GUY DAVID
STEPHEN SEMMES

Singular integrals and rectifiable sets in $\mathbb{R}^n$. Au-delà des graphes lipschitziens

Astérisque, tome 193 (1991)

<http://www.numdam.org/item?id=AST_1991__193__1_0>

193 ASTÉRISQUE

1991

SINGULAR INTEGRALS AND RECTIFIABLE SETS IN $\mathbb{R}^n$

Au-delà des graphes lipschitziens

Guy David, Stephen Semmes

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

Publié avec le concours du CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

A.M.S. Subjects Classification : 49F20, 42B40.

Introduction

There are a number of natural ways to look at the goals and results of this monograph. The first can be stated broadly as the problem of relating the geometry of a set E in $\mathbf{R}^n$ to the analysis of functions and linear operators on E. A specific question of this type that we shall be concerned with here is the following. Let E be a subset of $\mathbf{R}^n$ that has Hausdorff dimension d, $0 < d < n$. We equip E with d-dimensional Hausdorff measure restricted to it, and we assume that this measure is locally finite. Under what conditions on E is it true that plenty of singular integral operators are bounded on $L^2(E)$? Examples of the sort of singular integral operators that we have in mind are the Cauchy integral when $d = 1$ and $n = 2$ and the double-layer potential when $d = n - 1$.

It is known from the work of Coifman, McIntosh, and Meyer [**CMM**] that this is true when E is a Lipschitz graph. There are several more general conditions on E which are known to be sufficient to ensure the boundedness of lots of singular integral operators, but there has not been much progress on finding necessary conditions. Our main result provides geometrical characterizations of the sets E for which a fairly large class of singular integral operators are bounded on $L^2(E)$, at least if we make an auxiliary technical assumption on E (Ahlfors regularity). See Section 1 for the precise statement. Unfortunately we do not know at this time how to work with smaller classes of operators; for example, when $d = 1$ and $n = 2$ we would like to use only the boundedness of the Cauchy integral.

The geometrical conditions that arise in the aforementioned theorem can be thought of as quantitative analogues of the classical notion of rectifiability. Recall that E is said to be rectifiable if it is contained in the union of a countable family of Lipschitz images of $\mathbf{R}^d$, except for a set of d-dimensional Hausdorff measure zero. Rectifiability is a qualitative condition, and it is not strong enough to imply the boundedness of singular integral operators. Using real-variable methods as in [**D1**, **3**] it can be shown that various quantitative versions of rectifiability are strong enough to imply the boundedness of plenty of singular integral operators, and our theorem provides a converse to this.

There is a great deal of information available about rectifiable sets. See [**Fe**], [**Fl**], and [**Ma**], for instance. Not so much seems to be known about quantitative analogues of rectifiability. In particular there are many characterizations of rectifiability, and these give rise to many candidates for the notion of quantitative rectifiability, but the complete relationship between these various candidates is not at all clear. Our theorem provides some nontrivial equivalences between some of these conditions. Although this is a purely geometrical issue, it turns out that singular integral operators provide a useful tool for passing between some of these conditions.

Our main result also gives a higher-dimensional version of Peter Jones' travelling salesman theorem ([**J3**]). That is, we give two other conditions on E that are equivalent to the others, and which are roughly as follows. One of these conditions says that E is contained in a set that admits a nice parameterization by $\mathbf{R}^d$. The other condition is a bound on certain quantities that measure the extent to which E can be approximated by d-planes. Again, Section 1 should be consulted for the precise statement.

Although there are several ways of looking at what we are doing and what it means, there is an underlying common theme. To a large degree we are trying to produce methods for analyzing the geometry of sets, in much the same way that more traditional harmonic analysis (as in [**St**]) is concerned with the analysis of functions and operators. Some of the ideas of harmonic analysis make sense in this context, but mostly the techniques don't work so well, because of the absence of a linear structure. The methods that have grown out of Carleson's corona construction seem to be more cooperative in this geometrical setting.

In connection with the analogy with traditional harmonic analysis it is interesting to look at the theorem in Section 1 from the perspective of Littlewood-Paley theory. In some sense this theorem gives a Littlewood-Paley characterization of a class of good sets that is analogous to well-known results for Sobolev spaces. It turns out that this analogy is somewhat misleading, in that there are some other results in our geometrical context that do not have a natural counterpart for Sobolev spaces. Such a result is discussed just before Lemma 5.13, but its details will appear elsewhere.

The precise statements of our main results are given in Section 1, along with some background information and organizational details. It is perhaps worth mentioning now that there is a discussion of open problems in Section 21. In that section there is also some limited description of other work in this general area. More information of that nature can be found in [**D5**].

The authors would like to thank R.R. Coifman and P.W. Jones for many helpful discussions. Portions of this work were completed while the second author was visiting the first author at the Ecole Polytechnique and UCLA.

The second author is partially supported by the U.S. National Science Foundation and the Alfred P. Sloan Foundation.

Table of Contents

1. Statement of the main results

Let E be a subset of $\mathbf{R}^n$ with Hausdorff dimension d, $0 < d < n$. Unless explicitly stated otherwise, d will always be an integer. Assume that d-dimensional Hausdorff measure H^d is locally finite when restricted to E. Consider singular integral operators on E of the form

$$Tf(x) = p.v. \int_E K(x-y)\, f(y)\, dy, \tag{1.1}$$

where dy denotes $H^d\,|_E$, and where $K(x)$ is smooth on $\mathbf{R}^n \setminus \{0\}$, odd, and satisfies

$$|\, \nabla^j K(x) \,| \leq C(j) \,|\, x \,|^{-d-j}, \quad j = 0, 1, 2, \ldots \tag{1.2}$$

We would like to know what conditions on E are needed in order for (1.1) to define a bounded operator on $L^2(E)$, say.

For technical reasons it is better not to look for an L^2 estimate for (1.1) but rather an estimate like

$$\sup_{\epsilon>0} \int_E \Big|\, \int_{E\cap\{|x-y|>\epsilon\}} K(x-y)\, f(y)\, dy \,\Big|^2 dx \leq C(K) \int_E |\, f \,|^2\, dx, \tag{1.3}$$

for all $f \in L^2(E)$. This formulation avoids the problem of the existence of a principal value.

An important example of such a function $K(x)$ is the Cauchy kernel, i.e., $K(x) = \frac{1}{x_1 + ix_2}$ for $x \in \mathbf{R}^2$, $d = 1$, $n = 2$. This case is of course relevant for complex analysis; for instance, the L^2 boundedness for the associated operator is closely related to the analytic capacity of E and its subsets. (See

[G2], [Mu], [C1, 2].) A higher-dimensional analogue of the Cauchy kernel is $K(x) = \frac{x}{|x|^n}$, $d = n-1$. One way that this kernel arises is in connection with the double-layer potential, which can be expressed in terms of $K(x)$.

It is easy to see that an operator as in (1.1) is bounded on $L^2(E)$ if E is a smooth submanifold (e.g., $C^{1,\alpha}$, $\alpha > 0$) which is nice at ∞, or a subset of such a submanifold. This is also true if E is a Lipschitz graph over some d-plane, but it is much harder to prove. For the Cauchy kernel, this is a theorem of Coifman, McIntosh, and Meyer [CMM], improving an earlier result of Calderón (which covered the case of graphs of Lipschitz functions with small norm). The case of general kernels was derived from this result in [CDM], at least when $d = 1$, $n = 2$. The higher-dimensional case can easily be obtained from this using the method of rotation, just like the argument in Section 13 of [CMM].

The fact that Lipschitz graphs are O.K. for these operators shows that the smoothness of E is not the issue, but it is not obvious how wild E can be. The following two results help to clarify the situation.

The first says that if $d = 1$, E is a curve, and, say, $n = 2$ and K is the Cauchy kernel, then T is bounded on $L^2(E)$ if and only if there is a $C > 0$ such that

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

for all $x \in \mathbf{R}^2$, $R > 0$, where $B(x,R)$ denotes the ball with center x and radius R. Such curves are often called regular curves. This was proved in [D1].

The second result goes as follows. Let G be the Cantor set in $[0,1]$ obtained from the usual construction, except that you remove the middle half of the interval at each stage. Then $E = G \times G$ satisfies (1.4), and also

$$H^1(E \cap B(x,R)) \geq C^{-1}R$$

for all $x \in E$, $0 < R \leq 1$, but the Cauchy integral operator is not bounded on $L^2(E)$. This follows from [G1], see also [G2], [D2], [J1], and [Ma3].

These two results suggest that rectifiability plays a role here. To make this precise it is helpful to recall a couple of definitions and facts from geometric measure theory.

Let A be a subset of $\mathbf{R}^n$ with Hausdorff dimension d. We say that A is (countably) rectifiable if there is a countable family f_j of Lipschitz maps of $\mathbf{R}^d$ into $\mathbf{R}^n$ such that

$$H^d(A \setminus \Big(\bigcup_j f_j(\mathbf{R}^d)\Big)) = 0,$$

so that A is almost covered by the union of the images of the f_j's. It turns out that this is equivalent to requiring that A be almost covered by a countable family of d-dimensional Lipschitz graphs, or even C^1 submanifolds. A is said to be unrectifiable if $H^d(A \cap B) = 0$ for all rectifiable sets B. A basic fact is that any set A with $H^d(A) < \infty$ can be written as the union of a rectifiable set and an unrectifiable one.

For example, any subset of a curve of finite length is rectifiable, but it can be shown that $G \times G$ is unrectifiable (with $d = 1$ in both cases).

References for these topics include **[Ma]**, **[Fl]**, and **[Fe]**.

You might hope that if E is rectifiable, and if you have some control on $H^d \mid_E$ (like (1.4), when $d = 1$), then singular integrals have to be bounded on $L^2(E)$. This is not true, but for a good reason; rectifiability is a qualitative condition, while estimates on singular integrals are quantitative. It is not hard to build sets that are rectifiable, but which have pieces that approximate unrectifiable sets (like $G \times G$) on which singular integrals do not define bounded operators, in such a way that singular integral operators are not bounded on the set you've constructed.

Thus we need to look for quantitative notions of rectifiability. We give an example of such a notion after the following definition.

DEFINITION. *A set $E \subseteq \mathbf{R}^n$ is regular (with dimension d) if it is closed and if*

$$\frac{1}{C} R^d \leq H^d(E \cap B(x, R)) \leq C R^d \tag{1.5}$$

for all $x \in E$, $R > 0$, where C does not depend on x, R.

We shall assume throughout the rest of this paper that E is regular, and we shall often write $|A|$ for $H^d(A)$ when $A \subseteq E$.

It is not hard to show that if singular integrals like (1) define bounded operators on $L^2(E)$, then the right hand inequality in (1.5) must hold. (See **[D5]**, **[S1]**.) The left hand inequality should be thought of as a nondegeneracy condition. Note that it is translation and dilation invariant.

Notice also that if E is closed and if there is a measure σ supported on E for which the analogue of (1.5) holds, then σ must be equivalent in size to $H^d \mid_E$, and E must be regular. Thus we lose nothing by restricting ourselves to Hausdorff measure here.

DEFINITION. *E has BPLG (big pieces of Lipschitz graphs) if it is regular and if there exist C, $\epsilon > 0$ so that for every $x \in E$, $R > 0$ there is a*

d-dimensional Lipschitz graph Γ with constant $\leq C$ such that

$$H^d(E \cap B(x,R) \cap \Gamma) \geq \epsilon R^d. \tag{1.6}$$

When we say that Γ is a d-dimensional Lipschitz graph with constant $\leq C$ we mean that there is a d-plane P, an $(n-d)$-plane $P^\perp$ orthogonal to P, and a Lipschitz function $A : P \to P^\perp$ with norm $\leq C$ such that $\Gamma = \{p + A(p) : p \in P\}$.

This condition is a good example of what we mean by a quantitative version of rectifiability. It is not hard to see that E is rectifiable if it has BPLG, but the converse is not true; rectifiability only allows you to conclude (1.6) with an ϵ that depends on x, R.

Notice that if you fix C, ϵ and look at the class of sets having BPLG with constants C, ϵ, and which also satisfy (1.5) with this same C, then this class is invariant under translations, rotations, and dilations, and it is also closed in the Hausdorff topology on closed subsets on $\mathbf{R}^n$. (In this topology $E_j \to E$ if for every ϵ, $R > 0$ and all j sufficiently large we have that each point in $E_j \cap B(0,R)$ is within ϵ of an element of E, and vice-versa.) We do not know whether it is true that any class of rectifiable subsets of $\mathbf{R}^n$ with these same invariance and closure properties has to be contained in the class of sets that have BPLG.

It follows from [**D1**, **3**] that if the regular set E has BPLG, then the estimate (1.3) holds for all K as before. We do not know if the converse is true. However, there are some other "quantitative rectifiability" conditions similar to (and a priori weaker than) BPLG which also imply (1.3), and for which we are able to obtain a converse. We also obtain other analytical and geometrical characterizations of these sets.

THEOREM. *Let $E \subseteq \mathbf{R}^n$ be a regular d-dimensional set in $\mathbf{R}^n$. The following conditions* (C1)-(C7) *are equivalent.*

(C1) *If $K(x)$ is any smooth odd function on $\mathbf{R}^n \setminus \{0\}$ that satisfies* (1.2), *then* (1.3) *holds.*

Using a standard fact from Calderón-Zygmund theory (Cotlar's inequality — see [**JL**]) we have that (1.3) is equivalent to

$$\int_E \sup_{\epsilon>0} \left| \int_{E\cap\{|x-y|>\epsilon|\}} K(x-y) f(y) dy \right|^2 dx \leq C(K) \int_E |f|^2 \tag{1.7}$$

for all $f \in L^2(E)$. [To be honest we should admit that in order to apply the techniques of Calderón-Zygmund theory we should first observe that E is a space of homogeneous type, in the sense of [**CW**], when E is equipped with the measure $H^d \mid_E$ and the Euclidean distance.] Calderón-Zygmund theory also implies that these L^2 estimates are equivalent to their L^p counterparts, $1 < p < \infty$, etc. (See [**JL**] again.)

(C2) *For each smooth odd function* ψ *on* $\mathbf{R}^n$ *with compact support we have that*

$$\sum_{k=-\infty}^{\infty} \left| \int_E \psi_k(x-y)dy \right|^2 dx\, d\delta_{2^k}(t) \tag{1.8}$$

is a Carleson measure on $E \times \mathbf{R}_+$.

Here $\psi_k(x) = 2^{-kd}\psi(2^{-k}x)$ and $d\delta_s(t)$ denotes the Dirac mass at s in t. A Carleson measure on $E \times \mathbf{R}_+$ is a measure μ for which there is a $C > 0$ such that for every $x \in E$ and $R > 0$ we have

$$\int_0^R \int_{B(x,R)} d\mu \leq CR^d.$$

Thus Carleson measures are measures on $E \times \mathbf{R}_+$ that behave as though they are d-dimensional near $E \times \{0\} \cong E$.

This condition is quite natural despite its technical appearance. Because ψ is odd, $\int_E \psi_k(x-y)dy$ is zero if E is a d-plane; thus this quantity measures in some way how close E is to being a d-plane.

One can think of (C2) as a geometrical analogue of classical characterizations of various function spaces in terms of the size of expressions like $\left| \int_{\mathbf{R}^d} \psi_k(x-y)f(y)dy \right|$. This geometric Littlewood-Paley point of view is discussed somewhat more thoroughly in [**DS1**].

(C3) $\beta_1(x,t)^2 \frac{dx\,dt}{t}$ *is a Carleson measure on* $E \times \mathbf{R}_+$.

For $x \in E$, $t > 0$ we define $\beta_1(x,t)$ by

$$\beta_1(x,t) = \inf_P \frac{1}{t^d} \int_{E \cap B(x,t)} \frac{\operatorname{dist}(y,P)}{t}\, dy,$$

where the infimum is taken over all d-planes P. Thus $\beta_1(x,t)$ measures how well E can be approximated by a d-plane.

Peter Jones was the first person (to our knowledge) to look at this kind of condition. Actually, he worked with an L^∞ version, i.e., with $\beta_\infty(x,t)$, where

$$\beta_q(x,t) = \inf_P \left(\frac{1}{t^d} \int_{E\cap B(x,t)} \left(\frac{\operatorname{dist}(y,P)}{t} \right)^q dy \right)^{\frac{1}{q}} . \tag{1.9}$$

In [**J1**] he showed how to use the fact that Lipschitz graphs satisfy (C3) (with $\beta_\infty(x,t)$) to give a new approach to the estimates for Cauchy integrals on Lipschitz graphs. He later found a characterization of subsets of rectifiable curves in terms of a (related) quadratic condition on the β_∞'s. In particular he showed that subsets of regular curves can be characterized by a quadratic Carleson measure condition on the β_∞'s. Our results give analogues of this characterization for $d > 1$.

If we replace β_1 by β_q in (C3), then we still get an equivalent condition as long as $q < \frac{2d}{d-2}$ ($q \le \infty$ if $d = 1$). However, Jones and Fang have produced 3-dimensional Lipschitz graphs so that (C3) does not hold for β_∞.

There is a classical counterpart of (C3) for functions just as there was for (C2). Given a function f on $\mathbf{R}^d$, set

$$\gamma_q(x,t) = \inf_a \left(\frac{1}{t^d} \int_{B(x,t)} \left(\frac{|f(y)-a(y)|}{t} \right)^q dy \right)^{\frac{1}{q}} , \tag{1.10}$$

where now the infimum is over all affine functions. Notice that for Lipschitz functions the γ's for f are essentially equivalent to the β's for the graph of f.

The γ's can be used to characterize smoothness properties of f, e.g., whether f lies in a particular Sobolev space. (See [**Do**], for instance.) This is closely related to the corresponding results using second differences instead (see [**St**]), which are perhaps more familiar. In Section 19 we shall give a condition (C8) that is a geometrical version of a second-difference condition for functions, and we shall show that it is equivalent to the others.

(C4) *E admits a corona decomposition.*

The precise explanation of this condition is complicated and will be postponed until the next section. Roughly speaking it means that you can decompose $E \times \mathbf{R}_+$ into two pieces, the good and the bad parts, with the following properties. The bad part is not too big, in that it is controlled by a Carleson measure. The good part can be subdivided into stopping time regions on each of which E is well-approximated by a Lipschitz graph. There aren't too many of these regions, in that they satisfy a Carleson measure packing condition.

This condition plays a central role for us, acting as a bridge between (C1)–(C3) and (C5)–(C7). Although it is awkward to state, it carries a lot of useful information, and is not so hard to work with.

(C5) *E has very big pieces of bilipschitz images of $\mathbf{R}^d$ inside $\mathbf{R}^{n^*}$, $n^* = \max(n, 2d+1)$.*

This means that for every $\epsilon > 0$ there is an $M > 0$ so that for each $x \in E$ and $R > 0$ there is a mapping $\rho : \mathbf{R}^d \to \mathbf{R}^{n^*}$ which is bilipschitz with constant M, i.e.,

$$\frac{1}{M}|x-y| \leq |\rho(x) - \rho(y)| \leq M|x-y| \quad \text{for all } x, y \in \mathbf{R}^d, \tag{1.11}$$

and whose image almost contains $B(x, R) \cap E$, that is,

$$|\, E \cap B(x,R) \setminus \rho(\mathbf{R}^d) \,| \leq \epsilon R^d.$$

Here we identify $\mathbf{R}^n$ with a subset of $\mathbf{R}^{n^*}$ in the obvious way.

(C6) *E has big pieces of Lipschitz images of subsets of $\mathbf{R}^d$.*

This means that there exist ϵ, $M > 0$ so that for every $x \in E$, $R > 0$ there is a Lipschitz mapping ρ with norm $\leq M$ from the ball $B_d(0, R)$ in $\mathbf{R}^d$ into $\mathbf{R}^n$ such that

$$|\, E \cap B(x,R) \cap \rho(B_d(0,R)) \,| \geq \epsilon R^d.$$

It follows from the main result in **[J1]** that (C6) is equivalent to the condition you get by replacing $B(0, R)$ with a subset F of $B(0, R)$ and "Lipschitz" with "bilipschitz." [To apply that result it is useful to notice that Hausdorff measure is equivalent in size to Hausdorff content for subsets of a regular set.] Also, if $\rho : F \to \mathbf{R}^n$ is bilipschitz, $F \subseteq \mathbf{R}^d$, then you can extend ρ to a bilipschitz mapping on $\mathbf{R}^d$, at least if you replace $\mathbf{R}^n$ with $\mathbf{R}^{n^*}$, as we shall discuss in Section 17.

(C7) *There is an A_1-weight ω on $\mathbf{R}^d$ and an ω-regular mapping* $z : \mathbf{R}^d \to \mathbf{R}^{n+1}$ *whose image contains E.*

Recall that $\omega(x)$ is an A_1-weight on $\mathbf{R}^d$ if it is a positive locally-integrable function such that for each ball B,

$$\frac{1}{|B|}\int_B \omega \;\leq\; C \operatorname*{ess\,inf}_B \omega. \tag{1.12}$$

It is well-known that this implies that ω is an A_∞ weight, which can be characterized by the existence of C, $\delta > 0$ such that

$$\left(\frac{1}{|B|}\int_B \omega^{1+\delta}\right)^{\frac{1}{1+\delta}} \;\leq\; C\,\frac{1}{|B|}\int_B \omega \tag{1.13}$$

for all balls B. (See **[JL]** for basic facts about weights.)

As in **[D3]** we say that $z : \mathbf{R}^d \to \mathbf{R}^{n+1}$ is an ω-regular mapping if ω is an A_∞ weight, z has locally integrable distributional derivatives, $|\nabla z| \leq C\omega^{\frac{1}{d}}$ a.e., and

$$\omega(z^{-1}(B(y,R))) \;\leq\; CR^d \tag{1.14}$$

for all $y \in \mathbf{R}^{n+1}$, $R > 0$, where $\omega(A) = \int_A \omega$.

If $z(\cdot)$ is ω-regular, and B is any ball in $\mathbf{R}^d$, then

$$\operatorname{diam}(z(B)) \leq C\omega(B)^{\frac{1}{d}}. \tag{1.15}$$

This can be derived from $|\nabla z| \leq C\omega^{\frac{1}{d}}$ and (1.13) using standard results. Conversely, it is not hard to show that if (1.15) holds for an A_∞-weight ω, then z has locally integrable distributional derivatives and $|\nabla z| \leq C\omega^{\frac{1}{d}}$ a.e..

It is also not hard to show that if $z(\cdot)$ is ω-regular, then its image $\widetilde{E} = z(\mathbf{R}^d)$ is a regular set, and ω is equivalent in size to the pull-back of Hausdorff measure. Notice that $z(\cdot)$ is 1-regular if it is bilipschitz.

It is proven in **[D3]** that (C1) holds for $\widetilde{E} = z(\mathbf{R}^d)$ if $z(\cdot)$ is an ω-regular mapping with $\omega \in A_\infty$. Thus (C7) implies (C1) even if we weaken the requirement $\omega \in A_1$ to $\omega \in A_\infty$.

When $n \geq 2d$ we can replace $\mathbf{R}^{n+1}$ in (C7) by $\mathbf{R}^n$ and still have an equivalent condition. This is proved using the methods of **[D3]**, Section 5. The main interest in this observation comes from the case $d = 1$, $n =$

2. Notice that when $d = 1$ we can always take $w \equiv 1$, because we can reparameterize $z(\mathbf{R})$ by arclength.

We originally derived the equivalence of (C7) with (C1)–(C6) when $d = 1$ from Peter Jones' **[J3]** characterization of the subsets of regular curves in terms of his version of (C3). The equivalence of (C3) and (C7) when $d > 1$ provides a higher-dimensional version of his result, at least for regular sets.

This finishes the statement of conditions (C1)–(C7) (except for (C4)), whose equivalence is stated by the theorem. Notice that the ambient dimension n does not play a serious role.

Let us now describe the routing of implications that we follow in proving the theorem. It is relatively easy to show that (C1) implies (C2). The proofs that each of (C2) and (C3) imply (C4) are quite similar and they constitute the main step in the proof of the theorem. The proof that (C4) implies (C3) is pretty straightforward but messy. Both of (C5) and (C7) will be obtained from (C4) by direct constructions. Of course (C6) is a trivial consequence of (C5). It follows from **[D3]** that (C1) holds if any of (C5), (C6), or (C7) do. (In the case of (C6) we also use the result of **[J3]** as discussed above.) You can also derive (C1) directly from (C4), as in **[S4]**.

We should point out that our methods for deriving (C5) or (C7) from (C4) are quite constructive, although somewhat messy. The stopping-time argument given in Section 7 for producing a corona decomposition (if one exists) is both constructive and fairly simple, and one could imagine asking a computer to do it. The difficult part — proving that a corona decomposition does exist if (C2) or (C3) holds — is not the computer's problem.

A curious feature of our arguments is that we do not know how to pass from (C2) to (C1) analytically, without going through the geometry. One can look for analogues of some of the well-known methods for controlling singular integrals on $\mathbf{R}^n$ using square functions (via reproducing formulas, for example), but we have not been able to make anything like that work here. Similarly, it is not so clear how to pass from (C6) to (C5), or from (C5) to (C7), without going through singular integrals.

We also take up a version of the main theorem for fractional dimensional sets. Conditions (C1) and (C2) still make sense in this case, although (C3)–(C7) don't. We shall prove that if E is a d-dimensional regular set, d noninteger, then neither (C1) nor (C2) can hold.

The organization of the remaining sections is as follows. In Section 2 we cover some preliminary material and also give the precise defintion of a corona decomposition. We prove that (C1) implies (C2) in Section 3. We derive in Sections 4, 5, and 6 some geometrical consequences of (C2), and in particular we show in Section 5 that neither (C1) nor (C2) can hold if E has fractional dimension.

We set up in Sections 7 and 8 the initial machinery common to the proofs that (C4) holds if either (C2) or (C3) do. We also give an outline of the argument used to show that (C2) implies (C4) in Section 7, and the details are carried out in Sections 9, 10, and 11. The proof that (C3) implies (C4) is given in Sections 12, 13, and 14.

We show that (C4) implies (C3), (C5), and (C7) in Sections 15, 16 and 17, and 18, respectively. In Section 19 we state condition (C8), a variant of (C2) and (C3), and we indicate why it is equivalent to the other conditions. We give a counterexample in Section 20 to show that the "weak geometric lemma" (see Section 5) is not strong enough to imply rectifiability, even if E is regular. In the last section we discuss some open problems, concerning the theorem and its proof as well as other related topics. In so doing we also give some small and partial indications of other work in this general area. A more substantial overview can be found in [**D5**].

We should also indicate the interdependence of the sections. Section 2 is essential for most of what we do. Sections 3, 15, (16 + 17), 18, 20, and 21 are all independent of each other and also Sections 4–14, with minor exceptions. The proof that (C2) implies (C4) is given in Sections 4–11, while the proof that (C3) implies (C4) uses parts of Section 5 and also Sections 7, 8, 12, 13, 14, and 11, in that order. The details of Section 19 rely on the proof of the equivalence of (C3) with (C4).

2. Dyadic cubes and the corona decomposition

As in **[D4]** one can build a family of subsets of E that play much the same role that dyadic cubes do for $\mathbf{R}^d$. More precisely, there is a family of partitions Δ_j of E, $j \in \mathbf{Z}$, into "cubes" Q with the following properties:

(2.1) if $j \leq k, Q \in \Delta_j$, and $Q' \in \Delta_k$, then either $Q \cap Q' = \emptyset$ or $Q \subseteq Q'$;

(2.2) if $Q \in \Delta_j$, then $C^{-1}2^j \leq \operatorname{diam} Q \leq C2^j$ and $C^{-1}2^{jd} \leq |Q| \leq C2^{jd}$.

The cubes can also be built in such a way that they have relatively small boundary, like ordinary cubes in $\mathbf{R}^d$ do:

(2.3) if $Q \in \Delta_j$ and $\tau > 0$, then
$$| \{x \in Q_j : \operatorname{dist}(x, E \setminus Q) \leq \tau 2^j\} | \leq C\tau^{\frac{1}{C}} 2^{jd}.$$

Of course it is important that the constant C in (2.2) and (2.3) does not depend on j, Q, or τ.

The properties of the Q's and Δ_j's given in **[D4]** are not quite the same as those above, but the same kind of construction can still be used. For a slightly better proof, see **[D5]**. An extension of this result can be found in [C2].

We shall follow the standard practice of referring to the cubes that contain a given cube as its ancestors, referring to its subcubes in the next generation as its children, etc.

Let $\Delta = \cup \Delta_j$ denote the set of all our cubes. We can think of Δ as providing a discrete version of $E \times \mathbf{R}_+$, by letting $(x,t) \in E \times \mathbf{R}_+$ correspond to the $Q \in \Delta_j$ with $x \in Q$ and $2^j \leq t < 2^{j+1}$.

We say that E admits a corona decomposition if for each $\eta > 0$ (think of η as being small) there is a $C = C(\eta) > 0$ such that we can partition Δ into a good set $\mathcal{G}$ and a bad set $\mathcal{B}$ with the following features.

The bad set is not too large, in that it satisfies the Carleson measure packing condition

$$\sum_{\substack{Q \in \mathcal{B} \\ Q \subseteq R}} |Q| \leq C|R| \qquad \text{for all } R \in \Delta. \tag{2.4}$$

The good set $\mathcal{G}$ can be partitioned into a family $\mathcal{F}$ of subsets S of $\mathcal{G}$ such that:

(2.5) each S is coherent, which means that it has a maximal element $Q(S)$, and that if $Q \in S$, $Q' \in \Delta$, $Q \subseteq Q' \subseteq Q(S)$, then $Q' \in S$;

(2.6) from the viewpoint of each S, E is well-approximated by a d-dimensional Lipschitz graph Γ with constant $\leq \eta$, in the sense that for each $Q \in S$ we have

$$\operatorname{dist}(x, \Gamma) \leq \eta \operatorname{diam} Q \text{ whenever } x \in E, \operatorname{dist}(x, Q) \leq \operatorname{diam} Q;$$

(2.7) there aren't too many of the S's, in that they satisfy the packing condition

$$\sum_{\substack{S \in \mathcal{F} \\ Q(S) \subseteq R}} |Q(S)| \leq C|R| \qquad \text{for all } R \in \Delta.$$

There are a number of places where something like a corona decomposition has been used before. One is the work of Garnett and Jones **[GJ]** on the corona theorem for Denjoy domains. Although what they did is somewhat different in details it is quite similar in spirit. Another example is Peter Jones' proof of the L^2-boundedness of the Cauchy integral on regular curves **[J1]**, and later in his quadratic estimates on the β_∞'s for rectifiable curves **[J3]**. A corona decomposition also arose in **[S4]** for a certain class of hypersurfaces in $\mathbf{R}^n$, in connection with square function estimates for the Cauchy integral from Clifford analysis.

In each of these examples something like a corona decomposition was obtained by applying Carleson's corona construction to a function that somehow controlled the geometry. In our case we cannot apply the corona construction so directly, but we shall use many of the same ideas.

This notion of a corona decomposition is somewhat technical and complicated, but it is very useful. It includes enough control on the geometry of E to imply other things, and it is set up in such a way as to make it amenable to proving it using a stopping-time argument. In fact, there is sort of a universal stopping-time argument for deciding whether E admits a corona decomposition, which is described in Section 7. This argument is universal in the sense that it produces a corona decomposition for E whenever one exists.

Let us give an imprecise outline of this procedure. To get started you need to know that E satisfies the weak geometric lemma. This condition is defined in Section 5; roughly speaking, it means that for most cubes Q, E is well-approximated by a d-plane P_Q. The cubes for which this is not true are put into the bad set, and then you use a stopping-time argument to partition the good cubes into regions S for which (2.5) holds, the angle between P_Q and $P_{Q(S)}$ is small for $Q \in S$, and such that the minimal cubes Q of S either have a bad son or have angle $(P_Q, P_{Q(S)})$ being not too small. If you choose the parameters correctly (2.6) holds, and the hard part is to verify (2.7).

In most of the examples the verification of (2.7) works as follows. Your hypothesis is some sort of square function condition on E, such as (C2) or (C3). The main step is to show that if S has lots of minimal cubes with angle $(P_Q, P_{Q(S)})$ not too small, then there has to be a substantial contribution to the square function condition on E coming from S. To do this you push the contribution from S down to a square function estimate on the Lipschitz graph, and you can usually work with that using classical results. This gives you control on the Lipschitz graph (namely, control on the oscillation of its tangent plane) that permits you to show that S can't have too many minimal cubes of the above type.

It is natural to ask after seeing the definition of a corona decomposition whether the graph of a Lipschitz function A necessarily has one. Of course our main theorem says that it does, but it is not so difficult to prove this directly, by applying the corona construction to, say, the Poisson extension of ∇A. This is similar to the approach taken in [S4], although [S4] applies in more generality, and can certainly be simplified in this case.

There are a number of variations that we can make in the definition of a corona decomposition that would still yield an equivalent condition. For example, we could replace

$$\operatorname{dist}(x,\Gamma) \leq \eta \ \operatorname{diam} Q \quad \text{by} \quad \operatorname{dist}(x,\Gamma) \leq C \operatorname{diam} Q.$$

Because we shall not need this fact we shall content ourselves with merely an outline of the proof.

Suppose that E admits this weakened version of the corona decomposition, and let us show that we can find one of the stronger type. We replace each $S \in \mathcal{F}$ by a subregion $\widetilde{S}$ as follows. We require that if $Q \in \widetilde{S}$ then $R \in S$ for every cube R in the same generation Δ_j as Q which contains a point x such that $\operatorname{dist}(x,Q) \leq \operatorname{diam} Q$. We also require that if $Q \in \widetilde{S}$ then all its subcubes for the next m generations lie in S, where $m \approx \log \frac{1}{\eta}$, and that the cubes R as above also have this property. If we take all the cubes in $S \setminus \widetilde{S}$ and add them to $\mathcal{B}$ for each S, then the resulting augmentation of $\mathcal{B}$ still satisfies (2.4). (This can be checked using (2.3) and (2.7).) It is not difficult to then decompose the $\widetilde{S}$'s into subregions that satisfy (2.5), (2.7), and the stronger version of (2.6). (You have to decompose the $\widetilde{S}$'s because they do not have maximal elements.)

Similarly, in (2.6) we can replace

$$\operatorname{dist}(x,Q) \leq \operatorname{diam} Q \text{ by } \operatorname{dist}(x,Q) \leq k \operatorname{diam} Q$$

for any given $k > 1$. Indeed, if E admits a corona decomposition as above, we can remove the top m layers of each $S \in \mathcal{F}$, $m \approx \log k$, put them into $\mathcal{B}$, and reorganize what's left of each S into new coherent regions.

We also don't really have to require that the Lipschitz graphs Γ in (2.6) have small constant. Indeed, if E admits a corona decomposition where the Γ's merely have uniformly bounded constants, then we can build one where they have small constants by applying the corona decomposition to each of these Γ's to get new ones with small constants, and combining these corona decompositions for the Γ's into one for E.

It is also not hard to show that whether E admits a corona decomposition does not depend on the choice of Δ, as long as (2.1) and (2.2) are satisfied. (Our second variation of the corona decomposition — the one with the k — is helpful in this regard.)

Hopefully these variations give an indication of the flexibility of the notion of a corona decomposition. In practice it is particularly convenient that the only requirement on the bad set is (2.4).

3. From (C1) to (C2)

It is helpful to introduce an intermediate condition.

(C2$'$) *Given any smooth odd function ψ on $\mathbf{R}^n$ with compact support,*

$$\sum_{k=-\infty}^{\infty} \int_E \left| \int_E \psi_k(x-y)f(y)dy \right|^2 dx \leq C \int_E |f|^2 dx.$$

for all $f \in L^2(E)$.

That (C2$'$) implies (C2) is well-known, and is obtained by applying (C2$'$) to characteristic functions of balls. (The converse is not too difficult either; it is essentially a square function version of the $T(1)$ theorem.)

To prove that (C1) implies (C2) we use a familiar artiface. Let Ω denote the space of all sequences $\omega = \{\omega_j\}$, $j \in \mathbf{Z}$, of ± 1's, with the usual product topology and the product measure that gives each choice of ± 1 equal probability. Define $\epsilon_j : \Omega \to \{\pm 1\}$ by $\epsilon_j(\omega) = \omega_j$. As usual we observe that

$$\tag{3.1} \begin{aligned} &\int_\Omega \left| \sum_{j=-m}^{m} \int_E \epsilon_j(\omega)\psi_j(x-y)f(y)dy \right|^2 d\omega \\ &= \sum_{j=-m}^{m} \left| \int_E \psi_j(x-y)f(y)dy \right|^2 . \end{aligned}$$

for any m. This follows from the orthonormality of the ϵ_j's.

We want to apply (C1) to the kernel

$$K(x) = K_m(x,\omega) = \sum_{j=-m}^{m} \epsilon_j(\omega)\psi_j(x)$$

(which is certainly odd and satisfies (1.2)) to conclude that

$$\int_E \Big| \int_E K_m(x-y,\omega) f(y) dy \Big|^2 dx \le C \int_E |f|^2 dx. \tag{3.2}$$

Unfortunately, however, the way we stated (C1) only gives us (3.2) with a C that depends on m and ω. Using a "completeness" argument we shall show that (C1) actually does imply that (3.2) holds with a C that is independent of m and ω. Once we've done that it will follow immediately that (C1) implies (C2), because of (3.1).

