

Publ. Mat. **45** (2001), 265–333

REAL ANALYSIS, QUANTITATIVE TOPOLOGY, AND GEOMETRIC COMPLEXITY

STEPHEN SEMMES

Abstract

In this paper, we give an overview of some topics involving behavior of homeomorphisms and ways in which real analysis can arise in geometric settings.

Contents

1. Finite polyhedra and combinatorial parameterization problems	266
2. Mappings and distortion	275
3. The mathematics of good behavior much of the time, and the BMO frame of mind	282
4. Quantitative topology, and calculus on singular spaces	288
5. Uniform rectifiability	298
5.1. Smoothness of Lipschitz and bilipschitz mappings	304
5.2. Smoothness and uniform rectifiability	309
5.3. A class of variational problems	313
Appendices	
Appendix A. Fourier transform calculations	315
Appendix B. Mappings with branching	318
References	320

In general, there can be significant complications involved with homeomorphisms and their behavior. Some aspects of this are reviewed in Section 1; see also [Sem9]. On the other hand, there are ways in which looking at what happens on average, and variations of this, can have useful features, as in real analysis. Here we discuss some topics related to these themes. In Section 2 we consider a basic geometric question about

2000 *Mathematics Subject Classification.* 42B99.

Key words. BMO, bilipschitz mappings, uniform rectifiability.

The author was supported by the U.S. National Science Foundation.

mappings in the plane and the distortion of distances. The analytic notion of “bounded mean oscillation” arises from this, and is described in Section 3. Section 4 concerns relations between functions and integrals of their derivatives, as on Euclidean spaces. If one is working on something like a surface with a well-behaved parameterization by a Euclidean space, then properties of functions on the surface can be reduced to analogous questions on the Euclidean space, for which there are numerous classical results; here we consider situations in which this may not be available. Finally, Section 5 deals with “uniform rectifiability”, in which a surface can have nice properties on average, but not at all points. The appendices contain some supplements to the earlier sections.

This survey originated with the John J. Gergen Memorial Lectures at Duke University in January, 1998. The author would like to thank the Mathematics Department at Duke University for the opportunity to give these lectures.

1. Finite polyhedra and combinatorial parameterization problems

Fix a positive integer d , and let P be a d -dimensional polyhedron. We assume that P is a finite union of d -dimensional simplices, so that P has “pure” dimension d .

Problem 1.1. How can one tell if P is a PL (piecewise-linear) manifold? In other words, when is P locally PL-equivalent to \mathbf{R}^d at each point?

To be precise, P is locally PL-equivalent to \mathbf{R}^d at a point $x \in P$ if there is a neighborhood of x in P which is homeomorphic to an open set in \mathbf{R}^d through a mapping which is piecewise-linear.

This is really just a particular example of a general issue, concerning existence and complexity of parameterizations of a given set. Problem 1.1 has the nice feature that finite polyhedra and piecewise-linear mappings between them can, in principle, be described in finite terms.

Before we try to address Problem 1.1 directly, let us review some preliminary matters. It will be convenient to think of P as being like a simplicial complex, so that it is made up of simplices which are always either disjoint or meet in a whole face of some (lower) dimension. Thus we can speak about the vertices of P , the edges, the 2-dimensional faces, and so on, up to the d -dimensional faces.

Since P is a finite polyhedron, its local structure at any point is pretty simple. Namely, P looks like a cone over a $(d - 1)$ -dimensional polyhedron at every point. To make this precise, imagine that Q is some

finite polyhedron in some \mathbf{R}^n , and let z be a point in \mathbf{R}^n which is affinely-independent of Q , i.e., which lies in the complement of an (affine) plane that contains Q . (We can always replace \mathbf{R}^n with \mathbf{R}^{n+1} , if necessary, to ensure that there is such a point.) Let $c(Q)$ denote the set which consists of all rays in \mathbf{R}^n which emanate from z and pass through an element of Q . We include z itself in each of these rays. This defines the “cone over Q centered at z ”. It does not really depend on the choice of z , in the sense that a different choice of z leads to a set which is equivalent to the one just defined through an invertible affine transformation.

If x is a “vertex” of P , in the sense described above, then there is a natural way to choose a $(d - 1)$ -dimensional polyhedron Q so that P is the same as the cone over Q centered at x in a neighborhood of x . Let us call Q the *link* of P at x . (Actually, with this description Q is only determined up to piecewise-linear equivalence, but this is adequate for our purposes.)

Now suppose that x is not a vertex. One can still realize P as a cone over a $(d - 1)$ -dimensional polyhedron near x , but one can also do something more precise. If x is not a vertex, then there is a positive integer k and a k -dimensional face F of P such that x lies in the interior of F . In this case there is a $(d - k - 1)$ -dimensional polyhedron Q such that P is locally equivalent to $\mathbf{R}^k \times c(Q)$ near x , with x in P corresponding to a point (y, z) in $\mathbf{R}^k \times c(Q)$, where z is the center of $c(Q)$. This same polyhedron Q works for all the points in the interior of F , and we call Q the *link* of F .

Basic Fact 1.2. P is everywhere locally equivalent to \mathbf{R}^d if and only if all of the various links of P (of all dimensions) are piecewise-linearly equivalent to standard spheres (of the same dimension).

Here the “standard sphere of dimension m ” can be taken to be the boundary of the standard $(m + 1)$ -dimensional simplex.

Basic Fact 1.2 is standard and not hard to see. The “if” part is immediate, since one knows exactly what the cone over a standard sphere looks like, but for the converse there is a bit more to check. A useful observation is that if Q is a j -dimensional polyhedron whose cone $c(Q)$ is piecewise-linearly equivalent to \mathbf{R}^{j+1} in a neighborhood of the center of $c(Q)$, then Q must be piecewise-linearly equivalent to a standard j -dimensional sphere. This is pretty easy to verify, and one can use it repeatedly for the links of P of codimension larger than 1. (A well-known point here is that one should be careful not to use *radial* projections to investigate links around vertices, but suitable *pseudo-radial* projections,

to fit with the piecewise-linear structure, and not just the topological structure.)

A nice feature of Basic Fact 1.2 is that it sets up a natural induction in the dimensions, since the links of P always have dimension less than P . This leads to the following question.

Problem 1.3. If Q is a finite polyhedron which is a k -dimensional PL manifold, how can one tell if Q is a PL sphere of dimension k ?

It is reasonable to assume here that Q is a PL-manifold, because of the way that one can use Basic Fact 1.2 and induction arguments.

Problem 1.3 is part of the matter of the Poincaré conjecture, which would seek to say that Q is a PL sphere as soon as it is homotopy-equivalent to a sphere. This has been established in all dimensions except 3 and 4. (Compare with [RouS].) In dimension 4 the Poincaré conjecture was settled by M. Freedman [Fre] in the “topological” category (with ordinary homeomorphisms (continuous mappings with continuous inverses) and topological manifolds), but it remains unknown in the PL case. The PL case is equivalent to the smooth version in this dimension, and both are equivalent to the ordinary topological version in dimension 3. (A brief survey related to these statements is given in Section 8.3 of [FreQ].) Although the Poincaré conjecture is known to hold in the PL category in all higher dimensions (than 4), it does not always work in the smooth category, because of exotic spheres (as in [Mil1], [KerM]).

If the PL version of the Poincaré conjecture is true in all dimensions, then this would give one answer to the question of recognizing PL manifolds among finite polyhedra in Problem 1.1. Specifically, our polyhedron P would be a PL manifold if and only if its links are all homotopy-equivalent to spheres (of the correct dimension).

This might seem like a pretty good answer, but there are strong difficulties concerning complexity for matters of homotopy. In order for a k -dimensional polyhedron Q to be a homotopy sphere, it has to be simply connected in particular, at least when $k \geq 2$. In other words, it should be possible to continuously deform any loop in Q to a single point, or, equivalently, to take any continuous mapping from a circle into Q and extend it to a continuous mapping from a closed disk into Q . This extension can entail enormous complexity, in the sense that the filling to the disk might have to be of much greater complexity than the original loop itself.

This is an issue whose geometric significance is often emphasized by Gromov. To describe it more precisely it is helpful to begin with some related algebraic problems, concerning finitely-presented groups.

Let G be a group. A finite presentation of G is given by a finite list g_1, g_2, \dots, g_n of generators for G together with a finite set r_1, r_2, \dots, r_m of “relations”. The latter are (finite) words made out of the g_i 's and their inverses. Let us assume for convenience that the set of relations includes the inverses of all of its elements, and also the empty word. The r_j 's are required to be trivial, in the sense that they represent the identity element of G . This implies that arbitrary products of conjugates of the r_j 's also represent the identity element, and the final requirement is that if w is any word in the g_i 's and their inverses which represents the identity element in G , then it should be possible to obtain w from some product of conjugates of the r_j 's through cancellations of subwords of the form $g_i^{-1}g_i$ and $g_i g_i^{-1}$.

For instance, the group \mathbf{Z}^2 can be described by two generators a, b and one relation, $aba^{-1}b^{-1}$. As another concrete example, there is the (Baumslag-Solitar) group with two generators x, y and one relation $x^2yx^{-1}y^{-1}$.

Suppose that a group G and finite presentation of G are given and fixed, and let w be a word in the generators of G and their inverses. Given this information, how can one decide whether w represents the identity element in G ? This is called “the word problem” (for G). It is a famous result that there exist finite presentations of groups for which there is no algorithm to solve the word problem. (See [Man].)

To understand what this really means, let us first notice that the set of trivial words for the given presentation is “recursively enumerable”. This means that there is an algorithm for listing all of the trivial words. To do this, one simply has to have the algorithm systematically generate all possible conjugates of the relations, all possible products of conjugates of relations, and all possible words derived from these through cancellations as above. In this way the algorithm will constantly generate trivial words, and every trivial word will eventually show up on the list.

However, this does not give a finite procedure for determining that a given word is *not* trivial. A priori one cannot conclude that a given word is not trivial until one goes through the entire list of trivial words.

The real trouble comes from the cancellations. In order to establish the triviality of a given word w , one might have to make derivations through words which are enormously larger, with a lot of collapsing at the end. If one had a bound for the size of the words needed for at least one derivation of the triviality of a given word w , a bound in terms of an effectively computable (or “recursive”) function of the length of w , then the word problem would be algorithmically solvable. One could simply search through all derivations of at most a computable size.

This would not be very efficient, but it would be an algorithm. As it is, even this does not always work, and there are finitely-presented groups for which the derivations of triviality may need to involve words of nonrecursive size compared to the given word.

One should keep in mind that for a given group and a given presentation there is always *some* function $f(n)$ on the positive integers so that trivial words of length at most n admit derivations of their triviality through words of size no greater than $f(n)$. This is true simply because there are only finitely many words of size at most n , and so one can take $f(n)$ to be the maximum size incurred in some finite collection of derivations. The point is that such a function f may not be bounded by a recursive function. This means that f could be really huge, larger than any tower of exponentials, for instance.

The same kind of phenomenon occurs geometrically, for deciding whether a loop in a given polyhedron can be continuously contracted to a point. This is because any finite presentation of a group G can be coded into a finite polyhedron, in such a way that the group G is represented by the fundamental group of the polyhedron. This is a well-known construction in topology.

Note that while the fundamental group of a space is normally defined in terms of *continuous* (based) loops in the space and the continuous deformations between them, in the case of finite polyhedra it is enough to consider *polygonal* loops and deformations which are piecewise-linear (in addition to being continuous). This is another standard fact, and it provides a convenient way to think about complexity for loops and their deformations.

Although arbitrary finite presentations can be coded into finite polyhedra, as mentioned above, this is not the same as saying that they can be coded into compact *manifolds*. It turns out that this does work when the dimension is at least 4, i.e., for each $n \geq 4$ it is true that every finite presentation can be coded into a compact PL manifold of dimension n . This type of coding can be used to convert algorithmic unsolvability results for problems in group theory into algorithmic unsolvability statements in topology. For instance, there does not exist an algorithm to decide when a given finite presentation for a group actually defines the trivial group, and, similarly, there does not exist an algorithm for deciding when a given manifold (of dimension at least 4) is simply-connected. See [BooHP], [Mar1], [Mar2], [Mar3] for more information and results.

Let us mention that in dimensions 3 and less, it is *not* true that arbitrary finitely-presented groups can be realized as fundamental groups

of compact manifolds. Fundamental groups of manifolds are very special in dimensions 1 and 2, as is well known. The situation in dimension 3 is more complicated, but there are substantial restrictions on the groups that can arise as fundamental groups. As an aspect of this, one can look at restrictions related to Poincaré duality. In a different vein, the fundamental group of a 3-dimensional manifold has the property that all of its finitely-generated subgroups are finitely-presented. See [Sco], and Theorem 8.2 on p. 70 of [Hem1]. See also [Jac]. In another direction, there are relatively few abelian groups which can arise as subgroups of fundamental groups of 3-dimensional manifolds. See [Eps], [EvaM], Theorems 9.13 and 9.14 on p. 84f of [Hem1], and p. 67–69 of [Jac]. At any rate, it is a large open problem to know exactly what groups arise as fundamental groups of 3-dimensional manifolds.

See also [Thu] and Chapter 12 of [Eps+] concerning these groups. The book [Eps+] treats a number of topics related to computability and groups, and not just in connection with fundamental groups of 3-manifolds. This includes broad classes of groups for which positive results and methods are available. See [Far] as well in this regard.

Beginning in dimension 5, it is known that there is no algorithm for deciding when a compact PL manifold is piecewise-linearly equivalent to a standard (PL) sphere. This is a result of S. Novikov. See Section 10 of [VolKF], and also the appendix to [Nab]. Note that in dimensions less than or equal to 3, such algorithms do exist. This is classical for dimensions 1, 2; see [Rub1], [Rub2], [Tho] concerning dimension 3, and related problems and results.

Imagine that we have a connected PL manifold M of some dimension $n \geq 5$ whose equivalence to a standard sphere is true but “hard” to check. According to the solution of the Poincaré conjecture in these dimensions, M will be equivalent to an n -sphere if it is homotopy-equivalent to \mathbf{S}^n . For standard reasons of algebraic topology, this will happen exactly when M is simply-connected and has trivial homology in dimensions 2 through $n - 1$. (Specifically, this uses Theorem 9 and Corollary 24 on pages 399 and 405, respectively, of [Spa]. It also uses the existence of a degree-1 mapping from M to \mathbf{S}^n to get started (i.e., to have a mapping to which the aforementioned results can be applied), and the fact that the homology of M and \mathbf{S}^n vanish in dimensions larger than n , and are equal to \mathbf{Z} in dimension n . To obtain the degree-1 mapping from M to \mathbf{S}^n , one can start with any point in M and a neighborhood of that point which is homeomorphic to a ball. One then collapses the complement of that neighborhood to a point, which gives rise to the

desired mapping.) The vanishing of homology can be determined algorithmically, and so if the equivalence of M with an n -sphere is “hard” for algorithmic verification, then the problem must occur already with the simple-connectivity of M .

To determine whether M is simply-connected it is enough to check that a finite number of loops in M can be contracted to a point, i.e., some collection of generators for the fundamental group. If this is “hard”, then it means that the complexity of the contractions should be enormous compared to the complexity of M . For if there were a bound in terms of a recursive function, then one could reverse the process and use this to get an algorithm which could decide whether M is PL equivalent to a sphere, and this is not possible.

If M is a hard example of a PL manifold which is equivalent to an n -sphere, then any mapping from M to the sphere which realizes this equivalence must necessarily be of very high complexity as well. Because of the preceding discussion, this is also true for mappings which are homotopy-equivalences, or even which merely induce isomorphisms on π_1 , if one includes as part of the package of data enough information to justify the condition that the induced mapping on π_1 be an isomorphism. (For a homotopy equivalence, for instance, one could include the mapping f from M to the n -sphere, a mapping g from the n -sphere to M which is a homotopy-inverse to f , and mappings which give homotopies between $f \circ g$ and $g \circ f$ to the identity on the n -sphere and M , respectively.) This is because one could use the mapping to reduce the problem of contracting a loop in M to a point to the corresponding problem for the n -sphere, where the matter of bounds is straightforward.

Similar considerations apply to the problem of deciding when a finite polyhedron P is a PL manifold. Indeed, given a PL manifold M whose equivalence to a sphere is in question, one can use it to make a new polyhedron P by taking the “suspension” of M . This is defined by taking two points y and z which lie outside of a plane that contains M , and then taking the union of all of the (closed) line segments that go from either of y or z to a point in M . One should also be careful to choose y and z so that these line segments never meet, except in the trivial case of line segments from y and z to the same point x in M , with x being the only point of intersection of the two segments. (One can imagine y and z as lying on “opposite sides” of an affine plane that contains M .)

If M is equivalent to a sphere, then this operation of suspension produces a PL manifold equivalent to the sphere of 1 larger dimension, as one can easily check. If M is not PL equivalent to a sphere, then the

suspension P of M is not a PL manifold at all. This is because M is the link of P at the vertices y and z , by construction, so that one is back to the situation of Basic Fact 1.2.

Just as there are PL manifolds M whose equivalence with a sphere is hard, the use of the suspension shows that there are polyhedra P for which the property of being a PL manifold is hard to establish. Through the type of arguments outlined above, when PL coordinates exist for a polyhedron P , they may have to be of enormous complexity compared to the complexity of P itself. This works more robustly than just for PL coordinates, i.e., it applies to objects which are precise enough to deal with simple-connectivity of the links of P (which are local, as opposed to more global aspects of P). Again, this follows the discussion above.

We have focussed on piecewise-linear coordinates for finite polyhedra for the sake of simplicity, but similar themes of complexity come up much more generally, and in a number of different ways. In particular, existence and complexity of parameterizations is often related in a strong manner to the behavior of something like π_1 , sometimes in a localized form, as with the links of a polyhedron. For topology of manifolds in high dimensions, π_1 and the filling of loops with disks come up in the Whitney lemma, for instance. This concerns the separation of crossings of submanifolds through the use of embedded 2-dimensional disks, and it can be very useful for making some geometric constructions. (A very nice brief review of some of these matters is given in Section 1.2 of [DonK].) Localized π_1 -type conditions play a crucial role in taming theorems in geometric topology. Some references related to this are [Bin1], [Bin2], [Bin3], [Bur], [BurC], [Can1], [Can2], [Dave1], [Dave2], [Edw1], [Moi], [Rus1], [Rus2].

As another type of example, one has the famous “double suspension” results of Edwards and Cannon [Can1], [Can3], [Dave2], [Edw2]. Here one starts with a finite polyhedron H which is a manifold with the same homology as a sphere of the same dimension, and one takes the suspension (described above) of the suspension of H to get a new polyhedron K . The result is that K is actually homeomorphic to a sphere. A key point is that H is *not* required to be simply-connected. When $\pi_1(H) \neq 0$, it is not possible for the homeomorphism from K to a standard sphere to be piecewise-linear, or even Lipschitz continuous (which means that the distance between two points in the image is always bounded by a constant times the distance between their preimages in K). Concerning the latter, see [SieS]. Not much is known about the complexity of the homeomorphisms in this case. (We shall say a bit more about this in Section 5.)

If J is obtained as a single suspension of H , and if $\pi_1(H) \neq 0$, then J cannot be a topological manifold at all (at least if the dimension of H is greater than or equal to 2). Indeed, if M is a topological manifold of dimension n , then for every point p in M there are arbitrarily small neighborhoods U of p which are homeomorphic to an open n -ball, and $U \setminus \{p\}$ must then be simply-connected when $n \geq 3$. This does not work for the suspension J of H when $\pi_1(H) \neq 0$, with p taken to be one of the two cone points in J introduced in the suspension construction.

However, J has the advantage over H that it *is* simply-connected. This comes from the process of passing to the suspension (and the fact that H should be connected, since it has the same homology as a sphere). It is for this reason that the cone points of K do not have the same trouble as in J itself, with no small deleted neighborhoods which are simply-connected. The singularities at the cone points in J lead to trouble with the codimension-2 links in K , but this turns out not to be enough to prevent K from being a topological manifold, or a topological sphere. It *does* imply that the homeomorphisms involved have to distort distances in a very strong way, as in [SieS].

