\mathbb{R}^n$ there is never more than one such j anyway, but in general there can be some ambiguity, as we have pointed out before. The range of possible j's is always bounded, though, because of (2.17).) Set

$$*Q = \bigcup \{T \in \Delta_{j(Q)} : \ \operatorname{dist}(T, Q) \le \operatorname{diam} Q\} \,. \tag{6.7}$$

Since we are taking our mapping h to be defined only on Q_0, let us "truncate" $*Q$ slightly by setting

$$\widehat{Q} = *Q \cap Q_0 \,. \tag{6.8}$$

In practice this truncation is not very significant, in that "most" cubes $Q \in \Delta(Q_0)$ have $*Q \subseteq Q_0$. This is made precise by Lemma 2.34 in Section 2.6, and we shall use this type of observation later in the section.

Let σ be a small positive number. It will have a role like that of τ in Propositions 3.6, 4.2, and 5.5, but it will need to be moderately larger than τ for the results of this section. We shall be interested in cubes Q such that

$$(1+\sigma)^{-1} \frac{|h(Q)|}{|Q|} \le \frac{|h(\widehat{Q})|}{|\widehat{Q}|} \le (1+\sigma) \frac{|h(Q)|}{|Q|} \,. \tag{6.9}$$

When this is true, it will help us to make arguments like the one indicated at the beginning of the section, for controlling the overlap of

$h(R_1)$ and $h(R_1)$ when R_1 and R_2 are nearby cubes of similar size, even when R_1, R_2 do not have the same parent.

Put

$$\mathcal{G}(\sigma) = \{Q \in \Delta(Q_0) : \text{ (6.9) holds}\}\,. \tag{6.10}$$

Let τ be a small number. We shall specify it later, depending on σ, but we want to give it a name now, for the sake of applying Proposition 5.5. Also let $\delta > 0$ be fixed (but arbitrary). Using τ and δ we may apply Proposition 5.5 to get a family $\mathcal{F}_2$ of stopping-time regions and a family $\{Q_i\}_{i \in I}$ of cubes contained in Q_0. We shall refer to these freely in this section. Set

$$G_2 = \bigcup_{S \in \mathcal{F}_2} S\,. \tag{6.11}$$

This is the same as

$$G_2 = \{Q \in \Delta(Q_0) : \ Q \not\subseteq Q_i \text{ for any } i \in I\}\,, \tag{6.12}$$

by b) in Proposition 5.5.

Proposition 6.13. *Notation and assumptions as above. This includes Standing Assumptions* 3.1*, and the (notation of the) parameters τ and δ. Let $\sigma > 0$ be given (and arbitrary). If τ is sufficiently small, depending on σ, n, and the Ahlfors-regularity constant for M, then $G_2 \backslash \mathcal{G}(\sigma)$ is a Carleson set, with a constant that depends only on τ, δ, n, and the Ahlfors-regularity constant for M.*

More precisely, it will be enough to have $\tau \le 1$ and τ less than or equal to a constant multiple of σ, where the constant depends only on n and the Ahlfors-regularity constant for M.

In other words, (6.9) holds for nearly all the cubes in G_2, at least if the parameters are chosen correctly. This gives an improvement of the measure-preserving behavior in Proposition 4.2 that will be quite useful.

One can take this a bit further and refine the stopping-time regions so that they are wholly contained in $\mathcal{G}(\sigma)$ (modulo some exceptions which could be contained in a Carleson set). We shall not need this here, but similar refinements are discussed in [11, Part I, Section 3.2].

Remark 6.14. In David's construction in [6], one derives stronger measure-preserving behavior than in Proposition 6.13, in terms of overlaps of images of cubes which are in the same level (*i.e.*, same Δ_j), but which are not necessarily too close to each other (*e.g.*, subcubes of some $\widehat{Q}$ of roughly the same size). See [6, Step D, p. 79], for instance. We shall not need this stronger restriction here.

The rest of this section will be devoted to the proof of Proposition 6.13. We begin with some simple reductions.

First reduction 6.15. *It is enough to show that, for τ is sufficiently small (depending on σ, n, and the Ahlfors-regularity constant for M),*

$$S \setminus \mathcal{G}(\sigma) \tag{6.16}$$

is a Carleson set for every $S \in \mathcal{F}_2$, with a constant that depends only on n and the Ahlfors-regularity constant for M.

This follows from Lemma 2.50 in Section 2.6 and the fact that the collection of top cubes

$$Q(S), \qquad S \in \mathcal{F}_2 \ , \tag{6.17}$$

is a Carleson set, as in part e) of Proposition 5.5. Note that we have dropped the dependence on τ and δ for the Carleson constant for $S \setminus \mathcal{G}(\sigma)$. For an individual S the dependence on τ, δ is not needed, but it reappears in the end through the Carleson condition for (6.17).

Given $Q \in \Delta$, write $j(Q)$ for the largest value of j such that $Q \in \Delta_j$. For each $S \in \mathcal{F}_2$, set

$$\begin{aligned} S' = \{Q \in S : \ & T \in S \text{ whenever } T \in \Delta_{j(Q)} \\ & \text{and } \operatorname{dist}(T, Q) \le \operatorname{diam} Q\} \ . \end{aligned} \tag{6.18}$$

The cubes T on the right are the ones used in the definition (6.7) of $*Q$, and it will be easier for us to be able to work inside of S' instead of S.

Lemma 6.19. *For each $S \in \mathcal{F}_2$, $S \setminus S'$ is a Carleson set, and with a constant that depends only on n and the Ahlfors-regularity constant of M.*

Indeed, if $Q \in S$ but $Q \notin S'$, then it means that Q has a neighbor (in the sense of Section 2.6) which is not contained in S, at least if we take the parameter A in (2.30), (2.31) to be large enough. Thus the Carleson condition for $S \backslash S'$ follows from Lemma 2.58 (which also provides a Carleson condition for the union of $S \backslash S'$ for all $S \in \mathcal{F}_2$).

Combining Lemma 6.19 with First Reduction 6.15 we get the following.

Second Reduction 6.20. *To prove Proposition* 6.13*, it suffices to show that if τ is small enough, depending on σ, n, and the Ahlfors-regularity constant for M, then*

$$S' \setminus \mathcal{G}(\sigma) \tag{6.21}$$

is a Carleson set for every $S \in \mathcal{F}_2$, with a constant that depends only on τ, n, and the Ahlfors-regularity constant for M.

Let us now fix $S \in \mathcal{F}_2$.

Third Reduction 6.22. *It is enough to show that if τ is sufficiently small, depending on the usual parameters, then for every $Q \in S$ there is a measurable subset $D(Q) \subseteq Q$ such that*

$$|D(Q)| \geq \gamma \, |Q| \tag{6.23}$$

and

$$\begin{aligned} &\text{for each } x \in D(Q), \text{ there are at most } m \text{ cubes} \\ &R \in S' \setminus \mathcal{G}(\sigma) \text{ such that } R \subseteq Q \text{ and } x \in R, \end{aligned} \tag{6.24}$$

where m and γ are positive constants that depend only on n and the Ahlfors-regularity constant for M.

In other words, if we can always find such sets $D(Q)$, then $S' \backslash \mathcal{G}(\sigma)$ is a Carleson set, and with bounded constant. This follows from Lemma 2.28, modulo a minor point; for the statement of Lemma 2.28, we should have $D(Q)$ as above for all $Q \in \Delta$, and not just for $Q \in S$. This more limited range of Q's is all that one ever really needs anyway, but we can also check directly that suitable subsets $D(Q)$ exist when $Q \notin S$, as follows. (Note that we use S here, rather than S'. This is not a real issue, but just a bit simpler.)

Let $Q \in \Delta$, $Q \notin S$ be given. As usual, either Q is disjoint from $Q(S)$, Q is contained in $Q(S)$, or Q contains $Q(S)$ as a proper subset. If Q is disjoint from $Q(S)$, then there are no cubes $R \in S$ such that $R \subseteq Q$, and we can take $D(Q) = Q$. If $Q \subseteq Q(S)$, then again there are no cubes $R \in S$ such that $R \subseteq Q$, because Q does not lie in S and S is a stopping-time region. (See (2.22) in Section 2.4.) If Q contains $Q(S)$ as a proper subset, then we can take $D(Q)$ to be $Q \backslash Q(S)$. With this choice we have (6.23) (for a suitable γ) because of the basic properties of cubes (as in Section 2.3), and (6.24) holds with $m = 0$ because cubes in S are all contained in $Q(S)$ (as in (2.21)), and hence cannot intersect $D(Q) = Q \backslash Q(S)$.

This shows the validity of the third reduction. Let us now fix $Q \in S$, as in the third reduction. We want to divide up the relevant class of "bad cubes" into two types, as follows. Set

$$\mathcal{B}_1 = \{R \in S' \setminus \mathcal{G}(\sigma) : \ *R \subseteq Q\} \tag{6.25}$$

and

$$\mathcal{B}_2 = \{R \in S' \setminus \mathcal{G}(\sigma) : \ R \subseteq Q \text{ but } * R \not\subseteq Q\} . \tag{6.26}$$

Note that $\mathcal{B}_1 \cup \mathcal{B}_2$ contains every $R \in S' \backslash \mathcal{G}(\sigma)$ such that $R \subseteq Q$.

Lemma 6.27. *There is a constant C_2, which depends only on n and the Ahlfors-regularity constant for M, so that*

$$\sum_{R \in \mathcal{B}_2} |R| \le C_2 \, |Q| . \tag{6.28}$$

This follows from Lemma 2.34 in Section 2.6. More precisely, if $R \in \mathcal{B}_2$, then $R \subseteq Q$ but $*R \not\subseteq Q$, and this implies that R has a neighbor which is not contained in Q, at least if the constant A in (2.30), (2.31) is large enough. This follows easily from the definition (6.7) of $*R$. Thus $R \in \mathcal{C}_A(Q)$, where the latter is defined in (2.33). The packing condition (6.28) then follows from the general one in Lemma 2.34.

In effect this means that we only have to worry about $\mathcal{B}_1$. Here is a more precise statement.

Fourth Reduction 6.29. *It suffices to show that if τ is small enough, depending only on σ, n, and the Ahlfors-regularity constant for M, then there is a measurable subset $E(Q)$ of Q such that*

$$|E(Q)| \geq \frac{1}{2} |Q| \tag{6.30}$$

and

$$\text{there are no cubes in } \mathcal{B}_1 \text{ which intersect } E(Q) \,. \tag{6.31}$$

To prove that this is sufficient, we use Lemma 6.27. Given $x \in Q$, let $N_2(x)$ denote the number of cubes $R \in \mathcal{B}_2$ such that $x \in R$. As in (2.27), we have that

$$\sum_{R \in \mathcal{B}_2} |R| = \int_Q N_2(x) \, dx \,, \tag{6.32}$$

by Fubini's theorem. If C_2 is as in Lemma 6.27, then

$$|\{x \in Q : \ N_2(x) \geq 4 \, C_2\}| \leq \frac{1}{4} |Q| \,. \tag{6.33}$$

If there exists a subset $E(Q)$ of Q as in the fourth reduction, then we can set

$$D(Q) = \{x \in E(Q) : \ N_2(x) < 4 \, C_2\} \,, \tag{6.34}$$

and this does the job. Specifically,

$$|D(Q)| \geq \frac{1}{4} |Q| \,, \tag{6.35}$$

by (6.30) and (6.33). Also, if $x \in D(Q)$, then $N_2(x) < 4 \, C_2$, by construction, so that there are fewer than $4 \, C_2$ cubes $R \in \mathcal{B}_2$ with $x \in R$, and there are no cubes $R \in \mathcal{B}_1$ which contain x, as in (6.31). Thus we get (6.23) and (6.24), with $\gamma = 1/4$ and $m = 4 \, C_2$. This finishes the justification of the fourth reduction.

Thus it remains to show that we can actually find $E(Q) \subseteq Q$ as in the fourth reduction. This is really the heart of the matter, and indeed the preceding reductions work in a very general way. It is only for this last part that we need to choose τ carefully, or that we really

use the mapping h, or the definition of $\mathcal{G}(\sigma)$, or the information about the stopping-time region S from Proposition 5.5.

Set

$$U = \bigcup_{R \in \mathcal{B}_1} R \,. \tag{6.36}$$

According to the fourth reduction, we want to show that

$$|U| \leq \frac{1}{2} \, |Q| \tag{6.37}$$

when τ is small enough (*i.e.*, we can then take $E(Q) = Q \backslash U$.)

Given a cube $T \in \Delta$ and a number $\lambda \geq 1$, put

$$\lambda T = \{x \in M : \ \mathrm{dist}\,(x, T) \leq (\lambda - 1)\,\mathrm{diam}\, T\} \,. \tag{6.38}$$

We begin with the following covering lemma.

Lemma 6.39. *There is a family $\{R_j\}_{j \in J}$ of elements of $\mathcal{B}_1$ such that*

$$*R_j \cap *R_i = \varnothing \,, \qquad \text{when } j \neq i \,, \tag{6.40}$$

and

$$\bigcup_{R \in \mathcal{B}_1} R \subseteq \bigcup_{j \in J} \lambda R_j \ , \tag{6.41}$$

where λ depends only on n and the Ahlfors-regularity constant for M.

This is a version of the Vitali covering lemma, as in [27]. We include a proof for the sake of completeness, following the usual "greedy algorithm" to choose the R_j's. First take R_1 to be an element of $\mathcal{B}_1$ whose diameter is as large as possible. (Note that the possible diameters are bounded from above, since the elements of $\mathcal{B}_1$ are contained in Q by definition.) If $R_1, \ldots, R_\ell$ have been selected already, choose $R_{\ell+1} \in \mathcal{B}_1$ so that

$$*R_{\ell+1} \cap *R_i = \varnothing \,, \qquad \text{when } 1 \leq j \leq \ell \,, \tag{6.42}$$

and so that the diameter of $R_{\ell+1}$ is as large as possible. If there are no cubes in $\mathcal{B}_1$ which satisfy (6.42), then we simply stop and take $R_1, \ldots, R_\ell$ for $\{R_j\}_{j \in J}$. Otherwise we keep going.

This defines the family $\{R_j\}_{j \in J}$. If there are infinitely many R_j's, then their diameters must tend to 0. This is because they are all subcubes of Q, and because of the basic properties of cubes in Section 2.3. (There are only finitely many subcubes of Q which can lie in any given Δ_k, for instance.)

The first condition (6.40) holds automatically, by construction. As for (6.41), let R be any cube in $\mathcal{B}_1$, and let us show that

$$R \subseteq \lambda R_j \,, \tag{6.43}$$

for some j, at least if λ is large enough.

