SOME REMARKS ABOUT METRIC SPACES, SPHERICAL MAPPINGS, FUNCTIONS AND THEIR DERIVATIVES

Stephen Semmes

Abstract _

If $p \in \mathbf{R}^n$, then we have the radial projection map from $\mathbf{R}^n \setminus \{p\}$ onto a sphere. Sometimes one can construct similar mappings on metric spaces even when the space is nontrivially different from Euclidean space, so that the existence of such a mapping becomes a sign of approximately Euclidean geometry. The existence of such spherical mappings can be used to derive estimates for the values of a function in terms of its gradient, which can then be used to derive Sobolev inequalities, etc. In this paper we shall discuss these topics mostly in the context of metric doubling measures, which provides a nontrivial setting in which these mappings exist and can be used. This provides an alternative approach (or understanding) of the results in [**DS**], and a variation on the themes of [**Se4**].

1. Introduction.

Consider Euclidean space \mathbf{R}^n for a moment, and let p be any point in \mathbf{R}^n . Define a mapping $\pi_p : \mathbf{R}^n \setminus \{p\} \to \mathbf{S}^{n-1}$ by

(1.1)
$$\pi_p(x) = \frac{x-p}{|x-p|}$$

Now suppose that (M, d(x, y)) is a metric space. Assume that M is homeomorphic to \mathbf{R}^n , but we want to permit it to have very different geometry. Given a point $p \in M$, let us ask ourselves whether we can find a mapping $\theta_p : M \setminus \{p\} \to \mathbf{S}^{n-1}$ such that θ_p satisfies the "Lipschitz" condition

(1.2)
$$|\theta_p(x) - \theta_p(y)| \le C \frac{d(x,y)}{\min(d(x,p), d(y,p))}$$

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

for some constant C and all $x, y \in M \setminus \{p\}$, and such that θ_p is topologically nondegenerate in the sense of having nonzero degree, or degree 1, or being homotopic to π_p . (To talk about the degree in this case we use the fact that $M \setminus \{p\}$ has the same homotopy type as \mathbf{S}^{n-1} .)

The Lipschitz condition in (1.2) is the natural one in this context, it amounts to saying that θ_p is locally Lipschitz with norm bounded by a constant divided by the distance to the singularity p. Note that π_p has this property (with respect to the Euclidean metric).

The existence of such a mapping θ_p indicates approximate Euclidean behavior of the geometry of M near p. For Euclidean space itself we can simply take $\theta_p = \pi_p$, and if M is bilipschitz equivalent to a Euclidean space then we also get mappings like θ_p . We shall see that this is also true if M is quasisymmetrically equivalent to Euclidean space, although this case is a bit trickier.

Note however that there are plenty of examples of metric spaces whose geometry is very much like Euclidean geometry in some ways but which are still very different from Euclidean spaces. Different in the absence of a bilipschitz or quasisymmetric parameterization, for instance. See **[Se2]**, **[Se3]**. The existence of a mapping θ_p can capture some geometric structure that one might want even when there is not a good parameterization.

The existence of mappings like θ_p enjoy a naive geometric appeal, but they also have a concrete usefulness. Under suitable conditions they can permit us to control the value of a function at the point p in terms of the gradient of the function on the whole space, in a manner analogous to classical formulas on Euclidean space. Once one has such estimates one can then try to derive Sobolev inequalities, etc.

As a practical matter we shall focus here on the metric spaces which arise from metric doubling measures. This concept will be recalled in the next section. In $[\mathbf{DS}]$ it was proved that one has for metric doubling measures practically the same Sobolev and Poincaré inequalities as for Euclidean spaces. The observations of this paper provide a different approach to this result, or rather a better understanding of the approach in $[\mathbf{DS}]$, with clearer general principles.

This paper is also a variation on the theme of [Se4]. In [Se4] too we were constructing special mappings on metric spaces and then using these mappings to derive Sobolev inequalities. The present paper and [Se4] share common underlying principles, but the special mappings in [Se4] were slightly less beautiful than the ones given here, while the constructions in [Se4] were more flexible if more complicated and less sharp. The present paper will also be a kind of complement to [Se4], in that [Se4] was long and pretty thorough about technicalities and generality, whereas here we shall try to present the main points clearly and concisely in a special case while skipping technical but fairly straightforward discussions of the various consequences of the method.

We review metric doubling measures in the next section, and then discuss special mappings into spheres in Section 3 for non-Euclidean metrics. This construction is applied in Section 5 to control the value of a function in terms of an integral of its gradient, with bounds that are appropriate for the geometry of a metric doubling measure. Section 4 before that addresses a technical point about smoothing the maps obtained in Section 3. In Section 6 we explain the principles of this method more succinctly in terms of differential forms, and we describe some other ways to use these principles.

Some readers might find it pleasant to skip to Section 6 first.

I would like to thank Juha Heinonen and the referee for their comments and suggestions.

2. Metric doubling measures.

Definition 2.1. A *doubling measure* on \mathbb{R}^n is a positive Borel measure μ which is finite on compact sets, not identically zero, and which satisfies

(2.2)
$$\mu(2B) \le C\,\mu(B)$$

for some constant C > 0 and all (Euclidean) balls B in \mathbb{R}^n . Here 2B denotes the ball with the same center as B but twice the radius.

Given a doubling measure μ on \mathbf{R}^n we can associate a quasimetric $\delta(x, y) = \delta_{\mu}(x, y)$ on \mathbf{R}^n by

(2.3)
$$\delta(x,y) = \mu(B_{x,y})^{\frac{1}{n}},$$

where $B_{x,y}$ is the smallest (closed) ball which contains x and y. To say that $\delta(x, y)$ is a quasimetric means that it is nonnegative and symmetric, that it vanishes exactly on the diagonal, and that it satisfies

(2.4)
$$\delta(x,z) \le K \left(\delta(x,y) + \delta(y,z)\right)$$

for some constant K and all $x, y, z \in \mathbb{R}^n$. It is not hard to check (2.4) using the doubling condition on μ , and also that this quasimetric defines the standard topology on \mathbb{R}^n .