The first examples of finite polyhedra which are homeomorphic to each other but not piecewise-linearly equivalent were given by Milnor [Mil2]. See also [Sta]. This is the “failure of the Hauptvermutung”. These polyhedra are not PL manifolds, and it turns out that there are examples of compact PL manifolds which are homeomorphic but not piecewise-linearly equivalent too. See [Sie] for dimensions 5 and higher, and [DonK], [FreQ] for dimension 4. In dimensions 3 and lower, this does not happen [Moi], [Bin2]. The examples in [Mil2], [Sta], [Sie] involved non-PL homeomorphisms whose behavior is much milder than in the case of double-suspension spheres. There are general results in this direction for PL manifolds (and more broadly) in dimensions greater than or equal to 5. See [Sul1], [SieS]. Analogous statements fail in dimension 4, by [DonS].

Some other examples where homeomorphic coordinates do not exist, or necessarily have complicated behavior, even though the geometry behaves well in other ways, are given in [Sem4], [Sem5].

See [DaviS4], [Fu], [HeiS], [HeiY], [MülS], [Sem2], [Toro1], [Toro2] for some related topics concerning homeomorphisms and bounds for their behavior.

One can try to avoid difficulties connected to π_1 by using mappings with branching rather than homeomorphisms. This is discussed further in Appendix B.

Questions of algorithmic undecidability in topology have been revisited in recent years, in particular by Nabutovsky and Weinberger. See [NabW1], [NabW2], for instance, and the references therein.

2. Mappings and distortion

A very basic mechanism for limiting geometric complexity is to limit the way that distances can be distorted by a mapping.

If distances are distorted by only a small amount, then one might think of the mapping as being approximately “flat”. Let us look more closely at this, and see what actually happens.

Let δ be a small positive number, and let f be a mapping from the Euclidean plane \mathbf{R}^2 to itself. Given two points $x, y \in \mathbf{R}^2$, let $|x - y|$ denote the usual Euclidean distance between them. We shall assume that

$$(2.1) \quad (1 + \delta)^{-1}|x - y| \leq |f(x) - f(y)| \leq (1 + \delta)|x - y|$$

for all $x, y \in \mathbf{R}^2$. This says exactly that f does not ever shrink or expand distances by more than a factor of $1 + \delta$.

What does this really mean about the behavior of f ? A first point is that if δ were equal to 0, so that f does not distort distances at all, then f would have to be a “rigid” mapping. This means that f could be expressed as

$$(2.2) \quad f(x) = A(x) + b,$$

where b is an element of \mathbf{R}^2 and A is a linear mapping on \mathbf{R}^2 which is either a rotation or a combination of a rotation and a reflection. This is well known, and it is not hard to prove. For instance, it is not hard to show that the assumption that f preserve distances implies that f takes lines to lines, and that it preserve angles, and from there it is not hard to see that f must be of the form (2.2) as above.

If δ is not equal to zero, then one would like to say that f is approximately equal to a rigid mapping when δ is small enough. Here is a more precise statement. Let D be a (closed) disk of radius r in the plane, $D = \{x \in \mathbf{R}^2 : |x - w| \leq r\}$ for some $w \in \mathbf{R}^2$. There is a rigid mapping $T: \mathbf{R}^2 \rightarrow \mathbf{R}^2$, depending on D and f , such that

$$(2.3) \quad r^{-1} \sup_{x \in D} |f(x) - T(x)| \leq \text{small}(\delta),$$

where $\text{small}(\delta)$ depends only on δ , and not on D or f , and has the property that

$$(2.4) \quad \text{small}(\delta) \rightarrow 0 \quad \text{as} \quad \delta \rightarrow 0.$$

There are a number of ways to look at this. One can give direct constructive arguments, through basic geometric considerations or computations. In particular, one can derive explicit bounds for $\text{small}(\delta)$ in terms of δ . Statements of this kind are given in [Joh]. There are also abstract and inexplicit methods, in which one argues by contradiction using compactness and the Arzela-Ascoli theorem. (In some related but different contexts, this can be fairly easy or manageable, while explicit arguments and estimates are less clear.)

The presence of the factor of r^{-1} on the left side of (2.3) may not make sense at first glance, but it is absolutely on target, and indispensable. It reflects the natural *scaling* of the problem, and converts the left-hand side of (2.3) into a dimensionless quantity, just as δ is dimensionless. One can view this in terms of the natural *invariances* of the problem. Nothing changes here if we compose f (on either side) with a translation, rotation, or reflection, and the same is true if we make simultaneous dilations on both the domain and the range of equal amounts. In other words, if a is any positive number, and if we define $f_a: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ by

$$(2.5) \quad f_a(x) = a^{-1}f(ax),$$

then f_a satisfies (2.1) exactly when f does. The approximation condition (2.3) is formulated in such a way as to respect the same kind of invariances as (2.1) does, and the factor of r^{-1} accounts for the dilation-invariance.

This kind of approximation by rigid mappings is pretty good, but can we do better? Is it possible that the approximation works at the level of the *derivatives* of the mappings, rather than just the mappings themselves?

Here is another way to think about this, more directly in terms of distance geometry. Let us consider a simple mechanism by which mappings that satisfy (2.1) can be produced, and ask whether this mechanism gives everything. Fix a nonnegative number k , and call a mapping $g: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ *k-Lipschitz* if

$$(2.6) \quad |g(x) - g(y)| \leq k|x - y|$$

for all $x, y \in \mathbf{R}^2$. This condition is roughly equivalent to saying that the differential of g has norm less than or equal to k everywhere. Specifically, if g is differentiable at every point in \mathbf{R}^2 , and if the norm of its differential is bounded by k everywhere, then (2.6) holds, and this can be derived from the mean value theorem. The converse is not quite true, however, because Lipschitz mappings need not be differentiable everywhere. They are differentiable *almost everywhere*, in the sense of Lebesgue measure.

(See [Fed], [Ste1], [Sem8].) To get a proper equivalence one can consider derivatives in the sense of distributions.

If $f = S + g$, where S is a rigid mapping and g is k -Lipschitz, and if $k \leq 1/2$ (say), then f satisfies (2.1) with $\delta = 2k$. (More precisely, one can take $\delta = (1 - k)^{-1} - 1$.) This is not hard to check. When k is small, this is a much stronger kind of approximation of f by rigid mappings than (2.3) is. In particular, it implies that the differential of f is uniformly close to the differential of S .

To what extent can one go in the opposite direction, and say that if f satisfies (2.1) with δ small, then f can be approximated by rigid mappings in this stronger sense? Let us begin by looking at what happens with the differential of f at individual points. Let x be some point in \mathbf{R}^2 , and assume that the differential df_x of f at x exists. Thus df_x is a linear mapping from \mathbf{R}^2 to itself, and

$$(2.7) \quad f(x) + df_x(y - x)$$

provides a good approximation to $f(y)$ for y near x , in the sense that

$$(2.8) \quad |f(y) - \{f(x) + df_x(y - x)\}| = o(|y - x|).$$

One can also think of the differential as the map obtained from f by “blowing up” at x . This corresponds to the formula

$$(2.9) \quad df_x(v) = \lim_{t \rightarrow 0} t^{-1}(f(x + tv) - f(x)),$$

with t taken from positive real numbers.

It is not hard to check that df_x , as a mapping on \mathbf{R}^2 (with x fixed), automatically satisfies (2.1) when f does. Because the differential is already linear, standard arguments from linear algebra imply that it is close to a rotation or to the composition of a rotation and a reflection when δ is small, and with easy and explicit estimates for the degree of approximation.

This might sound pretty good, but it is actually much weaker than something like a representation of f as $S + g$, where S is a rigid mapping and g is k -Lipschitz with a reasonably-small value of k . If there is a representation of this type, then it means that the differential df_x of f is always close to the differential of S , which is *constant*, i.e., independent of x . The simple method of the preceding paragraph implies that df_x is always close to being a rotation or a rotation composed with a reflection, but a priori the choice of such a linear mapping could depend on x in a strong way. That is very different from saying that there is a single linear mapping that works for every x .

Here is an example which shows how this sort of phenomenon can happen. (See also [Joh].) Let us work in polar coordinates, so that a point z in \mathbf{R}^2 is represented by a radius $r \geq 0$ and an angle θ . We define $f: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ by saying that if x is described by the polar coordinates (r, θ) , then

$$(2.10) \quad f(x) \text{ has polar coordinates } (r, \theta + \epsilon \log r).$$

Here ϵ is a small positive number that we get to choose. Of course f should also take the origin to itself, despite the fact that the formula for the angle degenerates there.

Thus f maps each circle centered at the origin to itself, and on each such circle f acts by a rotation. We do not use a single rotation for the whole plane, but instead let the rotation depend logarithmically on the radius, as above. This has the effect of transforming every line through the origin into a logarithmic spiral. This spiral is very “flat” when ϵ is small, but nonetheless it does wrap around the origin infinitely often in every neighborhood of the origin.

It is not hard to verify that this construction leads to a mapping f that satisfies (2.1), with a δ that goes to 0 when ϵ does, and at an easily computable (linear) rate. This gives an example of a mapping that cannot be represented as $S + g$ with S rigid and g k -Lipschitz for a fairly small value of k (namely, $k < 1$). For if f did admit such a representation, it would not be able to transform lines into curves that spiral around a fixed point infinitely often; instead it would take a line L to a curve Γ which can be realized as the graph of a function over the line $S(L)$. The spirals that we get can never be realized as a graph of a function over any line. This is not hard to check.

This spiralling is not incompatible with the kind of approximation by rigid mappings in (2.3). Let us consider the case where D is a disk centered at the origin, which is the worst-case scenario anyway. One might think that (2.3) fails when we get too close to the origin (as compared to the radius of D), but this is not the case. Let T be the rotation on \mathbf{R}^2 that agrees with f on the boundary of D . If ϵ is small (which is necessary in order for the δ to be small in (2.1)), then T provides a good approximation to f on D in the sense of (2.3). In fact, T provides a good approximation to f at the level of their derivatives too on most of D , i.e., on the complement of a small neighborhood of the origin. The approximation of derivatives breaks down near the origin, but the approximation of values does not, as in (2.3), because f and T both take points near the origin to points near the origin.

This example suggests another kind of approximation by rigid mappings that might be possible. Given a disk D of radius r and a mapping f that satisfies (2.1), one would like to have a rigid mapping T on \mathbf{R}^2 so that (2.3) holds, and also so that

$$(2.11) \quad \frac{1}{\pi r^2} \int_D \|df_x - dT\| dx \leq \text{small}'(\delta),$$

where $\text{small}'(\delta)$ is, as before, a positive quantity which depends only on δ (and not on f or D) and which tends to 0 when δ tends to 0. Here dx refers to the ordinary 2-dimensional integration against area on \mathbf{R}^2 , and we think of $df_x - dT$ as a matrix-valued function of x , with $\|df_x - dT\|$ denoting its norm (in any reasonable sense).

In other words, instead of asking that the differential of f be approximated *uniformly* by the differential of a rigid mapping, which is not true in general, one can ask only that the differential of f be approximated by the differential of T *on average*.

This does work, and in fact one can say more. Consider the expression

$$(2.12) \quad P(\lambda) = \text{Probability}(\{x \in D : \|df_x - dT\| \geq \text{small}'(\delta) \cdot \lambda\}),$$

where λ is a positive real number. Here “probability” means Lebesgue measure divided by πr^2 , which is the total measure of the disk D . The inequality (2.11) implies that

$$(2.13) \quad P(\lambda) \leq \frac{1}{\lambda}$$

for all $\lambda > 0$. It turns out that there is actually a universal bound for $P(\lambda)$ with exponential decay for $\lambda \geq 1$. This was proved by John [Joh] (with concrete estimates).

Notice that uniform approximation of the differential of f by the differential of T would correspond to a statement like

$$(2.14) \quad P(\lambda) = 0$$

for all λ larger than some fixed (universal) constant. John’s result of exponential decay is about the next best thing.

As a technical point, let us mention that one can get exponential decay conditions concerning the way that $\|df_x - dT\|$ should be small most of the time in a kind of trivial manner, with constants that are not very good (at all), using the linear decay conditions with good constants, together with the fact that df is bounded, so that $\|df_x - dT\|$ is bounded. In the exponential decay result mentioned above, the situation is quite

different, and one keeps constants like those from the linear decay condition. This comes out clearly in the proof, and we shall see more about it later.

This type of exponential decay occurs in a simple way in the example above, in (2.10). (This also comes up in [Joh].) One can obtain this from the presence of $\epsilon \log r$ in the angle coordinate in the image. The use of the logarithm here is not accidental, but fits exactly with the requirements on the mapping. For instance, if one differentiates $\log r$ in ordinary Cartesian coordinates, then one gets a quantity of size $1/r$, and this is balanced by the r in the first part of the polar coordinates in (2.10), to give a result which is bounded.

It may be a bit surprising, or disappointing, that uniform approximation to the differential of f does not work here. After all, we did have “uniform” (or “supremum”) bounds in the hypothesis (2.1), and so one might hope to have the same kind of bounds in the conclusion. This type of failure of supremum bounds is quite common, and in much the same manner as in the present case. We shall return to this in Section 3.

How might one prove (2.11), or the exponential decay bounds for $P(\lambda)$? Let us start with a slightly simpler situation. Imagine that we have a rectifiable curve γ in the plane whose total length is only slightly larger than the distance between its two endpoints. If the length of γ were *equal* to the distance between the endpoints, then γ would have to be a straight line segment, and nothing more. If the length is slightly larger, then γ has to stay close to the line segment that joins its endpoints. In analogy with (2.11), we would like to say that the tangents to γ are nearly parallel, on average, to the line that passes through the endpoints of γ .

In order to analyze this further, let $z(t)$, $t \in \mathbf{R}$, $a \leq t \leq b$, be a parameterization of γ by arclength. This means that $z(t)$ should be 1-Lipschitz, so that

$$(2.15) \quad |z(s) - z(t)| \leq |s - t|$$

for all $s, t \in [a, b]$, and that $|z'(t)| = 1$ for almost all t , where $z'(t)$ denotes the derivative of $z(t)$. Set

$$(2.16) \quad \zeta = \frac{z(b) - z(a)}{b - a} = \frac{1}{b - a} \int_a^b z'(t) dt.$$

Let us compute

$$(2.17) \quad \frac{1}{b - a} \int_a^b |z'(s) - \zeta|^2 ds,$$

which controls the average oscillation of $z'(s)$. Let $\langle \cdot, \cdot \rangle$ denote the standard inner product on \mathbf{R}^2 , so that

$$(2.18) \quad \begin{aligned} |x - y|^2 &= \langle x - y, x - y \rangle = \langle x, x \rangle - 2\langle x, y \rangle + \langle y, y \rangle \\ &= |x|^2 - 2\langle x, y \rangle + |y|^2 \end{aligned}$$

for all $x, y \in \mathbf{R}^2$. Applying this with $x = z'(s)$, $y = \zeta$, we get that

$$(2.19) \quad \frac{1}{b-a} \int_a^b |z'(s) - \zeta|^2 ds = 1 - 2\frac{1}{b-a} \int_a^b \langle z'(s), \zeta \rangle ds + |\zeta|^2,$$

since $|z'(s)| = 1$ a.e., and ζ does not depend on s . The middle term on the right side reduces to

$$(2.20) \quad 2\langle \zeta, \zeta \rangle,$$

because of (2.16). Thus (2.19) yields

$$(2.21) \quad \frac{1}{b-a} \int_a^b |z'(s) - \zeta|^2 ds = 1 - 2|\zeta|^2 + |\zeta|^2 = 1 - |\zeta|^2.$$

On the other hand, $|z(b) - z(a)|$ is the same as the distance between the endpoints of γ , and $b - a$ is the same as the length of γ , since $z(t)$ is the parameterization of γ by arclength. Thus $|\zeta|$ is exactly the ratio of the distance between the endpoints of γ to the length of γ , by (2.16), and $1 - |\zeta|^2$ is a dimensionless quantity which is small exactly when the length of γ and the distance between its endpoints are close to each other (proportionately). In this case (2.21) provides precise information about the way that $z'(s)$ is approximately a constant on average. (These computations follow ones in [CoiMe2].)

One can use these results for curves for looking at mappings from \mathbf{R}^2 (or \mathbf{R}^n) to itself, by considering images of segments under the mappings. This does not seem to give the proper bounds in (2.11), in terms of dependence on δ , though. In this regard, see John's paper [Joh]. (Compare also with Appendix A.) Note that for curves by themselves, the computations above are quite sharp, as indicated by the equality in (2.21). See also [CoiMe2].

The exponential decay of $P(\lambda)$ requires more work. A basic point is that exponential decay bounds can be derived in a very general way once one knows (2.11) for *all* disks D in the plane. This is a famous result of John and Nirenberg [JohN], which will be discussed further in Section 3. In the present situation, having estimates like (2.11) for all disks D (and with uniform bounds) is quite natural, and is essentially automatic, because of the invariances of the condition (2.1) under translations and dilations. In other words, once one has an estimate like (2.11) for some

fixed disk D and all mappings f which satisfy (2.1), one can conclude that the same estimate works for all disks D , because of invariance under translations and dilations.

3. The mathematics of good behavior much of the time, and the BMO frame of mind

Let us start anew for the moment, and consider the following question in analysis. Let h be a real-valued function on \mathbf{R}^2 . Let Δ denote the Laplace operator, given by

$$(3.1) \quad \Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2},$$

where x_1, x_2 are the standard coordinates on \mathbf{R}^2 . To what extent does the behavior of Δh control the behavior of the other second derivatives of h ?

Of course it is easy to make examples where Δh vanishes at a point but the other second derivatives do not vanish at the same point. Let us instead look for ways in which the overall behavior of Δh can control the overall behavior of the other second derivatives.

Here is a basic example of such a result. Let us assume (for simplicity) that h is smooth and that it has compact support, and let us write ∂_1 and ∂_2 for $\partial/\partial x_1$ and $\partial/\partial x_2$, respectively. Then

$$(3.2) \quad \int_{\mathbf{R}^2} |\partial_1 \partial_2 h(x)|^2 dx \leq \int_{\mathbf{R}^2} |\Delta h(x)|^2 dx.$$

This is a well-known fact, and it can be derived as follows. We begin with the identity

$$(3.3) \quad \int_{\mathbf{R}^2} \partial_1 \partial_2 h(x) \partial_1 \partial_2 h(x) dx = \int_{\mathbf{R}^2} \partial_1^2 h(x) \partial_2^2 h(x) dx,$$

which uses two integrations by parts. On the other hand,

$$(3.4) \quad \begin{aligned} \int_{\mathbf{R}^2} |\Delta h(x)|^2 dx &= \int_{\mathbf{R}^2} (\partial_1^2 h(x) + \partial_2^2 h(x))^2 dx \\ &= \int_{\mathbf{R}^2} (\partial_1^2 h(x))^2 + 2 \partial_1^2 h(x) \partial_2^2 h(x) + (\partial_2^2 h(x))^2 dx. \end{aligned}$$

Combining this with (3.3) we get that

$$(3.5) \quad \int_{\mathbf{R}^2} |\Delta h(x)|^2 dx - 2 \int_{\mathbf{R}^2} |\partial_1 \partial_2 h(x)|^2 dx = \int_{\mathbf{R}^2} (\partial_1^2 h(x))^2 + (\partial_2^2 h(x))^2 dx.$$

This implies (3.2), and with an extra factor of 2 on the left-hand side, because the right side of (3.5) is nonnegative. (One can improve this to get a factor of 4 on the left side of (3.2), using the right-hand side of (3.5).)

In short, the L^2 norm of Δh always bounds the L^2 norm of $\partial_1 \partial_2 h$. There are similar bounds for L^p norms when $1 < p < \infty$. Specifically, for each p in $(1, \infty)$, there is a constant $C(p)$ such that

$$(3.6) \quad \int_{\mathbf{R}^2} |\partial_1 \partial_2 h(x)|^p dx \leq C(p) \int_{\mathbf{R}^2} |\Delta h(x)|^p dx$$

whenever h is a smooth function with compact support. This is a typical example of a “Calderón-Zygmund inequality”, as in [Duo], [GarcR], [Garn], [Jou], [Sar], [Ste1], [Ste2], [SteW], [StrT], [Torc]. Such inequalities do *not* work for $p = 1$ or ∞ , and the $p = \infty$ case is like the question of supremum estimates in Section 2. Note that the $p = 1$ and $p = \infty$ cases are closely connected to each other, because of duality (of spaces and operators); the operators Δ and $\partial_1 \partial_2$ here are equal to their own transposes, with respect to the standard bilinear form on functions on \mathbf{R}^2 (defined by taking the integral of the product of two given functions). In a modestly different direction, there are classical results which give bounds in terms of the norm for Hölder continuous (or Lipschitz) functions of order α , for every $\alpha \in (0, 1)$, instead of the L^p norm. To be explicit, given α , this norm for a function g on \mathbf{R}^2 can be described as the smallest constant A such that

$$(3.7) \quad |g(x) - g(y)| \leq A|x - y|^\alpha$$

for all $x, y \in \mathbf{R}^2$. One can view this as a $p = \infty$ situation, like the L^∞ norm for g , but with a positive order α of smoothness, unlike L^∞ . There is a variety of other norms and spaces which one can consider, and for which there are results about estimates along the lines of (3.6), but for the norm in question instead of the L^p norm.

