A QUANTITATIVE METRIC DIFFERENTIATION THEOREM

JONAS AZZAM AND RAANAN SCHUL

(Communicated by Tatiana Toro)

Abstract. The purpose of this note is to point out a simple consequence of some earlier work of the authors, “Hard Sard: Quantitative implicit function and extension theorems for Lipschitz maps”. For $f$, a Lipschitz function from a Euclidean space into a metric space, we give quantitative estimates for how often the pullback of the metric under $f$ is approximately a seminorm. This is a quantitative version of Kirchheim’s metric differentiation result from 1994. Our result is in the form of a Carleson packing condition.

1. Introduction

Consider a function from Euclidean space into a metric space, $f : \mathbb{R}^n \to \mathcal{M}$. Without loss of generality, we will assume $\mathcal{M} = \ell^\infty$, which will ease some notation below. Let $\Delta(\mathbb{R}^n)$ be a collection of dyadic cubes in $\mathbb{R}^n$ and $\Delta(R) = \{Q \in \Delta(\mathbb{R}^n), Q \subset R\}$. For $Q \in \Delta(\mathbb{R}^n)$, side$(Q)$ denotes its sidelength, $x_Q$ its center, and $3Q$ the cube with the same center but 3 times the sidelength. Define

$$md(Q) := \frac{1}{\text{side}(Q)} \inf \sup_{x, y \in Q} \left| |f(x) - f(y)| - \|x - y\| \right|,$$

where the infimum is taken over all seminorms $\|\cdot\|$ on $\mathbb{R}^n$ and $|\cdot|$ is the $\ell^\infty$ norm. A function $f$ as above is said to be $L$-Lipschitz if for all $x, y \in \mathbb{R}^n$,

$$\text{dist}(f(x), f(y)) \leq L|x - y|.$$

Our main theorem in this note is that the pullback of the distance function on $\mathcal{M}$ under an $L$-Lipschitz function $f : \mathbb{R}^n \to \mathcal{M}$ is well approximated by a seminorm on most scales and in most locations.

Theorem 1.1. Let $f : \mathbb{R}^n \to \mathcal{M}$ be an $L$-Lipschitz function. Let $\delta > 0$. Then for each $R \in \Delta(\mathbb{R}^n)$,

$$\sum \{\text{vol}(Q) : Q \in \Delta(R), md(3Q) > \delta L\} \leq C_{\delta, n} \cdot \text{vol}(R).$$

The constant $C_{\delta, n}$ does not depend on the metric space $\mathcal{M}$ or the function $f : \mathbb{R}^n \to \mathcal{M}$.

Received by the editors January 10, 2012 and, in revised form, May 22, 2012.

2010 Mathematics Subject Classification. Primary 26A16, 54E40.

Key words and phrases. Lipschitz map, metric differential, differentiability of Lipschitz maps, quantitative differentiation, Rademacher.

The first author was supported by RTG grant DMS-0838212.

The second author was supported by a fellowship from the Alfred P. Sloan Foundation and by NSF grant DMS 1100008.
An estimate as above is typically called a Carleson packing condition in harmonic analysis or geometric measure theory (cf. [DS93]). The name stems from the fact that a Carleson packing on a collection of cubes usually arises from the stronger property of the existence of a Carleson measure on $[0,1]^n \times (0,\infty)$ (see [Gar07] or [Ste93]). For example, Dorronsoro established in [Dor85] that if $f$ is Lipschitz and
\[
\Omega_f(x,t) := \inf \left\{ \int_{B(x,t)} \frac{|f(y) - A(y)|}{t} dy : A : [0,1]^n \to \mathbb{R}^m \text{ is affine} \right\},
\]
then $\Omega_f(x,t)^2 \frac{dx}{t}$ defines a Carleson measure. This implies Theorem 1.1 in the case that $f$ maps into a Euclidean space and both its domain and range are equipped with the Euclidean norm. In fact, more is true: the theorem still holds if we replace $\text{md}(Q)$ with $\alpha(Q)$, where $\alpha(Q) := \inf \|f - A\|_{L^\infty}(Q)$ and the infimum is taken over all affine mappings $A$ between the domain and range of $f$. In the argot of [DS93], Theorem 1.1 is the weak approximation of Lipschitz functions by affine maps, or the WALA property, which holds more generally for Lipschitz functions from uniformly rectifiable sets into a Euclidean space. For more information about Carleson packing conditions, the WALA property, and their applications to geometric measure theory, see [Jon90,DS91,DS93,DS00,AS12].

