

REGULAR MAPPINGS BETWEEN DIMENSIONS

G. DAVID AND S. SEMMES

Abstract

The notions of Lipschitz and bilipschitz mappings provide classes of mappings connected to the geometry of metric spaces in certain ways. A notion between these two is given by “regular mappings” (reviewed in Section 1), in which some non-bilipschitz behavior is allowed, but with limitations on this, and in a quantitative way. In this paper we look at a class of mappings called (s, t) -regular mappings. These mappings are the same as ordinary regular mappings when $s = t$, but otherwise they behave somewhat like projections. In particular, they can map sets with Hausdorff dimension s to sets of Hausdorff dimension t . We mostly consider the case of mappings between Euclidean spaces, and show in particular that if $f: \mathbf{R}^s \rightarrow \mathbf{R}^n$ is an (s, t) -regular mapping, then for each ball B in \mathbf{R}^s there is a linear mapping $\lambda: \mathbf{R}^s \rightarrow \mathbf{R}^{s-t}$ and a subset E of B of substantial measure such that the pair (f, λ) is bilipschitz on E . We also compare these mappings in comparison with “nonlinear quotient mappings” from [6].

CONTENTS

| | |
|--|-----|
| 1. Regular mappings | 370 |
| 2. Regular mappings between dimensions | 372 |
| 3. Ahlfors-regular spaces | 376 |
| 4. Regular mappings between Euclidean spaces | 380 |
| 5. More classical results about Lipschitz mappings | 382 |
| 6. (m, t) -Regular mappings between Euclidean spaces | 384 |
| 7. Proofs for Theorem 6.1 and Corollary 6.5 | 386 |
| 8. Comparisons with the co-Lipschitz property | 400 |
| 9. Noncollapsing mappings | 402 |
| 10. A variant of Theorem 6.1 | 407 |
| References | 415 |

The second author was supported by the U.S. National Science Foundation. Portions of this work were done during a visit of the second author to the Université de Paris-Sud. The authors are grateful to M. Gromov, concerning a suggestion and a reference. The authors would also like to thank the referee for his or her thoughtful comments and suggestions.

Let us begin with a brief review of “regular mappings” in the ordinary sense, and then proceed to a more general notion that allows for changes in dimension (like Hausdorff dimension).

1. Regular mappings

Let $(M, d(x, y))$ be a metric space. That is, M is a nonempty set, $d(x, y)$ is a nonnegative real-valued function on $M \times M$ which is symmetric in x and y , vanishes exactly when $x = y$, and satisfies the triangle inequality.

Let $(N, \rho(u, v))$ be another metric space, and let $f: M \rightarrow N$ be a mapping. We say that f is *Lipschitz* if there is a constant $C > 0$ such that

$$(1.1) \quad \rho(f(x), f(y)) \leq C d(x, y) \quad \text{for all } x, y \in M.$$

The mapping f is said to be *bilipschitz* if there is a constant $C' > 0$ such that

$$(1.2) \quad C'^{-1} d(x, y) \leq \rho(f(x), f(y)) \leq C' d(x, y) \quad \text{for all } x, y \in M.$$

The notions of Lipschitz and bilipschitz mappings provide ways of making comparisons between different metric spaces. If two metric spaces are *bilipschitz equivalent*—so that there is a bilipschitz mapping from one of the spaces onto the other one—then it means that the two spaces are practically the same in many respects. With Lipschitz mappings one has more flexibility, and the possibility that the spaces are quite different, because of compression of distances.

One might add conditions to prevent distances from being compressed too much, without going all the way to bilipschitzness. For instance, one might ask that the image of the domain under a given mapping have positive mass with respect to some measure, like Hausdorff measure (in some dimension). Compare with [14], starting in Chapter 11.

The notion of a *regular mapping* (as in the next definition) gives another way to put limits on the manner in which distances might be compressed by a Lipschitz mapping.

Definition 1.3. Let $(M, d(x, y))$ and $(N, \rho(u, v))$ be two metric spaces. A mapping $f: M \rightarrow N$ between them is said to be *regular* if it is Lipschitz, and if there is a constant $C > 0$ such that for every ball B in N it is possible to cover $f^{-1}(B)$ by at most C balls in M with radius equal to $C \cdot \text{radius}(B)$.

The notion of regular mappings originally came up in [8], in a slightly different form. See also [14], [29] concerning this version. We shall say a bit more about this in Section 3, just after Lemma 3.3.

For the record, when we refer to a “ball” in a metric space, this means an open ball, unless something else is indicated. In practice, this does not matter too much, though.

Remark 1.4. Recall that a metric space $(M, d(x, y))$ is said to be *doubling* if there is a constant C_1 so that every ball B in M can be covered by at most C_1 balls with radius equal to $\text{radius}(B)/2$. When this is true, one can iterate the condition to say that every ball can be covered by C_1^k balls with radius equal to $\text{radius}(B)/2^k$, $k \in \mathbf{Z}_+$.

If $(M, d(x, y))$ is doubling, then the definition of a regular mapping from M to another metric space N can be simplified, as follows. Instead of asking that $f^{-1}(B)$ admit a covering by at most C balls with radius equal to $C \cdot \text{radius}(B)$, as above, one can ask that $f^{-1}(B)$ admit a covering by a bounded number of balls in M with the same radius as B .

In this paper, metric spaces which are doubling will normally be the primary focus. This includes Euclidean spaces \mathbf{R}^n (with the standard Euclidean metrics), and subsets of these (with the induced metrics).

Examples 1.5. Consider the real line \mathbf{R} with its usual metric. The mapping $f: \mathbf{R} \rightarrow \mathbf{R}$ defined by $f(x) = |x|$ is regular. This is easy to verify. As an extension of this, consider a mapping from \mathbf{R} to \mathbf{R}^2 which maps the two half-lines $(-\infty, 0]$ and $[0, \infty)$ onto rays σ, τ in \mathbf{R}^2 emanating from the origin. Assume that the mapping is linear on $(-\infty, 0]$ and $[0, \infty)$, sending $0 \in \mathbf{R}$ to the origin in \mathbf{R}^2 , and that it has “unit speed”, i.e., the derivative is a vector of norm 1. This mapping is Lipschitz with constant 1, and also regular with bounded constant. In other words, the constant for the regularity condition can be taken to be bounded independently of the angle between σ and τ . As long as σ and τ are distinct, the mapping will be one-to-one, and in fact it will be bilipschitz. However, the bilipschitz constant tends to ∞ as the angle between σ and τ goes to 0, while the regularity constant remains bounded. In the limit, one recovers a copy of the first mapping mentioned above, i.e., $f(x) = |x|$ on \mathbf{R} .

More generally, one can look at locally rectifiable curves in \mathbf{R}^2 (or \mathbf{R}^n), together with parameterizations of them by arclength. These parameterizations define regular mappings exactly when the curves in question are *regular* in the sense of [7] (as well as [1]). This means that the amount of arclength measure of the curve inside a disk of radius r is

bounded by a constant times r . (Compare with Lemma 3.3.) In particular, with this class of examples, it is easy to see how regular mappings can cross themselves in the image numerous times, have cusps in the image, and so on.

Bilipschitz mappings are automatically regular, but the converse is not true in general, as is indicated by the preceding examples. To make a nicer comparison with the notion of regular mappings, one can reformulate the bilipschitz property as follows. If M and N are metric spaces, then a mapping $f: M \rightarrow N$ is bilipschitz if it is Lipschitz, and if $f^{-1}(B)$ is contained in a single ball in M of radius $C \cdot \text{radius}(B)$ for all balls B in N , where C is a constant that does not depend on B .

A basic property of bilipschitz mappings is that they do not increase or decrease Hausdorff measures of a set (in the domain of the mapping) by more than a bounded factor (which may depend on the dimension for the Hausdorff measure that one is using). This is a well-known and straightforward consequence of the definition of Hausdorff measures, as in [16], [17], [25]. The same statement is true for regular mappings, and by nearly the same argument. See also Section 12.1 of [14].

2. Regular mappings between dimensions

Definition 2.1. Let $(M, d(x, y))$ and $(N, \rho(u, v))$ be metric spaces. Also let s, t be nonnegative real numbers, with $s \geq t$. A mapping $f: M \rightarrow N$ is said to be (s, t) -regular if it is Lipschitz, and if there is a constant C so that the following holds: if B_1 is an arbitrary ball in N , and B_2 is an arbitrary ball in M with $\text{radius}(B_2) \geq \text{radius}(B_1)$, then

$$(2.2) \quad f^{-1}(B_1) \cap B_2$$

can be covered by a family of at most

$$(2.3) \quad C \cdot \left(\frac{\text{radius}(B_2)}{\text{radius}(B_1)} \right)^{s-t}$$

balls in M with radius equal to $C \cdot \text{radius}(B_1)$.

If $s = t$, then this is equivalent to the notion of regular mappings from Definition 1.3. In general, the increment $s - t$ allows for mappings which are more like projections, in which the dimension of the image is less than the dimension of the domain.

Examples 2.4. Consider \mathbf{R}^n and the real line \mathbf{R} , with their Euclidean metrics. The mapping

$$(2.5) \quad f: \mathbf{R}^n \rightarrow \mathbf{R} \quad \text{given by} \quad f(x) = |x|$$

is $(n, 1)$ -regular. This is not hard to verify. Also, standard orthogonal coordinate projections from \mathbf{R}^m onto \mathbf{R}^n , $m \geq n$, are (m, n) -regular.

It is easy to make more examples like these. One can also make examples by combining ones like these with bilipschitz mappings and ordinary regular mappings, with crossings and so forth, as in Examples 1.5.

As in Remark 1.4, if the metric space M is doubling, then one might as well look at coverings of (2.2) by balls with radius equal to $\text{radius}(B_1)$, rather than $C \cdot \text{radius}(B_1)$. For the purposes of this paper, it is reasonable to restrict one's attention to metric spaces that are doubling.

Note that the definition of (s, t) -regular mappings does not really depend on s and t separately, but only on the difference $s - t$. It is often nice to mention them explicitly anyway, and to take s to be the dimension of M (e.g., Hausdorff dimension). In common situations, t could be the dimension of $f(M)$ (as in Examples 2.4), but a priori one has to be a bit careful about this. The (s, t) -regularity condition automatically becomes weaker when t becomes smaller (or s becomes larger), so that t might be less than its optimal value. In particular, it can be less than the dimension of $f(M)$. In the context of this paper, t will often be the Hausdorff dimension of $f(M)$.

When $s = t$, a special feature occurs, which is that one has the class of bilipschitz mappings sitting inside the class of regular mappings in a distinguished way. Strictly speaking, when $s > t$, bilipschitz mappings are also (s, t) -regular, but this is somewhat degenerate. This does not happen if one asks that s be the Hausdorff dimension of the domain, and that t be the Hausdorff dimension of the image, for instance, since a bilipschitz mapping would preserve the Hausdorff dimension.

There are special classes of mappings among (s, t) -regular mappings that one might consider, though, and which can be natural when $s > t$. For instance, if $s - t \geq 1$, then one might look at (s, t) -regular mappings f with the additional property that $f^{-1}(u)$ be connected for every point in the image.

The authors originally considered the notion of (s, t) -regular mappings, and observed most of the results discussed in this paper, several years ago. We simply never got around to writing this up formally before. In the intervening time, some related matters have come up. One is the study of nonlinear quotients in [6]. There they consider conditions along the lines of "co-continuity" and "co-Lipschitzness", in which images of balls in the domain contain balls in the range, with estimates from below for the radii of the balls in the image, and with the centers of the balls in the domain and range matching up under the mapping.

These conditions are quite different from (s, t) -regularity, but there is also some overlap with (s, t) -regularity and related properties. We shall say more about this later, especially in Sections 8 and 9. As a basic instance of this, standard linear projections from \mathbf{R}^m onto \mathbf{R}^n give examples both for mappings that are Lipschitz and co-Lipschitz, as in [6], and for mappings that are (m, n) -regular. On the other hand, mappings like $x \rightarrow |x|$ do not behave well for the conditions in [6] (for balls centered at the origin), but they are accepted by the regularity conditions considered here.

One of the issues considered in [6] is the way that nonlinear quotient mappings from an *infinite*-dimensional Banach space onto another one can respect a substantial amount of the Banach space structure. For instance, one might hope that a nonlinear quotient (which is uniformly continuous and co-uniformly continuous, or Lipschitz and co-Lipschitz) could actually be realized as a linear quotient, or has some properties like that. Some results of this type are discussed in [6].

To put this into perspective, there are analogous questions about Banach spaces which are bilipschitz equivalent, or homeomorphic through mappings which are uniformly continuous and have uniformly continuous inverse. In this case, it is natural to ask whether the two spaces are then linearly isomorphic. See [6] for references for results related to this.

Let us emphasize that the nonlinear quotient conditions in [6] allow the fibers to have infinite dimension, and include linear quotient mappings from one Banach space onto another as a special case. With (s, t) -regular mappings, the fibers always have Hausdorff dimension less than or equal to $s - t$. This is not hard to verify, and there are a number of simple variants of this (concerning the $(s - t)$ -dimensional behavior of the fibers of an (s, t) -regular mapping).

The classes of nonlinear quotient mappings in [6] readily accommodate infinite-dimensional behavior, but they are quite nontrivial in purely finite-dimensional situations as well. See [6] for some aspects of this. In particular, there are examples in [6] (in finite and infinite dimensions) of mappings which are Lipschitz, co-uniformly continuous, and for which there is a nontrivial ball in a codimension-1 subspace of the domain which is sent to a single point in the image. More precisely, this is happening even though the image is not of dimension 1 (in which case it would be normal). This is quite different from what would happen with (s, t) -regular mappings, at least if $s - t$ is strictly less than the dimension of the domain minus 1.

In another direction, one might look at regular and (s, t) -regular mappings in comparison with *quasiregular mappings*, in the sense of [27],

[28]. See also [19], [20] in this regard. Actually, for this it is easier to start with the notion of *mappings of bounded length distortion*, in the sense of [24]. These are also called *BLD mappings*. BLD mappings are quasiregular mappings in which the Jacobian is locally bounded, and uniformly bounded away from 0. This fits better with the notion of regular mappings (as in Definition 1.3), in that the regularity condition for a mapping on a Euclidean space ensures that the differential of the mapping is uniformly bounded, at every point where it exists (because of the Lipschitz condition), and that the absolute value of the Jacobian is uniformly bounded away from 0.

Note that a regular mapping from \mathbf{R}^n to itself (say) does not have to have positive Jacobian, or Jacobian of constant sign. Changing of sign in the Jacobian occurs with the mapping $x \mapsto |x|$ on \mathbf{R} , for instance. For mappings between arbitrary metric spaces, something like having the Jacobian be of constant sign does not really make sense anyway. Positivity of the Jacobian is an important condition for quasiregular and BLD mappings, however. In particular, $x \mapsto |x|$ on the real line is not quasiregular or BLD, because of the change in sign of the Jacobian.

In general, quasiregular mappings are allowed to have Jacobians and differentials which are not bounded or bounded from below, even locally. The concepts of BLD and quasiregular mappings are closer than they might seem at first, in the sense that one can modify the geometry of the domain of a mapping, using the weight that comes from the Jacobian, to put oneself in the situation where the mapping has Jacobian equal to 1 (with respect to the new geometry). Compare with [20], especially Section 2.3.

Quasiregular mappings are also discussed in [6], in connection with the classes of nonlinear quotient mappings considered there. Note that quasiregular mappings are always *open* mappings (i.e., they send open sets to open sets), like the nonlinear quotient mappings considered in [6].