The $p = \infty$ version of (3.6) would say that there is a constant C such that

$$(3.8) \quad \sup_{x \in \mathbf{R}^2} |\partial_1 \partial_2 h(x)| \leq C \sup_{x \in \mathbf{R}^2} |\Delta h(x)|$$

whenever h is smooth and has compact support. In order to see that this is not the case, consider the function $h(x)$ given by

$$(3.9) \quad h(x) = x_1 x_2 \log(x_1^2 + x_2^2),$$

$x = (x_1, x_2)$. It is not hard to compute Δh and $\partial_1 \partial_2 h$ explicitly, and to see that Δh is bounded while $\partial_1 \partial_2 h$ is not. Indeed,

$$(3.10) \quad \partial_1 \partial_2 h(x) = \log(x_1^2 + x_2^2) + \text{bounded terms},$$

while the logarithm does not survive in Δh , because $\Delta(x_1 x_2) \equiv 0$.

This choice of h is neither smooth nor compactly supported, but these defects can be corrected easily. For smoothness we can consider instead

$$(3.11) \quad h_\epsilon(x) = x_1 x_2 \log(x_1^2 + x_2^2 + \epsilon),$$

where $\epsilon > 0$, and then look at what happens as $\epsilon \rightarrow 0$. To make the support compact we can simply multiply by a fixed cut-off function that does not vanish at the origin. With these modifications we still get a singularity at the origin as $\epsilon \rightarrow 0$, and we see that (3.8) cannot be true (with a fixed constant C that does not depend on h).

This is exactly analogous to what happened in Section 2, i.e., with a uniform bound going in but not coming out. Instead of a uniform bound for the output, we also have a substitute in terms of “mean oscillation”, just as before. To be precise, let D be any disk in \mathbf{R}^2 of radius r , and consider the quantity

$$(3.12) \quad \frac{1}{\pi r^2} \int_D |\partial_1 \partial_2 h(x) - \text{Average}_D(\partial_1 \partial_2 h)| dx,$$

where “Average $_D$ $\partial_1 \partial_2 h$ ” is the average of $\partial_1 \partial_2 h$ over the disk D , i.e.,

$$(3.13) \quad \text{Average}_D(\partial_1 \partial_2 h) = \frac{1}{\pi r^2} \int_D \partial_1 \partial_2 h(u) du.$$

Instead of (3.8), it is true that there is a constant $C > 0$ so that

$$(3.14) \quad \frac{1}{\pi r^2} \int_D |\partial_1 \partial_2 h(x) - \text{Average}_D(\partial_1 \partial_2 h)| dx \leq C \sup_{x \in \mathbf{R}^2} |\Delta h(x)|$$

for every disk D in \mathbf{R}^2 of radius r and every smooth function h with compact support. This is not too hard to prove; roughly speaking, the point is to “localize” the L^2 estimate that we had before. (Results of this nature are discussed in [Duo], [GarcR], [Garn], [Jou], [Sar], [Ste2], [Torc]. There are also replacements for (3.6) for $p = 1$, and even for $0 < p < 1$ (!); in this connection, see [Ste1], [SteW], [StrT], in addition to the references just mentioned.)

Let us formalize this estimate by defining a new space of functions, namely the space BMO of functions of *bounded mean oscillation*, introduced by John and Nirenberg in [JohN]. A locally-integrable function g on \mathbf{R}^2 is said to lie in BMO if there is a nonnegative number k such that

$$(3.15) \quad \frac{1}{\pi r^2} \int_D |g(x) - \text{Average}_D(g)| \, dx \leq k$$

for every disk D in \mathbf{R}^2 of radius r . In this case we set

$$(3.16) \quad \|g\|_* = \sup_D \frac{1}{\pi r^2} \int_D |g(x) - \text{Average}_D(g)| \, dx,$$

with the supremum taken over all disks D in \mathbf{R}^2 . This is the same as saying that $\|g\|_*$ is the *smallest* number k that satisfies (3.15). One refers to $\|g\|_*$ as the “BMO norm of g ”, but notice that $\|g\|_* = 0$ when g is equal to a constant almost everywhere. (The converse is also true.)

This definition may look a little crazy, but it works quite well in practice. Let us reformulate (3.14) by saying that there is a constant C so that

$$(3.17) \quad \|\partial_1 \partial_2 h\|_* \leq C \|\Delta h\|_\infty,$$

where $\|\phi\|_\infty$ denotes the L^∞ norm of a given function ϕ . In other words, although the L^∞ norm of $\partial_1 \partial_2 h$ is not controlled (for all h) by the L^∞ norm of Δh , the BMO norm of $\partial_1 \partial_2 h$ is controlled by the L^∞ norm of Δh .

Similarly, one of the main points in Section 2 can be reformulated as saying that if a mapping $f: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ distorts distances by only a small amount, as in (2.1), then the BMO norm $\|df\|_*$ of the differential of f is small (and with precise estimates being available).

In Section 2 we mentioned a stronger estimate with exponential decay in the measure of certain “bad” sets. This works for all BMO functions, and can be given as follows. Suppose that g is a BMO function on \mathbf{R}^2 with $\|g\|_* \leq 1$, and let D be a disk in \mathbf{R}^2 with radius r . As in (2.12), consider the “distribution function” $P(\lambda)$ defined by

$$(3.18) \quad P(\lambda) = \text{Probability}(\{x \in D : |g(x) - \text{Average}_D(g)| \geq \lambda\}),$$

where “Probability” means Lebesgue measure divided by the area πr^2 of D . Under these conditions, there is a universal bound for $P(\lambda)$ with exponential decay, i.e., an inequality of the form

$$(3.19) \quad P(\lambda) \leq B^{-\lambda} \quad \text{for } \lambda \geq 1,$$

where B is a positive number greater than 1, and B does not depend on g or D . This is a theorem of John and Nirenberg [JohN].

Although we have restricted ourselves to \mathbf{R}^2 here for simplicity, everything goes over in a natural way to Euclidean spaces of arbitrary dimension. In fact, there is a much more general framework of “spaces of homogeneous type” in which basic properties of BMO (and other aspects of real-variable harmonic analysis) carry over. See [CoiW1], [CoiW2], and compare also with [GarcR], [Ste2]. This framework includes certain Carnot spaces that arise in several complex variables, like the unit sphere in \mathbf{C}^n with the appropriate (noneuclidean) metric.

The exponential decay bound in (3.19) helps to make precise the idea that BMO functions are very close to being bounded (which would correspond to having $P(\lambda) = 0$ for all sufficiently large λ). The exponential rate of decay implies that BMO functions lie in L^p locally for all finite p , but it is quite a bit stronger than that.

A basic example of a BMO function is $\log|x|$. This is not hard to check, and it shows that exponential decay in (3.19) is sharp, i.e., one does not have superexponential decay in general. This example also fits with (3.10), and with the “rotational” part of the differential of the mapping f in (2.10).

In general, BMO functions can be much more complicated than the logarithm. Roughly speaking, the total “size” of the unboundedness is no worse than for the logarithm, as in (3.19), but the arrangement of the singularities can be more intricate, just as one can make much more complex singular examples than in (3.9) and (2.10). There are a lot of tools available in harmonic analysis for understanding how BMO functions behave. (See [Duo], [GarcR], [Garn], [Jou], [Sar], [Ste2], [StrT], [Torc], for instance.)

BMO functions show up all over the place. One can reformulate the basic scenario in this section with the Laplacian and $\partial_1\partial_2$ by saying that the pseudodifferential or singular integral operator

$$(3.20) \quad \frac{\partial_1\partial_2}{\Delta}$$

maps L^∞ to BMO, and this holds for similar operators (of order 0) much more generally (as in the references above). This will be discussed a bit further in Appendix A. Note that the nonlinear problem in Section 2 has a natural linearization which falls into this rubric. (See Appendix A.)

Sobolev embeddings provide another class of linear problems in which BMO comes up naturally. One might wish that a function g on \mathbf{R}^n that satisfies $\nabla g \in L^n(\mathbf{R}^n)$ (in the sense of distributions) were bounded or continuous, but neither of these are true in general, when $n > 1$. However, such a function g is always in BMO, and in the subspace VMO

(“vanishing mean oscillation”), in which the measurements of mean oscillation (as in the left side of (3.15) when $n = 2$) tend to 0 as the radius r goes to 0. This is a well-known analogue of continuity in the context of BMO. (See [BreN], [GarcR], [Garn], [Sar], [Sem8], [Ste2], [Torc].)

BMO arises in a lot of nonlinear problems, in addition to the one in Section 2. For instance, there are circumstances in which one might wish that the derivative of a conformal mapping in the complex plane were bounded, and it is not, but there are natural estimates in terms of BMO. More precisely, it is BMO for the *logarithm* of the derivative that comes up most naturally. This is closely related to BMO conditions for tangents to curves under certain geometric conditions. See [CoiMe1], [CoiMe2], [CoiMe3], [Davi1], [JeK1], [Pom1], [Pom2], [Pom3], [Sar], for instance. Some basic computations related to the latter were given in Section 2, near the end. In general dimensions (larger than 1), BMO shows up naturally as the logarithm of the density for harmonic measure for Lipschitz domains, and for the logarithm of Jacobians of quasiconformal mappings. See [Dah1], [Dah2], [JeK2], [Geh2], [Rei], [Ste2] and the references therein. In all dimensions, there are interesting classes of “weights”, positive functions which one can use as densities for modifications of Lebesgue measure, whose logarithms lie in BMO, and which in fact correspond to open subsets of BMO (for real-valued functions). These weights have good properties concerning L^p boundedness of singular integral and other operators, and they also show up in other situations, in connection with conformal mappings in the plane, harmonic measure, and Jacobians of quasiconformal mappings in particular, as above. See [Duo], [GarcR], [Garn], [Jou], [Sar], [Ste2], [StrT], [Torc] for information about these classes of weights.

There is a simple reason for BMO functions to arise frequently as some kind of logarithm. In many nonlinear problems there is a symmetry which permits one to multiply some quantity by a constant without changing anything in a significant way. (E.g., think of rescaling or rotating a domain, or a mapping, or multiplying a weight by a positive constant.) At the level of the logarithm this invariance is converted into a freedom to add constants, and this is something that BMO accommodates automatically.

To summarize a bit, there are a lot of situations in which one has some function that one would like to be bounded, but it is not, and for which BMO provides a good substitute. One may not expect at first to have to take measure theory into account, but then it comes up on its own, or works in a natural or reasonable way.

Before leaving this section, let us return to the John-Nirenberg theorem, i.e., the exponential decay estimate in (3.19). How might one try to prove this? The first main point is that one cannot prove (3.19) for a particular disk D using only a bound like (3.15) for that one disk. That would only give a rate of decay on the order of $1/\lambda$. Instead one uses (3.15) over and over again, for many different disks.

Here is a basic strategy. Assume that g is a BMO function with $\|g\|_* \leq 1$. First use (3.15) for D itself (with $k = 1$) to obtain that the set of points x in D such that

$$(3.21) \quad |g(x) - \text{Average}_D(g)| \geq 10,$$

is pretty small (in terms of probability). On the bad set where this happens, try to make a good covering by smaller disks on which one can apply the same type of argument. The idea is to then show that the set of points x in D which satisfy

$$(3.22) \quad |g(x) - \text{Average}_D(g)| \geq 10 + 10$$

is significantly smaller still, and by a definite proportion. If one can repeat this forever, then one can get exponential decay as in (3.19). More precisely, at each stage the size of the deviation of $g(x)$ from $\text{Average}_D(g)$ will increase by the *addition* of 10, while the decrease in the measure of the bad set will decrease *multiplicatively*.

This strategy is roughly correct in spirit, but to carry it out one has to be more careful in the choice of “bad” set at each stage, and in the transition from one stage to the next. In particular, one should try to control the difference between the average of g over one disk and over one of the smaller disks created in the next step of the process. As a practical matter, it is simpler to work with cubes instead of disks, for the way that they can be decomposed evenly into smaller pieces. The actual construction used is the “Calderón-Zygmund decomposition”, which itself has a lot of other applications. See [JohN], [Duo], [GarcR], [Garn], [Jou], [Sar], [Sem8], [Ste2], [StrT], [Torc] for more information.

4. Quantitative topology, and calculus on singular spaces

One of the nice features of Euclidean spaces is that it is easy to work with functions, derivatives, and integrals. Here is a basic example of this. Let f be a real-valued function on \mathbf{R}^n which is continuously differentiable and has compact support, and fix a point $x \in \mathbf{R}^n$. Then

$$(4.1) \quad |f(x)| \leq \frac{1}{\nu_n} \int_{\mathbf{R}^n} \frac{1}{|x-y|^{n-1}} |\nabla f(y)| dy,$$

where ν_n denotes the $(n - 1)$ -dimensional volume of the unit sphere in \mathbf{R}^n , and dy refers to ordinary n -dimensional volume.

This inequality provides a way to say that the values of a function are controlled by *averages* of its derivative. In this respect it is like Sobolev and isoperimetric inequalities, to which we shall return in a moment.

To prove (4.1) one can proceed as follows (as on p. 125 of [Ste1]). Let v be any element of \mathbf{R}^n with $|v| = 1$. Then

$$(4.2) \quad f(x) = - \int_0^\infty \frac{\partial}{\partial t} f(x + tv) dt,$$

by the fundamental theorem of calculus. Thus

$$(4.3) \quad |f(x)| \leq \int_0^\infty |\nabla f(x + tv)| dt.$$

This is true for every v in the unit sphere of \mathbf{R}^n , and by averaging over these v 's one can derive (4.1) from (4.3).

To put this into perspective, it is helpful to look at a situation where analogous inequalities make sense but fail to hold. Imagine that one is interested in inequalities like (4.1), but for 2-dimensional surfaces in \mathbf{R}^3 instead of Euclidean spaces themselves. Let S be a smoothly embedded 2-dimensional submanifold of \mathbf{R}^3 which looks like a 2-plane with a bubble attached to it. Specifically, let us start with the union of a 2-plane P and a standard (round) 2-dimensional sphere Σ which is tangent to P at a single point z . Then cut out a little neighborhood of z , and glue in a small “neck” as a bridge between the plane and the sphere to get a smooth surface S .

If the neck in S is very small compared to the size of Σ , then this is bad for an inequality like (4.1). Indeed, let x be the point on Σ which is exactly opposite from z , and consider a smooth function f which is equal to 1 on most of Σ (and at x in particular) and equal to 0 on most of P . More precisely, let us choose f so that its gradient is concentrated near the bridge between Σ and P . If f makes the transition from vanishing to being 1 in a reasonable manner, then the integral of $|\nabla f|$ on S will be very small. This is not hard to check, and it is bad for having an inequality like (4.1), since the left-hand side would be 1 and the right-hand side would be small. In particular, one could not have uniform bounds that would work for arbitrarily small bridges between P and Σ .

The inequality (4.1) is a relative of the usual Sobolev and isoperimetric inequalities, which say the following. Fix a dimension n again, and an exponent p that satisfies $1 \leq p < n$. Define q by $1/q = 1/p - 1/n$, so that $p < q < \infty$. The Sobolev inequalities assert the existence of a

constant $C(n, p)$ such that

$$(4.4) \quad \left(\int_{\mathbf{R}^n} |f(x)|^q dx \right)^{\frac{1}{q}} \leq C(n, p) \left(\int_{\mathbf{R}^n} |\nabla f(x)|^p dx \right)^{\frac{1}{p}}$$

for all functions f on \mathbf{R}^n that are continuously differentiable and have compact support. One can also allow more general functions, with ∇f interpreted in the sense of distributions.

The *isoperimetric inequality* states that if D is a domain in \mathbf{R}^n (which is bounded and has reasonably smooth boundary, say), then

$$(4.5) \quad \begin{aligned} & n\text{-dimensional volume of } D \\ & \leq C(n) ((n-1)\text{-dimensional volume of } \partial D)^{\frac{n}{n-1}}. \end{aligned}$$

This is really just a special case of (4.4), with $p = 1$ and f taken to be the characteristic function of D (i.e., the function that is equal to 1 on D and 0 on the complement of D). In this case ∇f is a (vector-valued) measure, and the right-hand side of (4.4) should be interpreted accordingly. Conversely, Sobolev inequalities for all p can be derived from isoperimetric inequalities, by applying the latter to sets of the form

$$(4.6) \quad \{x \in \mathbf{R}^n : |f(x)| > t\},$$

and then making suitable integrations in t .

The sharp version of the isoperimetric inequality states that (4.5) holds with the constant $C(n)$ that gives equality in the case of a ball. See [Fed]. One can also determine sharp constants for (4.4), as on p. 39 of [Aub].

Note that the choice of the exponent $n/(n-1)$ in the right side of (4.5) is determined by scaling considerations, i.e., in looking what happens to the two sides of (4.5) when one dilates the domain D by a positive factor. The same is true of the relationship between p and q in (4.4), and the power $n-1$ in the kernel on the right side of (4.1).

The inequality (4.1) is a basic ingredient in one of the standard methods for proving Sobolev and isoperimetric inequalities (but not necessarily with sharp constants). Roughly speaking, once one has (4.1), the rest of the argument works at a very general level of integral operators on measure spaces, rather than manifolds and derivatives. This is not quite true for the $p = 1$ case of (4.4), for which the general measure-theoretic argument gives a slightly weaker result. See Chapter V of [Ste1] for details. The slightly weaker result does give an isoperimetric inequality (4.5), and it is not hard to recover the $p = 1$ case of (4.4) from the

weaker version using a bit more of the localization properties of the gradient than are kept in (4.1). (See also Appendix C of [Sem6], especially Proposition C.14.)

The idea of these inequalities makes sense much more broadly than just on Euclidean spaces, but they may not always work very well, as in the earlier example with bubbling. To consider this further, let M be a smooth Riemannian manifold of dimension n , and let us assume for simplicity that M is unbounded (like \mathbf{R}^n). Let us also think of M as coming equipped with a distance function $d(x, y)$ with the usual properties ($d(x, y)$ is nonnegative, symmetric in x and y , vanishes exactly when $x = y$, and satisfies the triangle inequality). One might take $d(x, y)$ to be the geodesic distance associated to the Riemannian metric on M , but let us not restrict ourselves to this case. For instance, imagine that M is a smooth submanifold of some higher-dimensional \mathbf{R}^k , and that $d(x, y)$ is simply the ambient Euclidean distance $|x - y|$ inherited from \mathbf{R}^k . In general this could be much smaller than the geodesic distance.

We shall make the standing assumption that the distance $d(x, y)$ and the Riemannian geodesic distance are approximately the same, each bounded by twice the other, on sufficiently small neighborhoods about any given point in M . This ensures that $d(x, y)$ is compatible with quantities defined locally on M using the Riemannian metric, like the volume measure, and the length of the gradient of a function. Note that this local compatibility condition for the distance function $d(x, y)$ and the Riemannian metric is satisfied automatically in the situation mentioned above, where M is a submanifold of a larger Euclidean space and $d(x, y)$ is inherited from the ambient Euclidean distance. We shall also require that the distance $d(x, y)$ be compatible with the (manifold) topology on M , and that it be complete. This prevents things like infinite ends in M which asymptotically approach finite points in M with respect to $d(x, y)$.

The smoothness of M should be taken in the character of an a priori assumption, with the real point being to have bounds that do not depend on the presence of the smoothness in a quantitative way. Indeed, the smoothness of M will not really play an essential role here, but will be convenient, so that concepts like volume, gradient, and lengths of gradients are automatically meaningful.