Definition 2.5. A doubling measure μ on \mathbb{R}^n is said to be a *metric* doubling measure if there is a metric $\delta'(x, y)$ on \mathbb{R}^n and a constant C > 0so that

(2.6)
$$C^{-1}\delta(x,y) \le \delta'(x,y) \le C\,\delta(x,y)$$

for all $x, y \in \mathbf{R}^n$.

For the record, a metric is a symmetric nonnegative function which vanishes exactly on the diagonal and which satisfies the triangle inequality. It is the latter which really matters here, i.e., replacing K with 1 in (2.4). As a practical matter the triangle inequality is very useful because it permits one to construct plenty of Lipschitz functions. On a quasimetric space there may be no nonconstant Lipschitz functions, e.g., $(\mathbf{R}^n, |x - y|^2)$.

When n = 1, $\delta(x, y)$ itself satisfies the triangle inequality, but this can fail in a strong way in higher dimensions. For instance, the measure $\mu(x) = |x_1| dx$ on \mathbf{R}^2 is a doubling measure but not a metric doubling measure. The failure of the metric doubling measure condition stems from the fact that the density $|x_1|$ vanishes on the x_1 -axis, which would force all points on the x_1 -axis to have $\delta(\cdot, \cdot)$ distance zero from each other if $\delta(\cdot, \cdot)$ were equivalent in size to a true metric. More generally if μ is a metric doubling measure of the form $\mu = \omega(x) dx$, where $\omega(x)$ is, say, continuous, then ω cannot be identically zero on a rectifiable curve. On the other hand, metric doubling measures can have a density ω that vanishes to large order, even on a large set. The zero set of ω has to be highly disconnected or crinkled.

The notion of a metric doubling measure was introduced in $[\mathbf{DS}]$, in a slightly different but equivalent form and with a different name ("strong A_{∞} weights"). See also [Se1], [Se2], for more examples especially.

Note that if d(x, y) is any quasimetric on a set M, then there exists a metric $\rho(x, y)$ on M and constants C > 0 and $s \ge 1$ so that

(2.7)
$$C^{-1} \rho(x, y)^s \le d(x, y) \le C \rho(x, y)^s$$

for all $x, y \in M$. This is essentially the content of the proof of Theorem 2 in [**MS**]. Thus the quasimetric associated to a metric doubling measure always has a metric around, the issue is whether we can take s = 1 in (2.7).

The following is a basic and gorgeous fact about metric doubling measures.

Theorem 2.8. If n > 1 and μ is a metric doubling measure on \mathbb{R}^n , then μ and Lebesgue measure are mutually absolutely continuous.

Thus every metric doubling measure is of the form $\mu = \omega(x) dx$ for a locally integrable function $\omega(x)$. The weights $\omega(x)$ that arise this way are called strong A_{∞} weights. It turns out that they are always A_{∞} weights in the sense of harmonic analysis (see [Ga], [J] for the definition).

Note that doubling measures can be singular with respect to Lebesgue measure. It is quite amusing that the geometric property of being a metric doubling measure is sufficient to ensure absolute continuity.

Theorem 2.8 was observed in [DS], which is to say that it was observed in [DS] that Theorem 2.8 follows from an argument of Gehring [Ge]. See also Proposition 3.4 in [Se1].

If μ is a metric doubling measure on \mathbf{R}^n , then we get a metric space $(\mathbf{R}^n, \delta'(x, y))$ determined by μ . (Actually the metric $\delta'(\cdot, \cdot)$ is not determined uniquely, but it is determined up to size, meaning that any other such metric must be bounded above and below by constant multiples of $\delta'(\cdot, \cdot)$, and that is sufficient for our purposes.) What can we say about the geometry of this metric space? We can view Theorem 2.8 as a rectifiability result for this metric space, because it implies that this metric space is asymptotically Euclidean at almost all points (the Lebesgue points of $\omega(x)$). The A_{∞} condition that ω satisfies actually implies a kind of uniform rectifiability property of this metric space. The main result in [DS] said that one has the same kind of Sobolev-Poincaré inequalities for this space as if it were bilipschitz equivalent to \mathbf{R}^n equipped with the standard Euclidean metric. One of the main points of $[\mathbf{DS}]$ was to shed some light on the question of whether these metric spaces always are bilipschitz equivalent to standard Euclidean spaces. This is trivially true in dimension 1, but it turns out to be false in dimensions greater than or equal to 3. Dimension 2 is unknown. See [Se2].

Note that $(\mathbf{R}^n, \delta'(x, y))$ is bilipschitz equivalent to $(\mathbf{R}^n, |x - y|)$ if and only if there is a quasiconformal mapping $f : \mathbf{R}^n \to \mathbf{R}^n$ whose Jacobian is bounded above and below by constant multiples of ω . This is well known and not difficult to verify using the definitions.

3. Constructing special mappings.

Let d(x, y) be a metric on \mathbb{R}^n . In this section we shall assume that the identity is a quasisymmetric mapping from $(\mathbb{R}^n, d(x, y))$ onto \mathbb{R}^n equipped with the Euclidean metric. This means that there is a homeomorphism $\eta : [0, \infty) \to [0, \infty)$ such that

(3.1)
$$d(x,y) \le t \, d(x,z) \quad \text{implies} \quad |x-y| \le \eta(t) \, |x-z|$$

for all $x, y, z \in \mathbf{R}^n$ and t > 0. This implies in particular that d(x, y) determines the same topology as the Euclidean metric (and indeed it implies a similar condition with the roles of d(x, y) and the Euclidean metric reversed).

Examples 3.2. (a) If $d(x,y) = |x - y|^s$ for some $0 < s \le 1$, then d(x,y) satisfies (3.1). (It also satisfies (3.1) when s > 1, but it is not a metric then.)