We may as well assume that R is not itself among the R_j's, since otherwise (6.43) is trivially true. Now, there must be some $k \in J$ so that

$$*R \cap *R_k \neq \varnothing \,, \tag{6.44}$$

since otherwise R should have been chosen among the R_j's eventually. (The only other possibility is that R fails the competition for largest diameter among the available cubes. This cannot happen forever, since the diameters of the R_j's tends to zero when there are infinitely many of them, as indicated above.)

Let k be the smallest positive integer such that (6.44) holds. Thus

$$*R \cap *R_j = \varnothing \,, \qquad \text{when } j < k \,. \tag{6.45}$$

This means that R was itself a competing choice for the cube R_k, and so

$$\operatorname{diam} R \le \operatorname{diam} R_k \,, \tag{6.46}$$

since otherwise R should have been selected instead of R_k.

Because of (6.44) and (6.46), it is easy to see that R must be contained in λR_k when λ is sufficiently large (depending only on the usual constants, through (2.18)). This uses the fact that $\operatorname{diam} *R$ is bounded by a constant multiple of $\operatorname{diam} R$, and similarly for R_k, by the definition (6.7) of $*R$. This completes the proof of Lemma 6.39.

A simple consequence of Lemma 6.39 is that

$$\Big| \bigcup_{R \in \mathcal{B}_1} R \Big| \le C_1 \sum_{j \in J} |R_j| \,, \tag{6.47}$$

for a constant C_1 that depends only on n and the Ahlfors-regularity constant of M. What we really want to prove now is that $\sum_{j \in J} |R_j|$ can be made as small as we like compared to $|Q|$ by choosing τ sufficiently small.

If $R \in \mathcal{B}_1$, then $R \notin \mathcal{G}(\sigma)$, as in the definition (6.25) of $\mathcal{B}_1$. This is something that we have to use, and the next lemmas will facilitate that.

Lemma 6.48. *If $R \in \mathcal{B}_1$, then $\widehat{R} = *R$.*

Remember that $*R$ and $\widehat{R}$ are defined in (6.7) and (6.8), respectively. Since $R \in \mathcal{B}_1$, we have that R lies in S', as in (6.25). The definition (6.18) of S' ensures that $*R \subseteq Q(S)$, because it requires that all of the cubes used to make up $*R$ in (6.7) lie in S. We also have that $Q(S) \subseteq Q_0$, because of the way that the stopping-time region $S \in \mathcal{F}_2$ was chosen, in Proposition 5.5. (Indeed, property a) in Proposition 5.5 guarantees that $S \subseteq \Delta(Q_0)$, so that $Q(S) \subseteq Q_0$.) Thus we get that $*R \subseteq Q_0$, which exactly says that $\widehat{R} = *R$, by (6.8). This proves Lemma 6.48.

Lemma 6.49. *If $\tau \le \min\{1, \sigma/3\}$ and $R \in \mathcal{B}_1$, then*

$$\frac{|h(*R)|}{|*R|} < (1+\sigma)^{-1} \frac{|h(R)|}{|R|} \, . \tag{6.50}$$

If $R \in \mathcal{B}_1$, then $R \notin \mathcal{G}(\sigma)$, as in (6.25). Remember that $\mathcal{G}(\sigma)$ consists of the subcubes of Q_0 such that (6.9) holds, as in (6.10). We know from Lemma 6.48 that $\widehat{R} = *R$, and so we conclude that either (6.50) is true, or

$$\frac{|h(*R)|}{|*R|} > (1+\sigma) \frac{|h(R)|}{|R|} \, . \tag{6.51}$$

Let $j(R)$ be the largest integer such that $R \in \Delta_j$ (as before), and let $N(R)$ denote the collection of cubes $T \in \Delta_{j(R)}$ such that $\operatorname{dist}(T, R) \le \operatorname{diam} R$. Thus

$$*R = \bigcup_{T \in N(R)} T \, , \tag{6.52}$$

by the definition (6.7) of $*R$. We also have that $R \in S'$, since $R \in \mathcal{B}_1$ (see (6.25)), and this means that each $T \in N(R)$ also lies in S, by the definition (6.18) of S'. In particular, $R \in S$, and so

$$\frac{|h(T)|}{|T|} \le (1 + \tau)^2 \, \frac{|h(R)|}{|R|} \, , \tag{6.53}$$

for $T \in N(R)$, as in property c) in Proposition 5.5. Thus

$$\begin{aligned} |h(*R)| &\le \sum_{T \in N(R)} |h(T)| \\ &\le \sum_{T \in N(R)} (1 + \tau)^2 \, \frac{|h(R)|}{|R|} \, |T| \\ &= (1 + \tau)^2 \, \frac{|h(R)|}{|R|} \, |*R| \, . \end{aligned} \tag{6.54}$$

This uses (6.52), and the fact that the T's in N are pairwise disjoint (for the last step), since they all lie in the same Δ_j. (See (2.17).)

We have assumed in Lemma 6.49 that $\tau \le \min\{1, \sigma/3\}$, and this guarantees that

$$(1 + \tau)^2 \le 1 + \sigma \, . \tag{6.55}$$

Therefore (6.54) is incompatible with (6.51). As before, this implies that (6.50) must hold. This completes the proof of Lemma 6.49.

Lemma 6.56. *If $\tau \le \min\{1, \sigma/3\}$ and $R \in \mathcal{B}_1$, then*

$$\frac{|h(*R)|}{|*R|} < (1 + \sigma)^{-1} \, (1 + \tau)^2 \, \frac{|h(Q)|}{|Q|} \, . \tag{6.57}$$

Indeed, if $R \in \mathcal{B}_1$, then R and Q both lie in S, and we can use property c) in Proposition 5.5 to obtain that

$$\frac{|h(R)|}{|R|} \le (1 + \tau)^2 \, \frac{|h(Q)|}{|Q|} \, . \tag{6.58}$$

Thus we may convert (6.50) into (6.57), as desired.

From now on we assume that $\tau \leq \min\{1, \sigma/3\}$, as above. Let $\{R_j\}_{j\in J}$ be as Lemma 6.39. Thus $R_j \in \mathcal{B}_1$ for all j, and the sets $*R_j$ are pairwise disjoint. Put

$$V = \bigcup_{j\in J} *R_j \,. \tag{6.59}$$

From these features of the R_j's and Lemma 6.56 we conclude that

$$\begin{aligned} |h(V)| &\leq \sum_{j\in J} |h(*R_j)| \\ &\leq (1+\sigma)^{-1}(1+\tau)^2 \frac{|h(Q)|}{|Q|} \sum_{j\in J} |*R_j| \qquad (6.60) \\ &= (1+\sigma)^{-1}(1+\tau)^2 \frac{|h(Q)|}{|Q|} |V| \,. \end{aligned}$$

In effect this means that

$$\frac{|h(V)|}{|V|}$$

is a bit small compared to

$$\frac{|h(Q)|}{|Q|} \,,$$

since we get to choose τ to be small compared to σ. In the next lemma we look at the complement of V in Q.

Lemma 6.61.

$$|h(Q\backslash V)| \leq (1+\tau)^2 \frac{|h(Q)|}{|Q|} |Q\backslash V| \,.$$

Let J_0 be an arbitrary finite subset of J, and set

$$V_0 = \bigcup_{j\in J_0} *R_j \,. \tag{6.62}$$

To prove Lemma 6.61 we shall derive a similar inequality for V_0, and then pass to the limit (for the case when J is infinite).

Let $N(R)$ be as defined just before (6.52), so that $N(R)$ is a finite collection of cubes and $*R$ is the union of the cubes T in $N(R)$. Each

R_j, $j \in J$, lies in $\mathcal{B}_1$, by construction, and this implies that the cubes in $N(R_j)$ lie in S for all $j \in J$, as noted just after (6.52). Thus V_0 is the finite union of pairwise-disjoint cubes which are all subcubes of Q and which lie in S. (For the disjointness we are using (6.40) (see Lemma 6.39), although one could get it for free by passing to maximal cubes.)

Remember that Q itself lies in S, by assumption (just before (6.25)). We may now apply Lemma 5.2 with this choice of Q, and with $\{T_i\}$ taken to be the family of cubes

$$T \in \bigcup_{j \in J_0} N(R_j) \tag{6.63}$$

of which V_0 is composed. This yields a finite collection $\{W_\ell\}$ of pairwise-disjoint subcubes of Q such that each W_ℓ lies in S, and

$$Q \backslash V_0 = \bigcup_\ell W_\ell \, . \tag{6.64}$$

Because each W_ℓ lies in S, we have that

$$\frac{|h(W_\ell)|}{|W_\ell|} \le (1 + \tau)^2 \, \frac{|h(Q)|}{|Q|} \, , \tag{6.65}$$

for each ℓ, by part c) of Proposition 5.5. Therefore

$$\begin{aligned} |h(Q \backslash V_0)| &\le \sum_\ell |h(W_\ell)| \\ &\le (1 + \tau)^2 \, \frac{|h(Q)|}{|Q|} \, \sum_\ell |W_\ell| \\ &= (1 + \tau)^2 \, \frac{|h(Q)|}{|Q|} \, |Q \backslash V_0| \, , \end{aligned} \tag{6.66}$$

using the fact that the W_ℓ's are pairwise disjoint in the last step. On the other hand, $V_0 \subseteq V$ by definition, and so we may convert this to

$$|h(Q \backslash V)| \le (1 + \tau)^2 \, \frac{|h(Q)|}{|Q|} \, |Q \backslash V_0| \, . \tag{6.67}$$

Because this holds for any finite subset J_0 of J, we can "pass to the limit" (if necessary) to obtain that

$$|h(Q \backslash V)| \le (1 + \tau)^2 \, \frac{|h(Q)|}{|Q|} \, |Q \backslash V| \, . \tag{6.68}$$

This is exactly what we wanted for Lemma 6.61.

We are now almost finished. We can combine Lemma 6.61 with (6.60) to conclude that

$$\begin{aligned} |h(Q)| &\le |h(V)| + |h(Q\backslash V)| \\ &\le (1+\sigma)^{-1}(1+\tau)^2 \, \frac{|h(Q)|}{|Q|} \, |V| \\ &\quad + (1+\tau)^2 \, \frac{|h(Q)|}{|Q|} \, |Q\backslash V| \, . \end{aligned} \tag{6.69}$$

Note that $|h(Q)| > 0$, by d) in Proposition 5.5, and since $Q \in S$ and $S \in \mathcal{F}_2$ (by assumption). Thus we can divide through in (6.69) by $|h(Q)|/|Q|$ to get that

$$|Q| \le (1+\sigma)^{-1}(1+\tau)^2 \, |V| + (1+\tau)^2 \, |Q\backslash V| \, . \tag{6.70}$$

Substituting $|Q\backslash V| = |Q| - |V|$ we obtain that

$$|Q| \le ((1+\sigma)^{-1} - 1) \, (1+\tau)^2 \, |V| + (1+\tau)^2 \, |Q| \, . \tag{6.71}$$

Note that $(1+\sigma)^{-1} - 1 = -\sigma \, (1+\sigma)^{-1}$. Let us move the $|Q|$ on the left side of (6.71) over to the right, and bring the $|V|$ term from the right to the left, to get that

$$\frac{\sigma}{1+\sigma} \, (1+\tau)^2 \, |V| \le (2\,\tau + \tau^2) \, |Q| \, . \tag{6.72}$$

(This uses $(1+\tau)^2 - 1 = 2\,\tau + \tau^2$.) Under the assumption that $\tau \le 1$, we can simplify this to obtain

$$|V| \le \frac{\tau}{\sigma} \, 3 \, (1+\sigma) \, |Q| \, . \tag{6.73}$$

Remember that V is as in (6.59). In particular,

$$\sum_{j \in J} |R_j| \le |V| \, , \tag{6.74}$$

since the R_j's are pairwise disjoint, as in Lemma 6.39. Combining this with (6.47) and (6.73) we conclude that

$$\Big| \bigcup_{R \in \mathcal{B}_1} R \Big| \le C_1 \, |V| \le C_1 \left(\frac{\tau}{\sigma} \right) 3 \, (1+\sigma) \, |Q| \, , \tag{6.75}$$

where C_1 depends only on n and the Ahlfors-regularity constant for M.

We assumed before that $\tau \le \min\{1, \sigma/3\}$. We now require also that τ be small enough that

$$C_1 \left(\frac{\tau}{\sigma}\right) 3\,(1+\sigma) \le \frac{1}{2}\,. \tag{6.76}$$

Applying this to (6.75) yields

$$\Big|\bigcup_{R \in \mathcal{B}_1} R\Big| \le \frac{1}{2}\,|Q|\,. \tag{6.77}$$

This is exactly what we wanted for (6.37). In other words, we have now shown that the conditions described in the fourth reduction are true, and this completes the proof of Proposition 6.13.

7. "Weak" measure-preserving conditions.

We continue to use the notations and assumptions from Standing Assumptions 3.1, with $Q_0 \in \Delta$ and the Lipschitz mapping $h : Q_0 \longrightarrow N$ in particular. We shall also use the notations $\Delta(Q)$, $*Q$, and $\widehat{Q}$ from (4.1), (6.7), and (6.8).

Let numbers $\delta > 0$ and $A > 1$ be fixed but arbitrary. As usual, δ will be employed as a threshold for deciding when the image of a cube has very small measure or not. With the parameter A we have a notion of cubes being "neighbors", as in (2.30), (2.31) in Section 2.6, and we shall use this notion freely in this section (with the implicit dependence on the choice of A).

Definition 7.1 (The class $\mathcal{M}(\zeta)$)**.** *Let ζ be a positive number, normally small. We let $\mathcal{M}(\zeta) = \mathcal{M}_A(\zeta)$ denote the collection of cubes $Q \in \Delta(Q_0)$ with the following properties*:

a) $|h(Q)| \ge (1+\zeta)^{-1}\,\delta\,|Q|$,

b) *if $R \in \Delta$ is a neighbor of Q, then $R \subseteq Q_0$, and*

$$(1+\zeta)^{-1}\,\frac{|h(Q)|}{|Q|} \le \frac{|h(R)|}{|R|} \le (1+\zeta)\,\frac{|h(Q)|}{|Q|}\,, \tag{7.2}$$

c) *if $R \in \Delta$ is a neighbor of Q, then*

$$(1+\zeta)^{-1}\,\frac{|h(Q)|}{|Q|} \le \frac{|h(\widehat{R})|}{|\widehat{R}|} \le (1+\zeta)\,\frac{|h(Q)|}{|Q|}\,. \tag{7.3}$$

Roughly speaking, $\mathcal{M}(\zeta)$ consists of the cubes $Q \in \Delta(Q_0)$ for which one has good almost-measure-preserving behavior for cubes which are not too far from Q. In this section we shall be concerned with the idea that $\mathcal{M}(\zeta)$ should contain many or even "most" cubes in $\Delta(Q_0)$. This will be made precise in Proposition 7.8 below, after we account for the cubes with small images in the next definition. The information that Proposition 7.8 provides is somewhat simpler and weaker than the earlier stories with stopping-time regions, and we shall consider these matters further after the proof of the proposition.