Suppose for the moment that M is *bilipschitz equivalent* to \mathbf{R}^n equipped with the usual Euclidean metric. This means that there is a mapping ϕ from \mathbf{R}^n onto M and a constant k such that

$$(4.7) \quad k^{-1}|z - w| \leq d(\phi(z), \phi(w)) \leq k|z - w| \quad \text{for all } z, w \in \mathbf{R}^n.$$

In other words, ϕ should neither expand or shrink distances by more than a bounded amount. This implies that ϕ does not distort the corresponding Riemannian metrics or volume by more than bounded factors either, as one can readily show. In this case the analogues of (4.1), (4.4), and (4.5) all hold for M , with constants that depend only on the constants for \mathbf{R}^n and the distortion factor k . This is because any test of these inequalities on M can be “pulled back” to \mathbf{R}^n using ϕ , with the loss of information in moving between M and \mathbf{R}^n limited by the bilipschitz condition for ϕ .

This observation helps to make clear the fact that inequalities like (4.1), (4.4), and (4.5) do not really require much in the way of smoothness for the underlying space. Bounds on curvature are not preserved by bilipschitz mappings, just as bounds on higher derivative of functions are not preserved. Bilipschitz mappings can allow plenty of spiralling and corners in M (or approximate corners, since we are asking that M that be smooth a priori).

Although bilipschitz mappings are appropriate here for the small amount of regularity involved, the idea of a “parameterization” is too strong for the purposes of inequalities like (4.1), (4.4), and (4.5). One might say that these inequalities are like algebraic topology, but more quantitative, while parameterizations are more like homeomorphisms, which are always more difficult.

I would like to describe now some conditions on M which are strong enough to give bounds as in (4.1), but which are quite a bit weaker than the existence of a bilipschitz parameterization. First, let us explicitly write down the analogue of (4.1) for M . If x is any element of M , this analogue would say that there is a constant C so that

$$(4.8) \quad |f(x)| \leq C \int_M \frac{1}{d(x,y)^{n-1}} |\nabla f(y)| dVol(y)$$

for all continuously differentiable functions f with compact support on M , where $|\nabla f(y)|$ and the volume measure $dVol(y)$ are defined in terms of the Riemannian structure that comes with M .

The next two definitions give the conditions on M that we shall consider. These and similar notions have come up many times in various parts of geometry and analysis, as in [Ale], [AleV1], [AleV2], [As1], [As2], [As3], [CoiW1], [CoiW2], [Gro1], [Gro2], [HeiKo1], [HeiKo2], [HeiKo3], [HeiY], [Pet1], [Pet2], [Väi4].

Definition 4.9 (The doubling condition). A metric space $(M, d(x, y))$ is said to be *doubling* (with constant L_0) if each ball B in M with

respect to $d(x, y)$ can be covered by at most L_0 balls of half the radius of B .

Notice that Euclidean spaces are automatically doubling, with a constant L_0 that depends only on the dimension. Similarly, every subset of a (finite-dimensional) Euclidean space is doubling, with a uniform bound for its doubling constant.

Definition 4.10 (Local linear contractability). A metric space $(M, d(x, y))$ is said to be *locally linearly contractable* (with constant L_1) if the following is true. Let B be a ball in M with respect to $d(x, y)$, and with radius no greater than L_1^{-1} times the diameter of M . (Arbitrary radii are permitted when M is unbounded, as in the context of the present general discussion.) Then (local linear contractability means that) it is possible to continuously contract B to a point inside of $L_1 B$, i.e., inside the ball with the same center as B and L_1 times the radius.

This is a kind of quantitative and scale-invariant condition of local contractability. It prevents certain types of cusps or bubbling, for instance. Both this and the doubling condition hold automatically when M admits a bilipschitz parameterization by \mathbf{R}^n , with uniform bounds in terms of the bilipschitz constant k in (4.7) (and the dimension for the doubling condition).

Theorem 4.11. *If M and $d(x, y)$ are as before, and if $(M, d(x, y))$ satisfies the doubling and local linear contractability conditions with constants L_0 and L_1 , respectively, then (4.8) holds with a constant C that depends only on L_0 , L_1 , and the dimension n .*

This was proved in [Sem6]. Before we look at some aspects of the proof, some remarks are in order about what the conclusions really mean.

In general one cannot derive bounds for Sobolev and isoperimetric inequalities for M just using (4.8). One might say that (4.8) is only as good as the behavior of the volume measure on M . If the volume measure on M behaves well, with bounds for the measure of balls like ones on \mathbf{R}^n , then one can derive conclusions from (4.8) in much the same way as for Euclidean spaces. See Appendices B and C in [Sem6].

The doubling and local linear contractability conditions do not themselves say anything about the behavior of the volume on M , and indeed they tolerate fractal behavior perfectly well. To see this, consider the metric space which is \mathbf{R}^n as a set, but with the metric $|x - y|^\alpha$, where α is some fixed number in $(0, 1)$. This is a kind of abstract and higher-dimensional version of standard fractal snowflake curves in the plane. However, the doubling and local linear contractability conditions work

just as well for $(\mathbf{R}^n, |x - y|^\alpha)$ as for $(\mathbf{R}^n, |x - y|)$, just with slightly different constants.

How might one prove Theorem 4.11? It would be nice to be able to mimic the proof of (4.1), i.e., to find a family of rectifiable curves in M which go from x to infinity and whose arclength measures have approximately the same kind of distribution in M as rays in \mathbf{R}^n emanating from a given point. Such families exist (with suitable bounds) when M admits a bilipschitz parameterization by \mathbf{R}^n , and they also exist in more singular circumstances.

Unfortunately, it is not so clear how to produce families of curves like these without some explicit information about the space M in question. This problem was treated in a special case in [DaviS1], with M a certain kind of (nonsmooth) conformal deformation of \mathbf{R}^n . The basic idea was to obtain these curves from level sets of certain mappings with controlled behavior. When $n = 2$, for instance, imagine a standard square Q , with opposing vertices p and q . The boundary of Q can be thought of as a pair of paths α, β from p to q , each with two segments, two sides of Q . If τ is a function on Q which equals 0 on α and 1 on β (and is somewhat singular at p and q), then one can try to extract a family of paths from p to q in Q from the level sets

$$(4.12) \quad \{x \in \mathbf{R}^2 : \tau(x) = t\}, \quad 0 < t < 1.$$

For the standard geometry on \mathbf{R}^2 one can write down a good family of curves and a good function τ explicitly. For a certain class of conformal deformations of \mathbf{R}^2 one can make constructions of functions τ with approximately the same behavior as in the case of the standard metric, and from these one can get controlled families of curves.

These constructions of functions τ used the standard Euclidean geometry in the background in an important way. For the more general setting of Theorem 4.11 one needs to proceed somewhat differently, and it is helpful to begin with a different formulation of the kind of auxiliary functions to be used.

Given a point x in \mathbf{R}^n , there is an associated spherical projection $\pi_x : \mathbf{R}^n \setminus \{x\} \rightarrow \mathbf{S}^{n-1}$ given by

$$(4.13) \quad \pi_x(u) = \frac{u - x}{|u - x|}.$$

This projection is topologically nondegenerate, in the sense that it has degree equal to 1. Here the “degree” can be defined by restricting π_x to a sphere around x and taking the degree of this mapping (from an $(n-1)$ -dimensional sphere to another one) in the usual sense. (See [Mas],

[Mil3], [Nir] concerning the notion of degree of a mapping.) Also, this mapping satisfies the bound

$$(4.14) \quad |d\pi_x(u)| \leq C|u - x|^{-1}$$

for all $u \in \mathbf{R}^n \setminus \{x\}$, where $d\pi_x(u)$ denotes the differential of π_x at u , and C is some constant. One can write down the differential of π_x explicitly, and (4.14) can be replaced by an equality, but this precision is not needed here, and not available in general. The rays in \mathbf{R}^n that emanate from x are exactly the fibers of the mapping π_x , and bounds for the distribution of their arclength measures can be seen as a consequence of (4.14), using the “co-area theorem” [Fed], [Morg], [Sim].

One can also think of π_x as giving (4.1) in the following manner. Let ω denote the standard volume form on \mathbf{S}^{n-1} , a differential form of degree $n - 1$, and normalized so that

$$(4.15) \quad \int_{\mathbf{S}^{n-1}} \omega = 1.$$

Let λ denote the differential form on $\mathbf{R}^n \setminus \{0\}$ obtained by pulling ω back using π_x . Then (4.14) yields

$$(4.16) \quad |\lambda(u)| \leq C'|u - x|^{-n+1}$$

for all $u \in \mathbf{R}^n \setminus \{x\}$, where C' is a slightly different constant from before. In particular, λ is locally integrable across x (and smooth everywhere else). This permits one to take the exterior derivative of λ on all of \mathbf{R}^n in the (distributional) sense of currents [Fed], [Morg], and the result is that $d\lambda$ is the current of degree n which is a Dirac mass at x . More precisely, $d\lambda = 0$ away from x because ω is automatically closed (being a form of top degree on \mathbf{S}^{n-1}), and because the pull-back of a closed form is always closed. The Dirac mass at x comes from a standard Stokes’ theorem computation, which uses the observation that the integral of λ over any $(n - 1)$ -sphere in \mathbf{R}^n around x is equal to 1. (The latter is one way to formulate the fact that the degree of π_x is 1.)

This characterization of $d\lambda$ as a current on \mathbf{R}^n means that

$$(4.17) \quad \int_{\mathbf{R}^n} df \wedge \lambda = -f(x)$$

when f is a smooth function on \mathbf{R}^n with compact support. This yields (4.1), because of (4.16). (A similar use of differential forms was employed in [DaviS1].)

The general idea of the mapping π_x also makes sense in the context of Theorem 4.11. Let M , $d(y, z)$ be as before, and fix a point x in M . One

would like to find a mapping $\pi_x: M \setminus \{x\} \rightarrow \mathbf{S}^{n-1}$ which is topologically nondegenerate and satisfies

$$(4.18) \quad |d\pi_x(u)| \leq Kd(u, x)^{-1}$$

for some constant K and all $u \in M \setminus \{x\}$. Note that now the norm of the differential of π_x involves the Riemannian metric on M . For the topological nondegeneracy of π_x , let us ask that it have nonzero degree on small spheres in M that surround x in a standard way. This makes sense, because of the a priori assumption that M be smooth.

If one can produce such a mapping π_x , then one can derive (4.8) as a consequence, using the same kind of argument with differential forms as above. One can also find enough curves in the fibers of π_x , with control on the way that their arclength measures are distributed in M , through the use of co-area estimates. For this the topological nondegeneracy of π_x is needed for showing that the fibers of π_x connect x to infinity in M .

In the context of conformal deformations of \mathbf{R}^n , as in [DaviS1], such mappings π_x can be obtained as perturbations to the standard mapping in (4.13). This is described in [Sem7]. For Theorem 4.11, the method of [Sem6] does not use mappings quite like π_x , but a “stabilized” version from which one can draw similar conclusions. In this stabilized version one looks for mappings from M to \mathbf{S}^n (instead of \mathbf{S}^{n-1}) which are constant outside of a (given) ball, topologically nontrivial (in the sense of nonzero degree), and which satisfy suitable bounds on their differentials. These mappings are like snapshots of pieces of M , and one has to move them around in a controlled manner. This means moving them both in terms of location (the center of the supporting ball) and scale (the radius of the ball).

At this stage the hypotheses of Theorem 4.11 may make more sense. Existence of mappings like the ones described above is a standard matter in topology, except for the question of uniform bounds. The hypotheses of Theorem 4.11 (the doubling condition and local linear contractibility) are also in the nature of quantitative topology. Note, however, that the kind of bounds involved in the hypotheses of the theorem and the construction of mappings into spheres are somewhat different from each other, with bounds on the differentials being crucial for the latter, while control over moduli of continuity does not come up in the former. (The local linear contractibility condition restricts the overall distances by which points are displaced in the contractions, but not the sizes of the smaller-scale oscillations, as in a modulus of continuity.) In the end the

bounds for the differentials come about because the hypotheses of Theorem 4.11 permit one to reduce various constructions and comparisons to finite models of controlled complexity.

In the proof of Theorem 4.11 there are three related pieces of information that come out, namely (1) estimates for the behavior of functions on our space M in terms of their derivatives, as in (4.8), (2) families of curves in M which are well-distributed in terms of arclength measure, and (3) mappings to spheres with certain estimates and nondegeneracy properties. These three kinds of information are closely linked, through various dualities, but to some extent they also have their own lives. Each would be immediate if M had a bilipschitz parameterization by \mathbf{R}^n , but in fact they are more robust than that, and much easier to verify.

Indeed, one of the original motivations for [DaviS1] was the problem of determining which conformal deformations of \mathbf{R}^n lead to metric spaces (through the geodesic distance) which are bilipschitz equivalent to \mathbf{R}^n . The deformations are allowed to be nonsmooth here, but this does not matter too much, because of the natural scale-invariance of the problem, and because one seeks uniform bounds. This problem is the same in essence as asking which (positive) functions on \mathbf{R}^n arise as the Jacobian of a quasiconformal mapping, modulo multiplication by a positive function which is bounded and bounded away from 0.

Some natural necessary conditions are known for these questions, with a principal ingredient coming from [Geh2]. It was natural to wonder whether the necessary conditions were also sufficient. As a test for this, [DaviS1] looked at the Sobolev and related inequalities that would follow if the necessary conditions were sufficient. These inequalities could be stated directly in terms of the data of the problems, the conformal factor or prospective Jacobian. The conclusion of [DaviS1] was that these inequalities could be derived directly from the conditions on the data, independently of whether these conditions were sufficient for the existence of bilipschitz/quasiconformal mappings as above.

In [Sem5] it was shown that the candidate conditions are *not* sufficient for the existence of such mappings, at least in dimensions 3 and higher. (Dimension 2 remains open.) The simplest counterexamples involved considerations of localized fundamental groups, in much the same fashion as in Section 1. (Another class of counterexamples were based on a different mechanism, although these did not start in dimension 3.) These counterexamples are all perfectly well-behaved in terms of the doubling and local linear contractability properties, and in fact are much better than that.

Part of the bottom line here is that spaces can have geometry which behaves quite well for many purposes even if they do not behave so well in terms of parameterizations.

For some other aspects of “quantitative topology”, see [Ale], [AleV1], [AleV2], [Att1], [Att2], [BloW], [ChaF], [Che], [Fer1], [Fer2], [Fer3], [Fer4], [Geh1], [Gro1], [Gro2], [Gro3], [HeiY], [HeiS], [Luu], [Pet1], [Pet2], [TukV], [Väi2], [Väi3], [Väi4]. Related matters of Sobolev and other inequalities on non-smooth spaces come up in [HeiKo2], [HeiKo3], [HeiKo+], in connection with the behavior of quasiconformal mappings.

5. Uniform rectifiability

A basic fact in topology is that there are spaces which are *manifold factors* but not manifolds. That is, there are topological spaces M such that $M \times \mathbf{R}$ is a manifold (locally homeomorphic to a Euclidean space) but M is not. This can even happen for finite polyhedra, because of the double-suspension results of Edwards and Cannon. See [Dave2], [Edw2], [Kir] for more information.

Uniform rectifiability is a notion of controlled geometry that trades topology for estimates. It tolerates some amount of singularities, like holes and crossings, and avoids some common difficulties with homeomorphisms, such as manifold factors.

The precise definition is slightly technical, and relies on measure theory in a crucial way. In many respects it is analogous to the notion of BMO from Section 3. The following is a preliminary concept that helps to set the stage.

Definition 5.1 (Ahlfors regularity). Fix n and d , with n a positive integer and $0 < d \leq n$. A set E contained in \mathbf{R}^n is said to be (*Ahlfors*) *regular of dimension d* if it is closed, and if there is a positive Borel measure μ supported on E and a constant $C > 0$ such that

$$(5.2) \quad C^{-1}r^d \leq \mu(B(x, r)) \leq Cr^d$$

for all $x \in E$ and $0 < r \leq \text{diam } E$. Here $B(x, r)$ denotes the (open) ball with center x and radius r .

Roughly speaking, this definition asks that E behave like ordinary Euclidean space in terms of the distribution of its mass. Notice that d -planes satisfy this condition automatically, with μ equal to the ordinary d -dimensional volume. The same is true for compact smooth manifolds, and finite polyhedra which are given as unions of d -dimensional simplices (i.e., with no lower-dimensional pieces sticking off in an

isolated manner). There are also plenty of “fractal” examples, like self-similar Cantor sets and snowflake curves. In particular, the dimension d can be any (positive) real number.

A basic fact is that if E is regular and μ is as in Definition 5.1, then μ is practically the same as d -dimensional Hausdorff measure H^d restricted to E . Specifically, μ and H^d are each bounded by constant multiples of the other when applied to subsets of E . This is not hard to prove, and it shows that μ is essentially unique. Definition 5.1 could have been formulated directly in terms of Hausdorff measure, but the version above is a bit more elementary.

Let us recall the definition of a bilipschitz mapping. Let A be a set in \mathbf{R}^n , and let f be a mapping from A to some other set in \mathbf{R}^n . We say that f is k -bilipschitz, where k is a positive number, if

$$(5.3) \quad k^{-1}|x - y| \leq |f(x) - f(y)| \leq k|x - y|$$

for all $x, y \in A$.

Definition 5.4 (Uniform rectifiability). Let E be a subset of \mathbf{R}^n which is Ahlfors regular of dimension d , where d is a positive integer, $d < n$, and let μ be a positive measure on E as in Definition 5.1. Then E is *uniformly rectifiable* if there exists a positive constant k so that for each $x \in E$ and each $r > 0$ with $r \leq \text{diam } E$ there is a closed subset A of $E \cap \overline{B}(x, r)$ such that

$$(5.5) \quad \mu(A) \geq \frac{9}{10} \cdot \mu(E \cap \overline{B}(x, r))$$

and

$$(5.6) \quad \text{there is a } k\text{-bilipschitz mapping } f \text{ from } A \text{ into } \mathbf{R}^d.$$

In other words, inside of each “snapshot” $E \cap \overline{B}(x, r)$ of E there should be a large subset, with at least 90% of the points, which is bilipschitz equivalent to a subset of \mathbf{R}^d , and with a uniform bound on the bilipschitz constant. This is like asking for a controlled parameterization, except that we allow for holes and singularities.

Definition 5.4 should be compared with the classical notion of (countable) rectifiability, in which one asks that E be covered, except for a set of measure 0, by a countable union of sets, each of which is bilipschitz equivalent to a subset of \mathbf{R}^d . Uniform rectifiability implies this condition, but it is stronger, because it provides quantitative information at definite scales, while the classical notion really only gives asymptotic information as one zooms in at almost any point. See [Fal], [Fed], [Mat] for more information about classical rectifiability.

Normally one would be much happier to simply have bilipschitz coordinates outright, without having to allow for bad sets of small measure where this does not work. In practice bilipschitz coordinates simply do not exist in many situations where one might otherwise hope to have them. This is illustrated by the double-suspension spheres of Edwards and Cannon [Can1], [Can3], [Dave2], [Edw2], and the observations about them in [SieS]. Further examples are given in [Sem4], [Sem5].

The use of arbitrary scales and locations is an important part of the story here, and is very similar to the concept of BMO. At the level of a single snapshot, a fixed ball $\bar{B}(x, r)$ centered on E , the bad set may seem pretty wild, as nothing is said about what goes on there in (5.5) or (5.6). However, uniform rectifiability, like BMO, applies to all snapshots equally, and in particular to balls in which the bad set is concentrated. Thus, inside the bad set, there are in fact further controls. We shall see other manifestations of this later, and the same basic principle is used in the John-Nirenberg theorem for BMO functions (discussed in Section 3).

Uniform rectifiability provides a substitute for (complete) bilipschitz coordinates in much the same way that BMO provides a substitute for L^∞ bounds, as in Section 3. Note that L^∞ bounds and bilipschitz coordinates automatically entail uniform control over all scales and locations. This is true just because of the way they are defined, i.e., a bounded function is bounded in all snapshots, and with a uniform majorant. With BMO and uniform rectifiability the scale-invariance is imposed by hand.

