TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 337, Number 2, June 1993

QUANTITATIVE RECTIFIABILITY AND LIPSCHITZ MAPPINGS

GUY DAVID AND STEPHEN SEMMES

ABSTRACT. The classical notion of rectifiability of sets in $\mathbf{R}^n$ is qualitative in nature, and in this paper we are concerned with quantitative versions of it. This issue arises in connection with L^p estimates for singular integral operators on sets in $\mathbf{R}^n$. We give a criterion for one reasonably natural quantitative rectifiability condition to hold, and we use it to give a new proof of a theorem in [D3]. We also give some results on the geometric properties of a certain class of sets in $\mathbf{R}^n$ which can be viewed as generalized hypersurfaces. Along the way we shall encounter some questions concerning the behavior of Lipschitz functions, with regard to approximation by affine functions in particular. We shall also discuss an amusing variation of the classical Lipschitz and bilipschitz conditions, which allow some singularities forbidden by the classical conditions while still forcing good behavior on substantial sets.

1. INTRODUCTION

A set E in $\mathbf{R}^n$ with Hausdorff dimension d, $d = 1, 2, \dots, n-1$, is said to be rectifiable if it is contained in the union of a countable family of d-dimensional Lipschitz graphs, except for a set of (d-dimensional) Hausdorff measure zero. [By a d-dimensional Lipschitz graph we mean a set of the form

$$\{p + A(p) : p \in P\}, \tag{1.1}$$

where P is a d-plane, A is a mapping of P into an $(n-d)$-plane Q in $\mathbf{R}^n$ that is orthogonal to P, and A is Lipschitz, i.e., there is a constant $C > 0$ such that

$$|A(p_1) - A(p_2)| \leq C|p_1 - p_2| \tag{1.2}$$

for all p_1, $p_2 \in P$.] Rectifiability is a qualitative condition, in that it does not come with bounds. For some purposes it is more appropriate to work with analogues of the notion of rectifiability that are quantitative. This occurs, for instance, when one tries to find conditions on E under which there are L^2 estimates for certain singular integral operators on E.

Let us give an example of a natural quantitative rectifiability condition. First we state an auxiliary definition.

Received by the editors March 11, 1991.
1991 *Mathematics Subject Classification*. Primary 49Q15, 42B20; Secondary 28A75.
Key words and phrases. Quantitative rectifiability, Lipschitz mappings.
The second author is partially supported by the NSF, the Alfred P. Sloan Foundation, and the Marian and Speros Martel Foundation.

Definition 1.3. *A set E in $\mathbf{R}^n$ is said to be regular with dimension d if it is closed and if there is a constant $C_0 > 0$ such that*

$$C_0^{-1} R^d \leq H^d(E \cap B(x, R)) \leq C_0 R^d \tag{1.4}$$

for all $x \in E$ and $R > 0$.

Here H^d denotes d-dimensional Hausdorff measure, and $B(x, R)$ is the ball with center x and radius R. We shall often simply write $|A|$ for $H^d(A)$.

The requirement that E be regular with dimension d can be viewed as a quantitative version of the property of having upper and lower densities with respect to H^d that are positive and finite. Note that (1.4) behaves naturally under dilations, as do the sort of quantitative rectifiability conditions that we have in mind.

We shall assume throughout this paper that E is regular with dimension d, $0 < d < n$, and that d is an integer, although we shall sometimes state this explicitly for emphasis. We shall also let μ denote the restriction of H^d to E.

Definition 1.5. E has big pieces of Lipschitz graphs (BPLG) if E is regular with dimension d and if there are constants M, $\theta > 0$ such that for each $x \in E$ and $R > 0$ there is a Lipschitz graph Γ of the form (1.1) such that

$$|E \cap \Gamma \cap B(x, R)| \geq \theta R^d \quad \text{and} \tag{1.6}$$

$$A \text{ has Lipschitz norm } \leq M. \tag{1.7}$$

It is not hard to show that E is rectifiable if it has BPLG, and that the converse is false. The main point is that the constants θ and M are not allowed to vary with x and R in Definition 1.5. If they were, then all rectifiable sets would satisfy the conditions of Definition 1.5.

If E has BPLG, then one can say a lot about estimates for singular integral operators on E. (See [D1, 2].) For instance, if $k(x)$ is an odd C^∞ function on $\mathbf{R}^n \backslash \{0\}$ such that

$$|\nabla^j k(x)| \leq C(j)|x|^{-d-j} \tag{1.8}$$

for $j = 0, 1, 2, \ldots$, and all $x \in \mathbf{R}^n \backslash \{0\}$, then

$$T_* f(x) = \sup_{\varepsilon > 0} \left| \int_{E \backslash B(x, \varepsilon)} k(x - y) f(y) d\mu(y) \right| \tag{1.9}$$

defines a bounded operator on $L^2(E)$. Something like BPLG also arises in connection with harmonic measure estimates, as in [DJ].

Sufficient conditions for a set E to satisfy BPLG have been given in [D1, 3] and [DJ]. The following is one such condition.

Definition 1.10. E satisfies Condition B if E is regular with codimension 1 (i.e., $d = n - 1$) and if there is a $C > 0$ so that for each $x \in E$ and $R > 0$ there exist y, z in different connected components of $\mathbf{R}^n \backslash E$ such that

$$\begin{gathered} |x - y| \leq R, \quad |x - z| \leq R, \quad \operatorname{dist}(y, E) \geq C^{-1} R, \text{ and} \\ \operatorname{dist}(z, E) \geq C^{-1} R. \end{gathered} \tag{1.11}$$

The requirement that y and z lie in different complementary components of E is crucial; without it all regular sets would satisfy Condition B.

Notice that Lipschitz graphs satisfy Condition B, as do chord-arc curves in the plane. More generally any bilipschitz image of $\mathbf{R}^d$ inside $\mathbf{R}^{d+1}$ satisfies Condition B. (This uses in part a theorem of Vaisälä [V].) On the other hand Condition B prohibits cusps. It does allow a certain amount of pinching off, but Theorem 1.20 below implies that there cannot be too much pinching off.

It was proved in [D3] that E has BPLG if it satisfies Condition B. See also [DJ], where this result is strengthened and given a simpler proof. It was previously shown in [S1] that if E satisfies Condition B and if E has exactly two complementary components, then a smaller class of singular integrals define bounded operators on $L^2(E)$. (A different approach to this was later given in [S4], which permitted a larger class of singular integral operators.) A version of Condition B for sets of larger codimension was also shown to imply BPLG in [D3].

An interesting open problem is to find other simple geometrical conditions that imply BPLG. Specific questions of this nature can be produced by formulating appropriate quantitative versions of classical results about rectifiable sets (as in [F, Ma]). An example of such a question is whether the following condition implies BPLG. Roughly speaking this condition requires that for each $x \in E$ and $R > 0$ there is a substantial set of d-planes through x onto which $E \cap B(x, R)$ has large projection.

Definition 1.12. E has big projections in plenty of directions if E is regular and if there is a $\delta > 0$ so that for each $x \in E$ and $R > 0$ there is a d-plane P_0 containing x such that

$$\left|\Pi_P(E \cap B(x, R))\right| \geq \delta R^d \tag{1.13}$$

for all d-planes P through x with $\operatorname{dist}(P, P_0) \leq \delta$ (for some reasonable choice of distance function on the space of d-planes that contain x). Here Π_P denotes the orthogonal projection onto P.

It seems to be a hard problem to decide whether this condition implies BPLG. If this were true then it would follow directly that Condition B implies BPLG, and it would also give similar results for some higher codimensional versions of Condition B. Notice that E must have big projections in plenty of directions if E has BPLG.

There is a variation of this problem which is easily resolved in the negative. We say that E has big projections if the same thing as in Definition 1.12 is true, except that we only require (1.13) to hold for a single d-plane through x, rather than a large set of d-planes. This condition does not imply BPLG, or even rectifiability. [It is not hard to find counterexamples. Probably the simplest are given by "cranks". (See [Mu].) One-dimensional Cantor sets will also work.] However, we do have the following.

Theorem 1.14. *If E is regular, has big projections, and satisfies the weak geometric lemma (defined below), then E has BPLG.*

Before defining the weak geometric lemma we make a few comments.

The weak geometric lemma is a kind of mild regularity condition on E which says that $E \cap B(x, R)$ is often well-approximated by a d-plane. It is not a very strong condition, but experience indicates that it is very useful in an auxiliary role. Theorem 1.14 is an example of this. Other examples arise in [S4 and DS].

One of the good features of the weak geometric lemma is that it can sometimes be verified without too much difficulty under natural hypotheses. This occurs in [DS], and it will happen again here: we are going to show that E satisfies the weak geometric lemma whenever Condition B holds. This together with Theorem 1.14 will provide a new proof of the result that Condition B implies BPLG.

Fortunately the proof of Theorem 1.14 is not very complicated. It uses essentially the same argument as given by Peter Jones in [J2].

Let us now proceed to the definition of the weak geometric lemma. We first define a certain quantity $\beta(x, t)$. Given $x \in E$ and $t > 0$, set

$$\beta(x, t) = \inf_P \sup_{y \in E \cap B(x,t)} t^{-1} \operatorname{dist}(y, P), \tag{1.15}$$

where the infimum is taken over all d-planes P in $\mathbf{R}^n$. [Throughout this paper we use the word "d-plane" to mean any d-dimensional affine subspace, not necessarily one that passes through the origin.]

Definition 1.16. E satisfies the weak geometric lemma (WGL) if it is regular and if

$$\{(x, t) \in E \times \mathbf{R}_+ : \beta(x, t) > \varepsilon\}$$

is a Carleson set for every $\varepsilon > 0$.

A measurable subset A of $E \times \mathbf{R}_+$ is said to be a Carleson set if

$$\chi_A(x, t) d\mu(x) \frac{dt}{t}$$

is a Carleson measure on $E \times \mathbf{R}_+$. [Recall that a Carleson measure on $E \times \mathbf{R}_+$ is a measure λ such that there is a $C > 0$ for which

$$|\lambda|((E \cap B(x, R)) \times (0, R)) \leq CR^d$$

for all $x \in E$ and $R > 0$.] In rough terms this means that A behaves like a d-dimensional set inside the $(d+1)$-dimensional space $E \times \mathbf{R}_+$, at least near $E \times \{0\}$.

The WGL is a variant of a condition that arose in some work of Peter Jones. He showed in [J1] that

$$\beta(x, t)^2 d\mu(x) \frac{dt}{t} \text{ is a Carleson measure on } E \times \mathbf{R}_+ \tag{1.17}$$

if E is a one-dimensional Lipschitz graph, and he also showed how this fact could be useful in controlling the Cauchy integral operator on E. This condition (1.17) has become known as the geometric lemma, and it has become clear that it is very natural and important. For example, Jones has shown in [J3] that a one-dimensional regular set E satisfies (1.17) if and only if it is contained in a curve that is itself a regular set.

The WGL is much weaker than the geometric lemma. It is easy to produce examples of one-dimensional regular sets that satisfy the WGL but are not rectifiable. (See [DS, Section 20].) The reason for this is the following: although the WGL ensures that the set E is often locally well-approximated by a d-plane, it provides very little control on how much these d-planes can spin around as you zoom in on a point. Rectifiability requires that these d-planes do not spin

around to much, e.g., at the points for which there is an approximate tangent plane.

Note that the converse to Theorem 1.14 is true: if E has BPLG, then it certainly has big projections, and it also satisfies the WGL. In the original version of this paper we said that there was not a simple proof known of this last fact, and we pointed out that it could be derived from [D1, 2 and DS]. Since then Peter Jones, the referee, and the authors have found direct proofs. We shall include such a proof in a later publication, and we shall also treat some of the obvious variations of this result as well. (See Chapter IV.2 in [DS2].)

As mentioned earlier we are going to show that Condition B implies the WGL, and we are going to do this by a direct geometrical argument. To do this we shall use the following criterion for the WGL to hold.

Proposition 1.18. *E satisfies the WGL if it satisfies the local symmetry condition LS.*

By definition E satisfies LS if it is regular and if

$$A_\varepsilon = \{(x, t) \in E \times \mathbf{R}_+ : \text{ there exist } u, v \in E \cap B(x, t) \text{ such that } \operatorname{dist}(2u - v, E) \geq \varepsilon t\} \tag{1.19}$$

is a Carleson set for each $\varepsilon > 0$. One should think of A_ε as being the set of (x, t) such that E is not approximately symmetric about each point in $E \cap B(x, t)$.

Proposition 1.18 is a consequence of Proposition 5.5 in [DS]. Although [DS] is long, the proof of Proposition 5.5 is pretty short, and it is also self-contained.

In addition to proving that LS holds when Condition B does we are going to show that Condition B implies another good property. This property will again be saying that certain subsets of $E \times \mathbf{R}_+$ are Carleson sets. For the rest of the introduction we assume that $n = d + 1$.

For each $\varepsilon > 0$ we define H_ε to be the set of (x, t) in $E \times \mathbf{R}_+$ for which there are $y_1, y_2 \in B(x, t)$ such that $\operatorname{dist}(y_i, E) \geq \varepsilon t$ for $i = 1, 2$, y_1 and y_2 lie in the same component of $\mathbf{R}^{d+1} \backslash E$, and the line segment that joins y_1 to y_2 intersects E. (See Figure 1 for an example where this happens.) Thus H_ε measures the extent to which each complementary component of E is not approximately convex. This is, of course, closely related to measuring the extent to which E looks approximately like a d-plane.

There is one more family of subsets of $E \times \mathbf{R}_+$ that we are going to use, which is a minor variation of the H_ε's. Given $\varepsilon > 0$ let H'_ε be the set of (x, t) in $E \times \mathbf{R}_+$ for which there are $y_1, y_2 \in B(x, t)$ such that $d(y_i, E) \geq \varepsilon t$ for $i = 1, 2$, y_1 and y_2 belong to the same component of $\mathbf{R}^{d+1} \backslash E$, but $\frac{1}{2}(y_1 + y_2) \in E$. Clearly $H'_\varepsilon \subseteq H_\varepsilon$ for all ε.