(b) If μ is a doubling measure on \mathbb{R}^n , then the associated quasidistance $\delta(x, y)$ satisfies the quasisymmetry property (3.1). This is well known and not hard to derive from the definitions. It will not be a metric in general, though, but in the case of a metric doubling measure it will be equivalent in size to a metric.

As indicated in the introduction, we would like to build mappings which approximate π_p as defined in (1.1) but which respect the d(x, y)-geometry as in (1.2).

Theorem 3.3. Suppose that d(x, y) is a metric on \mathbb{R}^n which satisfies the quasisymmetry condition (3.1). Then for each $p \in \mathbb{R}^n$ and $\epsilon > 0$ we can find a mapping $\theta : \mathbb{R}^n \setminus \{p\} \to \mathbb{S}^{n-1}$ which satisfies

(3.4)
$$\left| \theta(x) - \frac{x-p}{|x-p|} \right| \le \epsilon \text{ and}$$

(3.5)
$$|\theta(x) - \theta(y)| \le C(\epsilon) \frac{d(x,y)}{\min(d(x,p), d(y,p))}$$

for all $x, y \in \mathbf{R}^n \setminus \{p\}$. Here $C(\epsilon)$ depends also on the function $\eta(t)$ in (3.1) and the dimension n but nothing else.

It is important here that d(x, y) be a metric and not a quasimetric, we need to have the triangle inequality in order to have an adequate supply of Lipschitz functions with respect to d(x, y).

Note that (3.4) (for a reasonable choice of ϵ) implies that θ is homotopic to π_p through continuous mappings from $\mathbf{R}^n \setminus \{p\}$ into \mathbf{S}^{n-1} .

To prove Theorem 3.3 we use the techniques of the Whitney Extension Theorem ([St]), namely Whitney decompositions and partitions of unity.

All constants in the argument that follows are permitted to depend on the function $\eta(t)$ from (3.1).

Let $p \in \mathbf{R}^n$ and $\epsilon > 0$ be given, with $\epsilon < .01$, say. We may as well assume that ϵ is as small as we want, the conclusion gets stronger as ϵ gets smaller.

Let $\{Q_i\}$ be a listing of the maximal dyadic cubes in $\mathbb{R}^n \setminus \{p\}$ such that

(3.6)
$$\operatorname{diam} Q_i \le \epsilon \operatorname{dist}(Q_i, p),$$

where "dist" is the distance in the Euclidean metric. The maximality of these cubes ensures that

(3.7)
$$3 \operatorname{diam} Q_i > \epsilon \operatorname{dist}(Q_i, p).$$

416

By standard reasoning the cubes have disjoint interiors and their union is all of $\mathbb{R}^n \setminus \{p\}$. (Note that we interpret "cube" to mean closed cube.) Let q_i denote the center of Q_i .

We want to have a partition of unity associated to $\{Q_i\}$, but we want this partition of unity to be well-behaved with respect to the geometry of d(x, y) rather than Euclidean geometry. We want the functions in the partition of unity to be Lipschitz with respect to d(x, y), for instance. We start by constructing auxiliary functions $\psi_i : \mathbf{R}^n \to \mathbf{R}$ with the properties that $0 \leq \psi_i \leq C$ on \mathbf{R}^n , $\psi_i \geq 1$ on Q_i , $\psi_i \equiv 0$ on $\mathbf{R}^n \setminus 2Q_i$, and the Lipschitz condition

(3.8)
$$|\psi_i(x) - \psi_i(y)| \le C \frac{d(x,y)}{\min(d(x,p), d(y,p))}$$

This constant C is permitted to depend on ϵ as well as $\eta(t)$. The denominator on the right side may look a bit odd, but it is convenient for deriving (3.5). We define $\psi_i(x)$ in terms of the metric $d(\cdot, \cdot)$ as follows. Given $x \in \mathbf{R}^n$ we start by taking its distance to $\mathbf{R}^n \setminus 2Q_i$ with respect to $d(\cdot, \cdot)$. This gives a function $\alpha_i(x)$ which is Lipschitz with respect to $d(\cdot, \cdot)$ with norm ≤ 1 and which vanishes on $\mathbf{R}^n \setminus 2Q_i$ but which has the wrong size. We then divide $\alpha_i(x)$ by the distance D_i with respect to $d(\cdot, \cdot)$ between Q_i and $\mathbf{R}^n \setminus 2Q_i$ and set $\psi_i(x) = \alpha_i(x)/D_i$. This function is at least 1 on Q_i . It is not hard to see that $\psi_i(x)$ is also bounded, using the quasisymmetry hypothesis. It remains to check the Lipschitz condition (3.8). If x and y both lie outside $2Q_i$, then the left side of (3.8) vanishes, and there is nothing to check. If exactly one of x and y lies in $2Q_i$ and the other lies outside $3Q_i$ —say $x \in 2Q_i$ and $y \in \mathbf{R}^n \setminus 3Q_i$ then $|x-y| \ge C^{-1} \epsilon |x-p|$ and the right side of (3.8) is bounded from below, because of the quasisymmetry condition. In this case (3.8) is automatic since ψ_i is bounded. Finally, if both x and y lie in $3Q_i$, then |x-p| and |y-p| are both comparable in size to the Euclidean distance between Q_i and $\mathbf{R}^n \setminus 2Q_i$ (by (3.6), (3.7), and the requirement that ϵ be small), and so d(x, p) and d(y, p) are both comparable in size to the distance D_i with respect to $d(\cdot, \cdot)$ between Q_i and $\mathbf{R}^n \setminus 2Q_i$, because of the quasisymmetry property. (Note that the constants here depend on ϵ , but we do not mind.) From here the Lipschitz condition (3.8) follows from $\psi_i = \alpha_i / D_i$ and the fact that α_i is Lipschitz with norm ≤ 1 with respect to $d(\cdot, \cdot)$.