It may be a little surprising that one can get anything new through concepts like BMO and uniform rectifiability. For instance, suppose that f is a locally-integrable function on \mathbf{R}^k , and that the averages

$$(5.7) \quad \frac{1}{\omega_k t^k} \int_{B(z,t)} |f(w)| dw$$

are uniformly bounded, independently of z and t . Here ω_k denotes the volume of the unit ball in \mathbf{R}^k , so that $\omega_k t^k$ is the volume of $B(z, t)$. This implies that f must itself be bounded by the same amount almost everywhere on \mathbf{R}^k , since

$$(5.8) \quad f(u) = \lim_{t \rightarrow 0} \frac{1}{\omega_k t^k} \int_{B(z,t)} f(w) dw$$

almost everywhere on \mathbf{R}^k . Thus a uniform bound for the size of the snapshots does imply a uniform bound outright. For BMO the situation is different because one asks only for a uniform bound on the mean *oscillation* in every ball. In other words, one also has the freedom to make renormalizations by additive constants when moving from place to

place, and this gives enough room for some unbounded functions, like $\log|x|$. Uniform rectifiability is like this as well, although with different kinds of “renormalizations” available.

These remarks might explain why *some* condition like uniform rectifiability could be useful or natural, but why the specific version above in particular? Part of the answer to this is that nearly all definitions of this nature are equivalent to the formulation given above. For instance, the $9/10$ in (5.5) can be replaced by any number strictly between 0 and 1. See [DaviS3], [DaviS5] for more information.

Another answer lies in a theme often articulated by Coifman, about the way that operator theory can provide a good guide for geometry. One of the original motivations for uniform rectifiability came from the “Calderón program” [Cal2], concerning the L^p -boundedness of certain singular operators on curves and surfaces of minimal smoothness. David [Davi2], [Davi3], [Davi5] showed that uniform rectifiability of a set E implies L^p -boundedness of wide classes of singular operators on E . (See [Cal1], [Cal2], [CoiDM], [CoiMcM] and the references therein for related work connected to the Calderón program.) In [DaviS3], a converse was established, so that uniform rectifiability of an Ahlfors-regular set E is actually equivalent to the boundedness of a suitable class of singular integral operators (inherited from the ambient Euclidean space \mathbf{R}^n). See also [DaviS2], [DaviS5], [MatMV], [MatP].

Here is a concrete statement about uniform rectifiability in situations where well-behaved parameterizations would be natural but may not exist.

Theorem 5.9. *Let E be a subset of \mathbf{R}^n which is regular of dimension d . If E is also a d -dimensional topological manifold and satisfies the local linear contractability condition (Definition 4.10), then E is uniformly rectifiable.*

Note that Ahlfors-regularity automatically implies the doubling condition (Definition 4.9).

Theorem 5.9 has been proved by G. David and myself. Now-a-days we have better technology, which allows for versions of this which are localized to individual “snapshots”, rather than using all scales and locations at once. See [DaviS8] (with some of the remarks in Section 12.3 of [DaviS8] helping to provide a bridge to the present formulation). We shall say a bit more about this, near the end of Subsection 5.3.

The requirement that E be a topological manifold is convenient, but weaker conditions could be used. For that matter, there are natural variations of local linear contractability too.

One can think of Theorem 5.9 and related results in the following terms. Given a compact set K , upper bounds for the d -dimensional Hausdorff measure of K together with lower bounds for the d -dimensional topology of K should lead to strong information about the geometric behavior of K . See [DaviS6], [DaviS8], [Sem3] for more on this.

To understand better what Theorem 5.9 means, let us begin by observing that the hypotheses of Theorem 5.9 would hold automatically if E were bilipschitz equivalent to \mathbf{R}^d , or if E were compact and admitted bilipschitz local coordinates from \mathbf{R}^d . Under these conditions, a test of the hypotheses of Theorem 5.9 on E can be converted into a similar test on \mathbf{R}^d , where it can then be resolved in a straightforward manner.

A similar argument shows that the hypotheses of Theorem 5.9 are “bilipschitz invariant”. More precisely, if F is another subset of \mathbf{R}^n which is bilipschitz equivalent to E , and if the hypotheses of Theorem 5.9 holds for one of E and F , then it automatically holds for the other.

Since the existence of bilipschitz coordinates implies the hypotheses of Theorem 5.9, we cannot ask for more than that in the conclusions. In other words, bilipschitz coordinates are at the high end of what one can hope for in the context of Theorem 5.9. The hypotheses of Theorem 5.9 do in fact rule out a lot of basic obstructions to the existence of bilipschitz coordinates, like cusps, fractal behavior, self-intersections and approximate self-intersections, and bubbles with very small necks. (Compare with Section 4, especially Theorem 4.11 and the discussion of its proof and consequences.) Nonetheless, it can easily happen that a set E satisfies the hypotheses of Theorem 5.9 but does not admit bilipschitz local coordinates. Double-suspension spheres provide spectacular counterexamples for this (using the observations of [SieS]). Additional counterexamples are given in [Sem4], [Sem5].

We should perhaps emphasize that the assumption of being a topological manifold in Theorem 5.9 does not involve bounds. By contrast, uniform rectifiability does involve bounds, which is part of the point. In the context of Theorem 5.9, the proof shows that the uniform rectifiability constants for the conclusion are controlled in terms of the constants that are implicit in the hypotheses, i.e., in Ahlfors-regularity, the linear contractability condition, and the dimension.

If bilipschitz coordinates are at the high end of what one could hope for, what happens if one asks for less? What if one asks for homeomorphic local coordinates with some control, but not as much? For instance, instead of bounding the “rate” of continuity through Lipschitz conditions

like

$$(5.10) \quad |f(x) - f(y)| \leq C|x - y|$$

(for some C and all x, y in the domain of f), one could work with Hölder continuity conditions, which have the form

$$(5.11) \quad |f(x) - f(y)| \leq C'|x - y|^\gamma.$$

Here γ is a positive number, sometimes called the Hölder “exponent”. As usual, (5.11) is supposed to hold simultaneously for all x and y in the domain of f , and with a fixed constant C' . When x and y are close to each other and γ is less than 1, this type of condition is strictly weaker than that of being Lipschitz. Just as $f(x) = |x|$ is a standard example of a Lipschitz function that is not differentiable at the origin, $g(x) = |x|^\gamma$ is a basic example of a function that is Hölder continuous of order γ , $\gamma \leq 1$, but not of any order larger than γ , in any neighborhood of the origin.

Instead of local coordinates which are bilipschitz, one could consider ones that are “bi-Hölder”, i.e., Hölder continuous and with Hölder continuous inverse. It turns out that double-suspension spheres do not admit bi-Hölder local coordinates when the Hölder exponent γ lies above an explicit threshold. Specifically, if P is an n -dimensional polyhedron which is the double-suspension of an $(n - 2)$ -dimensional homology sphere that is not simply connected, then there are points in P (along the “suspension circle”) for which bi-Hölder local coordinates of exponent $\gamma > 1/(n - 2)$ do not exist. This comes from the same argument as in [SieS]. More precisely, around these points in P , there do not exist homeomorphic local coordinates from subsets of \mathbf{R}^n for which the inverse mapping is Hölder continuous of order $\gamma > 1/(n - 2)$ (without requiring a Hölder condition for the mapping itself).

Given any positive number a , there are examples in [Sem5] so that local coordinates (at some points) cannot have their inverses be Hölder continuous of order a . These examples do admit bi-Hölder local coordinates (with a smaller exponent), and even “quasisymmetric” [TukV] coordinates, and they satisfy the hypotheses of Theorem 5.9. In [Sem4] there are examples which satisfy the hypotheses of Theorem 5.9, but for which no uniform modulus of continuity for local coordinate mappings and their inverses is possible (over all scales and locations).

5.1. Smoothness of Lipschitz and bilipschitz mappings.

Another aspect of uniform rectifiability is that it provides the same amount of “smoothness” as when there is a global bilipschitz parameterization. To make this precise, let us first look at the smoothness of Lipschitz and bilipschitz mappings.

A mapping $f: \mathbf{R}^d \rightarrow \mathbf{R}^n$ is Lipschitz if there is a constant C so that (5.10) holds for all x, y in \mathbf{R}^d . The space of Lipschitz mappings is a bit simpler than the space of bilipschitz mappings, because the former is a vector space (and even a Banach space) while the latter is not. For the purposes of “smoothness” properties, though, there is not really any difference between the two. Bilipschitz mappings are always Lipschitz, and anything that can happen with Lipschitz mappings can also happen with bilipschitz mappings (by adding new components, or considering $x + h(x)$ when $h(x)$ has Lipschitz norm less than 1 to get a bilipschitz mapping).

One should also not worry too much about the difference between Lipschitz mappings which are defined on all of \mathbf{R}^d , and ones that are only defined on a subset. Lipschitz mappings into \mathbf{R}^n that are defined on a subset of \mathbf{R}^d can always be extended to Lipschitz mappings on all of \mathbf{R}^d . This is a standard fact. There are also extension results for bilipschitz mappings, if one permits oneself to replace the image \mathbf{R}^n with a Euclidean space of larger dimension (which is not too serious in the present context).

For considerations of “smoothness” we might as well restrict our attention to functions which are real-valued, since the \mathbf{R}^n -valued case can always be reduced to that.

Two basic facts about Lipschitz functions on \mathbf{R}^d are that they are differentiable almost everywhere (with respect to Lebesgue measure), and that for each $\eta > 0$ they can be modified on sets of Lebesgue measure less than η (depending on the function) in such a way as to become *continuously* differentiable everywhere. See [Fed].

These are well-known results, but they do not tell the whole story. They are not quantitative; they say a lot about the asymptotic behavior (on average) of Lipschitz mappings at very small scales, but they do not say anything about what happens at scales of definite size.

To make this precise, let a Lipschitz mapping $f: \mathbf{R}^d \rightarrow \mathbf{R}$ be given, and fix a point $x \in \mathbf{R}^d$ and a radius $t > 0$. We want to measure how well f is approximated by an affine function on the ball $B(x, t)$. To do

this we define the quantity $\alpha(x, t)$ by

$$(5.12) \quad \alpha(x, t) = \inf_{A \in \mathcal{A}} \sup_{y \in B(x, t)} t^{-1} |f(y) - A(y)|.$$

Here \mathcal{A} denotes the (vector space) of affine functions on \mathbf{R}^d . The part on the right side of (5.12) with just the supremum (and not the infimum) measures how well the particular affine function A approximates f inside $B(x, t)$, and then the infimum gives us the best approximation by any affine function for a particular choice of x and t . The factor of t^{-1} makes $\alpha(x, t)$ scale properly, and be dimensionless. In particular, $\alpha(x, t)$ is uniformly bounded in x and t when f is Lipschitz, because we can take $A(y)$ to be the constant function equal to the value of f at x .

The smallness of $\alpha(x, t)$ provides a manifestation of the smoothness of f . For functions which are twice-continuously differentiable one can get estimates like

$$(5.13) \quad \alpha(x, t) = O(t),$$

using Taylor's theorem. If $0 < \delta < 1$, then estimates like

$$(5.14) \quad \alpha(x, t) = O(t^\delta)$$

(locally uniformly in x) correspond to Hölder continuity of the gradient of f of order δ . Differentiability almost everywhere of f implies that

$$(5.15) \quad \lim_{t \rightarrow 0} \alpha(x, t) = 0 \quad \text{for almost every } x.$$

This does not say anything about any particular t , because one does not know how long one might have to wait before the limiting behavior kicks in.

Here is a simple example. Let us take $d = 1$, and consider the function

$$(5.16) \quad g_\rho(x) = \rho \cdot \sin(x/\rho)$$

on \mathbf{R} . Here ρ is any positive number. Now, $g_\rho(x)$ is Lipschitz with norm 1 no matter how ρ is chosen. This is not hard to check; for instance, one can take the derivative to get that

$$(5.17) \quad g'_\rho(x) = \cos(x/\rho)$$

so that $|g'_\rho(x)| \leq 1$ everywhere. This implies that

$$(5.18) \quad |g_\rho(u) - g_\rho(v)| \leq |u - v|$$

for all u and v (and all ρ), because of the mean-value theorem, or the fundamental theorem of calculus. (One also has that $|g'_\rho(x)| = 1$ at some points, so that the Lipschitz norm is always equal to 1.)

If ρ is very small, then one has to wait a long time before the limit in (5.15) takes full effect, because $\alpha(x, t)$ will *not* be small when $t = \rho$. In fact, there is then a positive lower bound for $\alpha(x, t)$ that does not depend on x or ρ (assuming that t is taken to be equal to ρ). This is not hard to verify directly. One does not really have to worry about ρ here, because one can use scaling arguments to reduce the lower bound to the case where $\rho = 1$.

Thus a bound on the Lipschitz norm is not enough to say anything about when the limit in (5.15) will take effect. These examples work uniformly in x , so that one cannot avoid the problem by removing a set of small measure or anything like that.

However, there is something else that one can observe about these examples. Fix a ρ , no matter how large or small. The corresponding quantities $\alpha(x, t)$ will not be too small when t is equal to ρ , as mentioned above, but they will be small when t is either much smaller than ρ , or much larger than ρ . At scales much smaller than ρ , $g_\rho(x)$ is approximately affine, because the smoothness of the sine function has a chance to kick in, while at larger scales $g_\rho(x)$ is simply small outright compared to t (and one can take $A = 0$ as the approximating affine function).

In other words, for the functions $g_\rho(x)$ there is always a bad scale where the $\alpha(x, t)$'s may not be small, and that bad scale can be arbitrarily large or small, but the bad behavior is confined to approximately just one scale.

It turns out that something similar happens for arbitrary Lipschitz functions. The bad behavior cannot always be confined to a single scale—one might have sums of functions like the g_ρ 's, but with very different choices of ρ —but, on average, the bad behavior is limited to a bounded number of scales.

Let us be more precise, and define a family of functions which try to count the number of “bad” scales associated to a given point x . Fix a radius $r > 0$, and also a small number ϵ , which will provide our threshold for what is considered “small”. We assume that a Lipschitz function f on \mathbf{R}^d has been fixed, as before. Given $x \in \mathbf{R}^d$, define $N_r(x)$ to be the number of nonnegative integers j such that

$$(5.19) \quad \alpha(x, 2^{-j}r) \geq \epsilon.$$

These j 's represent the “bad” scales for the point x , and below the radius r .

It is easy to see that (5.15) implies that $N_r(x) < \infty$ for almost all x . There is a more quantitative statement which is true, namely that the average of $N_r(x)$ over any ball B in \mathbf{R}^d of radius r is finite and uniformly

bounded, independently of the ball B and the choice of r . That is,

$$(5.20) \quad r^{-d} \int_B N_r(x) dx \leq C(n, \epsilon^{-1} \|f\|_{\text{Lip}}),$$

where $C(n, s)$ is a constant that depends only on n and s , and $\|f\|_{\text{Lip}}$ denotes the Lipschitz norm of f . This is a kind of “Carleson measure condition”.

Before we get to the reason for this bound, let us consider some examples. For the functions $g_\rho(x)$ in (5.16), the functions $N_r(x)$ are simply uniformly bounded, independently of x , r , and ρ . This is not hard to check. Notice that the bound does depend on ϵ , i.e., it blows up as $\epsilon \rightarrow 0$. As another example, consider the function f defined by

$$(5.21) \quad f(x) = |x|.$$

For this function we have that $N_r(0) = \infty$ as soon as ϵ is small enough. This is because $\alpha(0, t)$ is positive and independent of t , so that (5.19) holds for all j when ϵ is sufficiently small. Thus $N_r(x)$ is not uniformly bounded in this case. In fact it has a logarithmic singularity near 0, with $N_r(x)$ behaving roughly like $\log(r/|x|)$, and this is compatible with (5.20) for B centered at 0. If one is far enough away from the origin (compared to r), then $N_r(x)$ simply vanishes, and there is nothing to do.

In general one can have mixtures of the two types of phenomena. Another interesting class of examples to consider are functions of the form

$$(5.22) \quad f(x) = \text{dist}(x, F),$$

where F is some nonempty subset of \mathbf{R}^d which is not all of \mathbf{R}^d , and $\text{dist}(x, F)$ is defined (as usual) by

$$(5.23) \quad \text{dist}(x, F) = \inf\{|x - z| : z \in F\}.$$

It is a standard exercise that such a function f is always Lipschitz with norm at most 1. Depending on the behavior of the set F , this function can have plenty of sharp corners, like $|x|$ has at the origin, and plenty of oscillations roughly like the ones in the functions g_ρ . In particular, the oscillations can occur at lots of different scales as one moves from point to point. However, one does not really have oscillations at different scales overlapping each other. Whenever the elements of F become dense enough to make a lot of oscillations, the values of f become small in compensation. (One can consider situations where F has points at regularly-spaced intervals, for instance.)

How might one prove an estimate like (5.20)? This is part of a larger story in harmonic analysis, called Littlewood-Paley theory, some of whose classical manifestations are described in [Ste1]. The present discussion is closer in spirit to [Dor] for the measurements of oscillation used, and indeed (5.20) can be derived from the results in [Dor].

There are stronger estimates available than (5.20). Instead of simply counting how often the $\alpha(x, t)$'s are larger than some threshold, as in the definition of $N_r(x)$ above, one can work with sums of the form

$$(5.24) \quad \left(\sum_{j=0}^{\infty} \alpha(x, 2^{-j}r)^q \right)^{\frac{1}{q}}.$$

The “right” choice of q is 2, but to get this one should modify the definition of $\alpha(x, t)$ (in most dimensions) so that the measurement of approximation of f by an affine function uses a suitable L^p norm, rather than the supremum. (That the choice of $q = 2$ is the “right” one reflects some underlying orthogonality, and is a basic point of Littlewood-Paley theory. At a more practical level, $q = 2$ is best because it works for the estimates for the $\alpha(x, t)$'s and allows reverse estimates for the size of the gradient of f in terms of the sizes of the $\alpha(x, t)$'s.)

In short, harmonic analysis provides a fairly thorough understanding of the sizes of the $\alpha(x, t)$'s and related quantities, and with quantitative estimates. This works for Lipschitz functions, and more generally for functions in Sobolev spaces.

There is more to the matter of smoothness of Lipschitz functions than this, however. The $\alpha(x, t)$'s measure how well a given function f can be approximated by an affine function on a ball $B(x, t)$, but they do not say too much about how these approximating affine functions might change with x and t . In fact, there are classical examples of functions for which the $\alpha(x, t)$'s tend to 0 uniformly as $t \rightarrow 0$, and yet the derivative fails to exist at almost every point. Roughly speaking, the affine approximations keep spinning around as $t \rightarrow 0$, without settling down on a particular affine function, as would happen when the derivative exists. (A faster rate of decay for the $\alpha(x, t)$'s, as in (5.14), would prevent this from happening.)

For Lipschitz functions, the existence of the differential almost everywhere implies that for almost every x the gradients of the approximating affine functions on $B(x, t)$ do not have to spin around by more than a finite amount as t goes between some fixed number r and 0. In fact, quantitative estimates are possible, in much the same manner as before. Again one fixes a threshold ϵ , and one can measure how many oscillations

of size at least ϵ there are in the gradients of the affine approximations as t ranges between 0 and r . For this there are uniform bounds on the averages of these numbers, just as in (5.20).

This type of quantitative control on the oscillations of the gradients of the affine approximations of f comes from Carleson’s Corona construction, as in [Garn]. This construction was initially applied to the behavior of bounded holomorphic functions in the unit disk of the complex plane, but in fact it is a very robust real-variable method. For example, the type of bound just mentioned in the previous paragraph (on the average number of oscillations of the gradients of the affine approximations as t goes from r to 0) is completely analogous to one for the boundary behavior of harmonic functions given in Corollary 6.2 on p. 348 of [Garn].

A more detailed discussion of the Corona construction in the context of Lipschitz functions can be found in Chapter IV.2 of [Davis5].

The Corona construction and the known estimates for affine approximations as discussed above provide a fairly complete picture of the “smoothness” of Lipschitz functions. They also provide an interesting way to look at “complexity” of Lipschitz functions, and one that is quite different from what is suggested more naively by the definition (5.10).

5.2. Smoothness and uniform rectifiability.

The preceding discussion of smoothness for Lipschitz and bilipschitz mappings has natural extensions to the geometry of sets in Euclidean spaces. Instead of approximations of functions by affine functions, one can consider approximations of sets by affine planes. Differentials of mappings correspond to tangent planes for sets.