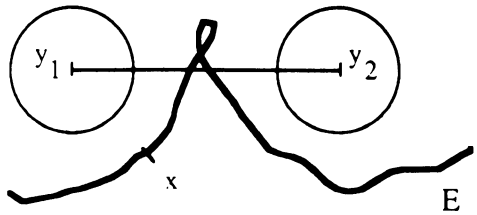

FIGURE 1

Theorem 1.20. *Suppose that E satisfies Condition* B. *Then H_ε is a Carleson set for all $\varepsilon > 0$, and E satisfies LS. Conversely, if E is regular, $d = n - 1$, and if H'_ε is a Carleson set for all $\varepsilon > 0$, then E satisfies Condition* B.

The fact that H_ε is a Carleson set for all $\varepsilon > 0$ was quite useful in [S3, 4]. It permitted the boundary values of some functions on $\mathbf{R}^n \backslash E$ to be controlled in terms of their second derivatives, by providing paths along which the derivatives could be integrated.

Theorems 1.14 and 1.20 and Proposition 1.18 combine to give a new (and not very complicated) proof of the fact that Condition B implies BPLG. To see this it suffices to check that Condition B forces E to have big projections. Let $x \in E$ and $R > 0$ be given, and let y, z be as in Definition 1.10. Take P to be the d-plane through x that is orthogonal to the line segment L that joins y to z. Then $\Pi_P(E \cap B(x, R))$ must contain the ball B in P centered at $\Pi_P(y)$ and having radius R/C (for some sufficiently large C). Indeed, every line parallel to L that goes through B must intersect two different components of $\mathbf{R}^{d+1} \backslash E$ inside $B(x, R)$, and must therefore intersect $E \cap B(x, R)$ as well. The fact that y and z lie in different components of $\mathbf{R}^{d+1} \backslash E$ is crucial here.

Although the result that Condition B implies BPLG provides very strong geometrical information about the sets E that satisfy Condition B, there is a lot that it does not tell us, particularly with regard to information that is partially topological in nature. Theorem 1.20 is a step in that direction, but it would be nice to go further. We would like to have some reasonable sense in which a set E that satisfies Condition B is approximately a submanifold, in a way that is not too degenerate.

To clarify this issue it is helpful to add hypotheses. Asking that $\mathbf{R}^{d+1} \backslash E$ have exactly two components would already be pretty good, but let us ask for more than that.

Definition 1.21. E is said to be a chord-arc surface if it satisfies Condition B, if $\mathbf{R}^{d+1} \backslash E$ has exactly two connected components, and if each of these components is an NTA domain in the sense of [**JK**].

Note that the assumption that the complementary domains are NTA includes the requirements of Condition B except for regularity. In the presence of Condition B the NTA assumptions reduce to the requirement that for each $A_0 > 0$ there is an $A_1 > 0$ so that if y_1, y_2 lie in the same complementary component of E, and if $\operatorname{dist}(y_i, E) \geq A_0^{-1}|y_1 - y_2|$ for $i = 1, 2$, then there is a curve joining y_1 to y_2 with diameter $\leq A_1|y_1 - y_2|$ and which stays at distance $\geq A_1^{-1}|y_1 - y_2|$ from E. Notice that Theorem 1.20 (the part about H_ε) implies that this is true for most y_1, y_2 if E satisfies Condition B.

One reason for the name "chord-arc surfaces" is that when $d = 1$ these are just the chord-arc curves in the plane, i.e., the curves for which the length of an arc is bounded by a constant times the distance between its endpoints. When $d > 1$ chord-arc surfaces satisfy an anologous condition, at least under some a priori smoothness assumptions, as we shall see in §6. Another reason for using this name is that these sets play much the same role for hypersurfaces in $\mathbf{R}^{d+1}$ as do chord-arc curves in the plane.

Unfortunately it is not at all clear to what extent chord-arc surfaces really are surfaces in the usual sense. For example, one could hope that when $d = 2$

there is some reasonable sense in which a chord-arc surface is a Riemann surface which may have an infinite number of handles, but for which the handles satisfy a Carleson-type packing condition. One could also seek results that say that E admits a well-behaved parameterization by $\mathbf{R}^d$ if it is a chord-arc surface that is homeomorphic to $\mathbf{R}^d$ and perhaps satisfies some other conditions. So far we have only been able to prove a result of this nature when $d = 2$.

Theorem 1.22. *Suppose that E is a two-dimensional simply-connected chord-arc surface. Assume a priori that E is smooth, and that $E \cup \{\infty\}$ is a smooth embedded submanifold of $S^3 \cong \mathbf{R}^3 \cup \{\infty\}$. Then there is a quasisymmetric embedding $\tau : \mathbf{R}^2 \to \mathbf{R}^3$ such that $\tau(\mathbf{R}^2) = E$, with estimates that do not depend quantitatively on the a priori smoothness assumptions on E.*

This can be viewed as a "large constant" version of Theorem 6.1 in [S2].

Recall that an embedding $\tau : \mathbf{R}^2 \to \mathbf{R}^3$ is said to be quasisymmetric if there is a continuous function $\eta : [0, \infty) \to [0, \infty)$ with $\eta(0) = 0$ such that

$$|\tau(y) - \tau(x)| \leq \eta(t)|\tau(z) - \tau(x)| \tag{1.23}$$

whenever $|y - x| \leq t|z - x|$. The theorem asserts the existence of such a τ with $\eta(\cdot)$ depending only on the constants that arise in the definition of a chord-arc surface.

Note that the Jacobian of τ is an A_∞ weight on $\mathbf{R}^2$ if τ is as in Theorem 1.22, with estimates that depend only on the chord-arc surface constants. This uses the same argument as in [G] for quasiconformal mappings on $\mathbf{R}^n$, together with the fact that E is regular.

The proof of Theorem 1.22 only works when $d = 2$ because it relies on the uniformization theorem. We do not know how to produce τ by a direct construction, which is closely related to the fact that we do not know how to do anything like this when $d > 2$. At least one of the authors thinks that the obvious analogue of Theorem 1.22 when $d > 2$ should be false, but neither author knows how to prove such a thing.

There is a partial converse to Theorem 1.22. If E is the image of a quasisymmetric embedding of $\mathbf{R}^2$ into $\mathbf{R}^3$, then E is certainly simply-connected, and it also has exactly two complementary components in $\mathbf{R}^3$ (by algebraic topology), each of which is an NTA domain. This last is due to J. Väisälä. (See (5.10) in [V].)

In connection with the results stated above we shall also consider the behavior of Lipschitz functions on E. For example, we shall encounter the problem of finding conditions on E under which there are estimates for affine approximation of Lipschitz functions on E like the estimates that are true when $E = \mathbf{R}^d$. We are also going to introduce a notion of "weakly Lipschitz functions", which will have content even when $E = \mathbf{R}^d$. This notion allows certain types of singularities that are forbidden to Lipschitz functions while still forcing some good behavior. In order to prove Theorem 1.22 we shall need to show in particular that the norm of a Lipschitz function on a chord-arc surface (which is assumed a priori to be smooth) is controlled by the L^∞ norm of its gradient.

Theorems 1.14, 1.20, and 1.22 are proved in §§2, 5, and 6, respectively. In §3 we discuss weakly Lipschitz and bilipschitz functions, and in §4 we give conditions on E under which the aforementioned estimates on affine approximation of Lipschitz functions are valid. These issues arise in §2 when we state

and prove a result that generalizes Theorem 1.14 by considering more general functions on E than just orthogonal projections.

We would like to thank J. C. Yoccoz for a helpful suggestion.

2. The proof of Theorem 1.14

We are going to derive Theorem 1.14 from a stronger result, Theorem 2.11 below. This stronger result will be modelled on the following theorem of Peter Jones [J2].

Theorem 2.1. *For each $d \geq 1$ and $\varepsilon > 0$ there is a constant $\tau > 0$ and an integer $M > 0$ for which the following is true. Let f be any function that maps a cube $Q_0 \subseteq \mathbf{R}^d$ into $\mathbf{R}^d$ and which is Lipschitz with norm ≤ 1. Then there exist compact sets F_j, $1 \leq j \leq M$, in Q_0 such that*

$$|f(x) - f(y)| \geq \tau|x - y| \quad \text{for all } x, y \in F_j \text{ and any } j, \tag{2.2}$$

and

$$|f(Q_0 \backslash (\cup F_j))| \leq \varepsilon |Q_0|. \tag{2.3}$$

Thus f is bilipschitz on each F_j, and the remainder set $Q_0 \backslash (\bigcup F_j)$ has small image under F. In particular, if $|f(Q_0)|$ is not too small, then the F_j's cannot all be too small.

The proof of Theorem 2.11 will be essentially the same as Jones' argument in [J2], but some complications arise as to its form. In [J2] a certain property of Lipschitz functions on $\mathbf{R}^d$ is used that is not true in general for Lipschitz functions on a regular set, and this property will have to be incorporated into the hypotheses of Theorem 2.11. Also, for the proof of Theorem 2.11, and, to a lesser extent, for its statement, it will be good to have an analogue of dyadic cubes for regular sets in $\mathbf{R}^n$. We discuss first this analogue of dyadic cubes.

Given a regular set E of dimension d in $\mathbf{R}^n$, it is possible to construct a family Δ_k, $k \in \mathbf{Z}$, of collections of measurable subsets of E with the following properties:

(2.4) for each $k \in \mathbf{Z}$, E is the disjoint union of the elements of Δ_k;

(2.5) if $k > k'$, $Q \in \Delta_k$, $Q' \in \Delta_{k'}$, then either $Q \cap Q' = \varnothing$ or $Q' \subseteq Q$;

(2.6) for each $k \in \mathbf{Z}$ and $Q \in \Delta_k$, Q contains the intersection of E with a ball centered on E having radius $\geq C^{-1}2^k$;

(2.7) for all $k \in \mathbf{Z}$ and $Q \in \Delta_k$ we have $\operatorname{diam} Q \leq C2^k$.

The constant C in (2.6) and (2.7) depends only on d, n, and the constant C_0 in (1.4). We shall write Δ for the union of the Δ_k's, and we shall refer to the elements of Δ as "cubes" in the context of subsets of E.

A construction of the Δ_k's can be found in [D3] or [D4]. (The reader is probably better off with [D4].) In fact the cubes can be built in such a way that they have small boundaries, in a certain sense that is often useful but irrelevant here. Relinquishing this requirement allows the construction of the cubes to be simplified substantially.

Of course there is no canonical family of Δ_k's for a given regular set. For our purposes it does not really matter which class of cubes we are using, as long as the properties above are satisfied. We shall assume for the rest of this paper

that a choice for a family of cubes Δ as above has been made whenever we are working with a regular set.

Notice that (2.6), (2.7), and (1.4) imply that

$$C^{-1}2^{kd} \leq |Q| \leq C2^{kd} \tag{2.8}$$

whenever $Q \in \Delta_k$. This implies in particular that a cube can have at most a bounded number of children; that is, if $Q \in \Delta_k$, then Q contains at most C elements of Δ_{k-1}.

Given $Q \in \Delta$ and $\lambda > 1$, set

$$\lambda Q = \{x \in E : \operatorname{dist}(x, Q) \leq (\lambda - 1) \operatorname{diam} Q\}.$$

This notation is convenient and will be used frequently.

Before stating Theorem 2.11 we need to introduce some notation concerning the extent to which a function can be approximated by affine functions. Let E be, as always, a regular set of dimension d in $\mathbf{R}^n$. Fix a cube Q_0 in E, and let $\Delta(Q_0)$ denote the set of cubes contained in Q_0. Let f be a function on Q_0 that takes values in $\mathbf{R}^d$.

Given $\alpha > 0$ (small) and $K > 0$ (large), let $A = A(f, \alpha, K)$ denote the set of cubes $Q \in \Delta(Q_0)$ for which there exists an affine function $a : \mathbf{R}^n \to \mathbf{R}^d$ such that $|\nabla a| \leq K$ and

$$\sup_{2Q \cap Q_0} |f - a| \leq \alpha \operatorname{diam} Q. \tag{2.9}$$

Thus A (for "approximable") consists of the cubes on which f is well-approximated by a nice affine function. (If we did not insist that $|\nabla a| \leq K$, we would be allowing f to consort with bad affine functions. This is not a serious issue when E is a d-plane.)

Define $NA = NA(f, \alpha, K)$ to be $\Delta(Q_0) \backslash A$, so that NA consists of the cubes on which f is not so well approximable. In practice we are going to need to have some control on NA, in the form of a bound on

$$|Q_0|^{-1} \sum_{Q \in NA(f, \alpha, K)} |Q| \tag{2.10}$$

for some $K > 0$ and a sufficiently small α.

When $E = \mathbf{R}^d$ and f is a Lipschitz function with norm 1, then it is known that there is a $K > 0$ such that (2.10) is at most $C(\alpha)$ for all $\alpha > 0$. We shall explain why this is true in Section 4, and we shall give other conditions on E under which this is still true.

Theorem 2.11. *Suppose that E is a d-dimensional regular set that satisfies the WGL, and let N_0, K, and $\varepsilon > 0$ be given. Then there exist positive constants α, τ, and M such that the following is true, with α depending only on d, n, ε, and the regularity constant for E, with τ depending on these constants and also on K, and with M depending on all these constants as well as N_0 and the constants from the WGL.*

Suppose that Q_0 is a cube in E and that $f : Q_0 \to \mathbf{R}^d$ is Lipschitz with norm 1. Assume also that (2.10) *is less than N_0. Then there are closed subsets F_j of Q_0, $1 \leq j \leq M$, such that* (2.2) *and* (2.3) *hold.*

In this theorem it is important that α does not depend on N_0, because in practice N_0 will depend on α. The only price you pay for increasing N_0 is that M must then be made larger.

The analogue of the theorem with cubes replaced by "balls" in E—i.e., sets of the form $E \cap B(x, R)$, $x \in E$—is also true. We used cubes in the statement of the theorem because we shall be using cubes in the proof.

As in [J2], there is a version of this result in the case where f takes values in $\mathbf{R}^m$ for some $m > d$. The only change required is that (2.3) should be replaced by the condition that $f(Q_0 \backslash (\bigcup F_j))$ has d-dimensional Hausdorff content less than $\varepsilon |Q_0|$. [Recall that d-dimensional Hausdorff content h_d is defined by

$$h_d(A) = \inf \sum r_j^d, \tag{2.12}$$

where the infimum is taken over all coverings of A by a sequence of balls B_j with radii r_j. (It is probably customary to include a normalizing constant in the definition of h_d, but we shall not bother.) Unlike Hausdorff measure, in this definition we allow coverings by balls that are not small. If A is contained inside a d-dimensional regular set, then the d-dimensional Hausdorff content is comparable to the Hausdorff measure.]