Assaf Naor and Sean Li have recently developed estimates similar to that of Theorem 1.1 to estimate the size of the largest cube for which a Lipschitz map of a function into a super-reflexive space is approximately affine [LN], thus strengthening an earlier infinitesimal result of Bates, Johnson, Lindenstrauss, Preiss, and Schechtman [BJL99]. Other authors who have been concerned with quantitative or coarse differentiation results include Jeff Cheeger, Alex Eskin, David Fisher, Irene Peng, and others (cf. [EFW07] and the references therein).

There are many other references of this type which we omit here. A common thread to these is that one may view such statements as coarse or quantitative versions of Rademacher’s theorem, which can be used to find a large scale where the function is approximately linear, or sufficiently regular in the appropriate sense. A Rademacher-type theorem for Lipschitz functions from Euclidean space into a general metric space was established by Kirchheim in terms of metric differentials.

**Theorem 1.2** ([Kir94]). If $f : [0,1]^n \to M$ is Lipschitz, where $M$ is a metric space, then for almost every $z \in [0,1]^n$, there is a seminorm $\|\cdot\|_z$ on $\mathbb{R}^n$ such that
\[
\left| |f(x) - f(y)| - \|x - y\|_z \right| = o(|x - z| + |y - z|).
\]

The seminorm $\|\cdot\|_z$ is called the metric differential of $f$ at $z$. In other words, for almost every point in the domain, the pullback of the metric under $f$ is well approximated by some seminorm on the domain. Theorem 1.1 now gives a more quantitative version of this result in the spirit of [Dor85]. The work below is a simple consequence of some techniques and ideas that were originally developed in and for [AS12].

In Section 2, we recall the definition of $\tilde{\beta}$-numbers from [Sch09] and the relevant lemmas from [AS12]. In Section 3 we prove the main theorem.
2. Preliminaries

For \( Q \in \Delta(\mathbb{R}^n) \), let \( Q^N \) denote the unique dyadic cube containing \( Q \) of sidelength \( 2^N \text{side}(Q) \). For functions \( A(t_1, ..., t_k) \) and \( B(t_1, ..., t_k) \), we write \( A \lesssim B \) if there is a constant \( C \) (independent of \( (t_1, ..., t_k) \)) such that \( A \leq CB \). We will also write \( A \lesssim_{t_i} B \) if the implied constant \( C \) depends on \( t_i \).

2.1. \( \beta \)-numbers. For a Lipschitz function \( f : [0, 1]^n \to \mathcal{M} \), define
\[
\partial^f(x, y, z) := |f(x) - f(y)| + |f(y) - f(z)| - |f(x) - f(z)|.
\]
For an interval \( I = [a, b] \subset \mathbb{R} \), let
\[
\tilde{\beta}_f(I)^2 \text{diam}(I) = \text{diam}(I)^{-3} \int_{x=a}^{x=b} \int_{y=x}^{y=b} \int_{z=y}^{z=b} \partial^f(x, y, z) \, dz \, dy \, dx.
\]
Identify \( \mathbb{R} \) with \( \{\mathbb{R}, 0, ..., 0\} \subset \mathbb{R}^n \), and let \( G_n \) be the group of all rotations of \( \mathbb{R} \) in \( \mathbb{R}^n \) equipped with its Haar measure \( d\mu \). Let \( dx \) be the \( n \) – dimensional Lebesgue measure on \( \mathbb{R}^n \cap g \mathbb{R} \), the orthogonal complement of \( g \mathbb{R} \) in \( \mathbb{R}^n \). For a cube \( Q \in \mathbb{R}^n \), define the quantity \( \tilde{\beta}_f^{(n)}(Q) \) by
\[
\tilde{\beta}_f^{(n)}(Q)^2 \text{side}(Q)^n = \int_{g \in G_n} \int_{x \in \mathbb{R}^n \cap g \mathbb{R}} \chi_{\{(x + g \mathbb{R}) \cap 7Q \geq \text{side}(Q)\}}(x + g \mathbb{R}) \, dx \, d\mu(g).
\]

**Theorem 2.1** ([Sch09]). For an \( L \)-Lipschitz function \( f : [0, 1]^n \to \mathcal{M} \) and \( N \) a fixed integer,
\[
\sum_{Q \in \Delta, Q \subset [0, 1]^n} \tilde{\beta}_f(3Q^N)^2 \text{vol}(Q) \lesssim_{N, n} L.
\]