Definition 7.4 (The class $\mathcal{SI}$)**.** *With $\delta > 0$ fixed, as above, we let $\mathcal{SI}$ denote the collection of cubes $Q \in \Delta(Q_0)$ for which there is a $W \in \Delta(Q_0)$ such that $Q \subseteq W$ and $|h(W)| < \delta\,|W|$.*

Lemma 7.5. *Put $\Sigma = \bigcup_{Q \in \mathcal{SI}} Q$. Then*

$$|h(\Sigma)| < \delta\,|\Sigma| \le \delta\,|Q_0| \tag{7.6}$$

(at least if $\mathcal{SI}$ is not empty).

The second inequality in (7.6) is trivial, since $\Sigma \subseteq Q_0$ by definition. As for the first inequality, let $\{T_\ell\}$ denote the collection of *maximal* elements of $\mathcal{SI}$. Thus every $Q \in \mathcal{SI}$ is a subcube of some T_ℓ, and the T_ℓ's are pairwise disjoint, by maximality. (This uses (2.17).) Maximality also ensures that $|h(T_\ell)| < \delta\,|T_\ell|$ for all ℓ (*i.e.*, if this inequality did not hold, then $T_\ell \in \mathcal{SI}$ would be (properly) contained in a cube (in $\mathcal{SI}$) for which it did hold, contradicting maximality.) From these observations, we have that Σ is the disjoint union of the T_ℓ's, and hence that

$$|h(\Sigma)| \le \sum_\ell |h(T_\ell)| < \sum_\ell \delta\,|T_\ell| = \delta\,|\Sigma|\,. \tag{7.7}$$

This proves Lemma 7.5.

Proposition 7.8. *For each $\zeta > 0$, the collection*

$$\mathcal{B} = \Delta(Q_0) \backslash (\mathcal{M}(\zeta) \cup \mathcal{SI}) \tag{7.9}$$

is a Carleson set, with a constant that depends only on ζ, δ, A, n, and the Ahlfors-regularity constant for M.

We shall derive this from propositions 5.5 and 6.13. To do this, first choose positive numbers σ and τ such that

$$(1+\sigma)\,(1+\tau)^2 \le 1+\zeta\,, \tag{7.10}$$

and so that τ is small enough compared to σ for the hypothesis of Proposition 6.13. These are the only conditions that we need to impose on σ and τ, so that they may be selected in such a way as to depend only on ζ, n, and the Ahlfors-regularity constant for M.

Using this choice of τ and the value of δ fixed above, we can apply Proposition 6.13 to get a certain family $\mathcal{F}_2$ of stopping-time regions contained in $\Delta(Q_0)$ and a family $\{Q_i\}_{i\in I}$ of subcubes of Q. Let G_2 denote the union of the stopping-time regions S in $\mathcal{F}_2$, as in (6.11).

Lemma 7.11. $\Delta(Q_0)\backslash G_2 \subseteq \mathcal{SI}$.

Indeed, if $Q \in \Delta(Q_0)\backslash G_2$, then $Q \subseteq Q_i$ for some $i \in I$, by part b) of Proposition 5.5. From f) in Proposition 5.5 we also have that $|h(Q_i)| < \delta\,|Q_i|$. This implies that $Q \in \mathcal{SI}$, from which Lemma 7.11 follows.

First Reduction 7.12. *In order to prove Proposition* 7.8*, it suffices to show that $G_2\backslash\mathcal{M}(\zeta)$ is a Carleson set, with suitable bounds for the Carleson constant.*

This is an immediate consequence of Lemma 7.11.

Let $\mathcal{G}(\sigma)$ be as in (6.10). Proposition 6.13 tells us that

$$G_2\backslash\mathcal{G}(\sigma) \text{ is a Carleson set} \tag{7.13}$$

(with suitable bounds). For each stopping-time region $S \in \mathcal{F}_2$, let S_A be as defined in (2.59), *i.e.*, as the set of cubes Q in S such that every neighbor of Q lies in S as well. Thus

$$\bigcup_{S\in\mathcal{F}_2} (S\backslash S_A) \text{ is a Carleson set}\,, \tag{7.14}$$

with bounds for the Carleson constant, because of Lemma 2.58 in Section 2.58 and part e) of Proposition 5.5.

Second Reduction 7.15. *In order to prove Proposition* 7.8*, it is enough to show the following.*

Let Q be a cube in G_2 that satisfies

$$Q \in S_A \text{ for some } S \in \mathcal{F}_2 \tag{7.16}$$

and

$$R \in \mathcal{G}(\sigma) \text{ whenever } R \in \Delta(Q_0) \text{ is a neighbor of } Q\,. \tag{7.17}$$

Then $Q \in \mathcal{M}(\zeta)$.

To see that this is sufficient, notice first that the collection of cubes $Q \in G_2$ which do not satisfy (7.16) is a Carleson set. This follows from (7.14) and the fact that G_2 is the union of the S's in $\mathcal{F}_2$, by definition (6.11). Similarly, the set of $Q \in G_2$ which satisfy (7.16) but fail to satisfy (7.17) is a Carleson set, because of (7.13) and Lemma 2.32 in Section 2.6. (More precisely, we use (7.16) to ensure that every neighbor of Q lies in G_2, and then the "bad" cubes R associated to the failure of (7.17) are accounted for by (7.13).) In short, the collection of exceptions to (7.16) and (7.17) among cubes in G_2 satisfies a Carleson condition. If the assertion in the second reduction is true, so that the cubes in G_2 which do fulfill (7.16) and (7.17) lie in $\mathcal{M}(\zeta)$, then we may conclude that $G_2 \backslash \mathcal{M}(\zeta)$ is a Carleson set, as required in the first reduction.

It remains to establish the assertion in the second reduction. Let Q be a cube in G_2 such that (7.16) and (7.17) hold. We want to show that Q also satisfies conditions a), b), and c) from Definition 7.1. To do this we simply read off the information that we have from propositions 5.5 and 6.13. Specifically, a) in Definition 7.1 is true because Q lies in some $S \in \mathcal{F}_2$, by (7.16), and hence satisfies d) in Proposition 5.5. As for b) in Definition 7.1, if R is a neighbor of Q, then Q and R both lie in the same stopping-time region $S \in \mathcal{F}_2$, because of (7.16). This implies that $R \subseteq Q_0$, by a) in Proposition 5.5, while (7.2) follows from c) in Proposition 5.5 and (7.10). This leaves c) in Definition 7.1. Let R be any neighbor of Q in Δ. Thus $R \subseteq Q_0$, as above, and Q, R satisfy (c) in Proposition 5.5, *i.e.*,

$$(1+\tau)^{-2}\,\frac{|h(Q)|}{|Q|} \le \frac{|h(R)|}{|R|} \le (1+\tau)^2\,\frac{|h(Q)|}{|Q|}\,. \tag{7.18}$$

On the other hand, R also lies in $\mathcal{G}(\sigma)$, because of (7.17). This means that R satisfies (6.9), *i.e.*,

$$(1+\sigma)^{-1}\,\frac{|h(R)|}{|R|} \le \frac{|h(\widehat{R})|}{|\widehat{R}|} \le (1+\sigma)\,\frac{|h(R)|}{|R|}\,. \tag{7.19}$$

Combining (7.19), (7.18), and (7.10) we get (7.3), as desired. This proves that Q satisfies each of the conditions a), b), and c) in Definition 7.1, so that the assertion in the second reduction is true.

This completes the proof of Proposition 7.8.

Let us compare the conclusions of Proposition 7.8 with the earlier results with stopping-time regions. It will be convenient to use the word "island" to refer to a subset of $\Delta(Q_0)$ which consists of a single cube $Q \in \Delta(Q_0)$ and all of the neighbors of Q in $\Delta(Q_0)$. In the context of Proposition 7.8, we are free to take the "neighborly" parameter $A > 1$ in (2.30), (2.31) as large as we want, so that the corresponding islands are also as large as we want. No matter the choice of A, $G_2 \backslash \mathcal{M}(\zeta)$ is always small in the sense of a Carleson condition. However, these islands are always "bounded", *e.g.*, they never involve more than a finite number of scales or a finite number of cubes. By contrast, the stopping-time regions in $\mathcal{F}_2$ from Proposition 5.5 can generally be much bigger than that.

To help make this precise, we begin with the following observation.

Lemma 7.20. *Let S be a stopping-time region in Δ, and fix a point $x \in Q(S)$. Then either $T \in S$ whenever $T \in \Delta$ satisfies $x \in T$ and $\operatorname{diam} T < \operatorname{diam} Q(S)$, or there are only finitely many cubes $Q \in S$ which contain x.*

Indeed, suppose that T is a cube in Δ such that $x \in T$. Then either $T \subseteq Q(S)$ or $Q(S) \subseteq T$, because of (2.17). Under the assumption that $\operatorname{diam} T < \operatorname{diam} Q(S)$, we must have $T \subseteq Q(S)$.

If $T \notin S$, then no subcube of T lies in S, by (2.22). On the other hand, any other cube which contains x must either contain T or be contained in T, by (2.17). Therefore, if $T \notin S$, then every cube $Q \in S$ which contains x as an element also contains T as a subcube. This implies that there are only finitely many elements of S which contain x (using also (2.16) and (2.18)). This proves Lemma 7.20.

Now suppose that we are in the situation of Proposition 5.5. Let $\{Q_i\}_{i \in I}$ be the family of cubes mentioned there, and set

$$D = Q_0 \setminus \bigcup_{i \in I} Q_i \, . \tag{7.21}$$

From f) in Proposition 5.5 we have that

$$\Big|h\Big(\bigcup_{i\in I} Q_i\Big)\Big| \le \sum_{i\in I} |h(Q_i)| < \delta \sum_{i\in I} |Q_i| = \delta \Big|\bigcup_{i\in I} Q_i\Big|\,, \tag{7.22}$$

using also the disjointness of the Q_i's. If

$$|h(Q_0)| \ge \delta\,|Q_0|\,, \tag{7.23}$$

then we get that $|D| > 0$.

Given a stopping-time region S, let us write $E(S)$ for the set of points $y \in Q(S)$ such that $T \in S$ whenever T is a cube which satisfies $y \in T$ and $\operatorname{diam} T < \operatorname{diam} Q(S)$.

Lemma 7.24. *The set D is contained in the union of the $E(S)$'s, $S \in \mathcal{F}_2$, except possibly for a set of measure* 0.

To prove this, let $N(x)$ denote the number of top cubes $Q(S)$, $S \in \mathcal{F}_2$, such that $x \in Q(S)$, where x is an element of Q_0. The average of $N(x)$ over Q_0 is bounded, because of the Carleson condition e) in Proposition 5.5 for the top cubes $Q(S)$, $S \in \mathcal{F}_2$, and the identity (2.27) in Section 2.5. In particular, $N(x)$ is finite for almost all x.

Fix $x \in D$ with $N(x) < \infty$. We want to show that x lies in $E(S)$ for some $S \in \mathcal{F}_2$.

There are infinitely many cubes $Q \in \Delta(Q_0)$ which contain x, and none of these cubes are contained in a Q_i, $i \in I$, by the definition (7.21) of D. Thus all of these cubes are contained in stopping-time regions in $\mathcal{F}_2$, by b) in Proposition 5.5. If $S_0 \in \mathcal{F}_2$ contains a cube Q with $x \in Q$, then $x \in Q(S_0)$, and there are only finitely many possibilities for S_0, since $N(x) < \infty$.

Thus there is an $S \in \mathcal{F}_2$ such that S contains infinitely many cubes which contain x as an element. This implies that $x \in E(S)$, by Lemma 7.20. This completes the proof of Lemma 7.24, since $N(x) < \infty$ for almost every $x \in D$.

To summarize a bit, the assumption (7.23) implies that $|D| > 0$, and then Lemma 7.24 tells us that D is covered, except for a set of measure 0, by the sets $E(S)$, $S \in \mathcal{F}_2$. In particular,

$$|E(S)| > 0\,, \tag{7.25}$$

for at least one $S \in \mathcal{F}_2$.

When a stopping-time region S is infinite, so that $E(S) \neq \varnothing$, it is already a lot larger than any single "island" (in the sense above), at least in some directions. The islands never reach all the way down to individual points in that way, *i.e.*, containing all sufficiently small cubes which contain a given point. When (7.25) holds, it means that the stopping-time region S is much larger still, going all the way down to a lot of points. Near points of density of $E(S)$, there will even be small cubes Q such that $E(S)$ contains nearly all of $\Delta(Q)$.

What does this mean in the context of propositions 5.5, 6.13, and 7.8? In all cases, one has a lot of good behavior in terms of almost preserving measure, and in about the same range of cubes (*i.e.*, the cubes in G_2, modulo perhaps some Carleson sets, which one can consider as small). The main difference is in the *scale factors* associated to the approximate measure-preserving behavior. In part c) of Proposition 5.5, one has a single scale factor for each stopping-time region, while in the context of Proposition 7.8, each island can have its own scale factor.

The latter is significantly weaker than the former. As one starts with some cube Q and shrinks down to individual points in Q_0, one can pass through infinitely many islands, and although the change in scale factors would normally be modest as one passes from one island to another, one could still have infinitely many oscillations of definite size over the infinitely many scales. In the context of Proposition 5.5 this cannot happen nearly as much, and indeed the number of oscillations is controlled by functions like $N(x)$ above, counting the number of top cubes $Q(S)$, $S \in \mathcal{F}_2$, which contain x. As in the proof of Lemma 7.24, this function is finite almost everywhere, and the Carleson condition for the collection of top cubes gives quantitative bounds. One even has exponential integrability for $N(x)$, for instance, as mentioned in Section 2.5. (See Lemma 2.28 and the discussion which follows it.)

This type of quantitative control for the oscillations is completely analogous to some applications of Carleson's Corona construction, as on [17, p. 348]. By contrast, the type of information provided by Proposition 7.8 is closer in spirit to the Bloch space and the Zygmund class in classical analysis. It is also similar to conditions like the WGL (weak geometric lemma) and WALA (weak approximation of Lipschitz functions by affine functions) discussed in [11]. See [11] for more about "weak" conditions like these, versus ones more like Carleson's Corona construction.

Although Proposition 7.8 does provide significantly less information than one gets from propositions 5.5 and 6.13, it is exactly what we

shall need for the applications to finding big bilipschitz pieces, as in [6], [19]. This is completely analogous to what happened in [19] (see also [10]), *i.e.*, "weak" conditions of approximation by affine mappings were sufficient, even though stronger information is provided by Littlewood-Paley theory (as in [14]) and Carleson's Corona construction.

With this section we finish the treatment of almost measure-preserving behavior in this paper. In the next section we review some general criteria for finding big bilipschitz pieces, and we discuss applications afterwards.