Let $N(m,\omega)$ denote the smallest constant for which (3.2) holds for all f. For each $m < \infty$ choose $\omega(m)$ so that

$$N(m,\omega(m)) = \max_\omega N(m,\omega) =: N(m).$$

We want to show that

$$\sup_m N(m) < \infty. \tag{3.3}$$

Suppose not. Choose m_j, $j = 1, 2, 3, \ldots$, such that

$$N(m_j) \ge 2^{m_{j-1}}. \tag{3.4}$$

Define $\omega(\infty) \in \Omega$ by $\omega_i(\infty) = \omega_i(m_j)$ whenever $m_{j-1} < |i| \le m_j$. It is easy to see that

$$N(m_j) = N(m_j,\omega(m_j)) \le 2N(m_j,\omega(\infty)) + Cm_{j-1}. \tag{3.5}$$

This uses only the fact that the number of i's such that $-m_j \le i \le m_j$ and $\omega_i(m_j) \ne \omega_i(\infty)$ is at most $2m_{j-1}+1$.

On the other hand, we have for any w that

$$\int_E \Big| \int_E K_m(x-y,\omega) f(y) dy \Big|^2 dx$$
$$\le C \int_E |f|^2 dx + C \sup_{\alpha>0} \int_E \Big| \int_{E\cap\{|x-y|>\alpha\}} K_\infty(x-y,\omega) f(y) dy \Big|^2 dx,$$

where C does not depend on ω, m, or f. This is not hard to check. Of course (C1) says that the right side is at most $C \int_E |f|^2 dx$, where now C depends on ω but not m or f.

Applying this to $\omega = \omega(\infty)$ and then recombining with (3.5) we find that there is a $C > 0$ not depending on j such that

$$N(m_j) \le C + Cm_{j-1}.$$

For j large enough this is incompatible with (3.4). The ensuing contradiction establishes (3.3).

4. (C2) implies a local symmetry condition

To simplify notations, we'll use from now on the convention that

$$\lambda Q = \{x \in E : \operatorname{dist}(x, Q) \leq (\lambda - 1) \operatorname{diam} Q\} \tag{4.1}$$

for $Q \in \Delta$ and $\lambda > 1$.

Given a small number τ, let $\mathcal{R}(\tau)$ denote the set of cubes $Q \in \Delta$ such that there exists two points $x,\, y \in 2Q$ with $\operatorname{dist}(2x - y, E) \geq \tau \operatorname{diam} Q$. Thus if $Q \notin \mathcal{R}(\tau)$, then for any $x,\, y \in 2Q$, the point $z = 2x - y$ is near E, and z is of course the point on the line through x and y which is opposite to y about x. In other words, if $Q \notin \mathcal{R}(\tau)$, then E is approximately symmetric near Q about each point in Q.

DEFINITION 4.2. *We say that E satisfies the local symmetry condition (LS) if for each $\tau > 0$ the set of cubes $\mathcal{R}(\tau)$ satisfies the packing condition*

$$\sum_{\substack{Q \in \mathcal{R}(\tau) \\ Q \subseteq R}} |Q| \leq C(\tau)|R| \qquad \textit{for all } R \in \Delta. \tag{4.3}$$

In the next section we'll show that if E satisfies (LS), then E is well-approximated by d-planes around most cubes Q.

PROPOSITION 4.4. *If E satisfies (C2), then E satisfies (LS).*

Fix $\tau > 0$. Our strategy for proving (4.3) will be to find a finite family of ψ's so that each cube in $\mathcal{R}(\tau)$ gives a substantial contribution to (1.8) for one of these ψ's. This will allow us to derive (4.3) from (1.8).

Let $Q \in \mathcal{R}(\tau)$ be given, and let $k \in \mathbf{Z}$ be such that $Q \in \Delta_k$. Let x, $y \in 2Q$ be such that $\operatorname{dist}(2x-y, E) \geq \tau \operatorname{diam} Q$. Denote by C_0 the constant in (2.2).

Set $y_0 = 2^{-k}(y-x)$ and $B = B(y_0, \tau/10C_0)$. Let $\psi = \psi_Q$ be an odd C^∞ function such that

$$\text{(4.5)} \qquad \operatorname{supp} \psi \subseteq B \cup (-B), \psi \geq 0 \text{ on } B, \text{ and } \psi \equiv 1 \text{ on } B(y_0, \tau/20C_0).$$

We can even find a finite family Ψ of functions so that for any $Q \in \mathcal{R}(\tau)$ we can take ψ_Q to be an element of Ψ. This is because $y_0 \in B(0, 10C_0) \setminus B(0, \tau/3C_0)$ independently of Q.

Let x' be any point of $B(x, 2^k\tau/40C_0)$. Because of our choice of ψ, $\psi_k(x'-u) \geq 0$, except perhaps when $u \in B(2x-y, 2^k\tau/5C_0)$, and in this case $u \notin E$ because of our assumptions. Also, if $u \in B(y, 2^k\tau/40C_0)$, then $\psi_k(x'-u) = 2^{-kd}$. Using our assumption that E is regular (1.5) we get that

$$\int_E \psi_k(x'-u)du \geq \tau^d/C \qquad \text{for all } x' \in B(x, \frac{2^k\tau}{40C_0}),$$

whence

$$\text{(4.6)} \qquad \int_{3Q} \Big| \int_E \psi_k(x'-u)du \Big|^2 dx' \geq C^{-1}\tau^{3d}2^{kd}.$$

For k and x' given, there are at most a bounded number of cubes Q such that $Q \in \Delta_k$ and $x' \in 3Q$. Hence, for any $R \in \Delta$,

$$\sum_{\substack{Q \in \mathcal{R}(\tau) \\ Q \subseteq R}} |Q| \leq C(\tau) \sum_{\substack{Q \in \mathcal{R}(\tau) \\ Q \subseteq R}} \int_{3Q} \Big| \int_E (\psi_Q)_k (x'-u)du \Big|^2 dx'$$

$$\leq C'(\tau) \sum_{\psi \in \Psi} \sum_{2^k \leq C \operatorname{diam} R} \int_{3R} \Big| \int_E \psi_k(x'-u)du \Big|^2 dx' \leq C''(\tau)|R|.$$

The last inequality uses (1.8). This proves the proposition.

REMARK 4.7: Our local symmetry condition implies a slightly stronger version of itself. For each $\tau > 0$ and each $k > 1$, let $\mathcal{R}(\tau, k)$ denote the set of cubes Q such that there exist x, $y \in kQ$ such that $\operatorname{dist}(2x-y, E) \geq \tau \operatorname{diam} Q$. If E satisfies our local symmetry condition, then

$$\text{(4.8)} \qquad \sum_{\substack{Q \in \mathcal{R}(\tau,k) \\ Q \subseteq R}} |Q| \leq C(\tau, k)|R|$$

for all $R \in \Delta$ and all τ, k. The proof is not difficult. (If $Q \in \mathcal{R}(\tau, k)$, then there is a not-too-distant ancestor of Q in $\mathcal{R}(\tau')$, if τ' is chosen suitably.)

5. The local symmetry condition (LS) implies the weak geometric lemma

For $Q \in \Delta$ and $1 \leq q \leq \infty$, define

$$\beta_q(Q) = \inf_P \left(\frac{1}{|Q|} \int_{2Q} \left(\frac{\operatorname{dist}(x, P)}{\operatorname{diam} Q} \right)^q dx \right)^{\frac{1}{q}}, \tag{5.1}$$

where the infimum is taken over all d-planes P. This is of course a minor variation of (1.9).

We say that E satisfies the weak geometric lemma if for each $\epsilon > 0$ we have

$$\sum_{\substack{\beta_\infty(Q) > \epsilon \\ Q \subseteq R}} |Q| \leq C(\epsilon)|R| \qquad \text{for all } R \in \Delta. \tag{5.2}$$

The name stems from the practice of saying that E satisfies the geometric lemma (of Peter Jones) if

$$\sum_{Q \subseteq R} \beta_\infty(Q)^2 |Q| \leq C|R| \qquad \text{for all } R \in \Delta. \tag{5.3}$$

Note that if E satisfies our version of the geometric lemma — i.e., (C3) — then it satisfies the weak geometric lemma. This can be readily derived from the fact that

$$\beta_\infty(x, t) \leq C \beta_1(x, 2t)^{\frac{1}{d+1}}. \tag{5.4}$$

To prove (5.4), let P be the d-plane that realizes the infimum in the definition of $\beta_1(x,2t)$, and let y be the point in $B(x,t)\cap E$ furthest from P. Set $D=\operatorname{dist}(y,P)$. If $D\le t$, then

$$\begin{aligned}\beta_1(x,2t) &\ge t^{-d-1}\int_{E\cap B(y,D/2)} \operatorname{dist}(x,P)\,dx \ge t^{-d-1}\mid E\cap B(y,D/2)\mid D\\ &\ge C^{-1}(t^{-1}D)^{d+1}\ge C^{-1}\beta_\infty(x,t)^{d+1}.\end{aligned}$$

If $D\ge t$, then $\beta_1(x,2t)\ge C$, and there's nothing to prove.

PROPOSITION 5.5. *Suppose that E satisfies (LS). For each $\epsilon>0$ let $\mathcal{G}(\epsilon)$ denote the set of cubes $Q\in\Delta$ such that there is a d-plane P_Q with the following two properties:*

$$\operatorname{dist}(x,P_Q)\le\epsilon\operatorname{diam}Q \qquad \textit{for all } x\in 2Q; \tag{5.6}$$

$$\textit{if } w\in P_Q \textit{ and } \operatorname{dist}(w,Q)\le\operatorname{diam}Q, \textit{ then } \operatorname{dist}(w,E)\le\epsilon\operatorname{diam}Q. \tag{5.7}$$

Let $\mathcal{B}(\epsilon)=\Delta\setminus\mathcal{G}(\epsilon)$ denote the complement of $\mathcal{G}(\epsilon)$. Then

$$\sum_{\substack{Q\in\mathcal{B}(\epsilon)\\ Q\subseteq R}}|Q|\le C(\epsilon)|R|$$

for all $R\in\Delta$ and all $\epsilon>0$.

Thus the conclusion tells us that E is well-approximated by d-planes in a stronger sense than the weak geometric lemma. Notice that the conclusions of Proposition 5.5 imply (LS).

To prove the proposition we shall first prove two lemmata. We shall also use these lemmata to prove that neither (C1) nor (C2) can hold if d is not an integer, and so for their statements and proofs we allow d to be noninteger.

LEMMA 5.8. *Fix $Q\in\Delta$, and let (d) denote the smallest integer greater than or equal to d. Then there exist $(d)+1$ points $y_0,\dots,y_{(d)}$ in Q such that $\operatorname{dist}(y_j,L_{j-1})\ge A^{-1}\operatorname{diam}Q$ for $j=1,\dots,(d)$, where L_k denotes the*

k-plane passing through $y_0, \ldots, y_k$, and where A depends only on d and the constant in (1.5).

This is obtained using an easy induction argument. Suppose we have found points y_0, $y_1, \ldots, y_j$ as above for some $j < (d)$. Suppose that there does not exist a suitable point y_{j+1}, so that

$$\operatorname{dist}(y, L_j) \leq A^{-1} \operatorname{diam} Q \qquad \text{for all } y \in Q.$$

Then it is not hard to see that we can cover Q by less than $(A+1)^j$ balls of radius $\frac{10n}{A} \operatorname{diam} Q$, say. From (1.5) we then get that

$$|Q| \leq C(A+1)^j \left(\frac{10n}{A} \operatorname{diam} Q\right)^d.$$

This is impossible if A is large enough, because $j < (d)$, and so y_{j+1} exists. This proves the lemma.

LEMMA 5.9. *Let M be a large integer, and set $k = 4(d+1)M$. Let $\mathcal{R}(\tau, k)$ be as in Remark 4.7, and assume that $Q \notin \mathcal{R}(\tau, k)$. If τ is small enough (depending on M), then given an integer ℓ, $1 \leq \ell \leq (d)+1$, $\ell+1$ points $y_0, \ldots, y_\ell$ of $2Q$, and $a_1, \ldots, a_\ell \in 2^\ell \mathbf{Z} \cap [-M, M]$, there is a point $z \in E$ such that*

$$\left| z - \{y_0 + \sum_{i=1}^{\ell} a_i(y_i - y_0)\} \right| \leq C(M, \ell)\tau \operatorname{diam} Q.$$

We prove this by induction on ℓ, beginning with $\ell = 1$. For notational convenience we take $y_0 = 0$.

The point of the proof is of course to use repeatedly the fact that $Q \notin \mathcal{R}(\tau, k)$. Taking $x = y_1$ and $y = 0$ in the definition of $\mathcal{R}(\tau, k)$ we see that there is a point $z_2 \in E$ such that $|z_2 - 2y_1| \leq \tau \operatorname{diam} Q$. Since $z_2 \in kQ$ we can take $x = z_2$ and $y = y_1$ to get a point $z_3 \in E$ such that $|z_3 - 2z_2 + y_1| \leq \tau \operatorname{diam} Q$, so that $|z_3 - 3y_1| \leq 3\tau \operatorname{diam} Q$. Repeating this argument we see that for $2 \leq j \leq M$ there is a $z_j \in E$ such that $|z_j - jy_1| \leq (2^{j-1} - 1)\tau \operatorname{diam} Q$. Taking $x = 0$ and $y = z_j$ in the definition of $\mathcal{R}(\tau, k)$ gives a point $z_{-j} \in E$ with $|z_{-j} + jy_1| \leq 2^{j-1}\tau \operatorname{diam} Q$. That takes care of the $\ell = 1$ case, with $C(M, 1) = 2^{M-1}$.

Of course, we could also have done the same thing with y_1 replaced with any of the other y_j's.

Assume now that the lemma holds for some ℓ, and let us prove it for $\ell+1$ (if $\ell+1 \leq (d)+1$). Let $a_1, \ldots, a_{\ell+1} \in 2^{\ell+1}\mathbf{Z} \cap [-M, M]$ be given. By induction hypothesis we can find w_1, $w_2 \in E$ so that

$$| w_1 - \sum_{j=1}^{\ell} \frac{1}{2} a_j y_j | \leq C(M, \ell)\tau \operatorname{diam} Q \tag{5.10}$$

$$| w_2 + a_{\ell+1} y_{\ell+1} | \leq 2^{M-1} \tau \operatorname{diam} Q \tag{5.11}$$

If τ is small enough, w_1 and w_2 both lie in kQ, we can find a point $z \in E$ such that $|z - 2w_1 + w_2| \leq \tau \operatorname{diam} Q$, so that from (5.10) and (5.11) we get

$$| z - \sum_{j=1}^{\ell+1} a_j y_j | \leq C(M, \ell+1)\tau \operatorname{diam} Q,$$

where $C(M, \ell+1) = 2C(M, \ell) + 2^{M-1} + 1$.

Let us now prove Proposition 5.5. For this we require d to be an integer again. Let $\epsilon > 0$ be given. Let M be large, to be chosen soon, depending on ϵ, and put $k = 4(d+1)M$ again. We shall choose τ after M, depending on both M and ϵ.

Fix $Q_0 \in \Delta$. Choose $Q \subseteq Q_0$ so that $\operatorname{diam} Q \sim \epsilon \operatorname{diam} Q_0$, and assume that $Q \notin \mathcal{R}(\tau, k)$. Let $y_0, \ldots, y_d$ be as in Lemma 5.8, and take $P = P_{Q_0}$ to be the d-plane that they span.

If τ is small enough, then (5.7) holds for Q_0. Indeed, every point on P which is at distance $\leq \operatorname{diam} Q_0$ from Q_0 is at distance $\leq \operatorname{diam} Q$ from a point of the form

$$y_0 + \sum_{j=1}^{d} a_j(y_j - y_0), \text{ with } a_j \in 2^d\mathbf{Z}, |a_j| \leq C\epsilon^{-1}.$$

If M is large enough ($M \geq C\epsilon^{-1}$), we can apply Lemma 5.9 to conclude that any such point is at distance $\leq C(M, d)\tau \operatorname{diam} Q$ from E. If τ is small enough, we see that every point of P which is at distance $\leq \operatorname{diam} Q_0$ from Q_0 is at distance $\leq C\epsilon \operatorname{diam} Q_0$ from E, so that (5.7) holds for Q_0 with ϵ replaced by $C\epsilon$.

Let us now show that (5.6) holds for Q_0 if we assume also that $Q_0 \notin \mathcal{R}(\tau, k)$. Suppose not; let $y_{d+1} \in 2Q_0$ be such that $\operatorname{dist}(y_{d+1}, P) \geq \epsilon \operatorname{diam} Q_0$.

Apply Lemma 5.9 to Q_0 and the points y_j; $1 \le j \le d+1$, to conclude that for each $a = (a_1, \dots, a_{d+1})$, $a_i \in 2^{d+1}\mathbf{Z} \cap [-M, M]$, there is a $z_a \in E$ so that if $x_a = y_0 + \sum_{i=1}^{d+1} a_i(y_i - y_0)$, then

$$| z_a - x_a | \le C(M, d+1)\tau \operatorname{diam} Q_0.$$

We want to use this to contradict the assumption that E is regular (1.5).

If $a \ne a'$, then $|x_a - x_{a'}| \ge \frac{1}{C}\epsilon \operatorname{diam} Q_0$. This is easily checked using the fact that for each $j \le d+1$, y_j is at distance $\ge \frac{1}{C}\epsilon \operatorname{diam} Q_0$ from the $(j-1)$-plane passing through $y_0, \dots, y_{j-1}$. By taking τ to be small enough, depending on M and ϵ, we get that

$$| z_a - z_{a'} | \ge \frac{1}{2C}\epsilon \operatorname{diam} Q_0 \quad \text{for } a \ne a'.$$

Thus the balls $B_a = B(z_a, \epsilon \operatorname{diam} Q_0/4C)$ are pairwise disjoint. From (1.5) we get

$$| kQ_0 | \ge \sum_a | B_a \cap E | \ge C^{-1}M^{d+1}(\epsilon \operatorname{diam} Q_0)^d$$

and also $|kQ_0| \le CM^d(\operatorname{diam} Q_0)^d$. This is impossible if we choose M to be much larger than ϵ^{-d}. This contradiction tells us that y_{d+1} does not exist, and so (5.6) holds for Q_0.

Thus we have proved that $Q_0 \in \mathcal{G}(C\epsilon)$ if $Q_0 \notin \mathcal{R}(\tau, k)$, if the associated cube $Q \notin \mathcal{R}(\tau, k)$, and if τ, k are chosen correctly. Proposition 5.5 now follows from Remark 4.7.

Let us now indicate why E cannot satisfy (C1) or (C2) if its dimension d is not an integer. It is enough to show that E cannot satisfy (LS). Suppose it did. Let Q_0 be a cube with $Q_0 \notin \mathcal{R}(\tau, k)$. We use an argument very similar to the one we just did, but with $\epsilon = 1$, $Q = Q_0$.

Use Lemma 5.8 to select points $y_0, y_1, \dots, y_{(d)}$ in Q_0. Given $a = (a_1, \dots, a_{(d)})$, $a_i \in 2^{(d)}\mathbf{Z} \cap [-M, M]$, let x_a, z_a, and B_a be as above (but with $\epsilon = 1$). If τ is small enough (depending on M), we obtain once again that the B_a's are disjoint, and hence

$$| kQ_0 | \ge \sum_a | B_a \cap E | \ge C^{-1}M^{(d)}(\operatorname{diam} Q_0)^d,$$

while $|kQ_0| \le CM^d(\operatorname{diam} Q_0)^d$. This is impossible if M is large enough and $(d) \ne d$.

REMARK 5.12: The weak geometric lemma is not strong enough to imply rectifiability. We shall give a counterexample in Section 20. However, a modification of an argument of Peter Jones [**J2**] shows that E has big pieces of Lipschitz graphs if it is regular, satisfies the weak geometric lemma, and has big projections. This last condition means that there is a $\theta > 0$ so that for each $x \in E$ and $R > 0$ there is a d-plane P so that $|\Pi(E \cap B(x, R))| \geq \theta R^d$, where Π is the orthogonal projection onto P. See [**DS3**] for more details.

Although the weak geometric lemma is not strong enough by itself to imply rectifiability, it seems to be a very useful intermediary condition. It certainly plays that role in proving that (C2) and (C3) imply (C4).

It turns out that (LS) is equivalent to (C1)–(C7). We hope that this is as big a surprise for the reader as it was for the authors. The proof will be given in a separate publication. When $d = 1$ there is a direct construction that shows that (LS) implies (C7). This combines with the arguments given here to provide a much simpler proof of the fact that (C1) or (C2) imply (C7) when $d = 1$. The argument for showing that (LS) implies (C1)–(C7) when $d > 1$ is much less direct and it relies in particular on the fact that (C1) implies the other conditions. Thus it does not enable us to dispense with what we are doing here and in the succeeding sections.

We conclude this section with an easy lemma which says that the good d-plane P_Q in Proposition 5.5 is almost unique.

LEMMA 5.13. *Let $Q \in \Delta$ be given, and suppose that P_1 and P_2 are two d-planes such that* $\operatorname{dist}(x, P_i) \leq \epsilon \operatorname{diam} Q$ *for all $x \in Q$, $i = 1, 2$. Then*

$$\operatorname{dist}(w, P_2) \leq C\epsilon \operatorname{diam} Q + C\epsilon \operatorname{dist}(w, Q) \qquad \text{for all } w \in P_1 \tag{5.14}$$

$$\operatorname{dist}(w, P_1) \leq C\epsilon \operatorname{diam} Q + C\epsilon \operatorname{dist}(w, Q) \qquad \text{for all } w \in P_2. \tag{5.15}$$

In particular, Angle $(P_1, P_2) \leq C\epsilon$.

Let $y_0, \ldots, y_d$ be the points in Q provided by Lemma 5.8, and let P denote the d-plane passing through them. It is not hard to prove (5.14) and (5.15) by comparing each of P_1 and P_2 to P using the fact that $\operatorname{dist}(y_j, P_i) \leq \epsilon \operatorname{diam} Q$ for $0 \leq j \leq d$, $j = 1, 2$. [An important point is that this last fact not only implies that all the points of P are close to P_i, but that the points of P_i are also close to P. To see this it is useful to notice that if $z_{j,i}$ are points in P_i close to y_j, then the $z_{j,i}$ generate P_i.]

6. Approximation of E in measure

From Propositions 4.4 and 5.5 we know that if E satisfies (C2) then most cubes can be well-approximated by d-planes in the sense of (5.6) and (5.7). This result can be strengthened, in that the Hausdorff measure on E will be well-approximated by a constant times Lebesgue measure on the d-plane. This is what we prove now.

Let k be a large constant. We denote by $\mathcal{G}(\epsilon, k)$ the set of cubes $Q \in \Delta$ such that there is a d-plane P_Q with the properties

$$\text{(6.1)} \qquad \operatorname{dist}(x, P_Q) \leq \epsilon \operatorname{diam} Q \qquad \text{for all } x \in kQ$$

$$\text{(6.2)} \quad \text{if } w \in P_Q \text{ and } \operatorname{dist}(w, Q) \leq k \operatorname{diam} Q, \text{ then } \operatorname{dist}(w, E) \leq \epsilon \operatorname{diam} Q.$$

If E satisfies the conclusion of Proposition 5.5, then $\Delta \backslash \mathcal{G}(\epsilon, k)$ satisfies the usual Carleson measure packing condition for all ϵ, k. This follows from the observation that if $Q \notin \mathcal{G}(\epsilon, k)$ and if Q^* is an ancestor of Q such that $k \operatorname{diam} Q \leq \operatorname{diam} Q^* \leq Ck \operatorname{diam} Q$, then $Q^* \notin \mathcal{G}(\epsilon/Ck)$, where $\mathcal{G}(\delta)$ is as in the statement of Proposition 5.5.

Given $Q \in \Delta$ and a d-plane P, define a measure $\mu_{Q,P}$ on P by

$$\text{(6.3)} \qquad \mu_{Q,P}(A) = | \Pi^{-1}(A) \cap (k/2)Q |$$

for all (Borel sets) $A \subseteq P$, where Π denotes the orthogonal projection onto P. In fancier language, $\mu_{Q,P}$ is obtained by taking the restriction of Hausdorff measure on E to $(k/2)Q$ and pushing it down to P using Π.

Let us call $\mathcal{H}(\epsilon, k)$ the subset of $\mathcal{G}(\epsilon, k)$ consisting of those cubes Q such that for each d-plane P with Angle $(P, P_Q) \leq \frac{1}{10}$ there is a constant $\lambda_{Q,P} > 0$ that satisfies

$$\text{(6.4)} \qquad \big| \, \mu_{Q,P}(A) - \lambda_{Q,P}|A| \, \big| \leq \epsilon^d (\operatorname{diam} Q)^d$$

for every cube $A \subseteq P$ with

$$\epsilon \operatorname{diam} Q \leq \operatorname{diam} A \leq \operatorname{diam} Q, \operatorname{dist}(A, \Pi(Q)) \leq \frac{k}{10} \operatorname{diam} Q.$$

PROPOSITION 6.5. *Suppose that E satisfies (C2). Then for every $\epsilon > 0$ and $k \geq 10$ there is a $C = C(\epsilon, k)$ such that*

$$\sum_{\substack{Q \notin \mathcal{H}(\epsilon,k) \\ Q \subseteq R}} |Q| \leq C\,|R| \qquad \text{for all } R \in \Delta. \tag{6.6}$$

By an easy covering argument one sees that (6.4) holds if

$$\big|\, \mu_{Q,P}(A) - \lambda_{Q,P}|A| \,\big| \leq \tau(\operatorname{diam} Q)^d \tag{6.7}$$

for every cube $A \subseteq P$ such that $\operatorname{dist}(A, \Pi(Q)) \leq (k/5) \operatorname{diam} Q$, A has sidelength exactly $\eta \operatorname{diam} Q$, and has sides parallel to a given (fixed) set of axes, provided that we choose η small enough with respect to ϵ and then choose τ small enough. To verify (6.7) it is certainly enough to show that

$$\big|\, \mu_{Q,P}(A) - \mu_{Q,P}(A') \,\big| \leq \tau(\operatorname{diam} Q)^d/2 \tag{6.8}$$

whenever A, A' are two cubes in P that satisfy the properties just listed and also $\operatorname{dist}(A, A') \geq \operatorname{diam} Q$.

Let us prove the proposition. Let τ_1 be much smaller than τ, to be specified later. We must show that most cubes lie in $\mathcal{H}(\epsilon, k)$, in the sense of (6.6); it is enough to show that most cubes in $\mathcal{G}(\tau_1, k)$ lie in $\mathcal{H}(\epsilon, k)$, by the remarks at the beginning of this section.

Fix $Q \in \mathcal{G}(\tau_1, k)$ and a d-plane P with Angle $(P, P_Q) \leq \frac{1}{10}$. We want to find conditions that imply (6.8) and such that the cubes that don't satisfy these conditions satisfy a Carleson measure packing condition. Of course we want these conditions to be given in terms of something controlled by (C2). The argument we use is similar to the proof of Proposition 4.4.

Let A, A' be given. Let p be the center of A and q the point of P_Q such that $\Pi(q) = p$, and define p' and q' similarly. By the version of (6.1) in this context, $\Pi^{-1}(A) \cap (k/2)Q$ is contained in the n-dimensional cube D centered at q whose projection is A. A similar result is true for A' and the corresponding n-cube D'. From the version of (6.2) in this context we get that there is a point $x_0 \in E$ at distance $\leq \tau_1 \operatorname{diam} Q$ from the midpoint between q and q'.

Let $j \in \mathbf{Z}$ be such that $Q \in \Delta_j$, so that $\operatorname{diam} Q \sim 2^j$. We want to choose a function ψ so that $2^{jd}\psi_j(x_0 - y)$ looks a lot like the characteristic function of D minus the characteristic function of D'. More precisely, we ask ψ to be odd, C^∞, compactly supported, and we want it to satisfy

(6.9a) $\psi_j(x_0 - y) = 2^{-jd}$ if $\operatorname{dist}(y, D) \leq C\tau_1 \operatorname{diam} Q$

(6.9b) $0 \leq \psi_j(x_0 - y) \leq 2^{-d}$ if $\operatorname{dist}(y, D) \leq 2C\tau_1 \operatorname{diam} Q$

(6.9c) $\psi_j(x_0 - y) = -2^{-jd}$ if $\operatorname{dist}(y, D') \leq C\tau_1 \operatorname{diam} Q$

(6.9d) $-2^{-jd} \leq \psi_j(x_0 - y) \leq 0$ if $\operatorname{dist}(y, D') \leq 2C\tau_1 \operatorname{diam} Q$

(6.9e) $\psi_j(x_0 - y) = 0$ if $\operatorname{dist}(y, D \cup D') \geq 2C\tau_1 \operatorname{diam} Q$.

We can even find a (fixed) finite family Φ of functions so that we can find a $\psi \in \Phi$ with these properties no matter what Q, P, A, and A' are, subject to the constraints imposed above. (We allow Φ to depend on τ_1 and all the other constants.)

An argument like the one used in the proof of Proposition 4.4 can be used to show that if E satisfies (C2), then $|\int_E \psi_j(x_0 - y)dy|$ is as small as we want except for a class $\mathcal{C}$ of Q's that satisfies a Carleson measure packing condition. (To be precise, $|\int_E \psi_j(x_0 - y)dy|$ should be as small as we want for all x_0 that arise from admissible choices of P, A, A'.) Because of (6.9), we get that if $Q \notin \mathcal{C}$, then

$$|E \cap D| \leq |E \cap G'| + \frac{1}{4}\tau(\operatorname{diam} Q)^d \tag{6.10}$$

$$|E \cap D'| \leq |E \cap G| + \frac{1}{4}\tau(\operatorname{diam} Q)^d, \tag{6.11}$$

where G is the cube concentric with D but whose sidelength has been increased by $4C\tau_1 \operatorname{diam} Q$, and similarly for G'.

From chasing definitions we see that (6.8) holds if (6.10), (6.11), and

$$|E \cap (G \setminus F)| + |E \cap (G' \setminus F')| \leq \frac{1}{4}\tau(\operatorname{diam} Q)^d \tag{6.12}$$

do. It is not hard to see that (6.12) is true if $Q \in \mathcal{G}(\tau, k)$ and τ_1 is small enough: because

$$E \cap (G \setminus F) \subseteq \{z : \operatorname{dist}(z, P_Q \cap (G \setminus F)) \leq 2\tau_1(\operatorname{diam} Q\},$$

we can cover $E \cap (G \setminus F)$ by less than $C\tau_1^{-(d-1)}$ balls of radius $3\tau_1 \operatorname{diam} Q$, and then use the regularity of E (1.5) to prove (6.12).

To summarize, we have shown that if $Q \in \mathcal{G}(\tau, k)$ and $Q \notin \mathcal{C}$, where $\mathcal{C}$ satisfies a Carleson measure packing condition, then (6.10), (6.11), and (6.12) hold, and they imply (6.8), and hence that $Q \in \mathcal{H}(\epsilon, k)$. This completes the proof of the proposition.

REMARK 6.13: In this proof, as in many places, we do not use the fact that we are dealing with d-dimensional Hausdorff measure on E; the same argument would work if we used any positive Borel measure supported on E and satisfying (1.5). (Of course such a measure must be comparable to Hausdorff measure.)

REMARK 6.14: The constants $\lambda_{P,Q}$ in (6.4) always satisfy $C^{-1} \leq \lambda_{P,Q} \leq C$. This is easily seen by applying (6.4) to a cube of diameter $\leq C \operatorname{diam} Q$ that contains $\Pi(Q)$.

7. Building the stopping-time regions, and some of their properties

In proving that (C2) or (C3) imply (C4) we shall use stopping-time arguments, and these arguments will be very similar to each other and have a substantial overlap. In this section we give a construction of the stopping-time regions that will be used in both cases.

Let E be given. We assume that E satisfies the weak geometric lemma. Suppose that we are given ϵ, δ, $0 < \epsilon < \delta$, both of which are as small as we want, and with δ/ϵ large. Let $k > 0$ be large, to be chosen later (but not depending on ϵ or δ). Suppose also that we are given a decomposition $\Delta = \mathcal{B} \cup \mathcal{G}$, where $\mathcal{B}$ satisfies (2.4), and for each cube $Q \in \mathcal{G}$ there is a d-plane P_Q such that (6.1) holds.

The reader should keep in mind that for us the only important property that $\mathcal{B}$ satisfies is (2.4). Thus we do not mind adding cubes to $\mathcal{B}$ as long as (2.4) is preserved.

LEMMA 7.1. *Under the preceeding assumptions we can find a new decomposition $\Delta = \mathcal{B}' \cup \mathcal{G}'$, where $\mathcal{G}' \subseteq \mathcal{G}$ and $\mathcal{B}'$ still satisfies (2.4), and where we can partition $\mathcal{G}'$ into a family $\mathcal{F}$ of stopping-time regions S such that each S satisfies (2.5) and also:*

(7.2) *if $Q \in S$, then Angle $(P_Q, P_{Q(S)}) \leq \delta$;*

(7.3) *if Q is a minimal cube of S, then at least one of the children of Q lies in $\mathcal{B}'$, or else Angle $(P_Q, P_{Q(S)}) \geq \delta/2$.*

Before proving this — which is not difficult — let us make a few remarks about how this fits into the big picture. We want to prove eventually that if E satisfies (C2) or (C3), then it admits a corona decomposition. To do this it is enough to show that for all sufficiently small ϵ, δ, the stopping-time regions S provided by the lemma also satisfy (2.6) and (2.7). We shall see in the next section that (2.6) is always true under the circumstances of the lemma, while (2.7) is harder to prove. We shall give some preliminary reductions in that direction after proving the lemma.

Let us prove the lemma. We start with a slightly simpler version of it. Given $R_0 \in \Delta$, let $\mathcal{G}(R_0)$ denote the subset of $\mathcal{G}$ of cubes contained in R_0. We first show that we can partition $\mathcal{G}(R_0)$ into a family $\mathcal{F}(R_0)$ of stopping-time regions with the above properties.

This is easy, because there is pretty much only one way to do it. Let Q_0 be an element of $\mathcal{G}(R_0)$ of maximal size (i.e., $Q_0 \in \Delta_j$ for j as large as possible). It is easy to see that there is a subset S of $\mathcal{G}(R_0)$ that has Q_0 as its maximal element and which satisfies (2.5), (7.2), and (7.3). (S can be built using the obvious stopping-time argument.) Remove S from $\mathcal{G}(R_0)$ and repeat the process: pick an element of $\mathcal{G}(R_0)$ with maximal size, and then build the associated stopping time region. Repeating this we get our partition $\mathcal{F}(R_0)$ of $\mathcal{G}(R_0)$.

For many purposes this localized version of the lemma is adequate, but it is not hard to prove the more global version either. To do this we need a sequence $\{R_j\}$ of cubes which are pairwise disjoint, whose union is all of E, and which have the property that for each ℓ there are at most C cubes in Δ_ℓ not contained in any of the R_j's. Once we have this sequence of cubes we set $\mathcal{G}' = \cup \mathcal{G}(R_j)$, $\mathcal{B}' = \Delta \setminus \mathcal{G}'$, and $\mathcal{F} = \cup \mathcal{F}(R_j)$, and it is not hard to check that these choices satisfy the conclusions of the lemma.

Let us indicate how to find such a sequence $\{R_j\}$. If $E = \mathbf{R}^d$, it is easy to write down such a sequence of dyadic cubes explicitly, and the general construction is in a similar spirit.

Fix a point $p_0 \in E$. For each $k \geq 0$ consider the set of cubes in Δ_k which intersect $B(p_0, 2^k)$ or which have a brother that intersects it. (Two cubes in Δ_k are called brothers if they have the same father.) If we now take the union over $k \geq 0$ of the cubes so selected, we get a sequence of cubes which have the desired properties except for being pairwise disjoint. The minimal elements (with respect to inclusion) of this sequence gives a new sequence having all the desired features.

That completes the proof of Lemma 7.1. We now give a preliminary reduction that is useful for checking (2.7).

Give $S \in \mathcal{F}$, let $m(S)$ denote the minimal cubes of S. Let $m_0(S)$ denote the set of minimal cubes of S which have at least one child in $\mathcal{B}'$, and let $m_1(S)$ denote the $Q \in m(S)$ with Angle $(P_Q, P_{Q(S)}) \geq \delta/2$. Thus $m(S) = m_0(S) \cup m_1(S)$, by (7.3).

Define $\mathcal{F}_0$, $\mathcal{F}_1$, and $\mathcal{F}_2$ as follows:

$$\mathcal{F}_0 = \{S \in \mathcal{F} : \Big| \bigcup_{Q \in m_0(S)} Q \Big| \geq |Q(S)|/4\}$$

$$\mathcal{F}_1 = \{S \in \mathcal{F} : \Big| \bigcup_{Q \in m_1(S)} Q \Big| \geq |Q(S)|/2\}$$

$$\mathcal{F}_2 = \{S \in \mathcal{F} : \Big| Q(S) \setminus \Big(\bigcup_{Q \in m(S)} Q \Big) \Big| \geq |Q(S)|/4\}$$

Clearly, then, $\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_1 \cup \mathcal{F}_2$.

LEMMA 7.4. *For $i = 0, 2$,*

$$\sum_{\substack{S \in \mathcal{F}_i \\ Q(S) \subseteq R}} |Q(S)| \leq C|R| \qquad \textit{for all } R \in \Delta. \tag{7.5}$$

Thus (2.7) holds if we can prove (7.5) for $i = 1$.