We should perhaps mention that quasiregular (and BLD) mappings are always *discrete*, i.e., the inverse-image of a point under the mapping is always a discrete set. This holds automatically for regular mappings in the sense of Definition 1.3.

The notions of quasiregular and BLD mappings involve having the domains and ranges of the mappings be of the same dimension. With nonlinear quotient mappings as in [6], and (s, t) -regular mappings as in Definition 2.1, one has related classes that allow for the domain and image to have different dimensions. Once the dimensions are permitted to be different, a number of things change, and it is not necessarily so

clear what one might want to view as analogues of quasiregular or BLD mappings (if there are any proper analogues).

One might also think of quasiregular and BLD mappings in connection with quasiconformal and bilipschitz mappings. That is, they are nearly the same, except for giving up the requirement of injectivity, even locally. When the dimensions of the domain and image are allowed to be different, this aspect changes too.

The class of (s, t) -regular mappings is also useful in [31]. See Section 16.3 in [31] in particular.

The main result of this paper will be stated in Section 6, and proved in Section 7. This concerns the way that, given a ball B , one can add $m - t$ components to an (m, t) -regular mapping $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$, to get a mapping which is bilipschitz on a subset of B of substantial measure (compared to the measure of B), with uniform bounds. The situation for ordinary regular mappings is reviewed before that, in Section 4, for which stronger assertions are known. Analogues of these stronger statements in the general case do not work, and this is discussed in Section 5. We also review in Sections 5 and 6 some more classical results about Lipschitz mappings, which help to indicate the differences between the two types of situations (where dimensions are preserved or not). Sections 8 and 9 contain some comparisons and extensions, in connection with co-Lipschitz mappings in particular, and Section 10 describes a more complicated version of the main result which is more global.

In the next section, we review the notion of Ahlfors-regular metric spaces, and mention a few facts related to them and (s, t) -regular mappings.

3. Ahlfors-regular spaces

Definition 3.1. Let $(M, d(x, y))$ be a metric space, and let s be a positive real number. We say that M is *Ahlfors-regular of dimension s* if it is complete, and if there is a constant $C > 0$ so that

$$(3.2) \quad C^{-1}r^s \leq H^s(\overline{B}(x, r)) \leq Cr^s$$

for all $x \in M$ and $r > 0$ such that $r \leq \text{diam } M$.

Here $H^s(A)$ denotes the s -dimensional Hausdorff measure of a set A (as in [16], [17], [25]), and $\overline{B}(x, r)$ denotes the closed ball in M with center x and radius r . For this definition, one might add the requirement that M have at least two elements, to avoid trivialities.

It is a standard fact that Ahlfors-regular metric spaces are always doubling. This is not too hard to show, and it is given in Lemma 5.1 on p. 19 of [14]. (There is a small adjustment needed for this, which is that one should allow arbitrary radii R in Lemma 5.1 in [14], and not just R 's with $R \leq \text{diam } M$. Alternatively, it is enough to consider only radii $R \leq \text{diam } M$ if one uses *closed* balls in M . We were careful to do this in Definition 3.1, and we should have done it in Definition 1.1 in [14]. This is only a small technical point, but one can give examples where it is an issue.)

For Ahlfors-regular spaces, the notions of regular and (s, t) -regular mappings can be given slightly different characterizations, as follows.

Lemma 3.3. *Let $(M, d(x, y))$ and $(N, \rho(u, v))$ be metric spaces, and let s, t be positive real numbers, with $s \geq t$. Assume that $(M, d(x, y))$ is Ahlfors-regular of dimension s . Then a mapping $f: M \rightarrow N$ is (s, t) -regular if and only if it is Lipschitz, and satisfies the following condition: there is a constant $C > 0$ so that if B_1 is a ball in N and B_2 is a ball in M , with $\text{radius}(B_2) \geq \text{radius}(B_1)$, then*

$$(3.4) \quad H^s(f^{-1}(B_1) \cap B_2) \leq C \left(\frac{\text{radius}(B_2)}{\text{radius}(B_1)} \right)^{s-t} \text{radius}(B_1)^s.$$

In the implications in this lemma, the constants that occur in the conclusions can be bounded in terms of constants that occur in the hypotheses (as usual).

When $s = t$, (s, t) -regularity becomes regularity in the sense of Definition 1.3, and the condition in (3.4) can be reduced to the requirement that

$$(3.5) \quad H^s(f^{-1}(B_1)) \leq C \text{radius}(B_1)^s$$

for all balls B_1 in N . This is because one can take the ball B_2 in M to be arbitrarily large, and the right-hand side of (3.4) does not depend on B_2 when $s = t$. The original definition of regular mappings in [8] was given in terms of (3.5), rather than the covering condition in Definition 1.3. This had its genesis in the case of “regular curves” (as in [7], and Examples 1.5), where (3.5) can be interpreted as saying that the amount of arclength of a given curve inside a ball B_1 is less than or equal to $C \cdot \text{radius}(B_1)$.

The case of $s = t$ in Lemma 3.3 is fairly standard, and is given explicitly in Lemma 12.6 on p. 103 in [14]. Lemma 3.3 can be proved in essentially the same manner. Let us briefly go through the argument.

If $f: M \rightarrow N$ is (s, t) -regular, then one can get the bound (3.4) directly from the definitions. That is, if B_1 and B_2 are as above, then one can cover $f^{-1}(B_1) \cap B_2$ by a constant times $(\text{radius}(B_2)/\text{radius}(B_1))^{s-t}$ balls with radius equal to $C_0 \cdot \text{radius}(B_1)$, by (s, t) -regularity, and each of these has H^s -measure less than or equal to a constant times $\text{radius}(B_1)^s$, by the upper bound in Ahlfors-regularity. This gives (3.4).

Conversely, suppose that $f: M \rightarrow N$ is Lipschitz and satisfies the condition in Lemma 3.3, and let us show that f is (s, t) -regular. Let balls B_1 and B_2 in N and M be given as before. We would like to show that $f^{-1}(B_1) \cap B_2$ can be covered by a constant times $(\text{radius}(B_2)/\text{radius}(B_1))^{s-t}$ balls with the same radius as B_1 .

The main point is to use mass bounds and simple covering arguments. By hypothesis, f satisfies the condition in Lemma 3.3, and we can apply this to the balls $2B_1$ and $2B_2$ to get that

$$(3.6) \quad H^s(f^{-1}(2B_1) \cap 2B_2) \leq C \left(\frac{2 \text{radius}(B_2)}{2 \text{radius}(B_1)} \right)^{s-t} (2 \text{radius}(B_1))^s.$$

Of course $f^{-1}(2B_1) \cap 2B_2$ contains $f^{-1}(B_1) \cap B_2$, and it does this with some “extra room”. Specifically, if f is Lipschitz with constant k , and if x is a point in $f^{-1}(B_1) \cap B_2$, then $f^{-1}(2B_1) \cap 2B_2$ contains the ball in M with center x and radius equal to $\min(k^{-1}, 1) \cdot \text{radius}(B_1)$.

From here one would like to obtain that $f^{-1}(B_1) \cap B_2$ can be covered by at most a constant times $(\text{radius}(B_2)/\text{radius}(B_1))^{s-t}$ balls with the same radius as B_1 . Suppose that A is any subset of $f^{-1}(B_1) \cap B_2$ such that

$$(3.7) \quad d(a, b) \geq \text{radius}(B_1) \quad \text{for all } a, b \in A, a \neq b.$$

This implies that the balls $B(a, \text{radius}(B_1)/2)$, $a \in A$, (in M) are pairwise disjoint. Hence the balls

$$(3.8) \quad B(a, \min(k^{-1}, 1/2) \text{radius}(B_1)), \quad a \in A,$$

are pairwise disjoint too. These balls are all contained in $f^{-1}(2B_1) \cap 2B_2$, as in the previous paragraph. Their union therefore has H^s measure bounded by the right-hand side of (3.6). The disjointness of these balls implies that the measure of their union is the equal to the sum of their measures. The measure of each ball is at least a constant times $\text{radius}(B_1)^s$, by the lower bound in Ahlfors-regularity. The combination of these upper and lower bounds implies that the total number of elements of A is bounded by a constant times $(\text{radius}(B_2)/\text{radius}(B_1))^{s-t}$ (independently of the choice of A).

On the other hand, if A is a maximal subset of $f^{-1}(B_1) \cap B_2$ which satisfies (3.7), then $f^{-1}(B_1) \cap B_2$ is covered by the balls $B(a, \text{radius}(B_1))$, $a \in A$. Indeed, if a point x lay in $f^{-1}(B_1) \cap B_2$ but not in any $B(a, \text{radius}(B_1))$, $a \in A$, then one could add x to the set A to get a larger set which still satisfies (3.7), and this would contradict the maximality of A .

Thus, by choosing A to be maximal, one can get a covering of $f^{-1}(B_1) \cap B_2$ with the required properties. This completes the proof of Lemma 3.3.

Remark 3.9. Although the statement of Lemma 3.3 asks that t be positive, the same assertion and proof work when $t = 0$. In this case, (3.4) becomes an inequality that holds automatically when M is Ahlfors-regular of dimension s . In other words, *any* Lipschitz mapping on M is (s, t) -regular when $t = 0$ and M is Ahlfors-regular of dimension s . This works more generally for spaces which are *semi-regular of dimension s* , in the terminology of Definition 5.6 on p. 24 of [14]. The latter is equivalent to *s -homogeneity* in the sense of [2]. See [23] for some additional topics related to this property (in connection with topological dimension in particular).

Here is another fact about regular mappings and Ahlfors-regular spaces.

Lemma 3.10. *Let $(M, d(x, y))$ and $(N, \rho(u, v))$ be metric spaces, and let $f: M \rightarrow N$ be a mapping which is regular (in the sense of Definition 1.3). If M is Ahlfors-regular of some dimension s , then $f(M)$ is Ahlfors-regular of dimension s as well.*

As usual, the constants involved in the Ahlfors-regularity of $f(M)$ are bounded by constants that depend only on ones implicit in the regularity assumption for f , and the Ahlfors-regularity condition for M .

This lemma arises in [8], in a slightly different form. See also Lemma 12.5 on p. 103 of [14]. The main point is simply that Hausdorff measure is preserved, to within bounded factors, by regular mappings (as mentioned in Section 1). This leaves the completeness condition in Definition 3.1, but in fact closed and bounded subsets of Ahlfors-regular spaces are compact (because of completeness and the fact that bounded sets are “totally bounded”, by the doubling property), and one can use this to get completeness of $f(M)$ (since compactness is preserved by taking images under continuous mappings).

The analogue of Lemma 3.10 for (s, t) -regular mappings does not work. That is, if $f: M \rightarrow N$ is (s, t) -regular, and M is Ahlfors-regular of dimension s , then it may not be true that $f(M)$ is Ahlfors-regular of dimension t .

As stated, this has no hope of being true, since an (s, t) -regular mapping is always (s, u) -regular when $u \leq t$. In any case, for (s, t) -regular mappings, one does not have a means by which to get upper bounds for the Hausdorff measure of sets in the image, as one has for ordinary regular mappings (the case when $s = t$). It is not hard to get *lower* bounds, however, just using the definitions; see Lemma 16.38 of [31]. If one knows upper bounds for the image in advance (e.g., if N satisfies the upper bounds associated to Ahlfors-regularity of dimension t), then one is in better shape. In particular, in this case the choice of t would be unique, and as large as possible.

4. Regular mappings between Euclidean spaces

In this section, we shall consider the special case where the metric spaces involved are Euclidean spaces. We shall also restrict ourselves to regular mappings (rather than (s, t) -regular mappings, to which we shall return later). We begin by recalling a result from [9].

Theorem 4.1. *Let m and n be positive integers, with $n \geq m$. If $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ is a regular mapping (using the usual Euclidean metrics on \mathbf{R}^m and \mathbf{R}^n), then for each ball B in \mathbf{R}^m there is a set $E \subseteq B$ such that*

$$(4.2) \quad H^m(E) \geq \delta H^m(B)$$

and

$$(4.3) \quad \text{the restriction of } f \text{ to } E \text{ is bilipschitz with constant } C.$$

Here δ and C are constants that depend only on m, n , and the constants involved in the regularity condition for f (and not on B or f).

Notice that if f is bilipschitz on a set E , then f is automatically bilipschitz on the closure of E as well (since f is continuous).

See Proposition 1 on p. 95 of [9] for this result. Note that there are stronger statements in [9] (from which this is derived), which permit one to obtain a “large bilipschitz piece”, as above, from a Lipschitz condition and information at a single location and scale, rather than an assumption like regularity. Namely, if B is a ball in \mathbf{R}^m , $f: B \rightarrow \mathbf{R}^n$ is Lipschitz, and $H^m(f(B)) \geq \eta H^m(B)$ for some $\eta > 0$, then one can find a set $E \subseteq B$ which satisfies (4.2) and (4.3) above, with constants that depend only on m, η , and the Lipschitz constant for f . See also [10].

In [21], alternate methods are given from which similar conclusions can be derived. The statements in [21] give a somewhat stronger conclusion; instead of having a single “large bilipschitz piece” in a ball B , as above, one gets that for each $\epsilon > 0$ it is possible to find subsets of B on which f is bilipschitz, and which cover all of B except for a set of measure less than $\epsilon H^m(B)$. Here the constants for the bilipschitz conditions, and the number of subsets of B that are used, are bounded in terms of constants that depend on ϵ , m , n , and the regularity constant for f .

As with [9], the results in [21] can be applied to a Lipschitz mapping at a single location and scale, without as restrictive an assumption as regularity. In general, the “bilipschitz pieces” cover all of B except for a set whose image has small measure. This set may not have small measure itself, but this is the case when the mapping is regular, since regular mappings preserve Hausdorff measures to within bounded factors.

Let us mention a corollary to Theorem 4.1, using the following definition.

Definition 4.4. Let $(N, \rho(u, v))$ be a metric space. We say that N is *uniformly rectifiable* if it is Ahlfors-regular of dimension n , for some positive integer n , and if there are constants $\theta, k > 0$ with the following property: for each $x \in N$ and $r > 0$ such that $r \leq \text{diam } N$, there is a subset A of the ball $B(x, r)$ with $H^n(A) \geq \theta r^n$ and a bilipschitz mapping from A into \mathbf{R}^n , where the bilipschitz constant for the mapping is bounded by k .

As a special case, N is uniformly rectifiable if it is bilipschitz equivalent to \mathbf{R}^n . Uniform rectifiability is more flexible than that, allowing partial parameterizations (at all scales and locations), as in Definition 4.4. See [11], [13] for some other characterizations of uniform rectifiability. See [16], [17], [25] concerning classical notions of rectifiability.

Corollary 4.5. *If $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ is regular, then $f(\mathbf{R}^m)$ is a uniformly rectifiable subset of \mathbf{R}^n .*

This is an easy consequence of Theorem 4.1 (and Lemma 3.10). One can get a slightly stronger conclusion than uniform rectifiability, which is that $f(\mathbf{R}^m)$ has “big pieces of Lipschitz graphs”. See Proposition 1 on p. 95 of [9].

An earlier analysis of subsets of \mathbf{R}^n which are images of Euclidean spaces under regular mappings was given in [8]. This analysis went in a similar direction as Theorem 4.1, if not with quite the same final conclusions. In particular, it was shown in [8] that large classes of singular

integral operators are bounded on L^p for Ahlfors-regular sets which are regular images of a Euclidean space.

See also [10] for more on these topics.

5. More classical results about Lipschitz mappings

Let $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ be a Lipschitz mapping. A basic fact is that such a mapping is always differentiable almost everywhere. See [17], [25], [30], [32], for instance.