8. Weak bilipschitz conditions.

Roughly speaking, in "weak bilipschitz conditions", one asks for approximate bilipschitz behavior at most locations and scales, where "most" is interpreted in terms of Carleson sets and packing conditions. Before we give precise definitions, let us set some notation.

The notations and assumptions in Standing Assumptions 3.1 will continue to be in force. Given a cube Q in M and a number $\lambda > 1$, put

$$\lambda\, Q = \{x \in M : \ \mathrm{dist}\,(x, Q) \le (\lambda - 1)\,\mathrm{diam}\, Q\}\,. \tag{8.1}$$

(This is the same as in (6.38), but we repeat it for convenience.)

If M has finite diameter, then let us agree to treat M itself as a cube in Δ. For instance, if necessary we can add $\Delta_{j_0} = \{M\}$ to the collections $\{\Delta_j\}_{j<j_0}$ from Section 2.3, or we can simply change Δ_{j_0-1} so that it consists exactly of M. This will not cause any trouble for (2.16)-(2.19), except perhaps for an adjustment to the constant in (2.18).

Lemma 8.2. *There is a constant $b \in (0, 1/10)$, depending only on n and the Ahlfors-regularity constant for M, so that if x and y are arbitrary distinct points in M, and Q is the smallest cube in M such that $x \in Q$ and $y \in 2\,Q$, then*

$$d(x, y) \ge 10\, b\, \mathrm{diam}\, Q\,. \tag{8.3}$$

This is easy to check. To put it another way, if $\mathrm{diam}\, Q$ is too large compared to $d(x, y)$, then one should be able to pass to a child Q' of Q and still have $x \in Q'$ and $y \in 2\,Q'$, in contradiction to minimality. This

uses the fact that the diameter of Q' is not too much smaller than the diameter of Q, because of (2.18).

Standing Assumption 8.4. *The constant b is chosen as in Lemma 8.2, and fixed.*

Let f be a mapping from M into another metric space $(N, \rho(u, v))$. Given a constant $k > 1$, let $\mathcal{BL}(k)$ denote the set of cubes $Q \in \Delta$ such that

$$\begin{aligned} &k^{-1}\, d(x, y) \le \rho(f(x), f(y)) \le k\, d(x, y)\,, \\ &\text{for all } x, y \in 2\,Q \text{ with } d(x, y) > b\,\mathrm{diam}\, Q\,. \end{aligned} \tag{8.5}$$

Definition 8.6 (Weakly bilipschitz mappings)**.** *A mapping f as above is said to be weakly bilipschitz if there is a constant k so that $\Delta \backslash \mathcal{BL}(k)$ is a Carleson set.*

This is taken from [10] (specifically, [10, Definition 3.5]).

If $\mathcal{BL}(k) = \Delta$, then it is easy to see that f is bilipschitz with constant k in the usual sense of (1.2). That is, given any pair of distinct points $x, y \in M$, one could take Q to be the minimal cube such that $x \in Q$ and $y \in 2\,Q$, and then apply (8.5) to get the bilipschitz condition for this particular pair of points.

In general a mapping could have a (limited) amount of singularities or folds and still be weakly bilipschitz, as in the examples discussed in [10]. Conversely, if a mapping $f : M \longrightarrow N$ is weakly bilipschitz, then the following is true. Fix $\varepsilon > 0$ and a cube Q. One can then find a bounded number of subsets $F_1, \ldots, F_\ell$ of Q such that the restriction of f to each F_i is bilipschitz with constant k, and so that $Q \backslash \bigcup_i F_i$ has measure less than $\varepsilon\, |Q|$. One can take ℓ to be bounded by a constant that depend only on ε, n, the Ahlfors-regularity constant for M, and the Carleson constant for $\Delta \backslash \mathcal{BL}(k)$ in Definition 8.6.

This assertion is given in [10, Proposition 3.11]. It is really just an abstraction of part of the argument in [19]. In the formulation in [10] the metric spaces M and N were taken to be subsets of Euclidean spaces, but this was not really needed. We shall say a bit about the proof in a moment, but let us first give another version of the same concepts which will be more directly applicable in this paper.

Fix a cube T in M, and suppose now that f is a mapping from T

into N. Let $\mathcal{BL}'(k)$ be the set of cubes $Q \in \Delta$ such that $Q \subseteq T$ and

$$\begin{aligned} &k^{-1}\, d(x,y) \le \rho(f(x), f(y)) \le k\, d(x,y) \\ &\text{for all } x, y \in 2\,Q \cap T \text{ with } d(x,y) > b \operatorname{diam} Q\,. \end{aligned} \tag{8.7}$$

This is practically the same as before, except that we take the localization to T into account.

Let D be a measurable subset of T, and let $\Delta(T, D)$ denote the collection of cubes Q in Δ such that $Q \subseteq T$ and $Q \cap D \neq \varnothing$ (as in (2.23)).

Definition 8.8 ((T, D)-weakly bilipschitz mappings)**.** *Notations as above. We say that $f : T \longrightarrow N$ is (T, D)-weakly bilipschitz if there is a constant k so that*

$$\mathcal{E} = \Delta(T, D) \setminus \mathcal{BL}'(k) \tag{8.9}$$

satisfies a packing condition, i.e.,

$$\sum_{Q \in \mathcal{E}} |Q| \le C_0\, |T|\,, \tag{8.10}$$

for some constant C_0.

In other words, f behaves roughly like a weakly bilipschitz mapping from the perspective of the subset D. The substitution of the packing condition (8.10) for the stronger requirement of being a Carleson set is not serious, and fits better with the conclusion that we are about to draw. It is also more compatible with the localization to (T, D) which is being made anyway (*i.e.*, which is already connected to focussing on a particular location and scale). For that matter, one could always replace D by a slightly smaller set to get a Carleson condition (and even a bit more than that), by throwing away the (small) set of points in D which are contained in a large number of cubes $Q \in \mathcal{E}$. (Compare with (2.27) in Section 2.5 and the related remarks there.)

Proposition 8.11. *Let $f : T \longrightarrow N$ and D be as above, with f a (T, D)-weakly bilipschitz mapping in particular. For each $\varepsilon > 0$ there exists a finite collection $F_1, \ldots, F_\ell$ of measurable subsets of D such that*

$$\Big| D \setminus \bigcup_{i=1}^{\ell} F_i \Big| < \varepsilon\, |T| \tag{8.12}$$

and so that the restriction of f to each F_i is k-bilipschitz, where k is as in Definition 8.8*, and where ℓ is bounded by a constant which depends only on ε, n, the Ahlfors-regularity constant for M, and the constant C_0 from Definition* 8.8.

One might say that $f : T \longrightarrow N$ behaves roughly like a branched covering on D, at least in measure-theoretic terms.

Proposition 8.11 is a minor variation on the themes of [19], [10], but we shall sketch some of the elements of the proof for the sake of clarity and completeness.

Given a cube $Q \in \Delta$, $Q \subseteq T$, let $j(Q)$ denote the largest value of j such that $Q \in \Delta_j$. Let $\widehat{Q}$ denote the union of the cubes $R \in \Delta_{j(Q)}$ such that $R \subseteq T$ and $R \cap 2\,Q \neq \varnothing$. (Note that this is slightly different from the notation in (6.8), but only slightly.)

Define $\widehat{N}(x)$ for $x \in T$ by

$$\widehat{N}(x) = \text{ the number of cubes } Q \in \mathcal{E} \text{ such that } x \in \widehat{Q}\,. \tag{8.13}$$

This is a measurable function, and Fubini's theorem yields

$$\int_T \widehat{N}(x)\,dx = \sum_{Q \in \mathcal{E}} |\widehat{Q}|\,, \tag{8.14}$$

as in (2.27) in Section 2.5. Here dx denotes H^n-measure on M. Combining this with (8.10) we get that

$$\int_T \widehat{N}(x)\,dx \le C\,C_0\,|T|\,, \tag{8.15}$$

where C depends only on n and the Ahlfors-regularity constant for M. This also uses the properties (2.18) and (2.16) of cubes, from Section 2.3, *i.e.*, to know that $|\widehat{Q}| \le C\,|Q|$ for any cube Q.

Put

$$E_\lambda = \{x \in T : \ \widehat{N}(x) > \lambda\}\,, \tag{8.16}$$

$\lambda > 0$. Thus

$$|E_\lambda| \le \frac{C\,C_0}{\lambda}\,|T|\,, \tag{8.17}$$

by (8.15) and the Tchebytchev inequality. In particular,

$$|E_\lambda| < \varepsilon\,|T| \tag{8.18}$$

if λ is large enough, *i.e.*, $\lambda > \varepsilon^{-1} C\, C_0$. We choose λ once and for all so that this is true, *e.g.*, $\lambda = \varepsilon^{-1} C\, C_0 + 1$.

It suffices now to find sets $F_1, \ldots, F_\ell \subseteq T$, with ℓ bounded as in the statement of Proposition 8.11, such that

$$D \backslash E_\lambda \subseteq \bigcup_{i=1}^{\ell} F_i \tag{8.19}$$

and so that the restriction of f to each F_i is k-bilipschitz. In other words, (8.19) implies (8.12).

To get the bilipschitz condition, it is enough to show that

$$\begin{aligned} &\text{if } x, y \text{ are distinct points in some } F_i, \text{ and if } Q \text{ is the smallest} \\ &\text{cube in } \Delta \text{ such that } x \in Q \text{ and } y \in 2\,Q, \text{ then } Q \in \mathcal{BL}'(k)\,. \end{aligned} \tag{8.20}$$

This follows from the definition (8.7) of $\mathcal{BL}'(k)$.

The rest of the proof consists of a coding argument for decomposing $D \backslash E_\lambda$ into a bounded number of subsets $F_1, \ldots, F_\ell$ which satisfy (8.20). The mapping $f : T \longrightarrow N$ and the underlying measure theory play no further role, and all that really matters are the cubes $Q \in \mathcal{E} = \Delta(T, D) \setminus \mathcal{BL}'(k)$ and the information that

$$\widehat{N}(x) \le \lambda\,, \qquad \text{when } x \in D \backslash E_\lambda\,. \tag{8.21}$$

The latter provides effective control on the way that the "bad cubes" in $\mathcal{E}$ can pile up around points that matter, *i.e.*, the elements of $D \backslash E_\lambda$. The required coding argument is practically the same as in [19], and it is reviewed also in [10] (at the end of Section 2). We omit the details, which involve only cosmetic differences from the treatments in [19], [10].

9. David's condition.

Standing Assumptions 3.1 continue to be in force in this section, with the Lipschitz mapping $h : Q_0 \longrightarrow N$ in particular.

David's condition is a slightly complicated assumption about $h : Q_0 \longrightarrow N$, and implicitly about the metric spaces M and N, which is

sufficient to enable one to find subsets of Q_0 of definite size (in terms of measure) on which h is bilipschitz, with uniform bounds, at least if one has a lower bound for the measure of the image of h at the start. This was one of the main results of [6], for which an alternate proof will be indicated in Section 10. In this section we shall show how almost-measure-preserving behavior in h can be converted into weak bilipschitz behavior (in the sense of Section 8) under the assumption of David's condition, which is a key portion of the argument described here.

The precise statement of David's condition is given below. Note that we use the notation $B_M(x, r)$ and $B_N(u, t)$ for balls in the metric spaces M and N, respectively. Keep in mind that we are using $|E|$ to denote $H^n(E)$ for subsets E of both M and N.

Condition 9.1 (David's condition)**.** *For every $C_1 > 0$ (perhaps large) and $\gamma > 0$ (perhaps small), there exist $C_2 > 0$ (large) and $\eta > 0$ (small) so that the following is true.*

Let $x \in Q_0$ and $j < j_0$ be given. Set

$$T_j(x) = \bigcup \{Q \in \Delta_j : \ Q \cap B_M(x, C_2\, 2^j) \neq \varnothing\} \,. \tag{9.2}$$

Assume that $T_j(x) \subseteq Q_0$, and that

$$|h(T_j(x))| \geq \gamma\, |T_j(x)| \,. \tag{9.3}$$

Under these assumptions, we should either have that

$$h(T_j(x)) \supseteq B_N(h(x), C_1\, 2^j) \,, \tag{9.4}$$

or that

$$\begin{aligned} &\text{there is a cube } W \in \Delta_j \text{ such that } W \subseteq T_j(x), \text{ and} \\ &|h(W)|\, |W|^{-1} \geq (1 + 2\,\eta)\, |h(T_j(x))|\, |T_j(x)|^{-1} \,. \end{aligned} \tag{9.5}$$

Condition 9.1 is by no means true in a general way, and indeed there are plenty of examples of Lipschitz mappings between Ahlfors-regular spaces of the same dimension such that the mappings have images with positive measure but are not bilipschitz on any subsets of positive measure. See [12] for some examples and more information about related topics. However, there are some significant situations in which one can show that Condition 9.1 holds, as in [6], [11], [13].

One of the basic scenarios for verifying Condition 9.1 is to take $N = \mathbb{R}^n$, and to show that if the condition failed, then M would have a kind of topological degeneracy (in dimension n). Some methods for doing this are given in [6], including the derivation of controlled deformations in which a substantial piece of M is displaced into a set of lower dimension. This is also discussed (and with somewhat more detail) in [13, Section 9].

Condition 9.1 is given as item (9) on [6, p. 77], and implicitly in [13, Main Lemma 8.7]. A modestly different version comes up in [11, Part II, Lemma 3.65]. Each of these formulations are slightly different from the others, and from the one above, in terms of the setting and background assumptions.

Remark 9.6. As in [11, Part II, Lemma 3.65], it is sometimes convenient to weaken Condition 9.1 slightly by replacing (9.4) with a requirement like

$$|B_N(h(x), C_1\, 2^j) \setminus h(T_j(x))| \le a\, 2^{jn}\,, \tag{9.7}$$

where a is a small number. This weaker version would work just as well for our arguments as (9.4), at least if a is small enough, depending on n, L (the bound for the Lipschitz constant of h from Standing Assumptions 3.1), the Ahlfors-regularity constant for M, and the parameter δ which is fixed in Standing Assumptions 9.8 below. We shall explain this further in Remark 9.56 (just after we use (9.4) in a proof).

Let now us turn to the main arguments of this section, in which we assume Condition 9.1, and see what we can get from it.

Standing Assumptions 9.8. *Let $\delta > 0$ be given, arbitrary, but fixed. This will be the choice of δ that we shall always use for Definition 7.1, and eventually for Proposition 7.8. Set*

$$\gamma = \frac{\delta}{2} \qquad \text{and} \qquad C_1 = 1\,. \tag{9.9}$$

We assume that Condition 9.1 is true, and apply it with these choices of γ and C_1. This yields positive constants C_2, η, which are now fixed as well.