The case $i = 2$ easily follows from the fact that the sets $Q(S) \setminus (\bigcup_{Q \in m(S)} Q)$ are pairwise disjoint in E, which is itself a consequence of the pairwise disjointness of the S's and (2.5). The case $i = 0$ can be derived without difficulty from the requirement that $\mathcal{B}'$ satisfy (2.4).

Let us say a few words more about how we'll show that (2.7) holds when E satisfies (C2). Similar ideas will be used for (C3).

For each $S \in \mathcal{F}$ we define a function $d(x)$ on $\mathbf{R}^n$ by

$$d(x) = \inf_{Q \in S} \{\operatorname{dist}(x, Q) + \operatorname{diam} Q\}. \tag{7.6}$$

This function encodes a lot of information about S, e.g., where its minimal cubes are very small. Also define an associated "summing region" $\sigma = \sigma(S) \subseteq E \times \mathbf{Z}$ by

$$\begin{aligned} \sigma = \{(x, \ell) \in E \times \mathbf{Z} :\ & x \in k_0 Q(S) \quad \text{and} \\ & \gamma_0 d(x) \leq 2^\ell \leq \operatorname{diam} Q(S)\}, \end{aligned} \tag{7.7}$$

where the constants k_0 (large) and γ_0 (small) will be chosen later. (They will not depend on ϵ, δ. Also, k will be chosen after k_0, and will be much larger.) One should think of σ as an enlarged and smeared-up version of $\{(x,\ell) : x \in Q \text{ for some } Q \in S \cap \Delta_\ell\}$. Given an odd smooth function ψ with compact support, set

$$J(S,\psi) = \int \sum_{(x,\ell)\in\sigma} \Big| \int_E \psi_\ell(x-y)dy \Big|^2 dx. \tag{7.8}$$

LEMMA 7.9. *Suppose that we are given k, ϵ, δ, k_0, and γ_0, and let $\mathcal{F}$ be as in Lemma 7.1. To prove that* (2.7) *holds if E satisfies* (C2), *it suffices to find a finite family Ψ of ψ's such that for some $\tau > 0$*

$$\sum_{\psi\in\Psi} J(S,\psi) \geq \tau|Q(S)| \quad \textit{whenever } S \in \mathcal{F}_1. \tag{7.10}$$

Here τ, Ψ are allowed to depend on all the constants above, but not on S.

This follows from Lemma 7.4 once we show that for any ψ,

$$\sum_{Q(S)\subseteq R} J(S,\psi) \ \leq \ C|R| \quad \text{for all } R \in \Delta.$$

This inequality is an immediate consequence of (C2) and the fact that $\sigma(S)$, $S \in \mathcal{F}$, have bounded overlap in $E \times \mathbf{Z}$. This last fact comes from the disjointness of the S's and a little definition-chasing: if $(x,\ell) \in \sigma(S)$, then there is a $Q \in S$ such that

$$\operatorname{dist}(x,Q) \leq \gamma_0^{-1}2^\ell, 2^\ell \leq \operatorname{diam} Q \leq \gamma_0^{-1}2^\ell.$$

For any given (x,ℓ), there are only a bounded number of cubes with these two properties, and so there are only a bounded number of S's with $(x,\ell) \in \sigma(S)$.

Thus to prove that (C2) implies (2.7) we want to show that if k_0 is large enough and γ_0, ϵ, δ, and ϵ/δ are small enough, then we can find Ψ such that (7.10) holds. The proof of this is complicated but the basic idea is fairly simple. We first show that there is a Lipschitz function A (on some d-plane) whose graph approximates E very well on the scale of $d(x)$, with errors on the order of ϵ. We use this approximation to push down estimates on J from E to the graph of A, proving eventually that the left side of (7.10) controls a square function applied to A, modulo terms that are small compared to δ, and that this square function controls the L^2 mean oscillation of ∇A. In other words, if the left side of (7.10) is very small, we show that ∇A must be almost constant, in a way that is incompatible with $S \in \mathcal{F}_1$. (If $S \in \mathcal{F}_1$, then the graph of A cannot be too flat, by definition of $\mathcal{F}_1$ and $m_1(S)$.)

The proof that (2.7) holds if (C3) does will be very similar in structure, the main difference being that we have to work with a different kind of square function.

REMARK 7.11: As we mentioned in Section 2, this kind of procedure will always produce a corona decomposition for E if there is one. That is, if E admits a corona decomposition, then it satisfies the weak geometric lemma, and we can apply Lemma 7.1 (with $\mathcal{G}$ taken to be exactly the set of Q's for which there is a P_Q such that (6.1) holds) to obtain the good regions $S \in \mathcal{F}$. The results of the next section imply that (2.6) holds, while Lemma 7.4 tells us that we need only check (7.5) for $i = 1$ to prove (2.7). This one can do using the assumption that E admits a corona decomposition, if the parameters are chosen correctly.

[The reason for this last assertion is that if $Q \in m_1(S)$, and if E has a corona decomposition with η chosen small enough (depending on δ), then it is not hard to show that there must be a cube R such that $Q \subseteq R \subseteq Q(S)$, and such that R is either a bad cube or a minimal or maximal cube for a stopping-time region associated to the corona decomposition of E with constant η.]

8. The construction of the approximating Lipschitz graph

Throughout this section we use the same assumptions and notations as in Section 7. Fix $S \in \mathcal{F}$ as in Lemma 7.1, and set $P = P_{Q(S)}$. Let $P^\perp$ be an $(n-d)$-plane orthogonal to P, and let Π and $\Pi^\perp$ denote the orthogonal projections onto P and $P^\perp$. We shall often identify P with $\mathbf{R}^d$, and in particular we equip P with dyadic cubes. We denote by L the diameter of $Q(S)$.

In addition to the function $d(x)$ defined in (7.6), we shall also use the function D defined on P by

$$(8.1) \qquad D(p) = \inf_{x \in \Pi^{-1}(P)} d(x) = \inf_{Q \in S} \{\operatorname{dist}(p, \Pi(Q)) + \operatorname{diam} Q\}.$$

PROPOSITION 8.2. *There is a Lipschitz function $A : P \to P^\perp$ with norm $\leq C\delta$ such that*

$$(8.3) \qquad \operatorname{dist}(x, (\Pi(x), A(\Pi(x)))) \leq C\epsilon\, d(x)$$

for all $x \in k_0 Q(S)$.

Thus the points of $k_0 Q(S)$ are close to the graph of A. As with Lemma 7.1, we only need to know that E satisfies the weak geometric lemma for this proposition.

It will be very important that we have an ϵ in (8.3) instead of a δ. When we later try to push square function estimates from E down to the graph of A, we will need to know that the errors are small compared to δ, and (8.3) is one of the reasons why.

Set $Z = \{z \in E : d(z) = 0\}$, so that $D(p) = 0$ iff $p \in \Pi(Z)$. We first define A on $\Pi(Z)$. To do this, we have no choice: we must prove that Π is 1-1 on Z, and that its inverse is Lipschitz. We shall even have use for the following more general result.

LEMMA 8.4. *If $x, y \in 10k_0Q(S)$ satisfy $|x-y| \geq 10^{-3}\min(d(x), d(y))$, then*

$$|\,\Pi^{\perp}(x) - \Pi^{\perp}(y)\,| \leq 2\delta\,|\,\Pi(x) - \Pi(y)\,|\,.$$

Assume that $|x-y| \geq 10^{-3}d(x)$. Let $Q \in S$ be such that

$$\mathrm{dist}(x, Q) + \mathrm{diam}\, Q \leq C|x-y|.$$

We can replace Q by one of its ancestors, if necessary, to get $\mathrm{diam}\, Q \sim |x-y|$. By our assumptions in Section 7, there is a d-plane P_Q for which (6.1) holds, and so

$$\mathrm{dist}(x, P_Q) + \mathrm{dist}(y, P_Q) \leq C\epsilon|x-y| \ll \delta|x-y|.$$

The lemma now follows from Angle $(P, P_Q) \leq \delta$.

From the lemma we see that

$$A(\Pi(z)) = \Pi^{\perp}(z) \quad \text{for } z \in Z \tag{8.5}$$

defines a 2δ-Lipschitz function on $\Pi(Z)$. To define A on the rest of P we'll use arguments from the proof of the Whitney extension theorem (see [**St**]), and in particular a variation of the Whitney decomposition of $P \setminus \Pi(Z)$.

For each $x \in P$ with $D(x) > 0$ and x not on the boundary of a dyadic cube, let R_x be the largest dyadic cube in P containing x and satisfying

$$\mathrm{diam}\, R_x \leq 20^{-1} \inf_{u \in R_x} D(u). \tag{8.6}$$

Let R_i, $i \in I$ be a relabelling of the set of all these cubes R_x without repetition. Thus the R_i's are pairwise disjoint, they cover $P \setminus \Pi(Z)$, and they do not intersect $\Pi(Z)$. [Here we use the convention that dyadic cubes are closed but are called disjoint if their interiors are disjoint.]

LEMMA 8.7. *If $10R_i \cap 10R_j \neq \emptyset$, then*

$$C^{-1}\,\mathrm{diam}\, R_j \leq \mathrm{diam}\, R_i \leq C\,\mathrm{diam}\, R_j.$$

It is clearly sufficient to check that

$$10\,\mathrm{diam}\, R_i \leq D(y) \leq 60\,\mathrm{diam}\, R_i \qquad \text{for all } y \in 10R_i. \tag{8.8}$$

Because D is Lipschitz with norm 1,

$$D(y) \geq \min_{u \in R_i} D(u) - 10 \operatorname{diam} R_i \geq 10 \operatorname{diam} R_i.$$

For the second inequality in (8.8) we use the fact that the father R of R_i fails (8.6). Thus there is a $z \in R$ such that $D(z) < 20 \operatorname{diam} R = 40 \operatorname{diam} R_i$, whence

$$D(y) \leq D(z) + 20 \operatorname{diam} R_i \leq 60 \operatorname{diam} R_i.$$

Let us proceed now to the construction of A on the ball $U_0 = P \cap B(\Pi(x_0), 2k_0 L)$, where x_0 is any fixed point of $Q(S)$, and L is still $\operatorname{diam} Q(S)$. For future use we set

$$U_j = P \cap B(\Pi(x_0), 2^{1-j} k_0 L) \qquad \text{for all } j \in \mathbf{N}. \tag{8.9}$$

We also restrict ourselves to the set I_0 of $i \in I$ for which R_i meets U_0. Given $i \in I_0$, let us choose a cube $Q(i) \in S$ such that

$$\begin{aligned} &C^{-1} \operatorname{diam} R_i \leq \operatorname{diam} Q(i) \leq C \operatorname{diam} R_i \quad \text{and} \\ &\operatorname{dist}(\Pi(Q(i)), R_i) \leq C \operatorname{diam} R_i. \end{aligned} \tag{8.10}$$

The existence of such a $Q(i)$ is not a problem. If p is any point of R_i, there is a cube $Q \in S$ such that

$$\operatorname{dist}(p, \Pi(Q)) + \operatorname{diam} Q \leq 2D(p) \leq 120 \operatorname{diam} R_i,$$

by definition of $D(p)$. We then take $Q(i)$ to be a suitable ancestor of Q (possibly even Q itself).

Note that a single cube Q may correspond in this way to more than one (but not too many) R_i.

Let B_i denote the affine function from P to $P^\perp$ whose graph is the d-plane $P_{Q(i)}$. Because of (7.2), the Lipschitz norm of B_i is $\leq 2\delta$.

For each i let $\tilde{\phi}_i$ be a C^2 bump function such that

$$\begin{aligned} &0 \leq \tilde{\phi}_i \leq 1, \tilde{\phi}_i = 1 \text{ on } 2R_i, \tilde{\phi}_i = 0 \text{ off } 3R_i, \text{ and} \\ &| \nabla^\ell \tilde{\phi}_i | \leq C (\operatorname{diam} R_i)^{-\ell} \quad \text{for } \ell = 1, 2. \end{aligned} \tag{8.11}$$

Because of Lemma 8.7, there are, for each i, at most C cubes R_j with $3R_i \cap 3R_j \neq \emptyset$; in particular the supports of the $\tilde{\phi}_i$ have bounded overlap. We can define a partition of unity for $V = \bigcup_{i \in I_0} 2R_i$ by

$$\phi_i(p) = \tilde{\phi}_i(p) \left\{ \sum_{j \in I_0} \tilde{\phi}_j(p) \right\}^{-1} \qquad \text{for } p \in V, \ i \in I_0. \tag{8.12}$$

Using Lemma 8.7 again we have that

$$(8.13)\qquad |\nabla^\ell \phi_i| \leq C(\operatorname{diam} R_i)^{-\ell}, \quad \ell = 1, 2.$$

We define A on V by

$$(8.14)\qquad A(p) = \sum_{i \in I_0} \phi_i(p) B_i(p).$$

Notice that $V \cap \Pi(Z) = \emptyset$ (by 8.8), and of course $U_0 \setminus \Pi(Z) \subseteq V$, so that (8.5) and (8.14) combined define A on U_0. Let us prove that A is Lipschitz with norm $C\delta$ on U_0.

We first check that the restriction of A to $2R_j$, $j \in I_0$, is 3δ-Lipschitz. Given p, $q \in 2R_j$ we have

$$(8.15)$$
$$\begin{aligned} |A(p) - A(q)| &\leq \Big|\sum_i \phi_i(p)\{B_i(p) - B_i(q)\}\Big| + \Big|\sum_i \{\phi_i(q) - \phi_i(p)\} B_i(q)\Big| \\ &\leq 2\delta\, |p - q|\, \{\sum_i \phi_i(p)\} + \Big|\sum_i \{\phi_i(p) - \phi_i(q)\}\{B_i(q) - B_j(q)\}\Big|. \end{aligned}$$

(In the last step we used the fact that $\sum_i (\phi_i(p) - \phi_i(q)) = 0$.) If $\phi_i(p) \neq 0$ or $\phi_i(q) \neq 0$, Lemma 8.7 gives $\operatorname{diam} R_i \sim \operatorname{diam} R_j$, and then we get from (8.13) that

$$(8.16)\qquad |\phi_i(p) - \phi_i(q)| \leq C (\operatorname{diam} R_j)^{-1} |p - q|.$$

To estimate $B_i(q) - B_j(q)$ we use the following lemma.

LEMMA 8.17. *If* $10R_i \cap 10R_j \neq \emptyset$, *then* $\operatorname{dist}(Q(i), Q(j)) \leq C \operatorname{diam} R_j$ *and*

$$(8.18)\qquad |B_i(q) - B_j(q)| \leq C\epsilon \operatorname{diam} R_j \quad \textit{for all } q \in 100R_j$$

To prove the first part, pick any $x \in Q(j)$ and $y \in Q(i)$. We may safely assume that $|x - y| \geq \frac{1}{3} \operatorname{diam} Q(j)$, and since $d(x) \leq \operatorname{diam} Q(j)$ by definition, we can apply Lemma 8.4 to obtain

$$|\Pi^\perp(x) - \Pi^\perp(y)| \leq |\Pi(x) - \Pi(y)|.$$

Because $|\Pi(x) - \Pi(y)| \leq C \operatorname{diam} R_j$ by (8.10) (and Lemma 8.7), we conclude that

$$\operatorname{dist}(Q(i), Q(j)) \leq |x - y| \leq C \operatorname{diam} R_j.$$

This implies that $Q(j) \subseteq kQ(i)$ (if k is large enough), and (8.18) now follows at once from Lemma 5.13 (with $Q = Q(j)$, $P_1 = P_{Q(j)}$, $P_2 = P_{Q(i)}$).

Combining (8.15), (8.16), and (8.18) we get

$$\begin{aligned} |A(p) - A(q)| &\leq 2\delta|p-q| + C(\operatorname{diam} R_j)^{-1}|p-q|\epsilon \operatorname{diam} R_j \\ &\leq 3\delta|p-q| \qquad \text{for } p, q \in 2R_j, \end{aligned} \tag{8.19}$$

if ϵ/δ is small enough. We used the fact that there are a bounded number of i's for which $\phi_i(p) - \phi_i(q) \neq 0$ for fixed p, q.

Next, let us show that

$$|A(p) - A(p_0)| \leq C\delta|p - p_0| \quad \text{if } p_0 \in \Pi(Z),\ p \in \bigcup_{j \in I_0} R_j. \tag{8.20}$$

Choose j so that $p \in R_j$, and pick $y \in Q(j)$. Thus we have

$$\begin{aligned} |A(p) - A(p_0)| \leq a_1 + a_2 + a_3 + a_4, \text{ where } & a_1 = |A(p) - B_j(p)|, \\ & a_2 = |B_j(p) - B_j(\Pi(y))|, a_3 = |B_j(\Pi(y)) - \Pi^{\perp}(y)|, \text{ and} \\ & a_4 = |\Pi^{\perp}(y) - A(p_0)|. \end{aligned}$$

Notice that $D(p) \leq |p - p_0|$, since $D(p_0) = 0$, and so $\operatorname{diam} R_j \leq |p - p_0|$. From the definition of A and Lemma 8.17 we get

$$a_1 \leq C\epsilon \operatorname{diam} R_j \leq C\epsilon|p - p_0|.$$

Next, $a_2 \leq 2\delta|p - \Pi(y)| \leq C\delta \operatorname{diam} R_j \leq C\delta|p - p_0|$, because of the 2δ-Lipschitzness of B_j and (8.10). The definition of B_j gives

$$a_3 \leq 2\operatorname{dist}(y, P_{Q(j)}) \leq 2\epsilon \operatorname{diam} Q(j) \leq C\epsilon|p - p_0|.$$

To estimate a_4 we apply Lemma 8.4 to $x = (p_0, A(p_0)) \in Z$ and y to get

$$a_4 = |\Pi^{\perp}(y) - \Pi^{\perp}(x)| \leq 2\delta|\Pi(y) - p_0| \leq C\delta|p - p_0|.$$

(Here we have used (8.10) for the last inequality.) Combining these various estimates gives (8.20).

Combining the fact that A is 2δ-Lipschitz on $\Pi(Z)$ with (8.19) and (8.20) it is easy to see that A is Lipschitz with norm $\leq C\delta$ on U_0. We can use the Whitney extension theorem to extend A from U_0 to a Lipschitz function on all of P with norm $\leq C\delta$.

We now turn to the proof of (8.3). If $x \in Z$, then x is on the graph of A, and there is nothing to prove. For the remaining case we'll use the following lemma, which will also be needed later.

LEMMA 8.21. *Let $p \in U_0$ and $r > 0$, $D(p) \le r \le k_0 L$, be given, and let $Q \in S$ be such that* $\operatorname{dist}(p, \Pi(Q)) \le Cr$ *and* $C^{-1} r \le \operatorname{diam} Q \le Cr$. *Then* $\Pi^{-1}(B(p,r)) \cap 2k_0 Q(S)$ *is contained in* $C_0 Q$, *where* C_0 *depends only on* C *and* k_0.

As a consequence,

$$\widetilde{C}_0^{-1} d(x) \le D(\Pi(x)) \le d(x) \quad \textit{for all } x \in k_0 Q(S),$$

where $\widetilde{C}_0$ *depends only on* k_0.

To prove the lemma, pick $x \in Q$ and let y be any point in $\Pi^{-1}(B(p,r)) \cap 2k_0 Q(S)$. If $|x - y| \le \operatorname{diam} Q$, then $y \in C_0 Q$, as promised. Otherwise, we can use the fact that $d(x) \le \operatorname{diam} Q$ to apply Lemma 8.4 to get

$$|\Pi^{\perp}(x) - \Pi^{\perp}(y)| \le |\Pi(x) - \Pi(y)| \le C \operatorname{diam} Q.$$

(The last inequality comes from our hypotheses concerning p, Q, and r.)

The second affirmation of the lemma is obvious when $\Pi(x) \in \Pi(Z)$; apply Lemma 8.4, for instance, to see that $\Pi^{-1}(\Pi(x)) \cap k_0 Q(S)$ contains only one point. Otherwise, if $\Pi(x) \notin \Pi(Z)$, take $p = \Pi(x)$ and $r = D(p)$. By definition of $D(p)$ there is a Q as above. The first part of the lemma tells us that $x \in C_0 Q$, whence $d(x) \le \widetilde{C}_0 D(\Pi(x))$. The other inequality follows from the definitions.

Coming back to (8.3), let $x \in k_0 Q(S)$ be such that $d(x) > 0$, and set $p = \Pi(x)$. The lemma tells us that $D(p) > 0$, and so p lies in some R_i. Applying the lemma with $r = D(p)$ and $Q = Q(i)$, we get that $x \in C_0 Q(i)$.

We'll choose k to be much larger than C_0, and so (6.1) gives

$$|\Pi^{\perp}(x) - B_i(\Pi(x))| \le 2\epsilon \operatorname{diam} Q(i) \le C\epsilon D(p) \le C\epsilon d(x),$$

and since $|B_i(\Pi(x)) - A(\Pi(x))| \le C\epsilon D(p)$, by Lemma 8.17 and the definition of A, we get $|\Pi^{\perp}(x) - A(\Pi(x))| \le C\epsilon d(x)$, as desired.

This completes the proof of Proposition 8.2. We end this section with one more estimate on A.

LEMMA 8.22. $|\nabla^2 A(u)| \le C\epsilon (\operatorname{diam} R_j)^{-1}$ *if* $u \in 2R_j$.

Indeed, if $\partial_\alpha \partial_\beta$ is any second partial derivative,

$$\begin{aligned}\partial_\alpha \partial_\beta A &= \partial_\alpha \partial_\beta \left(\sum \phi_i B_i\right) \\ &= \sum (\partial_\alpha \partial_\beta \phi_i) B_i + \sum (\partial_\alpha \phi_i)(\partial_\beta B_i) + \sum (\partial_\beta \phi_i)(\partial_\alpha B_i).\end{aligned}$$

(Since B_i is affine $\partial_\alpha \partial_\beta B_i = 0$.) Because $\sum \partial_\alpha \phi_i = \partial_\alpha(\sum \phi_i) = 0$, we have

$$\begin{aligned}\partial_\alpha \partial_\beta A = \sum \partial_\alpha \partial_\beta \phi_i (B_i - B_j) + \sum \partial_\alpha \phi_i (\partial_\beta B_i - \partial_\beta B_j) \\ + \sum \partial_\beta \phi_i (\partial_\alpha B_i - \partial_\alpha B_j).\end{aligned}$$

It is not hard to obtain the desired estimate from (8.13) and Lemma 8.17. (We also use the fact that $|\nabla B_i - \nabla B_j| \leq C\epsilon$ if $u \in \operatorname{supp}(\nabla \phi_i)$; this can be derived from Lemma 8.17, or proved using a very similar argument.)

9. Pushing square function estimates from E to the graph of A

In this section, and in the following two sections as well, we assume that E satisfies (C2), and we want to prove that E must then satisfy (C4). We follow the outline given at the end of Section 7, and we make the same assumptions and notations as in Sections 7 and 8.

To implement the program described in Section 7 we must first specify the subsets $\mathcal{G}$ and $\mathcal{B}$ of Δ. We take $\mathcal{G} = \mathcal{H}(\epsilon, k)$, as defined in Section 6, and $\mathcal{B} = \Delta \setminus \mathcal{G}$. As we said in Section 7, k is large and will be chosen later, but it will not depend on ϵ, δ.

Let $J = J(S, \psi)$ be as in (7.8), with ψ fixed. One of the things we have to do to carry out the program described in Section 7 is to transform an estimate on J into an estimate for a suitable square function applied to A, and it is this issue that we begin to take up now.

Let μ denote the measure on P obtained by pushing down $H^d \,|_{k_0 Q(S)}$ using Π, i.e.,

$$\mu(F) = |\Pi^{-1}(F) \cap k_0 Q(S)|.$$

Define $I = I(S, \psi)$ by

$$I = \int \sum_{(p,\ell)\in\sigma_1} \Big| \int_P \psi_\ell((p, A(p)) - (q, A(q)))\, d\mu(q) \Big|^2 dp, \tag{9.1}$$

where the "summing region" $\sigma_1 \subseteq P \times \mathbf{Z}$ is defined by

$$\sigma_1 = \{(p, \ell) \in P \times \mathbf{Z} : \; p \in U_4, \; C_1 \gamma_0 D(p) \leq 2^\ell \leq L\}, \tag{9.2}$$

U_4 is as in (8.9), L is still $\operatorname{diam} Q(S)$, $C_1 > 1$ will be chosen soon (large, and independent of ϵ, δ, and γ_0), and dp denotes Lebesgue measure on P.

PROPOSITION 9.3. *There is a constant $C \geq 0$, independent of δ and ϵ, such that*

$$I \leq CJ + C\epsilon^2|Q(S)|,$$

at least if we assume that supp $\psi \subseteq B(0, \frac{1}{10}k_0)$.

Set $a_\ell(x) = \int_E \psi_\ell(x-y)dy$ and

$$b_\ell(x) = \int_E \psi_\ell((\Pi(x),\ A(\Pi(x))) - (\Pi(y), A(\Pi(y))))dy.$$

We first want to replace a_ℓ by b_ℓ in the integral that defines J, with only small errors. Actually, we shall even ask for less than that, namely,

$$J_1 \leq 2J + C\epsilon^2|Q(S)|, \tag{9.4}$$

where

$$J_1 = \int \sum_{(x,\ell)\in\sigma_2} |b_\ell(x)|^2 dx, \tag{9.5}$$

$$\sigma_2 = \{(x,\ell) \in E \times \mathbf{Z} :\ x \in \frac{1}{2}k_0 Q(S),\ \gamma_0 d(x) \leq 2^\ell \leq L\}. \tag{9.6}$$

To see this, we first observe that for $(x,\ell) \in \sigma_2$ we have

$$|a_\ell(x) - b_\ell(x)| \leq C\|\nabla\psi\|_\infty \int_{E\cap B(x,\frac{1}{10}k_0 2^\ell)} \epsilon(d(x)+d(y))2^{-\ell}2^{-\ell d}dy.$$

If $(x,\ell) \in \sigma_2$ and y is in $E \cap B(x, \frac{1}{10}k_0 2^\ell)$, then $d(x) \leq \gamma_0^{-1}2^\ell$ and $d(y) \leq d(x) + \frac{1}{10}k_0 2^\ell \leq C2^\ell$. Thus we also have $|a_\ell(x) - b_\ell(x)| \leq C\epsilon$, whence

$$|a_\ell(x) - b_\ell(x)|^2 \leq C\epsilon^2 \int_{E\cap B(x,\frac{1}{10}k_0 2^\ell)} (d(x)+d(y))2^{-\ell}2^{-\ell d}dy.$$

Consequently,

$$\begin{aligned} J_1 &\leq 2J + 2\int \sum_{(x,\ell)\in\sigma_2} |b_\ell(x) - a_\ell(x)|^2 dx \\ &\leq 2J + C\epsilon^2 \int \sum_{(x,\ell)\in\sigma_2} \int_{E\cap B(x,\frac{1}{10}k_0 2^\ell)} (d(x)+d(y))2^{-\ell}2^{-\ell d}dydx \\ &\leq 2J + C\epsilon^2 \int \sum_{(x,\ell)\in\sigma_2} 2^{-\ell}d(x)dx + C\epsilon^2 \int_{y\in k_0Q(S)} \sum_{2^\ell \geq C^{-1}d(y)} 2^{-\ell}d(y)dy \\ &\leq 2J + C\epsilon^2 \int_{\frac{1}{2}k_0Q(S)} dx + C\epsilon^2 \int_{k_0Q(S)} dy \leq 2J + C\epsilon^2|Q(S)|, \end{aligned}$$

as desired.

Next we want to push our integrals down to P. Because the points y that appear in the definition of $b_\ell(x)$ always lie in $k_0Q(S)$ when $(x,\ell) \in \sigma_2$ we have that

$$b_\ell(x) = \int_P \psi_\ell((p, A(p)) - (q, A(q)))d\mu(q), \tag{9.7}$$

where $p = \Pi(x)$. Let us check that

$$J_2 = \int \sum_{(p,\ell)\in\sigma_3} |c_\ell(p)|^2 d\mu(p) \leq J_1, \tag{9.8}$$

where $c_\ell(p) = \int_P \psi_\ell((p, A(p)) - (q, A(q)))d\mu(q)$ and

$$\sigma_3 = \{(p,\ell) \in P \times \mathbf{Z} : \ p \in U_3, \ \frac{1}{2}C_1\gamma_0 D(p) \leq 2^\ell \leq L\}. \tag{9.9}$$

(Thus σ_3 is a little larger than the region σ_1 used to define I.)

Because of (9.7), we only have to check that if $(p,\ell) \in \sigma_3$ and $x \in \Pi^{-1}(p) \cap k_0Q(S)$ then $(x,\ell) \in \sigma_2$. Using (6.1) with $Q = Q(S)$ we see that such an x must lie in $\frac{1}{2}k_0Q(S)$. Also, if $\frac{1}{2}C_1$ is larger than the constant $\widetilde{C}_0$ from Lemma 8.21, then $d(x) \leq \frac{1}{2}C_1D(p) \leq \gamma_0^{-1}2^\ell$, and so $(x,\ell) \in \sigma_2$. This establishes (9.8).

To finish the proof of Proposition 9.3 we have to be able to replace the measure $d\mu(p)$ in J_2 by dp. Notice that $d\mu(p) \geq dp$ on $\Pi(Z)$, whence

$$\int_{\Pi(Z)} \sum_{2^\ell \leq L} |c_\ell(p)|^2 dp \leq J_2. \tag{9.10}$$

To control the rest we restrict our attention to each R_i for which $R_i \cap U_4 \neq \emptyset$ and compare $d\mu(p)$ to dp on R_i. Let $I_4 = \{i : \ R_i \cap U_4 \neq \emptyset\}$.

LEMMA 9.11. *If $i \in I_4$ and T is a cube satisfying $T \subseteq 10R_i$ and $\operatorname{diam} T \geq M\epsilon \operatorname{diam} R_i$ with M large enough (not depending on ϵ or δ), then*

$$\frac{1}{C}\int_T dp \leq \int_T d\mu(p) \leq C\int_T dp.$$

Here C does not depend on ϵ or δ.

To prove the lemma we first apply Lemma 8.21 to any $p \in R_i$ with $r = 60 \operatorname{diam} R_i$ and $Q = Q(i)$. We get that $\Pi^{-1}(10R_i) \cap k_0 Q(S)$ is contained in $C_0 Q(i)$ for some C_0. The restriction of μ to $10R_i$ is therefore the same as the restriction to $10R_i$ of the measure $\mu_{Q(i),P}$ introduced in Section 6 (see (6.3)), at least if $k > C_0$.

We chose our good set of cubes $\mathcal{G}$ to be $\mathcal{H}(\epsilon, k)$, so that (6.4) holds for each $Q \in \mathcal{G}$, for $Q = Q(i)$ in particular. We also pointed out in Remark 6.14 that the constant $\lambda_i = \lambda_{Q(i),P}$ satisfies $C^{-1} \leq \lambda_i \leq C$. Lemma 9.11 follows immediately from (6.4) and definition-chasing.

We want to use the lemma to control I in terms of J_2. Partition R_i into dyadic cubes $T_{i,j}$ such that

$$M\epsilon \operatorname{diam} R_i \;\leq\; \operatorname{diam} T_{i,j} \;\leq\; 2M\epsilon \operatorname{diam} R_i,$$

where M is as in the lemma. Let m denote the minimum of $|c_\ell(p)|$ on $T_{i,j}$. Straightforward estimates of the oscillation of $c_\ell(p)$ over $T_{i,j}$ yield

$$\begin{aligned}
\int_{T_{i,j}} |\, c_\ell(p) \,|^2 \, dp &\leq \int_{T_{i,j}} \left[m + C\epsilon \operatorname{diam} R_i 2^{-\ell}\right]^2 dp \\
&\leq Cm^2 \int_{T_{i,j}} d\mu(p) + C\epsilon^2 (\operatorname{diam} R_i)^2 2^{-2\ell} \int_{T_{i,j}} dp \\
&\leq C \int_{T_{i,j}} |\, c_\ell(p) \,|^2 \, d\mu(p) + C\epsilon^2 (\operatorname{diam} R_i)^2 2^{-2\ell} \int_{T_{i,j}} dp.
\end{aligned}$$

Summing this in j we get

$$\int_{R_i} |\, c_\ell(p) \,|^2 \, dp \leq C \int_{R_i} |\, c_\ell(p) \,|^2 \, d\mu(p) + C\epsilon^2 (\operatorname{diam} R_i)^2 2^{-2\ell} \int_{R_i} dp.$$

Before summing this in i we need some notation and an observation. Let $\mathcal{L}(i)$ denote the set of ℓ's such that $(p, \ell) \in \sigma_1$ for some $p \in R_i$. If $\ell \in \mathcal{L}(i)$, then $2^\ell \geq C_1 \gamma_0 D(p) \geq \frac{1}{2} C_1 \gamma_0 D(p')$ for any other $p' \in R_i$, and

$(p', \ell) \in \sigma_3$. Hence

$$\begin{aligned}
\int \sum_{\substack{(p,\ell)\in\sigma_1 \\ p\in P\setminus\Pi(Z)}} | c_\ell(p) |^2 \, dp &\leq \sum_{i\in I_4} \sum_{\ell\in\mathcal{L}(i)} \int_{R_i} | c_\ell(p) |^2 \, dp \\
&\leq C \sum_{i\in I_4} \sum_{\ell\in\mathcal{L}(i)} \int_{R_i} | c_\ell(p) |^2 \, d\mu(p) + C\epsilon^2 \sum_{i\in I_4} \sum_{\ell\in\mathcal{L}(i)} 2^{-2\ell}(\operatorname{diam} R_i)^2 \int_{R_i} dp \\
&\leq C \int \sum_{(p,\ell)\in\sigma_3} | c_\ell(p) |^2 \, d\mu(p) + C\epsilon^2 \sum_{i\in I_4} \int_{R_i} dp \\
&\leq C J_2 + C\epsilon^2 |Q(S)|.
\end{aligned}$$

(For the second to last inequality we used (8.8).)

This last estimate, combined with (9.10), (9.8), and (9.4), finishes the proof of Proposition 9.3.

10. Controlling a square function of A in terms of $J(S, \psi)$

According to the program outlined at the end of Section 7, we want to show that for suitable choices of ϵ, δ, etc., we can find a finite family Ψ of ψ's so that (7.10) holds. To do this we want to show that we can control square functions of A in terms of $J(S, \psi)$, modulo certain types of errors. Proposition 9.3 was a first step in this direction, but we need something better. The problem is with I in (9.1), in particular its nonlinear dependence on A and the appearance of $d\mu(q)$ instead of dq. In this section we show that if Ψ is chosen properly we can indeed control a more useful square function of A in terms of $\sum\limits_{\psi \in \Psi} J(\psi, S)$.

We continue to use the same assumptions and notations as in the preceeding 3 sections. Let us now choose the class Ψ that we shall work with.

Let P_0 and $P_0^\perp$ denote the translates of P and $P^\perp$ that pass through the origin. Pick $\nu \in C_c^\infty(P_0)$ which is radial, not identically zero, supported in $B(0, \frac{1}{20})$, zero on a neighborhood of the origin, and satisfies

$$\int_{P_0} \nu(p) f(p) dp = 0$$

for all polynomials f of degree ≤ 2.

We shall sometimes commit the following minor abuse of notation. Given a function on P_0, such as ν, and a function on P, such as A, we note by $\nu * A$ the function on P defined by

$$\int_P \nu(p - q) A(q) dq.$$

To simplify notations we allow our ψ's to be vector-valued. Write the generic element of $\mathbf{R}^n$ as (p,w), $p \in P_0$, $w \in P_0^\perp$. For our first ψ we take any odd, C^∞ function with compact support and values in $P_0^\perp$, and which satisfies

$$\psi(p,w) = \nu(p)w \qquad \text{for all } (p,w) \text{ such that } |w| \le |p|. \tag{10.1}$$

This function will give us control on A, but we also need a function to give us control on the measure $d\mu$. For our second ψ we take any function with values in P which is odd, C^∞, supported in $B(0,1)$, and satisfies

$$\psi(p,w) = \nu(p)p \qquad \text{for all } (p,w) \text{ with } |w| \le |p|. \tag{10.2}$$

We can make these choices in such a way that the family of all ψ's that arise is finite, by making sure that the set of all P_0's that arise is finite. This we can do, in a way that depends on ϵ. (We could also use a slightly different approach in which the family of ψ's is much smaller. This would complicate further the notations and presentation of our argument, but it would not present any serious problems.)

We take for our family Ψ the set of ψ's just described (in (10.1) and (10.2)) as well as some of their dilates. That is, we also take the functions $s^{jd}\psi(s^j x)$, ψ as above, where $s = 2^{\frac{1}{m_0}}$, $j = 1, 2, \ldots, m_0 - 1$, where m_0 is a large integer to be chosen. (It will not depend on E or any of the other constants; just n, d, and the choices of ψ's above.) We denote these functions by $\psi^j(x)$, and we take $\psi_\ell^j(x) = 2^{-\ell d}\psi^j\left(\frac{x}{2^\ell}\right)$ as always. Define ν^j, ν_ℓ^j similarly.

As we said before, we want to control a square function of A in terms $J(S,\psi^j)$, $\psi^j \in \Psi$. Set

$$I_1^j = I_1^j(S) = \int \sum_{(p,\ell)\in\sigma_4} 2^{-2\ell} \left| \int_P \nu_\ell^j(p-q)A(q)dq \right|^2 dp, \tag{10.3}$$

where

$$\begin{aligned}\sigma_4 = \{(p,\ell) \in P \times \mathbf{Z} : \operatorname{dist}(p, \Pi(Q(S)) \le 100L \quad &\text{and} \\ \gamma_0^{\frac{1}{2}} D(p) \le 2^\ell \le \theta L\}.&\end{aligned} \tag{10.4}$$

Here $\theta > 0$ is a small constant to be chosen later (before ϵ).