Before proving Theorem 2.11 let us explain why it implies Theorem 1.14. Suppose that E satisfies hypotheses of Theorem 1.14. Fix $x \in E$ and $R > 0$, and let P be a d-plane such that (1.13) holds with R replaced by $R/2$. Then there exist $\varepsilon > 0$ and $Q_0 \in \Delta$ such that $Q_0 \subseteq B(x, R) \cap E$, $\operatorname{diam} Q_0 \geq \varepsilon R$, and $|\Pi_P(Q_0)| \geq 2\varepsilon |Q_0|$. Apply Theorem 2.11, with $f = \Pi_P$. (Note that the hypotheses of Theorem 2.11 simplify substantially, since f is affine.) Then from (2.3) we obtain

$$|f(\bigcup F_j)| \;\geq\; |f(Q_0)| - \varepsilon |Q_0| \;\geq\; \varepsilon |Q_0|,$$

and hence $|f(F_i)| \geq M^{-1}\varepsilon |Q_0|$ for some i, $1 \leq i \leq M$. Because (2.2) holds on F_i, we get that F_i can be put into the form

$$F_i = \{p + A(p) : p \in f(F_i)\},$$

where A is a Lipschitz map of $f(F_i) \subseteq P$ into an $(n-d)$-plane Q orthogonal to P. By the Whitney extension theorem (see [St]) we can extend A to a Lipschitz map of P into Q. Because $|F_i| \geq |f(F_i)| \geq M^{-1}\varepsilon |Q_0|$, we conclude that E does indeed have BPLG, as desired.

The rest of this section will be devoted to the proof of Theorem 2.11.

It will be helpful for us to introduce a certain geometric constant, which will also be used in the next section. There is a $b > 0$ so that if x and y lie in E, and if Q is the smallest cube containing x such that $y \in 2Q$, then

$$|x - y| \;\geq\; 10b \operatorname{diam} Q. \tag{2.13}$$

We can choose b so that it depends only on d and the constants in (2.6) and (2.7).

Let N_0, K, and ε be given, as in Theorem 2.11. Let α and δ be small positive numbers, to be chosen later.

The first step of the proof is to break $\Delta(Q_0)$ up into two groups, of good cubes and bad cubes, with control on the number of bad cubes. The bad cubes come in three types, and we define $\mathscr{B}_j \subseteq \Delta(Q_0)$ for $j = 1$, 2, and 3 accordingly.

We put a cube $Q \in \Delta(Q_0)$ into $\mathscr{B}_1$ when

$$|f(Q)| \;\leq\; \frac{\varepsilon}{3}\, |Q|. \tag{2.14}$$

It is not hard to check that

$$\left| f\left(\bigcup_{\mathscr{B}_1} Q\right)\right| \le \frac{\varepsilon}{3}\left|\bigcup_{\mathscr{B}_1} Q\right| \le \frac{\varepsilon}{3}\,|Q_0|. \tag{2.15}$$

For the first inequality we used the fact that the union of the elements of $\mathscr{B}_1$ is the same as the union of the maximal elements, and that the maximal elements are pairwise disjoint.

Set $\mathscr{B}_2 = \{Q \in \Delta(Q_0)\colon \beta(Q) \ge \delta\}$, where

$$\beta(Q) = \inf_P \sup_{y \in 2Q} \operatorname{dist}(y, P)(\operatorname{diam} Q)^{-1} \tag{2.16}$$

is a minor variant of $\beta(x, t)$ in (1.15), with the infimum again taken over all d-planes P. It is not hard to check that the WGL implies that

$$\sum_{Q \in \mathscr{B}_2} |Q| \le C(\delta)|Q_0|. \tag{2.17}$$

We take $\mathscr{B}_3$ to be $NA(f, \alpha, K)$, so that

$$\sum_{Q \in \mathscr{B}_3} |Q| \le N_0|Q_0|, \tag{2.18}$$

by assumption.

We view $\bigcup_1^3 \mathscr{B}_j$ as the set of bad cubes, and we take $\mathscr{G} = \Delta(Q_0)\backslash(\bigcup_1^3 \mathscr{B}_j)$ to be the good cubes. The second step of the proof is to show that f satisfies a sort of weak bilipschitzness condition on the good cubes.

Lemma 2.19. *If α and δ are sufficiently small, then we can choose τ small enough so that for each $Q \in \mathscr{G}$ we have*

$$|f(x) - f(y)| \ge \tau|x - y|$$

whenever $x, y \in 2Q$ and $|x - y| \ge b \operatorname{diam} Q$.

Let $Q \in \mathscr{G}$ be given. Then $Q \in A(f, \alpha, K)$, and so there is an affine function $a : \mathbf{R}^n \to \mathbf{R}^d$ such that $|\nabla a| \le K$ and (2.9) holds. Since $\beta(Q) < \delta$ there is a d-plane P such that

$$\operatorname{dist}(z, P) \le \delta \operatorname{diam} Q \tag{2.20}$$

for all $z \in 2Q$. If we can show that there is a $\gamma = \gamma(\varepsilon, K)$ such that

$$|a(p) - a(q)| \ge \gamma|p - q| \tag{2.21}$$

for all $p, q \in P$, then the lemma follows easily from (2.9) and (2.20) if α and δ are small enough.

To prove (2.21) we use the fact that $|f(Q)| \ge \frac{\varepsilon}{3}|Q_0|$. The point is that because a is affine and we have a bound on $|\nabla a|$, (2.21) can fail only if a shrinks volumes a lot.

Fix an $x \in Q$, and set $B = B(x, \operatorname{diam} Q) \cap P$. Then B is a ball in P, and it has radius at least $\frac{1}{2}\operatorname{diam} Q$ if δ is small enough. The image $a(B)$ of B under a is an ellipsoid in $\mathbf{R}^d$. To prove (2.21) we need only get a lower bound on $|Q|^{-1}|a(B)|$.

From (2.9) and (2.20) we get that

$$f(Q) \subseteq \{z \in \mathbf{R}^d : \operatorname{dist}(z, a(B)) \leq C(\alpha + K\delta) \operatorname{diam} Q\},$$

and hence

$$|f(Q)| \leq |a(B)| + C(\alpha + K\delta)(\operatorname{diam} Q)^d \leq |a(B)| + \frac{\varepsilon}{10}|Q|$$

if α and δ are small enough. From this and the fact that $Q \notin \mathscr{B}_1$, so that (2.14) fails, we deduce that $|a(B)| \geq \frac{\varepsilon}{10}|Q|$. This implies (2.21) and proves the lemma.

The third step in the proof of Theorem 2.11 is to choose the subset of Q_0 that will eventually be subdivided into the F_j's. It is convenient to do this by specifying its complement in Q_0.

Choose α and δ, for the rest of the proof, to be small enough for the purposes of Lemma 2.19. Define $R_1 \subseteq Q_0$ by

$$R_1 = \bigcup_{\mathscr{B}_1} Q,$$

so that $|f(R_1)|$ is controlled by (2.15). Define R_2 as follows. Given any cube Q, let $\widehat{Q}$ denote the union of the cubes that lie in the same Δ_k as Q and which intersect $2Q$. Set

$$R_2 = \left\{ x \in Q_0 : \sum_{\mathscr{B}_2 \cup \mathscr{B}_3} \chi_{\widehat{Q}}(x) \geq L \right\},$$

where L is chosen so large that $|R_2| \leq \frac{\varepsilon}{3}|Q_0|$. It is possible to do this because

$$\int_{Q_0} \left(\sum_{\mathscr{B}_2 \cup \mathscr{B}_3} \chi_{\widehat{Q}}(x) \right) d\mu(x) \leq C \sum_{\mathscr{B}_2 \cup \mathscr{B}_3} |Q| \leq C(C(\delta) + N_0)|Q_0|,$$

by (2.17) and (2.18). Because f is Lipschitz with norm ≤ 1 we get that

$$|f(R_2)| \leq \frac{\varepsilon}{3}|Q_0|. \tag{2.22}$$

For the last step of the proof of Theorem 2.11 it suffices to find subsets $F_1, \ldots, F_M$ of Q_0 such that

$$Q_0 \backslash (R_1 \cup R_2) = \bigcup F_j \tag{2.23}$$

and (2.2) is satisfied. To get the F_j's promised by Theorem 2.11 you must then take the closures of these F_j's, but that does not disturb the other conditions. Because of Lemma 2.19 and the definition of b it is enough for $F_1, \ldots, F_M$ to satisfy (2.23) and the following property: if x and y lie in the same F_j, and if Q is the smallest cube containing x such that $y \in 2Q$, then $Q \in \mathscr{G}$.

A family of F_j's can be constructed using the same kind of coding argument as in [J2]. For the convenience of the reader we give now a brief sketch of this argument.

For each k and each cube Q_1 in $\Delta_k \cap \Delta(Q_0)$ let $\mathscr{F}(Q_1)$ be the set of cubes Q_2 in $\Delta_k \cap \Delta(Q_0)$ such that $Q_2 \neq Q_1$ and both Q_1 and Q_2 are contained in $\widehat{S}$ for some cube S in $\Delta_{k+l} \cap (\mathscr{B}_2 \cup \mathscr{B}_3)$. Here the constant l is chosen large

enough to ensure that $\operatorname{diam} Q_1 < b \operatorname{diam} S$. Notice that $\mathscr{F}(Q_1)$ has at most T elements, where T is a constant that depends only on the constants in (1.4), (2.6), and (2.7).

Let A be a set with $T+1$ elements. We view A as being an alphabet, and we shall refer to its elements as letters. An ordered finite set of letters from A will be called a word. We say that a word α begins with a word β if the length of β (call it m) is no greater than the length of α, and if the first m letters of α coincide with the corresponding letters of β.

It is possible to associate to each $Q \in \Delta(Q_0)$ a word $\alpha(Q)$ in such a way that this correspondence satisfies the following three properties. First, $\alpha(Q_0)$ is the empty word. Second, $\alpha(Q) = \alpha(Q^*)$ when $\mathscr{F}(Q) = \varnothing$, where Q^* denotes the parent of Q. Third, if $\mathscr{F}(Q) \neq \varnothing$, then $\alpha(Q)$ is obtained from $\alpha(Q^*)$ by adding a letter at the end, in such a way that if $Q' \in \mathscr{F}(Q)$, then:

(a) $\alpha(Q) \neq \alpha(Q')$ when $\alpha(Q)$ and $\alpha(Q')$ have the same length;
(b) $\alpha(Q)$ does not begin with $\alpha(Q')$ when $\alpha(Q)$ is longer than $\alpha(Q')$;
(c) $\alpha(Q')$ does not begin with $\alpha(Q)$ when $\alpha(Q)$ is shorter than $\alpha(Q')$.

The existence of such a correspondence can be obtained by a recursive procedure. That is, you define $\alpha(Q)$ for $Q \in \Delta_j \cap \Delta(Q_0)$ assuming that the correspondence has already been defined on $(\bigcup_{k>j} \Delta_k) \cap \Delta(Q_0)$, and you show that this can be done in such a way that the three properties above are preserved. We omit the details, which are pretty easy.

If $x \in Q_0 \backslash (R_1 \cup R_2)$, then there are fewer than L cubes Q that contain x and satisfy $\mathscr{F}(Q) \neq \varnothing$. For such an x we set $\alpha(x) = \alpha(Q)$, where Q contains x and is small enough so that $\alpha(Q) = \alpha(Q')$ whenever $Q' \subseteq Q$ and Q' contains x. Thus $\alpha(x)$ has less than L letters.

Let W denote the set of words with fewer than L letters, so that W has less than $2(T+1)^L$ elements. Given $w \in W$, we let F_w be the set of $x \in Q_0 \backslash (R_1 \cup R_2)$ such that $\alpha(x) = w$. We take the F_j's to simply be the F_w's enumerated by integers rather than elements of W.

It remains to check (2.2). Let $w \in W$ and x_1, $x_2 \in F_w$ be given. We may as well assume that $x_1 \neq x_2$. Let S be the smallest cube such that $x_1 \in S$ and $x_2 \in 2S$. If $S \in \mathscr{G}$, then the inequality in (2.2) follows from Lemma 2.19. Otherwise we must have $S \in \mathscr{B}_2 \cup \mathscr{B}_3$. Let k be such that $S \in \Delta_{k+l}$, and choose $Q_i \in \Delta_k$ so that $x_i \in Q_i$, $i = 1, 2$. Because $|x_1 - x_2| \geq b \operatorname{diam} S > \operatorname{diam} Q_1$, we conclude that $Q_1 \neq Q_2$, and hence $Q_2 \in \mathscr{F}(Q_1)$. It follows that $\alpha(x_1) \neq \alpha(x_2)$, because of the properties that $Q \mapsto \alpha(Q)$ must satisfy ((a), (b), and (c) above in particular). This contradicts the assumption that x_1, $x_2 \in F_w$, and the proof of Theorem 2.11 is now complete.

3. Weakly Lipschitz and bilipschitz mappings

Let E be a d-dimensional regular set in $\mathbf{R}^n$, and let Δ be a family of cubes in E, as in §2. For the purposes of this section the case of $E = \mathbf{R}^d$, $\Delta = \{\text{dyadic cubes}\}$ is already nontrivial and should be kept in mind. Let f be a function on E that takes values in $\mathbf{R}^m$ for some m. It is easy to see that f is Lipschitz if and only if

$$\sup_{x,y \in 2Q} |f(x) - f(y)| \leq C_1 \operatorname{diam} Q \tag{3.1}$$

for some constant C_1 and all $Q \in \Delta$.

Definition 3.2. f is weakly Lipschitz if there exist constants C_1, C_2 so that the set $\mathscr{B}(f)$ of cubes $Q \in \Delta$ for which (3.1) fails satisfies a Carleson packing condition with constant C_2.

This last means that

$$\sum_{\substack{Q \in \mathscr{B}(f) \\ Q \subseteq R}} |Q| \leq C_2 |R| \tag{3.3}$$

for all $R \in \Delta$.