Now suppose that $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ is Lipschitz, and that A is a subset of \mathbf{R}^m such that $H^m(f(A)) > 0$. In this case there are points in A at which f is differentiable, and such that the rank of the differential is m . This can be derived from the following two statements. If E denotes the set of points in \mathbf{R}^m at which f is *not* differentiable, then E has measure 0, and this implies that

$$(5.1) \quad H^m(f(E)) = 0$$

as well, since f is Lipschitz. On the other hand, if L denotes the set of points at which f is differentiable, but the differential of f has rank less than m , then

$$(5.2) \quad H^m(f(L)) = 0.$$

This is a well-known result, which can be established through covering arguments. See Theorem 7.6 on p. 103 of [25], or Theorem 3.2.3 on p. 243 of [17].

Thus, if $H^m(f(A)) > 0$, then A should not be wholly contained in the union of the sets E and L , since $H^m(f(E \cup L)) = 0$. Moreover,

$$(5.3) \quad H^m(f(A \setminus (E \cup L))) = H^m(f(A)).$$

This implies that A contains a set of positive measure on which f is differentiable, and where the differential always has rank m .

Here is another fact related to these.

Proposition 5.4. *Let $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ be a Lipschitz mapping. Then (in the notation above), $\mathbf{R}^m \setminus (E \cup L)$ can be covered by a countable family of sets, on each of which f is bilipschitz.*

This is contained in Lemma 3.2.2 on p. 242 of [17]. The latter gives some additional properties that one can have for f on the sets which are used to cover $\mathbf{R}^m \setminus (E \cup L)$.

Proposition 5.4 provides a more classical version of some of the themes in Section 4. It deals with similar types of properties, but in a fashion that does not give the same type of quantitative estimates.

We can use this as a kind of testing ground for questions about mappings between dimensions. Specifically, let $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ be a Lipschitz mapping again, and imagine now that we are in a situation where the expected dimension of the image of f is *less* than m .

For simplicity, let us just assume that $n < m$, and that we expect the image of f to be n -dimensional. We still have that f is differentiable almost everywhere. The largest that the rank of the differential of f could be is n , since f takes values in \mathbf{R}^n .

When the rank of the differential of f is equal to n , one is in pretty good shape. One can use results like Proposition 5.4 to say something about the behavior of f . We shall return to this in Section 6.

There is an important difference between this situation and the previous one, where the expected dimension of the image is equal to the dimension of the domain. In the previous case, one can determine (as above) that there are many points in the domain at which the differential of f exists and has rank m , simply given the information that the image of f has positive m -dimensional measure (and that f is Lipschitz). This type of argument does not work in the present circumstances.

Indeed, there are striking results to the effect that one can have mappings $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$, $m > n > 1$, such that

(5.5) f is C^1 (and better than that),

(5.6) the differential of f has rank $< n$ at all points in \mathbf{R}^m , and

(5.7) the image of f contains a nonempty open set in \mathbf{R}^n .

See [3], [5], [22], [33]. One can in fact ask that the differential of the mapping f have rank less than or equal to ℓ everywhere, with ℓ taken to be any integer in the range $1 \leq \ell < n$, and one can get a degree of smoothness (including C^1) which depends on m , n , and ℓ . In other words, $\ell = 1$ is the most impressive in terms of conditions on the differential of f , but then one does not get as much smoothness as when one allows larger ℓ 's.

These results are related to Sard's theorem, and the examples that show that sufficiently-strong smoothness hypotheses are needed (to conclude that images of critical sets have measure 0 in a given dimension). See Section 3.4 of [17] and [4], for instance. In the context of the preceding paragraph, *all* points in \mathbf{R}^m are critical points, which gives a more extreme situation.

6. (m, t) -Regular mappings between Euclidean spaces

Theorem 6.1. *Let m, n , and t be positive integers, with $t \leq m$. Suppose that $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ is an (m, t) -regular mapping, as in Definition 2.1. Then there are positive constants δ, C so that for each ball B in \mathbf{R}^m , there is a linear mapping $\lambda: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$ and a subset E of B with the following properties:*

$$(6.2) \quad H^m(E) \geq \delta H^m(B);$$

$$(6.3) \quad \text{the combined mapping } (f, \lambda): \mathbf{R}^m \rightarrow \mathbf{R}^n \times \mathbf{R}^{m-t} \\ \text{is bilipschitz on } E, \text{ with constant } C;$$

$$(6.4) \quad \text{the norm of } \lambda \text{ is } \leq C.$$

Here C depends only on m, n, t , and the constants implicit in the (m, t) -regularity condition for f (and not on B or the particular mapping f).

Note that it makes sense to think about the “norm of λ ”, as in (6.4), since λ is a linear mapping (which can be represented by a matrix). Exactly what norm one uses does not really matter, since it will only effect the bound to within a constant factor. For the standard “operator norm”, our choice of λ will have norm equal to 1, and λ will in fact be a composition of an orthogonal projection and an isometry onto \mathbf{R}^{m-t} .

Theorem 6.1 is analogous to Theorem 4.1, but for (m, t) -regular mappings rather than ordinary regular mappings. One might also think of the two of them as being similar to the implicit and inverse function theorems. In particular, one way to deal with the implicit function theorem is to add extra components to the given mapping to get one whose differential is invertible, and to which the inverse function theorem is applicable.

Unlike Theorem 4.1, Theorem 6.1 does not have such broad extensions to Lipschitz mappings in general, under mild conditions on the size of the image (as in [9], [21]). The results in [3], [22], [33], mentioned in the paragraph of (5.5), (5.6), (5.7), give strong limitations to this. Thus the assumption of (m, t) -regularity is more crucial here.

Although one does not have such broad results in this context as before, the proof of Theorem 6.1 is fairly robust, and admits a number of variants. In particular, let us point out that the result is local, in that one could start out with a mapping defined on a ball, for instance (rather than all of \mathbf{R}^m), and obtain analogous conclusions. A number of things like this will be clear from the proof. Another result along the lines of Theorem 6.1 will be discussed in Section 10.

In analogy with Corollary 4.5, we have the following.

Corollary 6.5. *Let m, n , and t be positive integers, with $t \leq m$. Assume that $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ is (m, t) -regular, and that the image $f(\mathbf{R}^m)$ is Ahlfors-regular of dimension t . Then $f(\mathbf{R}^m)$ is uniformly rectifiable (Definition 4.4), with bounds for the uniform rectifiability constants which depend only on the constants implicit in the hypotheses.*

As with Corollary 4.5, one also has the slightly stronger “big pieces of Lipschitz graphs” property. We shall say more about this in Section 7.

Note that Ahlfors-regularity of $f(\mathbf{R}^m)$ is an assumption in Corollary 6.5, while it is part of the conclusion in Corollary 4.5. This came up already in Section 3, with Lemma 3.10 and the remarks following it. As indicated there, one can get lower bounds for $f(\mathbf{R}^m)$ as in t -dimensional Ahlfors-regularity when f is (m, t) -regular, but not upper bounds. (One could use this to weaken the hypotheses of Corollary 6.5.)

We shall discuss proofs for Theorem 6.1 and Corollary 6.5 in Section 7. For the moment let us consider some more “classical” statements, in the spirit of Section 5.

Let $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ be a Lipschitz mapping. As before, f is automatically differentiable almost everywhere.

Assume that f is also (m, t) -regular. This implies that

$$(6.6) \quad \begin{aligned} \text{the rank of } df_x \text{ is at least } t, \\ \text{at any point } x \in \mathbf{R}^m \text{ where the differential exists.} \end{aligned}$$

It is not hard to verify this assertion, directly from the definitions.

For the rest of this discussion, we shall use only (6.6), and not the condition of (m, t) -regularity. Let $\lambda_1, \lambda_2, \dots, \lambda_j$ be a finite collection of linear mappings from \mathbf{R}^m to \mathbf{R}^{m-t} , which is sufficiently rich so that the following is true:

$$(6.7) \quad \begin{aligned} \text{if } \alpha: \mathbf{R}^m \rightarrow \mathbf{R}^n \text{ is a linear mapping with rank } \geq t, \text{ then} \\ (\alpha, \lambda_i): \mathbf{R}^m \rightarrow \mathbf{R}^n \times \mathbf{R}^{m-t} \text{ is injective for at least one } i. \end{aligned}$$

This is equivalent to saying that if P is any plane in \mathbf{R}^m of dimension less than or equal to $m - t$ (such as the kernel of α), then there is an i so that the kernel of λ_i is transverse to P . This can easily be arranged.

Lemma 6.8. *Under the conditions described above, there is a countable family of sets $\{E_\ell\}_\ell$ in \mathbf{R}^m such that $\bigcup_\ell E_\ell$ covers all of \mathbf{R}^m except for a set of measure 0, and for each ℓ there is an i , $1 \leq i \leq j$, so that*

$$(6.9) \quad (f, \lambda_i): \mathbf{R}^m \rightarrow \mathbf{R}^n \times \mathbf{R}^{m-t} \text{ is bilipschitz on } E_\ell.$$

This is analogous to Proposition 5.4, and it has much the same relationship to Theorem 6.1 as Proposition 5.4 has with Theorem 4.1. In particular, Theorems 4.1 and 6.1 give quantitative conclusions, while Proposition 5.4 and Lemma 6.8 do not, even if they involve similar kinds of basic structure.

It is not hard to derive Lemma 6.8 from Proposition 5.4. Specifically, one would apply Proposition 5.4 to the mappings $(f, \lambda_i): \mathbf{R}^m \rightarrow \mathbf{R}^n \times \mathbf{R}^{m-t}$, $1 \leq i \leq j$. Almost every element of \mathbf{R}^m is a point of differentiability for f , and the differential always has rank at least t , as in (6.6). Because of (6.7), for each point of differentiability $x \in \mathbf{R}^m$, there is at least one i so that the differential of (f, λ_i) will be injective at that point. This ensures that the sets provided by Proposition 5.4, on which the mappings (f, λ_i) are bilipschitz, cover almost all of \mathbf{R}^m , when one takes the union also over i . This gives Lemma 6.8.

7. Proofs for Theorem 6.1 and Corollary 6.5

To prove Theorem 6.1, we shall use some quantitative results about approximating Lipschitz functions by affine functions. We first review some aspects of this, before proceeding to the main part of the argument.

Let h be a real-valued function on \mathbf{R}^m , and let \mathcal{A} denote the set of real-valued affine functions on \mathbf{R}^m . Given $x \in \mathbf{R}^m$ and $r > 0$, define $\alpha(x, r)$ by

$$(7.1) \quad \alpha(x, r) = \inf_{A \in \mathcal{A}} \sup_{y \in \overline{B}(x, r)} \frac{|h(y) - A(y)|}{r}.$$

This quantity measures how well h can be approximated by an affine function on $\overline{B}(x, r)$. In particular, $\alpha(x, r) = 0$ exactly when h is equal to an affine function on $\overline{B}(x, r)$. If h is differentiable at x , then

$$(7.2) \quad \lim_{r \rightarrow 0} \alpha(x, r) = 0.$$

The converse to this last statement is not true in general, i.e., (7.2) may hold even if h is not differentiable at x . This is because the affine approximations $A(y)$ to h may move around significantly while r tends to 0.

Notice that if h is Lipschitz with constant C_0 , then

$$(7.3) \quad \alpha(x, r) \leq C_0 \quad \text{for all } x \in \mathbf{R}^m, r > 0.$$

This follows by taking $A(y)$ to be the constant function equal to $h(x)$ in (7.1). The boundedness of the $\alpha(x, r)$'s does not imply that h be Lipschitz, but corresponds instead to the ‘‘Zygmund class’’ of functions.

It is easy to check that $\alpha(x, r)$ is continuous in x and r when h is continuous, so that $\alpha(x, r)$ is measurable in particular.

Theorem 7.4. *Let h be a real-valued Lipschitz function on \mathbf{R}^m . Given $\epsilon > 0$, define a set $\mathcal{B}(\epsilon)$ by*

$$(7.5) \quad \mathcal{B}(\epsilon) = \{(x, r) \in \mathbf{R}^m \times (0, \infty) : \alpha(x, r) > \epsilon\}.$$

Then $\mathcal{B}(\epsilon)$ is a Carleson set in $\mathbf{R}^m \times (0, \infty)$, i.e., there is a constant C so that for each $z \in \mathbf{R}^m$ and $s > 0$ we have that

$$(7.6) \quad \int_{B(z,s)} \int_0^s \mathbf{1}_{\mathcal{B}(\epsilon)}(x, r) \frac{dx dr}{r} \leq C s^m.$$

Here $\mathbf{1}_{\mathcal{B}(\epsilon)}(x, r)$ denotes the indicator function of $\mathcal{B}(\epsilon)$, so that $\mathbf{1}_{\mathcal{B}(\epsilon)}(x, r)$ is equal to 0 when (x, r) does not lie in $\mathcal{B}(\epsilon)$, and is equal to 1 when $(x, r) \in \mathcal{B}(\epsilon)$. Also, dx and dr refer to ordinary Lebesgue measure on \mathbf{R}^m and $(0, \infty)$.

The constant C in (7.6) can be chosen so that it depends only on m , ϵ , and the Lipschitz constant for h .

In other words, $\alpha(x, r)$ is small most of the time, in the sense that the exceptional set $\mathcal{B}(\epsilon)$ is “sparse” for every $\epsilon > 0$, as in the Carleson condition above. A key point here is that the integral in (7.6) would diverge if one did not have the indicator function in the integrand, to restrict the integration to pairs (x, r) in $\mathcal{B}(\epsilon)$.

Theorem 7.4 can be derived from the results of [15], and indeed it is definitely weaker than the information provided by [15]. See also [21] and [13]. In the terminology of [13], the property in the conclusions of Theorem 7.4 is sometimes called the *WALA* (weak approximation of Lipschitz functions by affine functions), as in Definition 2.47 on p. 45 of [13], and the remarks just after it. The word “weak” is used to distinguish this condition from stronger quadratic estimates on the $\alpha(x, r)$ ’s and some variants of them (based on integral norms in (7.1) instead of the supremum), as in [15] and p. 18 of [13]. See also Remark 2.28 on p. 336 of [13], concerning a derivation of Theorem 7.4.

Corollary 7.7. *Let $h: \mathbf{R}^m \rightarrow \mathbf{R}$ be a Lipschitz function, and let $\epsilon > 0$ be given. Then there is a constant k so that for each $x \in \mathbf{R}^m$ and $r > 0$ there exist $x_1 \in \mathbf{R}^m$ and $r_1 > 0$ such that $\overline{B}(x_1, r_1) \subseteq \overline{B}(x, r)$, $r_1 \geq k^{-1}r$, and $\alpha(x_1, r_1) \leq \epsilon$. This constant k can be chosen so that it depends only on m , ϵ , and the Lipschitz constant for h .*

Thus, although $\alpha(x, r)$ might not be too small itself, there is always a pair (x_1, r_1) which is not too far from (x, r) in $\mathbf{R}^m \times (0, \infty)$ (in terms of hyperbolic geometry) such that $\alpha(x_1, r_1)$ is as small as one likes.

The conclusion of Theorem 7.4 is significantly stronger than that of Corollary 7.7, by saying that $\alpha(x, r)$ should be small most of the time, on the whole, and not just at regular intervals. We shall explain how Corollary 7.7 can be derived from Theorem 7.4 in a moment.

The property described in Corollary 7.7 and variants of it are discussed in [6], in the context of mappings between Banach spaces more generally. In particular, more direct proofs of Corollary 7.7 are given in [6], which apply to broader situations. The results of [6] also show that k can be taken to be independent of m .