PROPOSITION 10.5. *Assume that*

$$J(S, \psi^j) \leq \epsilon^2 |Q(S)| \tag{10.6}$$

for all $\psi^j \in \Psi$. *If* ϵ, δ, γ_0, θ, *and* k_0^{-1} *are small enough (how small* ϵ *has to be will depend on the other constants), then for* $j = 0, \dots, m_0 - 1$,

$$I_1^j \leq \left[C\epsilon^2 + C'(\gamma_0 + \theta^2)\delta^2\right] |Q(S)|, \tag{10.7}$$

where C, C' *do not depend on* ϵ, δ, *or* θ, *and* C' *does not depend on* γ_0.

We should perhaps point out that this proposition also depends on m_0 being sufficiently large, but this is not important because our choice of m_0 will not depend on the other constants.

We shall prove Proposition 10.5 in this section, and use it to show that (7.10) holds (with $\tau = \epsilon^2$) in the next section.

Suppose that (10.6) holds, so that Proposition 9.3 can be applied to conclude that

$$I = I(S, \psi^j) \leq C\epsilon^2 |Q(S)| \tag{10.8}$$

for each $\psi^j \in \Psi$. Let us write down explicitly what that means. For ψ as in (10.1), (10.8) becomes

$$\int \sum_{(p,\ell)\in\sigma_1} \left| \int_P s^j 2^{-\ell} \nu_\ell^j (p-q)(A(p) - A(q)) d\mu(q) \right|^2 dp \leq C\epsilon^2 |Q(S)|, \tag{10.9}$$

while for ψ as in (10.2) we get

$$\int \sum_{(p,\ell)\in\sigma_1} \left| \tilde{\nu}_\ell^j * \mu(p) \right|^2 dp \leq C\epsilon^2 |Q(S)|, \tag{10.10}$$

where $\tilde{\nu}(p) = \nu(p)p$. In both cases j runs from 0 to $m_0 - 1$.

To derive (10.7) from (10.9) we want to replace $d\mu(q)$ in (10.9) by dq with only acceptable errors incurred. To do this we shall use (10.10) to show that $d\mu$ can be approximated by dq. We shall also need the following lemma to help control the errors.

LEMMA 10.11. *Let r, M be given, $1 \le r < \frac{2d}{d-2}$ ($1 \le r \le \infty$ if $d = 1$), $M > 0$. For each $p \in P$ and $\ell \in \mathbf{Z}$ there is an affine function $A_{p,\ell} : P \to P^{\perp}$ such that*

$$\sum_{\ell} \left\{ 2^{-\ell d} \int\limits_{P \cap B(p, M2^{\ell})} \left| 2^{-\ell} [A(q) - A_{p,\ell}(q)] \right|^{r} dq \right\}^{\frac{2}{r}} dp \, d\delta_{2^{\ell}}(t) \tag{10.12}$$

is a Carleson measure on $P \times \mathbf{R}_{+}$ with norm $\le C(M) \|\nabla A\|_{\infty}^{2}$ (which is of course $\le C(M)\delta^{2}$).

See [**Do**] for a proof (or [**J1**] for $d = 1$). In [**Do**] the case of $\nabla A \in L^{\infty}$ is not discussed, but this lemma follows immediately from the results for $\nabla A \in L^{2}$.

Notice that the lemma is still true if we also require that $A_{p,\ell}$ be independent of p for p inside a dyadic cube with sidelength 2^{ℓ}. This is not hard to derive from the lemma. [One way to do this is to observe that if $M \ge 2\sqrt{d}$, Q is any dyadic cube in P with sidelength 2^{ℓ}, and if α is an affine function so that $|Q|^{-1} \int_{Q} |A - \alpha|$ is as small as possible, then for all $p \in Q$ we have

$$\sup_{Q} |A_{p,\ell} - \alpha| \le C 2^{-\ell d} \int\limits_{P \cap B(p, M2^{\ell})} |A - A_{p,\ell}|. \tag{10.13}$$

Thus we can replace $A_{p,\ell}$ by α without changing much.] From now on we assume that $A_{p,\ell}$ has this extra property.

There are two easy consequences of this observation that we shall use. The first is that

$$\sum_{\ell} \left| 2^{-\ell} [A(p) - A_{p,\ell}(p)] \right|^{2} dp \, d\delta_{2^{\ell}}(t) \tag{10.14}$$

is a Carleson measure on $P \times \mathbf{R}_{+}$. (Take $r = 2$ in (10.12) and use Fubini.) The second is that $A_{p,\ell}$ satisfies

$$\left| \nabla_{q} A_{p,\ell}(q) \right| \le C\delta, \ |A(p) - A_{p,\ell}(p)| \le C\delta 2^{\ell}. \tag{10.15}$$

We shall now use (10.10) to replace A in (10.9) by $A - A_{p,\ell}$. Afterwards we use Lemma 10.11 to control the errors that arise when we replace $d\mu$ by something more convenient.

LEMMA 10.16. *For* $j = 0, 1, \ldots, m_0 - 1$,

$$\int \sum_{(p,\ell)\in\sigma_1} \left| \int_P s^j 2^{-\ell} \nu_\ell^j(p-q) \left[A_{p,\ell}(p) - A_{p,\ell}(q)\right] d\mu(q) \right|^2 dp \le C\epsilon^2 |Q(S)|.$$

Indeed, because $A_{p,\ell}(q)$ is an affine function,

$$\int_P s^j 2^{-\ell} \nu_\ell^j(p-q) \left(A_{p,\ell}(p) - A_{p,\ell}(q)\right) d\mu(q)$$

is, for each p, ℓ, a linear combination of integrals of the form

$$\int_P s^j 2^{-\ell} \nu_\ell^j(p-q)(p_i - q_i) d\mu(q),$$

with coefficients that are dominated by $C\delta$ (because of (10.15)). Here p_i denotes the i^{th} component of p with respect to some basis. This integral is just the i^{th} component of $\tilde{\nu}_\ell^j * \mu(p)$, and so the lemma follows from (10.10).

Combining Lemma 10.16 with (10.9) yields

$$\int \sum_{(p,\ell)\in\sigma_1} \left| \int_P \alpha_{p,\ell}^j(q) d\mu(q) \right|^2 dp \le C\epsilon^2 |Q(S)|, \tag{10.17}$$

where

$$\alpha_{p,\ell}^j(q) = s^j 2^{-\ell} \nu_\ell^j(p-q) \left[A(p) - A(q) - A_{p,\ell}(p) + A_{p,\ell}(q)\right]. \tag{10.18}$$

The advantage of (10.17) over (10.9) is that we have better control over $\alpha_{p,\ell}(q)$ (coming from Lemma 10.11) than $A(p) - A(q)$, and this will be needed when we try to replace $d\mu(q)$ by something better.

To analyze μ using (10.10) we first need to understand the operator

$$Tf = \log s \left(\sum_{j=0}^{m_0-1} \sum_{\ell=-\infty}^{\infty} \tilde{\nu}_\ell^j * \tilde{\nu}_\ell^j * f \right). \tag{10.19}$$

Here we include a scalar product in the convolution $\tilde{\nu}_\ell^j * \tilde{\nu}_\ell^j$. (Remember that $\tilde{\nu}$ is vector-valued.) Let $\eta(\xi)$ be the Fourier transform of $\tilde{\nu} * \tilde{\nu}$.

Because each component of $\tilde{\nu}$ is real-valued and odd, each component of $\hat{\tilde{\nu}}$ is imaginary, and so $\eta(\xi) \le 0$ everywhere. Also, η is radial, because

ν is radial and $\tilde{\nu}(p) = \nu(p)p$. Since we required that ν have two vanishing moments we have that $|\eta(\xi)| \leq C|\xi|^2$. These properties (and the fact that η is rapidly decreasing) imply that

$$\int_0^\infty \eta(t\xi)\frac{dt}{t}$$

is a finite negative constant.

Set $\lambda(\xi) = \log s \sum_{j=0}^{m_0-1} \sum_{\ell=-\infty}^{\infty} \eta(s^{-j}2^\ell\xi)$, so that $(Tf)^\wedge = \lambda\hat{f}$. We can write λ as

$$\lambda(\xi) = \sum_{j=0}^{m_0-1} \sum_{\ell=-\infty}^{\infty} \int_{s^{-j}2^\ell}^{s^{-j+1}2^\ell} \eta(s^{-j}2^\ell\xi)\frac{dt}{t}.$$

As m_0 gets large, $\lambda(\xi)$ tends to $\int_0^\infty \eta(t\xi)\frac{dt}{t}$, and so our previous remarks now make it clear that there is a constant $a > 0$ so that

$$\frac{1}{2}a \leq -\lambda \leq \frac{3}{2}a$$

if m_0 is large enough. (Remember that $s = 2^{\frac{1}{m_0}}$.) We shall assume from now on that m_0 is sufficiently large for this to happen. Thus T is invertible, $(T^{-1}(f))^\wedge = \lambda^{-1}\hat{f}$, and T^{-1} can be written as a convolution singular integral operator such that the j^{th} derivatives of its kernel are dominated by $|p|^{-d-j}$.

Set $\zeta = T^{-1}(\tilde{\nu})$, so that ζ is also vector-valued. We have that

$$\text{(10.20)} \qquad \begin{aligned} &(a)\ |\nabla^j\zeta(p)| \leq C(j)(1+|p|)^{-d-2-j}, \quad j = 0, 1, \dots \\ &(b)\ \int \zeta(p)a(p)dp = 0 \quad \text{for all affine functions } a(p). \end{aligned}$$

These properties of ζ follow from the fact that $\tilde{\nu}$ is smooth, compactly supported, and satisfies (b), and the estimates on the kernel of T^{-1} noted above.

By definitions we have that

$$\text{(10.21)} \qquad \mu = \log s \sum_{j=0}^{m_0-1} \sum_{\ell=-\infty}^{\infty} \zeta_\ell^j * \tilde{\nu}_\ell^j * \mu,$$

where $\zeta_\ell^j(p) = s^{jd}2^{-\ell d}\zeta(s^j2^{-\ell}p)$, and where the series converges in the sense of distributions. [Again this includes a scalar product in $\zeta_\ell^j * \tilde{\nu}_\ell^j$.] We are going to use this formula to split μ into three pieces, of which one is smooth and the other two will produce contributions to (10.17) that can be controlled. The remaining part of (10.17) (coming from the smooth piece) will be used to control I_1 and thereby prove (10.7).

Let R_j, $j \in I$, be the family of cubes from Section 8, with the corresponding partition of unity $\{\phi_j\}$ of $P \setminus \Pi(Z)$.

Let us write $\mu = f + g + h$, where

$$\begin{aligned} f(p) &= \log s \sum_{j=0}^{m_0-1} \sum_{2^\ell > L} \zeta_\ell^j * \tilde{\nu}_\ell^j * \mu(p) \\ &\quad + \log s \sum_{j=0}^{m_0-1} \sum_{2^\ell \le L} \int_{P \setminus U_4} \zeta_\ell^j (p-w) \tilde{\nu}_\ell^j * \mu(w)dw, \end{aligned} \tag{10.22}$$

$$g(p) = \log s \sum_{j=0}^{m_0-1} \int \sum_{(w,\ell)\in\sigma_1} \zeta_\ell^j(p-w)\tilde{\nu}_\ell^j * \mu(w)dw, \tag{10.23}$$

$$h(p) = \log s \sum_{j=0}^{m_0-1} \int \sum_{(w,\ell)\in e} \zeta_\ell^j(p-w)\tilde{\nu}_\ell^j * \mu(w)dw. \tag{10.24}$$

For the definition of h we have used e to denote the set $\{(w,\ell) \in P \times \mathbf{Z} : w \in U_4, 2^\ell < C_1\gamma_0 D(w)\}$, where C_1 is as in the definition (9.2) of σ_1. These sums should be viewed as converging in the sense of distributions.

We want to replace the $d\mu(q)$ in (10.17) by $f(q)dq$, and so we have to control the corresponding contributions of g and h. The main point for the g term is that (10.10) allows us to control the L^2 norm of g.

LEMMA 10.25. *$\int_P g(p)^2 dp \le C\epsilon^2 |Q(S)|$, with C independent of ϵ, δ, and θ.*

This is proved using the usual duality argument. Let F be any function in $L^2(P)$. Then

$$\Big| \int_P Fg \Big| \le C \log s \sum_{j=0}^{m_0-1} \int \sum_{(w,\ell)\in\sigma_1} |\, \zeta_\ell^j * F(w) \,|\, |\, \tilde{\nu}_\ell^j * \mu(w) \,|\, dw.$$

Using (10.10) and Cauchy-Schwarz we can dominate this by

$$C\epsilon\ |Q(S)|^{\frac{1}{2}} \left(\log s \sum_{j=0}^{m_0-1} \sum_{\ell} \int_P |\ \zeta_\ell^j * F\ |^2\, dw \right)^{\frac{1}{2}} \leq C\epsilon\ |Q(S)|^{\frac{1}{2}}\ \|F\|_2.$$

The last inequality is a standard square function estimate, and in this case it can be derived simply from Plancherel. This proves the lemma.

Next we use this to control the contribution of g to (10.17).

LEMMA 10.26.

$$\int_P \sum_{\ell=-\infty}^{\infty} |\ \int_P \alpha_{p,\ell}^j(q) g(q) dq\ |^2\, dp\ \leq\ C\epsilon^2\ |Q(S)|,$$

where C does not depend on ϵ, δ, or θ.

Let $r \in (2, \frac{2d}{d-2})$ be arbitrary, let r' denote its conjugate exponent, and set

$$G_\ell^j(p) = \left\{ \int_P |\nu_\ell^j(p-q)|\ |g(q)|^{r'} dq \right\}^{\frac{1}{r}},$$

$$F_\ell^j(p) = |\ 2^{-\ell}\, [A(p) - A_{p,\ell}(p)]\ |, \tag{10.27}$$

$$H_\ell^j(p) = \left\{ \int_P |\nu_\ell^j(p-q)|\ |2^{-\ell}\, [A(q) - A_{p,\ell}(q)]\ |^r dq \right\}^{\frac{1}{r}}. \tag{10.28}$$

By Hölder's inequality, the quantity to be estimated is at most

$$C \int \sum G_\ell^j(p)^2 \left(F_\ell^j(p)^2 + H_\ell^j(p)^2 \right) dp. \tag{10.29}$$

We know from (10.12) and (10.14) that

$$\sum_\ell \left(F_\ell^j(p)^2 + H_\ell^j(p)^2 \right) dp d\delta_{2^\ell}(t) \tag{10.30}$$

is a Carleson measure with norm $\leq C\delta^2$. Because $r' < 2$, we can control the L^2 norm of the maximal function

$$G^j_*(p) = \sup_{|q-p|\leq 2^\ell} G^j_\ell(q)$$

by the L^2 norm of g, and so Carleson's inequality (see p236 of [**St**]) and Lemma 10.25 imply that (10.29) is $\leq C\epsilon^2\delta^2|Q(S)|$. Lemma 10.26 follows.

Next we want to control the contribution of h to (10.17). Set $\sigma_5 = \{(p,\ell) \in P \times \mathbf{Z} : p \in U_4,\ \gamma_0^{\frac{1}{2}} D(p) \leq 2^\ell \leq L\}$. (Recall that U_j was defined in (8.9).)

LEMMA 10.31. *There is a C independent of ϵ, δ, γ_0, and θ such that*

$$\int_{(p,\ell)\in\sigma_5} \sum \Big| \int_P \alpha^j_{p,\ell}(q)h(q)dq \Big|^2 dp \ \leq\ C\gamma_0\delta^2|Q(S)|.$$

The notation $h(q)dq$ is convenient but somewhat misleading, since h is a priori only a distribution and at best a measure. The idea behind the lemma is that although h is not very regular, it has a lot of local oscillation to help its integral against $\alpha^j_{p,\ell}$ (which is relatively smooth) to be pretty small.

Let us record a couple of simple estimates on $\alpha^j_{p,\ell}$. It is readily seen from the definition (10.18) of $\alpha^j_{p,\ell}$ (and also (10.15)) that

$$(10.32) \qquad \big|\, \alpha^j_{p,\ell}(q) \,\big| \leq\ C\delta 2^{-\ell d},\ |\nabla\alpha^j_{p,\ell}(q)| \ \leq\ C\delta 2^{-\ell(d+1)},$$

and also that $\operatorname{supp}\alpha^j_{p,\ell} \subseteq B(p,2^\ell)$.

Set $a = a^j_{p,\ell} = \int_P \alpha^j_{p,\ell}(q)h(q)dq$ for $(p,\ell) \in \sigma_5$. Thus

$$a = \log s \sum_{i=0}^{m_0-1} \int_{(w,k)\in e} \sum \int_P \alpha^j_{p,\ell}(q)\zeta^i_k(q-w)dq\ \tilde\nu^i_k * \mu(w)dw$$

by (10.24). We need to get estimates for the interior integral

$$AZ = AZ(j,p,\ell,i,k,w) = \int_P \alpha^j_{p,\ell}(q)\zeta^i_k(q-w)dq.$$

When $k \geq \ell$ or $|p-w| \geq 2^{\ell+1}$ we have

$$|AZ| \leq CH_\ell(p)\Phi_{k,\ell}(p-w), \tag{10.33}$$

where

$$H_\ell(p) = 2^{-\ell d} \int_{B(p,2^\ell)} 2^{-\ell}|A(q) - A_{p,\ell}(q)|dq + 2^{-\ell}|A(p) - A_{p,\ell}(p)|$$

and $\Phi_{k,\ell}(x) = 2^{2k}(2^k + 2^\ell + |x|)^{-d-2}$. This uses (10.20a). If $k \leq \ell$ and $|p-w| \leq 2^{\ell+1}$ then

$$|AZ| \leq C\delta 2^k 2^{-\ell(d+1)}. \tag{10.34}$$

This uses the bound (10.32) on $\nabla\alpha^j_{p,\ell}$ and both parts of (10.20). Let $AZ_1(p,\ell,k,w)$ denote the right side of (10.33) when $k \geq \ell$ or $|p-w| \geq 2^{\ell+1}$, and let it be zero otherwise, and let $AZ_2(p,\ell,k,w)$ denote the right side of (10.34) when $k \leq \ell$, $|p-w| \leq 2^{\ell+1}$, and zero otherwise.

Set $e_1 = \{(w,\ell) \in e : k \geq \ell \text{ or } |p-w| \geq 2^{\ell+1}\}$ and $e_2 = \{(w,\ell) \in e : k < \ell \text{ and } |p-w| < 2^{\ell+1}\}$, and split a into $a(1) + a(2)$ accordingly. Thus

$$a(1,2) \leq C \int \sum_{(w,k)\in e_{(1,2)}} AZ_{(1,2)}(p,\ell,k,w)2^{-kd} \int_{B(w,2^k)} d\mu(r)\, dw.$$

We want to show that

$$\int \sum_{(p,\ell)\in\sigma_5} |a(1,2)|^2 dp \leq C\gamma_0\delta^2|Q(S)|. \tag{10.35}$$

We start with $a(1)$.

By definitions

$$a(1) \leq CH_\ell(p) \int \sum_{(w,k)\in e_1} \int_{B(w,2^k)} 2^{k(2-d)}(2^k + 2^\ell + |p-w|)^{-d-2} d\mu(r)\, dw.$$

If $r \in B(w,2^k)$ and $(w,k) \in e_1$, then $2^k < C_1\gamma_0 D(w)$ [see the definition of e just after (10.24)] and $D(w) \leq |r-w| + D(r)$, and so $D(w) \leq 2D(r)$ if γ_0 is small enough. Thus $2^k < 2C_1\gamma_0 D(r)$. Also, $(w,k) \in e_1$ implies that $2^\ell \leq 2^k$ or $2^{\ell+1} \leq |p-w|$ and this implies that $2^k \geq 2^\ell$ or $|p-r| \geq 2^\ell$ when $|r-w| \leq 2^k$. Thus $(w,k) \in e_1$ and $|w-r| \leq 2^k$ imply that $(r,k) \in \tilde{e}_1$,

$\tilde{e}_1 = \{(r,k) \in P \times \mathbf{Z} : r \in U_3,\ 2^k < 2C_1\gamma_0 D(r)$, and either $2^k \geq 2^\ell$ or $|p-r| \geq 2^\ell\}$, whence

$$a(1) \leq CH_\ell(p) \int \sum_{(r,k)\in\tilde{e}_1} \int_{B(r,2^k)} 2^{k(2-d)}(2^k + 2^\ell + |p-r|)^{-d-2} dw\, d\mu(r)$$

$$\leq CH_\ell(p) \int \sum_{(r,k)\in\tilde{e}_1} 2^{2k}(2^k + 2^\ell + |p-r|)^{-d-2} d\mu(r).$$

To control this expression it will be convenient to break up the r-integral. Let $\{R_i\}$ be the cubes chosen in Section 8, which cover the set $\{q \in P : D(q) > 0\}$. Let I_3 denote the set of i such that $R_i \cap U_3 \neq \emptyset$. Then

$$a(1) \leq CH_\ell(p) \sum_{i\in I_3} \int \sum_{\substack{(r,k)\in\tilde{e}_1 \\ r\in R_i}} 2^{2k}(2^k + 2^\ell + |p-r|)^{-d-2} d\mu(r).$$

Consider $\int_{R_i} (2^k + 2^\ell + |p-r|)^{-d-2} d\mu(r)$. If $p \notin 2R_i$, then the integrand is roughly constant on R_i, and this integral is comparable to the one you get by replacing $d\mu(r)$ by dr, because of Lemma 9.11. If $p \in 2R_i$, then $10 \operatorname{diam} R_i \leq D(p)$, by (8.8), and of course $2^\ell \geq \gamma_0^{\frac{1}{2}} D(p)$ if $(p,\ell) \in \sigma_5$. From Lemma 9.11 we get that $\mu(T) \approx |T|$ for subcubes T of R_i of size 2^ℓ, at least if ϵ is small enough, depending on γ_0. [N.B.: Although we have to let the choice of ϵ depend on γ_0 here, the constant from Lemma 9.11 doesn't depend on anything, not ϵ or γ_0 in particular.] Because $(2^k + 2^\ell + |p-r|)^{-d-2}$ is roughly constant on the scale of 2^ℓ, we get that

$$\int_{R_i} (2^k + 2^\ell + |p-r|)^{-d-2} d\mu(r) \approx \int_{R_i} (2^k + 2^\ell + |p-r|)^{-d-2} dr$$

in this case as well.

Let us use this to control $a(1)$. If $(r,k) \in \tilde{e}_1$, $r \in R_i$, then $2^k \leq C\gamma_0 \operatorname{diam} R_i$ by (8.8), and so

$$a(1) \leq CH_\ell(p) \sum_{i\in I_3} \int_{R_i} \left(\sum_{2^k \leq C\gamma_0 \operatorname{diam} R_i} 2^{2k} \right) \left(2^\ell + |p-r|\right)^{-d-2} dr$$

$$\leq CH_\ell(p) \sum_{i\in I_3} \int_{R_i} \gamma_0^2 (\operatorname{diam} R_i)^2 \left(2^\ell + |p-r|\right)^{-d-2} dr.$$

Let $F_\ell(p)$ denote this last expression without the $CH_\ell(p)$ when $(p,\ell) \in \sigma_5$, and $F_\ell(p) = 0$ otherwise. To check (10.35) in this case we want to use Carleson's inequality and the fact that $\sum_\ell |H_\ell(p)|^2 dp d\delta_{2^\ell}(t)$ is a Carleson measure with norm $\leq C\delta^2$ (which comes from (10.12) and (10.14)). Thus we need to control the L^2 norm of

$$F_*(p) = \sup_{\substack{q,\ell \\ |q-p|\leq 2^\ell}} F_\ell(q).$$

The definition of F clearly gives $F_*(p) \leq C \sup_\ell F_\ell(p)$. Because $2^\ell \geq \gamma_0^{\frac{1}{2}} D(p)$ if $(p,\ell) \in \sigma_5$, we have

$$\operatorname{diam} R_i \leq D(r) \leq D(p) + |p-r| \leq \gamma_0^{-\frac{1}{2}}(2^\ell + |p-r|) \quad \text{when } r \in R_i,$$

so that

$$F_*(p) \leq C \sum_{i\in I_3} \int_{R_i} \gamma_0^2 \operatorname{diam} R_i^2 (\gamma_0^{\frac{1}{2}} \operatorname{diam} R_i + |p-r|)^{-d-2} dr.$$

To estimate the L^2 norm of F_* we integrate it against an arbitrary L^2-function G and observe that

$$\int F_*|G| \;\leq\; C \sum_{i\in I_3} \int_{R_i} \gamma_0 G^*(r) dr,$$

where $G^*(r)$ denotes the Hardy-Littlewood maximal function of G. Of course the right side is at most $C\gamma_0 |Q(S)|^{\frac{1}{2}}\ \|G\|_2$, and so $\int F_*^2 \leq C\gamma_0^2 |Q(S)|$. Thus (10.35) does follow in this case from Carleson's inequality.

Now we want to prove (10.35) for $a(2)$. We have

$$a(2) \leq C \int_{(w,k)\in e_2} \sum \delta 2^k 2^{-\ell(d+1)} 2^{-kd} \int_{B(w,2^k)} d\mu(r)\, dw.$$

If $|w-r| < 2^k$ and $(w,k) \in e$, then

$$D(w) \leq D(r) + |w-r| \leq D(r) + 2^k \leq D(r) + \frac{1}{2} D(w)$$

if γ_0 is small enough. Thus $D(w) \le 2D(r)$, and so $2^k \le 2C_1\gamma_0 D(r)$. Hence

$$\begin{aligned}
a(2) &\le C \int\limits_{w\in B(p,2^{\ell+1})} \sum_{\substack{2^k \le 2C_1\gamma_0 D(r)\\ k\le \ell}} \delta 2^k 2^{-\ell(d+1)} 2^{-kd} \int\limits_{B(w,2^k)} d\mu(r)\, dw \\
&\le C \int\limits_{B(p,2^{\ell+2})} \sum_{2^k \le 2C_1\gamma_0 D(r)} \delta 2^k 2^{-\ell(d+1)} d\mu(r) \\
&\le C \int\limits_{B(p,2^{\ell+2})} \delta\gamma_0 D(r) 2^{-\ell(d+1)} d\mu(r).
\end{aligned}$$

As before we want to replace $d\mu(r)$ by dr, and so we decompose $B(p,2^{\ell+2})$ using the R_i's. Using (8.8) we obtain

$$a(2) \le C \sum_i \int\limits_{B(p,2^{\ell+2})\cap R_i} \delta\gamma_0 (\operatorname{diam} R_i) 2^{-\ell(d+1)} d\mu(r).$$

If $B(p,2^{\ell+2}) \cap R_i \ne \emptyset$, then

$$\operatorname{diam} R_i \le \inf_{R_i} D(r) \le D(p) + \operatorname{dist}(p,R_i) \le D(p) + 2^{\ell+2} \le \gamma_0^{-\frac{1}{2}} 2^{\ell+1}$$

when $(p,\ell) \in \sigma_5$. In this case we obtain from Lemma 9.11 that

$$\mu(B(p,2^{\ell+2}) \cap R_i) \le C|B(p,2^{\ell+2}) \cap 2R_i|,$$

at least if ϵ is small enough. [As before, how small ϵ has to be depends on γ_0, but the constant doesn't.] This gives

$$\begin{aligned}
a(2) &\le C \sum_i \int\limits_{B(p,2^{\ell+2})\cap 2R_i} \delta\gamma_0 (\operatorname{diam} R_i) 2^{-\ell(d+1)} dr \\
&\le C \int\limits_{B(p,2^{\ell+2})} \delta\gamma_0 D(r) 2^{-\ell(d+1)} dr.
\end{aligned}$$

(We have used the fact that the $2R_i$'s have bounded overlap.)

Since $D(r) \le D(p) + |p-r| \le \gamma_0^{-\frac{1}{2}} 2^{\ell+1}$ if $(p,\ell) \in \sigma_5$ and $r \in B(p,2^{\ell+2})$, we conclude that in particular $a(2) \le C\delta\gamma_0^{\frac{1}{2}}$. Hence

$$a(2)^2 \le C\delta^2\gamma_0^{\frac{3}{2}} \int\limits_{B(p,2^{\ell+2})} D(r) 2^{-\ell(d+1)} dr.$$

Let us plug this into the left side of (10.35). We want to use Fubini to integrate in r last. Because $(p,\ell) \in \sigma_5$ and $r \in B(p, 2^{\ell+2})$ imply that $D(r) \leq \gamma_0^{-\frac{1}{2}} 2^{\ell+1}$ and $r \in U_3$ (since $2^\ell \leq L$), we get that

$$\int_{(p,\ell)\in\sigma_5} \sum a(2)^2 \leq C\delta^2\gamma_0^{\frac{3}{2}} \int_{U_3} \sum_{2^{\ell+1} \geq \gamma_0^{\frac{1}{2}} D(r)} \int_{B(r,2^{\ell+2})} 2^{-\ell(d+1)} dp\, D(r) dr$$

$$\leq C\delta^2\gamma_0^{\frac{3}{2}} \int_{U_3} \sum_{2^{\ell+1} \geq \gamma_0^{\frac{1}{2}} D(r)} 2^{-\ell} D(r) dr \leq C\delta^2\gamma_0 |Q(S)|.$$

This completes the proof of Lemma 10.31.

Combining (10.17) with Lemmas 10.26 and 10.31 we conclude that

$$\int_{(p,\ell)\in\sigma_5} \Big| \int_P \alpha^j_{p,\ell}(q) f(q) dq \Big|^2 dp \leq (C\epsilon^2 + C'\gamma_0\delta^2)|Q(S)|, \tag{10.36}$$

with C, C' independent of ϵ, δ, and θ, and C' also independent of γ_0. We now show that f is smooth on U_5, and then look at what happens when you replace $f(q)$ by $f(p)$.

Let us check that f satisfies

$$|f| \leq C,\ |\nabla f| \leq CL^{-1} \qquad \text{on } U_5. \tag{10.37}$$

The first term on the right side in (10.22) clearly satisfies these conditions. (Don't forget that μ has total mass $\leq C|Q(S)|$.) It is not hard to show that the second term also has these properties, using (10.20a).

If $p \in U_6$, $2^\ell \leq L$, then we get

$$\Big| \int_P \alpha^j_{p,\ell}(q)[f(p) - f(q)] dq \Big| \leq C\delta 2^\ell L^{-1}$$

from (10.37), (10.32), and supp $\alpha^j_{p,\ell} \subseteq B(p, 2^\ell)$. Because σ_4 (defined in (10.4)) is contained in $\sigma_5 \cap \{p \in U_6\}$ if k_0 is large enough, we conclude from (10.36) that

$$\int_{(p,\ell)\in\sigma_4} \sum \Big| \int_P \alpha^j_{p,\ell}(q) dq \Big|^2 |f(p)|^2 dp \leq (C\epsilon^2 + C'\gamma_0\delta^2)|Q(S)|$$

$$+ C \int_{(p,\ell)\in\sigma_4} \sum \delta^2 2^{2\ell} L^{-2} dp \leq [C\epsilon^2 + C'(\gamma_0 + \theta^2)\delta^2]\ |Q(S)|,$$

where C' is independent of γ_0, and both C, C' are independent of ϵ, δ, and θ. This reduces to

$$\int \sum_{(p,\ell)\in\sigma_4} 2^{-2\ell} \left| \int_P \nu_\ell^j(p-q)A(q)dq \right|^2 |f(p)|^2 dp \leq [C\epsilon^2 + C'(\gamma_0+\theta^2)\delta^2]\ |Q(S)|, \tag{10.38}$$

with C, C' as above, because of the definition (10.18) of $\alpha_{p,\ell}^j$ and the vanishing moments assumptions on ν.

Comparing this with (10.3) we see that Proposition (10.5) will follow if we can show that there is a $C_2>0$ so that

$$f(p) \geq C_2^{-1} \text{ for all } p\in P \text{ with } \operatorname{dist}(p,\Pi(Q(S))) \leq 100L. \tag{10.39}$$

Fix p. It suffices to show that there is a nonnegative smooth function η with integral 1, support contained in the ball with center p and radius $\left\{C_2\left(\sup_{U_5}|\nabla f| + L^{-1}\right)\right\}^{-1}$, and which satisfies

$$\int \eta(q)f(q)dq \geq 2C_2^{-1}.$$

Because of (10.37) we can choose η so that $\|\eta\|_\infty \leq C(C_2L^{-1})^d$, $\|\nabla\eta\|_\infty \leq C(C_2L^{-1})^{d+1}$. Since $d\mu(q)$ is equivalent to dq at the scale of $\epsilon \operatorname{diam} Q(S)$ (by (6.4), or Lemma 9.11, applied to $Q(S)$), we get that

$$\int \eta(q)d\mu(q) \geq 4C_2^{-1}$$

if C_2 is large enough and ϵ is small enough (ϵ depending on C_2, but not vice-versa). Thus it is enough to check that

$$\left| \int \eta(q)(g(q)+h(q))dq \right| \leq 2C_2^{-1}. \tag{10.40}$$

The contribution of g to (10.40) is certainly less than C_2^{-1} if ϵ is small enough, by Lemma 10.25. For the contribution of h we use the smoothness of η and the localization and cancellation properties (10.20) of ζ:

$$\begin{aligned}
\left| \int \eta(q)h(q)dq \right| &\leq \log s \sum_{j=0}^{m_0-1} \int \sum_{(w,\ell)\in e} |\zeta_\ell^j * \eta(w)|\ |\tilde{\nu}_\ell^j * \mu(w)| dw \\
&\leq C(C_2)\log s \sum_{j=0}^{m_0-1} \int \sum_{(w,\ell)\in e} 2^\ell L^{-d-1} |\tilde{\nu}_\ell^j * \mu(w)| dw \\
&\leq C(C_2) \sum_{2^\ell \leq C_1\gamma_0 L} 2^\ell L^{-d-1} \int_{U_3} d\mu(r) \leq C(C_2)\gamma_0.
\end{aligned}$$

[Remember that e was defined just after (10.24). We have used Fubini in the third inequality.] This is $\leq C_2^{-1}$ if γ_0 is small enough. [It is not hard to compute how $C(C_2)$ depends on C_2, but we don't need to know the answer.]

This proves (10.40), and finishes the proof of Proposition 10.5.

11. The end of the proof that (C2) implies (C4)

To finish the proof we have to show that (2.7) is true. As we have said before (see Lemma 7.9 and Proposition 10.5), it suffices to show that (10.6) implies that $S \notin \mathcal{F}_1$, which we do in this section. [Recall that $\mathcal{F}_1$ was defined just before Lemma 7.4.] The idea is to use (10.7) to control the oscillation of ∇A in a way that is incompatible with the lower bound on the oscillation of ∇A forced by $S \in \mathcal{F}_1$.

Assume that (10.6) holds. By Proposition 10.5,

$$\int_{(p,\ell)\in\sigma_4} \sum 2^{-2\ell} \mid \nu_\ell^j * A(p) \mid^2 dp \leq \tau |Q(S)|$$

for $j = 0, \ldots, m_0 - 1$, where $\tau = C\epsilon^2 + C'(\gamma_0 + \theta^2)\delta^2$, with C, C' as in the proposition. On the other hand, using Lemma 8.22 it is not hard to check that

$$\int_{U_0} \sum_{2^\ell \leq \gamma_1 D(p)} 2^{-2\ell} \mid \nu_\ell^j * A(p) \mid^2 dp \leq C\epsilon^2 |Q(S)|$$

for a γ_1 which is much larger than $\gamma_0^{\frac{1}{2}}$ ($\gamma_1 = (100n)^{-1}$ would be O.K.). [One way to do this is to observe that $\nu_\ell^j * A(p) = \nu_\ell^j * (A - A_p(p))$, where A_p is the linear Taylor approximation to A at p, and then estimate brutally using Lemma 8.22.] Thus if we set $V = \{x \in P : \operatorname{dist}(x, \Pi(Q(S))) \leq 100L\}$, then we can conclude that

$$\int_V \sum_{2^\ell \leq \theta L} 2^{-2\ell} \mid \nu_\ell^j * A(p) \mid^2 dp \leq 2\tau |Q(S)| \tag{11.1}$$

for $j = 0, \ldots, m_0 - 1$.

To use this estimate we need a reproducing formula involving the ν_ℓ^j's. As in the discussion between (10.19) and (10.20), if m_0 is large enough then we can find a function α on P_0 that satisfies (10.20) and also

$$f = \log s \sum_{j=0}^{m_0-1} \sum_{\ell=-\infty}^{\infty} \alpha_\ell^j * \nu_\ell^j * f$$

for all functions f in $L^2(P)$, say, where $\alpha_\ell^j = s^{jd} 2^{-\ell d} \alpha(s^j 2^{-\ell} p)$. Of course the convergence is taken in the sense of distributions. In particular we can apply this to the function A. [In dealing with convergence issues it is helpful to observe that we can choose A to have compact support.]