One can think of “embedding” the discussion for functions into one for sets by taking a function and replacing it with its graph. This is consistent with the correspondence between affine functions and d -planes, and between differentials and tangent planes.

How might one generalize the $\alpha(x, t)$ ’s (5.12) to the context of sets? Fix a set E in \mathbf{R}^n and a dimension $d < n$, and let $x \in E$ and $t > 0$ be given. In analogy with (5.12), consider the quantity $\beta(x, t)$ defined by

$$(5.25) \quad \beta(x, t) = \inf_{P \in \mathcal{P}_d} \sup\{t^{-1} \text{dist}(y, P) : y \in E \cap B(x, t)\}.$$

Here \mathcal{P}_d denotes the set of d -dimensional affine planes in \mathbf{R}^n , and $\text{dist}(y, P)$ is defined as in (5.23). In other words, we take a “snapshot” of E inside the ball $B(x, t)$, and we look at the optimal degree of approximation of E by d -planes in $B(x, t)$. The factor of t^{-1} in (5.25) makes

$\beta(x, t)$ a scale-invariant, dimensionless quantity. Notice that $\beta(x, t)$ is always less than or equal to 1, no matter the behavior of E , as one can see by taking P to be any d -plane that goes through x . The smoothness of E is reflected in how small $\beta(x, t)$ is.

If E is the image of a bilipschitz mapping $\phi: \mathbf{R}^d \rightarrow \mathbf{R}^n$, then there is a simple correspondence between the $\beta(x, t)$'s on E and the $\alpha(z, s)$'s for ϕ on \mathbf{R}^d . This permits one to transfer the estimates for the α 's on \mathbf{R}^d to estimates for the β 's on E , and one could also go backwards.

It turns out that the type of estimates that one gets for the β 's in this way when E is bilipschitz equivalent to \mathbf{R}^d also work when E is uniformly rectifiable. Roughly speaking, this because the estimates for the α 's and β 's are not uniform ones, but involve some kind of integration, and in a way which is compatible with the measure-theoretic aspects of the Definition 5.4. This is very much analogous to results in the context of BMO functions, especially a theorem of Strömberg. (See Chapter IV.1 in [DaviS5] for some general statements of this nature.)

To my knowledge, the first person to look at estimates like these for sets was P. Jones [Jon1]. In particular, he used the sharp quadratic estimates that correspond to Littlewood-Paley theory to give a new approach to the L^2 boundedness of the Cauchy integral operator on nonsmooth curves. Here “quadratic” means $q = 2$ in the context of (5.24).

In [Jon3], Jones showed how quadratic estimates on the β 's could actually be used to *characterize* subsets of rectifiable curves. The quadratic nature of the estimates, which come naturally from orthogonality considerations in Littlewood-Paley theory and harmonic analysis, can, in this context, be more directly linked to the ordinary Pythagorean theorem, as in [Jon3]. A completion of Jones' results for 1-dimensional sets in Euclidean spaces of higher dimension was given in [Oki].

Analogues of Jones' results for (Ahlfors-regular) sets of higher dimension are given in [DaviS3]. More precisely, if E is a d -dimensional Ahlfors-regular set in \mathbf{R}^n , then the uniform rectifiability of E is equivalent to certain quadratic Carleson measure conditions for quantities like $\beta(x, t)$ in (5.25). One cannot use $\beta(x, t)$ itself in general, with the supremum on the right side of (5.25) (there are counterexamples due to Fang and Jones), but instead one can replace the supremum with a suitable L^p norm for a range of p 's that depends on the dimension (and is connected to Sobolev embeddings). This corresponds to the situation for sharp estimates of quantities like $\alpha(x, t)$ in the context of Lipschitz functions, as in [Dor].

The problem of building parameterizations is quite different when $d > 1$ than in the 1-dimensional case. This is a basic fact, and a recurring theme of classical topology. Making parameterizations for 1-dimensional sets is largely a matter of ordering, i.e., lining up the points in a good way. For rectifiable curves there is a canonical way to regulate the “speed” of a parameterization, using arclength. In higher dimensions none of these things are true, although conformal coordinates sometimes provide a partial substitute when $d = 2$. See [DaviS4], [MülŠ], [HeiKo1], [Sem2], [Sem4]. In [DeTY] a different kind of “normalized coordinates” are discussed for $d = 3$, but the underlying partial differential equation is unfortunately not elliptic. Part of the point of uniform rectifiability was exactly to try to come to grips with the issue of parameterizations in higher dimensions.

Although this definition (5.25) of $\beta(x, t)$ provides a natural version of the $\alpha(x, t)$'s from (5.12), it is not the only choice to consider. There is a “bilateral” version, in which one measures both the distance from points in E to the approximating d -plane (as in (5.25)) as well as distances from points in the d -plane to E . Specifically, given a set E in \mathbf{R}^n , a point $x \in E$, a radius $t > 0$, and a d -plane P in \mathbf{R}^n , set

$$(5.26) \quad \text{Approx}(E, P, x, t) = \sup\{t^{-1} \text{dist}(y, P) : y \in E \cap B(x, t)\} \\ + \sup\{t^{-1} \text{dist}(z, E) : z \in P \cap B(x, t)\}$$

and then define the bilateral version of $\beta(x, t)$ by

$$(5.27) \quad b\beta(x, t) = \inf_{P \in \mathcal{P}_d} \text{Approx}(E, P, x, t).$$

This takes “holes” in E into account, which the definition of $\beta(x, t)$ does not. For instance, $\beta(x, t) = 0$ if and only if there is a d -plane P_0 such that every point in $E \cap B(x, t)$ lies in P_0 , while for $b\beta(x, t)$ to be 0 it should also be true that every point in $P_0 \cap B(x, t)$ lies in E (assuming that E is closed, as in the definition of Ahlfors regularity).

It turns out that the $b\beta(x, t)$'s behave a bit differently from the $\beta(x, t)$'s, in the following sense. Imagine that we do not look for something like sharp quadratic estimates, as we did before, but settle for cruder “thresholding” conditions, as discussed in Subsection 5.1. In other words, one might fix an $\epsilon > 0$, and define a function $N'_r(x)$ which counts the number of times that $b\beta(x, 2^{-j} r)$ is greater than or equal to ϵ , with $j \in \mathbf{Z}_+$ (as in the discussion around (5.19)). For uniformly rectifiable sets one has bounds on the averages of $N'_r(x)$ exactly as in (5.20), but now integrating over E instead of \mathbf{R}^d . A surprising fact is that the converse is also true, i.e., estimates like these for the $b\beta$'s are

sufficient to imply the uniform rectifiability of the set E , at least if E is Ahlfors-regular of dimension d . This was proved in [DaviS5].

In the context of functions, this type of thresholding condition is too weak, in that one can have the $\alpha(x, t)$'s going to 0 uniformly as $t \rightarrow 0$ for functions which are differentiable almost nowhere, as mentioned in Subsection 5.1. Similarly, there are Ahlfors-regular sets which are "totally unrectifiable" (in the sense of [Fal], [Fed], [Mat]) and have the $\beta(x, t)$'s tending to 0 uniformly as $t \rightarrow 0$. (See [DaviS3].) For the $b\beta$'s the story is simply different. On the other hand, the fact that suitable thresholding conditions on the $b\beta$'s are sufficient to imply uniform rectifiability relies heavily on the assumption that E be Ahlfors-regular, while mass bounds are part of the conclusion (rather than the hypothesis) in Jones' results, and no counterpart to the mass bounds are included in the above-mentioned examples for functions. One does have mass bounds for the examples in [DaviS3] (of totally-unrectifiable Ahlfors-regular sets for which the $\beta(x, t)$'s tend to 0 uniformly as $t \rightarrow 0$), and there the issue is more in the size of the holes in the set. The $b\beta$'s, by definition, control the sizes of holes. Note that this result for the $b\beta$'s does have antecedents for the classical notion of (countable) rectifiability, as in [Mat].

There are a number of variants of the $b\beta$'s, in which one makes comparisons with other collections of sets besides d -planes, like unions of d -planes, for instance. See [DaviS5].

Perhaps the strongest formulation of smoothness for uniformly rectifiable sets is the existence of a "Corona decomposition". This is a geometric version of the information that one can get about a Lipschitz function from the methods of Carleson's Corona construction (as mentioned in Subsection 5.1). Roughly speaking, in this condition one controls not only how often E is well-approximated by a d -plane, but how fast the d -planes turn as well. This can also be formulated in terms of good approximations of E by flat Lipschitz graphs.

Although a bit technical, the existence of a Corona decomposition is perhaps the most useful way of managing the complexity of a uniformly rectifiable set. Once one has a Corona decomposition, it is generally pretty easy to derive whatever else one would like to know. Conversely, in practice the existence of a Corona decomposition can be a good place to start if one wants to prove that a set is uniformly rectifiable.

In fact, there is a general procedure for finding a Corona decomposition when it exists, and one which is fairly simple (and very similar to Carleson's Corona construction). The difficult part is to show that this procedure works in the right way, with the correct estimates. Specifically, it is a stopping-time argument, and one does not want to have to

stop too often. This is a nice point, because in general it is not so easy to build something like a good parameterization of a set, even if one knows a priori that it exists. In this context, there are in principle methods for doing this.

See [DaviS2], [DaviS3], [DaviS5], [DaviS7], [Sem1] for more information about Corona decompositions of uniformly rectifiable sets and the way that they can be used. The paper [DaviS7] is written in such a way as to try to convey some of the basic concepts and constructions without worrying about why the theorems are true (which is much more complicated). In particular, the basic procedure for finding Corona decompositions when they exist is discussed. See [GarnJ], [Jon1], [Jon3] for some other situations in which Carleson’s Corona construction is used geometrically.

5.3. A class of variational problems.

Uniform rectifiability is a pretty robust condition. If one has a set which looks roughly as though it ought to be uniformly rectifiable, then there is a good chance that it is. This as opposed to sets which look roughly as though they should admit a well-behaved (homeomorphic) parameterization, and do not (as discussed before).

In this subsection we would like to briefly mention a result of this type, concerning a minimal surface problem with nonsmooth coefficients. Let $g(x)$ be a Borel measurable function on \mathbf{R}^n , and assume that g is positive, bounded, and bounded away from 0, so that

$$(5.28) \quad 0 < m \leq g(x) \leq M$$

for some constants m, M and all $x \in \mathbf{R}^n$. Let Q_0, Q_1 be a pair of (closed) cubes in \mathbf{R}^n , with sides parallel to the axes, and assume that Q_0 is contained in the interior of Q_1 .

Let U be an open subset of Q_1 which contains the interior of Q_0 . Consider an integral like

$$(5.29) \quad \int_{\partial U} g(x) \, d\nu_U(x),$$

where $d\nu_U$ denotes the measure that describes the $(n - 1)$ -dimensional volume of subsets of ∂U . This would be defined as in calculus when ∂U is at least a little bit smooth (like C^1), but in general one has to be more careful. One can simply take for $d\nu_U$ the restriction of $(n - 1)$ -dimensional Hausdorff measure to ∂U , but for technical reasons it is often better to define $d\nu_U$ using distributional derivatives of the characteristic function of U , as in [Giu]. For this one would work with sets U which have “finite

perimeter”, which means exactly that the distributional first derivatives of the characteristic function of U are measures of finite mass.

Here is one way in which this kind of functional, and the minimization of this kind of functional, can come up. Let F be a closed subset of \mathbf{R}^n . Imagine that one is particularly interested in domains U which have their boundary contained in F , or very nearly so. On the other hand, one might also wish to limit irregularities in the behavior of the boundary of U . For this type of situation one could choose g so that it is much smaller on F than on the complement of F , and then look for minimizers of (5.29) to find domains with a good balance between the behavior of ∂U and the desire to have it be contained (as much as possible) in F . (See [DaviS6] for an example of this.)

When do minimizers of (5.29) exist, and how do they behave? If one works with sets of finite perimeter, and if the function g is lower semi-continuous, then one can obtain the existence of minimizers through standard techniques (as in [Giu]). That is, one takes limits of minimizing sequences for (5.29) using weak compactness, and one uses the lower semi-continuity of g to get lower semi-continuity of (5.29) with respect to suitable convergence of the U 's. The latter ensures that the limit of the minimizing sequence is actually a minimum. Note that the “obstacle” conditions that U contain the interior of Q_0 and be contained in Q_1 prevents the minimization from collapsing into something trivial.

As to the behavior of minimizers of (5.29), one cannot expect much in the way of smoothness in general. For instance, if the boundary of U can be represented locally as the graph of a Lipschitz function, then U in fact minimizes (5.29) for a suitable choice of g . Specifically, one can take g to be a sufficiently small positive constant on ∂U , and to be equal to 1 everywhere else. That such a choice of g works is not very hard to establish, and more precise results are given in [DaviS6].

Conversely, minimizers of (5.29) are always Ahlfors-regular sets of dimension $n - 1$, and uniformly rectifiable. This is shown in [DaviS6], along with some additional geometric information which is sufficient to characterize the class of sets U which occur as minimizers for functionals of the form (5.29) (with g bounded and bounded away from 0). If a set U arises as the minimizer for some g , it is also a minimizer with g chosen as above, i.e., a small positive constant on ∂U and equal to 1 everywhere else.

The same regularity results work for a suitable class of “quasiminimizers” of the usual area functional, and one that includes minimizers for (5.29) as a special case.

Uniform rectifiability provides a natural level of structure for situations like this, where stronger forms of smoothness cannot be expected, but quantitative bounds are reasonable to seek. Note that properties of ordinary rectifiability always hold for boundaries of sets of finite perimeter, regardless of any minimizing or quasiminimizing properties. See [Giu].

Analogous results about regularity work for sets of higher codimension as well, although this case is more complicated technically. See [DaviS8] for more information. One can use this framework of minimization (with respect to nonsmooth coefficient functions g) as a tool for studying the structure of sets in \mathbf{R}^n with upper bounds on their d -dimensional Hausdorff measure and lower bounds for their d -dimensional topology. This brings one back to Theorem 5.9 and related questions, and in particular more “localized” versions of it.

To put it another way, minimization of functionals like these can provide useful means for obtaining “existence results” for approximate parameterizations with good behavior, through uniform rectifiability. See [DaviS6], [DaviS8]. Part of the motivation for this came from an earlier argument of Morel and Solimini [MoreS]. Their argument concerned the existence of curves containing a given set, with good properties in terms of the distribution of the arc-length measure of these curves, under more localized conditions on the given set (at all locations and scales). See Lemma 16.27 on p. 207 of [MoreS].

Appendix A. Fourier transform calculations

If $\phi(x)$ is an integrable function on \mathbf{R}^n , then its *Fourier transform* $\widehat{\phi}(\xi)$ is defined (for $\xi \in \mathbf{R}^n$) by

$$(A.1) \quad \widehat{\phi}(\xi) = \int_{\mathbf{R}^n} e^{i\langle x, \xi \rangle} \phi(x) dx.$$

Here $\langle x, \xi \rangle$ denotes the usual inner product for $x, \xi \in \mathbf{R}^n$, and $i = \sqrt{-1}$. Often one makes slightly different conventions for this definition—with some extra factors of π around, for instance—but we shall not bother with this.

A key feature of the Fourier transform is that it diagonalizes differential operators. Specifically, if ∂_k denotes the operator $\partial/\partial x_k$ on \mathbf{R}^n , then

$$(A.2) \quad (\partial_k \phi)(\xi) = i \xi_k \widehat{\phi}(\xi),$$

i.e., differentiation is converted into mere multiplication. For this one should either make some differentiability assumptions on ϕ , so that the

left side can be defined in particular, or one should interpret this equation in the sense of tempered distributions on \mathbf{R}^n . The Fourier transform also carries out this diagonalization in a controlled manner. That is, there is an explicit inversion formula (which looks a lot like the Fourier transform itself), and the Fourier transform preserves the L^2 norm of the function ϕ , except for a multiplicative constant, by the Plancherel theorem. See [Duo], [SteW], [Torc] for these and other basic facts about the Fourier transform.

Using Plancherel's theorem, it is very easy to give another proof of the L^2 estimate (3.2) from Section 3, and to derive many other inequalities of a similar nature. One can also use the Fourier transform to give a precise definition of the operator $R = \partial_j \partial_k / \Delta$, where Δ is the Laplace operator $\sum_{\ell=1}^n \partial_\ell^2$. Specifically, one can define it through the equation

$$(A.3) \quad (R\phi)\widehat{(\xi)} = \frac{\xi_j \xi_k}{|\xi|^2} \widehat{\phi}(\xi).$$

If $m(\xi)$ is any bounded function on \mathbf{R}^n , then

$$(A.4) \quad (T\phi)\widehat{(\xi)} = m(\xi) \widehat{\phi}(\xi)$$

defines a bounded operator on $L^2(\mathbf{R}^n)$. In general these operators are not bounded on L^p for any other value of p , but this is true for many of the operators that arise naturally in analysis. For instance, suppose that $m(\xi)$ is homogeneous of degree 0, so that

$$(A.5) \quad m(t\xi) = m(\xi) \quad \text{when } t > 0,$$

and that $m(\xi)$ is smooth away from the origin. Then the associated operator T is bounded on L^p for all p with $1 < p < \infty$. See [Duo], [GarcR], [Jou], [Ste1], [SteW], [StrT], [Torc]. Note that this criterion applies to the specific choice of $m(\xi)$ in (A.3) above.

For a multiplier operator as in (A.4) to be bounded on L^1 or L^∞ is even more exceptional than for L^p boundedness when $1 < p < \infty$. (See [SteW], [Torc].) For instance, if m is homogeneous, as above, and not constant, then the corresponding operator cannot be bounded on L^1 or L^∞ . However, if m is homogeneous and smooth away from the origin, then the operator T in (A.4) does determine a bounded operator from L^∞ into BMO. See [Duo], [GarcR], [Garn], [Jou], [Sar], [Ste2], [Torc]. In fact, T determines a bounded operator from BMO to itself.

Here is another example. Let ϕ now be a mapping from \mathbf{R}^2 to itself, with components ϕ_1, ϕ_2 . Consider the differential $d\phi$ of ϕ as a matrix-valued function, namely,

$$(A.6) \quad \begin{pmatrix} \partial_1\phi_1 & \partial_1\phi_2 \\ \partial_2\phi_1 & \partial_2\phi_2 \end{pmatrix}.$$

(Let us assume that ϕ is smooth enough that the differential is at least some kind of function when taken in the sense of distributions, although one can perfectly well think of $d\phi$ as a matrix-valued distribution.) Let A and S denote the antisymmetric and symmetric parts of $d\phi$, respectively, so that

$$(A.7) \quad A = \frac{d\phi - d\phi^t}{2}, \quad S = \frac{d\phi + d\phi^t}{2},$$

where $d\phi^t$ denotes the transpose of $d\phi$.

In this case of 2×2 matrices, the antisymmetric part A really contains only one piece of information, namely

$$(A.8) \quad \partial_1\phi_2 - \partial_2\phi_1.$$

It is not hard to check that this function can be reconstructed from the entries of S through operators of the form (A.4), using functions m which are homogeneous of degree 0 and smooth away from the origin. For this one should add some mild conditions on ϕ , like compact support, to avoid the possibility that S vanishes identically but A does not.

Under these conditions, we conclude that the L^p norm of A is always bounded by a constant multiple of the L^p norm of S , $1 < p < \infty$, and that the BMO norm of A is controlled by the L^∞ (or BMO) norm of S . (For the case of BMO norms, the possibility that S vanishes but A does not causes no trouble, because A will be constant in that case.)

This example is really a “linearized” version of the problem discussed in Section 2. Specifically, let us think of $f: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ as being of the form

$$(A.9) \quad f(x) = x + \epsilon\phi(x),$$

where ϵ is a small parameter. The extent to which f distorts distances is governed by the matrix-valued function $df^t df$, which we can write out as

$$(A.10) \quad df^t df = I + 4\epsilon S + \epsilon^2 d\phi^t d\phi.$$

Thus the linear term in ϵ is governed by S , while A controls the leading behavior in ϵ of the “rotational” part of df .

Appendix B. Mappings with branching

In general, there can be a lot of trouble with existence and complexity of homeomorphisms (with particular properties, like specified domain and range). If one allows mappings with *branching*, then the story can be very different.