**Remark 2.2.** If \( \partial^f(x, y, z) \) is small, this corresponds to \( f(x), f(y) \) and \( f(z) \) being close to lying on a geodesic. A crucial (and simple) observation in [AS12] was that if we define \( f(x) = (x, f(x)) \) as a map from \( Q_0 \) into \( \mathbb{R}^n \times \mathbb{R}^m \), then \( \partial^f(x, y, z) \) being small corresponds to \( f \) being approximately affine on \( x, y, z \). Thus, for \( \tilde{\beta}_f(Q) \) small enough, this gives that \( f \) is approximately affine when restricted to \( Q \). However, in the setting where \( f \) maps into a metric space, if we interpret \( \tilde{f} = (x, f(x)) \) as a map from \( \mathbb{R}^n \) into \( \mathbb{R}^n \oplus \ell^\infty \), equipped with the metric \( |(u, v)| = \sqrt{|u|^2 + |v|^2} \), then \( \tilde{\beta}_f(Q) \) being small corresponds to \( |f_Q \) being approximately homogeneous on all lines. We will use this to prove Theorem 1.1. One should compare the proof with the usual proof of Rademacher’s theorem or with the proof of Theorem 2 in [Kir94].

2.2. Approximate homogeneity. Let \( f : [0, 1]^n \to \mathcal{M} \) be 1-Lipschitz. Fix \( \alpha \in (0, 1) \) and \( N \in \mathbb{N} \) for now. The value for the constant \( N \) is determined in Lemma 2.4 with some dependencies, which are ultimately resolved in Proposition 3.2 where \( \alpha \) is also determined. For \( x, y \in 3Q \) with \( |x - y| \geq \text{side}(Q) \), define
\[
\sigma(x, y) = \inf_{x', y' \in L_{x, y} \cap 3Q^N, \text{mod}(x') \geq \text{side}(Q)} \frac{|f(x') - f(y')|}{|x' - y'|},
\]
where \( L_{x, y} \) is the line passing through \( x \) and \( y \).

We recall some lemmas whose proofs can be found in [AS12]. Their proofs can be read independently of the rest of this paper.
Lemma 2.3 ([AS12, Lemma 6.3 and the following discussion]). Let $0 < \alpha < 1$, $\epsilon' > 0$ and $N \in \mathbb{N}$. There is $\epsilon = \epsilon(N, \alpha, \epsilon') > 0$ such that for any cube $Q$ with $\tilde{\beta}_f(3Q^N) < \epsilon$ we have
\begin{equation}
|\frac{f(x) - f(y)}{|x - y|} - \sigma(x, y)| < \epsilon'
\end{equation}
for all $x, y \in Q$ such that $|x - y| > \text{diam}Q$.

In other words, if $\tilde{\beta}_f(3Q^N)$ is small, then $f$ is close to being homogeneous on lines. We remark that the constant $N$ here is only needed for consistency with the next lemma, Lemma 2.4.

Lemma 2.4 ([AS12, Corollary 6.5]). For any $\rho > 0$, there is $N = N(\rho, n)$ and $\epsilon = \epsilon(\alpha, \rho)$ such that if $\tilde{\beta}(3Q^N) < \epsilon$, and if $x, y \in Q$ are such that $|x - y| \geq \text{side}(Q)$, and $z \in \mathbb{R}^n$ is such that $|z| \leq \text{side}(Q)$, then
\begin{equation}
\sigma(x + z, y + z) \leq \sigma(x, y) + \rho
\end{equation}
and
\begin{equation}
|f(x + z) - f(y + z)| - |f(x) - f(y)| < \rho \text{side}(Q).
\end{equation}
Moreover, if $\rho$ is small enough, depending on $\alpha$, and $x_Q = f(x_Q) = 0$, then for every $x, y \in Q$,
\begin{equation}
|f(x + y)| \leq |f(x)| + |f(y)| + \alpha \text{side}(Q).
\end{equation}

Remark 2.5. It is easy to show that the function $\sigma(x, y)$ is continuous on \{$(x, y) \in Q \times Q : |x - y| \geq \text{side}Q$\}.

3. Proof of Theorem 1.1

We first recall an easy consequence of Carathéodory’s convex hull theorem [Mat02, Theorem 1.2.3, p. 6].