We define weak bilipschitzness similarly. Let $b > 0$ be as in § 2 (see (2.13)), but with the additional restriction that $b < \frac{1}{10}$. It is easy to check that f is bilipschitz in the ordinary sense iff there is a constant C_3 such that

$$\begin{aligned} &C_3^{-1} \operatorname{diam} Q \leq |f(x) - f(y)| \leq C_3 \operatorname{diam} Q \\ &\text{whenever } x, y \text{ lie in } 2Q \text{ and } |x - y| > b \operatorname{diam} Q \end{aligned} \tag{3.4}$$

for all $Q \in \Delta$.

Definition 3.5. f is weakly bilipschitz if there are constants C_3 and C_4 so that $\mathscr{B}'(f) = \{Q \in \Delta : \text{(3.4) fails}\}$ satisfies a Carleson packing condition with constant C_4.

Notice that (3.4) implies (3.1), with $C_1 = 10C_3$. The proof is simple. Fix u, $v \in Q$ with $|u - v| \geq \frac{9}{10} \operatorname{diam} Q$, so that $|f(u) - f(v)| \leq C_3 \operatorname{diam} Q$ by (3.4). If $x \in 2Q$, then either $|x - u| \geq b \operatorname{diam} Q$ or $|x - v| \geq b \operatorname{diam} Q$, and hence

$$\min(|f(x) - f(u)|, \ |f(x) - f(v)|) \leq C_3 \operatorname{diam} Q$$

by (3.4) again. From here it follows easily that (3.4) implies (3.1) with $C_1 = 10C_3$. In particular, weakly bilipschitz functions are weakly Lipschitz.

The reason for considering these notions of weak Lipschitzness and bilipschitzness is that they allow certain types of singularities which are not permitted by their classical counterparts, but they are not so weak to allow completely unrestrained behavior. We shall give in this section some examples of ways in which weakly Lipschitz or bilipschitz functions are well-behaved, largely in connection with Theorems 2.1 and 2.11, but first we describe some examples of weakly Lipschitz and bilipschitz functions. The verification of the various assertions that follow are left as exercises.

Examples. (a) Let H be a hyperplane in $\mathbf{R}^d$. If $f : \mathbf{R}^d \to \mathbf{R}$ is Lipschitz on the two components of $\mathbf{R}^d \backslash H$, then f is weakly Lipschitz on $\mathbf{R}^d$, no matter how it is defined on H, or how badly the boundary values of f on the two sides of H match up.

(b) There is nothing special about hyperplanes in (a), in that there is a version of (a) whenever H is a subset of $\mathbf{R}^d$ with the property that $\{Q \in \Delta : 2Q$ intersects $H\}$ satisfies a Carleson packing condition, as in (3.3). This happens in particular when H is a regular set with dimension less than d.

(c) Define $f : \mathbf{R} \to \mathbf{R}$ by $f(x) = |x|$. Then f is weakly bilipschitz. Thus a mapping can have a fold and still be weakly bilipschitz.

(d) Suppose that $f: \mathbf{R} \to \mathbf{R}^2$ agrees with the identity on $(-\infty, 0) \cup (1, \infty)$, while on $[0, 1]$ it agrees with some translation on $\mathbf{R}^2$. Then f is weakly bilipschitz, with constants that do not depend on the choice of the translation.

(e) Define $f: \mathbf{R} \to \mathbf{R}$ by $f(x) = x$ when $x \leq 0$, $f(x) = x - 2^j$ when $x \in [2^j, 2^{j+1}]$, $j \in \mathbf{Z}$. Then f is weakly bilipschitz. Thus the image of a weakly bilipschitz mapping can pile up on itself.

(f) Let $g: \mathbf{R} \to \mathbf{R}^2$, $g = (g_1, g_2)$, be given as follows. For g_1 we use the mapping defined in (e), while we let g_2 be anything which is constant on $(-\infty, 0]$ and on each $[2^j, 2^{j+1}]$, $j \in \mathbf{Z}$. Then g is also weakly bilipschitz. The point here is that we can choose g_2 in such a way that g is one-to-one but still has its image concentrating in certain regions in $\mathbf{R}^2$.

(g) Suppose that $f: \mathbf{R}^d \to \mathbf{R}^m$ is Lipschitz, and that there is a $C > 0$ so that

$$h_d(f(Q)) \geq C^{-1}|Q|$$

for all dyadic cubes Q. [Here h_d denotes Hausdorff content, whose definition is recalled in (2.12). When $m = d$ we can simply take Lebesgue measure.] Then f is weakly bilipschitz. Before explaining why this is true we give a generalization of this fact.

(h) Let E be a d-dimensional regular set in $\mathbf{R}^n$ that satisfies the WGL, and assume that $f: E \to \mathbf{R}^m$ be Lipschitz. Suppose that there is a $C > 0$ so that $h_d(f(Q)) \geq C^{-1}|Q|$ for all cubes Q in E. Then for each $K > 0$ there is an α so that if (2.10) is uniformly bounded over all cubes Q_0 in E, then E is weakly bilipschitz. The proof of this is obtained from the same sort of considerations as in the proof of Theorem 2.11, up to and including the proof of Lemma 2.19. Note that the required bounds on (2.10) are automatic for a certain class of sets E, including d-planes, as discussed in § 4.

(i) Suppose that $f: \mathbf{R}^d \to \mathbf{R}^m$ is a regular mapping, which means that it is Lipschitz and that there is a $C > 0$ so that $|f^{-1}(B)| \leq CR^d$ for all balls B contained in $\mathbf{R}^m$ that have radius R. Then f is weakly bilipschitz. Indeed, using the definition of Hausdorff content it is easy to check that

$$|f^{-1}(A)| \leq Ch_d(A)$$

for all $A \subseteq \mathbf{R}^m$, which implies that the criterion in (g) is satisfied.

(j) The obvious converse to (i) is false, i.e., a mapping can be Lipschitz and weakly bilipschitz without being regular. Here is an example, which is a modification of (e). Define $f: \mathbf{R} \to \mathbf{R}$ by $f(x) = x$ when $x \in (-\infty, 0]$, and $f(x) = 2^{j-1} - |x - (2^j + 2^{j-1})|$ when $x \in [2^j, 2^{j+1}]$, $j \in \mathbf{Z}$. Thus the restriction of f to each interval $[2^j, 2^{j+1}]$ is given by a folding map which equals zero at each endpoint. It is not hard to check that f is Lipschitz and weakly bilipschitz but not regular.

(k) We would like to be able to use the notions of weak Lipschitzness and bilipschitzness to produce natural generalizations of regular mappings. That is, we would like to weaken the requirements placed on a regular mapping without sacrificing too much in terms of good behavior of the image. We have not succeeded in doing this in a pleasing fashion, and in fact the following recipe can be used to produce examples that indicate that there probably is no nice way to do this.

Let $\{a_j\}_{j=0}^{\infty}$ be a sequence of real numbers such that $0 < a_j < \frac{1}{2}$ for all j. Suppose that $f: \mathbf{R} \to \mathbf{R}^2$ satisfies the following two conditions. First, $f(x) = x$

if $x < 1$ or if $x \in (2^j + a_j, 2^{j+1})$ for some $j \geq 0$. Second, the restriction of f to $[2^j, 2^j + a_j]$ is bilipschitz for each $j \geq 0$, with a constant that is uniformly bounded in j. Then it is not hard to check that f is weakly bilipschitz.

It is possible to choose f in such a way that it has the following additional properties: f is one-to-one; the closure of $f(\mathbf{R})$ is a one-dimensional regular set, and the complement of $f(\mathbf{R})$ in its closure has zero one-dimensional Hausdorff measure; $H^1(f(A)) \approx |A|$ for all $A \subseteq \mathbf{R}$; and the closure of $f(\mathbf{R})$ is a bad set in the sense of [DS]. The precise meaning of this last property is not easy to explain briefly, but basically it means that $f(\mathbf{R})$ is not rectifiable in any reasonable quantitative sense. In particular it means that it is not true that plenty of singular integrals as in (1.9) define bounded operators on L^2 of $f(\mathbf{R})$. On the other hand the image of a regular mapping is always a good set in the sense of [DS], by [D2, 3].

The main idea for producing examples of mappings f that have all these features is to make certain that $f(\bigcup_j [2^j, 2^j + a_j])$ has portions that are well-approximated by unrectifiable sets, like Cantor sets.

The various examples just discussed show that weakly Lipschitz and bilipschitz mappings can behave rather badly, and so the reader may wonder why it is possible to say anything good about them. Roughly speaking the point is that weakly bilipschitz mappings behave badly from the perspective of their images, but they behave rather well from the perspective of their domains. The rest of this section will be devoted to their good behavior.

Let $f : E \to \mathbf{R}^m$ be a weakly Lipschitz mapping with constants C_1 and C_2. Let $Z = Z(f)$ denote the set of $x \in E$ that lie in $2Q$ for some arbitrarily small cubes $Q \in \mathscr{B}(f)$, so that Z has measure zero. In general we cannot say much about f on Z. This is illustrated by Examples (a) and (b) above. We can, however, say something about f on the complement of Z. The next lemma is a simple result of this type. It will be convenient for us to formulate it in somewhat greater generality.

Lemma 3.6. *Let E be a d-dimensional regular set in $\mathbf{R}^n$, and let $C_1 > 0$ be given. Given $f : E \to \mathbf{R}^m$ let $\mathscr{B}(f)$ be the set of cubes for which* (3.1) *fails. Set $Z_1 = Z_1(f) = \{x \in E : Q \in \mathscr{B}(f)$ for all sufficiently small cubes that contain $x\}$. Then $H^d(f(A \backslash Z_1)) \leq C|A|$ for all measurable sets $A \subseteq E$, where C depends on C_1, d, m, n, and the regularity constants for E.*

Note that $Z_1 \subseteq Z$, and that Z has measure zero if f is weakly Lipschitz with constants C_1, C_2 for any $C_2 < \infty$. Of course Lemma 3.6 is not so interesting when $|Z_1| > 0$.

The proof of the lemma is fairly straightforward. Let $A \subseteq E$ be fixed. It suffices to show that there is a $C > 0$ so that for every ε, $\delta > 0$ there is a sequence of balls $\{B_i\}$ in $\mathbf{R}^m$ with radii $\{r_i\}$ such that $r_i \leq \delta$ for all i, $f(A \backslash Z) \subseteq \bigcup B_i$, and $\sum r_i^d \leq C(|A| + \varepsilon)$.

Let ε, $\delta > 0$ be given, and let U be an open subset of E such that $A \subseteq U$ and $|U| \leq |A| + \varepsilon$. Set $\mathscr{A} = \{Q \in \Delta : Q \subseteq U$ and $\operatorname{diam} Q \leq \delta\}$, and let $\{Q_i\}_{i=1}^{\infty}$ be an enumeration of the maximal cubes in $\mathscr{A} \backslash \mathscr{B}(f)$. Then

$$A \backslash Z_1 \subseteq \bigcup Q_i,$$

because every $x \in E \backslash Z_1$ is contained in a cube in $\Delta \backslash \mathscr{B}(f)$ which is as small as you wish.

For each i let B_i denote the smallest ball in $\mathbf{R}^m$ that contains $f(Q_i)$, and let r_i denote its radius. Then $f(A\backslash Z_1) \subseteq \bigcup B_i$,

$$r_i \leq C \operatorname{diam} Q_i \leq C\delta,$$

because each Q lies in $\Delta\backslash\mathscr{B}(f)$, and

$$\sum r_i^d \leq C \sum |Q_i| \leq C\left|\bigcup Q_i\right| \leq C|U| \leq C(|A| + \varepsilon),$$

since the Q_i's are pairwise disjoint. This proves Lemma 3.6.

Let us consider now Theorem 2.11 from the point of view of weakly Lipschitz functions. It is not hard to make appropriate modifications to the statement and proof of Theorem 2.11 to allow the Lipschitz condition to be replaced by weak Lipschitzness. However, we can even dispense with the Lipschitz condition altogether, because the required bound on (2.10) already gives an adequate substitute for weak Lipschitzness. The precise statement is as follows.

Proposition 3.7. *Suppose that E is a d-dimensional regular set that satisfies the WGL, and let m, N_0, K, and $\varepsilon > 0$ be given. Then there exist positive constants α, τ, and M such that the following is true, with α depending only on d, m, n, ε, and the regularity constant for E, with τ depending on these constants and also on K, and with M depending on all of these constants as well as N_0 and the constants from the WGL.*

Suppose that Q_0 is a cube in E and that $f\colon Q_0 \to \mathbf{R}^m$ is a mapping such that (2.10) *is less than N_0. Set $Y = \{x \in Q_0 : x$ lies in infinitely many cubes in $NA(f, \alpha, K)\}$. Then there exist closed subsets F_j of $Q_0\backslash Y$, $1 \leq j \leq M$, such that*

$$\tau|x-y| \leq |f(x)-f(y)| \leq \tau^{-1}|x-y| \quad \text{for all } x, y \in F_j \text{ and all } j, \text{ and} \tag{3.8}$$

$$h_d\left(f\left(Q_0 \setminus \left(Y \cup \left(\bigcup_j F_j\right)\right)\right)\right) \leq \varepsilon|Q_0|. \tag{3.9}$$

Note that $|Y| = 0$, since (2.10) is required to be finite.

We should perhaps mention that in the next section we shall encounter conditions on E under which there are bounds on (2.10) for all weakly Lipschitz functions. This permits a nicer formulation of Proposition 3.7 for such sets E.

The proof of Proposition 3.7 is almost the same as that of Theorem 2.11, and we shall only indicate some of the principal modifications of the proof that are needed.

Let f satisfy the conditions listed above. The assumption that (2.10) is less than N_0 implies a kind of weaker weak Lipschitz condition, because

$$\operatorname{diam} f(2Q \cap Q_0) \leq (K + \alpha) \operatorname{diam} Q \tag{3.10}$$

whenever $Q \in A(f, \alpha, K)$. In particular, a minor variation of Lemma 3.6 implies that $h_d(f(D\backslash Y)) \leq C|D|$ for all measurable sets $D \subseteq Q_0$.