Given $\zeta > 0$ and $A > 1$, let $\mathcal{M}(\zeta) = \mathcal{M}_A(\zeta)$ be as in Definition 7.1. Remember that A is used for deciding when two cubes are "neighbors",

as in (2.30), (2.31). In this section we shall generally write $\mathcal{M}_A(\zeta)$ instead of $\mathcal{M}(\zeta)$, to make explicit the dependence on A. Note that $\mathcal{M}_A(\zeta)$ implicitly involves δ too, but we shall not worry about that too much, since δ is fixed as above.

We want to show that if a cube $Q \subseteq Q_0$ lies in $\mathcal{M}_A(\zeta)$, and if A is large enough and ζ is small enough, then we have approximate bilipschitz behavior of our mapping $h : Q_0 \longrightarrow N$ at the location and scale of Q, in the sense of (8.7). Our first task will be to establish some lemmas which will give us access to the information in Condition 9.1 (*i.e.*, showing that its hypotheses are met).

In this endeavor, we shall be free to take A as large as we want, and ζ as small as we want. In particular, they may depend on C_2 and η. The price for this will come in the Carleson constant when we apply Proposition 7.8 at the end, in Section 10.

As before, we shall use the notation $\Delta(T)$ to denote the collection of cubes in Δ which are contained in a given cube T.

Lemma 9.10. *Let R be a cube in $\Delta(Q_0)$, with $R \in \Delta_j$, $j < j_0$, and let x be an element of R. If $R \in \mathcal{M}_A(\zeta)$, with A large enough, depending only on n, C_2, and the Ahlfors-regularity constant for M (and not on R or j), then $T_j(x) \subseteq Q_0$.*

This was one of the basic requirements in Condition 9.1 (just after (9.2)).

To prove the lemma, fix R and $x \in R$, and let Q be any cube in Δ_j which is contained in $T_j(x)$. If A is large enough, depending on C_2, n, and the Ahlfors-regularity constant for M, then Q and R are neighbors in the sense of (2.30), (2.31). This is easy to see. Part b) of Definition 7.1 then implies that $Q \subseteq Q_0$. (Note that Q and R have opposite roles here from what they were in Definition 7.1, with R now the element of $\mathcal{M}_A(\zeta)$.) Since this is true for all such Q, we have that $T_j(x) \subseteq Q_0$, as desired.

Next we want to show that (9.3) holds in the basic situations of concern.

Lemma 9.11. *Let R be a cube in $\Delta(Q_0)$, with $R \in \Delta_j$, $j < j_0$, and let x be a point in R. If $R \in \mathcal{M}_A(\zeta)$, with A large enough and ζ small enough, depending only on n, C_2, and the Ahlfors-regularity constant for M (and not on R or j), then (9.3) holds (with $\gamma = \delta/2$, as in (9.9)).*

Let R, j, and x be given as in the statement of the lemma. For each cube R_1, let $\widehat{R}_1$ be as defined in (6.8). We are interested in choosing R_1 so that

$$R \subseteq R_1 \subseteq Q_0 \qquad \text{and} \qquad \widehat{R}_1 \supseteq T_j(x)\,. \tag{9.12}$$

Because $T_j(x) \subseteq Q_0$, $R_1 = Q_0$ satisfies these conditions, but we would like to have R_1 be smaller than that. In fact we can (and do) choose R_1 so that (9.12) holds, and also

$$\operatorname{diam} R_1 \le C\, C_2 \operatorname{diam} R\,, \tag{9.13}$$

where C depends only on n and the Ahlfors-regularity constant for M (through the constants in (2.18)). That this is possible is easy to verify, using the assumption $T_j(x) \subseteq Q_0$, the definition (9.2) of $T_j(x)$, and the definition (6.8) of $\widehat{R}_1$ (which also relies on (6.7)).

If A is large enough, then R_1 is a neighbor of R. (See (2.30), (2.31).) The assumption that $R \in \mathcal{M}_A(\zeta)$ then ensures that

$$\frac{|h(\widehat{R}_1)|}{|\widehat{R}_1|} \ge (1+\zeta)^{-1}\, \frac{|h(R)|}{|R|}\,, \tag{9.14}$$

as in (7.3). This is pretty good, but we need to account for the image of $\widehat{R}_1 \backslash T_j(x)$ under h too, in order to get the desired lower bound (9.3) for $|h(T_j(x))|$.

What does $\widehat{R}_1 \backslash T_j(x)$ look like? Let us first check that

$$\widehat{R}_1 \text{ is a union of cubes in } \Delta_j\,. \tag{9.15}$$

Let $j(R_1)$ denote the largest integer such that $R_1 \in \Delta_{j(R_1)}$. By definition, $*R_1$ (as defined in (6.7)) is a union of cubes in $\Delta_{j(R_1)}$. This is also true of $\widehat{R}_1$, which is simply the intersection of $*R_1$ with Q_0, because of the usual property (2.17) of cubes and the fact that $R_1 \subseteq Q_0$, as in (9.12). (The latter implies that $j(R_1) \le j(Q_0)$.) Thus (9.5) holds with j replaced by $j(R_1)$. This implies (9.15) for j itself, since $j \le j(R_1)$ (because $R \subseteq R_1$), and using (2.16), (2.17) to say that cubes in Δ_k can always be realized as unions of cubes in Δ_j when $j \le k$.

From (9.15) we conclude that

$$\widehat{R}_1 \backslash T_j(x) \text{ is a union of cubes in } \Delta_j\,. \tag{9.16}$$

That is, $T_j(x)$ is a union of cubes in Δ_j by definition (see (9.2)), and we know from (2.17) that distinct cubes in Δ_j are necessarily disjoint. This permits us to derive (9.16) from (9.15).

If A is large enough, again depending only on n, C_2, and the Ahlfors-regularity constant for M, then every cube Q in Δ_j with $Q \subseteq \widehat{R}_1$ is a neighbor of R. This is the last condition on A that we shall impose. Since $R \in \mathcal{M}_A(\zeta)$, we conclude that

$$\frac{|h(Q)|}{|Q|} \le (1 + \zeta) \, \frac{|h(R)|}{|R|} \, , \tag{9.17}$$

for all such Q, by b) in Definition 7.1. (Note that the roles of R and Q here are again backwards from what they were in Definition 7.1. This does not matter for (7.2), which is symmetric in Q and R.)

Let $\mathcal{Z}$ denote the set of cubes $Q \in \Delta_j$ such that $Q \subseteq \widehat{R}_1 \backslash T_j(x)$. Thus $\widehat{R}_1 \backslash T_j(x)$ is the union of the cubes in $\mathcal{Z}$, by (9.16), and these cubes are pairwise disjoint, because of (2.17). Combining this with (9.17) we obtain that

$$\begin{aligned} |h(\widehat{R}_1 \backslash T_j(x))| &\le \sum_{Q \in \mathcal{Z}} |h(Q)| \\ &\le (1 + \zeta) \, \frac{|h(R)|}{|R|} \sum_{Q \in \mathcal{Z}} |Q| \\ &\le (1 + \zeta) \, \frac{|h(R)|}{|R|} \, |\widehat{R}_1 \backslash T_j(x)| \, . \end{aligned} \tag{9.18}$$

On the other hand, (9.14) implies that

$$(1 + \zeta)^{-1} \, \frac{|h(R)|}{|R|} \, |\widehat{R}_1| \le |h(\widehat{R}_1)| \le |h(T_j(x))| + |h(\widehat{R}_1 \backslash T_j(x))| \, . \tag{9.19}$$

This yields

$$(1 + \zeta)^{-1} \, \frac{|h(R)|}{|R|} \, |\widehat{R}_1| \le |h(T_j(x))| + (1 + \zeta) \, \frac{|h(R)|}{|R|} \, |\widehat{R}_1 \backslash T_j(x)| \, , \tag{9.20}$$

by (9.18). We can rewrite this as

$$(1 + \zeta)^{-1} \, \frac{|h(R)|}{|R|} \, (|\widehat{R}_1| - (1 + \zeta)^2 \, |\widehat{R}_1 \backslash T_j(x)|) \le |h(T_j(x))| \, , \tag{9.21}$$

and then as

$$(9.22)\quad (1+\zeta)^{-1}\,\frac{|h(R)|}{|R|}\,(|T_j(x)| - (2\,\zeta+\zeta^2)\,|\widehat{R}_1\backslash T_j(x)|) \le |h(T_j(x))|\,,$$

since

$$(9.23)\qquad |\widehat{R}_1| = |T_j(x)| + |\widehat{R}_1\backslash T_j(x)|$$

(because $T_j(x) \subseteq \widehat{R}_1$, as in (9.12).

We are assuming that $R \in \mathcal{M}_A(\zeta)$, and this implies that

$$(9.24)\qquad |h(R)| \ge (1+\zeta)^{-1}\,\delta\,|R|\,,$$

by a) in Definition 7.1. This permits us to simplify (9.22) to

$$(9.25)\qquad (1+\zeta)^{-2}\,\delta\,(|T_j(x)| - (2\,\zeta+\zeta^2)\,|\widehat{R}_1\backslash T_j(x)|) \le |h(T_j(x))|\,.$$

Next, there is a constant $C > 0$ so that

$$(9.26)\qquad |\widehat{R}_1\backslash T_j(x)| \le |\widehat{R}_1| \le C\,|T_j(x)|\,,$$

where C depends only on n and the Ahlfors-regularity constant for M. To see this, let us first check that

$$(9.27)\qquad \operatorname{diam} R_1 \le C'\,C_2\,2^j\,,$$

for some constant C' which depends only on n and the Ahlfors-regularity constant for M. This follows from (9.13) and the fact that $\operatorname{diam} R$ is bounded by a constant times 2^j, since $R \in \Delta_j$ by assumption. (See (2.18).) We also have that

$$(9.28)\qquad \operatorname{diam} \widehat{R}_1 \le C''\operatorname{diam} R_1\,,$$

where C'' depends only on n and the Ahlfors-regularity constant for M, by the definition (6.8), (6.7) of $\widehat{R}_1$ (and the usual properties of cubes). Thus the diameter of $\widehat{R}_1$ is bounded by a (geometric) constant times $C_2\,2^j$, and (9.26) then follows from the definition (9.2) of $T_j(x)$ and the Ahlfors regularity of M.

If ζ is sufficiently small, depending on n and the Ahlfors-regularity constant for M, then

$$(9.29)\qquad (1+\zeta)^{-2}\,(|T_j(x)| - (2\,\zeta+\zeta^2)\,|\widehat{R}_1\backslash T_j(x)|) \ge \frac{1}{2}\,|T_j(x)|\,,$$

because of (9.26). Plugging this into (9.25), we get that

$$\frac{1}{2}\,\delta\,|T_j(x)| \le |h(T_j(x))|\,. \tag{9.30}$$

We chose γ to be $\delta/2$, as in (9.9), and so (9.30) is the same as (9.3), which is exactly what we wanted. This completes the proof of Lemma 9.11.

The next lemma provides conditions under which (9.5) cannot occur (so that Condition 9.1 will lead us to (9.4)).

Lemma 9.31. *Let R be a cube in $\Delta(Q_0)$, with $R \in \Delta_j$, $j < j_0$. Suppose that $x \in R$ and that $T_j(x) \subseteq Q_0$. Let Q be a cube in Δ_j such that $Q \subseteq T_j(x)$. If $R \in \mathcal{M}_A(\zeta)$, with A large enough and ζ small enough, depending only on n, C_2, η, and the Ahlfors-regularity constant for M (and not on R, j, or Q), then*

$$\frac{|h(Q)|}{|Q|} < (1 + 2\,\eta)\,\frac{|h(T_j(x))|}{|T_j(x)|}\,. \tag{9.32}$$

We can prove Lemma 9.31 using practically the same estimates as for Lemma 9.11. Let R, j, x, and Q be given as in the statement of the lemma. Thus

$$\frac{|h(Q)|}{|Q|} \le (1 + \zeta)\,\frac{|h(R)|}{|R|} \tag{9.33}$$

as in (9.17), *i.e.*, Q is necessarily a neighbor of R under the conditions of the lemma, at least if A is large enough, and (9.33) follows then directly from the assumption $R \in \mathcal{M}_A(\zeta)$ and part b) of Definition 7.1.

On the other hand, we have that

$$(1 + \zeta)^{-1}\,(1 - (2\,\zeta + \zeta^2)\,C)\,\frac{|h(R)|}{|R|}\,|T_j(x)| \le |h(T_j(x))|\,, \tag{9.34}$$

because of (9.22) and (9.26). This constant C is the same as the one in (9.26), and depends only on n and the Ahlfors-regularity constant of M.

Combining (9.33) with (9.34), we obtain that

$$\frac{|h(Q)|}{|Q|} \le (1 + \zeta)^2\,(1 - (2\,\zeta + \zeta^2)\,C)^{-1}\,\frac{|h(T_j(x))|}{|T_j(x)|}\,. \tag{9.35}$$

If ζ is small enough, depending on n, η, and the Ahlfors-regularity constant for M, then (9.32) follows from this inequality. This proves Lemma 9.31.

Lemma 9.31 has the effect of neutralizing the parameter η from Condition 9.1, and we shall not have to deal with it again.

The next proposition gives the main conclusions of this section.

Proposition 9.36. *There are positive constants k, A_1, and ζ_1, depending only on n, η, C_2, L (which is the Lipschitz constant for our mapping h, as in Standing Assumptions* 9.31*), and the Ahlfors regularity constants for M, so that the following is true.*

Suppose that Q is a cube such that $Q \subseteq Q_0$ and $Q \in \mathcal{M}_{A_1}(\zeta_1)$. Then Q also lies in $\mathcal{BL}'(k)$, where the latter is defined as in (8.7). (*In* (8.7)*, one should replace T with Q_0, and f with our mapping h.*)

To prove this, let k, A_1 be large, and let ζ_1 be small, to be chosen soon. Let b be as in Lemma 8.2. Fix a cube $Q \in \Delta(Q_0)$ with $Q \in \mathcal{M}_{A_1}(\zeta_1)$. Also fix arbitrary points $x, y \in 2Q \cap Q_0$ (where $2Q$ is as defined in (8.1)) such that

$$d(x, y) > b \operatorname{diam} Q \,. \tag{9.37}$$

We want to show that, if k, A_1, and ζ_1 are chosen correctly, then

$$\rho(h(x), h(y)) \geq k^{-1} d(x, y) \,, \tag{9.38}$$

where $\rho(\cdot, \cdot)$ denotes the metric on the target space N. Once we do this, we shall be finished, because

$$\rho(h(x), h(y)) \leq L \, d(x, y) \tag{9.39}$$

holds automatically by the Lipschitz condition on h in Standing Assumptions 3.1. (Thus, to get (8.7), one should take k to be at least L.)