To derive Corollary 7.7 from Theorem 7.4, let h and ϵ be given as in the statement of the corollary, as well as x and r . Let us restrict our attention to x_1, r_1 such that $|x_1 - x| \leq r/2$ and $r_1 \leq r/2$. These conditions imply that $\overline{B}(x_1, r_1) \subseteq \overline{B}(x, r)$.

Let k be a positive number, greater than or equal to 2, to be chosen soon. Consider the integral

$$(7.8) \quad \int_{B(x, r/2)} \int_{k^{-1}r}^{r/2} \mathbf{1}_{\mathcal{B}(\epsilon)}(y, u) \frac{dy du}{u}.$$

The Carleson condition (7.6) implies that this integral is bounded by a constant times r^m , where the constant depends only on m, ϵ , and the Lipschitz constant for h , and not on k . To obtain the conclusions of Corollary 7.7, it suffices to show that there is a pair (x_1, r_1) in $\mathbf{R}^m \times (0, \infty)$ such that $x_1 \in B(x, r/2)$, $r_1 \in [k^{-1}r, r/2]$, and (x_1, r_1) does *not* lie in $\mathcal{B}(\epsilon)$ (so that $\alpha(x_1, r_1) \leq \epsilon$). If this were not the case, then the integral in (7.8) would reduce to

$$(7.9) \quad \int_{B(x, r/2)} \int_{k^{-1}r}^{r/2} \frac{dy du}{u}.$$

This would be too large, compared to the upper bound that we have, if k is taken large enough. How large k has to be depends only on m, ϵ , and the Lipschitz constant for h . This gives Corollary 7.7.

Let us be a bit more explicit about the conclusions of Corollary 7.7. One gets a pair (x_1, r_1) so that $\alpha(x_1, r_1) \leq \epsilon$, and this implies that there is an affine function $A: \mathbf{R}^m \rightarrow \mathbf{R}$ such that

$$(7.10) \quad |h(y) - A(y)| \leq \alpha(x_1, r_1)r_1 \leq \epsilon r_1 \quad \text{for all } y \in \overline{B}(x_1, r_1).$$

The norm of the gradient for A can be bounded in terms of the Lipschitz constant for h . This is something that comes out of standard proofs of Theorem 7.4, and one can also derive it directly. Specifically, one can bound the maximal oscillation of A on $\overline{B}(x_1, r_1)$ in terms of the corresponding oscillation for h , and the norm of the gradient of A is

bounded by its maximal oscillation on $\overline{B}(x_1, r_1)$ divided by r_1 , because it is linear.

Let us emphasize that the bound for the gradient of A that one gets does not depend on ϵ , i.e., it does not blow up when ϵ is small. The price for taking ϵ small comes in the constant k in Corollary 7.7, which controls how “far” from (x, r) one might have to go to get a pair (x_1, r_1) for which there is a good affine approximation.

Now let us prove Theorem 6.1. Let $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ be a mapping which is (m, t) -regular, as in the statement of Theorem 6.1. Let a ball B in \mathbf{R}^m be given.

Let ϵ be a small positive number, to be chosen later. We would like to apply Corollary 7.7 to get a ball $B_1 \subseteq B$ such that

$$(7.11) \quad \text{radius}(B_1) \geq k^{-1} \text{radius}(B),$$

and so that there is an affine mapping $A: \mathbf{R}^m \rightarrow \mathbf{R}^n$ which satisfies

$$(7.12) \quad |f(y) - A(y)| \leq \epsilon \text{radius}(B_1) \quad \text{for all } y \in B_1.$$

The only problem with this is that f takes values in \mathbf{R}^n now, rather than \mathbf{R} . However, it is easy to extend Theorem 7.4 and Corollary 7.7 to the case of \mathbf{R}^n -valued functions. For Theorem 7.4, one can obtain the \mathbf{R}^n -valued case by applying the version for real-valued functions to the components of an \mathbf{R}^n -valued mapping, and combining the information that one gets for them. The analogue of the set $\mathcal{B}(\epsilon)$ (defined in (7.5)) for the \mathbf{R}^n -valued mapping can be viewed as a subset of the union of corresponding sets for the individual components (with adjustments in the choice of ϵ for them to get ϵ in the end for f). The Carleson condition (7.6) for the larger set then follows from analogous conditions for the individual pieces, as desired for the conclusions of Theorem 7.4.

For that matter, common proofs of Theorem 7.4 extend easily enough to the vector-valued case, and would give better information about the constants. Once one has Theorem 7.4 for \mathbf{R}^n -valued functions, one can derive an \mathbf{R}^n -valued version of Corollary 7.7 from it in the same manner as before.

It is also possible to show that Corollary 7.7 directly implies a version of itself for vector-valued functions. See Proposition 2.2 of [6], which allows general Banach spaces in the domain. On the other hand, there are results in [6] concerning versions of Corollary 7.7 in which the domain is finite-dimensional, but the range is infinite-dimensional. There are results in [6] as well about trouble that occurs when both the domain and the range are infinite-dimensional.

At any rate, one can get an affine approximation on B_1 as in (7.11) and (7.12) for a suitable constant k . As before, in the remarks just after (7.10), the differential of A has bounded norm as well. This bound does not depend on ϵ , i.e., does not blow up as ϵ gets small.

From now on, let us imagine that

$$(7.13) \quad B_1 \quad \text{and} \quad A: \mathbf{R}^m \rightarrow \mathbf{R}^n$$

have been chosen and fixed, as above. (Note that B_1 and A depend on ϵ , though.) Our next task is to show that A is nondegenerate in a suitable way, when ϵ is small enough, using the (m, t) -regularity of f .

Let $L: \mathbf{R}^m \rightarrow \mathbf{R}^n$ denote the linear part of A . That is, $A(x) = a + L(x)$ for all $x \in \mathbf{R}^m$, where a is a (single) element of \mathbf{R}^n (namely, $A(0)$). Thus

$$(7.14) \quad L \text{ has bounded norm.}$$

This is the same as the statement above that the differential of A has bounded norm. As before, this bound does not depend on ϵ . This will be important later, in that how small ϵ should be taken to be will depend on the constant in this bound.

The following is a standard fact from linear algebra.

Lemma 7.15. *There is an orthonormal basis $\{v_i\}_{i=1}^m$ of \mathbf{R}^m such that the vectors $L(v_i)$ in \mathbf{R}^n , $1 \leq i \leq m$, are pairwise orthogonal.*

This is a kind of substitute for diagonalization of L . Note that some of the $L(v_i)$'s may be 0 (which is necessarily the case when $n < m$).

To prove Lemma 7.15, consider the mapping $L^t \circ L: \mathbf{R}^m \rightarrow \mathbf{R}^m$, where L^t denotes the transpose of L . This mapping is symmetric, and hence admits a diagonalization by an orthonormal basis $\{v_i\}_{i=1}^m$ of \mathbf{R}^m . From this it follows that the vectors $L(v_i)$ are orthogonal in \mathbf{R}^n , because the inner product between $L(v_i)$ and $L(v_j)$ in \mathbf{R}^n is the same as the inner product between $L^t \circ L(v_i)$ and v_j in \mathbf{R}^m , by definition of the transpose. This proves Lemma 7.15.

Lemma 7.16. *Let L be as above, and let $\{v_i\}_{i=1}^m$ be a basis for \mathbf{R}^m , with the properties described in Lemma 7.15. If ϵ is small enough, depending only on m and the (m, t) -regularity constants for f , then there are positive integers i_1, i_2, \dots, i_t with $1 \leq i_1 < i_2 < \dots < i_t \leq m$ such that*

$$(7.17) \quad |L(v_{i_\ell})| \geq C_0^{-1}, \quad 1 \leq \ell \leq t.$$

Here C_0 is a positive constant that depends only on m and the (m, t) -regularity constants for f (and not on ϵ in particular).

In other words, $|L(v_i)|$ is reasonably large for at least t choices of i . (It may be reasonably large for more than t choices of i .) In particular, the rank of L is at least t . This is analogous to (6.6).

A version of this lemma comes up in [21], with $t = m$ (so that one is getting lower bounds for all of L , rather than just in some directions). For this one can use simpler hypotheses, about the size of $f(B_1)$. For instance, if $t = n$ as well, then it is enough to have a bound from below for the Lebesgue measure of $f(B_1)$, compared to the measure of B_1 .

To prove the lemma, we start with the (m, t) -regularity of f , and try to convert it into a property for L . Let β be a ball in \mathbf{R}^n . If $\text{radius}(\beta) \leq \text{radius}(B_1)$, then we may apply the (m, t) -regularity of f (Definition 2.1) to obtain that

$$(7.18) \quad f^{-1}(\beta) \cap B_1 \text{ can be covered by } \leq C'(\text{radius}(B_1)/\text{radius}(\beta))^{m-t} \\ \text{balls in } \mathbf{R}^m \text{ with the same radius as } \beta,$$

where C' is a positive constant.

Let us assume now that $\text{radius}(\beta) > \epsilon \cdot \text{radius}(B_1)$ as well. Let $\tilde{\beta}$ denote the ball in \mathbf{R}^n with the same center as β , but with

$$(7.19) \quad \text{radius}(\tilde{\beta}) = \text{radius}(\beta) - \epsilon \cdot \text{radius}(B_1).$$

From (7.12) we get that

$$(7.20) \quad A^{-1}(\tilde{\beta}) \cap B_1 \subseteq f^{-1}(\beta) \cap B_1.$$

In other words, if $y \in B_1$ satisfies $A(y) \in \tilde{\beta}$, then $f(y) \in \beta$, because of (7.12) and the definition of $\tilde{\beta}$. Combining (7.20) with (7.18) gives

$$(7.21) \quad A^{-1}(\tilde{\beta}) \cap B_1 \text{ can be covered by } \leq C'(\text{radius}(B_1)/\text{radius}(\beta))^{m-t} \\ \text{balls in } \mathbf{R}^m \text{ with the same radius as } \beta.$$

The preceding observations apply to all balls β in \mathbf{R}^n such that

$$(7.22) \quad \epsilon \cdot \text{radius}(B_1) < \text{radius}(\beta) \leq \text{radius}(B_1).$$

In particular, there is no restriction on the center of β . As a result, the analogue of (7.21) with A replaced by L holds, since A and L differ only by a translation. Similarly, one can use translations in the domain and range to replace B_1 in (7.21) (and its analogue for L instead of A) with a ball which is centered at the origin, and has the same radius as before. This uses the fact that A is affine (and L is linear), and it would not work for general mappings like f .

One can also make a rescaling in the domain and range, so that $\text{radius}(B_1)$ is replaced by 1 in all occurrences, and for the same reasons of linearity. To summarize, we can convert (7.21) to

$$(7.23) \quad L^{-1}(\tilde{\beta}) \cap B(0, 1) \text{ can be covered by} \\ \leq C' \rho^{t-m} \text{ balls of radius } \rho \text{ in } \mathbf{R}^m,$$

where

$$(7.24) \quad \tilde{\beta} \text{ is any ball of radius } \rho - \epsilon \text{ in } \mathbf{R}^n, \text{ and} \\ \rho \text{ is a real number such that } \epsilon < \rho \leq 1.$$

At this point, we can take $\tilde{\beta}$ to be centered at the origin. As in Section 3, one might prefer to formulate (7.23) in terms of volume. This leads to

$$(7.25) \quad \text{Volume}(L^{-1}(B(0, \rho - \epsilon)) \cap B(0, 1)) \leq C'' \rho^t \quad \text{for all } \rho \in (\epsilon, 1],$$

where C'' is a suitable constant (depending on the regularity constant C' and the dimension m). (We are mistreating our notation somewhat here, in that $B(0, \rho - \epsilon)$ is supposed to be a ball in \mathbf{R}^n , while $B(0, 1)$ is a ball in \mathbf{R}^m .)

We want to go from (7.25) to conclusions like the ones in Lemma 7.16. Using the orthonormal basis $\{v_i\}_{i=1}^m$ for \mathbf{R}^m from Lemma 7.15, we can write $L^{-1}(B(0, \rho - \epsilon))$ down explicitly, as an ellipsoid. Namely,

$$(7.26) \quad L^{-1}(B(0, \rho - \epsilon)) = \left\{ x \in \mathbf{R}^m : \sum_{i=1}^m |\langle x, v_i \rangle|^2 |L(v_i)|^2 < (\rho - \epsilon)^2 \right\}.$$

We can write $B(0, 1)$ as

$$(7.27) \quad B(0, 1) = \left\{ x \in \mathbf{R}^m : \sum_{i=1}^m |\langle x, v_i \rangle|^2 < 1 \right\}.$$

If we define an ellipsoid $E(\rho)$ by

$$(7.28) \quad E(\rho) = \left\{ x \in \mathbf{R}^m : \sum_{i=1}^m |\langle x, v_i \rangle|^2 \max \left(1, \frac{|L(v_i)|^2}{(\rho - \epsilon)^2} \right) < 1 \right\},$$

then

$$(7.29) \quad E(\rho) \subseteq L^{-1}(B(0, \rho - \epsilon)) \cap B(0, 1).$$

Thus,

$$(7.30) \quad \text{Volume}(E(\rho)) \leq C'' \rho^t \quad \text{for all } \rho \in (\epsilon, 1],$$

by (7.25).

On the other hand, standard considerations give

$$(7.31) \quad \text{Volume}(E(\rho)) = c_m \prod_{i=1}^m \min \left(1, \frac{\rho - \epsilon}{|L(v_i)|} \right)$$

(when $\rho > \epsilon$). Here c_m is a constant which depends only on m (which is in fact the volume of the unit ball in \mathbf{R}^m). If $|L(v_i)| = 0$ for some i , then we interpret $(\rho - \epsilon)/|L(v_i)|$ as being $+\infty$, so that the minimum with 1 in (7.31) is 1.

Combining (7.30) and (7.31), we have that

$$(7.32) \quad \prod_{i=1}^m \min \left(1, \frac{\rho - \epsilon}{|L(v_i)|} \right) \leq c_m^{-1} C'' \rho^t \quad \text{for all } \rho \in (\epsilon, 1].$$

Given ρ and ϵ , let us write $\nu(\rho - \epsilon)$ for the number of integers i , $1 \leq i \leq m$, such that

$$(7.33) \quad \frac{(\rho - \epsilon)}{|L(v_i)|} < 1.$$

From (7.32) we get that

$$(7.34) \quad (\rho - \epsilon)^{\nu(\rho - \epsilon)} \leq c_m^{-1} C'' C_1^{\nu(\rho - \epsilon)} \rho^t \quad \text{for all } \rho \in (\epsilon, 1],$$

where C_1 is a bound for the norm of L as a linear mapping (as in (7.14)). (Thus $|L(v_i)| \leq C_1$ for each i .)

Let us restrict ourselves now to ρ 's such that $\rho \geq 2\epsilon$. In particular, $\rho - \epsilon$ is then greater than or equal to $\rho/2$. Using this, and combining the constants in (7.34), we obtain that

$$(7.35) \quad \rho^{\nu(\rho - \epsilon)} \leq C_2 \rho^t \quad \text{for all } \rho \in [2\epsilon, 1],$$

where C_2 can be chosen to depend on m and the regularity constants for f (which includes the Lipschitz constant for f), but nothing else. In particular, we can choose C_2 so that it does not depend on ρ or ϵ .

If $\rho \in [2\epsilon, 1]$ and $\nu(\rho - \epsilon) < t$, then (7.35) implies

$$(7.36) \quad C_2^{-1} \leq \rho.$$

Keep in mind that $\nu(\rho - \epsilon)$ and t are integers, so that $\nu(\rho - \epsilon) \leq t - 1$ when $\nu(\rho - \epsilon) < t$.