We decompose A into two pieces, $A = A_1 + A_2$, by

$$\begin{aligned} A_1(p) = &\log s \sum_{j=0}^{m_0-1} \sum_{2^\ell > \theta L} \alpha_\ell^j * \nu_\ell^j * A(p) \\ &+ \log s \sum_{j=0}^{m_0-1} \sum_{2^\ell \leq \theta L} \int_{P \setminus V} \alpha_\ell^j (p-q) \nu_\ell^j * A(q) dq, \end{aligned} \tag{11.2}$$

$$A_2(p) = \log s \sum_{j=0}^{m_0-1} \sum_{2^\ell \leq \theta L} \int_V \alpha_\ell^j (p-q) \nu_\ell^j * A(q) dq. \tag{11.3}$$

From (11.1) we get

$$\int_P |\nabla A_2|^2 \leq C\tau |Q(S)| \tag{11.4}$$

using a duality argument similar to the one in the proof of Lemma 10.25. Here C is a constant that doesn't depend on ϵ, δ, γ_0, or θ. (The same will be true of the other constants in this section unless stated otherwise.)

We would like to say that A_1 is smoother than A. This is true inside V, but not on $P \setminus V$. If we set $V_j = \{x \in P : \operatorname{dist}(x, Q(S)) \leq 100L2^{-j}\}$, then we do have that

$$|\nabla A_1| \leq C\delta, |\nabla^2 A_1| \leq C\delta\theta^{-1}L^{-1} \quad \text{on } V_1. \tag{11.5}$$

To check this it is convenient to let A_{11} and A_{12} denote the two terms on the right side of (11.2). It is not hard to show that A_{12} satisfies (11.5), using $\|\nabla A\| \leq C\delta$ and the fact that α satisfies (10.20a).

To control (11.5) for A_{11} it is helpful to make the following observations. We can write

$$\log s \sum_{j=0}^{m_0-1} \sum_{2^\ell > \theta L} \alpha_\ell^j * \nu_\ell^j = \phi_h,$$

where $h = [\log_2(\theta L)]$ and

$$\phi = \log s \sum_{j=0}^{m_0-1} \sum_{2^\ell > 1} \alpha_\ell^j * \nu_\ell^j.$$

By construction we also have that

$$\phi = (\text{Dirac mass at } 0) - \log s \sum_{j=0}^{m_0-1} \sum_{2^\ell \leq 1} \alpha_\ell^j * \nu_\ell^j,$$

with this last series converging in the sense of distributions. It is not hard to check that ϕ satisfies (10.20a), using the first formula for $\phi(p)$ when $|p| \leq 10$, and the second when $|p| > 10$. From here it is easy to see that $A_{11} = \phi_h * A$ satisfies (11.5).

We want to use these estimates on A_1 and A_2 to control how well A can be approximated by affine functions. First we consider the maximal function

$$N(A_2)(p) = \sup_B\{|B|^{-\frac{1}{d}}(|B|^{-1}\int_B |A_2 - m_B A_2|)\}, \tag{11.6}$$

where the supremum is taken over all balls B that contain p and have radius $\leq L$, and $m_B A_2 = \frac{1}{|B|}\int_B A_2$. It is well-known and not hard to see that

$$\int_P N(A_2)^2 \leq C \int_P |\nabla A_2|^2 \leq C_T |Q(S)|. \tag{11.7}$$

[You can use the Poincaré inequality

$$|B|^{-\frac{d+1}{d}} \int_B |A_2 - m_B A_2| \leq C \frac{1}{|B|} \int_B |\nabla A_2|$$

to reduce (11.7) to an estimate on the Hardy-Littlewood maximal function of ∇A_2.] We are going to use this together with the following estimate on the oscillation of A_2 on a ball B.

LEMMA 11.8. *Set* $\operatorname*{osc}_B A_2 = \sup_{p\in B} |A_2(p)-m_BA_2|$, *and let* r *denote the radius of* B. *If* $B \subseteq V_1$, *then* $\operatorname*{osc}_B A_2 \le Cr\{r^{-1}m_B(|A_2 - m_BA_2|)\}^{\frac{1}{d+1}}\delta^{\frac{d}{d+1}}$.

Let $q \in B$ be such that

$$|A_2(q) - m_BA_2| = \operatorname*{osc}_B A_2 =: \lambda.$$

Since $\|\nabla A_2\|_{L^\infty(B)} \le C\delta$ (because this is true for A, A_1), we have $|A_2(p) - m_BA_2| \ge \frac{\lambda}{2}$ when $p \in B$, $|p-q| \le \frac{\lambda}{2C\delta}$.

If $\frac{\lambda}{2C\delta} \le r$, then we get that

$$\int_B |A_2 - m_BA_2| \ge C^{-1}\frac{\lambda}{2}\left(\frac{\lambda}{2C\delta}\right)^d$$

whence $\lambda^{d+1} \le C\delta^d \int_B |A_2 - m_BA_2|$.

If $\frac{\lambda}{2C\delta} \ge r$, then $|A_2 - m_BA_2| \ge \frac{\lambda}{2}$ over a large portion of B, so that $m_B(|A_2 - m_BA_2|) \ge C^{-1}\lambda$. Using $\|\nabla A_2\|_{L^\infty(B)} \le C\delta$ we also have that

$$r^{-1}m_B(|A_2 - m_BA_2|) \le C\delta,$$

so that

$$\begin{aligned} C^{-1}\lambda \le m_B(|A_2 - m_BA_2|) &= r\{r^{-1}m_B(|A_2 - m_BA_2|)\}^{\frac{1}{d+1}+\frac{d}{d+1}} \\ &\le Cr\{r^{-1}m_B(|A_2 - m_BA_2|)\}^{\frac{1}{d+1}}\delta^{\frac{d}{d+1}}. \end{aligned}$$

Combining these two cases gives the lemma.

Now we are ready to look at how well A is approximated by affine functions. Let $B = B(p_0, r)$ be contained in V_1. We assume that $r \le r_0L$, where r_0 is a small number to be chosen soon.

LEMMA 11.9. *Set* $F = \{p \in V_2 : N(A_2)(p) \le \theta^{\frac{1}{2}}\delta\}$. *If* $B = B(p_0,r)$ *intersects* F *and* $r \le r_0L$, *then*

$$\sup_{p\in B} |A(p) - A(p_0) - \nabla A_1(p_0)\cdot(p-p_0)| \le C\{\theta^{\frac{1}{2(d+1)}} + r_0\theta^{-1}\}r\delta. \tag{11.10}$$

Indeed, if $p \in B$, then from (11.5) we obtain

$$\begin{aligned} &|A(p) - A(p_0) - \nabla A_1(p_0)\cdot(p-p_0)| \\ &\quad \le |A_2(p) - A_2(p_0)| + |A_1(p) - A_1(p_0) - \nabla A_1(p_0)\cdot(p-p_0)| \\ &\quad \le 2\operatorname*{osc}_B A_2 + C\delta\theta^{-1}L^{-1}r^2 \\ &\quad \le Cr\{r^{-1}m_B(|A_2 - m_BA_2|)\}^{\frac{1}{d+1}}\delta^{\frac{d}{d+1}} + C\delta\theta^{-1}r_0r \\ &\quad \le Cr\{N(A_2)(u)\}^{\frac{1}{d+1}}\delta^{\frac{d}{d+1}} + C\delta\theta^{-1}r_0r \end{aligned}$$

for any $u \in B$. Choosing $u \in B \cap F$, we get that this is at most $Cr\theta^{\frac{1}{2(d+1)}}\delta + C\delta\theta^{-1}r_0 r$, as desired.

We shall choose θ small enough, and then r_0 small enough, so that this last quantity is less than $\eta\, r\delta$, for some small η that will be chosen soon.

Let H_B be the d-plane which is the graph of the affine function

$$a_B(p) = A(p_0) + \nabla A_1(p_0) \cdot (p - p_0).$$

If B is as in the lemma, Γ is the graph of A, and the constants are chosen as we just explained, then

$$\sup_{x \in \Gamma \cap \Pi^{-1}(B)} r^{-1} \operatorname{dist}(x, H_B) \leq \eta\delta. \tag{11.11}$$

We want to show that these estimates on A are incompatible with $S \in \mathcal{F}_1$. Recall that $S \in \mathcal{F}_1$ if

$$\Big| \bigcup_{Q \in m_1(S)} Q \Big| \geq |Q(S)|/2, \tag{11.12}$$

where $m_1(S)$ is the set of minimal cubes Q of S such that Angle $(P_Q, P) \geq \delta/2$. We are going to show that if $Q \in m_1(S)$, then $\Pi(Q)$ cannot intersect F, or even get too close, by comparing the H_B's to the P_Q's. We first need to check that the cubes in $m_1(S)$ cannot be too large.

LEMMA 11.13. *Given $r_0 > 0$ and a constant $M > 0$, there is an $\epsilon_0 > 0$ so that if $\epsilon \leq \epsilon_0$ then Angle $(P_Q, P) \leq \delta/100$ for all $Q \in S$ with* $\operatorname{diam} Q \geq r_0 L/M$.

This is quite easy: given $Q \in S$, let $Q^1, Q^2, \ldots, Q^T = Q(S)$ denote the successive ancestors of Q, i.e., Q^1 is the father of Q, Q^2 the father of Q^1, etc. By definitions, all these cubes must satisfy (6.1), and so Lemma 5.13 gives Angle $(P_{Q^i}, P_{Q^{i+1}}) \leq C\epsilon$. Thus

$$\text{Angle } (P_Q, P) \leq C\, \epsilon\, T \leq C\epsilon \log (2L/\operatorname{diam} Q),$$

from which the lemma follows easily.

LEMMA 11.14. *If $Q \in m_1(S)$, then* $\operatorname{dist}(\Pi(Q), F) > \operatorname{diam} Q$.

Suppose not. Pick any $x_Q \in Q$, set $p_0 = \Pi(x_Q)$ and $B = B(p_0, r)$, $r = 10 \operatorname{diam} Q$. Then B intersects F, and $r \leq r_0 L$ because of Lemma 11.13 (if ϵ is sufficiently small). Notice also that $B \subseteq V_1$.

If $x \in Q$, then

$$|x - (\Pi(x), A(\Pi(x)))| \leq C\epsilon d(x) \leq C\epsilon \operatorname{diam} Q$$

by Proposition 8.2 and the definition (7.6) of $d(x)$, and so $\operatorname{dist}(x, H_B) \leq C\eta\delta r$ (if ϵ is small enough) because of (11.11). Invoking Lemma 5.13 again we get that Angle $(P_Q, H_B) \leq C\eta\delta < \frac{\delta}{100}$ if we choose η (and also ϵ) to be small enough.

Let Q^* be the largest ancestor of Q such that $10 \operatorname{diam} Q^* \leq r_0$. Let B^* be the ball centered at p_0 with radius $10 \operatorname{diam} Q^*$. The same argument gives Angle $(P_{Q^*}, H_{B^*}) \leq \frac{\delta}{100}$. Furthermore, $H_{B^*} = H_B$, since the function a_B depends only on p_0.

Thus Angle $(P_Q, P_{Q^*}) \leq \frac{\delta}{50}$, and Lemma 11.13 tells us that Angle $(P_{Q^*}, P) \leq \frac{\delta}{100}$ (if ϵ is small enough). These two inequalities force $Q \notin m_1(S)$, and the lemma is proved.

We are now at the final step of the proof that (10.7) is incompatible with $S \in \mathcal{F}_1$. Because of Lemma 11.14, it suffices to show that (11.12) implies that there is a $Q \in m_1(S)$ with $\operatorname{dist}(\Pi(Q), F) \leq \operatorname{diam} Q$. (Of course we continue to assume that (10.6) holds.)

We first choose a convenient covering of the set $X = \bigcup_{Q \in m_1(S)} Q$. For each $Q \in m_1(S)$, pick a point $x_Q \in Q$, and consider the ball $B_Q = (x_Q, C \operatorname{diam} Q)$. If C is large enough, we can find a subset $\mathcal{T}$ of $m_1(S)$ so that $\{B_Q : Q \in \mathcal{T}\}$ covers X and the balls $B(x_Q, 3 \operatorname{diam} Q)$, $Q \in \mathcal{T}$, are pairwise disjoint. (This follows from the well-known argument used in the proof of the covering lemma on p9 of [**St**].) Hence

$$|X| \leq \sum_{Q \in \mathcal{T}} |E \cap B_Q| \leq C \sum_{Q \in \mathcal{T}} (\operatorname{diam} Q)^d. \tag{11.15}$$

Because $B(x_Q, 3 \operatorname{diam} Q)$ are disjoint for $Q \in \mathcal{T}$, we can apply Lemma 8.4 (with $x = x_Q, y = x_{Q'}$) to conclude that the balls $D_Q = P \cap B(\Pi(x_Q), \operatorname{diam} Q)$, $Q \in \mathcal{T}$, are also disjoint. In particular $\left| \bigcup_{\mathcal{T}} D_Q \right| \geq |X|/C$.

All of these balls D_Q are contained in V_2, and none of them intersects F, because of Lemma 11.14. Since $F = \{p \in V_2 : N(A_2) \leq \theta^{\frac{1}{2}}\delta\}$, we conclude from (11.7) that

$$|X| \leq C \left| \bigcup_{\mathcal{T}} D_Q \right| \leq C|V_2 \setminus F| \leq C\theta^{-1}\delta^{-2}\tau|Q(S)|. \tag{11.16}$$

Recall that $\tau = C\epsilon^2 + C'(\gamma_0 + \theta^2)\delta^2$, where C, C' are independent of ϵ, δ, and θ, and C' is also independent of γ_0. If we choose θ small enough, then γ_0 and r_0, and then ϵ, we get that $|X| \leq \frac{1}{3}|Q(S)|$, so that (11.12) can't hold.

Thus we have shown that if S satisfies (10.6), then $S \notin \mathcal{F}_1$, at least if we chose the various parameters correctly. From the arguments in Section 7 it follows that (2.7) holds, and of course (2.6) was proved in Section 8. This completes the proof that (C2) implies (C4).

12. The proof that (C3) implies (C4): preliminary discussion

The argument will be very similar to the one used to show that (C2) implies (C4).

We assume throughout this section and the next two that E satisfies (C3). As we remarked at the beginning of Section 5, E must satisfy the weak geometric lemma. Let ϵ, δ be given, as small as we want, with ϵ/δ as small as we want also. Let $k > 0$ be large, to be chosen later, but not depending on ϵ or δ.

Let $\mathcal{G}$ be the set of cubes for which (6.1) holds, and set $\mathcal{B} = \Delta \setminus \mathcal{G}$. We can apply Lemma 7.1 to get the stopping-time regions, and to prove that (C4) holds we need only verify (2.7), since the rest of the requirements are established by Lemma 7.1 and Section 8. Lemma 7.4 tells us once again that we can reduce (2.7) to the corresponding estimate for $\mathcal{F}_1$.

To prove the packing condition for $\mathcal{F}_1$ we use a variation of Lemma 7.9, whose particulars are as follows. Let S be one of the stopping-time regions, and set

$$X = \{(x,t) \in E \times \mathbf{R}^+ : x \in kQ(S),\ k^{-1}d(x) \leq t \leq kL\}, \tag{12.1}$$

where $d(x)$ and $L(= \operatorname{diam} Q(S))$ are as before. To prove the packing condition for $\mathcal{F}_1$ it suffices to show that there is an $\eta > 0$ such that

$$\iint\limits_X \beta_1(x,kt)^2 \frac{dx\,dt}{t} \geq \eta |Q(S)| \qquad \text{when } S \in \mathcal{F}_1, \tag{12.2}$$

at least if we choose k_0, k, ϵ, δ correctly. Here η is allowed to depend on ϵ, δ or k, but not S.

The proof of (12.2) will be carried out in two steps, in much the same way as before. In the first step we show how the left side of (12.2) controls a similar quantity on the graph of A (A as in Section 8). In the second step we use these estimates on A to show that if the left side of (12.2) is small enough, then $S \notin \mathcal{F}_1$.

13. Pushing estimates on $\beta_1(x,t)$ from E down to the approximating Lipschitz graph

Let k_0 be large, to be chosen later; k_0 will not depend on ϵ, δ, and it will be chosen before k, in such a way that k will be much bigger than k_0. Let A, P, Π, R_i, $Q(i)$, etc. be as in Section 8, Proposition 8.2 in particular.

Given $p \in U_0$ and $t > 0$, set

$$\gamma(p,t) = \inf_a t^{-d} \int\limits_{B(p,t)} \frac{1}{t}|A(u) - a(u)|du, \tag{13.1}$$

where the infimum is taken over all affine functions $a : P \to P^{\perp}$. It is easy to see that if we set

$$\tilde{\gamma}(p,t) = \inf_M t^{-d} \int\limits_{B(p,t)} \frac{1}{t}\operatorname{dist}((u, A(u)), M)du, \tag{13.2}$$

where the infimum is taken over all d-planes M, then

$$\frac{1}{2}\tilde{\gamma}(p,t) \leq \gamma(p,t) \leq 2\tilde{\gamma}(p,t). \tag{13.3}$$

This uses the fact that A is Lipschitz with small norm. Of course $\tilde{\gamma}_1(p,t)$ is essentially the same as the analogue of $\beta_1(x,t)$ for the graph Γ of A.

This section is devoted to proving the following estimate.

LEMMA 13.4. *Set $T = k_0L/10$ and*

$$\tau = L^{-d} \int_0^T \int\limits_{U_2} \gamma(p,t)^2 \frac{dpdt}{t}. \tag{13.5}$$

If k/k_0 is large enough, then

$$\tau \leq C\epsilon^2 + CL^{-d} \iint\limits_X \beta_1(x, kt)^2 \frac{dxdt}{t}, \tag{13.6}$$

where C does not depend on ϵ or δ.

[Recall that X was defined in (12.1).] We restricted ourselves to U_2 so that $B(p,t)$ stays inside U_1 if $p \in U_2$ and $t < T$.

We first estimate $\gamma(p,t)$ for small t. Let R_i, I, I_0, etc., be as in Section 8, and let I_2 denote the set of $i \in I_0$ such that $R_i \cap U_2 \neq \emptyset$.

LEMMA 13.7.

$$\sum_{i \in I_2} \int_0^{\operatorname{diam} R_i} \int\limits_{R_i} \gamma(p,t)^2 \frac{dpdt}{t} \leq C\epsilon^2 L^d.$$

The main ingredients for this are Lemma 8.22 and Taylor's theorem. From the latter we get

$$\gamma(p,t) \leq Ct \sup_{u \in B(p,t)} |\nabla^2 A(u)|,$$

and so using Lemma 8.22 we get

$$\sum_{i \in I_2} \int_0^{\operatorname{diam} R_i} \int\limits_{R_i} \gamma(p,t)^2 \frac{dpdt}{t} \leq C\epsilon^2 \sum_i \int_0^{\operatorname{diam} R_i} \int\limits_{R_i} (\operatorname{diam} R_i)^{-2} t dpdt$$
$$\leq C\epsilon^2 \sum_i \int\limits_{R_i} dp \leq C\epsilon^2 |Q(S)| \leq CL^d.$$

Now we need to estimate $\gamma(p,t)$ when $p \in \Pi(Z)$ or when $p \in R_i$ but $t > \operatorname{diam} R_i$. Both cases will be covered if we assume that $t > D(p)/60$, because of (8.8).

Let $p \in U_2$ and t be given, $1/60D(p) \leq t \leq T$. Choose $z \in E$ such that $z \in Q(S)$, $|p - \Pi(z)| \leq Ct$. We want to control $\gamma(p,t)$ in terms of $\beta_1(z, \frac{1}{10}kt)$. Let $P_{p,t}$ be a d-plane for which the infimum in the definition of $\beta_1(z, \frac{1}{10}kt)$ is achieved. Taking $M = P_{p,t}$ in (13.2) we get that

$$\gamma(p,t) \leq 2t^{-d-1} \int\limits_{B(p,t)} \operatorname{dist}((u, A(u)), P_{p,t}) du \leq a + \sum_{i \in I(p,t)} a_i, \tag{13.8}$$

where $a = 2t^{-d-1} \int_{B(p,t)\cap\Pi(Z)} \operatorname{dist}((u, A(u)), P_{p,t})du$, $I(p,t)$ is the set of indices $i \in I_0$ for which $R_i \cap B(p,t) \neq \emptyset$, and

$$a_i = 2t^{-d-1} \int\limits_{B(p,t)\cap R_i} \operatorname{dist}((u, A(u)), P_{p,t})du.$$

Let us estimate a first. If $u \in B(p,t) \cap \Pi(Z)$, then $x = (u, A(u)) \in Z$, and it follows from Lemma 8.4 that $|x - z| \leq Ct$ and that x is the only point in $\Pi^{-1}(u) \cap k_0 Q(S)$. We can lift the integral defining a from $\Pi(Z)$ to $Z \subseteq E$ to get that

$$a \leq 2t^{-d-1} \int\limits_{\Pi^{-1}(B(p,t))\cap k_0 Q(S)} \operatorname{dist}(x, P_{p,t})dx \leq C\beta_1(z, \frac{1}{10}kt). \tag{13.9}$$

Next we estimate the a_i. Fix $i \in I(p,t)$. We have

$$a_i \leq C(b_i + c_i), \tag{13.10}$$

where $b_i = t^{-d-1} \int_{R_i \cap B(p,t)} \operatorname{dist}((u, A(u)), P_{Q(i)})du$ and

$$c_i = t^{-d-1}(\operatorname{diam} R_i)^d \sup\{\operatorname{dist}(w, P_{p,t}) : w \in P_{Q(i)},\ \operatorname{dist}(w, Q(i)) \leq C \operatorname{diam} R_i\}.$$

[The set of w's over which this supremum is taken is large enough because $\operatorname{dist}(R_i, \Pi(Q_i)) \leq C \operatorname{diam} R_i$.] Because $P_{Q(i)}$ is the graph of B_i,

$$b_i \leq t^{-d-1} \int\limits_{R_i \cap B(p,t)} |A(u) - B_i(u)|du \leq C\epsilon\, t^{-d-1}(\operatorname{diam} R_i)^{d+1}, \tag{13.11}$$

since $|A(u) - B_i(u)| \leq C\epsilon \operatorname{diam} R_i$ by Lemma 8.17 and the definition (8.14) of A. We can control c_i as follows.

Lemma 13.12.

$$\begin{aligned} c_i \leq{} & C\epsilon\, t^{-d-1} (\operatorname{diam} R_i)^{d+1} \\ & + Ct^{-d-1} (\operatorname{diam} R_i)^d \left\{ |2Q(i)|^{-1} \int\limits_{2Q(i)} \operatorname{dist}(x, P_{p,t})^{\frac{1}{3}}\, dx \right\}^3. \end{aligned}$$

This lemma is similar to Lemma 5.13, and their proofs are quite similar also. If $y_0, y_1, \ldots, y_d$ are the $d+1$ points promised by Lemma 5.8 with $Q = Q(i)$, then

$$c_i \le C\, t^{-d-1} \left(\operatorname{diam} R_i\right)^d \sup_{0 \le j \le d} \left\{ \operatorname{dist}(y_j, P_{p,t}) + \operatorname{dist}\left(y_j, P_{Q(i)}\right) \right\}.$$

To check this we compare $P_{p,t}$ and $P_{Q(i)}$ with the d-plane L_d generated by $y_0, y_1, \ldots, y_d$. The distance of points on $P_{Q(i)}$ to L_d is controlled by $\sup_{0 \le j \le d} \operatorname{dist}(y_j, P_{Q(i)})$. [To see this it is helpful to observe that because this quantity is small compared to $\operatorname{diam} Q(i)$, the points z_j in $P_{Q(i)}$ closest to y_j generate $P_{Q(i)}$.] It is easy to show that $\sup_{0 \le j \le d} \operatorname{dist}(y_j, P_{p,t})$ controls the distance of points on L_d to $P_{p,t}$. Combining these two facts gives the inequality. Since $Q(i) \in S$ satisfies (6.1) we have

$$c_i \le C\, t^{-d-1} \left(\operatorname{diam} R_i\right)^d \sup_{0 \le j \le d} \left\{ \operatorname{dist}(y_j, P_{p,t}) \right\} + C\epsilon\, t^{-d-1} (\operatorname{diam} R_i)^{d+1}.$$

Moreover, this inequality remains true if each y_j is replaced by any $\tilde{y}_j \in E$ such that $|y_j - \tilde{y}_j| \le C_0^{-1} \operatorname{diam} R_i$, provided C_0 is large enough. (Indeed, such $\tilde{y}_j$'s would still satisfy the same properties that the y_j's do.) The lemma follows by taking cubic roots of the inequality, averaging over such $\tilde{y}_j$'s, and using the regularity assumption (1.5).

We want to combine these various inequalities with (13.8) to get a good estimate for $\gamma(p,t)$. For this it is important that we took a cubic root in Lemma 13.12, instead of just an ordinary average, so that we can control the overlap of the $2Q(i)$'s which enters into the sum in i.

Let $J(i)$ be the subset of $I(p,t)$ composed of the j's such that $\operatorname{diam} Q(j) \le \operatorname{diam} Q(i)$ and $2Q(j) \cap 2Q(i) \ne \emptyset$. Set $N_i(x) = \sum_{j \in J(i)} \chi_{2Q(j)}(x)$. Then $\sum N_i(x)^{-2} \le C$ for all x (because $\sum \frac{1}{m^2} < \infty$), and

$$\int_{2Q(i)} N_i(x) \le \sum_{j \in J(i)} |2Q(j)| \le C \sum_{j \in J(i)} |R_j| \le C |Q(i)|,$$

because the R_j's are disjoint and stay at distance $\le C \operatorname{diam} R_i$ from $\Pi(Q_i)$.

From Hölder's inequality we now get that

$$\begin{aligned}
&\{|2Q(i)|^{-1} \int\limits_{2Q(i)} \operatorname{dist}(x, P_{p,t})^{\frac{1}{3}} N_i(x)^{-\frac{2}{3}} N_i(x)^{\frac{2}{3}} dx\}^3 \\
&\le \{|2Q(i)|^{-1} \int\limits_{2Q(i)} \operatorname{dist}(x, P_{p,t}) N_i(x)^{-2} dx\}\{|2Q(i)|^{-1} \int\limits_{2Q(i)} N_i(x) dx\}^2 \\
&\le C(\operatorname{diam} R_i)^{-d} \int\limits_{2Q(i)} \operatorname{dist}(x, P_{p,t}) N_i(x)^{-2} dx.
\end{aligned}$$

Putting this into Lemma 13.12, and then using (13.8), (13.9), (13.10), and (13.11) we obtain

$$\begin{aligned}
\gamma(p,t) \le\, & C\beta_1\left(z, \frac{1}{10}kt\right) + C\epsilon\, t^{-d-1} \sum_{i \in I(p,t)} (\operatorname{diam} R_i)^{d+1} \\
& + C\, t^{-d-1} \sum_{i \in I(p,t)} \int\limits_{2Q(i)} \operatorname{dist}(x, P_{p,t}) N_i(x)^{-2} dx.
\end{aligned} \tag{13.13}$$

Because $\sum N_i(x)^{-2}$, the last term is at most

$$C\, t^{-d-1} \int\limits_{\substack{\cup 2Q(i) \\ i \in I(p,t)}} \operatorname{dist}(x, P_{p,t}) dx. \tag{13.14}$$

Let us show that this can be dominated by $\beta_1(z, \frac{1}{10}kt)$. First note that $\operatorname{diam} R_i \le Ct$ if $i \in I(p,t)$, because $D(u) \le D(p) + t \le Ct$ on $B(p,t)$. Thus $\operatorname{dist}(\Pi(Q_i), \Pi(z)) \le Ct$ (because of the various definitions), and this implies that $Q_i \subseteq B(z, Ct)$. [Pick $x \in Q(i)$. By definition, $d(x) \le \operatorname{diam} Q(i) \le Ct$, and so if $|x - z|$ is much bigger than Ct, then Lemma 8.4 implies that

$$|\Pi^\perp(x) - \Pi^\perp(z)| \;\le\; 2\delta |\Pi(x) - \Pi(z)|.$$

Because the right side is at most Ct, we get that $|x - z| \le Ct$.] Now it is immediate that (13.14) is dominated by $C\beta_1(z, \frac{1}{10}kt)$, and so

$$\gamma(p,t) \le C\beta_1\left(z, \frac{1}{10}kt\right) + C\epsilon\, t^{-d-1} \sum_{i \in I(p,t)} (\operatorname{diam} R_i)^{d+1}. \tag{13.15}$$

Of course the same argument used to prove (13.15) also works with z replaced by any $w \in E \cap B(z,t)$ and so

$$\begin{aligned}\gamma(p,t)^2 \le & Ct^{-d} \int_{B(z,t)\cap E} \beta_1\left(w, \frac{1}{10}kt\right)^2 dw \\ & + C\{\epsilon\, t^{-d-1} \sum_{i \in I(p,t)} (\operatorname{diam} R_i)^{d+1}\}^2.\end{aligned} \tag{13.16}$$

To finish the proof of Lemma 13.4 we want to use this to estimate

$$\tau_1 = L^{-d} \int_{U_2} \int_{C^{-1}D(p)}^{T} \gamma(p,t)^2 \frac{dt}{t} dp.$$

Using (13.16) we have that $\tau_1 \le C(a+b)$, with

$$a = L^{-d} \int_{U_2} \int_{C^{-1}D(p)}^{T} t^{-d} \int_{B(z(p,t),t)\cap E} \beta_1\left(w, \frac{1}{10}kt\right)^2 dw \frac{dt}{t} dp,$$

$$b = \epsilon^2 L^{-d} \int_{U_2} \int_{C^{-1}D(p)}^{T} t^{-2(d+1)} \{\sum_{i\in I(p,t)} (\operatorname{diam} R_i)^{d+1}\}^2 \frac{dt}{t} dp,$$

where $z(p,t) = z$ is the point in $Q(S)$ we chose shortly after proving Lemma 13.7.

Consider a first. By definitions, for any (w,p,t) that arise in the integral, we have $|\Pi(w) - p| \le Ct$ and $w \in k_0 Q(S)$. Thus

$$\begin{aligned} a &\le L^{-d} \int_{k_0 Q(S)} \int_{C^{-1}D(p)}^{T} t^{-d} \left(\int_{p \in B(\Pi(w),Ct)} dp \right) \beta_1\left(w, \frac{1}{10}kt\right)^2 \frac{dw\,dt}{t} \\ &\le CL^{-d} \int_{k_0 Q(S)} \int_{C^{-1}D(p)}^{T} \beta_1\left(w, \frac{1}{10}kt\right)^2 dw \frac{dt}{t} \\ &\le CL^{-d} \iint_X \beta_1(x,kt)^2 \frac{dx\,dt}{t}. \end{aligned}$$

To estimate b recall that $\operatorname{diam} R_i \le Ct$ if $i \in I(p,t)$ (because $D(u) \le D(p) + t \le Ct$ on $B(p,t)$, and using (8.6)). Thus

$$\sum_{i\in I(p,t)} (\operatorname{diam} R_i)^{d+1} \le Ct \sum_{i \in I(p,t)} (\operatorname{diam} R_i)^d \le Ct^{d+1}.$$

This gives us that

$$
\begin{aligned}
b &\le C\epsilon^2 L^{-d} \int_{U_2} \int_{C^{-1}D(p)}^{T} \sum_{i\in I(p,t)} (\operatorname{diam} R_i)^{d+1} \frac{dtdp}{t^{d+2}} \\
&\le C\epsilon^2 L^{-d} \sum_{i\in I_1} (\operatorname{diam} R_i)^{d+1} \int_{C^{-1}\operatorname{diam} R_i}^{T} \int_{\operatorname{dist}(p,R_i)\le t} \frac{dpdt}{t^{d+2}}.
\end{aligned}
$$

In interchanging the order of the sum and the integral, we have used the facts that $\operatorname{dist}(p,R_i) \le t$ when $i \in I(p,t)$, and that $i \in I(p,t)$ and $t \ge C^{-1}D(p)$ imply $t \ge C^{-1}\operatorname{diam} R_i$, and we have let I_1 denote the set of $i \in I_0$ such that $R_i \cap U_1 \ne \emptyset$. From here we easily get that

$$
\begin{aligned}
b &\le C\epsilon^2 L^{-d} \sum_{i\in I_1} (\operatorname{diam} R_i)^{d+1} \int_{t>C^{-1}\operatorname{diam} R_i} \{\operatorname{diam} R_i + t\}^d \frac{dt}{t^{d+2}} \\
&\le C\epsilon^2 L^{-d} \sum_{i\in I_1} (\operatorname{diam} R_i)^d \le C\epsilon^2.
\end{aligned}
$$

Combining this with our estimate for a we have

$$
\tau_1 \le CL^{-d} \iint_X \beta_1(x,kt)^2 \frac{dxdt}{t} + C\epsilon^2.
$$

This and Lemma 13.7 give us the estimate (13.6) we wanted for τ, and so Lemma 13.4 is proved.

14. The end of the proof that (C3) implies (C4)

According to Section 12, it suffices to show that (12.2) holds for η small enough. Thus, for example, it suffices to show that if

$$\iint_X \beta_1(x, kt)^2 \frac{dxdt}{t} \leq \epsilon^2 |Q(S)| \tag{14.1}$$

then $S \notin \mathcal{F}_1$.

The proof of this is very similar to the argument given in Section 11 in proving that (C2) implies (C4): we use Section 13 to turn (14.1) into an estimate on A, then apply Littlewood-Paley theory to this estimate to show that ∇A doesn't oscillate much, and then check that that prevents many minimal cubes of S from lying in $m_1(S)$.

This step in the proof that (C3) implies (C2) is in fact so similar to the corresponding step for (C2) that we can simply reduce to our earlier argument. Assume that (14.1) holds for a fixed S. By Lemma 13.4,

$$\int_0^L \int_{U_2} \gamma(p,t)^2 \frac{dpdt}{t} \leq C\epsilon^2 |Q(S)|.$$

If ν_ℓ^j is as described relatively early in Section 10, then $2^{-\ell}|\nu_\ell^j * A(p)| \leq C\gamma(p,t)$ whenever $2^{\ell-1} \leq t \leq 2^\ell$, because of the vanishing moments assumption on ν. Therefore, just like for (11.1),

$$\int_V \sum_{2^\ell \leq L} 2^{-2\ell} |\nu_\ell^j * A(p)|^2 \leq C\epsilon^2 |Q(S)|, \tag{14.2}$$

where V is as it was for (11.1). This is even better than (11.1), since there are no θ, γ_0 around.

Once we have this version of (11.1), we can use exactly the same argument as in Section 11 to show that $S \notin \mathcal{F}_1$ if ϵ, δ are chosen properly. This completes the proof that (C3) implies (C4).

Of course if we were to do the argument over again there are some relatively minor changes we could profitably make. For example, we could replace the discrete square function estimate (14.2) by a continuous one, namely

$$\int_V \int_0^L |\nu_t * A(p)|^2 \frac{dpdt}{t^3} \leq C\epsilon^2 |Q(S)|,$$

and then we could use the reproducing formula

$$f = c \int_0^\infty \nu_t * \nu_t * f \frac{dt}{t}$$

instead of the more complicated discrete one we used before. There are other simplifications that could be made, stemming from the absence of θ in (14.2) and from the stronger nature of $\gamma(p,t)$ as compared to $\nu_t * A(p)$, but the main ideas are still the same.

15. (C4) implies (C3)

As promised in Section 1, we shall actually prove that if E satisfies (C4), then

$$\beta_r(x,t)^2 \frac{dx\,dt}{t} \tag{15.1}$$

is a Carleson measure on $E \times \mathbf{R}_+$ for all $r < \frac{2d}{d-2}$ if $d \geq 2$, $r \leq \infty$ if $d = 1$. The idea is that the existence of a corona decomposition will allow us to reduce to the case of Lipschitz graphs, for which we can use Lemma 10.14.

Let $\eta > 0$ be given, small, as in the definition of a corona decomposition. (The smallness of η will be a convenience that will not play a major role.) Let $\mathcal{B}$, $\mathcal{G}$, $\mathcal{F}$, $S \in \mathcal{F}$ be as in Section 2.

Given $Q \in \Delta$, let $\widehat{Q}$ denote the set of $(x,t) \in E \times \mathbf{R}_+$ such that $x \in Q$ and

$$(10C)^{-2} \operatorname{diam} Q \leq t \leq \operatorname{diam} Q,$$

where C is as in (2.2). It is easy to check that $E \times \mathbf{R}_+ = \bigcup_{Q \in \Delta} \widehat{Q}$. Given $S \in \mathcal{F}$, set $\widehat{S} = \bigcup_{Q \in S} \widehat{Q}$.

LEMMA 15.2. *Let $a(x,t)$ be a bounded, nonnegative function on $E \times \mathbf{R}_+$. Then $a(x,t)dx\frac{dt}{t}$ is a Carleson measure if $a(x,t)\chi_{\widehat{S}}(x,t)\frac{dx\,dt}{t}$ is for each $S \in \mathcal{F}$, with uniformly bounded norm.*

Decompose a into $a_0 + a_1$, where $a_0 = a\chi_{\widehat{\mathcal{B}}}$, $\widehat{\mathcal{B}} = \bigcup_{Q \in \mathcal{B}} \widehat{Q}$. It is easy to check that $a_0(x,t)\frac{dx\,dt}{t}$ is a Carleson measure, because of (2.4).