Let $j(Q)$ be the largest integer such that $Q \in \Delta_{j(Q)}$, and let j_1 be the largest integer at most $j(Q)$ which satisfies

$$T_{j_1}(x) \cap T_{j_1}(y) = \varnothing \,. \tag{9.40}$$

It is not hard to see that

$$0 \leq j(Q) - j_1 \leq C \,, \tag{9.41}$$

for some constant C that depends only on n, C_2, and the Ahlfors-regularity constant for M. (Note that b depends only on n and the Ahlfors-regularity constant for M, as in Lemma 8.2.) This follows from the definition (9.2) of $T_j(\cdot)$ and the usual properties of cubes (especially (2.17)).

Claim 9.42. *If A_1 is large enough, and ζ_1 is small enough, depending on n, C_2, and the Ahlfors-regularity constant for M, then $T_{j_1}(x), T_{j_1}(y) \subseteq Q_0$ and*

$$h(T_{j_1}(x)) \supseteq B_N(h(x), 2^{j_1}), \quad h(T_{j_1}(y)) \supseteq B_N(h(y), 2^{j_1}) \,. \tag{9.43}$$

To see this, let R_x and R_y denote cubes in Δ_{j_1} such that $x \in R_x$ and $y \in R_y$. Let A, ζ be positive numbers which are chosen large and small enough, respectively, so that Lemmas 9.10, 9.11, and 9.31 hold. If A_1 is large enough, and ζ_1 is small enough, depending on these choices of A and ζ (which themselves depend only on acceptable parameters), then we have that

$$R_x, R_y \in \mathcal{M}_A(\zeta) \,. \tag{9.44}$$

This is not too hard to prove. Indeed, (9.41) and the fact that $x, y \in 2\,Q$ imply that R_x and R_y are neighbors of Q, with a constant which depends only on n, C_2, and the Ahlfors-regularity constant for M. Any cube which is a neighbor of R_x or R_y is then a neighbor of Q, but with the "neighborly" constant increased in a controlled fashion. This permits one to derive the requirements in Definition 7.1 for (9.44) to hold from the corresponding features for Q coming from the hypothesis that $Q \in \mathcal{M}_{A_1}(\zeta_1)$.

Once one has (9.44), one can apply lemmas 9.10 and 9.11 to conclude that the "hypotheses" of Condition 9.1 hold for x, j_1 and y, j_1 (*i.e.*, j_1 instead of j). In particular, Lemma 9.10 implies that $T_{j_1}(x), T_{j_1}(y) \subseteq Q_0$, as asserted in Claim 9.42. The "conclusions" of Condition 9.1 then imply that one of (9.4) and (9.5) must hold for each of x, j_1 and y, j_1. From Lemma 9.31 (applied to R_x and R_y) we know that (9.5) is not possible in either case, and so we are left with (9.4) for both x, j_1 and y, j_1. In other words, (9.43) holds, which is exactly what we wanted. This proves Claim 9.42.

From now on we assume that A_1 is at least large enough, and ζ_1 at least small enough, for the purposes of Claim 9.42.

Let us assume now that (9.38) does not hold, so that

$$\rho(h(x), h(y)) < k^{-1}\, d(x, y)\,. \tag{9.45}$$

We want to derive a contradiction.

Claim 9.46. *If k is large enough, depending on n, C_2, and the Ahlfors-regularity constants for M, then*

$$h(T_{j_1}(x)) \supseteq B_N(h(y), 2^{j_1 - 1})\,. \tag{9.47}$$

Note that there is an x on the left side of (9.47), and a y on the right side. In other words, we can convert (9.43) into a condition of overlapping images. This is quite straightforward. To derive (9.47) from (9.43), we simply need to know that

$$\rho(h(x), h(y)) \le 2^{j_1 - 1}\,. \tag{9.48}$$

On the other hand, $d(x, y) \le 3 \operatorname{diam} Q$, since $x, y \in 2\,Q$ by assumption (as stated at the beginning of the proof of Proposition 9.36), and $\operatorname{diam} Q$ is bounded by a (somewhat large) constant multiple of 2^{j_1}, because of (9.41) and the usual properties of cubes (namely, (2.17)). Thus (9.45) implies that $\rho(h(x), h(y))$ is less that k^{-1} times a constant multiple of 2^{j_1}, which ensures that (9.48) holds when k is sufficiently large. This proves Claim 9.46.

From now on we assume that k is large enough for Claim 9.46 to work. We shall not need to impose any further conditions on k, so it may now be chosen and fixed, once and for all.

Claim 9.49. *If A_1 is large enough, depending only on n, C_2, L, and the Ahlfors-regularity constant for M, and if $\zeta_1 \le 1$ (say), then*

$$|h(T_{j_1}(x)) \cap h(T_{j_1}(y))| \ge c\, |h(Q)|\,, \tag{9.50}$$

where c depends only on n, C_2, L, and the Ahlfors-regularity constant for M.

To prove the claim we use (9.43) and (9.47). Note that if N happens to be Ahlfors regular of dimension n, then Claim 9.49 would follow

immediately, and could be simplified slightly, but because we are not assuming this we have to do a bit more work.

Let $W \in \Delta(Q_0)$ be a cube which contains x. If $\operatorname{diam} W < L^{-1}\, 2^{j_1 - 1}$, then we have that

$$h(W) \subseteq B_N(h(x), 2^{j_1 - 1}) \,, \tag{9.51}$$

since h is L-Lipschitz by hypothesis (as in Standing Assumptions 3.1). In particular we have that

$$h(W) \subseteq h(T_{j_1}(x)) \cap h(T_{j_1}(y)) \,, \tag{9.52}$$

by (9.43) and (9.47).

On the other hand, we can choose W to be as large as possible subject to the conditions above, and this ensures that

$$\operatorname{diam} W \ge C'^{-1}\, L^{-1}\, 2^{j_1 - 1} \,, \tag{9.53}$$

where C' depends only on n and the Ahlfors-regularity constant for M. If A_1 is large enough, depending on n, L, C_2, and the Ahlfors-regularity constant for M, then this implies that W must be a neighbor of Q, using A_1 for the neighborly constant in (2.30), (2.31). For this assertion we also use (9.41) and the fact that x lies in $2Q$, by construction. Because $Q \in \mathcal{M}_{A_1}(\zeta_1)$, we may apply b) in Definition 7.1 to obtain that

$$\frac{|h(W)|}{|W|} \ge (1 + \zeta_1)^{-1}\, \frac{|h(Q)|}{|Q|} \,. \tag{9.54}$$

We can convert this into

$$|h(W)| \ge C''^{-1}\, |h(Q)| \,, \tag{9.55}$$

where C'' depends only on n, L, C_2, and the Ahlfors-regularity constant for M, because of (9.53) and (9.41). Claim 9.49 now follows by combining (9.52) and (9.55).

Remark 9.56. We mentioned before in Remark 9.6 that it is sometimes convenient to weaken Condition 9.1 by replacing (9.4) with a condition like (9.7), and that such a change would be innocuous for the purposes of this section. Indeed, it is only in the proof of Claim 9.49 that this modification would have more than a cosmetic affect. The crucial point would come in (9.52); now we would only be able to say that $h(W)$

minus a set of small measure is contained in $h(T_{j_1}(x)) \cap h(T_{j_1}(y))$. More precisely, by choosing the constant a in Remark 9.6 correctly, one could still say that at least half of $h(W)$ (in terms of measure) is contained in $h(T_{j_1}(x)) \cap h(T_{j_1}(y))$, and this would be perfectly fine for Claim 9.49. One would simply have to change the constant in (9.50) slightly.

Let us be a bit careful about why one can choose a in this way (*i.e.*, with a depending only on the constants given at the beginning). From a) and b) in Definition 7.1 we have the lower bound

$$\frac{|h(W)|}{|W|} \ge (1 + \zeta_1)^{-1} \, \delta \ge \frac{\delta}{2} \, . \tag{9.57}$$

As in (9.53), we also have a lower bound for the diameter of W, in terms of a constant multiple of 2^{j_1}. This leads to the lower bound

$$|W| \ge C_0^{-1} \, L^{-n} \, 2^{j_1 n} \, , \tag{9.58}$$

where C_0 depends only on n and the Ahlfors-regularity constant for M. Thus

$$|h(W)| \ge C_0^{-1} \, \frac{\delta}{2} \, L^{-n} \, 2^{j_1 n} \, . \tag{9.59}$$

Therefore, if we choose a so that

$$a \le C_0^{-1} \, \frac{\delta}{4} \, L^{-n} \, , \tag{9.60}$$

then we would be able to use (9.51) and (9.7) to conclude that

$$h(T_{j_1}(x)) \cap h(T_{j_1}(y)) \tag{9.61}$$

contains at least half of the elements of $h(W)$, in terms of measure, instead of (9.52) as before. This type of condition on a is completely acceptable, in that it depends only on the parameters that are given to us in advance, rather than anything like C_2 or η from Condition 9.1, which would lead to circles in the argument. (The constant C_2 does come into the later estimates, as in the proof of Claim 9.49, but this does not matter, because it is not connected to the choice of a.)

To finish the proof of Proposition 9.36, we want to use Claim 9.49 to get a contradiction. The basic idea is that the substantial overlap in (9.50) is incompatible with the approximate preservation of measure,

in much the same way as in the discussion at the beginning of Section 6.

Much of the argument will be similar to the one in the proof of Lemma 9.11. We begin by choosing a cube Q_1 such that

$$Q \subseteq Q_1 \subseteq Q_0 \qquad \text{and} \qquad T_{j_1}(x), T_{j_1}(y) \subseteq \widehat{Q}_1 \,, \tag{9.62}$$

where $\widehat{Q}_1$ is defined through (6.8) and (6.7). The cube Q_0 already satisfies these conditions, but by choosing Q_1 as small as possible we can guarantee that

$$\operatorname{diam} Q_1 \le C \operatorname{diam} Q \,, \tag{9.63}$$

where C depends only on n and the Ahlfors-regularity constant for M. One does not need C to depend on C_2 – although that would not really matter – because the requirement that $T_{j_1}(x)$ and $T_{j_1}(y)$ be disjoint, as in (9.40), ensures that the diameters of $T_{j_1}(x)$ and $T_{j_1}(y)$ are not too large compared to the diameter of Q. This also uses the definition (9.2) to know that x and y are roughly in the middle of $T_{j_1}(x)$ and $T_{j_1}(y)$, so that the distance between x and y controls the diameters of $T_{j_1}(x)$ and $T_{j_1}(y)$ up to a bounded factor. (A key occurrence of C_2 is in the choice of j_1, as in (9.41), which goes in sort of the opposite direction from the choice of Q_1.)

From (9.62) and (9.63) it follows that Q_1 is a neighbor of Q, with constant A_1, at least if A_1 is large enough. Since $Q \in \mathcal{M}_{A_1}(\zeta_1)$, we may apply c) in Definition 7.1 to get that

$$(1 + \zeta_1)^{-1} \, \frac{|h(Q)|}{|Q|} \le \frac{|h(\widehat{Q}_1)|}{|\widehat{Q}_1|} \,. \tag{9.64}$$

We would like to get an upper bound for $|h(\widehat{Q}_1)|$ which contradicts this lower bound.

From (9.40) and (9.62) we know that $T_{j_1}(x)$ and $T_{j_1}(y)$ are disjoint subsets of $\widehat{Q}_1$. Set

$$E = \widehat{Q}_1 \setminus (T_{j_1}(x) \cup T_{j_1}(y)) \,. \tag{9.65}$$

Thus

$$h(\widehat{Q}_1) = h(E) \cup h(T_{j_1}(x)) \cup h(T_{j_1}(y)) \,, \tag{9.66}$$

and

$$|h(\widehat{Q}_1)| \le |h(E)| + |h(T_{j_1}(x))| + |h(T_{j_1}(y))| - |h(T_{j_1}(x) \cap T_{j_1}(y))| \,. \tag{9.67}$$

(Compare with (6.5).)

By definition, $T_{j_1}(x)$ and $T_{j_1}(y)$ are unions of cubes in Δ_{j_1}. (See (9.2).) Let us check that

$$\widehat{Q}_1 \text{ is a union of cubes in } \Delta_{j_1} \,. \tag{9.68}$$

Let $j(Q_1)$ denote the largest integer such that $Q_1 \in \Delta_{j(Q_1)}$. Thus $j(Q_1) \ge j(Q)$, since $Q_1 \supseteq Q$, as in (9.62). We also have that

$$\widehat{Q}_1 \text{ is a union of cubes in } \Delta_{j(Q_1)} \,, \tag{9.69}$$

by the construction of $\widehat{Q}_1$ in (6.8) and (6.7). (Strictly speaking, we use here the fact that $Q_1 \subseteq Q_0$, to know that $j(Q_1) \le j(Q_0)$, where $j(Q_0)$ is the largest integer such that $Q_0 \in \Delta_{j(Q_0)}$.) Since $j(Q_1) \ge j(Q)$, as noted above, and $j(Q) \ge j_1$, as in (9.41), we have that $j(Q_1) \ge j_1$. Thus (9.68) follows from (9.69), by the usual properties of cubes (namely, (2.16) and (2.17)).

Since $T_{j_1}(x)$, $T_{j_1}(y)$, and $\widehat{Q}_1$ are all unions of cubes in Δ_{j_1}, and since the cubes in Δ_{j_1} are pairwise disjoint by (2.16), we conclude that E is also a union of cubes in Δ_{j_1}.