Lemma 3.1. Let $K$ be a set in $\mathbb{R}^n$ and $\overline{\text{co}}K$ its closed convex hull. Then $\overline{\text{co}}K = \overline{K}$, where
\[
\overline{K} := \left\{ \sum_{j=0}^{n} a_j x_j : x_j \in K, 0 \leq a_j \leq 1, \sum_{j=0}^{n} a_j = 1 \right\}.
\]
That is, $\overline{\text{co}}K$ is the closure of all convex combinations of at most $n + 1$ points from $K$.

We are now ready to state and prove the main tool.

Proposition 3.2. Let $\delta > 0$ and $f : [0, 1]^d \rightarrow \mathcal{M}$ be $1$-Lipschitz. Then there is $\epsilon > 0$ and $N \in \mathbb{N}$ such that if $\tilde{\beta}_f(3Q^N) < \epsilon$, then $\text{md}(Q) < \delta$.

Proof. When rescaling the domain and range by the same factor, the Lipschitz constant of $f$, $\text{md}$, and $\tilde{\beta}_f$ remain unchanged. Thus we may assume that $Q$ is centered at zero with $\text{side}(Q) = 1$ and $f(0) = 0$. Fix $\alpha > 0$, which will be specified later, and pick $\epsilon > 0$ and $N$ so that the conclusions of Lemmas 2.3 and 2.4 hold for $\epsilon' = \rho = \alpha$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Let $\sigma(x) = \sigma(0, x)$. Set
\[ C = \{ x \in Q : \sigma(x)|x| \leq 1 \} \]
and
\[ D = \{ x \in Q : \sigma(x)|x| = 1 \text{ or } \sigma(x) = 0 \}. \]

For $t > 0$ and $A \subseteq \mathbb{R}^n$, define $tA = \{ tx : x \in A \}$.

**Claim.** For $\alpha' = 2(n + 1)\alpha$,
\[ (3.1) \quad (1 - \alpha')\overline{co}D \subseteq C \subseteq \overline{co}D. \]

The right containment is clear by the definition of $D$ and $\sigma(x)$ being constant on any given 1 dimensional vector space (in particular, $C$ is a star-shaped subset of $\overline{co}D$). To see the left containment, we proceed as follows.

Note that $C$ is closed because $\sigma$ is continuous (see Remark 2.5). Thus by Lemma 3.1, it suffices to show that $(1 - \alpha')\overline{co}D \subseteq C$. In this vein, let $x \in (1 - \alpha')\overline{co}D$, so that
\[ x = \sum_{j=0}^{n} a_j x_j \in \overline{D}, \text{ where } x_j \in (1 - \alpha')D, \ a_j \geq 0, \text{ and } \sum_{j=0}^{n} a_j = 1. \]

Let $0 \leq k \leq n$ be the maximal integer so that (after reordering the $x_j$’s)
\[ j \leq k \implies |x_j| \leq \alpha. \]

Then by (2.3) and the definition of $\sigma$,
\[
\begin{align*}
|x|\sigma(x) & \leq |f(x) - f(0)| = \left| f \left( \sum_{j=0}^{n} a_j x_j \right) \right| \\
& \leq (n + 1)\alpha + \sum_{j=k+1}^{n} |f(a_j x_j)| \\
& \leq (n + 1)\alpha + (k + 1)\alpha + \sum_{j=k+1}^{n} |f(a_j x_j)| \\
& \leq 2(n + 1)\alpha + \sum_{j=k+1}^{n} \sigma(a_j x_j)|a_j x_j| \\
& \leq 2(n + 1)\alpha + \sum_{j=k+1}^{n} a_j \sigma(x_j)|x_j| \\
& \leq 2(n + 1)\alpha + \sum_{j=0}^{n} a_j (1 - \alpha') \leq 1,
\end{align*}
\]

where in the last two inequalities we used the facts that $\sigma \leq 1$ since $f$ is 1-Lipschitz, $\sum_{j=0}^{n} a_j = 1$, and $\alpha' = 2(n + 1)\alpha$. This proves the claim.

Let
\[ \|x\| = \inf \{ t \geq 0 : x \in t\overline{co}D \}. \]

By Kolmogorov’s theorem [Kol34] (see also [Bou87, Section II, p. 20, Prop. 23]), since $\overline{co}D$ is convex and $-\overline{co}D = \overline{co}D$ (since $-D = D$), we have that $\| \cdot \|$ is a seminorm. By (3.1),
\[ (3.3) \quad \|x\| \leq \sigma(x)|x| \leq \frac{1}{1 - \alpha'}\|x\|. \]