On the other hand, we are allowed to choose ϵ to be as small as we like. For the purposes of Lemma 7.16, we ask that

$$(7.37) \quad 2\epsilon < C_2^{-1}.$$

This is the only condition that we need to impose on ϵ , i.e., in connection with the hypothesis in Lemma 7.16 that ϵ be small enough. (There is no problem with consistency here, since C_2 does not depend on ϵ .)

If we choose ρ in the range $[2\epsilon, C_2^{-1})$, then (7.36) does not hold, and so the hypotheses of (7.36) should not hold. This yields

$$(7.38) \quad \nu(\rho - \epsilon) \geq t.$$

Going back to the definition of $\nu(\rho - \epsilon)$ (in the sentence containing (7.33)), we get that

$$(7.39) \quad \rho/2 \leq \rho - \epsilon \leq |L(v_i)| \text{ for at least } t \text{ choices of } i, \\ 1 \leq i \leq m, \text{ when } \rho \in [2\epsilon, C_2^{-1}).$$

If ϵ is small enough so that (7.37) holds, then the interval $[2\epsilon, C_2^{-1})$ is nonempty. By choosing ρ to be less than C_2^{-1} , but close to it, one obtains the conclusions of Lemma 7.16 (when (7.37) holds). (In the end, one can take the constant C_0 in (7.17) to be $2C_2$, for instance.) This completes the proof of Lemma 7.16.

Remark 7.40. It is only here, in the proof of Lemma 7.16, that the assumption of (m, t) -regularity of the mapping $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ is invoked, beyond the requirement that f be Lipschitz, for the proof of Theorem 6.1. Of course, the conclusions of Lemma 7.16 will be needed in the rest of the argument, and so the (m, t) -regularity hypothesis will be implicitly employed there as well, but it will not be called up separately again.

The (m, t) -regularity assumption was used here to get (7.18) (and hence (7.21), and so on). In the end, for a particular choice of initial ball B (from the statement of Theorem 6.1, and a few lines before (7.11)), we really only need to apply this condition on a bounded range of locations and scales (compared to the radius of B). One can see this from the proof, and the way that ρ was chosen finally. Notice, however, that this range of possible scales and locations includes the range from which B_1 is chosen, as in (7.11) and (7.12), after B is fixed. In particular, it depends on the constant k from Corollary 7.7, and hence on the (m, t) -regularity constants for f .

This should be compared with the “equidimensional” case, where $m = t$, as in Section 4, and with the discussion in Section 5. In other words, although the equidimensional case is generally more “stable”, with substantially less than regularity needed for many assertions, one does have some limits on the way that the (m, t) -regularity condition is needed when $m > t$.

From now on, let us assume that ϵ is small enough for Lemma 7.16 to be applied. Let us also assume that i_1, i_2, \dots, i_t are as in the conclusions of Lemma 7.16.

Let us write Q_1 for the t -plane in \mathbf{R}^m which is spanned by $v_{i_1}, v_{i_2}, \dots, v_{i_t}$. Let Q_2 denote the orthogonal complement of Q_1 in \mathbf{R}^m , which is the same as taking the span of the v_j 's, where j ranges among the integers between 1 and m (inclusively) which are not one of the i_ℓ 's, $\ell = 1, 2, \dots, t$. In particular, Q_1 has dimension t , and Q_2 has dimension $m - t$.

Let P denote the plane in \mathbf{R}^n which is spanned by $L(v_{i_1}), L(v_{i_2}), \dots, L(v_{i_t})$. These vectors are all orthogonal, as in Lemma 7.15, and they are all nonzero, by Lemma 7.16. Thus

$$(7.41) \quad \text{the dimension of } P \text{ is equal to } t.$$

In fact,

$$(7.42) \quad \begin{aligned} \text{the restriction of } L \text{ to } Q_1 \text{ is an invertible mapping onto } P, \\ \text{with the norm of the inverse bounded by } C_0. \end{aligned}$$

Here C_0 is as in Lemma 7.16. The bound in (7.42) follows from the one in Lemma 7.16, together with the orthogonality of the v_i 's and $L(v_i)$'s.

We are now ready to choose a linear mapping $\lambda: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$, for the purposes of Theorem 6.1. Specifically, λ should be a linear mapping which satisfies the following conditions:

$$(7.43) \quad \begin{aligned} \lambda \equiv 0 \text{ on } Q_1, \text{ and the restriction} \\ \text{of } \lambda \text{ to } Q_2 \text{ is an isometry onto } \mathbf{R}^{m-t}. \end{aligned}$$

In other words, λ is the composition of an orthogonal projection of \mathbf{R}^m onto Q_2 , and an isometry from Q_2 onto \mathbf{R}^{m-t} .

Let π denote the orthogonal projection from \mathbf{R}^n onto P .

Lemma 7.44. *Let C_0 be as in Lemma 7.16, and let λ, π , etc., be as above. If $\epsilon \leq \min(1, C_0^{-1})/2$, then the combined mapping $(\pi \circ f, \lambda)$, which maps \mathbf{R}^m into $P \times \mathbf{R}^{m-t}$, has the following property: the image of the ball B_1 in \mathbf{R}^m under $(\pi \circ f, \lambda)$ contains a ball with radius equal to $(\min(1, C_0^{-1})/2) \cdot \text{radius}(B_1)$ in $P \times \mathbf{R}^{m-t}$. (Remember that B_1 is as in (7.11), (7.12).)*

Recall that ϵ is also assumed to be small enough for the purposes of Lemma 7.16. Note that the constant C_0 from Lemma 7.16 does not depend on ϵ , so that there is no consistency problem with asking that ϵ be smaller than $\min(1, C_0^{-1})/2$.

To prove Lemma 7.44, consider first the corresponding question where f is replaced with our linear mapping $L: \mathbf{R}^m \rightarrow \mathbf{R}^n$. In this case, the combined mapping $(\pi \circ L, \lambda): \mathbf{R}^m \rightarrow P \times \mathbf{R}^{m-t}$ is also linear. In fact this mapping is invertible, and the norm of its inverse is $\leq \max(1, C_0)$.

To see this, let us identify \mathbf{R}^m with the Cartesian product of Q_1 and Q_2 , using the usual orthogonal decomposition of vectors in \mathbf{R}^m into their components in Q_1 and Q_2 . By construction, $\lambda \equiv 0$ on Q_1 . Also, $\pi \circ L = 0$ everywhere on Q_2 . Indeed, if $w \in \mathbf{R}^m$ lies in Q_2 , then it is a linear combination of v_j 's with j not among i_1, i_2, \dots, i_t , so that $L(w)$ is a linear combination of $L(v_j)$'s, with the same j 's. This implies that $L(w)$ is orthogonal to $L(v_{i_\ell})$ for all ℓ , $1 \leq \ell \leq t$, which means exactly that $L(w)$ is orthogonal to the t -plane P . Hence $\pi(L(w)) = 0$ when $w \in Q_2$, as desired.

In short, $(\pi \circ L, \lambda): \mathbf{R}^m \rightarrow P \times \mathbf{R}^{m-t}$ is essentially the same as the Cartesian product of $\pi \circ L: Q_1 \rightarrow P$ and $\lambda: Q_2 \rightarrow \mathbf{R}^{m-t}$. The question of invertibility of the combined mapping reduces to the invertibility of the individual pieces. We chose λ so that it is an isometry from Q_2 onto \mathbf{R}^{m-t} , and hence this part is fine. As for $\pi \circ L$, notice that it is equal to L on Q_1 , since P is the image of Q_1 under L by definition, and π projects \mathbf{R}^n onto P . We already know that L is invertible as a linear mapping from Q_1 to P , with norm bounded by C_0 , as in (7.42).

Thus we obtain that $(\pi \circ L, \lambda): \mathbf{R}^m \rightarrow P \times \mathbf{R}^{m-t}$ is invertible, with the norm of its inverse bounded by $\max(1, C_0)$. This implies that the image of B_1 under $(\pi \circ L, \lambda)$ contains a ball in $P \times \mathbf{R}^{m-t}$ with radius equal to $\min(1, C_0^{-1}) \cdot \text{radius}(B_1)$.

Now let $A: \mathbf{R}^m \rightarrow \mathbf{R}^n$ be the affine mapping from before, as in (7.12). By definition, L is the linear part of A . (See the lines just before (7.14).) In particular, A and L differ simply by a translation in the image. The same is true of $\pi \circ A$ and $\pi \circ L$. Thus we conclude that the image of B_1 under $(\pi \circ A, \lambda)$ contains a ball of radius $\min(1, C_0^{-1}) \cdot \text{radius}(B_1)$ in $P \times \mathbf{R}^{m-t}$.

It remains to go from A to our original mapping $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$. The difference between A and f is bounded by $\epsilon \cdot \text{radius}(B_1)$ on B_1 , as in (7.12). This implies that the difference between $(\pi \circ A, \lambda)$ and $(\pi \circ f, \lambda)$ is bounded by $\epsilon \cdot \text{radius}(B_1)$ on B_1 as well.

At this point, the fact that the image of B_1 under $(\pi \circ f, \lambda)$ contains a ball in $P \times \mathbf{R}^{m-t}$ with radius equal to $\min(1, C_0^{-1})/2$ times $\text{radius}(B_1)$ when $\epsilon \leq \min(1, C_0^{-1})/2$ can be derived in a standard way from degree theory, as in [26]. Specifically, if we think of $(\pi \circ A, \lambda)$ as a mapping from B_1 into $P \times \mathbf{R}^{m-t}$, then this mapping has degree equal to 1 or -1 everywhere on the image of B_1 . The value of 1 or -1 depends on how orientations are chosen in the domain and in the image, and whether $(\pi \circ A, \lambda)$ preserves or reverses orientations. (Keep in mind that $(\pi \circ A, \lambda)$ is an affine homeomorphism from \mathbf{R}^m onto $P \times \mathbf{R}^{m-t}$.)

To be more precise, the degree of $(\pi \circ A, \lambda)$ as a mapping from B_1 into $P \times \mathbf{R}^{m-t}$ is defined for all points y in $P \times \mathbf{R}^{m-t}$ which do not lie in the image of ∂B_1 under $(\pi \circ A, \lambda)$ (at least if orientations have been selected for the domain and range spaces). The degree is automatically 0 on the complement of the closure of the image of B_1 , but this will not matter too much here.

Similarly, if we think of $(\pi \circ f, \lambda)$ as a mapping from B_1 into $P \times \mathbf{R}^{m-t}$, then the degree of f is defined for points y in $P \times \mathbf{R}^{m-t}$ that do not lie in the image of ∂B_1 under $(\pi \circ f, \lambda)$. Note that the image of ∂B_1 under $(\pi \circ f, \lambda)$ is compact, since ∂B_1 is compact and $(\pi \circ f, \lambda)$ is continuous. The degree is locally constant on the complement of the image of ∂B_1 under $(\pi \circ f, \lambda)$, and is thus constant on the components of this set. The degree is zero at points which are not in the image of $(\pi \circ f, \lambda)$ (on B_1). These are general properties for degrees of (continuous) mappings.

In addition to $(\pi \circ f, \lambda)$ and $(\pi \circ A, \lambda)$, one can look at convex combinations of these two mappings (as mappings from B_1 into $P \times \mathbf{R}^{m-t}$). Another general property of the degree is that it does not change under continuous deformations of mappings, as long as the point y in the range at which the degree is being evaluated does not lie in the image of the boundary of the domain (B_1 in this case) under any of the mappings under consideration.

In our situation, $(\pi \circ f, \lambda)$ and convex combinations of $(\pi \circ f, \lambda)$ and $(\pi \circ A, \lambda)$ all differ from $(\pi \circ A, \lambda)$ on $\overline{B_1}$ by at most $\epsilon \cdot \text{radius}(B_1)$, because of (7.12). Using this, one gets that the degree of $(\pi \circ f, \lambda)$ is the same as the degree of $(\pi \circ A, \lambda)$ at any point $y \in P \times \mathbf{R}^{m-t}$ such that the distance from y to the image of ∂B_1 under $(\pi \circ A, \lambda)$ is greater than $\epsilon \cdot \text{radius}(B_1)$.

We have already mentioned that the image of B_1 under $(\pi \circ A, \lambda)$ contains a ball in $P \times \mathbf{R}^{m-t}$ of radius $\min(1, C_0^{-1}) \cdot \text{radius}(B_1)$. Let us call this ball β . In particular, the image of ∂B_1 under $(\pi \circ A, \lambda)$ lies in the complement of β , since $(\pi \circ A, \lambda)$ is a homeomorphism (of \mathbf{R}^m onto $P \times \mathbf{R}^{m-t}$). The degree of $(\pi \circ A, \lambda)$ is equal to 1 or to -1 at all points of β , as before.

Let β' denote the set of points in β which lie at distance $> \epsilon \cdot \text{radius}(B_1)$ from the complement of β . Combining the statements in the previous paragraphs, we get that $(\pi \circ f, \lambda)$ and $(\pi \circ A, \lambda)$ have the same degree at every point y in β' . We know that the degree of $(\pi \circ A, \lambda)$ at these points is 1 or -1 , and so the same is true for $(\pi \circ f, \lambda)$.

In particular, these points lie in the image of B_1 under $(\pi \circ f, \lambda)$, since otherwise the degree would be 0. In short, β' is contained in the image of B_1 under $(\pi \circ f, \lambda)$. The conclusions of Lemma 7.44 follow easily from

this, when $\epsilon \leq \min(1, C_0^{-1})/2$, because of the definitions of β and β' . This completes the proof of Lemma 7.44.

At this stage, it is easy to finish the proof of Theorem 6.1. Let us now fix $\epsilon > 0$, once and for all, and small enough for the purposes of Lemmas 7.16 and 7.44. Although ϵ should be small enough for these lemmas, one should not take it to be too small, because some of the other estimates deteriorate as ϵ gets small. Thus one might as well take it to be as large as possible, subject to the conditions of Lemmas 7.16 and 7.44.

The main point now is that we can apply results from [9], [21] (as mentioned in Section 4) to the mapping $(\pi \circ f, \lambda): B_1 \rightarrow P \times \mathbf{R}^{m-t}$. Note that $P \times \mathbf{R}^{m-t}$ is essentially the same as \mathbf{R}^m , since P has dimension t (as in (7.41)).

Specifically, in applying the results from [9], [21], we are using two pieces of information. The first is that $(\pi \circ f, \lambda): B_1 \rightarrow P \times \mathbf{R}^{m-t}$ is Lipschitz, with bounded constant. The constant for $(\pi \circ f, \lambda)$ is less than or equal to the sum of 1 and the Lipschitz constant for f , because π and λ are both Lipschitz with constant 1, by their definitions (as an orthogonal projection and an orthogonal projection composed with an isometry). The second piece of information is that the (m -dimensional) Lebesgue measure of the image of B_1 under $(\pi \circ f, \lambda)$ is greater than or equal to a constant times the Lebesgue measure of B_1 . This follows from Lemma 7.44.

Under these conditions, we obtain that there is a subset E of B_1 such that the measure of E is greater than or equal to a constant times the measure of B_1 , and such that the restriction of $(\pi \circ f, \lambda)$ to E is bilipschitz. The constants in these two properties are controlled in terms of the constants mentioned in the previous paragraph, and the dimension m .

This is exactly what we want for Theorem 6.1, except for two points. The first is that we should compare E with our original ball B (mentioned a few lines before (7.11)), rather than with B_1 . As in (7.11) and the line preceding it, B_1 is contained in B , and the radius of B_1 is greater than or equal to a constant times the radius of B . Thus E is contained in the original ball B , and the measure of E is bounded from below by a constant times the measure of B . Note that the constant from (7.11) depends on ϵ , and hence this constant for the lower bound for the measure of E in terms of the measure of B does too. This is the place where taking ϵ to be too small would have an effect.