Let $\mathcal{U}(X)$ denote the collection of cubes $W \in \Delta_{j_1}$ such that $W \subseteq X$, where X is $T_{j_1}(x)$, $T_{j_1}(y)$, $\widehat{Q}_1$, or E. Then

$$\begin{aligned} &|h(E)| + |h(T_{j_1}(x))| + |h(T_{j_1}(y))| \\ &\le \sum_{W \in \mathcal{U}(E)} |h(W)| + \sum_{W \in \mathcal{U}(T_{j_1}(x))} |h(W)| + \sum_{W \in \mathcal{U}(T_{j_1}(y))} |h(W)| \,, \end{aligned} \tag{9.70}$$

since E, $T_{j_1}(x)$, and $T_{j_1}(y)$ are each given by the union of the cubes in $\mathcal{U}(E)$, $\mathcal{U}(T_{j_1}(x))$, and $\mathcal{U}(T_{j_1}(y))$. On the other hand, the collections $\mathcal{U}(E)$, $\mathcal{U}(T_{j_1}(x))$, and $\mathcal{U}(T_{j_1}(y))$ are pairwise disjoint, since E, $T_{j_1}(x)$, and $T_{j_1}(y)$ are pairwise disjoint (by (9.65) and (9.40)). We also have that

$$\mathcal{U}(\widehat{Q}_1) = \mathcal{U}(E) \cup \mathcal{U}(T_{j_1}(x)) \cup \mathcal{U}(T_{j_1}(y)) \,. \tag{9.71}$$

In other words, every cube $W \in \Delta_{j_1}$ which is a subset of $\widehat{Q}_1$ must also be a subset of one of E, $T_{j_1}(x)$, and $T_{j_1}(y)$. This uses the fact that $\widehat{Q}_1$ is the union of E, $T_{j_1}(x)$, and $T_{j_1}(y)$ (by (9.65)), the earlier observation that each of E, $T_{j_1}(x)$, and $T_{j_1}(y)$ is a union of cubes in Δ_{j_1}, and the pairwise-disjointness of cubes in Δ_{j_1}. Because of (9.71) and the disjointness of the union in (9.71) we may convert (9.70) into

$$|h(E)| + |h(T_{j_1}(x))| + |h(T_{j_1}(y))| \le \sum_{W \in \mathcal{U}(\widehat{Q}_1)} |h(W)| \,. \tag{9.72}$$

If A_1 is large enough, depending on n, C_2, and the Ahlfors-regularity constant, then the cubes $W \in \mathcal{U}(\widehat{Q}_1)$ are all neighbors of Q, with neighborly constant A_1. (This is the last time that we impose a condition on A_1, and so it may now be chosen and fixed, once and for all.) This permits us to apply b) in Definition 7.1 to get that

$$\frac{|h(W)|}{|W|} \le (1 + \zeta_1) \, \frac{|h(Q)|}{|Q|} \,, \tag{9.73}$$

for all $W \in \mathcal{U}(\widehat{Q}_1)$. Thus (9.72) leads to

$$|h(E)| + |h(T_{j_1}(x))| + |h(T_{j_1}(y))| \le (1 + \zeta_1) \, \frac{|h(Q)|}{|Q|} \sum_{W \in \mathcal{U}(\widehat{Q}_1)} |W| \,. \tag{9.74}$$

The cubes W in $\mathcal{U}(\widehat{Q}_1)$ are pairwise disjoint, since they all lie in Δ_{j_1}, and they are also subsets of $\widehat{Q}_1$, by definition of $\mathcal{U}(\widehat{Q}_1)$. Thus we conclude that

$$|h(E)| + |h(T_{j_1}(x))| + |h(T_{j_1}(y))| \le (1 + \zeta_1) \, \frac{|h(Q)|}{|Q|} \, |\widehat{Q}_1| \,. \tag{9.75}$$

Combining (9.75) with (9.67) yields

$$|h(\widehat{Q}_1)| \le (1 + \zeta_1) \, \frac{|h(Q)|}{|Q|} \, |\widehat{Q}_1| - |h(T_{j_1}(x) \cap T_{j_1}(y))| \,. \tag{9.76}$$

We can also apply (9.50) to get that

$$|h(\widehat{Q}_1)| \le (1 + \zeta_1) \, \frac{|h(Q)|}{|Q|} \, |\widehat{Q}_1| - c \, |h(Q)| \,, \tag{9.77}$$

where c depends only on n, C_2, L, and the Ahlfors-regularity constant for M.

Let us rewrite (9.77) as

$$\frac{|h(\widehat{Q}_1)|}{|\widehat{Q}_1|} \le (1+\zeta_1)\,\frac{|h(Q)|}{|Q|} - c\,\frac{|h(Q)|}{|\widehat{Q}_1|}\,. \tag{9.78}$$

Putting this upper bound together with the lower bound in (9.64), we obtain that

$$(1+\zeta_1)^{-1}\,\frac{|h(Q)|}{|Q|} \le (1+\zeta_1)\,\frac{|h(Q)|}{|Q|} - c\,\frac{|h(Q)|}{|\widehat{Q}_1|}\,. \tag{9.79}$$

This is equivalent to

$$(1+\zeta_1)^{-1} \le (1+\zeta_1) - c\,\frac{|Q|}{|\widehat{Q}_1|}\,. \tag{9.80}$$

(Note that $|h(Q)| > 0$, since $Q \in \mathcal{M}_{A_1}(\zeta_1)$. See a) in Definition 7.1.)

On the other hand,

$$|\widehat{Q}_1| \le D\,|Q|\,, \tag{9.81}$$

for some constant D which depends only on n and the Ahlfors-regularity constant for M, by (9.63) and the usual properties of cubes (*i.e.*, (2.18)). Thus (9.80) leads to

$$(1+\zeta_1)^{-1} \le (1+\zeta_1) - c\,D^{-1}\,. \tag{9.82}$$

This gives us the desired contradiction if we choose ζ_1 small enough, depending on c and D. In other words, $c\,D^{-1}$ is a positive number of definite size, depending only on n, C_2, L, and the Ahlfors-regularity constant for M, and since we are free to choose ζ_1 as small as we want, depending on these parameters, we may choose it so that (9.82) is not true.

This completes the proof of Proposition 9.36. We shall make use of it in the next section.

10. A summarizing theorem.

Again the provisions of Standing Assumptions 3.1 are in force. Thus M and N are metric spaces, with M Ahlfors-regular of dimension n, $\{\Delta_j\}_{j<j_0}$ is a sequence of families of cubes in M, as in Section 2.3, Q_0 is a fixed cube in M, and $h : Q_0 \longrightarrow N$ is an Lipschitz mapping with norm at most L. Also, we use $|A|$ to denote n-dimensional Hausdorff measure of A, whether A is contained in M or N. (To some extent one could allow other measures besides Hausdorff measure, as in Remark 3.5.)

Theorem 10.1. *Notations and assumptions as above. Suppose also that Condition* 9.1 *holds. Let $\delta > 0$ be arbitrary, but fixed, and set*

$$\mathcal{SI}_1 = \{W \in \Delta : \ W \subseteq Q_0, \, |h(W)| \le \delta |W|\} \, . \tag{10.2}$$

Also put

$$D = Q_0 \setminus \bigcup_{W \in \mathcal{SI}_1} W \, . \tag{10.3}$$

Then

$$|h(Q_0 \setminus D)| < \delta \, |Q_0 \setminus D| < \delta \, |Q_0| \tag{10.4}$$

(at least if $Q_0 \setminus D \neq \varnothing$), and h is (Q_0, D)-weakly bilipschitz, in the sense of Definition 8.8. *The constants for the (Q_0, D)-weak bilipschitz property may be taken to depend only on n, L, δ, the Ahlfors-regularity constant for M, and the constants C_2, η in Condition* 9.1, *associated to $C_1 = 1$ and $\gamma = \delta/2$.*

In particular, for each $\varepsilon > 0$ one can find measurable subsets $F_1, \ldots, F_\ell$ of D such that

$$\left| D \setminus \bigcup_{i=1}^{\ell} F_i \right| < \varepsilon \, |Q_0| \tag{10.5}$$

and so that the restriction of h to each F_i is bilipschitz, with ℓ and the bilipschitz constants bounded by quantities which depend only on ε and the parameters mentioned above.

Note that

$$\left| h(Q_0) \setminus \bigcup_{i=1}^{\ell} h(F_i) \right| < (L^n \, \varepsilon + \delta) \, |Q_0| \, , \tag{10.6}$$

because of (10.4), (10.5), and (2.3). This is like the situation in [19]; the $h(F_i)$'s account for all of $h(Q_0)$ except for a subset of small measure. If $h(Q_0)$ itself is of small measure, then Theorem 10.1 does not really contain any information, but as soon as the measure of the image is of definite size, one obtains substantial subsets of Q_0 on which h is bilipschitz, as in [6].

To prove the theorem we basically only have to concatenate pieces from the previous sections. The inequality (10.4) is quite automatic, and we have done this type of calculation several times before. (See Lemma 7.5, for instance.) For the weak bilipschitz condition, choose k as in Proposition 9.36. Let $\mathcal{BL}'(k)$ be as in Section 8 (see (8.7)), and let $\Delta(Q_0, D)$ denote the collection of cubes in M which are contained in M and which intersect D. We want to show that

$$\Delta(Q_0, D) \backslash \mathcal{BL}'(k) \text{ is a Carleson set} \, , \tag{10.7}$$

which is slightly stronger than needed for Definition 8.8.

According to Proposition 9.36,

$$\mathcal{M}_{A_1}(\zeta_1) \subseteq \mathcal{BL}'(k) \tag{10.8}$$

when A_1 is large enough and ζ_1 is small enough (and with suitable bounds). To establish (10.7), it is therefore enough to show that

$$\Delta(Q_0, D) \backslash \mathcal{M}_{A_1}(\zeta_1) \text{ is a Carleson set} \, . \tag{10.9}$$

This we can get from Proposition 7.8. (The A-parameter was left implicit in Proposition 7.8, *i.e.*, we used the notation $\mathcal{M}(\zeta)$ instead of $\mathcal{M}_A(\zeta)$.) More precisely, Proposition 7.8 provides a Carleson condition for

$$\Delta(Q_0) \backslash (\mathcal{M}_{A_1}(\zeta_1) \cup \mathcal{SI}) \, , \tag{10.10}$$

where $\mathcal{SI}$ was defined (in Definition 7.4) to be the set of cubes $Q \subseteq Q_0$ such that $Q \subseteq W$ for some $W \in \mathcal{SI}_1$. If Q is a cube in $\Delta(Q_0, D)$, then

Q intersects D, and therefore cannot be contained in some $W \in \mathcal{SI}_1$. Thus

$$\Delta(Q_0, D) \backslash \mathcal{M}_{A_1}(\zeta_1) \subseteq \Delta(Q_0) \backslash (\mathcal{M}_{A_1}(\zeta_1) \cup \mathcal{SI}) \,, \tag{10.11}$$

and so (10.9) follows from the Carleson condition for (10.10).

This proves that h is (Q_0, D)-weakly bilipschitz in the sense of Definition 8.8, and with suitable bounds. The last part of Theorem 10.1 follows directly from this and Proposition 8.11. This completes the proof of Theorem 10.1.

Remark 10.12. Let us mention a simple extension of the preceding theorem and proof. We want to "weaken" Condition 9.1, by allowing some exceptions which are controlled by a packing condition, in the following manner.

Let Q_0 be a fixed cube in M, as in Condition 9.1, and let $\mathcal{B}$ be a collection of cubes in Δ which are also subcubes of Q_0. These will be the set of "bad" cubes for Condition 9.1. Given a point $x \in Q_0$ and an integer $j < j_0$, we shall allow the pair (x, j) to be "excused" from the provisions of Condition 9.1 if there is a cube $Q \in \mathcal{B}$ such that

$$\text{there is a } W \in \mathcal{B} \text{ such that } x \in W \text{ and } W \in \Delta_j \,. \tag{10.13}$$

To prevent this from happening too often, we ask that $\mathcal{B}$ satisfy a packing condition of the form

$$\sum_{W \in \mathcal{B}} |W| \le C' \, |Q_0| \,. \tag{10.14}$$

We should be a bit more careful here about the role of $\mathcal{B}$ in Condition 9.1. We allow $\mathcal{B}$ and the constant C' in (10.14) to depend on C_1 and γ, but not on anything else. In other words, $\mathcal{B}$ and C' are in roughly the same category as C_2 and η, and should be given at the same time. The rest of Condition 9.1 remains the same as before, except that the pairs (x, j) which satisfy (10.13) are excused from the conclusions of Condition 9.1. One could rephrase this by treating (10.13) as a third alternative, in addition to the two (9.4), (9.5) that are already there.

If Condition 9.1 is weakened in this manner, then the conclusions of Theorem 10.1 still hold, and with essentially the same proof. The main point is that Proposition 9.36 should be modified, to say that the cube Q in the statement of Proposition 9.36 either lies in $\mathcal{BL}'(k)$, as before,

or is a neighbor of a cube in $\mathcal{B}$, for a suitable choice of constant A in the neighbor-conditions (2.30), (2.31). This choice of A may depend on C_2, as well as n and the Ahlfors-regularity constant for M. With this adjustment, Proposition 9.36 is derived in practically the manner as in Section 9. The only difference comes in the justification of Claim 9.42, which is where Condition 9.1 was used. Now the weaker version of Condition 9.1 would imply that either the given cube Q is a neighbor of a cube in $\mathcal{B}$ (for a sufficiently large neighbor-constant A), or that the same conclusions as in Claim 9.42 are true, so that the rest of the argument can be finished in the same way as before. (In particular, Condition 9.1 was not employed again after Claim 9.42.)

For the proof of Theorem 10.1, one would then replace (10.7) with the requirement that $\Delta(Q_0, D)\backslash\mathcal{BL}'(k)$ satisfy a packing condition, and one would replace (10.8) with

$$\mathcal{BL}'(k)\backslash\mathcal{M}_{A_1}(\zeta_1) \text{ satisfies a packing condition}. \tag{10.15}$$

The latter is what one would get from the modified version of Proposition 9.36, and it is sufficient for the conclusions of Theorem 10.1, for the same reasons as before.

As usual, one should not be too concerned with the difference between Carleson and packing conditions here, and indeed in practice the set $\mathcal{B}$ of bad cubes could well be a Carleson set anyway. One could also consider weakenings of the packing condition, as may be appropriate in some circumstances, but this is easy to analyze and we shall not pursue it here.

11. Some technical extensions.

In this section we would like to record some modest refinements of the assertions in this paper, concerning the Lipschitz condition on our initial mapping $h : Q_0 \longrightarrow N$ from Standing Assumptions 3.1.

Instead of the Lipschitz condition (3.2), consider the requirements that

$$h : Q_0 \longrightarrow N \text{ be continuous} \tag{11.1}$$

and

$$H^n(h(E)) \le L^n\, H^n(E), \qquad \text{for all } E \subseteq Q_0\ . \tag{11.2}$$

Each of these properties holds automatically when h is Lipschitz with constant L, using (2.3) to get (11.2). (For (11.2), one might also be interested in different measures besides Hausdorff measure, as in Remark 3.5).

Notice first that Lemma 2.4 works for mappings which satisfy (11.1) and (11.2) instead of being Lipschitz. Indeed, all that one really needs for Lemma 2.4 is that the given mapping send compact sets to compact sets (which follows from continuity), and that sets of measure 0 are sent to sets of measure 0 (which is a special case of (11.2)).

With (11.1) and (11.2) instead of the Lipschitz condition (3.2) for h, the statement and proof of Proposition 3.6 go through exactly as before. The main points are that one still has (3.12), now by fiat, and that Lemma 3.16 continues to work under these conditions. Specifically, (3.19) in Lemma 3.16 still follows from (3.18), by (11.2), and the finiteness of $|f(Q)|$ mentioned just after (3.22) also follows from (11.2). (In the context of the proof of Proposition 3.6, Lemma 3.16 was always applied with the mapping f taken to be h, and with the cube Q a subcube of our original cube Q_0.)

None of this should be considered surprising, as Proposition 3.6 relies only on very general principles, in which the mapping h plays little role. (One really only uses h to get the subadditive measure $E \longmapsto |h(E)|$ for subsets of M.)

Similarly, the extension of Proposition 3.6 to propositions 4.2 and 5.5 does not involve h at all, beyond what is incorporated into Proposition 3.6 already, and so they also work in this more general setting. In Proposition 6.13, the mapping h does participate slightly, but only for subadditivity of $E \longmapsto |h(E)|$ again. This does not take place until the last parts of the proof, after all of the initial reductions. For Proposition 7.8 the mapping h plays essentially no active role either, and the proof in effect provides merely a different view of some of the information given in Proposition 6.13.