With these observations most of the changes needed in the proof of Theorem 2.11 are primarily cosmetic. For example, in (2.14) $|f(Q)|$ should be

replaced by $h_d(f(Q))$, and similar substitutions are needed throughout. With these modifications the analogue of Lemma 2.19 is still true, and with the same proof. Very few changes are needed after that, but there are two that are somewhat significant. The first is that (2.22) should be replaced by $h_d(f(R_2\backslash Y)) \leq \frac{\varepsilon}{3}|Q_0|$. The second is that you have to verify the second inequality in (3.8), whereas before there was only the first one. For this you use the fact that (3.10) holds when $Q \in A(f, \alpha, K)$, and hence when $Q \in \mathscr{G}$, where $\mathscr{G}$ is as in the proof of Theorem 2.11.

The techniques of the proofs of Theorem 2.1 and Theorem 2.11 can be used to derive some amusing properties of weakly Lipschitz and bilipschitz functions, as in the next two propositions.

Proposition 3.11. *Let E be a d-dimensional regular set in $\mathbf{R}^n$, and suppose that $f : E \to \mathbf{R}^m$ is weakly bilipschitz. Then for every $\varepsilon > 0$ there exist M, $N > 0$ so that the following is true, with M and N depending only on n, d, ε, m, the regularity constant for E, and the weak bilipschitz constants for f.*

For every $Q_0 \in \Delta$ there exist closed subsets F_j of Q_0, $1 \leq j \leq M$, such that

$$|Q_0\backslash(\bigcup F_j)| \leq \varepsilon|Q_0| \tag{3.12}$$

and

$$\begin{aligned} &N^{-1}|x-y| \leq |f(x)-f(y)| \leq N|x-y| \\ &\text{for all } x, y \in F_j, \ j = 1, \dots, M. \end{aligned} \tag{3.13}$$

Proposition 3.14. *Proposition* 3.11 *remains true if we replace "weakly bilipschitz" by "weakly Lipschitz" and* (3.13) *by*

$$|f(x)-f(y)| \leq N|x-y| \quad \text{for all } x, y \in F_j, \ j = 1, \dots, M. \tag{3.15}$$

These two results are proved using arguments that are similar to but much simpler than the proof of Theorem 2.11. When proving Proposition 3.11, for instance, you do not need anything like Lemma 2.19, because its conclusion is essentially contained in the definition of weak bilipschitzness. You also do not need anything like $\mathscr{B}_1$ or $\mathscr{B}_2$, but you do need the analogue of $\mathscr{B}_3$ with $NA(f, \alpha, K)$ replaced by $\mathscr{B}(f)$. What remains of the proof of Theorem 2.11 (which is mostly the coding argument) should be pretty much left as it is. The proof of Proposition 3.14 is essentially the same; you simply do not keep track of the lower bounds on $|f(x)-f(y)|$.

In Propositions 3.11 and 3.14 we could just as well have allowed f to take values in any metric space, with only notational changes in the statements and proofs. This is not the case for Theorem 2.11 or Proposition 3.7, because of the role of affine functions.

Proposition 3.14 is rather nice because it says that weakly Lipschitz functions really do look like ordinary Lipschitz functions on substantial sets. It is also nice because there are examples where the phenomenon described in the conclusion does occur, e.g., Examples (a) and (b) above.

A natural question to ask is whether it is possible to improve the conclusions of Proposition 3.14, in such a way that there is some balance required between how badly the $f|_{F_j}$'s match up on the common boundary and how much common boundary there is. In Examples (a) and (b) the common boundary was small enough so that it did not matter how bad the match-ups were.

4. Estimates for affine approximations to Lipschitz functions

Let E be a d-dimensional regular set, as usual. In this section we give conditions on E under which bounds on (2.10) are automatic for Lipschitz functions on E. For this issue we may as well restrict ourselves to real-valued functions on E.

Let f be a real-valued function on E. Given $K > 0$ and $p \in [1, \infty]$ define $\alpha_{K,p}(x, t)$ on $E \times \mathbf{R}_+$ by

$$\alpha_{K,p}(x, t) = \inf_{\substack{a \\ |\nabla a| \le K}} t^{-1} \left(t^{-d} \int_{B(x,t)\cap E} |f - a|^p d\mu \right)^{\frac{1}{p}}, \tag{4.1}$$

with the usual modifications when $p = \infty$. Here the infimum is taken over all affine functions $a : \mathbf{R}^n \to \mathbf{R}$ with $|\nabla a| \le K$.

We want to find conditions on E under which the following is true:

(4.2) There is a $K > 0$ so that for every $\varepsilon > 0$ there is a $C(\varepsilon) > 0$ such that $\{(x, t) \in E \times \mathbf{R}_+ : \alpha_{K,\infty}(x, t) > \varepsilon\}$ is a Carleson set with constant $C(\varepsilon)$ whenever f is a Lipschitz function on E with norm ≤ 1.

It is easy to check that if (4.2) is true, then (2.10) is uniformly bounded for all $\alpha > 0$ and all Lipschitz mappings of E into $\mathbf{R}^m$ with norm ≤ 1. In this event the hypotheses of Theorem 2.11 simplify accordingly.

In practice (4.2) is not checked directly, but rather a condition like the following is verified:

(4.3) There exist $C, K > 0$ so that $\alpha_{K,1}(x, t)^2 d\mu(x) \frac{dt}{t}$ is a Carleson measure on $E \times \mathbf{R}_+$ with norm $\le C$ whenever f is a Lipschitz function on E with norm ≤ 1.

[We are restricting ourselves to $p = 1$ here for simplicity. For some (other) purposes it is desirable to consider larger p's.] In order to show that (4.3) implies (4.2) it suffices to check that

$$\alpha_{K,\infty}(x, t) \le C(K) \alpha_{K,1}(x, 2t)^{\frac{1}{d+1}} \tag{4.4}$$

for all Lipschitz functions with norm ≤ 1. This is not difficult. Fix x and t, and let a be an affine function on $\mathbf{R}^n$ with $|\nabla a| \le K$ such that the infimum in (4.1) is attained with $p = 1$ and t replaced by $2t$. For each $y \in B(x, t)$ and $\lambda \in (0, 1)$ we have

$$\begin{aligned} |f(y) - a(y)| &\le C(\lambda t)^{-d} \int_{B(y,\lambda t)\cap E} \{|f(y) - a(y) - f(x) + a(x)| \\ &\qquad\qquad + |f(x) - a(x)|\} d\mu(x) \\ &\le C(K+1)\lambda t + C\lambda^{-d} \alpha_{K,1}(x, 2t) t. \end{aligned}$$

We have used here the fact that $f - a$ is Lipschitz with norm $\le K + 1$. Hence

$$\alpha_{K,\infty}(x, t) \le C(K+1)\lambda + C\lambda^{-d} \alpha_{K,1}(x, 2t).$$

From here (4.4) follows easily, by choosing $\lambda = \alpha_{K,1}(x, 2t)^{\frac{1}{d+1}}$.

The next three propositions provide increasingly general conditions on E that ensure that (4.3), and hence (4.2), holds.

Proposition 4.5. *If E is a d-plane, then* (4.3) *is true.*

This is basically well-known. Here are some highlights of a proof. The first step is to reduce to the case where $K = \infty$, that is, to show that if (4.3) holds with $K = \infty$, then it holds for some finite K. The main point is that if an affine function a approximates f well on $E \cap B(x, t)$, and if f has Lipschitz norm ≤ 1, then $a|_E$ cannot have large Lipschitz norm. It is important here that E is a d-plane, so that $a|_E$ can be considered to be affine in its own right. The estimate when $K = \infty$ can be derived easily from an analogous result for functions f that satisfy $\int_E |\nabla f|^2 \leq 1$ instead of the Lipschitz condition. This result is contained in Theorem 2 of [Do].

Proposition 4.6. *If E is a Lipschitz graph, then* (4.3) *holds.*

This is a simple consequence of the preceding result. Suppose that E is given by (1.1), and let f be a function on E which is Lipschitz with norm ≤ 1. Let A, P be as in (1.1), and let us identify $\mathbf{R}^n$ with $P \times Q$ in the obvious way, where Q is the $(n-d)$-plane orthogonal to P that passes through the origin. Define F on P by

$$F(p) = f(p, A(p)),$$

so that F is Lipschitz on P. The key observation is that if $a(p)$ is an affine function on P that approximates F well (on some ball in P, say), then $\tilde{a}(p, q) = a(p)$ is an affine function on $\mathbf{R}^n$ that approximates f well on the corresponding part of E. This allows you to reduce to Proposition 4.5.

Proposition 4.7. *Suppose that E satisfies one of the equivalent conditions* (C1)–(C8) *in* [DS]. (*This happens in particular when E has BPLG.*) *Then* (4.3) *holds.*

We shall not explain here what the conditions (C1)–(C8) are. The reader should consult [DS] for this and also to decipher the following sketch of the proof of Proposition 4.7. We hope that the gentle reader will forgive us for not being more forthcoming with the details, but we suspect that full disclosure would not lead to increased happiness for most readers. (*Added in proof*: The authors' feelings of guilt have culminated in the inclusion of a detailed proof of this proposition in §4.1 of Part III of the monograph [DS2].)

The argument for proving Proposition 4.7 is very similar to the one used in [DS] to show that (C4)—the existence of a corona decomposition for E—implies (C3), the higher-dimensional version of Peter Jones' geometric lemma. The details of the statement that E admits a corona decomposition are complicated, but basically it means that E can be well-approximated by a family of Lipschitz graphs. More precisely it says that you can decompose Δ into $\mathscr{B} \cup \mathscr{G}$, the bad cubes and the good cubes, with $\mathscr{B}$ satisfying a Carleson packing condition, and with $\mathscr{G}$ further subdivided into a family of stopping-time regions. There are not too many of these regions, in that they satisfy a Carleson packing condition, and from the perspective of each of these regions E is well-approximated by a Lipschitz graph.

To prove (4.3) you take a Lipschitz function f on E, and, for each one of these good stopping-time regions S, you approximate f by a Lipschitz function f_S on the corresponding Lipschitz graph Γ_S. It is somewhat technical but basically straightforward to derive the estimate in (4.3) for f from the corresponding estimates for each f_S (which come from Proposition 4.6) and

the Carleson packing conditions on the bad cubes and the good stopping-time regions. The computations are substantially simpler if you are willing to settle for (4.2) instead of (4.3).

A special case of Proposition 4.7 occurs when we assume that E satisfies Condition B, since we know that E must then have BPLG. If we also require that the complement of E have exactly two components, then there is a fairly direct proof of the existence of a corona decomposition for E in [S4]. (Otherwise you have to go through the fact that E has BPLG, and the proof that BPLG implies the existence of a corona decomposition is quite roundabout and complicated. *Added in proof*: A more direct proof of this last fact is given in Chapter 2 of Part IV of [DS2].) If we further make some a priori smoothness assumptions, then the estimates of Proposition 4.7 and related results can be proved rather easily using square function estimates for the Cauchy integral instead of the existence of a corona decomposition. See [S3].

If E is a chord-arc surface and we make suitable a priori smoothness assumptions, then the methods of [S3] actually give that the Lipschitz norm of a function on E is controlled by the L^∞ norm of its distributional gradient on E, just as for $\mathbf{R}^d$. See Proposition 6.12.

Finally we take up the issue of knowing when the analogue of (4.2) for weakly Lipschitz functions is true.

Proposition 4.8. *Suppose that* (4.2) *is true for* E. *Then for every* C_1, $C_2 < \infty$ *there is a* $K < \infty$ *so that for each* $\varepsilon > 0$ *there exists* $C = C(\varepsilon) < \infty$ *such that* $\{(x, t) : \alpha_{K,\infty}(x, t) \geq \varepsilon\}$ *is a Carleson set in* $E \times \mathbf{R}_+$ *with constant* $\leq C$ *whenever* $f : E \to \mathbf{R}$ *is weakly Lipschitz with constants* C_1, C_2.

The proof of Proposition 4.8 has very much the same flavor as some of the arguments used in [DS]. Once again we beg the reader's forgiveness for not presenting a complete proof, but merely the outline below, and for not having one that is over quickly. (The reader who is interested in completing the proof might find it helpful to read §3.2 in Part I of [DS2]. Also, some of the techniques used in §4.1 of Part III of [DS2] are relevant for this argument as well.)

Let a weakly Lipschitz function $f: E \to \mathbf{R}$ be given, let $\mathscr{B}(f)$ be as in Definition 3.2, and set $\mathscr{G}(f) = \Delta \backslash \mathscr{B}(f)$. It is not so hard to show that Δ can be decomposed into $\mathscr{G}_1(f) \cup \mathscr{B}_1(f)$, where $\mathscr{G}_1(f) \subseteq \mathscr{G}(f)$, $\mathscr{B}_1(f)$ still satisfies a Carleson packing condition like (3.3), and where $\mathscr{G}_1(f)$ can itself be decomposed into a union of families S_j of cubes which satisfy the following properties:

(4.9) each S_j has a maximal cube Q_j;

(4.10) if $Q, Q' \in \Delta$, $Q \subseteq Q' \subseteq Q_j$, and $Q \in S_j$, then $Q' \in S_j$;

(4.11) the Q_j's also satisfy a Carleson packing condition like (3.3).

The proof of the existence of these decompositions is very similar to the proof of Lemma 7.1 in [DS], and we omit the details. Note that this would just be a standard and simple stopping time argument if we were only trying to decompose $\Delta(Q_0)$ for some fixed Q_0 rather than all of Δ. Some minor adjustments are needed to decompose all of Δ, and they are given in the proof of Lemma 7.1 in [DS].

The next step is to show that for each j there is a Lipschitz function $f_j : E \to \mathbf{R}$ that approximates f well from the point of view of S_j. More precisely,

the f_j's have Lipschitz norm bounded by a constant that depends only on the weak Lipschitz constants of f, and for each $Q \in S_j$ we have

$$|f - f_j| \leq C \operatorname{diam} Q \text{ on } \tfrac{3}{2} Q.$$

The proof of this is neither difficult nor pleasantly brief, and so we omit it. [The reader who attempts it should be warned to take into account the fact that there could be minimal cubes in S_j that are close together but which have very different sizes.]