As a basic example of this, there is a classical result originating with Alexander to the effect that any oriented pseudomanifold of dimension n admits an orientation-preserving branched covering over the n -sphere. Let us state this more carefully, and then see how it is proved.

Let M be a finite polyhedron. We assume that M is given as a finite union of n -dimensional simplices that meet only in their boundary faces (so that M is really a simplicial complex). To be a pseudomanifold means that every $(n - 1)$ -dimensional face in M arises as the boundary face of exactly *two* n -dimensional simplices. In effect this says that M looks like a manifold away from its codimension-2 skeleton (the corresponding statement for the codimension-1 skeleton being automatic). For the present purposes it would be enough to ask that every $(n - 1)$ -dimensional face in M arise as the boundary face of *at most two* n -dimensional simplices, which would be like a “pseudomanifold with boundary”.

An *orientation* for an n -dimensional pseudomanifold M means a choice of orientation (in the usual sense) for each of the constituent n -dimensional simplices in M , with compatibility of orientations of adjacent n -dimensional simplices along the common $(n - 1)$ -dimensional face. In terms of algebraic topology, this means that the sum of the n -dimensional simplices in M , with their orientations, defines an n -dimensional *cycle* on M .

For the purposes of the Alexander-type result, it will be convenient to think of the n -sphere as consisting of two standard simplices S_1 and S_2 glued together along the boundary. This is not quite a polyhedron in the usual (affine) sense, but one could easily repair this by subdividing S_1 or S_2 . We also assume that S_1 and S_2 have been oriented, and have opposite orientations relative to their common boundary.

To define a mapping from M to the n -sphere one would like to simply identify each of the constituent n -dimensional simplices in M with S_1 or S_2 in a suitable manner. Unfortunately, this does not work, even when $n = 2$, but the problem can be fixed using a *barycentric subdivision* of M . Recall that the *barycenter* of a simplex (embedded in some vector space) is the point in the interior of the simplex which is the average of the vertices of the simplex. The *set of barycenters* for M means the set of barycenters of all of the constituent simplices in M (viewed as a

simplicial complex), of all dimensions, including 0. In particular, the set of barycenters for M includes the vertices of M (which are themselves 0-dimensional simplices, and their own barycenters). The barycentric subdivision of M is a refinement of M as a simplicial complex whose vertices are exactly the set of barycenters of M . In other words, the set M as a whole does not change, just its decomposition into simplices, which is replaced by a finer decomposition.

Here is a precise description of the simplices in the barycentric subdivision of M . Let s_0, s_1, \dots, s_k be a finite sequence of simplices in M , with each s_i an i -dimensional simplex which is a face of s_{i+1} (when $i < k$). Let $b(s_i)$ denote the barycenter of s_i . Then $b(s_0), b(s_1), \dots, b(s_k)$ are affinely independent, and hence determine a k -dimensional simplex. The simplices that arise in this manner are precisely the ones used for the barycentric subdivision of M . (See p. 123 of [Spa] for more details.)

Let \tilde{V} denote the set of all vertices in the barycentric subdivision of M . This is the same as the set of points which arise as barycenters of simplices in the original version of M , and in particular we have a natural mapping from \tilde{V} to the integers $\{0, 1, \dots, n\}$, defined by associating to each point b in \tilde{V} the dimension of the simplex from which it was derived.

If T is a k -dimensional simplex in the barycentric subdivision of M , then the mapping from \tilde{V} to $\{0, 1, \dots, n\}$ just described induces a one-to-one correspondence between the $k + 1$ vertices of T and the set $\{0, 1, \dots, k\}$. This follows easily from the definitions.

We are now ready to define our mapping from M to the n -sphere. There are exactly $n + 1$ vertices in our realization of the n -sphere as the gluing of S_1 and S_2 . Let us identify these vertices with the integers from 0 to n . Thus our mapping from \tilde{V} to $\{0, 1, \dots, n\}$ can now be interpreted as a mapping from the vertices of the barycentric subdivision of M to the vertices of the n -sphere.

This mapping between vertices admits a canonical linear extension to each k -dimensional simplex, $k < n$, in the barycentric subdivision of M . For the n -dimensional simplices the extension is uniquely determined once one chooses S_1 or S_2 for the image of the simplex. Because of the orientations, there is only one natural choice of S_1 or S_2 for each n -dimensional simplex T , namely the one so that the linear mapping from T onto S_j is orientation-preserving.

In the end we get a mapping from the barycentric subdivision of M to the n -sphere which preserves orientations and which defines an affine isomorphism from each n -dimensional simplex T in the domain onto one of S_1 and S_2 . This uses the fact that our initial mapping between vertices

was always one-to-one on the set of vertices in any given simplex in the domain, by construction.

This completes the proof. We should emphasize that the singularities of the mapping from M to the n -sphere —i.e., the places where it fails to be a local homeomorphism— are confined to the codimension-2 skeleton of the barycentric subdivision of M . This is because of the orientation and pseudomanifold conditions, which ensure that if a point in M lies in the interior of an $(n - 1)$ -dimensional simplex in the barycentric subdivision of M , then the (two) adjacent n -simplices at that point are not sent to the same S_j in the image.

The idea of branching also makes sense for mappings that are not piecewise-linear, and there are well-developed notions of “controlled geometry” in this case, as with the classes of quasiregular mappings and mappings of bounded length distortion. See [HeiKiM], [MarRiV1], [MarRiV2], [MarRiV3], [MarV], [Res], [Ric1], [Väi1], [Vuo], for instance. In [HeiR1], [HeiR2] there are examples where branching maps of controlled geometry can be constructed but suitable homeomorphisms either do not exist or must distort distances more severely.

Sullivan [Sul2], [Sul3] has proposed some mechanisms by which the existence of local (controlled) branching maps can be deduced, and some ideas for studying obstructions to controlled homeomorphic coordinates.

See [Gut+], [HeiKi], [MarRyV] for some recent results about branching and regularity conditions under which it does not occur. A broader and more detailed discussion of mappings with branching can be found in [HeiR2]. For some real-variable considerations of mappings which may branch but enjoy substantial geometric properties, see [Davi4], [Jon2], [DaviS4].

References

- [Ale] P. ALESTALO, Uniform domains of higher order, *Ann. Acad. Sci. Fenn. Ser. A I Math. Dissertationes* **94** (1994), 48.
- [AleV1] P. ALESTALO AND J. VÄISÄLÄ, Uniform domains of higher order II, *Ann. Acad. Sci. Fenn. Ser. A I Math.* **21(2)** (1996), 411–437.
- [AleV2] P. ALESTALO AND J. VÄISÄLÄ, Uniform domains of higher order III, *Ann. Acad. Sci. Fenn. Ser. A I Math.* **22(2)** (1997), 445–464.
- [As1] P. ASSOUD, *Espaces métriques, plongements, facteurs*, Thèse de doctorat, U.E.R. Mathématique, Université Paris XI, Orsay, France, 1977.

- [As2] P. ASSOUD, Étude d'une dimension métrique liée à la possibilité de plongements dans \mathbf{R}^n , *C. R. Acad. Sci. Paris Sér. A-B* **288(15)** (1979), A731–A734.
- [As3] P. ASSOUD, Plongements lipschitziens dans \mathbf{R}^n , *Bull. Soc. Math. France* **111(4)** (1983), 429–448.
- [Att1] O. ATTIE, Quasi-isometry classification of some manifolds of bounded geometry, *Math. Z.* **216(4)** (1994), 501–527.
- [Att2] O. ATTIE, *Quasi-isometry classification of coarse Hadamard manifolds*, DIMACS Technical Report 98-47, October 1998.
- [Aub] T. AUBIN, “*Nonlinear analysis on manifolds. Monge-Ampère equations*”, Grundlehren der Mathematischen Wissenschaften **252**, Springer-Verlag, New York, 1982.
- [Bin1] R. H. BING, A surface is tame if its complement is 1-ULC, *Trans. Amer. Math. Soc.* **101** (1961), 294–305.
- [Bin2] R. H. BING, “*The geometric topology of 3-manifolds*”, American Mathematical Society Colloquium Publications **40**, American Mathematical Society, Providence, RI, 1983.
- [Bin3] R. H. BING, “*Collected papers*”, Vol. 1, 2, American Mathematical Society, Providence, RI, 1988.
- [BloW] J. BLOCK AND S. WEINBERGER, Large scale homology theories and geometry, in: “*Geometric topology*” (Athens, GA, 1993), AMS/IP Stud. Adv. Math., 2.1, Amer. Math. Soc., Providence, RI, 1997, pp. 522–569.
- [BooHP] W. W. BOONE, W. HAKEN AND V. POENARU, On recursively unsolvable problems in topology and their classification, in: “*Contributions to Math. Logic*” (Colloquium, Hannover, 1966), North-Holland, Amsterdam, 1968, pp. 37–74.
- [BreN] H. BREZIS AND L. NIRENBERG, Degree theory and BMO I. Compact manifolds without boundaries, *Selecta Math. (N.S.)* **1(2)** (1995), 197–263.
- [Bur] C. E. BURGESS, Embeddings of surfaces in Euclidean three-space, *Bull. Amer. Math. Soc.* **81(5)** (1975), 795–818.
- [BurC] C. E. BURGESS AND J. W. CANNON, Embeddings of surfaces in E^3 , *Rocky Mountain J. Math.* **1(2)** (1971), 259–344.
- [Cal1] A.-P. CALDERÓN, Cauchy integrals on Lipschitz curves and related operators, *Proc. Nat. Acad. Sci. U.S.A.* **74(4)** (1977), 1324–1327.
- [Cal2] A.-P. CALDERÓN, Commutators, singular integrals on Lipschitz curves and applications, in: “*Proceedings of the International Congress of Mathematicians*” (Helsinki, 1978), Acad. Sci. Fennica, Helsinki, 1980, pp. 85–96.

- [Can1] J. W. CANNON, The characterization of topological manifolds of dimension $n \geq 5$, in: “*Proceedings of the International Congress of Mathematicians*” (Helsinki, 1978), Acad. Sci. Fennica, Helsinki, 1980, pp. 449–454.
- [Can2] J. W. CANNON, The recognition problem: what is a topological manifold?, *Bull. Amer. Math. Soc.* **84(5)** (1978), 832–866.
- [Can3] J. W. CANNON, Shrinking cell-like decompositions of manifolds. Codimension three, *Ann. of Math. (2)* **110(1)** (1979), 83–112.
- [ChaF] T. A. CHAPMAN AND S. C. FERRY, Approximating homotopy equivalences by homeomorphisms, *Amer. J. Math.* **101(3)** (1979), 583–607.
- [Che] J. CHEEGER, Finiteness theorems for Riemannian manifolds, *Amer. J. Math.* **92** (1970), 61–74.
- [CoiDM] R. R. COIFMAN, G. DAVID AND Y. MEYER, La solution de conjecture de Calderón, *Adv. in Math.* **48(2)** (1983), 144–148.
- [CoiMcM] 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. (2)* **116(2)** (1982), 361–387.
- [CoiMe1] R. R. COIFMAN AND Y. MEYER, Le théorème de Calderón par les “méthodes de variable réelle”, *C. R. Acad. Sci. Paris Sér. A-B* **289(7)** (1979), A425–A428.
- [CoiMe2] R. R. COIFMAN AND Y. MEYER, Une généralisation du théorème de Calderón sur l’intégrale de Cauchy, in “*Fourier analysis*” (Proc. Sem., El Escorial, 1979), Asoc. Mat. Española **1**, Madrid, 1980, pp. 87–116.
- [CoiMe3] R. R. COIFMAN AND Y. MEYER, Lavrentiev’s curves and conformal mapping, Institut Mittag-Leffler, Report no. 5, 1983.
- [CoiW1] R. R. COIFMAN AND G. WEISS, “*Analyse harmonique non-commutative sur certains espaces homogènes*”, Lecture Notes in Mathematics **242**, Springer-Verlag, Berlin, 1971.
- [CoiW2] R. R. COIFMAN AND G. WEISS, Extensions of Hardy spaces and their use in analysis, *Bull. Amer. Math. Soc.* **83(4)** (1977), 569–645.
- [Dah1] B. E. J. DAHLBERG, Harmonic functions in Lipschitz domains, in: “*Harmonic analysis in Euclidean spaces*”, Part 1, Proc. Sympos. Pure Math. **35**, Amer. Math. Soc., Providence, R.I., 1979, pp. 313–322.

- [Dah2] B. E. J. DAHLBERG, Real analysis and potential theory, in: “*Proceedings of the International Congress of Mathematicians*”, Vol. 1, 2 (Warsaw, 1983), PWN, 1984, pp. 953–959.
- [Dave1] R. J. DAVERMAN, Embeddings of $(n - 1)$ -spheres in Euclidean n -space, *Bull. Amer. Math. Soc.* **84(3)** (1978), 377–405.
- [Dave2] R. J. DAVERMAN, “*Decompositions of manifolds*”, Pure and Applied Mathematics **124**, Academic Press Inc., Orlando, FL, 1986.
- [Davi1] G. DAVID, Courbes de Lavrentiev et intégrales singulières, Thèse de 3^o cycle, Université de Paris-Sud, 1981.
- [Davi2] G. DAVID, Opérateurs intégraux singuliers sur certaines courbes du plan complexe, *Ann. Sci. École Norm. Sup. (4)* **17(1)** (1984), 157–189.
- [Davi3] G. DAVID, Opérateurs d’intégrale singulière sur les surfaces régulières, *Ann. Sci. École Norm. Sup. (4)* **21(2)** (1988), 225–258.
- [Davi4] G. DAVID, Morceaux de graphes lipschitziens et intégrales singulières sur une surface, *Rev. Mat. Iberoamericana* **4(1)** (1988), 73–114.
- [Davi5] G. DAVID, “*Wavelets and singular integrals on curves and surfaces*”, Lecture Notes in Mathematics **1465**, Springer-Verlag, Berlin, 1991.
- [DaviS1] G. DAVID AND S. SEMMES, Strong A_∞ weights, Sobolev inequalities and quasiconformal mappings, in “*Analysis and partial differential equations*”, Lecture Notes in Pure and Appl. Math. **122**, Dekker, New York, 1990, pp. 101–111.
- [DaviS2] G. DAVID AND S. SEMMES, Harmonic analysis and the geometry of subsets of \mathbf{R}^n , *Publ. Mat.* **35(1)** (1991), 237–249.
- [DaviS3] G. DAVID AND S. SEMMES, Singular integrals and rectifiable sets in \mathbf{R}^n : Beyond Lipschitz graphs, *Astérisque* **193** (1991), 152.
- [DaviS4] G. DAVID AND S. SEMMES, Quantitative rectifiability and Lipschitz mappings, *Trans. Amer. Math. Soc.* **337(2)** (1993), 855–889.
- [DaviS5] G. DAVID AND S. SEMMES, “*Analysis of and on uniformly rectifiable sets*”, Mathematical Surveys and Monographs **38**, American Mathematical Society, Providence, RI, 1993.

- [DaviS6] G. DAVID AND S. SEMMES, Quasiminimal surfaces of codimension 1 and John domains, *Pacific J. Math.* **183**(2) (1998), 213–277.
- [DaviS7] G. DAVID AND S. SEMMES, Sets and their approximations by surfaces, Preprint (1997).
- [DaviS8] G. DAVID AND S. SEMMES, Uniform rectifiability and quasiminimizing sets of arbitrary codimension, *Mem. Amer. Math. Soc.* **144**(687) (2000), viii+132.
- [DeTY] D. M. DETURCK AND D. YANG, Existence of elastic deformations with prescribed principal strains and triply orthogonal systems, *Duke Math. J.* **51**(2) (1984), 243–260.
- [DonK] S. K. DONALDSON AND P. B. KRONHEIMER, “*The geometry of four-manifolds*”, The Clarendon Press, Oxford University Press, New York, 1990.
- [DonS] S. K. DONALDSON AND D. P. SULLIVAN, Quasiconformal 4-manifolds, *Acta Math.* **163**(3–4) (1989), 181–252.
- [Dor] J. R. DORRONSORO, A characterization of potential spaces, *Proc. Amer. Math. Soc.* **95**(1) (1985), 21–31.
- [Duo] J. DUOANDIKOETXEA, “*Fourier analysis*”, American Mathematical Society, Providence, RI, 2001.
- [Edw1] R. D. EDWARDS, Demension theory I, in: “*Geometric topology*” (Proc. Conf., Park City, Utah, 1974), Lecture Notes in Math. **438**, Springer, Berlin, 1975, pp. 195–211.
- [Edw2] R. D. EDWARDS, The topology of manifolds and cell-like maps, in: “*Proceedings of the International Congress of Mathematicians*” (Helsinki, 1978), Acad. Sci. Fennica, Helsinki, 1980, pp. 111–127.
- [Eps] D. B. A. EPSTEIN, Projective planes in 3-manifolds, *Proc. London Math. Soc.* (3) **11** (1961), 469–484.
- [Eps+] D. B. A. EPSTEIN, J. W. CANNON, D. F. HOLT, S. V. F. LEVY, M. S. PATERSON AND W. P. THURSTON, “*Word processing in groups*”, Jones and Bartlett Publishers, Boston, MA, 1992.
- [EvaM] B. EVANS AND L. MOSER, Solvable fundamental groups of compact 3-manifolds, *Trans. Amer. Math. Soc.* **168** (1972), 189–210.
- [Fal] K. J. FALCONER, “*The geometry of fractal sets*”, Cambridge Tracts in Mathematics **85**, Cambridge University Press, Cambridge, 1986.
- [Far] B. FARB, Automatic groups: a guided tour, *Enseign. Math.* (2) **38**(3–4) (1992), 291–313.

- [Fed] H. FEDERER, “*Geometric measure theory*”, Die Grundlehren der mathematischen Wissenschaften **153**, Springer-Verlag New York Inc., New York, 1969.
- [Fer1] S. C. FERRY, Homotoping ε -maps to homeomorphisms, *Amer. J. Math.* **101(3)** (1979), 567–582.
- [Fer2] S. C. FERRY, Counting simple-homotopy types of topological manifolds, Preprint.
- [Fer3] S. C. FERRY, Topological finiteness theorems for manifolds in Gromov-Hausdorff space, *Duke Math. J.* **74(1)** (1994), 95–106.
- [Fer4] S. C. FERRY, Controlled topology and the characterization of manifolds, *Amer. Math. Soc.* (to appear).
- [Fre] M. H. FREEDMAN, The topology of four-dimensional manifolds, *J. Differential Geom.* **17(3)** (1982), 357–453.
- [FreQ] M. H. FREEDMAN AND F. QUINN, “*Topology of 4-manifolds*”, Princeton Mathematical Series **39**, Princeton University Press, Princeton, NJ, 1990.
- [Fu] J. H. G. FU, Bi-Lipschitz rough normal coordinates for surfaces with an L^1 curvature bound, *Indiana Univ. Math. J.* **47(2)** (1998), 439–453.
- [GarcR] J. GARCÍA-CUERVA AND J. L. RUBIO DE FRANCIA, “*Weighted norm inequalities and related topics*”, North-Holland Mathematics Studies **116**, North-Holland Publishing Co., Amsterdam, 1985.
- [Garn] J. B. GARNETT, “*Bounded analytic functions*”, Pure and Applied Mathematics **96**, Academic Press Inc., New York, 1981.
- [GarnJ] J. B. GARNETT AND P. W. JONES, The corona theorem for Denjoy domains, *Acta Math.* **155(1–2)** (1985), 27–40.
- [Geh1] F. W. GEHRING, The Hausdorff measure of sets which link in Euclidean space, in: “*Contributions to analysis (a collection of papers dedicated to Lipman Bers)*”, Academic Press, New York, 1974, pp. 159–167.
- [Geh2] F. W. GEHRING, The L^p -integrability of the partial derivatives of a quasiconformal mapping, *Acta Math.* **130** (1973), 265–277.
- [Giu] E. GIUSTI, “*Minimal surfaces and functions of bounded variation*”, Monographs in Mathematics **80**, Birkhäuser Verlag, Basel, 1984.