Thus we get our functions ψ_i . For the partition of unity we take

(3.9)
$$\phi_i = \frac{\psi_i}{\sum_j \psi_j}.$$

More precisely, although the denominator on the right hand side vanishes at p, we interpret ϕ_i to also vanish at p. Thus $0 \le \phi_i \le \psi_i \le C$, since

 $\sum_{j} \psi_{j} \geq 1$ on $\mathbf{R}^{n} \setminus \{p\}$. Notice that $\psi_{j}(x) \neq 0$ for only finitely many j for any given x, that $\phi_{i} \equiv 0$ on $\mathbf{R}^{n} \setminus 2Q_{i}$, and that

(3.10)
$$\sum_{i} \phi_{i} \equiv 1 \quad \text{on} \quad \mathbf{R}^{n} \setminus \{p\}.$$

Let us check that ϕ_i enjoys the Lipschitz condition

(3.11)
$$|\phi_i(x) - \phi_i(y)| \le C \frac{d(x,y)}{\min(d(x,p), d(y,p))}$$

We want to use (3.8). To do this we need to check that for each fixed point z the number of j's for which $\psi_j(z) \neq 0$ is uniformly bounded. If $\psi_j(z) \neq 0$, then $z \in 2Q_j$, and (3.6), (3.7) imply that diam Q_j is comparable in size to |z - p|. Thus all the Q_j 's with $z \in 2Q_j$ have approximately the same diameter, and there can only be boundedly many of them because they have disjoint interiors. Thus $\psi_j(z) \neq 0$ for only a bounded number of j's for any given z, and then (3.11) follows from (3.8) and the definition (3.9) of ϕ_i by a simple computation.

Let us now define a mapping $h : \mathbf{R}^n \setminus \{p\} \to \mathbf{R}^n$ by

(3.12)
$$h(x) = \sum_{i} \frac{q_i - p}{|q_i - p|} \phi_i(x).$$

The idea is that this is a good approximation to $\frac{x-p}{|x-p|}$. Specifically, we have that

(3.13)
$$\left| h(x) - \frac{x-p}{|x-p|} \right| \le 20 \epsilon$$

for all $x \in \mathbf{R}^n \setminus \{p\}$. To see this we write

(3.14)
$$\left| h(x) - \frac{x-p}{|x-p|} \right| \le \sum_{i} \left| \frac{q_i - p}{|q_i - p|} - \frac{x-p}{|x-p|} \right| \phi_i(x).$$

If $\phi_i(x) \neq 0$, then $x \in 2Q_i$, whence $|x-q_i| \leq 2 \operatorname{diam} Q_i \leq 2 \epsilon \operatorname{dist}(Q_i, p) \leq 2 \epsilon |q_i - p|$, by (3.6). Since ϵ is small we conclude that

(3.15)
$$\left|\frac{q_i - p}{|q_i - p|} - \frac{x - p}{|x - p|}\right| \le 20 \epsilon.$$

(There is nothing special about 20 here, but there is no point in worrying about the right constant.) This implies (3.13), because of (3.10).

Define $\theta : \mathbf{R}^n \setminus \{p\} \to \mathbf{S}^{n-1}$ by $\theta(x) = \frac{h(x)}{|h(x)|}$. This is well defined, because (3.13) ensures that |h(x)| is close to 1. From (3.13) we also get (3.4), but with an extra constant factor on the right hand side of (3.4) that does not bother us. It remains to verify the Lipschitz condition (3.5). It is enough to show that the analogous estimate holds for hinstead of θ . For h we have that

(3.16)
$$|h(x) - h(y)| \le \sum_{i} |\phi_i(x) - \phi_i(y)|.$$

If $\phi_i(x) \neq 0$, then $x \in 2Q_i$, and for each x this can hold only for a bounded number of *i*'s. Thus there are only boundedly many nonzero terms in (3.16). We can use (3.11) for each of them, and in the end we obtain the analogue of (3.5) for h.

This completes the proof of Theorem 3.3.

4. Metric doubling measures, special mappings, and smooth approximations.

Let $\mu = \omega(x) dx$ be a metric doubling measure on \mathbf{R}^n , n > 1. (Remember Theorem 2.8. We restrict ourselves to n > 1 here not because the n = 1 case does not work, but because it works in a degenerate manner that is not very interesting.) From μ we get a quasimetric $\delta(x, y)$ on \mathbf{R}^n as in (2.3), and the assumption that μ is a metric doubling measure implies that $\delta(x, y)$ is comparable in size to a true metric, as in (2.6). This metric satisfies the quasisymmetry condition (3.1), because $\delta(x, y)$ does, and so we may apply Theorem 3.3 to it. We conclude that for each $p \in \mathbf{R}^n$ we can find a mapping θ as in Theorem 3.3, but with d(x, y)replaced with $\delta(x, y)$.

Fix $p \in \mathbf{R}^n$ and $\epsilon > 0$, and let θ be as in Theorem 3.3. In practical computation we would be much happier if θ were C^1 , so that we would not have to think about applying standard rules of calculus, but of course we cannot take θ to be C^1 in general. We can approximate θ by smooth functions in the usual way, and this provides enough control for many applications. We shall deal with this technical point in this section.

Fix a nonnegative smooth function $\phi(x)$ on \mathbf{R}^n with support contained in the unit ball and $\int_{\mathbf{R}^n} \phi(x) \, dx = 1$. Define $\phi_t(x)$ for t > 0 by $\phi_t(x) = t^{-n}\phi(x/t)$. Set $\theta_t = \phi_t * \theta$, where * denotes convolution. This is a smooth function, but it is only a good approximation to θ away from the point p. That is, we are interested in topological properties of θ , its topological nontriviality as a mapping into the sphere in particular, and so we really want an approximation to θ which takes values in $\mathbf{R}^n \setminus \{0\}$. Precisely because of this topological nontriviality we cannot do this across p, make a smooth approximation across p with values in $\mathbf{R}^n \setminus \{0\}$, and so this approximation should be taken seriously away from p only. **Lemma 4.1.** If L > 1 and $|x - p| \ge 10 L t$, then

(4.2)
$$\left|\theta_t(x) - \frac{x-p}{|x-p|}\right| \le \epsilon + L^{-1}.$$

This is easy to check using (3.4) and the properties of ϕ . It implies that

(4.3)
$$\frac{1}{2} \le \left|\theta_t(x)\right| \le 2$$

when $\epsilon \leq 1/10$ (say) and |x - p| is large enough compared to t.