Once you have the f_j's it is not hard to control $\alpha_{K,\infty}(x, t)$ for f in terms of the corresponding quantity for f_j when, say, $x \in Q$ and $\frac{1}{2} \operatorname{diam} Q \leq t \leq \operatorname{diam} Q$ for some $Q \in S_j$. We can use (4.2) to control the $\alpha_{K,\infty}(x, t)$'s for the f_j's, and it is not hard to show that this gives the right sort of estimate for f.

5. The proof of Theorem 1.20

Throughout this section we take $n = d + 1$.

To prove Theorem 1.20 we are going to define a family of subsets G_ε of $E \times \mathbf{R}_+$ and show that they are always Carleson sets for any regular set E. We shall then show that this implies that the sets H_ε and A_ε as in §1 must also be Carleson sets if E satisfies Condition B.

For each $\varepsilon > 0$ let G_ε denote the set of $(x, t) \in E \times \mathbf{R}_+$ for which there exist three points $y_0, y_1, y_2 \in B(x, t)$ such that $\operatorname{dist}(y_i, E) \geq \varepsilon t$ for $i = 0, 1, 2$, y_0 lies on the line segment that joins y_1 to y_2, but y_0 does not lie in the same component of $\mathbf{R}^{d+1} \backslash E$ as either y_1 or y_2. (See Figure 2 for examples of how this can occur.) Thus G_ε notices bubbles in particular.

Proposition 5.0. *If E is regular, then G_ε is a Carleson set for all $\varepsilon > 0$.*

To prove this we first define a variant G'_ε of G_ε and show that it is always a Carleson set.

Fix a unit vector v in $\mathbf{R}^{d+1}$. Let V denote the line through the origin that passes through v. We shall think of $\mathbf{R}^{d+1}$ as being $V \oplus V^\perp$. We are going to refer to lines parallel to V as being "vertical", and we use v to orient V and to give meaning to the words "up", "down", "above", "below", etc. Thus x is considered to be above y if the projection of x on V is above the projection of y on V, in the obvious sense.

Given $\varepsilon > 0$, we define $G'_\varepsilon = G'_\varepsilon(v)$ to be the set of $(x, t) \in E \times \mathbf{R}^{d+1}$ for which there exist u_0, u_1, u_2 in $\mathbf{R}^{d+1}$ with the following properties: u_0, u_1, and u_2 all lie on a single vertical line, with u_0 above u_1 and below u_2;

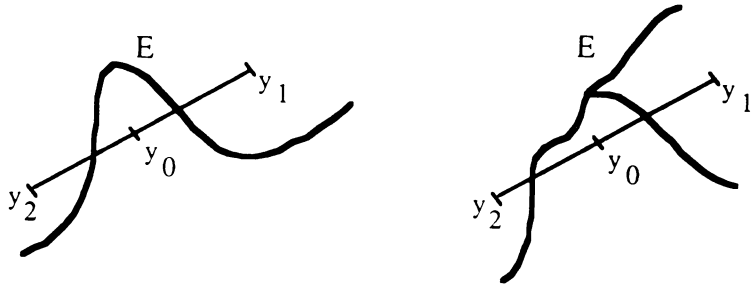

Figure 2

$\operatorname{dist}(u_i, E) \geq \varepsilon t$ for $i = 0, 1, 2$; and u_0 lies in a connected component of $\mathbf{R}^{d+1} \backslash E$ that is different from each of the complementary components that contain u_1 and u_2. The only significant difference between this and G_ε is that we specify in advance the direction of the line that contains the three points.

Lemma 5.1. *If E is regular, then G'_ε is a Carleson set for all $\varepsilon > 0$.*

Fix $\varepsilon > 0$. We may as well require that $\varepsilon < \frac{1}{10}$. We must show that there is a $C = C(\varepsilon)$ so that for each $x \in E$ and $R > 0$ we have that

$$\iint_A d\mu(y) \frac{dt}{t} \leq CR^d,$$

where $A = G'_\varepsilon \cap \{(E \cap B(x, R)) \times (0, R]\}$. We may as well assume that $x = 0$ and $R = 1$, by standard rescaling arguments. Thus it suffices to prove that

$$\sum_{m=0}^{\infty} |A_m| \leq C, \tag{5.2}$$

where $A_m = \{y \in E \cap B(0, 1) :$ there is a $t \in [2^{-m-1}, 2^{-m}]$ such that $(y, t) \in G'_\varepsilon\}$.

For each $m \geq 0$ cover A_m by the balls $B(y, 2^{-m+1})$, $y \in A_m$. We can find a finite subcovering by balls $D_i = D_i^m = B(x_i^m, 2^{-m+1})$, where x_i^m, $i \in I_m$, is a finite sequence of points in A_m, in such a way that

$$\sum_{i \in I_m} \chi_{D_i^m} \leq C$$

for each m. It is unnecessary to disturb Besicovitch's covering lemma to find the D_i^m's, because all the balls in the original covering have the same radius (for each m).

In order to prove Lemma 5.1 we are going to associate to each D_i^m a measurable set $S(D_i^m) \subseteq E$ with bounds on the overlaps of the $S(D_i^m)$'s, in such a way that (5.2) will reduce to an easy estimate for $\left| \bigcup_{i,m} S(D_i^m) \right|$.

Let D be one of the D_i^m's for some $m \geq 0$ and $i \in I_m$. Because the center $x = x_i^m$ of D lies in A_m, there exist three points u_0, u_1, and u_2 in $B(x, 2^{-m}) = \frac{1}{2}D$ with the following three properties:

(5.3) u_0, u_1, and u_2 all lie on the same vertical line L, with u_0 above u_1 and below u_2;

(5.4) the balls $B_i = B(u_i, \varepsilon 2^{-m-1})$, $i = 0, 1, 2$, do not intersect E;

(5.5) if U_i denotes the component of $\mathbf{R}^{d+1} \backslash E$ that contains B_i, $i = 0, 1, 2$, then $U_1 \neq U_0$ and $U_2 \neq U_0$.

Let T be the piece of vertical tube composed of all points z in $\mathbf{R}^{d+1}$ which are at distance at most $\varepsilon 2^{-m-2}$ from L and whose vertical coordinates lie between the vertical components of u_0 and u_2, as in Figure 3. Notice that $T \subseteq D$.

We define $S(D) \subseteq E \cap D$ to be the set of points $z \in E \cap T$ such that the open vertical segment below z that joins z to $\frac{1}{2}B_0$ does not intersect E. In other words, z is the first point in E on some vertical ray that emanates up from $\frac{1}{2}B_0$.

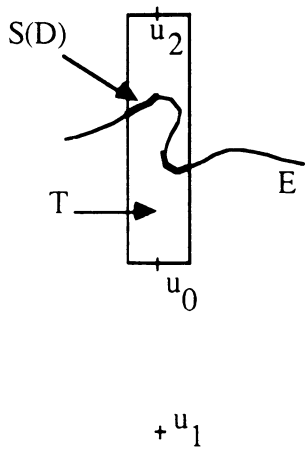

FIGURE 3

Notice that each vertical line segment that joins $\frac{1}{2}B_0$ to $\frac{1}{2}B_2$ must hit E at least once, by (5.5), and so it must hit $S(D)$ too. Therefore

$$|S(D)| \geq C^{-1}\varepsilon^d 2^{-md}. \tag{5.6}$$

[In order to quell the tremors of author number two let us mumble a few words of measurability. We can view $S(D)$ as the graph of a real-valued function defined on the intersection of $\frac{1}{2}B_0$ with the hyperplane through u_0 that is orthogonal to the vertical direction. It is easy to check that this function is lower semicontinuous, using the fact that E is closed, which implies that $S(D)$ is a Borel set.]

The next lemma will permit us to control the overlap of the $S(D_i^m)$'s.

Lemma 5.7. *There is an integer $M > 0$ that depends only on E and ε such that $S(D_i^m) \cap S(D_{i'}^{m'}) = \varnothing$ whenever $m' \leq m - M$, $i \in I_m$, and $i' \in I_{m'}$.*

Suppose not. Let $D = D_i^m$, $D' = D_{i'}^{m'}$ be as in the lemma, and suppose that there is a $z \in S(D) \cap S(D')$ despite the fact that $m' \leq m - M$ for a large M. We shall show that this leads to a contradiction if M is large enough.

Let B_0, B_1, and B_2 (respectively B_0', B_1', and B_2') denote the balls associated to D (respectively D') as above. If M is large enough, then the radii of the balls B_j', $j = 0, 1, 2$, are much larger than 2^{-m}. This forces the center of B_2' to be above B_2. [This is because the vertical ray emanating up from z intersects $\frac{1}{2}B_2'$ and $\frac{1}{2}B_2$, while the downward-pointing vertical ray meets $\frac{1}{2}B_0$. If the center of B_2' were not above B_2, then the center of B_2' would get too close to z and B_2' would contain B_2 and B_0, violating (5.5). See Figure 4.] Similarly the center of B_0' must lie below B_1. [Again, if this were not true, then B_0' would contain B_1 and B_0, contradicting (5.5).]

Let W denote the vertical line segment below z that joins z to $\frac{1}{2}B_0'$. This segment must pass through B_0 and B_1, and so it must meet E somewhere in between, by (5.5). Thus z is not the first point on $W \cap E$ on the way up from $\frac{1}{2}B_0'$, which contradicts the assumption that $z \in S(D')$. This proves Lemma 5.7.

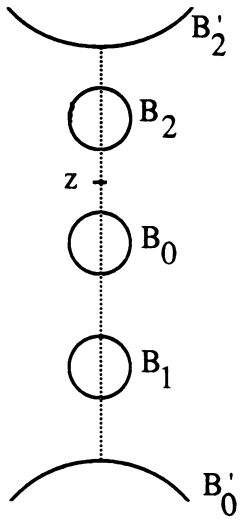

FIGURE 4

Using Lemma 5.7 we conclude that no point in E can lie in more than C of the $S(D_i^m)$'s. Indeed, no point of E can lie in $S(D_i^m)$ for more than M values of m, by Lemma 5.7, and for each fixed m no element of E can lie in $S(D_i^m)$ for more than a bounded number of i's. This last follows from $S(D_i^m) \subseteq D_i^m$ and the fact that we chose the D_i^m's so that they have bounded overlap for fixed m.

Let us verify (5.2). We have

$$\sum_{m=0}^{\infty} |A_m| \leq \sum_{m=0}^{\infty} \sum_{i \in I_m} |E \cap D_i^m| \leq C \sum_{m=0}^{\infty} \sum_{i \in I_m} |S(D_i^m)|$$

by (5.6), and our control on the overlaps of the $S(D_i^m)$'s now gives

$$\sum_{m=0}^{\infty} |A_m| \leq C \left| \bigcup_m \bigcup_{i \in I_m} S(D_i^m) \right| \leq C|E \cap B(0, 2)| \leq C.$$

This finishes the proof of Lemma 5.1.

Let us now use Lemma 5.1 to prove Proposition 5.0. We have to check that G_ε must be a Carleson set for each $\varepsilon > 0$. The only difference between G_ε and G'_ε is that G'_ε is defined in terms of a fixed, prescribed direction, while the definition of G_ε does not require the directions of the lines to be specified in advance. Given $\varepsilon > 0$ we have that

$$G_\varepsilon \subseteq \bigcup_j G'_{\varepsilon/2}(v_j) \tag{5.8}$$

for any sufficiently thick set of unit vectors v_j in $\mathbf{R}^{d+1}$. This is not hard to verify; the point is that the definition of G_ε leaves us some room to move around. Because we can find a finite number (depending on ε and d) of v_j's that are thick enough in the sphere for (5.8) to hold, the fact that G_ε is a Carleson set follows from the corresponding result for G'_ε.

Next we use Proposition 5.0 to prove the first part of Theorem 1.20. Assume that E satisfies Condition B.

We would like to show that for each $\varepsilon > 0$ there is an $\eta > 0$ so that $H_\varepsilon \subseteq G_\eta$, which implies that H_ε is a Carleson set. Fix $\varepsilon > 0$, and suppose that $(x, t) \in H_\varepsilon$. Thus there exist y_1, $y_2 \in B(x, t)$ such that $\operatorname{dist}(y_i, E) \geq \varepsilon t$ for $i = 1$, 2, y_1 and y_2 lie in the same component of $\mathbf{R}^{d+1} \backslash E$, but the line segment that joins y_1 to y_2 intersects E at a point y, say. Because E satisfies Condition B we can find a point $u_0 \in B(y, \varepsilon t/10)$ such that $\operatorname{dist}(u_0, E) \geq \varepsilon t/10C$ and u_0 belongs to a different component of $\mathbf{R}^{d+1} \backslash E$ from y_1 and y_2. It is easy to see that there must be points u_1, u_2 such that $u_j \in B(y_j, \varepsilon t/2)$, $j = 1$, 2, and such that u_0 lies on the line segment that joins u_1 to u_2. This implies that $(x, 2t) \in G_\eta$ with $\eta = \varepsilon/20C$. This is not quite what we wanted, but it is good enough.

The proof that E satisfies LS is quite similar. Let $\varepsilon > 0$ be given, and let A_ε be as in (1.19). We are going to show that

$$A_\varepsilon \subseteq \{(x, t) : (x, 10t) \in G_\eta\} \tag{5.9}$$

for some sufficiently small η, from which it will follow immediately that A_ε is a Carleson set.

Fix $(x, t) \in A_\varepsilon$, and let u, v be as in (1.19). Set $w = 2u - v$. Because E satisfies Condition B we can find a point y_0 in $B(u, \varepsilon t/10)$ which does not lie in the same component of $\mathbf{R}^{d+1} \backslash E$ as does w, and whose distance to E is at least $C^{-1}\varepsilon t$. Similarly we can find a point y_2 in $B(v, \varepsilon t/10)$ whose distance to E is at least $C^{-1}\varepsilon t$ and which does not lie in the same complementary component of E as y_0. Finally we choose $y_1 \in B(w, \varepsilon t/2)$ so that y_0 lies on the segment that joins y_1 to y_2. Notice that $\operatorname{dist}(y, E) \geq \varepsilon t/2$ and that y_0 and y_1 lie in different components of $\mathbf{R}^{d+1} \backslash E$. It follows from all this that $(x, 10t)$ lies in G_η if η is small enough. This proves (5.9).