- [Gro1] M. GROMOV, “*Structures métriques pour les variétés riemanniennes*”, (J. Lafontaine and P. Pansu, eds.), Textes Mathématiques **1**, CEDIC, Paris, 1981.
- [Gro2] M. GROMOV, “*Metric structures for Riemannian and non-Riemannian spaces*”, Progress in Mathematics **152**, Birkhäuser Boston Inc., Boston, MA, 1999.
- [Gro3] M. GROMOV, Quantitative homotopy theory, in “*Prospects in mathematics*” (Princeton, NJ, 1996), Amer. Math. Soc., Providence, RI, 1999, pp. 45–49.
- [Gut+] V. Y. GUTLYANSKIĬ, O. MARTIO, V. I. RYAZANOV AND M. VUORINEN, On local injectivity and asymptotic linearity of quasiregular mappings, *Studia Math.* **128(3)** (1998), 243–271.
- [HanH] B. HANSON AND J. HEINONEN, An n -dimensional space that admits a Poincaré inequality but has no manifold points, *Proc. Amer. Math. Soc.* **128(11)** (2000), 3379–3390.
- [HeiKi] J. HEINONEN AND T. KILPELÄINEN, BLD-mappings in $W^{2,2}$ are locally invertible, *Math. Ann.* **318(2)** (2000), 391–396.
- [HeiKiM] J. HEINONEN, T. KILPELÄINEN AND O. MARTIO, “*Non-linear potential theory of degenerate elliptic equations*”, Oxford Mathematical Monographs, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1993.
- [HeiKo1] J. HEINONEN AND P. KOSKELA, Definitions of quasiconformality, *Invent. Math.* **120(1)** (1995), 61–79.
- [HeiKo2] J. HEINONEN AND P. KOSKELA, From local to global in quasiconformal structures, *Proc. Nat. Acad. Sci. U.S.A.* **93(2)** (1996), 554–556.
- [HeiKo3] J. HEINONEN AND P. KOSKELA, Quasiconformal maps in metric spaces with controlled geometry, *Acta Math.* **181(1)** (1998), 1–61.
- [HeiKo+] J. HEINONEN, P. KOSKELA, N. SHANMUGALINGAM AND J. TYSON, Sobolev classes of banach space-valued functions and quasiconformal mappings, Preprint.
- [HeiR1] J. HEINONEN AND S. RICKMAN, Quasiregular maps $\mathbf{S}^3 \rightarrow \mathbf{S}^3$ with wild branch sets, *Topology* **37(1)** (1998), 1–24.
- [HeiR2] J. HEINONEN AND S. RICKMAN, Geometric branched covers between generalized manifolds, *Duke Math. J.* (to appear).
- [HeiS] J. HEINONEN AND S. SEMMES, Thirty-three yes or no questions about mappings, measures, and metrics, *Conform. Geom. Dyn.* **1** (1997), 1–12 (electronic).

- [HeiY] J. HEINONEN AND S. S. YANG, Strongly uniform domains and periodic quasiconformal maps, *Ann. Acad. Sci. Fenn. Ser. A I Math.* **20(1)** (1995), 123–148.
- [Hem1] J. HEMPEL, “3-Manifolds”, *Ann. of Math. Studies* **86**, Princeton University Press, Princeton, N. J., 1976.
- [Hem2] J. HEMPEL, Residual finiteness for 3-manifolds, in: “Combinatorial Group Theory and Topology” (Alta, Utah, 1984) (S. Gersten and J. Stallings, eds.), *Annals of Mathematics Studies* **111**, Princeton University Press, Princeton, NJ, 1987, pp. 379–396.
- [Jac] W. JACO, “Lectures on three-manifold topology”, CBMS Regional Conference Series in Mathematics **43**, American Mathematical Society, Providence, R.I., 1980.
- [JeK1] D. S. JERISON AND C. E. KENIG, Hardy spaces, A_∞ , and singular integrals on chord-arc domains, *Math. Scand.* **50(2)** (1982), 221–247.
- [JeK2] D. S. JERISON AND C. E. KENIG, Boundary value problems on Lipschitz domains, in: “Studies in partial differential equations”, MAA Stud. Math. **23**, Math. Assoc. America, Washington, DC, 1982, pp. 1–68.
- [Joh] F. JOHN, Rotation and strain, *Comm. Pure Appl. Math.* **14** (1961), 391–413.
- [JohN] F. JOHN AND L. NIRENBERG, On functions of bounded mean oscillation, *Comm. Pure Appl. Math.* **14** (1961), 415–426.
- [Jon1] P. W. JONES, Square functions, Cauchy integrals, analytic capacity, and harmonic measure, in: “Harmonic analysis and partial differential equations” (El Escorial, 1987), *Lecture Notes in Math.* **1384**, Springer, Berlin, 1989, pp. 24–68.
- [Jon2] P. W. JONES, Lipschitz and bi-Lipschitz functions, *Rev. Mat. Iberoamericana* **4(1)** (1988), 115–121.
- [Jon3] P. W. JONES, Rectifiable sets and the traveling salesman problem, *Invent. Math.* **102(1)** (1990), 1–15.
- [Jou] J.-L. JOURNÉ, “Calderón-Zygmund operators, pseudodifferential operators and the Cauchy integral of Calderón”, *Lecture Notes in Mathematics* **994**, Springer-Verlag, Berlin, 1983.
- [KerM] M. A. KERVAIRE AND J. W. MILNOR, Groups of homotopy spheres. I, *Ann. of Math. (2)* **77** (1963), 504–537.
- [Kir] R. C. KIRBY, “The topology of 4-manifolds”, *Lecture Notes in Mathematics* **1374**, Springer-Verlag, Berlin, 1989.

- [Luu] J. LUUKKAINEN, Assouad dimension: antifractal metrization, porous sets, and homogeneous measures, *J. Korean Math. Soc.* **35(1)** (1998), 23–76.
- [Man] Y. I. MANIN, “*A course in mathematical logic*”, Graduate Texts in Mathematics **53**, Springer-Verlag, New York, 1977.
- [Mar1] A. A. MARKOV, The insolubility of the problem of homeomorphy, *Dokl. Akad. Nauk SSSR* **121** (1958), 218–220.
- [Mar2] A. A. MARKOV, Unsolvability of certain problems in topology, *Dokl. Akad. Nauk SSSR* **123** (1958), 978–980.
- [Mar3] A. A. MARKOV, Insolubility of the problem of homeomorphy, in: “*Proc. Internat. Congress Math. 1958*”, Cambridge Univ. Press, New York, 1960, pp. 300–306.
- [MarRiV1] O. MARTIO, S. RICKMAN AND J. VÄISÄLÄ, Definitions for quasiregular mappings, *Ann. Acad. Sci. Fenn. Ser. A I No.* **448** (1969), 40.
- [MarRiV2] O. MARTIO, S. RICKMAN AND J. VÄISÄLÄ, Distortion and singularities of quasiregular mappings, *Ann. Acad. Sci. Fenn. Ser. A I No.* **465** (1970), 13.
- [MarRiV3] O. MARTIO, S. RICKMAN AND J. VÄISÄLÄ, Topological and metric properties of quasiregular mappings, *Ann. Acad. Sci. Fenn. Ser. A I*(**488**) (1971), 31.
- [MarRyV] O. MARTIO, V. I. RYAZANOV AND M. VUORINEN, BMO and injectivity of space quasiregular mappings, *Math. Nachr.* **205** (1999), 149–161.
- [MarV] O. MARTIO AND J. VÄISÄLÄ, Elliptic equations and maps of bounded length distortion, *Math. Ann.* **282(3)** (1988), 423–443.
- [Mas] W. S. MASSEY, “*A basic course in algebraic topology*”, Graduate Texts in Mathematics **127**, Springer-Verlag, New York, 1991.
- [Mat] P. MATTILA, “*Geometry of sets and measures in Euclidean spaces*”, Cambridge Studies in Advanced Mathematics **44**, Cambridge University Press, Cambridge, 1995.
- [MatMV] P. MATTILA, M. S. MELNIKOV AND J. VERDERA, The Cauchy integral, analytic capacity, and uniform rectifiability, *Ann. of Math. (2)* **144(1)** (1996), 127–136.
- [MatP] P. MATTILA AND D. PREISS, Rectifiable measures in \mathbf{R}^n and existence of principal values for singular integrals, *J. London Math. Soc. (2)* **52(3)** (1995), 482–496.
- [Mil1] J. W. MILNOR, On manifolds homeomorphic to the 7-sphere, *Ann. of Math. (2)* **64** (1956), 399–405.

- [Mil2] J. W. MILNOR, Two complexes which are homeomorphic but combinatorially distinct, *Ann. of Math. (2)* **74** (1961), 575–590.
- [Mil3] J. W. MILNOR, “*Topology from the differentiable viewpoint*”, The University Press of Virginia, Charlottesville, Va., 1965.
- [Moi] E. E. MOISE, “*Geometric topology in dimensions 2 and 3*”, Graduate Texts in Mathematics **47**, Springer-Verlag, New York, 1977.
- [MoreS] J.-M. MOREL AND S. SOLIMINI, “*Variational methods in image segmentation*”, Progress in Nonlinear Differential Equations and their Applications **14**, Birkhäuser Boston Inc., Boston, MA, 1995.
- [Morg] F. MORGAN, “*Geometric measure theory*”, Academic Press Inc., Boston, MA, 1988.
- [MülŠ] S. MÜLLER AND V. ŠVERÁK, On surfaces of finite total curvature, *J. Differential Geom.* **42(2)** (1995), 229–258.
- [Nab] A. NABUTOVSKY, Einstein structures: existence versus uniqueness, *Geom. Funct. Anal.* **5(1)** (1995), 76–91.
- [NabW1] A. NABUTOVSKY AND S. WEINBERGER, Algorithmic unsolvability of the triviality problem for multidimensional knots, *Comment. Math. Helv.* **71(3)** (1996), 426–434.
- [NabW2] A. NABUTOVSKY AND S. WEINBERGER, Algorithmic aspects of homeomorphism problems, in: “*Tel Aviv Topology Conference: Rothenberg Festschrift (1998)*”, *Contemp. Math.* **231**, Amer. Math. Soc., Providence, RI, 1999, pp. 245–250.
- [Nir] L. NIRENBERG, “*Topics in Nonlinear Functional Analysis*”, Courant Institute of Mathematical Sciences, New York, 1974.
- [Oki] K. OKIKIOLU, Characterization of subsets of rectifiable curves in \mathbf{R}^n , *J. London Math. Soc. (2)* **46(2)** (1992), 336–348.
- [Pet1] P. PETERSEN, A finiteness theorem for metric spaces, *J. Differential Geom.* **31(2)** (1990), 387–395.
- [Pet2] P. PETERSEN, Gromov-Hausdorff convergence of metric spaces, in: “*Differential geometry: Riemannian geometry*”, Part 3 (Los Angeles, CA, 1990), *Proc. Sympos. Pure Math.* **54**, Amer. Math. Soc., Providence, RI, 1993, pp. 489–504.

- [Pom1] C. POMMERENKE, Schlichte funktionen und analytische funktionen von beschränkter mittlerer oszillation, *Comment. Math. Helv.* **52(4)** (1977), 591–602.
- [Pom2] C. POMMERENKE, On univalent functions, Bloch functions and VMOA, *Math. Ann.* **236(3)** (1978), 199–208.
- [Pom3] C. POMMERENKE, Boundary behaviour of conformal mappings, in: “*Aspects of contemporary complex analysis*” (Proc. NATO Adv. Study Inst., Univ. Durham, Durham, 1979), Academic Press, London, 1980, pp. 313–331.
- [Rei] H. M. REIMANN, Functions of bounded mean oscillation and quasiconformal mappings, *Comment. Math. Helv.* **49** (1974), 260–276.
- [Res] Y. G. RESHETNYAK, “*Space mappings with bounded distortion*”, Translations of Mathematical Monographs **73**, American Mathematical Society, Providence, RI, 1989.
- [Ric1] S. RICKMAN, “*Quasiregular mappings*”, *Ergebnisse der Mathematik und ihrer Grenzgebiete (3)* **26**, Springer-Verlag, Berlin, 1993.
- [Ric2] S. RICKMAN, Construction of quasiregular mappings, in: “*Quasiconformal mappings and analysis*” (Ann Arbor, MI, 1995), Springer, New York, 1998, pp. 337–345.
- [RouS] C. P. ROURKE AND B. J. SANDERSON, “*Introduction to piecewise-linear topology*”, *Ergebnisse der Mathematik und ihrer Grenzgebiete* **69**, Springer-Verlag, New York, 1972.
- [Rub1] J. H. RUBINSTEIN, An algorithm to recognize the 3-sphere, in: “*Proceedings of the International Congress of Mathematicians*”, Vol. 1, 2 (Zürich, 1994), Birkhäuser, Basel, 1995, pp. 601–611.
- [Rub2] J. H. RUBINSTEIN, Polyhedral minimal surfaces, Heegaard splittings and decision problems for 3-dimensional manifolds, in “*Geometric topology*” (Athens, GA, 1993), AMS/IP Stud. Adv. Math., 2.1, Amer. Math. Soc., Providence, RI, 1997, pp. 1–20.
- [Rus1] T. B. RUSHING, “*Topological embeddings*”, Pure and Applied Mathematics **52**, Academic Press, New York, 1973.
- [Rus2] T. B. RUSHING, Hausdorff dimension of wild fractals, *Trans. Amer. Math. Soc.* **334(2)** (1992), 597–613.
- [Sar] D. SARASON, “*Function theory on the unit circle*”, Notes for lectures given at a Conference at Virginia Polytechnic Institute and State University, Blacksburg, Va., 1978, 19–23.

- [Sco] G. P. SCOTT, Finitely generated 3-manifold groups are finitely presented, *J. London Math. Soc. (2)* **6** (1973), 437–440.
- [Sem1] S. SEMMES, Analysis vs. geometry on a class of rectifiable hypersurfaces in \mathbf{R}^n , *Indiana Univ. Math. J.* **39(4)** (1990), 1005–1035.
- [Sem2] S. SEMMES, Chord-arc surfaces with small constant II. Good parameterizations, *Adv. Math.* **88(2)** (1991), 170–199.
- [Sem3] S. SEMMES, Finding structure in sets with little smoothness, in: “*Proceedings of the International Congress of Mathematicians*”, Vol. 1, 2 (Zürich, 1994), Birkhäuser, Basel, 1995, pp. 875–885.
- [Sem4] S. SEMMES, Good metric spaces without good parameterizations, *Rev. Mat. Iberoamericana* **12(1)** (1996), 187–275.
- [Sem5] S. SEMMES, On the nonexistence of bi-Lipschitz parameterizations and geometric problems about A_∞ -weights, *Rev. Mat. Iberoamericana* **12(2)** (1996), 337–410.
- [Sem6] S. SEMMES, Finding curves on general spaces through quantitative topology, with applications to Sobolev and Poincaré inequalities, *Selecta Math. (N.S.)* **2(2)** (1996), 155–295.
- [Sem7] S. SEMMES, Some remarks about metric spaces, spherical mappings, functions and their derivatives, *Publ. Mat.* **40(2)** (1996), 411–430.
- [Sem8] S. SEMMES, Metric spaces and mappings seen at many scales, appendix in: “*Metric Structures for Riemannian and Non-Riemannian Spaces*”, Birkhäuser, Basel, 1999, pp. 875–885.
- [Sem9] S. SEMMES, Some topics concerning homeomorphic parameterizations, *Publ. Mat.* **45(1)** (2001), 3–67.
- [Sie] L. C. SIEBENMANN, Topological manifolds, in: “*Actes du Congrès International des Mathématiciens*”, Tome 2 (Nice, 1970), Gauthier-Villars, Paris, 1971, pp. 133–163.
- [SieS] L. C. SIEBENMANN AND D. P. SULLIVAN, On complexes that are Lipschitz manifolds, in: “*Geometric topology*” (Proc. Georgia Topology Conf., Athens, Ga., 1977), Academic Press, New York, 1979, pp. 503–525.
- [Sim] L. SIMON, “*Lectures on geometric measure theory*”, Proceedings of the Centre for Mathematical Analysis, Australian National University **3**, Australian National University Centre for Mathematical Analysis, Canberra, 1983.

- [Spa] E. H. SPANIER, *“Algebraic topology”*, McGraw-Hill Book Co., New York, 1966.
- [Sta] J. R. STALLINGS, On infinite processes leading to differentiability in the complement of a point, in: *“Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse)”*, Princeton Univ. Press, Princeton, N.J., 1965, pp. 245–254.
- [Ste1] E. M. STEIN, *“Singular integrals and differentiability properties of functions”*, Princeton Mathematical Series **30**, Princeton University Press, Princeton, N.J., 1970.
- [Ste2] E. M. STEIN, *“Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals”*, Princeton Mathematical Series **43**, Princeton University Press, Princeton, NJ, 1993.
- [SteW] E. M. STEIN AND G. WEISS, *“Introduction to Fourier Analysis on Euclidean Spaces”*, Princeton Mathematical Series **32**, Princeton University Press, Princeton, NJ, 1971.
- [StrT] J.-O. STRÖMBERG AND A. TORCHINSKY, *“Weighted Hardy spaces”*, Lecture Notes in Mathematics **1381**, Springer-Verlag, Berlin, 1989.
- [Sul1] D. P. SULLIVAN, Hyperbolic geometry and homeomorphisms, in: *“Geometric topology”* (Proc. Georgia Topology Conf., Athens, Ga., 1977), Academic Press, New York, 1979, pp. 543–555.
- [Sul2] D. P. SULLIVAN, Exterior d , the local degree, and smoothability, in: *“Prospects in topology”* (Princeton, NJ, 1994), Princeton Univ. Press, Princeton, NJ, 1995, pp. 328–338.
- [Sul3] D. P. SULLIVAN, On the foundation of geometry, analysis, and the differentiable structure for manifolds, Lezioni Guido Fubini, Scientific pamphlet, Institute for Scientific Interchange Foundation, Torino, Italy, 1996; also appears in *Topics in Low-Dimensional Topology in honor of Steve Armentrout* (A. Banyaga, H. Movahedi-Lankarani, and R. Wells, eds.), World Scientific, 1999, pp. 89–92.
- [Tho] A. THOMPSON, Algorithmic recognition of 3-manifolds, *Bull. Amer. Math. Soc. (N.S.)* **35(1)** (1998), 57–66.
- [Thu] W. P. THURSTON, *“Three-dimensional geometry and topology”*, Vol. 1, Princeton Mathematical Series **35**, Princeton University Press, Princeton, NJ, 1997.

- [Torc] A. TORCHINSKY, “*Real-Variable Methods in Harmonic Analysis*”, Pure and Applied Mathematics **123**, Academic Press, Inc., Orlando, FL, 1986.
- [Toro1] T. TORO, Surfaces with generalized second fundamental form in L^2 are Lipschitz manifolds, *J. Differential Geom.* **39(1)** (1994), 65–101.
- [Toro2] T. TORO, Geometric conditions and existence of bilipschitz parameterizations, *Duke Math. J.* **77(1)** (1995), 193–227.
- [TukV] P. TUKIA AND J. VÄISÄLÄ, Quasisymmetric embeddings of metric spaces, *Ann. Acad. Sci. Fenn. Ser. A I Math.* **5(1)** (1980), 97–114.
- [Väi1] J. VÄISÄLÄ, A survey of quasiregular maps in \mathbf{R}^n , in: “*Proceedings of the International Congress of Mathematicians*” (Helsinki, 1978), Acad. Sci. Fennica, Helsinki, 1980, pp. 685–691.
- [Väi2] J. VÄISÄLÄ, Quasisymmetric embeddings in Euclidean spaces, *Trans. Amer. Math. Soc.* **264(1)** (1981), 191–204.
- [Väi3] J. VÄISÄLÄ, The wall conjecture on domains in Euclidean spaces, *Manuscripta Math.* **93(4)** (1997), 515–534.
- [Väi4] J. VÄISÄLÄ, Metric duality in Euclidean spaces, *Math. Scand.* **80(2)** (1997), 249–288.
- [VolKF] I. A. VOLODIN, V. E. KUZNECOV AND A. T. FOMENKO, The problem of the algorithmic discrimination of the standard three-dimensional sphere, *Uspehi Mat. Nauk* **29(5(179))** (1974), 71–168.
- [Vuo] M. VUORINEN, “*Conformal geometry and quasiregular mappings*”, Lecture Notes in Mathematics **1319**, Springer-Verlag, Berlin, 1988.

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

Primera versió rebuda el 15 d'octubre de 2000,
 darrera versió rebuda el 20 d'abril de 2001.