The second point is that we get a bilipschitz condition for $(\pi \circ f, \lambda)$ on E , rather than for (f, λ) . Of course f and λ are already Lipschitz, and so the only issue is with the lower bounds in the bilipschitz condition. For

this the bilipschitzness of $(\pi \circ f, \lambda)$ is a stronger property than for (f, λ) , since π is Lipschitz. (We shall make use of this stronger information in a moment.) Thus one gets bilipschitzness for (f, λ) on E as well.

This completes the proof of Theorem 6.1.

Now let us prove Corollary 6.5. Let $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ be given, where f is (m, t) -regular, and assume that the image $f(\mathbf{R}^m)$ is Ahlfors-regular of dimension t . We would like to show that $f(\mathbf{R}^m)$ is uniformly rectifiable (with dimension t), with suitable bounds for the constants involved.

Fix a point $z \in f(\mathbf{R}^m)$ and a radius $r > 0$. Since $z \in f(\mathbf{R}^m)$, there is an element x of \mathbf{R}^m such that $f(x) = z$. Also, let C be the Lipschitz constant for f . We have that

$$(7.45) \quad f(B(x, C^{-1}r)) \subseteq B(z, r),$$

from these choices.

Let us now apply Theorem 6.1, with the ball B in \mathbf{R}^m taken to be $B(x, C^{-1}r)$, and with this mapping f . Thus we get a linear mapping $\lambda: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$ and a set $E \subseteq B$ such that the measure of E is greater than or equal to a constant times the measure of B , and (f, λ) is bilipschitz on E , with a bounded constant. As in the proof of Theorem 6.1, we can take λ to be a composition of an orthogonal projection and an isometry onto \mathbf{R}^{m-t} .

For each point $u \in \mathbf{R}^{m-t}$, $\lambda^{-1}(u)$ is an affine plane in \mathbf{R}^m of dimension t . Put

$$(7.46) \quad E_u = E \cap \lambda^{-1}(u).$$

On each E_u , f is bilipschitz (with bounded constant), because of the bilipschitz condition for (f, λ) on E itself. On the other hand, there exist u 's in \mathbf{R}^{m-t} such that the t -dimensional measure of E_u is bounded from below by a constant times $(C^{-1}r)^t$. This follows easily from a Fubini theorem argument, since the measure of E is bounded from below by a constant times the measure of B , and since the relevant u 's (for which $E_u \neq \emptyset$) lie in $\lambda(B)$, which is a ball in \mathbf{R}^{m-t} with the same radius as B .

The bilipschitz property for f on E_u then gives the kind of "big bilipschitz piece" for $f(\mathbf{R}^m)$ in $B(z, r)$ which is required in the definition of uniform rectifiability, as in Definition 4.4. This completes the proof of Corollary 6.5.

Remark 7.47. Under the assumptions of Corollary 6.5, one can get a slightly stronger conclusion for $f(\mathbf{R}^m)$, which is that it has "big pieces of Lipschitz graphs" (BPLG). This property is defined in practically the same manner as uniform rectifiability was, except that instead of asking that $f(\mathbf{R}^m) \cap B(z, r)$ have a subset of substantial size which is

bilipschitz equivalent to a subset of \mathbf{R}^t , with uniform bounds, we want to have a substantial subset of $f(\mathbf{R}^m) \cap B(z, r)$ which lies in the graph of a Lipschitz mapping (over some t -dimensional plane in \mathbf{R}^n), again with uniform bounds.

The BPLG condition for $f(\mathbf{R}^m)$, under the hypotheses of Corollary 6.5, can be established through nearly the same argument as above, for the uniform rectifiability of $f(\mathbf{R}^m)$. The main difference is the following. Instead of the bilipschitzness of (f, λ) on the set E , as provided by the statement of Theorem 6.1 and used in the argument above, one employs the stronger feature of bilipschitzness for $(\pi \circ f, \lambda)$ on E , where π is an orthogonal projection of \mathbf{R}^n onto an t -dimensional subspace. This was given in the proof of Theorem 6.1, and was indicated at the very end of the proof in particular. This leads to bilipschitzness for $\pi \circ f$ on the slices E_u , rather than just for f itself. Once one has this, it is not hard to get a big piece of a Lipschitz graph for $f(\mathbf{R}^m)$, in the given ball in the image, as before. This is analogous to the situation in [9], for ordinary regular mappings.

Remark 7.48. The conditions “ $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ is (m, t) -regular” and “ $f(\mathbf{R}^m)$ is Ahlfors-regular of dimension t ” each make sense for positive real numbers t , whether or not t is an integer. However, if one assumes both conditions at the same time, as in the context of Corollary 6.5, then that implies that t is an integer. Indeed, if f is (m, t) -regular, then it is Lipschitz in particular, and hence differentiable almost everywhere on \mathbf{R}^m . Let x be a point of differentiability of f . One can show that the rank of the differential df_x of f at x should be less than or equal to t , under the condition that $f(\mathbf{R}^m)$ be Ahlfors-regular of dimension t . Similarly, the assumption that f be (m, t) -regular implies that the rank of df_x should be at least t , as in (6.6). Thus t should be equal to the rank of df_x , and this is automatically an integer.

8. Comparisons with the co-Lipschitz property

Proposition 8.1. *Let m and t be positive integers, with $m \geq t$. Suppose that $f: \mathbf{R}^m \rightarrow \mathbf{R}^t$ is an (m, t) -regular mapping. Then there is a constant $C > 0$ so that for every ball B in \mathbf{R}^m there is a ball B' in \mathbf{R}^t such that*

$$(8.2) \quad f(B) \supseteq B'$$

and

$$(8.3) \quad \text{radius}(B') \geq C^{-1} \text{radius}(B).$$

This constant C can be chosen to depend only on m , t , and the constants involved in the (m, t) -regularity condition for f .

The conclusion of this proposition is somewhat close to the co-Lipschitz property in [6]. The difference is that the ball B' is not required to be centered at the image of the center of B under f . Let us refer to the condition in the conclusion of the proposition as the “non-centered co-Lipschitz property”.

Proposition 8.1 would not work in general if one did ask that the center of B' be the image of the center of B under f . For example, consider the mapping $f: \mathbf{R} \rightarrow \mathbf{R}$ given by $f(x) = |x|$ for all x . It is easy to see that this mapping is regular (or (m, t) -regular, with $m = t = 1$), as in Examples 1.5. This mapping satisfies the conclusions of Proposition 8.1, but this would not be true with the additional condition on the center of B' . Specifically, the additional condition does not work when B is centered at the origin in \mathbf{R} .

One can make similar examples in other dimensions, covering all pairs of positive integers (m, t) with $m \geq t$. Thus Proposition 8.1 is reasonably sharp, as a statement which gives part of co-Lipschitzness. In Section 9 we shall go in the opposite direction, and look at some consequences of conditions like the co-Lipschitz property.

The proof of Proposition 8.1 is approximately contained in the proof of Theorem 6.1 in Section 7 already. In particular, Lemma 7.44 provides a conclusion which is close to the one being sought here. The main difference is that what was \mathbf{R}^n is now \mathbf{R}^t . This implies that what was the t -dimensional plane P in \mathbf{R}^n before (as in (7.41) and the lines just before it) is now simply $\mathbf{R}^n = \mathbf{R}^t$ itself. Similarly, what was the projection $\pi: \mathbf{R}^n \rightarrow P$ (as defined just before Lemma 7.44) is now the identity mapping on \mathbf{R}^t . With these changes, Lemma 7.44 gives almost exactly the result desired for Proposition 8.1. (One has just to be a little careful about the choice of ϵ , and the relationship of B_1 in Lemma 7.44 to the original ball B . These points are essentially the same as in Section 7, for the last part of the proof of Theorem 6.1 (after Lemma 7.44) in particular.)

For the present purposes, one could also simplify Lemma 7.44 and its proof. Instead of working with the mapping $(\pi \circ f, \lambda)$, as before, one can use the restriction of f to the t -dimensional plane Q_1 (defined a couple of paragraphs before the statement of Lemma 7.44). One would then want to show that $f(B_1 \cap Q_1)$ contains a ball in \mathbf{R}^t with radius greater than or equal to a constant times the radius of B_1 . One can do this through much the same argument as before, using degree theory, applied to f as a mapping from $B_1 \cap Q_1$ to \mathbf{R}^t (and the affine approximation of f).

9. Noncollapsing mappings

Let us begin with an auxiliary definition. Given a set $E \subseteq \mathbf{R}^n$ and a nonnegative number t , define the t -dimensional Hausdorff content of E , $H_{\text{con}}^t(E)$, as follows. For any sequence of sets $\{A_j\}_j$ in \mathbf{R}^n , consider the sum

$$(9.1) \quad \sum_j (\text{diam } A_j)^t.$$

To get $H_{\text{con}}^t(E)$, one takes the infimum of this sum over all sequences $\{A_j\}_j$ of sets which cover E .

If $H^t(E)$ denotes the ordinary t -dimensional Hausdorff measure of E , then

$$(9.2) \quad H_{\text{con}}^t(E) \leq H^t(E)$$

automatically. This is because $H^t(E)$ is defined in terms of the same kind of sums (9.1), but with more restrictions on the coverings $\{A_j\}_j$. Namely, the A_j 's would be required to have diameter less than a positive number δ , where one takes the limit as $\delta \rightarrow 0$ at the end, after first taking the infimum of the sums (9.1) over these coverings. However, it is true that

$$(9.3) \quad H_{\text{con}}^t(E) = 0 \quad \text{implies} \quad H^t(E) = 0.$$

This is not hard to verify. (If $H_{\text{con}}^t(E) = 0$, then the coverings of E that one gets consist of sets with small diameter anyway.)

If E is contained in a set which is Ahlfors-regular of dimension t , then one does have bounds of the form

$$(9.4) \quad H^t(E) \leq C H_{\text{con}}^t(E).$$

This is easy to check, just using the definitions. It works as well if one only has the upper bound in (3.2) (with s replaced with t).

Hausdorff content behaves in essentially the same manner under Lipschitz and bilipschitz mappings as Hausdorff measure, i.e., it does not increase by more than a constant factor for Lipschitz mappings (depending on the dimension t and the Lipschitz constant), and it does not decrease by more than a constant factor either for bilipschitz mappings. This follows easily from the definition.

Definition 9.5. Let m and n be positive integers, and let t be a positive real number with $t \leq m$. Suppose that $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ is a mapping which is Lipschitz. We say that f is (m, t) -noncollapsing if there is a constant C

so that

$$(9.6) \quad H_{\text{con}}^t(f(B)) \geq C^{-1}(\text{radius}(B))^t$$

for all balls B in \mathbf{R}^m .

There are obviously some natural generalizations and variants of this. In particular, one might consider the condition in which t -dimensional Hausdorff measure is used in (9.6) instead of Hausdorff content. If the image set $f(\mathbf{R}^m)$ is contained in a set which is Ahlfors-regular of dimension t (or satisfies the corresponding upper bounds from (3.2)), then the two conditions are equivalent, i.e., using Hausdorff content or Hausdorff measure.

If $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ is (m, t) -regular, then it is (m, t) -noncollapsing. This is not too hard to verify, directly from the definitions. That is, if the t -dimensional Hausdorff content of the image of a ball B were ever too small, then it would mean that there is a covering of $f(B)$ for which a sum like (9.1) would be too small. This covering could then be converted into one for the ball B itself, using the (m, t) -regularity property, in a way that would give a contradiction. Specifically, it would lead to the m -dimensional Lebesgue measure of B being too small compared to $(\text{radius}(B))^m$. (Similar observations are described in more detail in [31].)

If t is an integer, and $f: \mathbf{R}^m \rightarrow \mathbf{R}^t$ is Lipschitz and co-Lipschitz (in the sense of [6]), then f is (m, t) -noncollapsing as well. Indeed, in this case $f(B)$ will contain a ball in \mathbf{R}^t with radius $\geq C^{-1} \text{radius}(B)$ for some constant C (that does not depend on B), so that the t -dimensional Lebesgue measure will be bounded from below by a constant times $(\text{radius}(B))^t$. The same will be true of the t -dimensional Hausdorff content, as in (9.4).

This argument also works if f is Lipschitz and satisfies the non-centered co-Lipschitz property from Section 8. In other words, one does not need to know anything about the center of the ball in \mathbf{R}^t which is contained in $f(B)$, but only the lower bound for the radius.

Conversely, if $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ is Lipschitz and (m, t) -noncollapsing, then one has the same kinds of conclusions as for (m, t) -regular mappings in Sections 6, 7, and 8. We can state the main part of this as follows.

Proposition 9.7. *Theorem 6.1, Corollary 6.5, and Proposition 8.1 still hold if one asks that $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ be Lipschitz and (m, t) -noncollapsing, instead of (m, t) -regular.*

Corollary 9.8. *If t is an integer, and if $f: \mathbf{R}^m \rightarrow \mathbf{R}^t$ is Lipschitz, then f is (m, t) -noncollapsing if and only if it satisfies the non-centered co-Lipschitz property.*

The “if” part of the corollary was mentioned before the statement of Proposition 9.7, and the “only if” part follows from the extension of Proposition 8.1 to (m, t) -noncollapsing mappings indicated in Proposition 9.7.

Let us look at the reasons why Theorem 6.1, Corollary 6.5, and Proposition 8.1 extend to the case of Lipschitz mappings that are (m, t) -noncollapsing, as well as the other observations from Sections 6, 7, and 8. The main point is that affine approximations of (m, t) -noncollapsing mappings satisfy the same kind of nondegeneracy properties as for (m, t) -regular mappings. For this we shall use the following lemma.

Lemma 9.9. *Let E be a subset of \mathbf{R}^n . Assume that*

$$(9.10) \quad \text{diam } E \leq kR$$

for some positive real numbers k and R , and that P is a d -dimensional plane in \mathbf{R}^n such that

$$(9.11) \quad \text{dist}(z, P) \leq \tau R \quad \text{for all } z \in E.$$

Here τ is another positive real number, which we assume to be less than 1. (Normally τ will be small, k will be bounded, and R will reflect an arbitrary choice of scale.) If t is a positive real number with $t \geq d$, then

$$(9.12) \quad H_{\text{con}}^t(E) \leq C R^t \tau^{t-d}.$$

Here C is a constant which may depend on k , d , and t , but does not depend on R or τ .

To see this, one first observes that

$$(9.13) \quad E \text{ can be covered by } \leq C_1 \tau^{-d} \text{ balls of radius } 2\tau R,$$

where C_1 is a constant which depends only on k and d . This covering can be obtained as follows. Let π denote the orthogonal projection of \mathbf{R}^n onto P . Thus $\pi(E)$ is a subset of P with diameter which is less than or equal to the diameter of E . Because $\text{diam } E \leq kR$ and P has dimension d , one can cover $\pi(E)$ with $\leq C_1 \tau^{-d}$ balls in P of radius τR , where C_1 depends only on k and d . To get the covering indicated in (9.13), one takes the balls in \mathbf{R}^n with the same centers as these balls in P , but with radii $2\tau R$ instead of τR . This gives a covering of E itself, because of the assumption (9.11).

Once one has a covering as in (9.13), the estimate (9.12) follows immediately from the definition of the Hausdorff content. This proves Lemma 9.9.

Notice that the same argument would work when $t < d$. However, in this case one can simply use the diameter bound (9.10) to get that $H_{\text{con}}^t(E) \leq (kR)^t$. One does not need the power of τ in (9.12), which would now be negative, and not helping the estimate.