On the other hand

$$a_1 \leq \sum_{S \in \mathcal{F}} a\chi_{\widehat{S}}$$

by definitions. Fix $x \in E$ and $R > 0$, and consider

$$\sum_{S \in \mathcal{F}} \int_{y \in E \cap B(x,R)} \int_0^R a(y,t)\chi_{\widehat{S}}(y,t)dy\frac{dt}{t}. \tag{15.3}$$

Let Σ^1 denote the sum over the $S \in \mathcal{F}$ for which $Q(S) \cap B(x,R) \neq \emptyset$ and $\operatorname{diam} Q(S) \leq R$, and let Σ^2 be the sum over the S's such that $\widehat{S} \cap ((B(x,R) \cap E) \times (0,R)) \neq \emptyset$ and $\operatorname{diam} Q(S) > R$. We may replace the sum in (15.3) by $\Sigma^1 + \Sigma^2$.

By assumptions we have that

$$\begin{aligned}
\sum{}^1 \int_0^R \int_{B(x,R)\cap E} & a(y,t)\chi_{\widehat{S}}(y,t)\frac{dydt}{t} \\
&\leq \sum{}^1 \int_0^{\operatorname{diam} Q(S)} \int_{Q(S)} a(y,t)\chi_{\widehat{S}}(y,t)\frac{dydt}{t} \\
&\leq C \sum{}^1 |Q(S)| \leq CR^d.
\end{aligned}$$

The last inequality comes from (2.7). To control Σ^2 we observe that Σ^2 only involves a bounded number of S's. This is because the S's are disjoint, and the ones that arise in Σ^2 must each contain a cube Q for which $Q \cap B(x,R) \neq \emptyset$ and $\operatorname{diam} Q \approx R$. This and our hypotheses on a yield

$$\sum{}^2 \int_0^R \int_{B(x,R)} a(y,t)\chi_{\widehat{S}}(y,t)dy\frac{dt}{t} \leq C \sum{}^2 R^d \leq CR^d.$$

This proves the lemma.

In view of the lemma we need only control the $\beta_r(x,t)$'s on each $\widehat{S}$. Fix $S \in \mathcal{F}$, and let Γ be the Lipschitz graph over the d-plane P promised in (2.6). Let $P^\perp$ be the $(n-d)$-plane that passes through the origin and is orthogonal to P, and let $A : P \to P^\perp$ be the Lipschitz function such that $\Gamma = \{p + A(p) : p \in P\}$. Let Π, $\Pi^\perp$ be the orthogonal projections onto P, $P^\perp$ so that $x = \Pi(x) + \Pi^\perp(x)$. Define $d(x)$ as in (7.6).

Set

$$\gamma_r(p,t) = \inf_a \left\{ t^{-d} \int_{B(p,t)} |t^{-1}[A(u) - a(u)]|^r du \right\}^{\frac{1}{r}} \tag{15.4}$$

for $p \in P$, $t > 0$, where the infimum is taken over all affine functions $a : P \to P^\perp$. We want to control the β_r's on E in terms of the γ_r's.

Let $(x,t) \in \widehat{S}$ be given, so that $t \leq L = \operatorname{diam} Q(S)$, and let p be any point in P such that $|p - \Pi(x)| \leq t$. Let a be an affine function for which the infimum in the definition of $\gamma_r(p, 100t)$ is attained, and let $H = H_{p,t}$ denote its graph. For $y \in E$, $|y - x| \leq t$, we have

$$\begin{aligned} \operatorname{dist}(y,H) &\leq |\Pi^\perp(y) - A(\Pi(y))| + |A(\Pi(y)) - a(\Pi(y))| \\ &\leq C\eta d(y) + |A(\Pi(y)) - a(\Pi(y))|. \end{aligned} \tag{15.5}$$

(This uses $\operatorname{dist}(y,\Gamma) \leq C\eta d(y)$ which can be derived from (2.6).)

We want to use this to control $\beta_r(x,t)$ in terms of $\gamma_r(p, 100t)$. For $r = \infty$ we get that

$$\beta_\infty(x,t) \leq C\eta \left(\sup_{y \in B(x,t) \cap E} t^{-1} d(y) \right) + \gamma_\infty(p, 100t) \tag{15.6}$$

for all $p \in P$ with $|p - \Pi(x)| \leq t$. For $r < \infty$ we cannot estimate so quickly, because we might get a very singular measure when we push Hausdorff measure on E down to P, and so we don't want to simply integrate (15.5).

Instead we have to smear up (15.5) first. Let $\delta \in (0,1)$ be a small positive number, to be chosen later, and let B_y denote the ball in P with center $\Pi(y)$ and radius $\delta d(y)$. Because of (15.5) we have for $u \in B_y$ that

$$\operatorname{dist}(y,H) \leq C\eta d(y) + C\delta d(y) + |A(u) - a(u)|.$$

(We have used here the fact that A is Lipschitz.) Taking r^{th} powers and averaging over B_y gives

$$\operatorname{dist}(y,H)^r \leq C(\eta + \delta)^r d(y)^r + C[\delta d(y)]^{-d} \int_{B_y} |A - a|^r du. \tag{15.7}$$

[When $d(y) = 0$ we interpret the average over B_y as being the value at y. Similar liberal interpretations are required below. Alternatively, one can treat the case $d(y) = 0$ separately, by simply integrating (15.5).] Hence

$$\begin{aligned} \int_{B(x,t)\cap E} \operatorname{dist}(y,H)^r dy &\leq C(\eta+\delta)^r \int_{B(x,t)\cap E} d(y)^r dy \\ &\quad + C \int_{B(x,t)\cap E} [\delta d(y)]^{-d} \int_{B_y} |A(u) - a(u)|^r du dy, \end{aligned} \tag{15.8}$$

(Don't forget that $y \in E$, but $u \in P$.)

We want to use Fubini to simplify the last term, but first we need some preliminary information. We begin by observing that $d(x) \leq Ct$ because $(x,t) \in \widehat{S}$, and so $d(y) \leq Ct$ if $y \in B(x,t)$. In particular, $\delta d(y) \leq t$, if δ is small enough. Next, let us check that if $y,\ z \in B(x,t) \cap E$ satisfy $B_y \cap B_z \neq \emptyset$, then $d(y)$ and $d(z)$ are comparable, and that $|y-z| \leq d(y)$. Indeed, we certainly have

$$|\Pi(y) - \Pi(z)| \ \leq \ C\delta(d(y) + d(z)),$$

and also

$$|y - (\Pi(y), A(\Pi(y)))| \ \leq \ C \operatorname{dist}(y,\Gamma) \leq \ C\eta d(y),$$

and similarly for z. Hence

$$|y-z| \ \leq \ C(\delta+\eta)(d(y)+d(z)).$$

If η and δ are small enough we can conclude that

$$\frac{1}{2}d(y) \leq d(z) \leq 2d(y),$$

because $d(\cdot)$ is Lipschitz.

With these observations we can reverse the order of integration in the last term in (15.8) to obtain

$$\begin{aligned}
&\int\limits_{B(x,t)\cap E} [\delta d(y)]^{-d} \int\limits_{B_y} |A(u) - a(u)|^r du dy \\
&\qquad \leq \int\limits_{B(p,10t)} |A(u)-a(u)|^r \left(\int\limits_{\substack{y\in B(x,t)\cap E \\ u\in B_y}} [\delta d(y)]^{-d}\, dy \right) du \\
&\qquad \leq C \int\limits_{B(p,10t)} |A(u)-a(u)|^r du.
\end{aligned}$$

Hence

$$\beta_r(x,t) \leq C \left(\frac{1}{t^d} \int\limits_{B(x,t)\cap E} \left(t^{-1}d(y)\right)^r dy \right)^{\frac{1}{r}} + C\gamma_r(p, 100t).$$

Combining this with (15.6) we get that

$$\beta_r(x,t) \le C\left(\frac{1}{t^d}\int\limits_{B(x,t)\cap E}\left(t^{-1}d(y)\right)^r dy\right)^{\frac{1}{r}} + C\inf_{p\in B(\Pi(x),t)}\gamma_r(p,100t) \tag{15.9}$$

for any $r \le \infty$, $(x,t) \in \widehat{S}$, $t \le \operatorname{diam} Q(S)$. Let us use this to check that $\beta_r(x,t)^2\chi_{\widehat{S}}(x,t)\frac{dxdt}{t}$ is a Carleson measure if $r < \frac{2d}{d-2}$ when $d \ge 2$, $r \le \infty$ when $d = 1$.

Fix $z \in E$, $R > 0$. To control

$$\int_0^R \int\limits_{B(z,R)\cap E} \beta_r(x,t)^2\chi_{\widehat{S}}(x,t)\frac{dxdt}{t} \tag{15.10}$$

we consider separately the contributions coming from the two terms on the right side of (15.9), starting with the second. It is dominated by

$$\int_0^R \int\limits_{B(z,R)\cap E} \chi_{\widehat{S}}(x,t)\frac{1}{t^d}\int\limits_{B(\Pi(x),t)} \gamma_r(p,100t)^2 dpdx\frac{dt}{t}. \tag{15.11}$$

Of course we want to use Fubini's theorem.

It is easy to reduce to the case where $z \in Q(S)$ and $R \le \operatorname{diam} Q(S)$. It is also not hard to check that the set of x in $Q(S)$ such that $(x,t) \in \widehat{S}$ and $|\Pi(x) - p| \le t$ (for given p, t) is contained in a ball of radius Ct, and hence has measure $\le Ct^d$. This uses (2.6) and the fact that $d(x) \le Ct$ if $(x,t) \in \widehat{S}$. Thus by reversing the order of integration we see that (15.11) is at most

$$C\int_0^R \int\limits_{B(\Pi(z),2R)} \gamma_r(p,100t)^2\,\frac{dpdt}{t},$$

and this is $\le CR^d$ by Lemma 10.11.

To show that the contribution to (15.10) from the first term on the right side of (15.9) is at most CR^d it is enough to prove that for any cube $T \in S$,

$$\sum_{\substack{Q\in S\\ Q\subseteq T}} \lambda_r(Q)^2|Q| \le C|T| \tag{15.12}$$

if $r < \frac{2d}{d-2}$ ($r \leq \infty$ when $d = 1$), where

$$\lambda_r(Q) = \left(\frac{1}{|Q|} \int_{2Q} (d(y)/\operatorname{diam} Q)^r \, dy \right)^{\frac{1}{r}}.$$

[To see that this is sufficient it is helpful to notice that the first term on the right side of (15.9) is $\leq C\lambda_r(Q)$ when $(x,t) \in \widehat{Q}$.] To verify (15.12) it is useful to introduce a collection of cubes which is like the set of minimal cubes in S except that it is more regular, in a certain sense, and it is easier to compute $d(x)$ in terms of this collection.

Let $\widetilde{Q}(S)$ denote the union of the cubes lying in the same generation Δ_j of cubes as $Q(S)$ and which intersect $2Q(S)$. Let $n(S)$ denote the cubes $N \subseteq \widetilde{Q}(S)$ which are minimal among the cubes that also satisfy

$$200C^2 \operatorname{diam} N \geq \inf_{x \in N} d(x), \tag{15.13}$$

where C is as in (2.2). If N is one of these minimal cubes, then none of its children satisfy (15.13), and so

$$100 \operatorname{diam} N \leq \inf_{x \in N} d(x).$$

This implies that

$$\sup_{x \in N} d(x) \leq 2 \inf_{x \in N} d(x),$$

since $d(x)$ is Lipschitz with norm 1. Notice that $d(y) = 0$ when $y \in \widetilde{Q}(S)$ but y does not lie in any of the N's. Of course the N's are disjoint.

Let us use these cubes N to prove (15.12). Given Q, let $\widetilde{Q}$ denote the union of the cubes in the same generation Δ_j as Q that intersect $2Q$. For $r = \infty$ we have that

$$\sum_{\substack{Q \in S \\ Q \subseteq T}} \lambda_\infty(Q)^2 |Q| \leq C \sum_{\substack{Q \in S \\ Q \subseteq T}} \left(\sup_{\substack{N \subseteq \widetilde{Q} \\ N \in n(S)}} \frac{\operatorname{diam} N}{\operatorname{diam} Q} \right)^2 |Q| \leq C \sum_{\substack{Q \in S \\ Q \subseteq T}} \sum_{\substack{N \subseteq \widetilde{Q} \\ N \in n(S)}} \frac{|N|^{2/d}}{|Q|^{2/d}} |Q|.$$

When $d = 1$ this is at most

$$C \sum_{\substack{N \subseteq \widetilde{T} \\ N \in n(S)}} \left(\sum_{Q : \widetilde{Q} \supseteq N} |Q|^{-1} \right) |N|^2 \leq C \sum_{\substack{N \subseteq \widetilde{T} \\ N \in n(S)}} |N| \leq C|T|.$$

Now suppose that $r < \infty$. We may as well assume that $r > 2$, in which case

$$\sum_{\substack{Q\in S\\ Q\subseteq T}} \lambda_r(Q)^2|Q| \leq C \sum_{\substack{Q\in S\\ Q\subseteq T}} \left(\sum_{\substack{N\subseteq\widetilde{Q}\\ N\in\overline{n}(S)}} \left(\frac{|N|}{|Q|}\right)^{\frac{r}{d}+1} \right)^{\frac{2}{r}} |Q|$$

$$\leq C \sum_{\substack{Q\in S\\ Q\subseteq T}} \sum_{\substack{N\subseteq\widetilde{Q}\\ N\in\overline{n}(S)}} \left(\frac{|N|}{|Q|}\right)^{\frac{2}{r}(\frac{r}{d}+1)} |Q|$$

$$\leq C \sum_{\substack{N\subseteq\widetilde{T}\\ N\in\overline{n}(S)}} \sum_{Q:\widetilde{Q}\supseteq N} |Q|^{1-\frac{2}{r}-\frac{2}{d}}|N|^{\frac{2}{r}+\frac{2}{d}}.$$

If $r < \frac{2d}{d-2}$, then $1 - \frac{2}{d} < \frac{2}{r}$, and so the inner sum converges and this is

$$\leq C \sum_{\substack{N\subseteq\widetilde{T}\\ N\in\overline{n}(S)}} |N| \leq C|T|.$$

This proves (15.12), which is the last step in the proof that (C4) implies that (15.1) is a Carleson measure for $r < \frac{2d}{d-2}$ when $d \geq 2$, $r \leq \infty$ when $d = 1$.

16. The main step in the proof that (C4) implies (C5)

PROPOSITION 16.1. *Let E be a regular set that satisfies (C4). Then for any $\epsilon > 0$ there is an $M > 0$ so that for all $x \in E$ and $r > 0$ there is a compact set $F \subseteq E \cap B(x,r)$ and a mapping $f : F \to \mathbf{R}^d$ such that*

$$|E \cap B(x,r) \setminus F| \leq \epsilon r^d \tag{16.2}$$

$$M^{-1}|x-y| \leq |f(x) - f(y)| \leq M|x-y| \qquad \text{for all } x,y \in F. \tag{16.3}$$

Once we prove this proposition all that remains in showing that (C4) implies (C5) is to find a bilipschitz extension of $f^{-1} : f(F) \to F$ to a embedding of $\mathbf{R}^d$ into $\mathbf{R}^{n*}$. This will be taken up in the next section.

To prove the proposition we first construct F, and then the definition of f will become natural. The general idea of the proof is that for a good region S in Δ as in Section 2 it is easy to choose a map of $Q(S)$ onto $\mathbf{R}^d$ that is bilipschitz at scales larger than the minimal cubes of S, because of (2.6) (the approximation of $Q(S)$ by a Lipschitz graph), and we can try to patch these maps together to get f. The total amount of patching required will be controlled by (2.7).

Let $x \in E$ and $r > 0$ be given, and let m_0 be such that $2^{m_0} \leq r < 2^{m_0+1}$. Let R_0 be the union of the cubes in Δ_{m_0} that intersect $B(x,r)$. Call these cubes $(Q_{0,j})$, $j \in J$; note that there are a bounded number of them. Decompose the set of cubes $Q \in \Delta$ such that $Q \subseteq R_0$ into a bad set and a family of good regions $S \subseteq \Delta$ in such a way that (2.4), (2.5), (2.6) and (2.7) hold, with η small, to be chosen later (depending only on the geometry of E, not on ϵ). It is easy to obtain this decomposition from the one provided by (C4).

In the construction of the set F we shall sometimes need to remove from cubes the part which is too close to the boundary. For each $Q \in \Delta_m$ set

$$\sigma(Q) = \{x \in Q : \operatorname{dist}(x, E \setminus Q) \leq \tau 2^m\}, \tag{16.4}$$

where τ is a small constant that we shall choose soon (depending on ϵ). Thus $\sigma(Q)$ is the part of Q which is very close to the boundary, and the measure of $\sigma(Q)$ is controlled by (2.3).

Applying (2.3) with a not-too-small choice of τ already gives the existence, for each cube $Q \in \Delta$, of a "center" $c_Q \in Q$ such that

$$\operatorname{dist}(c_Q, E \setminus Q) \geq \frac{1}{C_0} \operatorname{diam} Q, \tag{16.5}$$

with C_0 independent of ϵ.

We still need a little more notation before defining the set F. We call a cube $Q \subseteq R_0$ a "transition cube" if it is a bad cube, the top cube in one of the good regions S, or a minimal cube in one of the S's. We denote by T the set of transition cubes, and by $T_j = T \cap \Delta_j$ the ones of generation j. For each transition cube Q, let $\ell(Q)$ denote the number of ancestors of Q that are transition cubes. For instance, the $Q_{0,j}$'s are transition cubes such that $\ell(Q_{0,j}) = 0$.

Our set F is defined by

$$F = E \cap B(x,r) \cap \left(\bigcup_{Q \in T} \sigma(Q) \right)^c \cap \left(\bigcup_{\substack{Q \in T \\ \ell(Q) \geq L}} Q \right)^c, \tag{16.6}$$

where L is a very large constant whose value will be chosen very soon.

Notice that

$$\sum_{Q \in T} |Q| \leq C|R_0|,$$

because of (2.4) and (2.7). Using (2.3) we can choose τ so small that

$$\sum_{Q \in T} |\sigma(Q)| \leq C\tau^{\frac{1}{C}} |R_0| < \epsilon r^d/2.$$

Similarly, if L is large enough, then $A = \bigcup_{\substack{Q\in T\\ \ell(Q)\geq L}} Q$ satisfies $|A| < \epsilon r^d/2$, because $\sum_{Q\in T} \chi_Q \geq L$ on A, so that

$$|A| \leq L^{-1} \sum_{Q\in T} |Q| \leq CL^{-1}|R_0|.$$

Thus F satisfies (16.2).

As a first approximation of the map f on F we shall define a function g on the set T of transition cubes with values in the set of cubes in $\mathbf{R}^d$. Our mapping will have the following properties:

$$\text{for each } Q \in T_m, g(Q) \text{ is a cube of } \mathbf{R}^d \text{ with diameter } C_1^{-\ell(Q)}2^m, \text{ where } C_1 \text{ is a large constant (to be specified),} \tag{16.7}$$

$$\text{if } Q,\ Q' \in T,\ Q \subseteq Q', \text{ then } g(Q) \subseteq \tfrac{1}{2}g(Q'). \tag{16.8}$$

Let us start with the cubes $Q_{0,j}$, $j \in J$. We pick the cubes $g(Q_{0,j})$ to be cubes of size 2^{m_0}, at mutual distances between 2^{m_0} and $C2^{m_0}$. This is possible if C is large enough.

In general we define g recursively as follows. Suppose $g(Q)$ has already been defined for some $Q \in T_m$. Assume first that Q is a bad cube or the minimal cube of some good region S. Let $\mathcal{C}(Q)$ denote the set of children in Q, i.e., the set of cubes in Δ_{m-1} which are subcubes of Q. Notice that they are all either bad cubes or top cubes of some other good region, and that $\ell(R) = \ell(Q)+1$ if $R \in \mathcal{C}(Q)$. If C_1 is sufficiently large, we can easily choose cubes $g(R)$, $R \in \mathcal{C}(Q)$, such that each $g(R)$ has diameter $C_1^{-\ell(Q)-1}2^{m-1}$, is contained in $\frac{1}{2}g(Q)$, and also

$$\operatorname{dist}(g(R), g(R')) \geq 2^{m-1}C_1^{-\ell(Q)-1} \tag{16.9}$$

for all R, $R' \in \mathcal{C}(Q)$, $R \neq R'$

Assume now that Q is the top cube of a good region S. We want to define $g(R)$ for the minimal cubes R of S, but first we need some more notation. Let Γ_Q be as in (2.6), so that Γ_Q is a Lipschitz graph over a d-plane P_Q, and let Π_Q be the orthogonal projection onto P_Q. Let ϕ_Q be an affine mapping from P_Q to $\mathbf{R}^d$ such that $\phi_Q(\Pi_Q(Q)) \subseteq \frac{1}{3}g(Q)$ and

$$C^{-1}C_1^{-\ell(Q)}|p-q| \leq |\phi_Q(p) - \phi_Q(q)| \leq C_1^{-\ell(Q)}|p-q| \tag{16.10}$$

for all p, $q \in P_Q$.

If R is a minimal cube of S, and m is such that $R \in T_m$, then we take $g(R)$ to be a cube centered at $\phi_Q(\Pi_Q(c_R))$ with diameter $C_1^{-\ell(Q)-1}2^m$, where c_R is the "center" of R that we chose before (so that (16.5) holds). Clearly $g(R) \subseteq \frac{1}{2}g(Q)$.

We need to estimate $\operatorname{dist}(g(R), g(R'))$ for any pair R, R' of minimal cubes of S. Observe that

$$|c_R - c_{R'}| \geq C^{-1}\{\operatorname{diam} R + \operatorname{diam} R' + \operatorname{dist}(R, R')\},$$

because of (16.5). Applying (2.6) to the smallest cube $Q' \subseteq Q$ that contains c_R and has diameter $\geq |c_R - c_{R'}|$ we get

$$\operatorname{dist}(c_R, \Gamma_Q) + \operatorname{dist}(c_{R'}, \Gamma_Q) \leq 2\eta \operatorname{diam} Q' \leq C\eta |c_R - c_{R'}|.$$

If η is small enough this implies that

$$|\Pi_Q(c_R) - \Pi_Q(c_{R'})| \geq \frac{1}{2}|c_R - c_{R'}|,$$

since Γ_Q is a Lipschitz graph with constant $\leq \eta$ over P_Q. Combining these estimates with (16.10) we see that if C_1 is large enough, then

$$\begin{aligned} &C_1^{-\ell(Q)-1}\{\operatorname{diam} R + \operatorname{diam} R' + \operatorname{dist}(R, R')\} \\ &\leq \operatorname{dist}(g(R), g(R')) \\ &\leq C_1^{-\ell(Q)}\{\operatorname{diam} R + \operatorname{diam} R' + \operatorname{dist}(R, R')\}. \end{aligned} \tag{16.11}$$

Now that we have constructed g we are ready to define f on F. If $x \in F$, then it can only belong to a finite number of transition cubes; let $Q(x)$ be the smallest one. By the definitions of F and transition cubes, $Q(x)$ is the top cube of some good region $S(x)$. [Otherwise all the children of $Q(x)$ are also transition cubes, and they are disjoint from F.] We take

$$f(x) = \phi_{Q(x)}\left(\Pi_{Q(x)}(x)\right),$$

with the notations above. Note that $f(x) \in g(Q(x))$.

We want to check the bilipschitz condition (16.3). Let x, $y \in F$ be given, and let Q_1 be the largest cube in R_0 that contains x but not y.

Suppose first that Q_1 lies in a stopping time region S, but is not its top cube $Q(S)$. In this case $y \in Q(S)$ too. Suppose further that there are two minimal cubes R, R' of S containing x, y, respectively. Of course

$$|x - y| \leq \operatorname{diam} R + \operatorname{diam} R' + \operatorname{dist}(R, R'),$$

but because $x \notin \sigma(R)$, $y \notin \sigma(R')$ (by definition of F), we also have

$$|x-y| \geq \frac{\tau}{C}\operatorname{diam} R + \frac{\tau}{C}\operatorname{diam} R' + \frac{1}{2}\operatorname{dist}(R, R').$$

For this pair x, y (16.3) follows from these inequalities, (16.11), and from $f(x) \in g(R)$, $f(y) \in g(R')$ and $\ell(R) = \ell(R') \leq L$.

If x is not contained in any minimal cube of S, then $f(x) = \phi_Q(\Pi_Q(x))$, where $Q = Q(S)$ is the top cube of S. The argument used to establish (16.11) also gives

$$\begin{aligned} C_1^{-\ell(Q)-1}&\{\operatorname{diam} R' + \operatorname{dist}(x, R')\} \\ &\leq \operatorname{dist}(f(x), g(R')) \\ &\leq C_1^{-\ell(Q)}\{\operatorname{diam} R' + \operatorname{dist}(x, R')\} \end{aligned}$$

for such an x, and (16.3) is deduced as before. The same argument works if y is not contained in a minimal cube of S, or if neither x nor y is.

We are left now with the case where Q_1 is a transition cube which is not a minimal cube of some good region. This means that the father Q_2 of Q_1 is also a transition cube, and it is not the top cube of some good region. Because Q_2 contains y, we can bound $|f(x)-f(y)|$ from above using $f(x)$, $f(y) \in g(Q_2)$, and we can bound it from below by applying (16.9) to $R = Q_1$ and R' = the brother of Q_1 that contains y [since $f(x) \in g(R)$, $f(y) \in g(R')$]. These bounds imply (16.3), because

$$\frac{\tau}{C}\operatorname{diam} Q_1 \leq |x-y| \leq C \operatorname{diam} Q_1$$

(since $x \notin \sigma(Q_1)$, $y \in Q_2 \setminus Q_1$).

Thus we have shown that F and f satisfy (16.2) and (16.3). If F happens not to be closed, then we can replace it by its closure. This completes the proof of Proposition 16.3.

17. An extension theorem

PROPOSITION 17.1. *Let $A \subseteq \mathbf{R}^d$ be a closed set. Suppose that $f : A \to \mathbf{R}^n$ satisfies*

$$C_0^{-1}|x-y| \le |f(x)-f(y)| \le C_0|x-y| \tag{17.2}$$

for all $x, y \in A$. Assume also that $n \ge 2d+1$. Then there is an extension $g : \mathbf{R}^d \to \mathbf{R}^n$ of f such that

$$M^{-1}|x-y| \le |g(x)-g(y)| \le M|x-y| \tag{17.3}$$

for all $x, y \in \mathbf{R}^d$. Here M depends only on d, n, and C_0.

That (C4) implies (C5) follows from this and Proposition 16.1. (Of course the f here corresponds to the f^{-1} from Proposition 16.1.)

It is not clear that the condition $n \ge 2d+1$ is optimal. If we want g to be a small perturbation of a Lipschitz extension of f that is given in advance, then we do have to have $n > 2d$.

The proposition is not hard to prove using the techniques of **[D3]**, Section 4. (Section 5 of **[D3]** is relevant to the question of the optimality of $n \ge 2d+1$.) Because the proof requires a rather large amount of notations, and no new ideas, we shall omit it here, and content ourselves with a slightly weaker result.

PROPOSITION 17.4. *Let $A \subseteq \mathbf{R}^d$ be closed, and suppose that $f : A \to \mathbf{R}^n$ satisfies (17.2). Then there exist an integer m, a constant M, and an extension $g : \mathbf{R}^d \to \mathbf{R}^m$ of f such that (17.3) holds. The constants m, M depend only on d, n, and C_0.*

This result, together with Proposition 16.1, still prove that (C4) implies (C5), only with a different value of n^*. It is of course still possible to deduce (C1) from this weaker version of (C5).

Let us prove the proposition. By the Whitney extension theorem we know that we can extend f to a map from $\mathbf{R}^d$ to $\mathbf{R}^n$ which is Lipschitz with norm $\leq C_0'$, where C_0' depends only on C_0 and d, and $C_0' \geq C_0$. Let f denote this extension also.

We also need a Whitney decomposition of $\Omega = \mathbf{R}^d \setminus A$. Let Q_i, $i \in I$, be the maximal dyadic cubes in Ω with

$$\operatorname{diam} Q_i \leq 10^{-3} \operatorname{dist}(Q_i, A).$$

For each $i \in I$ choose $\phi_i : \mathbf{R}^d \to \mathbf{R}^{d+1}$ such that:

$$\text{(17.5)} \quad \phi_i \text{ is supported in } 3Q_i, |\phi_i| \leq C_1 \operatorname{diam} Q_i, \text{ and } |\nabla \phi_i| \leq C_1;$$

$$\text{(17.6)} \qquad |\phi_i(x) - \phi_i(y)| \geq C_1^{-1}|x-y| \quad \text{for } x, y \in 2Q_i;$$

$$\text{(17.7)} \qquad |\phi_i(x)| \leq \frac{1}{2} \operatorname{diam} Q_i, \qquad \text{for } x \notin 2Q_i;$$

$$\text{(17.8)} \qquad |\phi_i(x)| \geq \operatorname{diam} Q_i, \qquad \text{for } x \in Q_i.$$

The constant C_1 is a geometric constant that depends only on d.

It is quite easy to construct these functions ϕ_i. We added one more coordinate so that (17.8) can be satisfied too.

The idea of the proof is to add the functions ϕ_i to f, to increase $f(x) - f(y)$ when necessary. We shall put the functions ϕ_i in different dimensions, so that they never interfere with each other, or with f.

For each $i \in I$ let $V(i)$ denote the set of indices $j \in I$ such that

$$\text{(17.9)} \qquad 10^{-2} \operatorname{diam} Q_i \leq \operatorname{diam} Q_j \leq 10^2 \operatorname{diam} Q_i, \text{ and}$$

$$\text{(17.10)} \qquad \operatorname{dist}(Q_i, Q_j) \leq 10^6 (C_0')^2 \{\operatorname{diam} Q_i + \operatorname{diam} Q_j\}.$$

Thus $j \in V(i)$ if and only if $i \in V(j)$, and there is an integer N such that $V(i)$ has less than N elements for each i. Let X be a set with N elements. There is a function $a : I \to X$ such that $a(i) \neq a(j)$ whenever $j \in V(i)$ but $j \neq i$. [One way to find such a function is to arrange the elements of I in a sequence and then define $a(i)$ recursively in such a way that this property always holds.]

Set $m = n + 1 + (d+1)N$, and identify $\mathbf{R}^m$ with $\mathbf{R}^n \times \mathbf{R} \times (\mathbf{R}^{d+1})^X$. We define $g : \mathbf{R}^d \to \mathbf{R}^m$ by

$$\text{(17.11)} \qquad g_0 = f, \quad \text{where } g_0 \text{ is the } \mathbf{R}^n \text{ coordinate of } g,$$

(17.12) $g_1(x) = \operatorname{dist}(x, A)$, where g_1 denotes the $\mathbf{R}$-coordinate of g, and

$$g_\alpha(x) = \sum_{\substack{i\in I\\ a(i)=\alpha}} \phi_i(x), \quad \alpha \in X \tag{17.13}$$

Clearly g is M-Lipschitz for some M. To check the other part of (17.3) let x, $y \in \mathbf{R}^d$ be given, $x \neq y$, and let us distinguish a few cases.

Suppose first that x and y both lie in Ω, $x \in Q_i$, $y \in Q_j$. We may assume that $\operatorname{diam} Q_i \geq \operatorname{diam} Q_j$.

<u>Case 1</u> $\operatorname{dist}(Q_i, Q_j) \geq 10^5 (C_0')^2 \operatorname{diam} Q_i$.

Thus Q_i and Q_j are much further away from each other than they are from A (by the way we chose the Q_i's). We shall use (17.2) to find a lower bound for $|g_0(x) - g_0(y)|$. Choose u, $v \in A$ as close as possible to x, y, so that

$$|x-u| \leq 3\cdot 10^3 \operatorname{diam} Q_i, \ |y-v| \leq 3\cdot 10^3 \operatorname{diam} Q_j \leq 3\cdot 10^3 \operatorname{diam} Q_i.$$

Then $|u-v| \geq \frac{1}{2}|x-y|$, and

$$\begin{aligned} |f(x)-f(y)| &\geq |f(u)-f(v)| - |f(x)-f(u)| - |f(y)-f(v)| \\ &\geq C_0^{-1}|u-v| - C_0'|x-u| - C_0'|y-v| \\ &\geq \frac{1}{2}C_0^{-1}|x-y| - 6\cdot 10^3 C_0' \operatorname{diam} Q_i. \end{aligned}$$

By assumption the last term is at most

$$\frac{6}{100}(C_0')^{-1}\operatorname{dist}(Q_i, Q_j) \leq \frac{6}{100}(C_0)^{-1}|x-y|,$$

and so $|f(x)-f(y)| \geq (10C_0)^{-1}$, as desired.

<u>Case 2</u> $\operatorname{dist}(Q_i, Q_j) \leq 10^5 (C_0')^2 \operatorname{diam} Q_i$, and $\operatorname{diam} Q_j \leq \frac{1}{5}\operatorname{diam} Q_i$.

Thus $\operatorname{dist}(Q_i, Q_j)$ is not too large compared to the distance of Q_i to A, but Q_j is closer to A than Q_i is. In this case we have

$$\begin{aligned} &g_1(x) \geq \operatorname{dist}(Q_i, A) \geq 10^3 \operatorname{diam} Q_i, \\ &g_1(y) \leq \operatorname{dist}(Q_j, A) + \operatorname{diam} Q_j \leq 3\cdot 10^3 \operatorname{diam} Q_j \leq \frac{3}{5}10^3 \operatorname{diam} Q_i, \end{aligned}$$

by definition of our Whitney cubes. This provides an adequate lower bound on $|g(x)-g(y)|$, because $|x-y| \leq 2\cdot 10^5 (C_0')^2 \operatorname{diam} Q_i$, by assumption.

Case 3 $\operatorname{dist}(Q_i, Q_j) \leq 10^5 (C_0')^2 \operatorname{diam} Q_i$, and $\operatorname{diam} Q_j > \frac{1}{5} \operatorname{diam} Q_i$.

Thus the distances between Q_i, Q_j, and A are roughly comparable to each other. Under these conditions we have $k \in V(i)$ if $3Q_k$ meets Q_j; thus $g_{a(i)}(y) = \phi_i(y)$, because of (17.13) and the properties of $a(\cdot)$. We have $g_{a(i)}(x) = \phi_i(x)$ for the same reason. We can bound $|g_{a(i)}(x) - g_{a(i)}(y)|$ from below using (17.6) when $y \in 2Q_i$, and (17.7) and (17.8) when $y \notin 2Q_i$. This lower bound is good enough because $|x - y| \leq 2 \cdot 10^5 (C_0')^2 \operatorname{diam} Q_i$, by assumption.

These three cases take care of the situation where x, y both lie in Ω. If $x \in \Omega$ and $y \in A$, then we can use the same arguments as in Cases 1 and 2 above, depending on whether $\operatorname{dist}(y, Q_i)$ is larger or smaller than $10^5 (C_0')^2 \operatorname{diam} Q_i$. When both x and y lie in A we can apply (17.2) directly. This completes the proof of Proposition 17.4.

18. The proof that (C4) implies (C7)

The idea, roughly speaking, is that the corona decomposition permits us to realize E as a subset of a countable union of Lipschitz graphs, with a lot of control on how these graphs fit together, and we shall build our w-regular mapping by connecting up pieces of these Lipschitz graphs.

Before beginning our construction we first address a technical point that will be needed later, concerning the "connectedness" of Δ. A finite sequence $Q_1, \ldots, Q_\ell$ of cubes in Δ will be called a path if for each j, $1 \leq j \leq \ell - 1$, Q_j is either a son or the father of Q_{j+1}. A subset of Δ is said to be connected if any two points can be joined by a path in the subset.

A useful fact which is easy to verify is that Δ is connected if and only if every pair of elements has a common ancestor.

If Δ is the set of dyadic cubes in $\mathbf{R}^n$, then Δ is not connected; it has 2^n components. In general Δ will always have only finitely many components. However, it is not hard to modify Δ slightly in such a way that Δ is connected, as follows.

Let Δ be as in Section 2, and let $\widetilde{\Delta}$ be the set of subsets of E defined by $\widetilde{\Delta} = \bigcup_{j=-\infty}^{\infty} \widetilde{\Delta}_j$, where $Q \in \widetilde{\Delta}_j$ if $Q \in \Delta_j$ and Q does not intersect $B(0, 2^j)$, or if Q is the union of the $R \in \Delta_j$ that intersect $B(0, 2^j)$. Clearly each $\widetilde{\Delta}_j$ is a partition of E, and $\widetilde{\Delta}$ still satisfies (2.2) and (2.3). It is not hard to check that $\widetilde{\Delta}$ also satisfies (2.1); the main point is the trivial fact that if $Q \in \Delta(j)$, $Q' \in \Delta(k)$, $j \leq k$, $Q \subseteq Q'$, and if Q intersects $B(0, 2^j)$, then Q' intersects $B(0, 2^k)$. Of course $\widetilde{\Delta}$ is also connected.

In view of this we may as well assume that Δ is connected, since we can replace it by $\widetilde{\Delta}$ otherwise. This change in Δ will not upset the condition that E admits a corona decomposition. Indeed, we noted in Section 2 that the existence of a corona decomposition does not depend on the choice of Δ as long as (2.1) and (2.2) are satisfied. In this particular case it is even easier to transfer a corona decomposition for Δ to one for $\widetilde{\Delta}$. For example, you can take all the cubes in Δ that were changed and put them into $\mathcal{B}$ without disturbing (2.4), and then take the S's with the changed cubes removed and reorganize them into regions that satisfy (2.5), (2.6), and (2.7).