To summarize, we have that the statements and proofs of Propositions 3.6, 4.2, 5.5, 6.13, and 7.8 work just as well with the Lipschitz condition (3.2) for h replaced with (11.1) and (11.2). Actually, we never even need (11.1) here. It is nice to have, for the sake of measurability of the images, as in Lemma 2.4 and its extension mentioned above, but one does not really need measurability for these propositions. This is because Hausdorff measure H^n is defined as an outer measure on all subsets of N, and subadditivity of $E \longmapsto |h(E)|$ is all that was ever used for these propositions. (Measurability of the images is needed for

the *interpretation* of measure-preserving behavior mentioned near the beginning of Section 6, concerning approximate disjointness of images of disjoint sets, but this did not enter into the proofs of the propositions. It does play a role in later applications, as in (9.67) in the proof of Proposition 9.36.)

For the work in sections 9 and 10, it is important to have Lipschitz control on h, and not just (11.1) and (11.2). However, the nature of this control can be weakened, as in the following notion.

Definition 11.3 (Pseudo-Lipschitz condition)**.** *Let M, N, and Q_0 be as usual, in Standing Assumptions* 3.1*. We shall call $h : Q_0 \longrightarrow N$ pseudo-Lipschitz with constant L if h is continuous on Q_0, and if there is a subset Y of Q_0 such that the restriction of h to Y is Lipschitz with constant L, and such that*

$$H^n(h(Q_0 \backslash Y)) = 0\,. \tag{11.4}$$

In other words, h is L-Lipschitz on Y, and h is completely degenerate on $Q_0 \backslash Y$, in the sense of (11.4). This comes up naturally in some situations, where one starts with some mapping and tries to "clean it up" by collapsing portions that are not essential (*e.g.*, which are not needed for some topological purpose). In doing this it may not be convenient or possible to keep track of the Lipschitz behavior of the given mapping on the whole domain, but, as in the pseudo-Lipschitz property, one may not need a bound on the parts where collapsing takes place.

If $h : Q_0 \longrightarrow N$ is pseudo-Lipschitz with constant L, then it automatically satisfies (11.1) and (11.2), the latter by (2.3) and (11.4). In particular, Propositions 3.6, 4.2, 5.5, 6.13, and 7.8 continue to hold in these circumstances.

We want to extend the results of sections 9 and 10 to the case where $h : Q_0 \longrightarrow N$ is pseudo-Lipschitz with constant L as well. Let us first make some observations about Section 8, and weak bilipschitz conditions.

Let M and N be as usual (in Standing Assumptions 3.1), and fix a cube T in M. Also fix a measurable subset S of T. Instead of a mapping $f : T \longrightarrow N$, as in Section 8, let us take f to be defined only on S. (If it happens to be defined on all of T, then we simply forget about the part that is not in S.)

Fix a constant $b > 0$ as in Lemma 8.2, and let k be a number greater than 1. Define $\mathcal{BL}'_0(k)$ to be the collection of cubes $Q \in \Delta$ such

that $Q \subseteq T$ and

$$k^{-1}\, d(x,y) \le \rho(f(x), f(y)) \le k\, d(x,y) \quad \text{for all } x, y \in 2\,Q \cap S \text{ with } d(x,y) > b \operatorname{diam} Q\,. \tag{11.5}$$

This is the same as (8.7), except that we restrict ourselves to points x, y in S.

Definition 11.6 ((T, D, S)-weakly bilipschitz mappings)**.** *Given a measurable set $D \subseteq T$, we say that $f : S \longrightarrow N$ is (T, D, S)-weakly bilipschitz if exactly the same conditions hold as in Definition* 8.8*, except that $\mathcal{BL}'(k)$ is replaced with $\mathcal{BL}'_0(k)$.*

If $f : S \longrightarrow N$ is (T, D, S)-weakly bilipschitz, then exactly the same conclusions as in Proposition 8.11 are true, except that the F_i's should now be subsets of $D \cap S$, and the D in (8.12) should be replaced with $D \cap S$. This can be proved in practically the same manner as before, except that now one only worries about points in S. More precisely, the choice of "bad set" E_λ of points to remove can be made in exactly the same manner as before, in (8.16). One then wants to show that

$$(D \cap S) \backslash E_\lambda \tag{11.7}$$

can be covered by a bounded number of sets F_i on which f is k-bilipschitz, as in the statement around (8.19). As before, to get the bilipschitz condition, it suffices to choose the F_i's so that (8.20) holds. Thus one only needs to choose the F_i's so that they cover $(D \cap S) \backslash E_\lambda$ and satisfy (8.20) (and so that there are only boundedly many of them), and this can be accomplished through exactly the same kind of coding argument as for Proposition 8.11.

Let us now proceed to the material in sections 9 and 10. We use the same notations and assumptions as before, including the ones in Standing Assumptions 3.1, except that $h : Q_0 \longrightarrow N$ is now required to be pseudo-Lipschitz with constant L, instead of the usual Lipschitz condition (3.2). Let $Y \subseteq Q_0$ be as in Definition 11.3. We may as well assume that Y is relatively closed in Q_0, since otherwise we can simply replace it with its relative closure.

We do not change the formulation of David's condition (Condition 9.1) in this context, and we also keep Standing Assumptions 3.1 as they are. The weakening of David's condition described in Remark 9.6 would also work fine here.

The statements and proofs of lemmas 9.10, 9.11, and 9.31, carry over to this context without trouble, in much the same way as for propositions 3.6, 4.2, 5,5, 6.13, and 7.8 before. For this the pseudo-Lipschitz condition could be replaced with (11.1) and (11.2) (or just (11.2)).

For Proposition 9.36, one should be a bit more careful. In the statement of Proposition 9.36 one should replace $\mathcal{BL}'(k)$ with $\mathcal{BL}'_0(k)$ as defined above, around (11.5). (For the definition of $\mathcal{BL}'_0(k)$, one should now take the cube T to be Q_0, f to be h, and S to be Y.)

The proof of the modified version of Proposition 9.36 begins the same way as before, except that x and y should be chosen in $2\,Q \cap Y$, rather than $2\,Q \cap Q_0$. The pseudo-Lipschitz condition on h still gives (9.39) in that case, since x and y lie in Y.

Proceeding with the earlier proof, the choices of $j(Q)$ and j_1 remain the same as before, and one can derive Claim 9.42 from lemmas 9.10, 9.11, and 9.31 in exactly the same manner as in Section 9.

For Claim 9.49 we have to be a bit more careful. If N happens to be Ahlfors regular of dimension n, then there is nothing to do, and Claim 9.49 would follow directly from Claim 9.42. Without that assumption we can try to argue as before, but now (9.51) need not be true, since we only have a pseudo-Lipschitz condition for h. Instead of (9.51) we have that

$$h(W \cap Y) \subseteq B_N(h(x), 2^{j_1 - 1}) \,, \tag{11.8}$$

because h is L-Lipschitz on Y. This is practically as good as (9.51), since

$$|h(W \backslash Y)| = 0 \,, \tag{11.9}$$

by (11.4), and therefore

$$|h(W \cap Y)| = |h(W)| \,. \tag{11.10}$$

From here the argument for Claim 9.49 is nearly the same as in Section 9. One should replace W in the left side of (9.52) with $W \cap Y$, but the measure-theoretic computations do not change, because of (11.10). (Similarly, the considerations of Remark 9.56 extend to the present circumstances as well.)

The rest of the proof of Proposition 9.36 carries over without much incident, and indeed would work under the assumptions of (11.1) and

(11.2) in place of the (stronger) pseudo-Lipschitz property. The one slightly delicate point is that for this part of the argument we do need the measurability of the images of measurable sets under h (for (9.67)), and so we do use now the continuity condition (11.1).

This takes care of the material in Section 9. As for Section 10, Theorem 10.1, goes over with only minor adjustments. Specifically, instead of the (Q_0, D)-weak bilipschitz property for h in the conclusions of Theorem 10.1, one would get (Q_0, D, Y)-weak bilipschitzness, in the sense of Definition 11.6, and with the same estimates as before. In (10.5) one should replace D with $D \cap Y$, as in the variant of Proposition 8.11 for (T, D, S)-weak bilipschitzness discussed above (in the paragraph containing (11.7)). These are the changes that one should make to the statement of Theorem 10.1, and then the proof is almost exactly the same as in Section 10. One simply has to use $\mathcal{BL}'_0(k)$ (from (11.5)) instead of $\mathcal{BL}'(k)$, and employ the versions of propositions 7.8, 8.11, and 9.36 adapted to this context (*i.e.*, with the pseudo-Lipschitz condition instead of the Lipschitz condition).

Thus Theorem 10.1 can be extended to the case where h is pseudo-Lipschitz with constant L instead of Lipschitz. Note that the inequality (10.6) does still work in this case, *i.e.*, without having to replace D with $D \cap Y$, since $|h(Q_0 \backslash Y)| = 0$ (as part of the pseudo-Lipschitz condition, as in (11.4)). In particular, given a suitable lower bound for $|h(Q_0)|/|Q_0|$. we may conclude that there is a subset of Q_0 of definite size on which h is bilipschitz, and with uniform bounds.

Of course David's method in [6] also works perfectly well in this context. (Since we happen to be here, though, it is convenient to go through the verifications for the arguments described in this paper.)

Remark 11.11. As one last comment, let us mention that if N is Ahlfors regular of dimension n, then one could simplify these extensions slightly, as follows. One might as well split the issue of the upper bounds in the bilipschitz conditions off from the lower bounds, and just concentrate on the latter. If one does have upper bounds, as in the pseudo-Lipschitz condition, one can simply add that on afterwards, separately.

In concentrating on the lower bounds on the bilipschitz conditions, one should drop the upper bounds from most of the hypotheses and conclusions in the various statements. Thus, instead of weak bilipschitz conditions in Section 8, one would work with similar conditions which involve only lower bounds in (8.5), and (8.7), and only lower bounds

in Proposition 8.11. Analogous changes would be made to Proposition 9.36 and Theorem 10.1. With these modifications it would be enough to use (11.1) and (11.2), rather than the pseudo-Lipschitz condition (at least if N is Ahlfors regular of dimension n, to avoid trouble with Claim 9.49 in the proof of Proposition 9.36.)

Acknowledgements. The author is grateful to G. David for some helpful comments and suggestions.

References.

[1] Assouad, P., *Espaces Métriques, Plongements, Facteurs.* Thèse de Doctorat (1977), Université de Paris XI.

[2] Assouad, P., Étude d'une dimension métrique liée à la possibilité de plongement dans $\mathbb{R}^n$. *C. R. Acad. Sci. Paris* **288** (1979), 731-734.

[3] Assouad, P., Plongements Lipschitziens dans $\mathbb{R}^n$. *Bull. Soc. Math. France* **111** (1983), 429-448.

[4] Christ, M., Lectures on Singular Integral Operators. *CBMS Regional Conference Series in Math.* **77** (1990) American Mathematical Society.

[5] David, G., Opérateurs d'intégrale singulière sur les surfaces régulières. *Ann. Sci. École. Norm. Sup.* **21** (1988), 225-258.

[6] David, G., Morceaux de graphes lipschitziens et intégrales singulières sur un surface. *Revista Mat. Iberoamericana* **4** (1988), 73-114.

[7] David, G., *Wavelets and Singular Integrals on Curves and Surfaces.* Lecture Notes in Math. **1465** (1991), Springer-Verlag.

[8] David, G., Jerison, D., Lipschitz approximations to hypersurfaces, harmonic measure, and singular integrals. *Indiana Univ. Math. J.* **39** (1990), 831-845.

[9] David, G., Semmes, S., *Singular Integrals and Rectifiable Sets in* $\mathbb{R}^n$: *au-delà des graphes lipschitziens.* Astérisque **193**, Société Mathématique de France, 1991.

[10] David, G., Semmes, S., Quantitative rectifiability and Lipschitz mappings. *Trans. Amer. Math. Soc.* **337** (1993), 855-889.

[11] David, G., Semmes, S., *Analysis of and on Uniformly Rectifiable Sets.* American Mathematical Society, 1993.

[12] David, G., Semmes, S., *Fractured Fractals and Broken Dreams: Self-Similar Geometry through Metric and Measure.* Oxford University Press, 1997.

[13] David, G., Semmes, G., Uniform rectifiability and quasiminimizing sets of arbitrary codimension. *Mem. Amer. Math. Soc.* **687** (2000).

[14] Dorronsoro, J., A characterization of potential spaces. *Proc. Amer. Math. Soc.* **95** (1985), 21-31.

[15] Falconer, K., *The Geometry of Fractal Sets.* Cambridge University Press, 1984.

[16] Federer, H., *Geometric Measure Theory.* Springer-Verlag, 1969.

[17] Garnett, J., *Bounded Analytic Functions.* Academic Press, 1981.

[18] Journé, J.-L., *Calderón-Zygmund Operators, Pseudodifferential Operators, and the Cauchy Integral of Calderón.* Lecture Notes in Math. **994** (1983), Springer-Verlag.

[19] Jones, P., Lipschitz and bi-Lipschitz functions. *Revista Mat. Iberoamericana* **4** (1988), 115-122.

[20] Mattila, P., *Geometry of Sets and Measures in Euclidean Spaces.* Cambridge University Press, 1995.

[21] Semmes, S., Analysis vs. geometry on a class of rectifiable hypersurfaces in $\mathbb{R}^n$. *Indiana Univ. Math. J.* **39** (1990), 1005-1035.

[22] Semmes, S., Good metric spaces without good parameterizations. *Revista Mat. Iberoamericana* **12** (1996), 187-275.

[23] Semmes, S., On the nonexistence of bilipschitz parameterizations and geometric problems about A_∞ weights. *Revista Mat. Iberoamericana* **12** (1996), 337-410.

[24] Semmes, S., *Mappings and spaces.* In *Quasiconformal Mappings and Analysis*: *A Collection of Papers Honoring F. W. Gehring.* Edited by P. Duren et al., Springer-Verlag (1998), 347-368.

[25] Semmes, S., *Metric Spaces and Mappings Seen at Many Scales.* Appendix in *Metric Structures for Riemannian and Non-Riemannian Spaces.* M. Gromov et al., Birkhäuser, 1999.

[26] Semmes, S., Real analysis, quantitative topology, and geometric complexity. Preprint.

[27] Stein, E. M., *Singular Integrals and Differentiability Properties of Functions.* Princeton University Press, 1970.

Recibido: 30 de abril de 1.998

Stephen Semmes
Department of Mathematics
Rice University
Houston, Texas 77251, U.S.A.
semmes@math.rice.edu

The author is grateful to the U.S. National Science Foundation for support.