Now let us return to the discussion of mappings between Euclidean spaces. Suppose that $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ is Lipschitz and (m, t) -noncollapsing. Let $x \in \mathbf{R}^m$ be a point at which f is differentiable (which includes almost all points in \mathbf{R}^m , since f is Lipschitz). Then the rank of the differential of f at x is greater than or equal to t . To see this, suppose to the contrary that the rank of the differential at some point x is strictly less than t . Let B be a small ball which is centered at x . We would like to apply Lemma 9.9, with $E = f(B)$ and $R = \text{radius}(B)$. In this case (9.10) holds automatically, with k equal to twice the Lipschitz constant for f .

Let P be the plane in \mathbf{R}^n which is the image of \mathbf{R}^m under the affine mapping α defined by

$$(9.14) \quad \alpha(y) = f(x) + df_x(y - x).$$

In other words, α is the affine mapping which approximates f well near x , in the sense that

$$(9.15) \quad \lim_{y \rightarrow x} \frac{|f(y) - \alpha(y)|}{|y - x|} = 0,$$

because of differentiability. With this choice of P , the condition (9.11) holds with τ as small as one would like (for $E = f(B)$), when the ball B is sufficiently small.

If d is the dimension of P , then d is the rank of the differential of f at x , which we are assuming is less than t . The estimate (9.12) then applies to say that

$$(9.16) \quad H_{\text{con}}^t(f(B)) \leq C R^t \tau^{t-d},$$

where C depends only on d , t , and the Lipschitz constant for f . This inequality contradicts the (m, t) -noncollapsing condition for f , when τ is small enough.

This shows that the rank of the differential of f is always at least t . More generally, the following is true. Let B be a ball in \mathbf{R}^m , and suppose that f is well-approximated by an affine mapping $A: \mathbf{R}^m \rightarrow \mathbf{R}^n$ in a

ball B . This means that

$$(9.17) \quad |f(y) - A(y)| \leq \epsilon \operatorname{radius}(B) \quad \text{for all } y \in B$$

(as in (7.12)), where ϵ is a small positive number. Let L denote the linear part of A , as in Section 7, so that $A(x) = a + L(x)$ for some $a \in \mathbf{R}^n$ and all $x \in \mathbf{R}^m$. If ϵ is small enough, depending only on t and the Lipschitz and (m, t) -noncollapsing constants for f , then L satisfies the same kind of t -dimensional nondegeneracy conditions (with bounds) as in Lemma 7.16. This can be shown with the same kinds of arguments as above, using also the orthonormal basis $\{v_i\}_{i=1}^m$ for \mathbf{R}^m provided by Lemma 7.15.

To be more precise, suppose that one does not have a good lower bound for $|L(v_i)|$ for at least t choices of i , as in (7.17) in Lemma 7.16. This means that there is an integer $d < t$ such that $|L(v_i)|$ is small for all but d choices of i . Let Q_0 denote the span of d choices of v_i in \mathbf{R}^m which cover all of the v_j 's for which $|L(v_j)|$ is not too small, and let P_0 denote the image of Q_0 under the affine mapping A . Then each point in $f(B)$ lies within $\epsilon \cdot \operatorname{radius}(B)$ of a point in $A(B)$, because of (9.17), and each point in $A(B)$ lies close to P_0 (compared to $\operatorname{radius}(B)$), because of our assumption about the small values of $|L(v_k)|$ (when v_k is not among the d choices of v_i of which Q_0 is composed). In other words, the small values of $|L(v_k)|$ lead to an estimate like (9.11), with $P = P_0$ and $E = f(B)$ again. If ϵ and the values of $|L(v_k)|$ for v_k not in Q_0 are small enough, then Lemma 9.9 gives an upper bound for the Hausdorff content for $f(B)$ which would be too small for the (m, t) -noncollapsing condition, in much the same manner as before.

In this way one can get the same kind of t -dimensional nondegeneracy conclusions as in Lemma 7.16. One can also use an argument more like the one employed for Lemma 7.16 before, in which the (m, t) -noncollapsing condition for f on B is converted into a similar condition for the affine approximation A , and then its linear part L , and then one looks at the ellipsoid which is the image of B under L . In the end, the two arguments are about the same anyway.

Once one has the analogue of Lemma 7.16 in this setting, the rest is practically the same as before. That is, it was only for getting Lemma 7.16 that we really needed the (m, t) -regularity assumption in Section 7. This was mentioned in Remark 7.40.

The same is true for the arguments in Section 8, which were closely based on the ones in Section 7. In Section 6 some remarks were made about more “classical” statements for the behavior of (m, t) -regular mappings, as in Lemma 6.8. For these all that was really needed was the

fact that the differential has rank at least t at any point where it exists (as in (6.6)), and we have already looked at this in the present setting.

In short, this is why the results and observations about (m, t) -regular mappings from Sections 6, 7, and 8 carry over to mappings that are Lipschitz and (m, t) -noncollapsing. This includes Theorem 6.1, Corollary 6.5, and Proposition 8.1, as in the statement of Proposition 9.7. In particular, these assertions apply to mappings from \mathbf{R}^m to \mathbf{R}^t which are Lipschitz and co-Lipschitz, in the sense of [6] (since co-Lipschitz implies (m, t) -noncollapsing in this case).

Although mappings that are Lipschitz and (m, t) -noncollapsing have several properties in common with (m, t) -regular mappings in this way, it is not true that Lipschitz (m, t) -noncollapsing mappings are (m, t) -regular. A counterexample to this is given by an example in [12], namely, Example (j) on p. 869 of [12]. This occurs already for mappings on the real line. The basic construction uses “tent” mappings. Given a closed interval $I = [a, b]$ in \mathbf{R} , define a corresponding “tent” mapping $t_I(x)$ by setting

$$(9.18) \quad t_I(x) = \frac{b-a}{2} - \left| x - \frac{a+b}{2} \right|.$$

In other words, this mapping vanishes at the endpoints a, b of I , it takes the value $(b-a)/2 = |I|/2$ at the midpoint of I , and it is linear on the two halves of I (between the endpoints and the midpoint). By combining a lot of “tent” mappings like this, on dyadic intervals $[2^j, 2^{j+1}]$, $[-2^{j+1}, -2^j]$, $j \in \mathbf{Z}$, for instance, one can get mappings from the real line to itself which are Lipschitz and $(1, 1)$ -noncollapsing, but not $(1, 1)$ -regular. In particular, the mapping would take the value 0 at infinitely many points, while a $(1, 1)$ -regular mapping should take any given value only finitely many times.

10. A variant of Theorem 6.1

In Theorem 6.1, one chooses a location and scale in \mathbf{R}^m , as represented by a ball B , and then there is a linear mapping $\lambda: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$ so that the combined mapping $(f, \lambda): \mathbf{R}^m \rightarrow \mathbf{R}^n \times \mathbf{R}^{m-t}$ is bilipschitz on a subset E of B which is of substantial proportion (in terms of Lebesgue measure). The choice of λ depends on B , even if the estimates do not.

Instead of this, one might try to find a single mapping $g: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$ so that the combined mapping $(f, g): \mathbf{R}^m \rightarrow \mathbf{R}^n \times \mathbf{R}^{m-t}$ has some kind of good behavior uniformly over all scales and locations at once. For this one should not be able to take g to be linear, in general, but one would still look for something like a Lipschitz condition.

As a basic example, $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ might be obtained from a bilipschitz mapping by throwing away $m - t$ coordinates. The idea would then be to try to recover those $m - t$ coordinates, or some reasonable versions of them.

There are results of this nature that one can get, and which will be discussed in this section. For simplicity (and brevity), we shall only give an outline of some of the main points, rather than precise statements and arguments (which could take a while).

A reasonable framework for these issues is provided by the notions of “weakly Lipschitz” and “weakly bilipschitz” mappings from [12]. These notions can be described roughly as follows. First, let us reformulate the usual Lipschitz condition by saying that a mapping $h: \mathbf{R}^m \rightarrow \mathbf{R}^k$ is Lipschitz if there is a constant C so that

$$(10.1) \quad \text{diam } h(B) \leq C \text{ diam } B$$

for all balls B in \mathbf{R}^m . It is easy to see that this is equivalent to the usual Lipschitz condition. For a weakly Lipschitz mapping, one asks that (10.1) hold for “most” balls B in \mathbf{R}^m , where “most” means that the exceptional balls $B(x, r)$ are parameterized by a set of pairs (x, r) in $\mathbf{R}^m \times (0, \infty)$ which is a Carleson set (as defined in the statement of Theorem 7.4).

For the exceptional balls, no condition at all is placed on the mapping. The idea is that the fact that the set of exceptional balls is small, in the sense of the Carleson condition, makes up for this.

The lack of restriction on the exceptional balls has the effect of allowing arbitrary behavior of the mapping on reasonably “thin” subsets of \mathbf{R}^m . For instance, if P is any plane in \mathbf{R}^m (of dimension strictly less than m), then the set of all balls $B(x, r)$ which intersect P corresponds to a Carleson set of pairs (x, r) in $\mathbf{R}^m \times (0, \infty)$. This is not hard to check, and it implies that a weakly Lipschitz mapping can have arbitrary behavior along such a plane P .

As a variant of this, suppose that P is a hyperplane in \mathbf{R}^m , so that $\mathbf{R}^m \setminus P$ consists of two components. If a given mapping h is Lipschitz on each of these two complementary half-planes, then it is automatically weakly Lipschitz on all of \mathbf{R}^m . This holds no matter how much the two Lipschitz mappings may behave differently along the common boundary. Note that h would be Lipschitz on all of \mathbf{R}^m under these conditions if it were continuous along the common boundary P .

Now let us consider bilipschitz conditions. Let us say that a mapping $h: \mathbf{R}^m \rightarrow \mathbf{R}^k$ is “approximately bilipschitz” on a ball B if it satisfies

a condition like (10.1), and also

$$(10.2) \quad |h(x) - h(y)| \geq C^{-1} |x - y|$$

whenever $x, y \in B$ satisfy $|x - y| \geq (\text{diam } B)/2$.

In other words, (10.2) provides the kind of lower bound for $|h(x) - h(y)|$ that one has for bilipschitz mappings, but only for x, y in B which are not too close together compared to the size of B . (The precise choice of $(\text{diam } B)/2$ in (10.2) is not too important, though.)

It is easy to check that h is bilipschitz in the usual sense exactly when it satisfies this kind of approximate bilipschitz condition for all balls B in \mathbf{R}^m , and with a uniform choice for the constant C (in (10.1) and (10.2)). As before, we shall say that h is *weakly bilipschitz* if there is a constant C so that the same condition of approximate bilipschitzness holds on “most” balls B , where “most” means that the collection of exceptional balls $B(z, r)$ should correspond to a set of pairs (z, r) in $\mathbf{R}^m \times (0, \infty)$ which is a Carleson set. For these exceptional balls, one does not ask for any condition on h .

As for weakly Lipschitz mappings, weakly bilipschitz mappings can have arbitrary behavior along “thin” subsets of \mathbf{R}^m , such as planes (of dimension strictly less than m). In particular, one can obtain weakly bilipschitz mappings by combining ordinary bilipschitz mappings on two half-spaces as before, without any conditions on how they might match up on the common boundary. Of course, one can have singular behavior for weakly Lipschitz or bilipschitz mappings which is more diffuse than that, but this construction illustrates some of the basic points.

This formulation of weak bilipschitzness is slightly different from the one [12], but the difference is not significant. (E.g., one uses cubes in [12] instead of balls, and the definition is given in a way that accommodates more general spaces.)

Although weakly Lipschitz and bilipschitz mappings can have essentially arbitrary behavior on sufficiently thin sets in \mathbf{R}^m , their average behavior on sets of positive measure is more like that of ordinary Lipschitz and bilipschitz mappings. Results of this nature are given in [12] (and have their genesis in arguments in [21]).

With weakly Lipschitz and bilipschitz mappings one has more flexibility for making constructions than for ordinary Lipschitz and bilipschitz mappings. In the present setting, one gets a more global version of Theorem 6.1, in which the conclusion is that there is a weakly Lipschitz mapping $g: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$ so that the combined mapping $(f, g): \mathbf{R}^m \rightarrow \mathbf{R}^n \times \mathbf{R}^{m-t}$ is weakly bilipschitz.

This applies to mappings $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ which are (m, t) -regular, as in Theorem 6.1, and more generally to mappings f which are Lipschitz and (m, t) -noncollapsing, as in Section 9. In the special case where $m = t$, one does not need an extra mapping g , and the conclusion is that f itself is weakly bilipschitz. This case is discussed in [12] (and is closely related to [9], [21]). (See Examples (g) and (i) on p. 869 of [12].)

In general, when $t < m$, how might one produce such a complementary mapping g ? The basic idea is to use mappings λ as in Theorem 6.1 as the initial ingredients, and then to combine these mappings at different scales and locations to get g . One has to do these things with some care, and in particular one should not try to combine too many of these initial mappings. Otherwise, the estimates will not work properly, like the Carleson conditions in the weak Lipschitz and bilipschitz properties.

One could try to do this by iterating the kind of information that one gets in Theorem 6.1. Instead of this, let us indicate a more direct method, in which one uses arguments like those in the proof of Theorem 6.1, together with some extra information.

Part of Section 7 already fits nicely with the present discussion. Namely, Theorem 7.4 already provides for the existence of good affine approximations for our mapping $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ (which is Lipschitz by assumption) on all balls B in \mathbf{R}^m , except for a collection of balls $B(x, r)$ corresponding to a Carleson set in $\mathbf{R}^m \times (0, \infty)$.

As in Lemma 7.16, we also have good bounds for the t -dimensional nondegeneracy of the linear parts of the affine approximations that come from Theorem 7.4, at least if the parameter ϵ in Theorem 7.4 is chosen small enough. How small ϵ needs to be depends only on the dimensions and (m, t) -regularity constants for f (or Lipschitz and (m, t) -noncollapsing constants, as in Section 9).

Suppose that B is a ball for which one has a good affine approximation for f like this. If ϵ is small enough, one can then get approximate bilipschitzness on B , in the sense of (10.1) and (10.2), by combining f with a linear mapping from \mathbf{R}^m to \mathbf{R}^{m-t} which complements the linear part of the affine approximation to f on B in a suitable way. We did something very similar to this in Section 7, with the choice of λ in (7.43), and in Lemma 7.44. In particular, in first part of the proof of Lemma 7.44, we saw that the linear mapping $(\pi \circ L, \lambda): \mathbf{R}^m \rightarrow P \times \mathbf{R}^{m-t}$ is invertible, with a bound for the norm of its inverse. (Here L is the linear part of the affine approximation to f on B , π is a certain projection, and P is a t -dimensional plane in \mathbf{R}^n .) If ϵ is small enough, depending only on suitable constants, then this leads to the desired approximate bilipschitz property of f on B , and with uniform bounds for the constants.

For the present purposes, the problem with this is that the choice of the linear mapping $\lambda: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$ depends on B . However, there is some extra information that we can bring in, coming from Carleson's Corona construction. This has the effect of saying that the affine approximations to f can be chosen in such a way that their linear parts do not change too often, while still maintaining a good degree of approximation to f , as in Theorem 7.4. The precise statement for this is a little bit complicated, but it basically says that the set of pairs (x, r) in $\mathbf{R}^m \times (0, \infty)$ around which the linear parts of the affine approximations have to change is a Carleson set.