Let us now proceed with the construction of an w-regular mapping whose image contains E. Let $\eta > 0$ be small, to be chosen later. Its value will only depend on geometric constants. Let $\mathcal{B}$, $\mathcal{G}$, $\mathcal{F}$, S be as in Section 2. Given $S \in \mathcal{F}$, (2.6) tells us that there is a Lipschitz graph Γ over a d-plane $P = P_S$ which well-approximates E with respect to S. Let $\Pi = \Pi_S$ denote the orthogonal projection onto P, and let $\Gamma(S)$ denote the part of Γ whose projection onto P is the closed ball $B(\Pi(x), 20 \operatorname{diam} Q(S))$, where x is any point of $Q(S)$.

We observe first that E is contained in the union of the $\Gamma(S)$'s, $S \in \mathcal{F}$, except possibly for a set of measure zero. To see this let $Q_j(y)$ denote the element of Δ_j that contains y. Because of (2.5), $Q_j(y) \in \mathcal{B}$ for at most finitely many j's, except for a set of y's of measure zero. Similarly, (2.7) tells us that for almost all y, $Q_j(y)$ is a $Q(S)$ for only finitely many j. Hence for a.a. y there is an $S \in \mathcal{F}$ such that $y \in Q(S)$ and every $Q \subseteq Q(S)$ with $Q \ni y$ lies in S. For such an y we have $y \in \Gamma(S)$ because of (2.6).

We are going to modify the $\Gamma(S)$'s to make it easier to connect them together. We also introduce an additional dimension and work in $\mathbf{R}^{n+1}$ instead of $\mathbf{R}^n$; this will be needed to ensure that the mapping $z : \mathbf{R}^d \to \mathbf{R}^{n+1}$ that we define eventually is indeed regular. (Of course we identify $\mathbf{R}^n$ with the subspace of $\mathbf{R}^{n+1}$ where the last coordinate is zero.)

Fix a stopping time region S. In addition to the notation that we've already recalled, let $d(x)$ be as in (7.6), and let $P^\perp$ be the $(n-d)$-plane that passes through 0 and is orthogonal to P in $\mathbf{R}^n$. Let $A : P \to P^\perp$ be the function whose graph is Γ. We shall construct a surface $\Gamma_1(S)$ which is vaguely reminiscent of a pair of trousers with many legs (see Figure 1), and which is better to work with than $\Gamma(S)$. Doing this requires quite a bit of notation, unfortunately.

Let $c_1(S)$ denote the $(d-1)$-sphere in P with center $\Pi(x)$ and radius $20 \operatorname{diam} Q(S)$. (Remember that $x \in Q(S)$ is the point selected in the definition of $\Gamma(S)$.) Also let $b_1(S)$ be the closed d-ball in P bounded by $c_1(S)$, and let $a_1(S)$ be the $(d-1)$-sphere in $\mathbf{R}^{n+1}$ that is parallel to $c_1(S)$

but which is centered at

$$(\Pi(x), A(\Pi(x)),\ 2\operatorname{diam} Q(S)) \in P \times P^{\perp} \times \mathbf{R} \cong \mathbf{R}^{n+1}.$$

(Here we identify $P \times P^{\perp}$ with $\mathbf{R}^n$ via the map that sends (p, q) to $p + q$.)

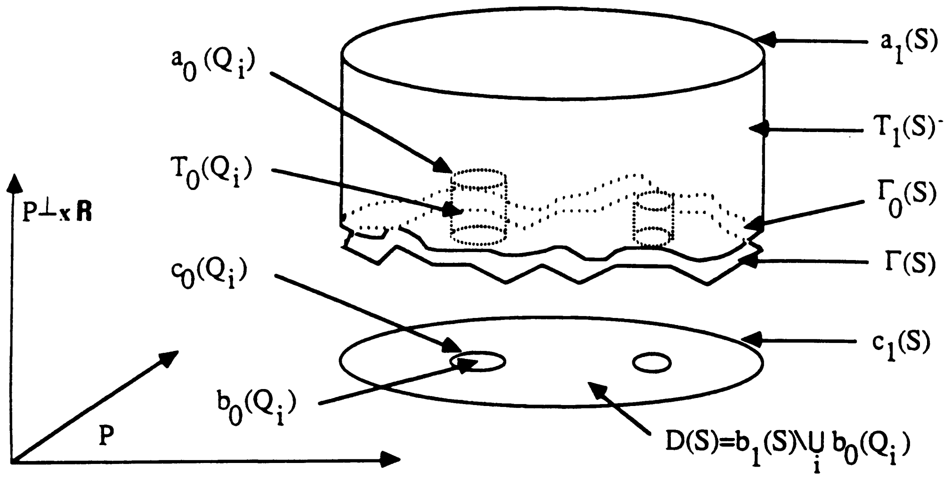

Figure 1. A symbolic picture of $\Gamma_1(S)$. The vertical direction contains both $P^{\perp}$ and the $(n+1)^{\text{th}}$-co-ordinate axis.

For each minimal cube Q of S we make the following construction. For each child Q_i of Q select a point $x_i \in Q_i$ at distance $\geq 10 \operatorname{diam} Q_i / C_0$ from $E \setminus Q_i$. We can do this if C_0 is large enough, because of (2.3). Let $c_0(Q_i)$ be the $(d-1)$-sphere contained in P with center $\Pi(x_i)$ and radius $2 \operatorname{diam} Q_i / C_0$, and let $b_0(Q_i)$ be the open ball it encloses.

Let $I(S)$ denote the set of all Q_i's, where Q runs over all minimal cubes of S. If η is small enough, we get (using (2.6)) that

$$\text{dist}\,(b_0(Q_i), b_0(Q_j)) \geq \frac{1}{C_0}\,[\text{diam}\,Q_i + \text{diam}\,Q_j] \tag{18.1}$$

$$\text{dist}\,(b_0(Q_i), c_1(S)) \geq \text{diam}\,Q(S) \tag{18.2}$$

for all $i,\, j \in I(S)$, $i \neq j$. We also let $a_0(Q_i)$ be the $(d-1)$ sphere in $\mathbf{R}^{n+1}$ that is parallel to $c_0(Q_i)$ but is centered at

$$(\Pi(x_i),\ A(\Pi(x_i)),\ 2\,\text{diam}\,Q_i).$$

Set $D(S) = b_1(S) \setminus (\bigcup_{R \in I(S)} b_0(R))$ and

$$\Gamma_0(S) = \{(u, A(u), d'(u)) : u \in D(S)\} \subseteq \mathbf{R}^{n+1},$$

where $d'(u) = d((u, A(u)))$, $d(x)$ as in (7.6). Almost every $y \in E$ lies in $\Gamma_0(S)$ for some S. Indeed, for almost every y there is an S such that $y \in Q(S)$ and such that $Q \in S$ whenever $Q \subseteq Q(S)$, $y \in Q$, and for such an S we have $y \in \Gamma$ and $d(y) = 0$. We also have $\Pi(y) \in D(S)$ because of the way we chose the $b_0(Q_i)$, and because of (2.6).

We define $\Gamma_1(S)$ by taking $\Gamma_0(S)$ and adding some pieces to it; to wit,

$$\Gamma_1(S) = \Gamma_0(S) \cup T_1(S) \cup \left(\bigcup_{Q_i \in I(S)} T_0(Q_i) \right),$$

where $T_1(S)$ and $T_0(R)$ are defined as follows. We take $T_1(S)$ to be the "tube" obtained by joining each point $(u, A(u), d'(u))$, $u \in c_1(S)$, to

$$(u, A(\Pi(x)),\ 2\,\text{diam}\,Q(S)) \in a_1(S)$$

by a straight line. Similarly, if $Q_i \in I(S)$, we take $T_0(Q_i)$ to be the tube obtained by joining $(u, A(u), d'(u))$ to

$$(u, A(\Pi(x_i)),\ 2\,\text{diam}\,Q_i) \in a_0(Q_i)$$

by a straight line for each $u \in c_0(Q_i)$.

It is not hard to see from our construction that there is a bilipschitz map of $D(S)$ onto $\Gamma_1(S)$ which sends $c_1(S)$ to $a_1(S)$ and $c_0(Q_i)$ to $a_0(Q_i)$ for all $Q_i \in I(S)$.

Next we want to do a similar (but simpler) construction for cubes Q in $\mathcal{B}$. To simplify notations we associate to each $Q \in \mathcal{B}$ a new stopping-time region $S = S(Q)$, where S has Q as its unique element (and so $Q(S) = Q$). We let $\overline{\mathcal{F}}$ denote the union of $\mathcal{F}$ with $\{S \subseteq \Delta : S = S(Q), Q \in \mathcal{B}\}$. Given $Q \in \mathcal{B}$, select any $x \in Q$ and any d-plane P that passes through x, and let $c_1(S)$ be the $(d-1)$-sphere in P with center x and radius $20 \operatorname{diam} Q$; as before, we take $b_1(S)$ to be the closed ball in P enclosed by $c_1(S)$.

Let Q_i be the children of Q, and let $I(S)$ denote the set of these Q_i. For each $Q_i \in I(S)$ choose a point $x_i \in b_1(S)$, and let $c_0(Q_i)$ be the $(d-1)$-sphere in P centered at x_i with radius $2 \operatorname{diam} Q_i/C_0$. The reader will most likely be unsurprised to learn that $b_0(Q_i)$ denotes the open d-ball enclosed by $c_0(Q_i)$. If C_0 is large enough, we can choose the x_i in such a way that (18.1) and (18.2) hold. [In this case we do not need to take the points x_i inside $\Pi(Q_i)$; when Q is a bad cube, we don't have to respect the geometry, because the geometry is bad.]

We now set $D(S) = b_1(S) \setminus (\bigcup_{Q_i \in I(S)} b_0(Q_i))$ also in this case. We view $D(S)$ both as a subset of P and as a subset of $\mathbf{R}^n$. We take $\Gamma_1(S)$ to be the translation of $D(S)$ inside $\mathbf{R}^{n+1}$ by $2 \operatorname{diam} Q(S)$ in the last variable:

$$\Gamma_1(S) = \{(u, 2 \operatorname{diam} Q(S)) \in \mathbf{R}^{n+1} : u \in D(S)\}.$$

We let $a_1(S)$ denote the "exterior" boundary of $\Gamma_1(S)$, and, for $Q_i \in I(S)$, we take $a_0(Q_i)$ to be the image of $c_0(Q_i)$ by the same translation of $2 \operatorname{diam} Q(S)$ in the last variable.

Thus, just like when $S \in \mathcal{F}$, $\Gamma_1(S)$ is a "nice" surface with boundary $a_1(S) \cup (\bigcup_{Q_i \in I(S)} a_0(Q_i))$, and there is a bilipschitz map of $D(S)$ onto $\Gamma_1(S)$.

We can now glue the various $\Gamma_1(S)$'s together without much difficulty. For $S \in \overline{\mathcal{F}}$ let $T_2(S)$ be a "tube" that joins the sphere $a_1(S)$ to $a_0(Q(S))$. In writing $a_0(Q(S))$ we must keep in mind that $Q(S)$ is not only the top cube in S, but it is also the son of a bad cube, or of a minimal cube of an element of $\mathcal{F}$. We can choose $T_2(S) \subseteq \mathbf{R}^{n+1}$ in such a way that $C^{-1} \operatorname{diam} Q(S) \leq y_{n+1} \leq C \operatorname{diam} Q(S)$ for all $y \in T_2(S)$, and also so that there is a mapping z_S of $D(S)$ onto $\Gamma_2(S) = \Gamma_1(S) \cup T_2(S)$ with the following properties:

(18.3) z_S is 1-regular, uniformly in S, and in particular, $C^{-1} \leq |\nabla z_S| \leq C$ a.e.;

(18.4) the restriction of z_S to $c_1(S)$ is an affine function that sends the sphere $c_1(S)$ onto $a_0(Q(S))$, and, for each $i \in I(S)$, the restriction of z_S to $c_0(Q_i)$ is an affine map onto $a_0(Q_i)$. These affine maps are all compositions of translations, dilations, and rotations.

A few comments are perhaps in order concerning (18.3). We can split $D(S)$ into the union of a not-too-large spherical shell near the outer boundary and the remaining subregion of $D(S)$. We can map the latter onto $\Gamma_1(S)$ in a bilipschitz fashion. We can't always map the spherical shell onto $T_2(S)$ in a bilipschitz manner, because it could be that $a_1(S)$ and $a_0(Q(S))$ intersect, but we can certainly do it using a 1-regular mapping.

Although our definition of w-regular mappings considered only the case of maps defined on all of $\mathbf{R}^d$, it can easily be extended to subdomains of $\mathbf{R}^d$. In the case of 1-regular maps, for example, one has to decide if the condition $|\nabla z| \leq C$ a.e. should be replaced by the requirement that z be Lipschitz on its domain. Fortunately in our case this is not an issue; we can certainly build z_S so that it is Lipschitz, and in any case $\nabla z \in L^\infty$ implies z is Lipschitz on a domain like $D(S)$.

Note that (18.3) implies that $T_2(S)$ is chosen so that $\Gamma_2(S)$ satisfies the regularity condition

$$C^{-1}R^d \leq |B(x,R) \cap \Gamma_2(S)| \leq CR^d$$

whenever $x \in \Gamma_2(S)$ and $0 < R \leq \operatorname{diam} Q(S)$, and that $\operatorname{diam} \Gamma_2(S) \leq C \operatorname{diam} Q(S)$.

Set $\widetilde{E} = \bigcup_{S \in \overline{\mathcal{F}}} \Gamma_2(S)$. We certainly have that $E \setminus \widetilde{E}$ has measure 0; we already checked this for $E \setminus (\bigcup_{S \in \overline{\mathcal{F}}} \Gamma_0(S))$. Before constructing a parameterization of $\widetilde{E}$ let us check that it satisfies (1.5).

We begin by showing that for any $x \in \mathbf{R}^{n+1}$ and $R > 0$ we have

$$|B(x,R) \cap \widetilde{E}| \leq CR^d, \tag{18.5}$$

starting with the case where $x_{n+1} = 0$.

Let $S \in \overline{\mathcal{F}}$ be such that $\Gamma_2(S) \cap B(x,R) \neq \emptyset$. Assume first that $S = S(Q)$ for some $Q \in \mathcal{B}$. Then $\operatorname{diam} Q \leq CR$; otherwise, we have $y_{n+1} > R$ for all $y \in \Gamma_2(S)$, so that $y \notin B(x,R)$. This implies that

$Q \subseteq B(x, CR)$, and so

$$|B(x,R) \cap \left(\bigcup_{S \in \overline{\mathcal{F}} \setminus \mathcal{F}} \Gamma_2(S) \right)| \leq \sum_{\substack{Q \in \mathcal{B} \\ Q \subseteq B(x,CR)}} |\Gamma_2(S)| \leq \sum_{\substack{Q \in \mathcal{B} \\ Q \subseteq B(x,CR)}} C|Q| \leq CR^d,$$

by (2.4).

Similarly, we can bound by CR^d the total mass of the $\Gamma_2(S)$'s for which $S \in \mathcal{F}$, $Q(S) \subseteq B(x, CR)$. This leaves the $S \in \mathcal{F}$ such that $\Gamma_2(S) \cap B(x, R) \neq \emptyset$ but $Q(S) \not\subseteq B(x, CR)$. For each such S we certainly have

$$|\Gamma_2(S) \cap B(x,R)| \leq CR^d;$$

let us show that there is only a bounded number of these S's. To do this it suffices to show that there is a $Q \in S$ such that $\operatorname{dist}(x, Q) \leq CR$ and $R \leq \operatorname{diam} Q \leq CR$, because there are only a bounded number of such Q's, and because the S's are disjoint subsets of Δ.

For one of these S's we must have $\operatorname{diam} Q(S) \geq R$, since otherwise $Q(S) \subseteq B(x, CR)$ (if C is large enough). Because of this we are reduced to finding $Q \in S$ with $\operatorname{dist}(x, Q) \leq CR$ and $\operatorname{diam} Q \leq CR$; we can replace Q by an ancestor if necessary to get $\operatorname{diam} Q \geq R$. We may as well assume that $\operatorname{diam} Q(S) \geq CR$, since otherwise we can take $Q = Q(S)$.

To find such a Q we use the assumption that $\Gamma_2(S) \cap B(x, R) \neq \emptyset$. If $R^{-1} \operatorname{diam} Q(S)$ is large enough we have $T_2(S) \cap B(x, R) = \emptyset$, and so $B(x, R)$ must intersect $\Gamma_1(S)$. Suppose that $B(x, R)$ intersects $\Gamma_0(S)$, and that $y \in \mathbf{R}^{n+1}$ lies in the intersection. Write $y = (y', y_{n+1})$, $y' \in \mathbf{R}^n$. Then $|y_{n+1}| \leq R$, and $y_{n+1} = d(y')$, where $d(\cdot)$ is as in (7.6) (for this S). Thus $d(y') \leq R$, $y' \in B(x, R) \cap \mathbf{R}^n$, and the definition of $d(y')$ provides us with the Q that we want.

Suppose now that there is a $y = (y', y_{n+1})$ in $B(x, R) \cap T_1(S)$. By definition $T_1(S)$ is a union of line segments, and there is a $\lambda \in [0, 1]$ such that

$$y = \lambda(u, A(u), d'(u)) + (1 - \lambda)(u, A(\Pi(x_S)),\ 2 \operatorname{diam} Q(S))$$

for some $u \in c_1(S)$, where x_S denotes the point selected in the definition of $\Gamma(S)$ (which we called x before). Because $|y_{n+1}| \leq R$ and $R^{-1} \operatorname{diam} Q(S)$ is big, we must have that $d'(u) \leq R$ and that

$$1 - \lambda \leq R(2 \operatorname{diam} Q(S))^{-1}.$$

This implies that $|y' - (u, A(u))| \leq CR$, by using the fact that

$$|A(u) - A(\Pi(x_S))| \leq 20 \operatorname{diam} Q(S)$$

in the above formula for y. Since $d'(u) = d(u, A(u)) \leq R$ and $d(\cdot)$ is Lipschitz, we have that $d(y') \leq CR$, and so we are in the same situation as before.

The only remaining possibility is that there is a $y = (y', y_{n+1})$ in $B(x,R) \cap T_0(Q_i)$ for some $Q_i \in I(S)$. By definition y lies on the line segment that joins $(u, A(u), d'(u))$ to $(u, A(\Pi(x_i)), 2 \operatorname{diam} Q_i)$ for some $u \in c_0(Q_i)$. If $\operatorname{diam} Q_i \leq CR$, then we can take $Q = Q_i$; if not, the same argument as the one we just used implies that $d(y') \leq CR$, and this again provides us with the sort of Q that we want.

Combining all these cases we get that (18.5) holds when $x_{n+1} = 0$. Let us show that this is still true if $x_{n+1} \neq 0$. We may as well assume that $x_{n+1} > 0$, since $y_{n+1} \geq 0$ whenever $y \in \widetilde{E}$. We can also require that $x_{n+1} > 2R$, since the other possibility can easily be reduced to the case where $x_{n+1} = 0$. Write $x = (x', x_{n+1})$.

Let $S \in \overline{\mathcal{F}}$ be such that $\Gamma_2(S)$ intersects $B(x,R)$. If $S = S(Q)$ for some $Q \in \mathcal{B}$, then $\operatorname{diam} Q \sim x_{n+1}$, and $\operatorname{dist}(x', Q) \leq C x_{n+1}$, and there are only a bounded number of Q's like that. If $S \in \mathcal{F}$, then $\operatorname{diam} Q(S) \geq C^{-1} x_{n+1}$, and an argument like the one we gave in the $x_{n+1} = 0$ case shows that there can only be a bounded number of these S's as well. (For each such S you prove that there is a Q in S with $\operatorname{diam} Q \sim x_{n+1}$ and $\operatorname{dist}(x', Q) \leq C x_{n+1}$.) Since

$$|B(x,R) \cap \Gamma_2(S)| \leq CR^d$$

for any $S \in \overline{\mathcal{F}}$, we get that (18.5) is true in this case also. Altogether now we have proved (18.5) for all $x \in \mathbf{R}^{n+1}$.

Actually, our proof gives a bit more: for each $x \in \mathbf{R}^{n+1}$ and $R > 0$ we have

$$\sum_{S \in \overline{\mathcal{F}}} |B(x,R) \cap \Gamma_2(S)| \leq CR^d. \tag{18.6}$$

This controls the overlapping of the $\Gamma_2(S)$'s.

Let's check the lower bound, i.e., that

$$|\widetilde{E} \cap B(x,R)| \geq C^{-1} R^d \tag{18.7}$$

whenever $x \in \widetilde{E}$, $R > 0$.

Notice that $\{y \in E : y_{n+1} = 0\}$ is exactly E modulo a set of measure zero. Indeed, we have

$$\widetilde{E} \cap \mathbf{R}^n = \bigcup_{S \in \overline{\mathcal{F}}} (\Gamma_0(S) \cap \mathbf{R}^n)$$

and $\Gamma_0(S) \cap \mathbf{R}^n \subseteq E$ by definition, and we have already seen that $E \setminus \widetilde{E}$ has measure zero. From this it follows that (18.7) holds if $x_{n+1} = 0$, since we are assuming that E satisfies (1.5).

Suppose now that $x = (x', x_{n+1})$, $x_{n+1} > 0$. Choose $S \in \overline{\mathcal{F}}$ so that $x \in \Gamma_2(S)$. By construction we have

$$x_{n+1} + \operatorname{dist}(x', Q(S)) \leq C \operatorname{diam} Q(S).$$

If R is large enough compared to $\operatorname{diam} Q(S)$, then there is a point $y \in Q(S)$ such that $y \in B(x, \frac{1}{2}R)$. Hence $B(x, R) \supseteq B(y, \frac{1}{2}R)$, and since $y \in E$ and $\widetilde{E} \setminus E$ has measure zero, we can derive (18.7) from (1.5) again. If $R \leq C \operatorname{diam} Q(S)$, then we have

$$|B(x, R) \cap \widetilde{E}| \geq |B(x, R) \cap \Gamma_2(S)| \geq C^{-1} R^d.$$

This completes the proof that $\widetilde{E}$ satisfies (1.5). We now construct our parameterization $z(\cdot)$ of $\widetilde{E}$.

For each $S \in \overline{\mathcal{F}}$ we are going to choose a mapping $h_S : D_S \to \mathbf{R}^d$ which is the composition of a translation, rotation, and a dilation. We shall define $z(\cdot)$ on $h_S(D(S))$ by $z(x) = z_S \circ h_S^{-1}(x)$, where $z_S : D(S) \to \Gamma_2(S)$ is the map we chose earlier (satisfying (18.3) and (18.4)). We need to choose the h_S's in such a way that we can do this coherently.

Fix $S_0 \in \overline{\mathcal{F}}$, and let h_{S_0} be any map of $D(S_0)$ into $\mathbf{R}^d$ that is the composition of a translation, rotation, and a nontrivial dilation. For each $Q_i \in I(S_0)$ there is exactly one choice of $h_{S(Q_i)}$ [$S(Q_i)$ is the element of $\overline{\mathcal{F}}$ whose top cube is Q_i] such that $h_{S_0}(c_0(Q_i)) = h_{S(Q_i)}(c_1(S(Q_i)))$ and such that $z_{S(Q_i)} \circ h_{S(Q_i)}^{-1}$ agrees with $z_{S_0} \circ h_{S_0}^{-1}$ on that set. By repeating this procedure we choose h_S for all $S \in \overline{\mathcal{F}}$ with $Q(S) \subseteq Q(S_0)$. We can also run this process backwards to choose h_S for all $S \in \overline{\mathcal{F}}$ with $Q(S) \supseteq Q(S_0)$, and then run it forwards again to get all $S \in \overline{\mathcal{F}}$. We do indeed reach all $S' \in \overline{\mathcal{F}}$, because of the connectedness of Δ.

Set $\Omega = \bigcup_{S \in \overline{\mathcal{F}}} h_S(D_S)$. (Remember that each D_S is closed, by construction.) Thus $\Omega \subseteq \mathbf{R}^d$, and we can define $z : \Omega \to \mathbf{R}^{n+1}$ by $z(x) = z_S \circ h_S^{-1}(x)$ if $x \in h_S(D(S))$. This is well-defined by construction. In particular, the interiors of the $h_S(D_S)$ are pairwise disjoint.

We need to understand Ω better. By definitions, $h_S(D_S) = \beta_S \setminus \gamma_S$, where β_S is a closed ball and γ_S is a countable union of open balls. One can check that $\gamma_S \subseteq \frac{1}{10}\beta_S$; this follows from the corresponding inclusion back on D_S, and the fact that h_S does not distort relative size. Hence if B is a ball in γ_S, $10B \subseteq \beta_S$. Our construction also gives that for S, $S' \in \overline{\mathcal{F}}$,

$$\begin{aligned} &\beta_S \subseteq \beta_{S'} \text{ when } Q(S) \subseteq Q(S'), \\ &\beta_S \cap \beta_{S'} = \emptyset \text{ when } Q(S) \cap Q(S') = \emptyset. \end{aligned} \tag{18.8}$$

These observations imply that $\beta_S \subseteq \frac{1}{10}\beta_{S'}$ when S, $S' \in \overline{\mathcal{F}}$, $Q(S) \subseteq Q(S')$, and $S \neq S'$, and also that

$$10\beta_S \subseteq \beta_{S'}. \tag{18.9}$$

It now follows that $\bigcup_{S \in \overline{\mathcal{F}}} \beta_S = \mathbf{R}^d$. Indeed, if S_j is any sequence in $\overline{\mathcal{F}}$ such that $Q(S_j) \subseteq Q(S_{j+1})$, $S_j \neq S_{j+1}$, for each j, then $\beta_{S_j} \supseteq (10)^{j-1}\beta_{S_1}$.

Next let us show that $M = \mathbf{R}^d \setminus \Omega$ has measure zero. Let $S \in \overline{\mathcal{F}}$ be given. Let $S_{1,k}$ be an enumeration of the $S(Q)$'s with $Q \in I(S)$. Repeat this process; if we have chosen $S_{j,k} \in \overline{\mathcal{F}}$, $k = 1, \dots, N(j)$, for a given j, we let $S_{j+1,\ell}$ be an enumeration of the $S(Q)$'s that arise from $Q \in I(S_{j,k})$, $k = 1, \dots, N(j)$. This process may terminate after a finite number of steps. Set $\beta_{j,k} = \beta_{S_{j,k}}$.

From (18.9) we obtain

$$\sum_k |\beta_{j+1,k}| \leq \frac{1}{10} \sum_k |\beta_{j,k}|. \tag{18.10}$$

In particular, $\bigcap_j \left(\bigcup_k \beta_{j,k} \right)$ has measure zero. This set contains $M \cap \beta_S$, and so we conclude that M has measure zero, since S is arbitrary.

We can extend $z(\cdot)$ so that it is defined on all of $\mathbf{R}^d$, not just Ω. Fix $x \in M$, and choose $S \in \overline{\mathcal{F}}$ so that $x \in \beta_S$. Let $S_{j,k}$ be as above, and for $j = 1, 2, \dots$, let $k(j)$ be the index such that $x \in \beta_{j,k(j)}$. (Notice that there must be infinitely many $S_{j,k}$'s that contain x if $x \in M$.) Then $\overline{Q(S_{j,k(j)})}$ is a decreasing sequence of closed sets whose diameters tend to zero, and so their intersection consists of a single point, which we take to be $z(x)$.

Observe that $z(M)$ has measure zero; if $y \in z(M)$, then $y \in \overline{Q(S)}$ for infinitely many $S \in \overline{\mathcal{F}}$, and the set of such y has measure zero because of (2.4) and (2.7).

Let us check that this extension of $z(\cdot)$ to $\mathbf{R}^d$ is continuous. By construction,

$$\operatorname{diam}\Gamma_2(S) \leq C \operatorname{diam} Q(S), \operatorname{dist}(\Gamma_2(S), Q(S)) \leq C \operatorname{diam} Q(S). \tag{18.11}$$

Combining this with (18.8) it is easy to see that

$$\operatorname{diam} z(\beta_S) \leq C \operatorname{diam} Q(S), \quad \operatorname{dist}(z(\beta_S), Q(S)) \leq C \operatorname{diam} Q(S). \tag{18.12}$$

From this it follows easily that $z(\cdot)$ is continuous even at points of M.

This is a pretty good time to show that $z(\mathbf{R}^d) \supseteq E$. We already know that $z(\Omega) = \widetilde{E}$ contains almost every point in E; the same argument shows that $z(\beta_S \setminus M) = \cup\{\Gamma_2(S') : Q(S') \subseteq Q(S)\}$ contains almost all points in $Q(S)$. Hence $z(\overline{\beta}_S)$ contains all points with positive lower density in $Q(S)$, because $z(\overline{\beta}_S)$ is compact. From here and (1.5) it follows easily that $z(\mathbf{R}^d) = \cup z(\overline{\beta}_S) \supseteq E$.

It remains to show that $z(\cdot)$ is ω-regular for some A_1-weight ω on $\mathbf{R}^d$. By definitions $z(\cdot)$ is locally Lipschitz on Ω, and so $\nabla z(x)$ is defined almost everywhere.

Let us show that for any $x \in \widetilde{E}$ and any $R > 0$ we have

$$C^{-1}|B(x,R) \cap \widetilde{E}| \leq \int_{z^{-1}(B(x,R))} |\nabla z(y)|^d dy \leq C|B(x,R) \cap \widetilde{E}|. \tag{18.13}$$

Notice first that we don't have to worry about M or $z(M)$, since they both have measure zero. We know that the analogue of (18.13) holds for z, $\widetilde{E}$ replaced by $z_S \circ h_S^{-1}$, $\Gamma_2(S)$, because each z_S is 1-regular, and because h_S is just a composition of a translation, rotation, and a dilation. Hence

$$\begin{aligned}\int_{z^{-1}(B(x,R))} |\nabla z(y)|^d dy &= \sum_{S \in \overline{\mathcal{F}}} \int_{z^{-1}(B(x,R)) \cap (\beta_S \setminus \gamma_S))} |\nabla z(y)|^d dy \\ &\approx \sum_{S \in \overline{\mathcal{F}}} |B(x,R) \cap \Gamma_2(S)|.\end{aligned}$$

The desired estimate now follows from (18.6) and (18.7).

Next we want to check that $\omega(x) = |\nabla z(x)|^d$ is an A_1-weight. Let $B = B(x,R)$ be a given ball in $\mathbf{R}^d$. Choose S so that $B \subseteq \overline{\beta}_S$ and so that $Q(S)$ is minimal with this property. Let $\delta(S) \in \mathbf{R}_+$ denote the factor by which h_S dilates; this makes sense, because h_S is a composition of a translation, rotation, and dilation.

Observe that

$$|\nabla z(x)| \geq (C\delta(S))^{-1} \text{ a.e. on } \beta_S. \tag{18.14}$$

To see this notice that

$$C^{-1}\delta(\widetilde{S}) \leq \delta(S(Q)) \leq \frac{1}{2}\delta(\widetilde{S}) \tag{18.15}$$

whenever $\widetilde{S} \in \overline{\mathcal{F}}$ and $Q \in I(\widetilde{S})$, because $h_{\widetilde{S}}(c_0(Q)) = h_{S(Q)}(c_1(S(Q))$ and $c_0(Q)$ is much smaller than $c_1(S(Q))$, although not excessively smaller. In particular $\delta(S') \leq \delta(S)$ if $Q(S') \subseteq Q(S)$, and (18.14) now follows from this, (18.3), and the definition of z $[z = z_S \circ h_S^{-1}$ on $h_S(D(S))]$.

To finish the proof that ω is an A_1-weight we need to show that

$$\omega(B) \leq C\delta(S)^{-d}R^d. \tag{18.16}$$

In view of (18.13) and (18.5) we need only show that $z(B)$ is contained in a ball of radius $C\delta(S)^{-1}R$.

Set $\alpha_S = \beta_S \setminus \gamma_S$ and $B_0 = B \cap \{\alpha_S \cup (\bigcup_{Q_i \in I(S)} \alpha_{S(Q_i)})\}$. Of course B intersects α_S, since otherwise S is not minimal. Let y_0 be a point in this intersection.

We know that z is Lipschitz on B_0, with norm $\leq C\delta(S)^{-1}$, because of (18.3), the remarks concerning (18.3) given after (18.4), and (18.15). Hence $z(B_0) \subseteq B(y_0, C\delta(S)^{-1}R)$.

Suppose that $B \setminus B_0$ is not empty. Let $Q_i \in I(S)$ and $Q_{1,j} \in I(S(Q_i))$ be such that B intersects $\alpha_{S(Q_{1,j})}$. Since $\alpha_{S(Q_{1,j})} \subseteq \frac{1}{10}\beta_{S(Q_i)}$ and B intersects the complement of $\beta_{S(Q_i)}$, we have that

$$R \geq \frac{1}{2} \text{ radius } \beta_{S(Q_i)} \geq C^{-1}\delta(S) \operatorname{diam} Q_i.$$

Because we also have $\operatorname{diam} z(\beta_{S(Q_i)}) \leq C \operatorname{diam} Q_i$ from (18.12), we conclude that

$$z(\beta_{S(Q_i)}) \subseteq \{y : \operatorname{dist}(y, z(B_0)) \leq C\delta(S)^{-1}R\} \subseteq B(y_0, C'\delta(S)^{-1}R)$$

if C' is large enough. It now follows that $z(B) \subseteq B(y_0, C'\delta(S)^{-1}R)$. This proves (18.16), and hence that $\omega \in A_1$.

Combining (18.16) and (18.14) we get that $\omega(B)^{\frac{1}{d}}$ is comparable to $\delta(S)^{-1}R$, and putting this fact together with the one we just proved gives

$$\operatorname{diam} z(B) \leq C\omega(B)^{\frac{1}{d}} \tag{18.17}$$

for all balls $B \subseteq \mathbf{R}^d$. This implies that z has a locally integrable distributional gradient on $\mathbf{R}^d$, not just on Ω. Indeed, the distributional directional derivatives of z can be given as weak limits of $t^{-1}(z(x+tv) - z(x))$, and (18.17) implies that for $0 < t \leq 1$ these functions are uniformly dominated by

$$\sup_{0<R\leq 1} \left(\frac{1}{R^d} \int\limits_{B(x,R)} \omega \right)^{\frac{1}{d}}.$$

Thus the distributional directional derivatives must be locally integrable. This argument also gives $|\nabla z| \leq C\omega^{\frac{1}{d}}$ a.e., but of course $|\nabla z|^d = \omega$ a.e. by definition of ω.

This completes the proof that (C4) implies (C7), and hence of the theorem stated in Section 1.

19. A variant of (C2) and (C3)

Set

$$(19.1)\quad sy_r(x,t) = \left(\frac{1}{t^{2d}} \int\limits_{B(x,t)\cap E} \int\limits_{B(x,t)\cap E} (t^{-1}\operatorname{dist}(2y-w,E))^r dydw \right)^{\frac{1}{r}}$$

for $1 \le r \le \infty$. This measures the extent to which E is symmetric about each of its points. In particular, if $sy_r(x,t) = 0$ for all x, t, then E must be symmetric about all of its elements. The variant of (C2) and (C3) that we shall consider is the following:

(C8) $sy_1(x,t)^2 \frac{dxdt}{t}$ *is a Carleson measure on* $E \times \mathbf{R}_+$.

This condition is equivalent to the others, and this is still true if we replace $r = 1$ by any r, $1 \le r < \frac{2d}{d-2}$ ($1 \le r \le \infty$ if $d = 1$). We are going to indicate the proofs of these results in this section.

One can think of $sy_1(x,t)$ as a geometrical version of an average of second differences of a function. When E is a Lipschitz graph it is not hard to make precise the relationship between $sy_1(x,t)$ and averages of second differences of the function being graphed.

Let us first check that (C8), with $r = 1$, implies our local symmetry condition (LS) from Section 4. Let $\tau > 0$ be given, as in Definition 4.2. Suppose that $Q \in \mathcal{R}(\tau)$, so that there are y, $w \in 2Q$ such that

$$\operatorname{dist}(2y-w,E) \ge \tau \operatorname{diam} Q.$$

If we can show that this implies that

$$(19.2)\qquad sy_1(x,t) \ge C^{-1}\tau^{2d+1}$$

for any $x \in Q$ and any t such that $3 \operatorname{diam} Q \le t \le 4 \operatorname{diam} Q$, then it is easy to derive (4.3) from (C8).

Fix any such x and t. It is easy to check (19.2) using the definition (19.1) and the fact that

$$\operatorname{dist}(2y' - w', E) \ge \frac{\tau}{10} \operatorname{diam} Q$$

for any $y' \in B(y, \frac{1}{100}\tau t)$, $w' \in B(w, \frac{1}{100}\tau t)$.

From here it is not too difficult (but rather tedious) to prove that (C8) with $r = 1$ implies (C4) by modifying the arguments used to show that (C2) and (C3) each imply (C4). We omit the details, because they are somewhat messy, and because we now know that (LS) itself implies (C1)–(C7), as mentioned just before Lemma 5.13.

Assume now that E satisfies (C4) and consequently the conclusion of Proposition 5.5 as well, and let us show that (C8) must hold, and even the stronger version where $r = 1$ is replaced by $r < \frac{2d}{d-2}$ ($r \le \infty$ if $d = 1$).

Let $\eta > 0$ be small and let $k > 0$ be large, to be specified later. Let $\mathcal{B}$, $\mathcal{G}$, $\mathcal{F}$, $S \in \mathcal{F}$ be as in the definition of a corona decomposition.