Lemma 4.4. $\theta_t \to \theta$ as $t \to 0$ uniformly on compact subsets of $\mathbf{R}^n \setminus \{p\}$.

This is easy to check, using the uniform continuity of θ on such compact sets.

Next let us bound the derivative of θ_t . Define ω_t by $\omega_t(x) = t^{-n} \int_{B(x,t)} \omega(y) \, dy$.

Lemma 4.5. If |x - p| > 10t then

(4.6)
$$|\nabla \theta_t(x)| \le C \, \frac{\omega_t(x)^{\frac{1}{n}}}{\delta(x,p)},$$

where C depends on $\mu = \omega(x) dx$, n, and ϵ , but not on p, x, or t.

By standard arguments (including the observation that $\nabla \phi$ has integral zero) we have that

(4.7)
$$\begin{aligned} |\nabla \theta_t(x)| &= \left| \int_{\mathbf{R}^n} (\nabla \phi_t) (x - y) \, \theta(y) \, dy \right| \\ &\leq C \, t^{-n-1} \int_{B(x,t)} |\theta(y) - \theta(x)| \, dy. \end{aligned}$$

We can bound the right hand side using (3.5) (in which we may replace d(x, y) with $\delta(x, y)$). In this case $\delta(x, p)$ and $\delta(y, p)$ are comparable in size, since |x - y| < |x - p|/10 (and the doubling/quasisymmetry condition), and so we may as well replace the denominator in (3.5) with $\delta(x, p)$. On the other hand $t^{-1} \delta(x, y) \leq C \omega_t(x)^{\frac{1}{n}}$ for all $y \in B(x, t)$, by

the definition (2.3) of $\delta(x, y)$ and the doubling condition for μ . Plugging these bounds into (4.7) we get that

(4.8)
$$|\nabla \theta_t(x)| \le C t^{-n} \int_{B(x,t)} \omega_t(x)^{\frac{1}{n}} \, \delta(x,p)^{-1} \, dy$$

Now there is no y in the integrand, and the right side collapses to the right side of (4.6). This proves Lemma 4.5.

Now set $h_t(x) = \frac{\theta_t(x)}{|\theta_t(x)|}$. This is well defined as soon as $\epsilon \leq 1/10$ (say) and |x - p| is large enough compared to t. Again we have that $h_t \to \theta$ as $t \to 0$ uniformly on compact subsets of $\mathbf{R}^n \setminus \{p\}$. We also have that $h_t(x)$ is smooth and that

(4.9)
$$|\nabla h_t(x)| \le C \frac{\omega_t(x)^{\frac{1}{n}}}{\delta(x,p)}$$

as soon as $\epsilon \leq 1/10$ and |x - p| is large enough compared to t. There is also a version of (4.2) for h_t .

In conclusion we have mappings h_t taking values in \mathbf{S}^{n-1} which are much like our original mapping θ , except that they are smooth, and that they have some minor defects, such as being defined only on the complement of a neighborhood of p, and the inequality (3.5) has to be replaced by (4.9). This is not a real problem, it is just not quite as nicely geometric.

5. Metric doubling measures and estimates for functions in terms of their gradient.

Given a smooth function with compact support on \mathbb{R}^n , one can control the values of the function in terms of integrals of the gradient against suitable potentials. This is well known and easy. (See (18) on p. 125 of **[St]**, for instance.) In this section we derive a similar estimate for metric doubling measures. We restrict ourselves to n > 1 again, not because the n = 1 doesn't work, but because it works in an uninteresting manner.

Let $\mu = \omega(x) dx$ be a metric doubling measure on \mathbb{R}^n , and let $\delta(x, y)$ be the quasimetric associated to μ as in (2.3).

Theorem 5.1. Let f be a C^1 function on \mathbb{R}^n with compact support, and fix $p \in \mathbb{R}^n$. Then

(5.2)
$$|f(p)| \le C \int_{\mathbf{R}^n} \frac{1}{\delta(x,p)^{n-1}} \,\omega(x)^{-\frac{1}{n}} \,|\nabla f(x)| \,\omega(x) \,dx.$$

Here C depends on μ and n but not on f or p.

Note that this is standard when $\omega \equiv 1$ (so that $\delta(x, y) = c |x - y|$), as in (18) on p. 125 of [**St**]. For metric doubling measures this was derived in [**DS**], through a construction which is basically equivalent to the one below, but using slightly different principles.

The integrand in (5.2) may seem to be written in a slightly ridiculous manner, with the double appearance of ω , but this is done to make the "Riemannian" nature of the expression more apparent. That is, we are really working with a non-smooth conformal deformation of the Euclidean metric, for which $\omega(x) dx$ is the natural volume, and for which $\omega(x)^{-\frac{1}{n}} |\nabla f(x)|$ is the correct norm for $\nabla f(x)$. One can show that $\delta(x, y)$ is equivalent in size to its associated geodesic distance (see [**DS**], [**Se1**]).

Notice how (5.2) becomes uninteresting when n = 1.

We are assuming here that f is C^1 and has compact support really for convenience. These assumptions could be weakened in standard ways, but this is a technical matter and not the central point.

Let us now prove Theorem 5.1. Let f and p be given. Assume first that f is constant on a neighborhood of p, on B(p, r), say.