Carleson's Corona construction is actually more directly concerned with the behavior of averages of bounded measurable functions. In the context of affine approximations for a Lipschitz mapping f , one would look at the Corona construction in connection with the differential of f , which gives a bounded measurable function. (The differential is a matrix-valued function, but that is okay.) An excellent reference concerning the Corona construction and related results is [18], especially Chapter VIII. Theorem 6.1 and Section 6 in general in Chapter VIII of [18] are particularly relevant and useful here. A version of this is also reviewed in Chapter 2 of Part IV of [13], especially Section 2.2 there. A translation to the setting of Lipschitz functions and affine approximations of them is also provided there.

At any rate, with this one is able to get some information about the dependence of the "complementary" linear mappings $\lambda = \lambda_B: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$ on the ball B . In particular, one finds that these mappings do not have to be changed too often, as one varies the ball B (in terms of both its center and its radius). The occasions when the λ_B 's have to be changed is controlled by a Carleson set of pairs (x, r) in $\mathbf{R}^m \times (0, \infty)$. This is a crucial point for making more global constructions, with weakly Lipschitz and bilipschitz mappings.

We shall not go into details about the constructions involved, but let us mention a few of the main points. The first is that it is more convenient to work with dyadic cubes instead of balls. Imagine that Q_0 is some dyadic cube, and that $\lambda_0: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$ is a linear mapping associated to it, as above. In particular, let us imagine that the combined mapping $(f, \lambda_0): \mathbf{R}^m \rightarrow \mathbf{R}^n \times \mathbf{R}^{m-t}$ is approximately bilipschitz around Q_0 , in the same sense as described before for balls, in (10.1) and (10.2). (Actually, one would normally ask for behavior like this on something like the double of Q_0 , as in [12].)

The information that we have about the λ 's not changing too often implies that this linear mapping λ_0 will normally work well not only

for Q_0 , but for many of its dyadic subcubes too. In general, λ_0 would not work for *all* of the dyadic subcubes of Q_0 , however. As one goes “down” through the locations and scales, one would have to “stop” at various subcubes of Q_0 . One would then want to start over again, for the purpose of constructing a global mapping $g: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$ from the various λ 's.

For some cubes there is trouble, simply because f does not have any sufficiently-good affine approximation (on the double of the cube, say) with which to work. As before, we know that this does not happen too often, with the exceptional locations and scales controlled by a Carleson condition. When one runs into cubes like these, one simply skips over them, without worrying about it too much, and goes on to cubes for which good affine approximations do exist.

Thus one gets to cubes for which there are sufficiently-good affine approximations (on the double of the cube), and hence for which there are corresponding complementary linear mappings $\lambda: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$. Imagine that we are working inside of the cube Q_0 , from before, and we have now arrived to a dyadic cube Q_1 contained in Q_0 , for which the linear mapping $\lambda_0: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$ that we have already does *not* work so well. That is, λ_0 does not provide a good complement to f in terms of having approximate bilipschitz properties around Q_1 , even if it might do so at larger scales in Q_0 . Assume however that a new linear mapping $\lambda_1: \mathbf{R}^m \rightarrow \mathbf{R}^{m-t}$ does behave well in this way. We would like to combine λ_1 with λ_0 on Q_1 , in such a manner as to keep the good properties of λ_0 at other scales and locations in Q_0 where λ_0 works fine, while bringing in the new mapping λ_1 for use inside of Q_1 .

To be more precise, we shall focus (in a moment) on what happens for a single cube Q_1 inside Q_0 like this, but normally there will be many cubes in this situation. These “stopping places” inside Q_0 will also occur at many different scales in general. I.e., the earlier choice of λ_0 will stop working at variable scales and locations in Q_0 . The cubes Q_1 that arise in this manner will have disjoint interiors, by construction. (Essentially one takes them to be as large as possible.)

It turns out that one does not have to worry too much about what happens across cubes like the Q_1 's, inside Q_0 . This is because the notions of weakly Lipschitz and weakly bilipschitz functions allow for sets of locations and scales on which one does not have information about the given mapping, at least if these sets satisfy Carleson conditions. We shall return to this later. At any rate, the basic point is that the problems that occur in going across the Q_1 's can be included in exceptional sets of locations and scales like these. In particular, one does not have to try

to smooth out these transitions, and this is a useful feature in working with weakly Lipschitz and bilipschitz mappings.

One does have to be careful about what happens inside the individual Q_1 's, and what happens in the original cube Q_0 at locations and scales which lie above the ones described by the Q_1 's. To see what happens for these issues, let us fix a single Q_1 , as before, and just look at it. In the actual construction, one would deal with all of the Q_1 's in the same way, in parallel.

Let c_{Q_1} denote the center of Q_1 . On Q_1 , let us imagine replacing our original mapping $\lambda_0(x)$ with

$$(10.3) \quad \tilde{\lambda}_0(x) = \lambda_0(c_{Q_1}) + (\lambda_1(x) - \lambda_1(c_{Q_1})).$$

From the point of view of scales larger than Q_1 , this function still looks like $\lambda_0(x)$. To make this precise, notice that

$$(10.4) \quad \sup_{x \in Q_1} |\lambda_0(x) - \lambda_0(c_{Q_1})| \leq C \operatorname{diam} Q_1$$

and

$$(10.5) \quad \sup_{x \in Q_1} |\lambda_1(x) - \lambda_1(c_{Q_1})| \leq C \operatorname{diam} Q_1$$

for a suitable constant C (which in fact can normally be taken to be 1). This is because the norms of λ_0 and λ_1 as linear mappings are bounded (and can be taken to be bounded by 1), by the way that the λ 's are chosen. Combining (10.4) and (10.5), we get that

$$(10.6) \quad \sup_{x \in Q_1} |\tilde{\lambda}_0(x) - \lambda_0(x)| \leq 2C \operatorname{diam} Q_1.$$

Thus $\tilde{\lambda}_0(x)$ and $\lambda_0(x)$ are nearly the same on Q_1 , up to errors which are comparable to $\operatorname{diam} Q_1$ (which would be small at larger scales).

On the other hand, $\tilde{\lambda}_0(x)$ is also close to $\lambda_1(x)$ on Q_1 , in the sense that

$$(10.7) \quad \tilde{\lambda}_0(x) - \lambda_1(x) \quad \text{is constant on } Q_1.$$

For the purposes of Lipschitz and bilipschitz conditions, $\tilde{\lambda}_1(x)$ is essentially the same as $\lambda_1(x)$, because of this.

This is the basic method by which one puts mappings like the λ 's on top of each other for this construction. Here one should modify λ_0 in this manner for all of the Q_1 's in Q_0 , in parallel, as mentioned earlier. In each of the Q_1 's, one repeats the process, with Q_1 in the role that Q_0 had before, and using $\tilde{\lambda}_0(x)$, as above. For the internal properties of Q_1 , this is practically the same as $\lambda_1(x)$, because of (10.7).

One can organize this construction in such a way as to build a global mapping on \mathbf{R}^m . As one goes through successive repetitions of the procedure, one makes many modifications along these lines. At each step, an individual modification is localized to some cube, and the size of the modification is bounded in terms of the diameter of this cube. This ensures that the modifications add up in reasonable ways as one goes through the generations of the construction, i.e., they are controlled by geometric series.

(Let us mention that one never runs up to infinity in scales, but instead one starts on some dyadic cubes in \mathbf{R}^m and goes down the scales inside them. These starting cubes are chosen in such a way that every dyadic cube in \mathbf{R}^m is contained in one of them, except for a collection of cubes that satisfies a Carleson condition. Basically, the exceptional cubes can be chosen to lie near a single point, like the origin. Because of the Carleson condition, one does not have to worry about the cubes in \mathbf{R}^m that are skipped in this way.)

For Lipschitz and bilipschitz conditions, these modifications do cause trouble at some locations and scales. In particular, one can have trouble along the boundaries of cubes like the Q_1 's above. More precisely, in replacing $\lambda_0(x)$ with $\tilde{\lambda}_0(x)$ as above, one has good properties away from the Q_1 's, at scales larger than the Q_1 's in Q_0 , and inside the individual Q_1 's, but one does not normally maintain good properties for measurements that cross the boundary of Q_1 (i.e., as in Lipschitz or bilipschitz conditions).

The main point is then to control the total collection of scales and locations for which this kind of trouble occurs, in terms of Carleson conditions. This uses the Carleson conditions that we have from the beginning, and discussed before, i.e., the Carleson conditions for how often the λ 's need to be changed, and for the exceptional scales and locations at which sufficiently-good affine approximations of the original Lipschitz mapping f do not exist (so that one may not have suitable λ 's to begin with). These conditions control how often the Q_0 's and Q_1 's come up, which is to say, the cubes at which one starts and stops, as well as some nearby cubes over which one might skip. The estimates for the total collection mentioned above also uses some elementary observations about obtaining Carleson conditions for sets of locations and scales associated to other such sets which are already known to satisfy Carleson conditions.

For instance, if some set of locations and scales satisfies a Carleson condition, then so do collections of locations and scales which are not too far from the ones in the first set. Also, if $\{Q_i\}_i$ is a family of dyadic cubes

in \mathbf{R}^m which satisfies a Carleson condition, then one also has Carleson conditions for sets of locations and scales that lie near the boundaries of the Q_i 's. In other words, this includes locations and scales near the locations and scales of the Q_i 's themselves, and also ones where the scales are much smaller than that, as long as the locations are close to the boundaries of the Q_i 's. This observation uses the fact that the Q_i 's have reasonably "small boundaries". In particular, the set of locations and scales near the boundary of a single dyadic cube satisfies a Carleson condition. This is easy to check.

This gives an outline of how the construction works. Note that this argument does not use topological degree theory, unlike the one before (in the proof of Lemma 7.44). One makes up for this by using affine approximations of f in stronger ways.

References

- [1] L. AHLFORS, Zur theorie der überlagerungsflächen, *Acta Math.* **65** (1935), 157–194.
- [2] P. ASSOUD, Plongements lipschitziens dans \mathbf{R}^n , *Bull. Soc. Math. France* **111**(4) (1983), 429–448.
- [3] S. M. BATES, On the image size of singular maps. I, *Proc. Amer. Math. Soc.* **114**(3) (1992), 699–705.
- [4] S. M. BATES, Toward a precise smoothness hypothesis in Sard's theorem, *Proc. Amer. Math. Soc.* **117**(1) (1993), 279–283.
- [5] S. M. BATES, On smooth rank-1 mappings of Banach spaces onto the plane, *J. Differential Geom.* **37**(3) (1993), 729–733.
- [6] S. M. BATES, W. B. JOHNSON, J. LINDENSTRAUSS, D. PREISS AND G. SCHECHTMAN, Affine approximation of Lipschitz functions and nonlinear quotients, *Geom. Funct. Anal.* **9**(6) (1999), 1092–1127.
- [7] G. DAVID, Opérateurs intégraux singuliers sur certaines courbes du plan complexe, *Ann. Sci. École Norm. Sup. (4)* **17**(1) (1984), 157–189.
- [8] G. DAVID, Opérateurs d'intégrale singulière sur les surfaces régulières, *Ann. Sci. École Norm. Sup. (4)* **21**(2) (1988), 225–258.
- [9] G. DAVID, Morceaux de graphes lipschitziens et intégrales singulières sur une surface, *Rev. Mat. Iberoamericana* **4**(1) (1988), 73–114.
- [10] G. DAVID, "Wavelets and singular integrals on curves and surfaces", Lecture Notes in Mathematics **1465**, Springer-Verlag, Berlin, 1991.

- [11] G. DAVID AND S. SEMMES, Singular integrals and rectifiable sets in \mathbf{R}^n : Beyond Lipschitz graphs, *Astérisque* **193** (1991), 152.
- [12] G. DAVID AND S. SEMMES, Quantitative rectifiability and Lipschitz mappings, *Trans. Amer. Math. Soc.* **337(2)** (1993), 855–889.
- [13] G. DAVID AND S. SEMMES, “*Analysis of and on uniformly rectifiable sets*”, Mathematical Surveys and Monographs **38**, American Mathematical Society, Providence, RI, 1993.
- [14] G. DAVID AND S. SEMMES, “*Fractured fractals and broken dreams. self-similar geometry through metric and measure*”, Oxford Lecture Series in Mathematics and its Applications **7**, The Clarendon Press Oxford University Press, New York, 1997.
- [15] J. R. DORRONSORO, A characterization of potential spaces, *Proc. Amer. Math. Soc.* **95(1)** (1985), 21–31.
- [16] K. J. FALCONER, “*The geometry of fractal sets*”, Cambridge Tracts in Mathematics **85**, Cambridge University Press, Cambridge, 1986.
- [17] H. FEDERER, “*Geometric measure theory*”, Springer-Verlag New York Inc., New York, 1969.
- [18] J. B. GARNETT, “*Bounded analytic functions*”, Pure and Applied Mathematics **96**, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1981.
- [19] J. HEINONEN AND I. HOLOPAINEN, Quasiregular maps on Carnot groups, *J. Geom. Anal.* **7(1)** (1997), 109–148.
- [20] J. HEINONEN AND S. RICKMAN, Quasiregular maps $\mathbf{S}^3 \rightarrow \mathbf{S}^3$ with wild branch sets, *Topology* **37(1)** (1998), 1–24.
- [21] P. W. JONES, Lipschitz and bi-Lipschitz functions, *Rev. Mat. Iberoamericana* **4(1)** (1988), 115–121.
- [22] R. KAUFMAN, A singular map of a cube onto a square, *J. Differential Geom.* **14(4)** (1979), 593–594.
- [23] J. LUUKKAINEN, Assouad dimension: antifractal metrization, porous sets, and homogeneous measures, *J. Korean Math. Soc.* **35(1)** (1998), 23–76.
- [24] O. MARTIO AND J. VÄISÄLÄ, Elliptic equations and maps of bounded length distortion, *Math. Ann.* **282(3)** (1988), 423–443.
- [25] P. MATTILA, “*Geometry of sets and measures in Euclidean spaces*”, Cambridge Studies in Advanced Mathematics **44**, Cambridge University Press, Cambridge, 1995.
- [26] L. NIRENBERG, “*Topics in nonlinear functional analysis*”, Courant Institute of Mathematical Sciences, New York University, New York, 1973–1974.

- [27] Y. G. RESHETNYAK, “*Space mappings with bounded distortion*”, Translations of Mathematical Monographs **73**, American Mathematical Society, Providence, RI, 1989.
- [28] S. RICKMAN, “*Quasiregular mappings*”, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)] **26**, Springer-Verlag, Berlin, 1993.
- [29] S. SEMMES, On the nonexistence of bilipschitz parametrizations and geometric problems about A_∞ weights, *Rev. Mat. Iberoamericana* **12** (1996), 337–410.
- [30] S. SEMMES, Metric spaces and mappings seen at many scales, Appendix in “*Metric Structures for Riemannian and non-Riemannian Spaces*”, by M. Gromov et al., Progress in Mathematics **152**, Birkhäuser Boston, Inc., Boston, MA, 1999.
- [31] S. SEMMES, Some novel types of fractal geometry, Preprint.
- [32] E. M. STEIN, “*Singular integrals and differentiability properties of functions*”, Princeton Mathematical Series **30**, Princeton University Press, Princeton, N.J., 1970.
- [33] Y. YOMDIN, Surjective mappings whose differential is nowhere surjective, *Proc. Amer. Math. Soc.* **111**(1) (1991), 267–270.

G. David:

Institut universitaire de France et
Département de Mathématiques
Bât. 425
Université de Paris-Sud
91405 Orsay Cedex
France
E-mail address: Guy.David@math.u-psud.fr

S. Semmes:

Department of Mathematics
Rice University
Box 1892
Houston TX 77251–1892
U.S.A.
E-mail address: semmes@math.rice.edu

Primera versió rebuda el 6 de juliol de 1999,
darrera versió rebuda el 22 de febrer de 2000.