Let $\mathcal{G}(\eta, k)$ be as defined in the beginning of Section 6. We want to modify the corona decomposition slightly, so that all the good cubes also lie in $\mathcal{G}(\eta, k)$. Thus we replace $\mathcal{B}$ by $\mathcal{B} \cup (\Delta \setminus \mathcal{G}(\eta, k))$ and $\mathcal{G}$ by $\mathcal{G} \cap \mathcal{G}(\eta, k)$. By assumptions (2.4) is not disturbed, but (2.5)–(2.7) notice the change. However, if for each $S \in \mathcal{F}$ you subdivide $S \cap \mathcal{G}(\eta, k)$ into maximal regions that satisfy (2.5), then the resulting family of regions also satisfies (2.6) and (2.7). Thus we may as well assume that our corona decomposition has $\mathcal{G} \subseteq \mathcal{G}(\eta, k)$.

The argument we use to show that (C8) holds is quite similar to the one in Section 15. We shall use the same notations also, and the reader may find it helpful to review that section.

By Lemma 15.2 it suffices to show that, for $r < \frac{2d}{d-2}$,

$$sy_r(x,t)^2 \chi_{\widehat{S}}(x,t) \frac{dx\,dt}{t} \tag{19.3}$$

is a Carleson measure for each $S \in \mathcal{F}$, with norm bounded independently of S. We may as well replace $\widehat{S}$ here by $\widehat{S}_1$, where $\widehat{S}_1 = \{(x,t) \in \widehat{S} : t \le 10^{-3} \operatorname{diam} Q(S)\}$; the contribution to (19.3) coming from $\widehat{S} \setminus \widehat{S}_1$ can easily be controlled using $sy_r(x,t) \le C$.

Fix $S \in \mathcal{F}$ and $r < \frac{2d}{d-2}$. As in Section 15 we let Γ be the Lipschitz graph over a d-plane P as promised in (2.6). We also let $P^\perp$ be the $(n-d)$-plane that passes through the origin and which is orthogonal to P, and we let $A : P \to P^\perp$ be the function whose graph is Γ. We denote by Π and $\Pi^\perp$ the usual projections of $\mathbf{R}^n$ onto P and $P^\perp$, and we define $d(x)$ and $D(p)$ by (7.6) and (8.1), as always.

In order to show that (19.3) is a Carleson measure we are going to estimate $sy_r(x,t)$ in terms of an L^r-average of second differences of A, plus error terms arising from the approximation of E by Γ.

Fix $x,\ t \in \widehat{S}_1$. Given $y \in \mathbf{R}^n$, set $\check{y} = (\Pi(y), A(\Pi(y))) \in \Gamma$. If $y \in 2Q(S)$ we have $|y - \check{y}| \le C\eta d(y)$, because of (2.6). Hence

$$sy_r(x,t) \le \left(t^{-2d} \int_{B(x,t)\cap E} \int_{B(x,t)\cap E} \left(t^{-1} \operatorname{dist}(2\check{y} - \check{w}, E)\right)^r dy dw \right)^{\frac{1}{r}} + C\eta \left(t^{-d} \int_{B(x,t)\cap E} \left(t^{-1} d(y)\right)^r dy \right)^{\frac{1}{r}}. \tag{19.4}$$

The first term on the right side of (19.4) can be dominated by

$$\left(t^{-2d} \int_{B(p,2t)\cap P} \int_{B(p,2t)\cap P} \left(t^{-1} \operatorname{dist}(2\hat{q} - \hat{s}, E)\right)^r d\lambda(q) d\lambda(s) \right)^{\frac{1}{r}} \tag{19.5}$$

where $p = \Pi(x)$, $\hat{q} = (q, A(q))$, $\hat{s} = (s, A(s))$, and λ is the measure obtained by pushing $H^d\,\big|_{2Q(S)}$ down to P using Π.

We want to replace λ by Lebesgue measure in (19.5). Of course λ is comparable to Lebesgue measure on $\Pi(Z)$, and so we need only look at $P \setminus \Pi(Z)$, which is the set of points $q \in P$ such that $D(q) > 0$. In this region we have that λ is comparable to Lebesgue measure at the scale of $5\eta D(q)$, at least if you stay close to $\Pi(Q(S))$. More precisely, we claim that

$$C^{-1}(\eta D(q))^d \le \lambda(B(q, 5\eta D(q))) \le C(\eta D(q))^d \tag{19.6}$$

for all $q \in P$ such that

$$\operatorname{dist}(q, \Pi(Q(S))) \le \frac{1}{10} \operatorname{diam} Q(S). \tag{19.7}$$

(Actually, we need (19.7) only for the first inequality in (19.6).)

Before checking (19.6) let us record and verify the useful fact that

$$C^{-1}d(u) \leq D(\Pi(u)) \leq d(u) \quad \text{when } u \in 2Q(S). \tag{19.8}$$

We have seen this sort of thing before, in Lemma 8.21, but the circumstances were somewhat different there. The second inequality in (19.8) is a direct consequence of the definitions, and so we need only concern ourselves with the first. Let $u \in 2Q(S)$ be given, and choose $v \in Q(S)$ such that

$$|\,\Pi(u) - \Pi(v)\,| \; + d(v) \leq 2D(\Pi(u)).$$

This we can do, because of the definitions of $D(\cdot)$ and $d(\cdot)$ in (8.1) and (7.6). We can control $|u-v|$ by projecting onto Γ, as follows:

$$\begin{aligned}|u-v| &\leq |u - (\Pi(u))^{\wedge}| + |(\Pi(u))^{\wedge} - (\Pi(v))^{\wedge}| + |v - (\Pi(v))^{\wedge}| \\ &\leq C\eta d(u) + C|\Pi(u) - \Pi(v)| + C\eta d(v) \\ &\leq C\eta d(u) + CD(\Pi(u)).\end{aligned}$$

It is easy to obtain the first half of (19.8) from these inequalities together with $d(u) - d(v) \leq |u-v|$, if η is small enough.

Now let us check the right side of (19.6). It suffices to show that

$$\Pi^{-1}(B(q, 5\eta D(q)) \cap 2Q(S)) \subseteq B(\hat{q}, C\eta D(q)). \tag{19.9}$$

Let u be an element of the left side of (19.9). Then $u \in 2Q(S)$, and so

$$|u - (\Pi(u))^{\wedge}| \;\leq\; C\eta d(u).$$

From (19.8) we have

$$\begin{aligned}d(u) &\leq CD(\Pi(u)) \\ &\leq CD(q) + C|\Pi(u) - q| \\ &\leq CD(q),\end{aligned}$$

since u lies in the left side of (19.9). Thus

$$\begin{aligned}|u - \hat{q}| &\leq |u - (\Pi(u))^{\wedge}| + |(\Pi(u))^{\wedge} - \hat{q}| \\ &\leq C\eta D(q) + C|\Pi(u) - q| \\ &\leq C\eta D(q).\end{aligned}$$

This proves (19.9), and also the right side of (19.6).

Next we verify the left side of (19.6). By definition of $D(q)$ we can find a $Q \in S$ such that

$$\operatorname{dist}(q, \Pi(Q)) + \operatorname{diam} Q \leq 2D(q). \tag{19.10}$$

By replacing Q by an ancestor if necessary we may also assume that

$$D(q) \leq C \operatorname{diam} Q. \tag{19.11}$$

Because $Q \in S$, Q is a good cube, and in particular $Q \in \mathcal{G}(\eta, k)$. Thus Q satisfies (6.2), with ϵ replaced by η. We want to use this to produce a point in $\frac{3}{2}Q(S)$ whose projection onto P lies near q. We shall then use that and the regularity of E to get the first inequality in (19.6).

Let us begin by observing that

$$\operatorname{Angle}(P, P_Q) \leq C\eta. \tag{19.12}$$

This is a minor variation of Lemma 5.13; the point is that every element of Q is within $C\eta \operatorname{diam} Q$ of both P_Q and Γ, and Γ is a Lipschitz graph over P with constant $\leq \eta$. The argument used to prove Lemma 5.13 can also be applied in this case.

Let w be the element of P_Q such that $\Pi(w) = q$, and let $\widetilde{Q}$ be the projection of Q onto P_Q. Then

$$\operatorname{dist}(w, \widetilde{Q}) \leq (1 + C\eta) \operatorname{dist}(q, \Pi(Q)) + C\eta \operatorname{diam} Q. \tag{19.13}$$

Indeed, let $y \in Q$ be such that $|q - \Pi(y)| = \operatorname{dist}(q, \Pi(Q))$, and let $\tilde{y}$ be the projection of y onto P_Q. Then $|y - \tilde{y}| \leq \eta \operatorname{diam} Q$, and

$$\begin{aligned}
\operatorname{dist}(w, \widetilde{Q}) &\leq |w - \tilde{y}| \\
\text{(by (19.12))} \qquad &\leq (1 + C\eta)|\Pi(w) - \Pi(\tilde{y})| \\
&\leq (1 + C\eta)\{|q - \Pi(y)| + |\Pi(y) - \Pi(\tilde{y})|\} \\
&\leq (1 + C\eta)\{\operatorname{dist}(q, \Pi(Q)) + \eta \operatorname{diam} Q\}.
\end{aligned}$$

In particular we have $\operatorname{dist}(w, Q) \leq C \operatorname{diam} Q$, by (19.10) and (19.11). Because $w \in P_Q$ and (6.2) holds with ϵ replaced by η, we get that there is a point $u \in E$ such that

$$|u - w| \leq \eta \operatorname{diam} Q. \tag{19.14}$$

(This is the place where we need k to be reasonably large.)

From (19.14) and (19.10) we obtain

$$|\Pi(u) - q| \leq 2\eta D(q), \tag{19.15}$$

and hence

$$\Pi^{-1}(B(q, 5\eta D(q))) \supseteq B(u, \eta D(q)).$$

To prove the first inequality in (19.6) it is enough to show that

$$B(u, \eta D(q)) \cap E \subseteq 2Q(S), \tag{19.16}$$

since $\lambda(B(q, 5\eta D(q))) = |\Pi^{-1}(B(q, 5\eta D(q))) \cap 2Q(S)|$ (by definition of λ).

To prove (19.16) it suffices to check that $\operatorname{dist}(u, Q(S)) \leq \frac{3}{2} \operatorname{diam} Q(S)$ if η is small enough, since $\eta D(q) \leq \frac{11}{10}\eta \operatorname{diam} Q(S) \leq \frac{1}{2} \operatorname{diam} Q(S)$ if η is small enough (and if q satisfies (19.7)). Using (19.14) and (19.13) we have

$$\begin{aligned}
\operatorname{dist}(u, Q(S)) &\leq |u - w| + \operatorname{dist}(w, Q(S)) \\
&\leq \eta \operatorname{diam} Q + \operatorname{dist}(w, Q) \\
&\leq 2\eta \operatorname{diam} Q + \operatorname{dist}(w, \widetilde{Q}) \\
&\leq C\eta \operatorname{diam} Q + (1 + C\eta) \operatorname{dist}(q, \Pi(Q)).
\end{aligned}$$

This is less than $\frac{3}{2} \operatorname{diam} Q(S)$ if η is small enough and q satisfies (19.7). [We are using here the fact that $Q \subseteq Q(S)$.]

This finishes the proof of (19.6). Let us use this to analyze (19.5).

It is not hard to show that

$$(19.5) \leq C \left(t^{-2d} \int_{B(p,3t)\cap P} \int_{B(p,3t)\cap P} \left(t^{-1} \operatorname{dist}(2\hat{q} - \hat{s}, E)\right)^r dq ds \right)^{\frac{1}{r}}$$

$$+ C \left(t^{-d} \int_{B(p,3t)\cap P} (\eta t^{-1} D(q))^r d\lambda(q) \right)^{\frac{1}{r}}. \tag{19.17}$$

We shall only sketch the argument. The first step is to find a covering of $\Pi(2Q(S))$ by $\Pi(Z)$ and a countable family of balls that have bounded overlap and which are of the form $B(a, 5\eta D(a))$. This is similar to the story about the R_i's in Section 8, and it is not difficult. To derive (19.17) you use this covering to break up the integrals in (19.5), and then you control the pieces separately using the following three facts. First, the oscillation of $\operatorname{dist}(2\hat{q}-\hat{s}, E)$ in q or s over a ball $B(a, 5\eta D(a))$ is bounded by $C\eta D(a)$. Second, we know that λ and Lebesgue measure are comparable on $\Pi(Z)$, and that they give the balls $B(a, 5\eta D(a))$ comparable mass. Third, since $(x,t) \in \widehat{S}$ we have $Ct \geq d(x) \geq D(p)$. This implies that if $B(p, 2t)$ intersects $B(a, 5\eta D(a))$, then $Ct \geq D(a)$. This ensures that our covering is thick enough to be useful for controlling the integrals over $B(p, 2t)$. More precisely, it implies that $B(a, 5\eta D(a)) \subseteq B(p, 3t)$ if $B(a, 5\eta D(a))$ intersects $B(p, 2t)$ and η is small enough.

The last term in (19.17) is dominated by

$$C\left(t^{-d}\int_{B(x,Ct)\cap E}(\eta t^{-1}d(y))^r dy\right)^{\frac{1}{r}}.$$

To show this it is enough to check that

$$\Pi^{-1}(B(p,10t)) \cap 2Q(S) \subseteq B(x, Ct), \tag{19.18}$$

because of (19.8) and the definition of λ. The proof of (19.18) is similar to that of (19.9), and we omit the details. It is important to use the fact that $Ct \geq d(x)$, which holds because $(x,t) \in \widehat{S}$.

Putting these inequalities back into (19.4) we obtain

$$\begin{aligned} sy_r(y,t) \leq{}& C\left(t^{-2d}\int_{B(p,3t)\cap P}\int_{B(p,3t)\cap P}\left(t^{-1}\operatorname{dist}(2\hat{q}-\hat{s},E)\right)^r dq ds\right)^{\frac{1}{r}} \\ &+ C\eta\left(t^{-d}\int_{B(x,Ct)\cap E}(t^{-1}d(y))^r dy\right)^{\frac{1}{r}}. \end{aligned} \tag{19.19}$$

To deal with this we observe that

$$\operatorname{dist}(2\hat{q}-\hat{s}, E) \leq |2\hat{q}-\hat{s}-(2q-s)^\wedge| + \operatorname{dist}((2q-s)^\wedge, E).$$

Of course $|2\hat{q}-\hat{s}-(2q-s)^\wedge| = |2A(q)-A(s)-A(2q-s)|$, and so its contribution to the first term on the right of (19.24) is at most

$$C\left(t^{-2d}\int\limits_{B(p,3t)\cap P}\int\limits_{B(p,3t)\cap P}\left(t^{-1}|A(2q-s)+A(s)-2A(q)|\right)^r dqds\right)^{\frac{1}{r}}$$

$$\leq C\gamma_r(p,6t)\leq C\inf_{B(\Pi(x),t)}\gamma_r(\cdot,10t),$$

where $\gamma_r(p,t)$ is as in (15.4). Thus

$$sy_r(x,t)\leq C\inf_{B(\Pi(x),t)}\gamma_r(\cdot,10t)+C\left(t^{-d}\int\limits_{B(p,9t)\cap P}(t^{-1}\operatorname{dist}(\hat{u},E))^r du\right)^{\frac{1}{r}}$$

$$+C\eta\left(t^{-d}\int\limits_{B(x,Ct)\cap E}(t^{-1}d(y))^r dy\right)^{\frac{1}{r}}. \tag{19.20}$$

Let us show that the middle term on the right can be controlled by the last term. To see this we first check that

$$\left(t^{-d}\int\limits_{B(p,9t)\cap P}(t^{-1}\operatorname{dist}(\hat{u},E))^r du\right)^{\frac{1}{r}}$$

$$\leq C\left(t^{-d}\int\limits_{B(p,10t)\cap P}(t^{-1}\operatorname{dist}(\hat{u},E))^r d\lambda(u)\right)^{\frac{1}{r}} \tag{19.21}$$

$$+C\left(t^{-d}\int\limits_{B(p,10t)\cap P}(\eta t^{-1}D(u))^r d\lambda(u)\right)^{\frac{1}{r}}.$$

This can be proved using the same sort of covering argument as used to get (19.17): dq and $d\lambda(q)$ are comparable on the scale of $5\eta D(u)$, and $\operatorname{dist}(\hat{q},E)$ oscillates by at most $C\eta D(u)$ on a ball of radius $5\eta D(u)$, because it is Lipschitz. It is also important that $D(p) \leq d(x) \leq Ct$ and that $t \leq 10^{-3} \operatorname{diam} Q(S)$, which are true because $(x,t) \in \widehat{S}_1$. The second condition is needed to ensure that (19.7) holds in the relevant cases.

Next we convert the integrals on the right side of (19.21) into integrals on E. Using (19.18) and the definition of λ we obtain

(19.22) left side of (19.21)

$$\leq C \left(t^{-d} \int_{B(x,Ct)\cap 2Q(S)} (t^{-1} \operatorname{dist}((\Pi(y), A(\Pi(y))), E))^r dy \right)^{\frac{1}{r}}$$

$$+ \; C \left(t^{-d} \int_{B(x,Ct)\cap E} (\eta t^{-1} d(y))^r dy \right)^{\frac{1}{r}} .$$

We have also used the fact that $D(\Pi(y)) \leq d(y)$. The first term on the right is dominated by the second, because of (2.6), and because all the relevant y's lie in E.

From (19.22) it follows that the middle term on the right side of (19.20) is indeed dominated by the last term, so that

$$(19.23)\quad sy_r(x,t) \leq C \inf_{B(\Pi(x),t)} \gamma_r(\cdot, 10t) + C\eta \left(t^{-d} \int_{B(x,Ct)\cap E} (t^{-1} d(y))^r dy \right)^{\frac{1}{r}} .$$

The remainder of the proof that (19.17) is a Carleson measure with bounded norm if $r < \frac{2d}{d-2}$ is exactly like the corresponding step in Section 15 (beginning at (15.9)).

This finishes the proof of the result stated at the beginning of the section, that (C8) is equivalent to (C1)–(C7), even if $r = 1$ is replaced by any r such that $1 \leq r < \frac{2d}{d-2}$.

20. A counterexample

Although the weak geometric lemma (5.2) is a useful auxiliary condition, it is not strong enough to imply rectifiability properties of the set in question. We shall now construct a 1-dimensional set E in $\mathbf{R}^2$ which satisfies the weak geometric lemma but not much else. On the other hand, we shall show in [**DS3**] that a d-dimensional regular set that satisfies the weak geometric lemma and another geometrical condition (big projections on some d-planes) has big pieces of Lipschitz graphs, and satisfies (C1)–(C8) in particular.

We use a minor modification of the Van Koch snowflake. The set E will be obtained as the limit of a sequence E_n of sets, with each E_n being the union of 4^n line segments of length 4^{-n}.

Given a sequence $\{\alpha_n\}$ of small real numbers we construct the E_n's recursively according to the following recipe. We take E_0 to be the unit interval on the x-axis. Suppose E_{n-1} has been constructed. To construct E_n we replace each line segment L of E_{n-1} by four segments L_1, L_2, L_3, L_4 with the following properties. (See Figure 2.)

(20.1) The length of L_i is 4^{-n}, $i = 1, \ldots, 4$.

(20.2) The endpoint of L_i is the initial point of L_{i+1}, $i = 1, 2, 3$.

(20.3) The L_i's make the angles 0, α_n, $\pi - \alpha_n$, and 0, respectively, with L.

(20.4) The midpoint of L is also the midpoint of the segment that joins the initial point of L_1 to the endpoint of L_4.

The α_n's are allowed to be quite arbitrary, for the moment anyway, although we do ask that they be small enough so that the various segments do not cross each other. ($|\alpha_n| < \frac{1}{100}$ for all n will do.)

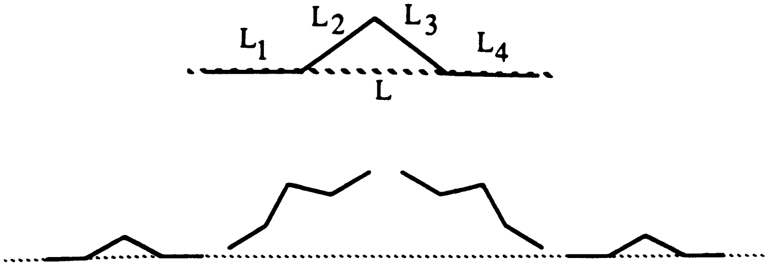

Figure 2

Let E be the set obtained in the limit from the E_n's, using the Hausdorff metric, for instance. It is not hard to see that E is (locally) regular; if B is a ball centered on E with radius 4^{-n}, then the total length of $E_n \cap B$ is about 4^{-n}, and taking further generations does not alter this significantly.

It is also not hard to check that E satisfies the weak geometric lemma if and only if α_n tends to 0.

If $\Sigma \alpha_n^2 < \infty$, then E is contained in a curve of finite length: the distance from the initial point of one of the segments of E_{n-1} to the initial point of the corresponding piece of E_n is about $\alpha_n^2 4^{-n}$, and similarly for the endpoints, and so you can connect the various pieces of E together and get a curve which contains E and has finite length (and is even chord-arc).

If $\Sigma \alpha_n^2 = \infty$, then the curve you get from the method just described has infinite length. Moreover, any curve that contains E must have infinite length. This follows from the theorem of Peter Jones in **[J3]** and the fact that the measure $\mu = \beta_1(x,t)^2 \frac{dx\,dt}{t}$ on $E \times \mathbf{R}_+$ [as in (C3)] satisfies

$$\mu(B \times (0,R)) = \infty$$

for any ball centered on E and any $R > 0$.

Even more is true. If $\Sigma \alpha_n^2 = +\infty$, then $|\Gamma \cap E| = 0$ for any rectifiable curve Γ, so that E is totally unrectifiable in the sense of geometric measure theory. It suffices to check this for Lipschitz graphs, because any rectifiable curve is contained in a countable union of C^1 curves, except perhaps for a set of length zero. (See **[Fe]**, for example.)

Suppose that Γ is a Lipschitz graph and $|E \cap \Gamma| = \tau > 0$. For each n, E is naturally divided into 4^n parts $E(n,i)$, $1 \leq i \leq n$. This is easily seen from the construction of E; these 4^n parts of E correspond to the 4^n

components of E_n. Because $\sum_i |\Gamma \cap E(n,i)| = \tau$, there is an i such that

$$|\Gamma \cap E(n,i)| \geq \tau 4^{-n} = \tau |E(n,i)|.$$

Since any $E(n,i)$ can be obtained from any $E(n,j)$ by a rigid motion, we conclude that for each n, j there is a Lipschitz graph $\Gamma_{n,j}$ such that

$$|\Gamma_{n,j} \cap E(n,j)| \geq \tau |E(n,j)|.$$

In other words, we can use the self-similarity of E to pass from the positivity of $|E \cap \Gamma|$ to E having big pieces of Lipschitz graphs. In particular E must satisfy (C6), which implies that E is contained in a regular curve and that the measure μ is a Carleson measure, both of which we know to be false. This proves that $|E \cap \Gamma| = 0$ for all Lipschitz graphs Γ.

21. Some open problems

There are three omissions in the theorem that are particularly glaring. The first is that we don't know if we can restrict ourselves to a "small" collection of kernels in (C1), e.g., to $K(x) = \frac{x_i}{|x|^{d+1}}$, $i = 1, \ldots, n$. A partial result in this direction is given in [**Fg**]. A related result is in [**Ma2**].

We also don't know so much about restricting ourselves to small sets of ψ's in (C2). Our arguments in Sections 4,5, and 6 (for proving Proposition 6.5) relied heavily on much flexibility in the choice of ψ's. However, we really didn't need so many ψ's in Sections 9, 10, and 11, although it simplified the proof substantially. We used the fact that we could choose ψ's that satisfy (10.1) and (10.2) in order to obtain (10.9) and (10.10). We could have used much less special (and much smaller) families of ψ's to get versions of (10.9) and (10.10) with ϵ^2 replaced by $\epsilon^2 + \delta^4$. Roughly speaking, the reason for this is that you can split $\psi((p, A(p)) - (q, A(q)))$ into the part that is linear in A and the remainder, and the latter is controlled by $C\delta^2$ using the fact that the Lipschitz norm of A is $\leq C\delta$.

The second omission is that we do not know whether "E has BPLG" is equivalent to our other conditions. (Of course it implies (C6).) In some sense (C5) and (C7) are not so far from this, since we know from [**D4**] that images of $\mathbf{R}^d$ under bilipschitz or ω-regular maps in $\mathbf{R}^n$ have BPLG.

The third is that the theorem does not say anything about the case when E is not regular. On the other hand, the theorem of Jones [**J3**] characterizing the subsets of planar curves with finite length in terms of quadratic conditions on the β_∞'s does not need the set to be regular. Of course for this it helps that you are working with β_∞ instead of β_q for $q < \infty$, and, as we've pointed out, Jones and Fang found a counterexample to show that such quadratic conditions on the β_∞'s need not hold for higher-dimensional Lipschitz graphs.

A natural conjecture concerning non-regular sets that might be provable using current technology is that qualitative versions of (C1)–(C3) imply that E is rectifiable. For example, if $H^d(E) < \infty$ and if

$$\sup_{\epsilon>0} \Big| \int_{E\cap\{|x-y|>\epsilon\}} K(x-y)f(y)dy \Big|$$

is finite a.e. on E for all K as before and all f in some reasonably rich class (e.g., bounded measurable functions with compact support), then is it true that E is rectifiable?

Mattila [**Ma2**] has obtained a result related to this problem. Roughly speaking he gives a characterization of rectifiability of one-dimensional sets in the plane in terms of the existence of principal values for the Cauchy kernel.

The natural qualitative versions of (C2) and (C3) are given in terms of the finiteness a.e. of square functions. For (C2) this would be the requirement that for all ψ as in (C2),

$$\sum_{\ell\leq 0} 2^{-\ell d} \int_{B(x,2^\ell)\cap E} \Big| \int_E \psi_\ell(y-z)dz \Big|^2 dy < \infty$$

for a.e. $x \in E$. The counterpart to (C3) is

$$\int_0^1 t^{-d} \int_{B(x,t)\cap E} \beta_1(y,t)^2 dy \frac{dt}{t} < \infty$$

for a.e. $x \in E$. One can formulate a qualitative version of (C8) similarly.

This possible relationship between the finiteness a.e. of these square functions and the rectifiability of E is analogous to the classical results that characterize the existence of limits or derivatives a.e. by the finiteness a.e. of square functions, as in Chapters 7 and 8 of [**St**]. There are already results in geometric measure theory that are reminiscent of the characterizations of a.e. existence of limits or derivatives in terms finiteness a.e. of maximal functions. For example, there are results that relate rectifiability (which is equivalent to the existence a.e. of approximate tangent planes) to the existence of cones at a.e. point in the set which contain most of the mass of the set asymptotically as you shrink down to the given point.

There are other questions like these which show up in connection with harmonic measure estimates and the Ahlfors distortion theorem. (See

Chapter 1 of **[B]**, especially p 33–34, or **[BCGJ]**.) One formulation goes as follows. Let Γ be a Jordan curve in the plane, not necessarily rectifiable, with complementary components Ω_1, Ω_2. Given $z \in \Gamma$, $t > 0$, let $\theta_i(z,t)$ denote the length of the largest component of $\partial B(z,t) \cap \Omega_i$, $i = 1, 2$. Thus if Γ is a line, $\theta_i(z,t) = \pi$, and in general

$$\epsilon(z,t) = \max_{i=1,2} |\pi - \theta_i(z,t)|$$

measures how close Γ is to being a line.

Let E be a closed subset of Γ of finite length. If

$$\int_0^1 \epsilon(z,t)^2 \frac{dt}{t} < \infty$$

a.e. on E, does that imply that E is rectifiable, as a 1-dimensional set in the plane?

Notice that $\epsilon(z,t)$ is quite similar to $sy_q(x,t)$ (defined in Section 19); $sy_q(x,t)$ is an integrated version of $\epsilon(z,t)$. However, we do not know how to show that quadratic conditions on $\epsilon(x,t)$ imply rectifiability properties of E, even if we make quantitative versions of these assumptions, i.e., that E is regular and $\epsilon(z,t)^2 dz \frac{dt}{t}$ is a Carleson measure on $E \times \mathbf{R}_+$. The problem is that $\epsilon(z,t)$ is not stable enough to apply the methods of this paper.

Another basic issue is to find and understand more simple geometrical sufficient conditions on E for singular integrals to be bounded on E. Some things are known (see **[D4]**, **[DJ]**, **[DS3]**, **[S1, 3, 5]**), but it is quite easy to generate questions of this type that we do not know how to answer.

Here is an example. Let E be a d-dimensional regular set in $\mathbf{R}^n$. Suppose that there is a constant C so that for each $x \in E$, $R > 0$ there is a relatively open subset U of E which is homeomorphic to a ball in $\mathbf{R}^d$ and satisfies

$$E \cap B(x,R) \subseteq U \subseteq E \cap B(x,CR).$$

Does this imply that (C1) holds for E?

This can be thought of as a higher-dimensional version of the chord-arc condition for curves. When $d = 2$ it is shown in **[S2]** that if you make a priori smoothness assumptions about E, then there is a quasisymmetric parameterization of E (by a plane) which is ω-regular for some $\omega \in A_\infty$, with estimates that do not depend on the a priori assumptions. We don't know anything for $d > 2$; for $d = 2$ we also don't have any direct geometrical understanding of this condition. (The result was proved by using uniformization to get a conformal map from the plane to E, and then estimating extremal length.)

A productive method for generating questions of this sort is to take a book on geometric measure theory ([**Fe**] for instance, or [**Ma**]), look at some of the known results concerning rectifiability, and try to find quantitative versions. For example, it is known that A is rectifiable if and only if $H^d(\Pi_V B) > 0$ for all measurable subsets B of A with $H^d(B) > 0$, and almost all d-planes V, where $\Pi_V B$ denotes the orthogonal projection of B onto V. We would like to have a version of this theorem with estimates.

Another group of problems pertains to (C7) and the existence of good parameterizations. An obvious question is whether in (C7) we can do any better than A_1 weights, e.g., is (C7) equivalent to being able to find a 1-regular mapping whose image contains E? At least one of the authors thinks that the answer should be no, but neither knows how to prove it. We don't know any good geometrical invariants to help us distinguish between 1-regular and ω-regular mappings.

We observed in the introduction that (C7) is equivalent to the version of itself with A_1 replaced by A_∞. Is there a more direct way to see this? One could hope that an ω-regular mapping for $w \in A_\infty$ could be "reparameterized" somehow to get an $\tilde{\omega}$-mapping, $\tilde{\omega} \in A_1$. An optimist might hope that this could be done by composing with a quasiconformal mapping on $\mathbf{R}^d$. This leads us to the old and difficult problem of understanding which A_∞ weights can arise as the Jacobian of quasiconformal mappings, modulo multiplication by a function which is bounded and bounded away from zero. The paper [**DS2**] is related to this problem.

One can also ask for a characterization of the A_∞ weights ω for which there is an ω-regular mapping. This is probably easier than the corresponding problem for quasiconformal mappings. We do know some partial results for this question, including the fact that there does exist an ω-regular mapping when ω is an A_1 weight.

BIBLIOGRAPHY

[B] C.J. Bishop, *Harmonic measures supported on curves*, Dissertation, Univ. of Chicago, 1987.

[BCGJ] ______, L. Carleson, J.B. Garnett and P.W. Jones, *Harmonic measures supported on curves*, Pac. J. Math. **138** (1989), 233–236.

[C1] F.M. Christ, CBMS Lectures, 1989.

[C2] __________, *A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral*, to appear.

[CW] R.R. Coifman and G. Weiss, *Extensions of Hardy spaces and their uses in analysis*, Bull. Amer. Math. Soc. **83** (1977), 569–645.

[CDM] R.R. Coifman, G. David and Y. Meyer, *La solution des conjectures de Calderón*, Adv. in Math. **48** (1983), 144–148.

[CMM] R.R. Coifman, A. McIntosh and Y. Meyer, *L'intégrale de Cauchy définit un opérateur borné sur L^2 pour les courbes lipschitziennes*, Ann. of Math. **116** (1982), 361–388.

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

[D2] ________, *Une minoration de la norme de l'opérateur de Cauchy sur les graphes lipschitziennes*, Trans. Amer. Math. Soc. **302** (1987), 741–750.

[D3] ________, *Opérateurs d'integrale singuliere sur les surfaces régulières*, Ann. Sci. Ec. Norm. Sup. (4) **21** (1988), 225–258.

[D4] ________, *Morceaux de graphes lipschitziennes et intégrales singulières sur un surface*, Rev. Mat. Iberoamericana **4** (1988), 73–114.

[D5] ________, <u>Wavelets and Singular Integrals on Curves and Surfaces</u>, to appear (Lecture Notes in Math, Springer-Verlag).

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

[DS1] ______ and S. Semmes, *Harmonic analysis and the geometry of subsets of* $\mathbf{R}^n$, to appear in the proceedings of the conference held in honor of J.L. Rubio de Francia (El Escorial, Spain, June, 1989).

[DS2] ____________________, *Strong A_∞-weights, Sobolev inequalities, and quasiconformal mappings*, in <u>Analysis and Partial Differential Equations</u>, edited by C. Sadosky, Lecture Notes in Pure and Applied Math., vol 122, Marcel Dekker, 1990.

[DS3] ____________________, *Quantitative rectifiability and Lipschitz mappings*, preprint.

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

[Fe] H. Federer, Geometric Measure Theory, Springer-Verlag, 1969.

[Fg] X. Fang, *The Cauchy integral of Calderón and analytic capacity*, Thesis, Yale, 1990.

[Fl] K.J. Falconer, The Geometry of Fractal Sets, Cambridge University Press, 1984.

[G1] J. Garnett, *Positive length but zero analytic capacity*, Proc. Amer. Math. Soc. **21** (1970), 696–699.

[G2] __________, Analytic Capacity and Measure, Lecture Notes in Math. **297**, Springer-Verlag, 1972.

[G3] __________, Bounded Analytic Functions, Academic Press, 1981.

[GJ] ______ and P.W. Jones, *The corona theorem for Denjoy domains*, Acta. Math. **155** (1985), 29–40.

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

[JL] J.L. Journé, Calderón-Zygmund Operators, Pseudo-Differential Operators and the Cauchy Integral of Calderón, Lecture Notes in Math. **994**, Springer-Verlag, 1983.

[Ma] P. Mattila, Lecture Notes on Geometric Measure Theory, Departmento de Matemáticas, Universidad de Extremadura, 1986.

[Ma2] __________, *Cauchy singular integrals and rectifiability of measures in the plane*, preprint.

[Ma3] __________, *A class of sets with positive length and zero analytic capacity*, Ann. Acad. Sci. Fenn. **10** (1985), 387–395.

[Mu] T. Murai, A Real-Variable Method for the Cauchy Transform, and Analytic Capacity, Lecture Notes in Math. **1307**, Springer-Verlag, 1988.

[S1] S. Semmes, *Chord-arc surfaces with small constant I*, to appear in Adv. in Math.

[S2] __________, *Chord-arc surfaces with small constant II: good parameterizations*, ibid.

[S3] __________, *A criterion for the boundedness of singular integrals on hypersurfaces*, Trans. Amer. Math. Soc. **311** (1989), 501-513.

[S4] __________, *Analysis vs geometry on a class of rectifiable hypersurfaces in* $\mathbf{R}^n$, to appear, Indiana Math. Journal.

RÉSUMÉ

L'objet de ce texte est l'étude de relations entre certaines propriétés analytiques d'un sous-ensemble E de $\mathbb{R}^n$ (notamment, la continuité-L^2 d'opérateurs définis par des noyaux singuliers comme le noyau de Cauchy ou le potentiel de double-couche) et des propriétés plus géométriques de E . Nous supposerons toujours que E est un ensemble de dimension d régulier au sens d'Ahlfors, c-à-d. tel que, pour toute boule $B(x,r)$ centrée sur E , la mesure de Hausdorff d-dimensionelle de $E \cap B(x,r)$ est comprise entre $C^{-1} r^d$ et $C r^d$. Le résultat principal est l'équivalence de diverses conditions, analytiques ou géométriques, portant sur E .

La première condition est la continuité sur $L^2(E)$ de l'opérateur d'intégrale singulière défini par tout noyau $K(x-y)$, où K est impaire et a des dérivées d'ordre j inférieures à $C|x-y|^{-d-j}$. La plupart des conditions géométriques peuvent être vues comme des formes plus fortes, et quantifiées, de rectifiabilité. Par exemple, l'une des conditions équivalentes est que E est contenu dans une surface Γ admettant un paramétrage "ω-régulier" (en dimension d=1, cela veut dire que Γ est une courbe régulière au sens d'Ahlfors ; dire que E est rectifiable signifierait seulement que E est contenu, à un ensemble de longueur nulle près, dans une union dénombrable de courbes rectifiables). D'autres conditions sont obtenues en mesurant, de diverses manières, l'écart entre l'intersection de E avec chaque boule centrée sur E et un plan affine de dimension d , et en demandant une certaine intégrabilité du résultat obtenu. Ce point de vue est inspiré du résultat de P. Jones sur le problème du voyageur de commerce. On peut aussi voir certaines des conditions équivalentes comme des analogues de conditions de Littlewood-Paley, ou de différences symétriques, utilisées pour décrire la régularité des fonctions.

Les techniques utilisées sont des techniques de variable réelle. La partie la plus délicate de la démonstration repose sur un argument de temps d'arrêt très proche de la construction dite "de la couronne" introduite par L. Carleson, où l'on cherche à bien approximer E par des graphes lipschitziens sur des régions de $E \times \mathbb{R}^+$ les plus grandes possibles.