We want to compute using Stokes' theorem. It will be convenient to use the language of differential forms. Let λ denote the standard volume form on \mathbf{S}^{n-1} , normalized to have total volume 1. Set $\epsilon = 1/10$, say, let θ be as in Theorem 3.3, and let h_t be as in Section 4. We would like to do calculus with θ , using it to pull back λ to get a differential form on $\mathbf{R}^n \setminus \{p\}$ etc., but to avoid technical nuisances it is more pleasant to work with the h_t 's. Let $\eta_t = h_t^*(\lambda)$ be the pull-back of λ by h_t , a smooth differential form of degree n - 1 that is defined on $\mathbf{R}^n \setminus B(p, r/2)$ when tis small enough. Using (4.9) and the definition of the pull-back we get that

(5.3)
$$|\eta_t(x)| \le C \left(\frac{\omega_t(x)^{\frac{1}{n}}}{\delta(x,p)}\right)^{n-1}$$

for x in the domain of h_t , where C does not depend on p, x, or t, and where ω_t is as in Section 4 (defined just before Lemma 4.5).

From Stokes' theorem we have that

(5.4)
$$\int_{\partial B(p,r)} f \eta_t = \int_{\mathbf{R}^n \setminus B(p,r)} df \wedge \eta_t$$

except perhaps for a sign error. This uses the fact that $d\eta_t = 0$ on the domain of h_t , since d commutes with the operation of pulling back, and since λ is automatically closed (because it has top degree on \mathbf{S}^{n-1}). Since f is constant on B(p, r) we have that

(5.5)
$$\int_{\partial B(p,r)} f \eta_t = f(p) \int_{\partial B(p,r)} \eta_t$$

Let us check that

(5.6)
$$\int_{\partial B(p,r)} \eta_t = 1$$

when t is sufficiently small (except perhaps for a sign error that comes from being sloppy about orientations). When t is sufficiently small we have that

(5.7)
$$\left| h_t(y) - \frac{y-p}{|y-p|} \right| \le \frac{2}{10},$$

because of (3.4) and the fact that h_t converges uniformly to θ on compact sets that do not contain p. Thus h_t defines a smooth mapping from $\partial B(p,r)$ into \mathbf{S}^{n-1} which is a small ($\leq .2$) perturbation of the obvious radial projection. This implies that the restriction of h_t to $\partial B(p,r)$ is homotopic to the obvious radial projection through smooth mappings of $\partial B(p,r)$ into \mathbf{S}^{n-1} . It is well-known that the integral in (5.6) remains constant under such homotopies (see [**N**]), and so (5.6) follows from its counterpart with h_t replaced by the obvious radial projection, in which case it is just a consequence of the usual change of variables formula.

Combining (5.6), (5.5), and (5.4) we get that

(5.8)
$$f(p) = \int_{\mathbf{R}^n \setminus B(p,r)} df \wedge \eta_t.$$

Applying (5.3) now we get that

(5.9)
$$|f(p)| \le C \int_{\mathbf{R}^n \setminus B(p,r)} |\nabla f(x)| \left(\frac{\omega_t(x)^{\frac{1}{n}}}{\delta(x,p)}\right)^{n-1} dx$$

Sending $t \to 0$ we get (5.2) in this case. (There is no problem with the convergence, because of the definition of ω_t , the fact that the other quantities are bounded on the range of integration, and the compactness of the support of f.)

Let us now prove (5.2) in the general case, assuming that it holds when f is constant on a neighborhood of p. Given r > 0 let $\chi_r(x)$ be

a smooth function on \mathbb{R}^n which vanishes off of B(p, r), takes only the value 1 on B(p, r/2), and satisfies $|\nabla \chi_r(x)| \leq C r^{-1}$ for some C and all x and r. Set $f_r(x) = f(x) (1 - \chi_r(x)) + f(p) \chi_r(x)$, so that each f_r is a C^1 function which is constant on B(p, r/2). From (5.2) we have that

(5.10)
$$|f(p)| = |f_r(p)|$$

 $\leq C \int_{\mathbf{R}^n \setminus B(p,r/2)} \frac{1}{\delta(x,p)^{n-1}} \,\omega(x)^{-\frac{1}{n}} |\nabla f_r(x)| \,\omega(x) \, dx.$

This implies that

(5.11)
$$|f(p)| \leq C \int_{\mathbf{R}^n} \frac{1}{\delta(x,p)^{n-1}} \,\omega(x)^{-\frac{1}{n}} |\nabla f(x)| \,\omega(x) \, dx$$

 $+ C \int_{\mathbf{R}^n \setminus B(p,r/2)} \frac{1}{\delta(x,p)^{n-1}} \,\omega(x)^{-\frac{1}{n}} |\nabla (f-f_r)(x)| \,\omega(x) \, dx.$

We want to check that the last term goes to 0 as $r \to 0$. We have that $f - f_r = (f - f(p))\chi_r$, and this is a function which vanishes outside B(p,r) and whose gradient is uniformly bounded, independently of r (since f is C^1). Thus our error term is bounded by

(5.12)
$$C \int_{B(p,r)\setminus B(p,r/2)} \frac{1}{\delta(x,p)^{n-1}} \,\omega(x)^{-\frac{1}{n}} \,\omega(x) \, dx$$

For $x \in B(p,r) \setminus B(p,r/2)$ we have that $\delta(x,p)$ is comparable to $(\int_{B(p,r)} \omega(y) \, dy)^{\frac{1}{n}}$, because of the doubling condition for μ . From Hölder's inequality we have that

(5.13)
$$\int_{B(p,r)} \omega(x)^{\frac{n-1}{n}} dx \leq C \left(\int_{B(p,r)} \omega(x) dx \right)^{\frac{n-1}{n}} \left(\int_{B(p,r)} dx \right)^{\frac{1}{n}} \leq C r \left(\int_{B(p,r)} \omega(x) dx \right)^{\frac{n-1}{n}}.$$

Thus we conclude that (5.12) is O(r), and so the same is true of the last term in (5.11). Therefore this error term goes to 0 in the limit as $r \to 0$, and (5.2) follows from (5.11).