It remains to establish the last statement in Theorem 1.20. We shall see from the proof that we actually need much less than the requirement that H'_ε be a Carleson set for all $\varepsilon > 0$ to conclude that E satisfies Condition B.

Let $\varepsilon > 0$ be small, to be chosen soon. Let $x \in E$ and $r > 0$ be given. Because H'_ε is a Carleson set, there is a constant $a > 0$ (small) so that there is a $z \in E \cap B(x, r/2)$ and a $t \in [ar, r/2]$ such that $(z, t) \notin H'_\varepsilon$. Here a depends on the Carleson constant for H'_ε (and hence on ε), but not on x or r.

Since $(z, t) \notin H'_\varepsilon$, there does not exist a pair of points y_1, y_2 in $\mathbf{R}^{d+1} \backslash E$ with the following properties:

$$\begin{aligned} &y_1, y_2 \in B(z, t), \quad \operatorname{dist}(y_i, E) \geq \varepsilon t \text{ for } i = 1, 2, \\ &\text{and } \tfrac{1}{2}(y_1 + y_2) \in E; \end{aligned} \tag{5.10}$$

$$y_1 \text{ and } y_2 \text{ lie in the same component of } \mathbf{R}^{d+1} \backslash E. \tag{5.11}$$

Thus if we can find y_1, y_2 for which (5.10) holds, then they must lie in different complementary components of E. This will imply that E satisfies Condition B, since x and r are arbitrary.

Let us check that there exist y_1, y_2 such that (5.10) holds if $\varepsilon > 0$ is small enough, using the fact that E is regular. Let u_i, $i \in I$, be a maximal set of

points in $B(z, t)$ such that

$$|u_i - u_j| > 4\varepsilon t \quad \text{and} \quad |u_i - v_j| > 4\varepsilon t$$

when $i \neq j$, where $v_j = 2z - u_j$ is the point symmetric to u_j about z. Clearly I has at least $C^{-1}\varepsilon^{-d-1}$ elements. Let J denote the set of $i \in I$ such that $B(u_i, \varepsilon t)$ or $B(v_i, \varepsilon t)$ intersects E. Then J has at most $C\varepsilon^{-d}$ elements, because the doubles of all these balls are disjoint, because

$$|E \cap B(u_i, 2\varepsilon t)| + |E \cap B(v_i, 2\varepsilon t)| \geq C^{-1}\varepsilon^d t^d$$

when $i \in J$, and because $|E \cap B(z, 2t)| \leq Ct^d$. Therefore $I \backslash J$ is nonempty if ε is small enough, which implies the existence of y_1 and y_2 that satisfy (5.10).

This completes the proof of Theorem 1.20.

Remark 5.12. There are other conditions besides the local symmetry condition LS that would serve just as well here, i.e., as a convenient criterion for the WGL. One of these is the local convexity condition LC. We say that E satisfies LC if

$$A'_\varepsilon = \Big\{(x, t) \in E \times \mathbf{R}_+ : \text{ there are two points } u, v \in E \cap B(x, t) \text{ such that } \operatorname{dist}\left(\frac{u+v}{2}, E\right) \geq \varepsilon t\Big\}$$

is a Carleson set for all $\varepsilon > 0$. There is a version of Proposition 1.18 for this condition, and it is also true that Condition B implies LC. The proofs are similar to those of the analogous results for LS.

6. The proof of Theorem 1.22

Let E be as in the hypotheses of Theorem 1.22. E inherits a Riemannian metric from $\mathbf{R}^3$, and in particular a conformal structure. The a priori assumptions on E and the simple connectivity of E imply that $E \cup \{\infty\}$ is diffeomorphic to S^2, and that the conformal structure on E extends to a smooth conformal structure on $E \cup \{\infty\}$. The uniformization theorem implies that $E \cup \{\infty\}$ equipped with this conformal structure is conformally equivalent to S^2 with the standard conformal structure. Thus there is a conformal diffeomorphism τ of $\mathbf{R}^2$ onto E, and it suffices to show that τ is quasisymmetric with appropriate bounds.

A criterion for the existence of these bounds is given in Theorem 5.4 of [S2]. (This result was proved using estimates for extremal length.) Using this criterion and the fact that E is regular we are reduced to showing that E satisfies the following two conditions:

$$|x - y| \leq d_E(x, y) \leq A|x - y| \quad \text{for all } x, y \in E, \tag{6.1}$$

where d_E denotes the geodesic distance on E

(i.e., the length of the shortest path that joins a given pair of points);

(6.2) there is a $K > 0$ so that for each $x \in E$ and $R > 0$ there is an open set $D \subseteq E$ that is homeomorphic to a disk and which satisfies $E \cap B(x, R) \subseteq D \subseteq E \cap B(x, KR)$.

As usual the constants A and K should depend only on the constants involved in the definition of a chord-arc surface, and in particular they should not depend quantitatively on our a priori smoothness assumptions.

The proof that (6.1) holds will be given later in this section. For this the assumptions that $d = 2$ and that E is simply-connected are irrelevant.

Let us now discuss the proof of (6.2). Fix $x \in E$ and $R > 0$. Set

$$G = \{y \in E : d_E(y, x) < R\}.$$

This is an open subset of E that contains $E \cap B(x, A^{-1}R)$ and is contained in $B(x, R)$, but it is not likely to be simply connected. To fix this we fill in the holes.

Let D be the subset of E obtained by taking G and filling in the holes, i.e., $E \backslash D$ is the unbounded component of $E \backslash G$. Then D contains G, D is open and connected, and it is also simply connected, since E is homeomorphic to $\mathbf{R}^2$ and since $E \backslash D$ is connected. Thus D is a topological disk, and it remains to check that

$$D \subseteq B(x, KR) \tag{6.3}$$

for some $K > 0$.

The idea for proving this is that if it were false, then some piece of E would have to be bubbling off (or protruding excessively from) G, and that that would violate the requirement that E be a chord-arc surface. It is not clear how to carry out this approach directly at the geometrical level, and we shall instead use an analytical argument based on a Poincaré inequality that can be proved using the methods of [S3].

We first define a function $f: E \to \mathbf{R}$ which measures the extent of the bubbling off and to which we shall apply the Poincaré inequality. Set $f = 0$ on $E \backslash D$, and define f on D by

$$f(y) = \min(2R, d_E(x, y)) - R.$$

We have defined f in such a way that the derivative of f is supported on a set whose diameter is $\leq CR$, but so that f itself is $\geq R$ on a big set if (6.3) is not true.

Lemma 6.4. *f is Lipschitz, and in fact*

$$|f(y) - f(z)| \leq d_E(y, z) \tag{6.5}$$

for all y, $z \in E$.

It clearly suffices to check (6.5) when $y \in D$ and $z \in E \backslash D$. The main point behind the proof is that $d_E(x, u) = R$ when $u \in \partial D$, because of the way G is defined.

Let γ be a curve on E that joins y to z and has minimal length. It is not hard to see that γ must intersect G, since y either lives in G or in a complementary component of G other than $E \backslash D$. Let w be any point in $G \cap \gamma$, and observe that

$$d_E(y, w) + d_E(w, z) = d_E(y, z).$$

Clearly $|f(y) - f(w)| \leq d_E(y, w)$. Also,

$$|f(z) - f(w)| = |f(w)| = R - d_E(x, w) \leq d_E(w, z),$$

since $R \leq d_E(x, z) \leq d_E(x, w) + d_E(w, z)$. Hence

$$\begin{aligned} |f(y) - f(z)| &\leq |f(y) - f(w)| + |f(w) - f(z)| \\ &\leq d_E(y, w) + d_E(w, z) = d_E(y, z), \end{aligned}$$

as asserted.

Lemma 6.6. $\int_E |\nabla_T f|^p \le CR^2$, $1 \le p \le 2$.

Here $\nabla_T f$ denotes the (tangential) gradient of f as a function on E. The "T" is there to emphasize that this is not the usual gradient for functions on $\mathbf{R}^{d+1}$.

The previous lemma implies that $|\nabla_T f| \le 1$, a.e., and so it suffices to show that $\nabla_T f = 0$ a.e. outside $B(x, CR)$ if C is large enough. This is true because f is locally constant on $\{y \in E : d_E(y, x) > 2R\}$, as can be readily verified from the definitions of f, G, and D.

The version of the Poincaré inequality that we shall use is the following. Given $p \in [1, 2]$, $t > 0$, and a Lipschitz function $g : E \to \mathbf{R}$ with compact support, we have that

$$\begin{aligned}\int_{E\cap B(x,t)} t^{-d} \int_{E\cap B(x,t)} (t^{-1}|g(y) - g(z)|)^p d\mu(y) d\mu(z) \\ \le C \int_E |\nabla_T g|^p d\mu,\end{aligned} \tag{6.7}$$

where, as always, the constant C is not allowed to depend on our a priori smoothness assumptions in a quantitative way. This can be proved using the techniques of [S3], and we will sketch the proof later in the section. The $p = 2$ case of (6.7) can be derived directly from the main theorem in [S3], but for the present purposes we can use any $p \in [1, 2]$ except $p = 2$. Note that we do not need E to be simply connected for (6.7), nor do we need $d = 2$. We have written t^{-d} in (6.7) instead of t^{-2} to emphasize this point.

Let us now derive (6.3) from (6.7). Let $\rho > 10R$ be such that

$$\{z \in D : d_E(z, x) > \rho\} \tag{6.8}$$

is nonempty. If no such ρ exists we take $K = 11A$ in (6.3) and there is nothing to do. Otherwise we want to show that ρ cannot be too large compared to R.

Choose $w \in E\backslash D$ such that $d_E(w, x) = \rho$. Such a point exists because $E\backslash D$ is connected, unbounded, and has its boundary contained in $\overline{G}$. The key feature of w is that

$$B(w, (2A)^{-1}\rho) \cap E \subseteq E\backslash D. \tag{6.9}$$

Indeed, the left side is contained in $\{z \in E : d_E(z, w) < \rho/2\}$, which is a connected subset of E that intersects $E\backslash D$ (since it contains w) and which is disjoint from G (because it is contained in $\{z \in E : d_E(z, x) > \rho/2\}$). Thus f vanishes on the left side of (6.9).

Similarly, we can find a point $u \in D$ such that $d_E(x, u) = \rho$, since (6.8) is nonempty, and since D is connected and contains G. We also have that

$$B(u, (2A)^{-1}\rho) \cap E \subseteq D. \tag{6.10}$$

This is because the left side of (6.10) is contained in $U = \{z \in E : d_E(z, u) < \rho/2\}$, which is a connected set that is disjoint from G, since it is contained in $V = \{z \in E : d_E(x, z) > \rho/2\}$. If U intersected the complement of D, it would have to be contained in $E\backslash D$, which is impossible, because $u \in D$. We also get that $f(z) = R$ when z lies in the left side of (6.10), because the latter is contained in V.

These observations imply that

$$\int_{E\cap B(w,(2A)^{-1}\rho)} \int_{E\cap B(u,(2A)^{-1}\rho)} |f(y)-f(z)|^p d\mu(y) d\mu(z) \tag{6.11}$$

is $\geq C^{-1}R^p\rho^4$. On the other hand (6.7) (with $t = 2A\rho$) and Lemma 6.6 imply that (6.11) is at most $C\rho^{2+p}R^2$. Using this with any $p \in [1, 2)$ yields $\rho \leq CR$, which is what we wanted. This proves (6.3).

It remains to prove (6.1) and (6.7). These two facts are very closely related, because (6.1) is essentially equivalent to proving the following.

Proposition 6.12. *Let E be a chord-arc surface in $\mathbf{R}^{d+1}$, and assume a priori that $E \cup \{\infty\}$ is a smooth submanifold in $S^{d+1} \cong \mathbf{R}^{d+1} \cup \{\infty\}$. Then*

$$|h(y) - h(z)| \leq C \, \|\nabla_T h\|_\infty |y - z| \tag{6.13}$$

for any Lipschitz function $h\colon E \to \mathbf{R}$, where C depends only on d and the constants that arise in the definition of a chord-arc surface.

Let us explain why this is essentially equivalent to (6.1). We can derive (6.1) from Proposition 6.12 by applying (6.13) with $h(y) = d_E(x, y)$, since h is Lipschitz and $|\nabla_T h| \leq 1$ a.e. Conversely, if we have (6.1), then we can obtain (6.13) in the case where h is C^1 simply by using the fact that

$$|h(\gamma(0)) - h(\gamma(1))| \leq \sup_{0\leq t\leq 1} |\nabla_T h(\gamma(t))| \times (\text{length of } \gamma)$$

for any path $\gamma(t)$ in E. You can pass from the C^1 case to the general Lipschitz case by standard approximation arguments.

Proposition 6.12, like (6.7), can be obtained using the techniques of [S3], which we are about to review. Before entering into the details let us first observe that we may as well restrict our attention to smooth functions, because of standard approximation arguments.

Let $g\colon E \to \mathbf{R}$ be a compactly supported smooth function. Let $G\colon \mathbf{R}^{d+1}\backslash E \to \mathbf{R}$ be the double layer potential of g, so that

$$G(x) = a(d) \int_E \sum_{j=1}^{d+1} \frac{x_j - y_j}{|x-y|^{d+1}} \, n_j(y) g(y) \, d\mu(y), \tag{6.14}$$

where $a(d)$ is some nonzero real constant, and $n_j(y)$, $1 \leq j \leq d+1$, are the components of some smooth choice of unit normal to E.

We are taking here a slightly different approach from that of [S3]. If we were following [S3] more closely we would take G to be the Cauchy integral of g, where "Cauchy integral" is defined in terms of the Cauchy kernel from Clifford analysis. We have in fact taken G to be the real part of the Cauchy integral of g. We are intentionally trying to minimize the role of Clifford analysis here to keep the exposition simple. Although the Clifford analysis is in some sense never essential, it does make some of the computations easier.