Let $x, y \in Q$. Since $f$ is 1-Lipschitz, we have $\sigma \leq 1$. By (3.3)
\[ (3.4) \quad \|x - y\| \leq \sigma(x - y)|x - y| \leq |x - y|. \]
Suppose $|x - y| > \alpha$; then
\[
|f(x) - f(y)| \leq |f(x - y)| + \alpha \leq \sigma(x - y)|x - y| + 2\alpha \leq \frac{1}{1 - \alpha^2}\|x - y\| + 2\alpha,
\]

and by Lemma 2.4,
\[
(3.5) \|x - y\| \leq \sigma(x - y)|x - y| \leq (\sigma(x, y) + \rho)|x - y| \leq |f(x) - f(y)| + \alpha\sqrt{n}
\]
(recall that side$Q = 1$ so that diam$Q = \sqrt{n}$ and earlier we picked $\rho = \alpha$).

If $|x - y| \leq \alpha$, then, by (3.4), $\|x - y\| \leq \alpha$. Furthermore, since $f$ is 1-Lipschitz, $|f(x) - f(y)| \leq |x - y| \leq \alpha$. Thus, the difference between $\|x - y\|$ and $|f(x) - f(y)|$ is at most $\alpha$.

Combining the above estimates and using the fact that $\|x - y\| \leq |x - y| \leq$ diam$Q = \sqrt{n}$, we have that for all $x, y \in Q$,
\[
\|x - y\| - \alpha \leq |f(x) - f(y)| \leq \frac{\|x - y\|}{1 - \alpha^2} + (2 + \sqrt{n})\alpha \leq \|x - y\| + \frac{\sqrt{n}\alpha'}{1 - \alpha^2} + (2 + \sqrt{n})\alpha.
\]

By choosing $\alpha$ so that $\frac{\sqrt{n}\alpha'}{1 - \alpha^2} + 2\alpha + \sqrt{n}\alpha < \delta$ (recall that $\alpha' = 2(n + 1)\alpha$), we conclude the proof.

**Proof of Theorem 1.1** Let $\delta > 0$ be given and $R \in \Delta$. By rescaling and translating the domain, we may assume $R = [0, 1]^n$. Note that $\frac{md}{L}$ is invariant under such a transformation. By rescaling the metric, we may also assume that the Lipschitz constant of $f$ is 1 and, again, that this will not affect $\frac{md}{L}$. Then if $\epsilon = \epsilon(\delta)$ is as in the previous proposition, we have by Theorem 2.1

\[
(3.6) \sum \{\text{vol}(Q) : Q \in \Delta([0, 1]^n), \text{md}(Q) > \delta\}
\leq \sum \{\text{vol}(Q) : Q \in \Delta([0, 1]^n), \tilde{f}(3Q^N) > \epsilon\}
\leq \frac{1}{\epsilon^2} \sum_{Q \in \Delta([0, 1]^n)} \tilde{f}(3Q^N)^2 \text{vol}(Q) \lesssim_{N, n} \frac{1}{\epsilon^2} \lesssim_\delta 1.
\]

By the standard $\frac{1}{3}$-trick (cf. [Oki92, pp. 339-340]), any cube $3Q$ of side$Q \leq \frac{1}{6}$ is contained in a cube of the form $R + v$, where $v \in \{0, \pm \frac{1}{3}1^n \}$, $R \in \Delta([0, 1]^n)$, and side$R \lesssim_n$ side$Q$. Note that $\text{md}(3Q) \lesssim_n \text{md}(R + v)$ and vol$(Q) \sim_n$ vol$(R)$. Hence, we can apply (3.6) to $f$ with respect to each of these translated grids to obtain
\[
\sum \{\text{vol}(3Q) : Q \in \Delta([0, 1]^n), \text{md}(Q) > \delta\}
\lesssim_n 1 + \sum_{v \in \{0, \pm \frac{1}{3}1^n \}} \sum \{\text{vol}(R) : R \in \Delta([0, 1]^n), \text{md}(R + v) > \delta\}
\lesssim_{n, \delta} 1 + \sum_{v \in \{0, \pm \frac{1}{3}1^n \}} 1 \lesssim_n 1.
\]
ACKNOWLEDGEMENTS

The authors thank John Garnett for his helpful discussions, Sean Li for pointing out [LN], and Matthew Badger for carefully reading a draft. Part of this note was written while the first author was visiting the Mathematical Sciences Research Institute.

REFERENCES


Department of Mathematics, University of Washington, Box 354350, Seattle, Washington 98195-4350

Department of Mathematics, Stony Brook University, Stony Brook, New York 11794-3651