This completes the proof of Theorem 5.1.

Once one has (5.2) one can derive Sobolev inequalities with respect to ω in practically the same manner as for Euclidean spaces. See [**DS**] for more information.

Incidentally, although we have chosen to obtain information from the special mappings in Section 3 through "calculus" in this section, there is a more geometric alternative. One can use the mappings to produce a nice family of curves from the given point p and extending to infinity, roughly like the sphere of rays from p in the Euclidean case. These curves can be found in the sets $\theta^{-1}(\sigma)$, $\sigma \in \mathbf{S}^{n-1}$, with estimates coming from the coarea theorem. See [**DS**] and [**Se4**] for more information about this approach.

6. Variations on the theme.

Let us rephrase the story of this paper more conceptually in terms of currents, which are essentially differential forms with distributional coefficients. This will make it easier to describe afterwards some variations of the methods and results.

Let $\pi_p : \mathbf{R}^n \setminus \{p\} \to \mathbf{S}^{n-1}$ be as in (1.1), and let λ be the standard volume form on \mathbf{S}^{n-1} , normalized to have total volume = 1. Let α be the smooth (n-1)-form on $\mathbf{R}^n \setminus \{p\}$ which is the pull back of λ under π_p . This pull-back has a singularity at p, and satisfies

(6.1)
$$|\alpha(x)| \le C |x-p|^{-n+1}$$

on $\mathbf{R}^n \setminus \{p\}$. Since α is the pull-back of a closed form (any (n-1)-form on \mathbf{S}^{n-1} is closed), α is also closed on $\mathbf{R}^n \setminus \{p\}$, i.e., $d\alpha = 0$ there. However, (6.1) says that the coefficients of α are locally integrable across the singularity, and so we can extend α across the singularity as a current. When we do this we cannot say immediately that this extension is closed even across the singularity, and indeed we have

$$(6.2) d\alpha = [p],$$

where [p] denotes the current of integration over the zero-dimensional manifold $\{p\}$. This equation means that

(6.3)
$$\int_{\mathbf{R}^n} df \wedge \alpha = f(p)$$

for any smooth function f on \mathbb{R}^n with compact support. The reader should not worry too much about whether the signs are correct in (6.2) and (6.3), for the author is concerned not at all.

These facts are not hard to check. Indeed one can even write down an explicit formula for α and compute everything directly. One can also verify these assertions more abstractly, using general properties of the construction instead of explicit formulae. This is in effect what we did in

the previous section. There we were working with the mapping θ instead of π_p , and although θ is not given so explicitly as π_p , it does enjoy many of the same properties. For θ we have the minor inconvenience of its lack of smoothness, but this can be managed with little difficulty, by smoothing it for instance, as we did before.

We never quite said before that what we were doing was deriving versions of (6.2) and (6.3) for θ instead of π_p , but that is really what was happening.

For conceptual purposes this language is clearer, although in practical use it is sometimes simpler to work a little differently, as in Section 5.

Let us remain in this language for the moment and consider some variations on the theme.

Instead of looking for a formula for f(p) in terms of the gradient of f we can take two points p and q and look for a formula for f(p) - f(q). Of course one can simply take the difference of two formulas like (6.3), but it is nicer to have a formula that does not require f to have compact support, indeed a formula which is local and works for any smooth function defined on a ball that contains p and q.

How can we get such a formula, along the lines of the construction just described? Instead of building a mapping π_p as above we should look for a mapping from $\mathbf{R}^n \setminus \{p, q\}$ into \mathbf{S}^{n-1} which is constant outside of a ball that contains p and q, which behaves like π_p near p, and which behaves in a manner like that of π_q near q, but with an extra reflection included in order to reverse the local degree at q. It is not difficult to construct such mappings, although they are not so easily written as π_p .

Once one has such a mapping one can pull back the volume form λ on \mathbf{S}^{n-1} to get a smooth (n-1)-form β on $\mathbf{R}^n \setminus \{p,q\}$ which has singularities at each of p and q like α has at p (6.1). This form β is also closed, and if we extend it as a current across the singularities we get that

(6.4)
$$d\beta = [p] - [q] \quad \text{and} \quad$$

(6.5)
$$\int_{\mathbf{R}^n} df \wedge \beta = f(p) - f(q),$$

modulo sign errors. Because our mapping is chosen to be constant outside a certain ball, β vanishes off of the same ball, and for (6.5) we do not need f to be defined on all of \mathbf{R}^n , etc.

Again this construction does not really depend on specific choices or explicit computations, just some basic properties of the mapping in \mathbf{S}^{n-1} that we build. Instead of a smooth mapping adapted to Euclidean geometry we can build mappings adapted to the geometry of a metric doubling measure on \mathbb{R}^n . That is, we approximate the mapping for Euclidean geometry by one which is locally Lipschitz with respect to the new metric, with suitable estimates near the singularities. We can make this approximation in the same manner as in Section 3, and with the same estimates at the singularities. This approximation will have the same topological properties as the one for Euclidean geometry, because we can build it in such a way as to be a small perturbation of the mapping for Euclidean geometry in the supremum norm, so as to be homotopic to the Euclidean mapping through mappings into \mathbf{S}^{n-1} . Thus we can maintain the local degrees near the singularities, and this nondegeneracy gives rise to formulas like (6.5), estimates like (5.2), etc., in the same manner as in Section 5.

This story of the two singular points is closer to the development given in [**DS**] and [**Se4**]. For applications to analysis the two-point formulas are generally more useful than the one-point formulas, and more flexible. For instance one can derive Poincaré-type inequalities by averaging the two-point formulas, and the locality of the two-point formulas is very convenient. The one-point estimates can be derived from the two-point estimates by sending the other point to infinity, but the direct constructions are cleaner.