There are two nontrivial features of (6.14) that will be crucial for us. The first deals with boundary values. Let Ω_+ and Ω_- denote the two components of $\mathbf{R}^{d+1}\backslash E$, with Ω_+ the one that our choice of the normal vector points into. It is well known (and not too difficult to check) that the restriction of G to Ω_+ or Ω_- has a C^1 extension to the boundary. (This uses our a priori smoothness

assumptions on g and E.) Let G_+ and G_- denote the boundary values of G on E from Ω_+ or Ω_-. Then

$$g = G_+ - G_-. \tag{6.15}$$

The second feature of (6.14) is that derivatives of G can be estimated in terms of $\nabla_T g$. More precisely,

$$|\nabla^j G(x)| \le C(j) \int_E \frac{1}{|x-y|^{d+j-1}} |\nabla_T g(y)| \, d\mu(y) \tag{6.16}$$

for all $x \in \mathbf{R}^{d+1} \backslash E$ and $j = 1, 2$. This follows from Lemma 1.4 in [S3], easy estimates for the Cauchy kernel, and the fact that G is the real part of the Cauchy integral of g. [This is an example of one of those times at which Clifford analysis makes the computations easier.]

The main idea for proving (6.7) and Proposition 6.12 is to control g using (6.15) and by simply integrating the estimates on the derivatives of G in (6.16). It is amusing to notice that we do not need more subtle estimates on the double layer potential or the Cauchy integral. In [S3] some square function estimates for the Cauchy integral were also used, but for the present purposes they are not needed.

Let us now prove (6.7). Fix $x \in E$ and $t > 0$. Choose $x_+ \in \Omega_+$ and $x_- \in \Omega_-$ such that $|x - x_\pm| \le t$, $\operatorname{dist}(x_\pm, E) \ge C^{-1} t$. It suffices to show that

$$\int_{E \cap B(x,t)} (t^{-1} |G_+(y) - G(x_+)|)^p d\mu(y) \le C \int_E |\nabla_T g|^p d\mu, \quad 1 \le p \le 2, \tag{6.17}$$

and similarly for G_-. Indeed, once you have this, you get

$$\int_{E \cap B(x,t)} (t^{-1} |g(y) - (G(x_+) - G(x_-))|)^p d\mu(y) \le C \int_E |\nabla_T g|^p d\mu$$

from (6.15). To get (6.7) you use the fact that the left side of (6.7) is bounded by

$$C \int_{E \cap B(x,t)} (t^{-1} |g(y) - \gamma|)^p d\mu(y) + C \int_{E \cap B(x,t)} (t^{-1} |g(z) - \gamma|)^p d\mu(z)$$

for any constant γ (e.g., $\gamma = G(x_+) - G(x_-)$), and you observe that these last two integrals are the same.

Because Ω_+ is, by assumption, an NTA domain, we can find for each $y \in E \cap B(x, t)$ a nice path between y and x_+. That is, we can find a function $\eta_+ : (0, 1] \to \Omega_+$ such that

$$C^{-1} st \le |\eta_+(s) - y| \le C \operatorname{dist}(\eta_+(s), E) \le C^2 st, \tag{6.18}$$

$$\eta_+(1) = x_+, \tag{6.19}$$

and

$$\left| \frac{d}{ds} \eta_+(s) \right| \le Ct. \tag{6.20}$$

Hence

$$\begin{aligned} |G_+(y) - G(x_+)| &= \left| \int_0^1 \frac{d}{ds} (G(\eta_+(s))) \, ds \right| \le C \int_0^1 t |\nabla G(\eta_+(s))| \, ds \\ &\le C \int_0^1 \int_E \frac{t}{|\eta_+(s) - w|^d} |\nabla_T g(w)| d\mu(w) \, ds. \end{aligned}$$

Using (6.18) it is easy to check that

$$|\eta_+(s) - w| \approx |y - w| + st.$$

Applying Fubini's theorem we get

$$\begin{aligned} |G_+(y) - G(x_+)| &\leq C \int_E \left(\int_0^1 \frac{t}{(|y-w| + st)^d}\, ds \right) |\nabla_T g(w)|\, d\mu(w) \\ &\leq C \int_E L_t(y, w) |\nabla_T g(w)|\, d\mu(w), \end{aligned} \tag{6.21}$$

where $L_t(y, w) = t|y - w|^{-d}$ when $|y - w| \geq t$, and $L_t(y, w) = |y - w|^{1-d}$ when $|y - w| \leq t$. (Actually, this is correct only when $d > 1$; when $d = 1$ minor modifications would have to be made.) It is not hard to derive (6.17) from (6.21) using standard techniques. (In fact the same result holds for $2 \leq p < \infty$, but with a constant that blows up as $p \to \infty$.) A similar argument gives (6.17) with G_+ and x_+ replaced by G_- and x_-. This completes the proof of (6.7).

Let us now prove Proposition 6.12. Let h be as in Proposition 6.12, and fix $y, z \in E$, $y \neq z$. Set $t = |y - z|$. We may assume that h is smooth, as we have already observed.

It is convenient to approximate h by functions with compact support. Let $R > t$ be arbitrary and large, and let θ_R be a smooth function on $\mathbf{R}^{d+1}$ that satisfies $\theta_R(x) = 0$ when $|x - y| \geq 2R$, $\theta_R(x) = 1$ when $|x - y| \leq R$, and $0 \leq \theta_R(x) \leq 1$ and $|\nabla \theta_R(x)| \leq CR^{-1}$ for all x. Define $g: E \to \mathbf{R}$ by $g = \theta_R h$, so that

$$|\nabla_T g| \leq \theta_R\, |\nabla_T h| + HR|\nabla \theta_R|. \tag{6.22}$$

Here H is a horrible constant that depends on our a priori assumptions on E and h. We are not allowed to have H appear in the final estimates; we shall make it go away by sending R to ∞.

Let G be as in (6.14) again, and let $x_+ \in \Omega_+$ and $x_- \in \Omega_-$ be such that $|y - x_\pm| \leq t$ and $\operatorname{dist}(x_\pm, E) \geq C^{-1}t$. Observe that (6.16) and (6.22) imply that

$$|\nabla^2 G(u)| \leq C \operatorname{dist}(u, E)^{-1}\, \|\nabla_T h\|_\infty + CHR^{-1} \tag{6.23}$$

for all $u \in \mathbf{R}^{d+1} \backslash E$ such that $|y - u| \leq R/2$.

We want to show that

$$|G_+(y) - [G(x_+) + \nabla G(x_+) \cdot (y - x_+)]| \leq C\|\nabla_T h\|_\infty t + CHR^{-1}t^2. \tag{6.24}$$

The proof of this uses only (6.23) and the fact that Ω_+ is NTA, and it is very similar to the computations around (2.4)–(2.7) in [S3]. We shall repeat the argument here, for completeness.

There is a sequence of points y_j, $j = 1, 2, 3, \ldots$, in Ω_+ such that $y_1 = x_+$ and, for some $\rho < 1$,

$$|y_j - y| \leq C\rho^j t, \quad \operatorname{dist}(y_j, E) \geq C^{-1}\rho^j t, \quad \text{and} \tag{6.25}$$

$$|y_j - y_{j+1}| \leq \frac{1}{10} \operatorname{dist}(y_j, E). \tag{6.26}$$

Here ρ depends only on the NTA constants of Ω_+. We are going to prove (6.24) by comparing the affine Taylor approximations of G at the y_j's.

Specifically, we have that

$$
\begin{aligned}
&G_+(y) - G(x_+) - \nabla G(x_+)\cdot(y - x_+) \\
&\quad = \sum_{j=1}^{\infty}\{G(y_{j+1}) - G(y_j) - \nabla G(x_+)\cdot(y_{j+1} - y_j)\} \\
&\quad = \sum_{j=1}^{\infty}\{G(y_{j+1}) - G(y_j) - \nabla G(y_j)\cdot(y_{j+1} - y_j)\} \\
&\qquad + \sum_{j=1}^{\infty}(\nabla G(y_j) - \nabla G(x_+))\cdot(y_{j+1} - y_j).
\end{aligned} \tag{6.27}
$$

Using (6.26), (6.23), (6.25), and calculus we get that

$$
\begin{aligned}
&|G(y_{j+1}) - G(y_j) - \nabla G(y_j)\cdot(y_{j+1} - y_j)| \\
&\qquad \le C\rho^{2j}t^2(\rho^{-j}t^{-1}\|\nabla_T h\|_\infty + HR^{-1}).
\end{aligned} \tag{6.28}
$$

Similarly we have that

$$
\begin{aligned}
|\nabla G(y_j) - \nabla G(x_+)| &\le \sum_{i=1}^{j-1}|\nabla G(y_{i+1}) - \nabla G(y_i)| \\
&\le C\sum_{i=1}^{j-1}\rho^i t(\rho^{-i}t^{-1}\|\nabla_T h\|_\infty + HR^{-1}) \\
&\le Cj\|\nabla_T h\|_\infty + CtHR^{-1},
\end{aligned}
$$

and hence

$$
\begin{aligned}
&\left|\sum_{j=1}^{\infty}(\nabla G(y_j) - \nabla G(x_+))\cdot(y_{j+1} - y_j)\right| \\
&\qquad \le C\sum_{j=1}^{\infty}(j\|\nabla_T h\|_\infty + tHR^{-1})\rho^j t \le Ct\|\nabla_T h\|_\infty + Ct^2HR^{-1}.
\end{aligned}
$$

Using this and (6.28) in (6.27) gives (6.24).

A similar argument gives the analogue of (6.24) with y replaced by z. Combining these we obtain

$$
|G_+(y) - G_+(z) - \nabla G(x_+)\cdot(y - z)| \le C\|\nabla_T h\|_\infty t + CHR^{-1}t^2. \tag{6.29}
$$

The corresponding fact with " + " replaced by " − " is also true, and the combination of the two with (6.15) yields

$$
|g(y) - g(z) - (\nabla G(x_+) - \nabla G(x_-))\cdot(y - z)| \le C\|\nabla_T h\|_\infty t + CHR^{-1}t^2. \tag{6.30}
$$

We are almost finished, but we need to explain why

$$
|\nabla G(x_+) - \nabla G(x_-)| \le C\int_E \frac{t}{(|y - w| + t)^{d+1}}|\nabla_T g(w)|d\mu(w) \tag{6.31}
$$

is true. The reason is the same as for (6.16). Because G is the real part of the Cauchy integral of g, (6.31) reduces via Lemma 1.4 in [S3] to easy estimates for the Cauchy kernel.

From (6.31) and (6.22) we get

$$|\nabla G(x_+) - \nabla G(x_-)| \leq C\|\nabla_T h\|_\infty + CHR^{-1}t.$$

Combining this with (6.30) gives

$$|g(y) - g(z)| \leq C\|\nabla_T h\|_\infty t + CHR^{-1}t^2. \tag{6.32}$$

By definitions we have $|g(y) - g(z)| = |h(y) - h(z)|$ and $t = |y - z|$, and so (6.13) follows from (6.32) by sending R to ∞. This finishes the proof of Proposition 6.12.

References

[D1] G. David, *Opérateurs intégraux singuliers sur certaines courbes du plan complexe*, Ann. Sci. Ecole Norm. Sup. **17** (1984), 157–189.

[D2] ——, *Opérateurs d'intégrale singuliere sur les surfaces régulières*, Ann. Sci. Ecole Norm. Sup. (4) **21** (1988), 225–258.

[D3] ——, *Morceaux de graphes lipschitziennes et intégrales singulières sur un surface*, Rev. Mat. Iberoamericana **4** (1988), 73–114.

[D4] ——, *Wavelets and singular integrals on curves and surfaces*, Lecture Notes in Math., vol. 1465, Springer-Verlag,, 1991.

[DJ] G. David and D. Jerison, *Lipschitz approximations to hypersurfaces, harmonic measure, and singular integrals*, Indiana Math. J. **39** (1990), 831–845.

[DS] G. David and S. Semmes, *Singular integrals and rectifiable sets in* $\mathbf{R}^n$: *au-delà des graphes lipschitziens*, Astérisque **193** (1991).

[DS2] ——, *Analysis of and on uniformly rectifiable sets*, manuscript, 1992.

[Do] J. R. Dorronsoro, *A characterization of potential spaces*, Proc. Amer. Math. Soc. **95** (1985), 21–31.

[F] H. Federer, *Geometric measure theory*, Springer-Verlag, 1969.

[G] F. Gehring, *The* L^p*-integrability of the partial derivatives of a quasiconformal mapping*, Acta Math. **130** (1973), 265–277.

[JK] D. Jerison and C. Kenig, *Boundary behavior of harmonic functions in nontangentially accessible domains*, Adv. in Math. **46** (1982), 80–147.

[J1] P. W. Jones, *Square functions, Cauchy integrals, analytic capacity, and harmonic measure*, Harmonic Analysis and Partial Differential Equations, edited by J. Garcia-Cuerva, Lecture Notes in Math., vol. 1384, Springer-Verlag, 1989.

[J2] ——, *Lipschitz and bi-Lipschitz functions*, Rev. Mat. Iberoamericana **4** (1988), 115–122.

[J3] ——, *Rectifiable sets and the travelling salesman problem*, Invent. Math. **102** (1990), 1–15.

[Ma] P. Mattila, *Lecture notes on geometric measure theory*, Departmento de Matemáticas, Universidad de Extremadura, 1986.

[Mu] T. Murai, *A real-variable method for the Cauchy transform, and analytic capacity*, Lecture Notes in Math., vol. 1307, Springer-Verlag, 1988.

[S1] S. Semmes, *A criterion for the boundedness of singular integrals on hypersurfaces*, Trans. Amer. Math. Soc. **311** (1989), 501–513.

[S2] ——, *Chord-arc surfaces with small constant* II: *good parameterizations*, Adv. in Math. **88** (1991), 170–199.

[S3] ——, *Differentiable function theory on hypersurfaces in* $\mathbf{R}^n$ (*without bounds on their smoothness*), Indiana Math. J. **39** (1990), 985–1004.

[S4] ——, *Analysis vs geometry on a class of rectifiable hypersurfaces in* $\mathbf{R}^n$, Indiana Math. J. **39** (1990), 1005–1035.

[St] E.M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, 1970.

[V] J. Väisälä, *Invariants for quasisymmetric, quasimöbius, and bilipschitz maps*, J. Analyse Math. **50** (1988), 201–233.

MATHÉMATIQUES—BÂT. 425, UNIVERSITÉ DE PARIS-SUD, 91405 ORSAY, FRANCE
E-mail address: gdavid@anh.matups.fr

DEPARTMENT OF MATHEMATICS, RICE UNIVERSITY, HOUSTON, TEXAS 77251
E-mail address: semmes@rice.edu