This construction for the two-point formulas gives practically the same result as the construction in $[\mathbf{DS}]$. The constructions themselves are not quite the same though, and the approach here isolates better the topological properties needed and the principles involved. This makes it easier to make similar constructions in other contexts. For instance the constructions in $[\mathbf{Se4}]$ are based on similar principles, but this similarity is less obvious if one tries to compare $[\mathbf{Se4}]$ and $[\mathbf{DS}]$ directly. In the setting of $[\mathbf{Se4}]$, the geometry was not given so explicitly, it was not such a simple perturbation of Euclidean geometry, and the mappings could not be obtained as simple perturbations. It was also better to work there with one-parameter families of nontrivial mappings to \mathbf{S}^n which were constant off of balls, rather than individual radial projections into \mathbf{S}^{n-1} , but this is more of a variation on the theme rather than a fundamentally different principle. (One can also see it in terms of stabilizing the space by taking the Cartesian product with the line or an interval.)

As another variation on the theme it is nicer to keep track of the dependence on the point p, in the following sense. Instead of fixing p and defining the mapping $\pi_p : \mathbf{R}^n \setminus \{p\} \to \mathbf{S}^{n-1}$ as before, let us define a

mapping $\pi : \mathbf{R}^n \times \mathbf{R}^n \setminus \Delta \to \mathbf{S}^{n-1}, \Delta = \text{the diagonal} = \{(x, x) : x \in \mathbf{R}^n\},\$ by

(6.6)
$$\pi(x,y) = \frac{x-y}{|x-y|}$$

Again we can pull back the volume form λ to a smooth (n-1)-form on $\mathbf{R}^n \times \mathbf{R}^n \backslash \Delta$. This form is closed, and it has a singularity at the diagonal like that of α near p (6.1). We can extend it to a locally integrable current across the singularity, and when we do that we get a current γ such that

$$(6.7) d\gamma = [\Delta$$

(modulo a possible sign error), where $[\Delta]$ denotes the current of integration over the diagonal Δ . Keep in mind that the exterior derivative dhere involves now derivatives in both sets of variables. This current γ provides the "kernel" for a familiar "homotopy operator" for d; that is an operator T acting on differential forms which satisfies dT + T d = I. (I learned this view of homotopy operators from Reese Harvey and John Polking [**HP**].) Remember that one cannot solve the equation $d\phi = \psi$ in general, one must require the compatibility condition that $d\psi = 0$, and homotopy operators provide solutions to the d-equation when the compatibility condition is fulfilled.

Again this construction does not really depend on the precise formula for π , only on properties that are more stable than that. Given a metric on \mathbf{R}^n which is quasisymmetrically equivalent to the Euclidean metric we can build an analogous mapping which approximates π in the supremum norm, hence is homotopic to π through mappings from $\mathbf{R}^n \times \mathbf{R}^n \setminus \Delta$ into \mathbf{S}^{n-1} and so has approximately the same topological properties as π , and which has approximately the same analytical properties as π with respect to the new metric. In particular one can do this for metrics associated to metric doubling measures. One can then build a homotopy operator for d whose kernel satisfies the right bounds associated to the given metric doubling measure, analogous to (5.2). Once one has a kernel with the right bounds, one can get bounds on the corresponding operator.

In summary, in analysis we are often looking for operators which provide some kind of inverse to a differential operator, like the exterior derivative d. On \mathbf{R}^n we can sometimes see these operators as being constructed from differential forms which satisfy certain properties, and these differential forms can sometimes be obtained from special mappings. In this way problems of calculus become problems of geometry. The problems of geometry are frequently stable under perturbations, their topological aspects being stable in particular, so that if we change the geometry we might be able to modify the mapping, in such a way as to be able to do everything for the new geometry. And in such a way as to solve problems of calculus for the new geometry which are much more mysterious when seen in purely analytical terms.

References

- [DS] G. DAVID AND S. SEMMES, Strong A_{∞} -weights, Sobolev inequalities, and quasiconformal mappings, in "Analysis and Partial Differential Equations," (C. Sadosky, ed.), Lecture Notes in Pure and Applied Mathematics **122**, Marcel Dekker, 1990, pp. 101–111.
- [Ga] J. GARNETT, "Bounded Analytic Functions," Academic Press, 1981.
- [Ge] F. W. GEHRING, The L^p integrability of the partial derivatives of a quasiconformal mapping, *Acta Math.* **130** (1973), 265–277.
- [HP] R. HARVEY AND J. POLKING, Fundamental solutions in complex analysis Part I. The Cauchy Riemann operator, *Duke Math. J.* 46 (1979), 253–300.
- [J] J. L. JOURNÉ, "Calderón-Zygmund Operators, Pseudodifferential Operators, and the Cauchy Integral of Calderón," Lecture Notes in Math. 994, Springer-Verlag, 1983.
- [MS] R. MACIAS AND C. SEGOVIA, Lipschitz functions on spaces of homogeneous type, Adv. Math. 33 (1979), 257–270.
- [N] L. NIRENBERG, Topics in Nonlinear Functional Analysis, Courant Institute of Mathematical Sciences (1974).
- [Se1] S. SEMMES, Bilipschitz mappings and strong A_{∞} weights, Ann. Acad. Sci. Fenn. Ser. A I 18 (1993), 211–248.
- [Se2] S. SEMMES, On the nonexistence of bilipschitz parameterizations and geometric problems about A_{∞} weights, *Rev. Mat. Iberoamericana* (1996).
- [Se3] S. SEMMES, Good metric spaces without good parameterizations, *Rev. Mat. Iberoamericana* 12 (1996), 187–275.
- [Se4] S. SEMMES, Finding Curves on General Spaces through Quantitative Topology with Applications for Sobolev and Poincaré Inequalities, *Selecta Math.* (1996), (to appear).

[St] E. M. STEIN, "Singular Integrals and Differentiability Properties of Functions," Princeton University Press, 1970.

> Department of Mathematics Rice University Houston, TX 77251-1892 U.S.A.

Primera versió rebuda el 21 de Desembre de 1995, darrera versió rebuda el 10 de Juny de 1996