H. Movahedi-Lankarani, Department of Mathematics, Penn State - Altoona Campus, Altoona, PA 16601-3760

ON THE THEOREM OF RADEMACHER

ABSTRACT. It is shown that there exist Cantor subsets X of $\mathbb{R}^N$ and bi-Lipschitz maps $f : X \to f(X) \subset \mathcal{H}$, where $\mathcal{H}$ is an infinite dimensional Hilbert space, such that f is not strongly differentiable at any point of X. Furthermore, for each such X and f the image $M = f(X)$ has the property that for any $N \geq 1$ and any differentiable map $F : [0,1]^N \to \mathcal{H}$ with $dF(x)$ nonsingular for all $x \in [0,1]^N$, the set $F^{-1}(M)$ is a finite set. Hence, f can agree with a nonsingular differentiable map at most on a finite set.

I. Introduction. It follows directly from the classical Lebesgue Differentiability Theorem that if $f : \mathbb{R} \to \mathbb{R}$ is locally Lipschitz, then it is differentiable almost everywhere with respect to Lebesgue measure. A more general (and more difficult) result is the following theorem of Rademacher.

Theorem. *(Rademacher) Let $\mathcal{U}$ be a nonempty open subset of $\mathbb{R}^N$ and suppose that $f : \mathcal{U} \to \mathbb{R}^m$ is locally Lipschitz. Then f is (strongly) differentiable almost everywhere with respect to Lebesgue measure.*

This differentiability theorem holds also for functions with values in a Hilbert space. (See for instance [1].) Observe however that the domains of these functions are open in $\mathbb{R}^N$. This leads to the following natural question:

Suppose that X is a compact (e.g. of Lebesgue measure zero) subset of $\mathbb{R}^N$ and $f : X \to \mathbb{R}^m$ is locally Lipschitz. Does it follow that f has to be (strongly) differentiable at some point $x \in X$?

For the case of sets of measure zero in the line and $m = 1$ the problem is solved. For any set $x \subset \mathbb{R}$ of measure zero Zahorski [8] constructed a real-valued Lipschitz function that is not differentiable (with respect to X) at any point of X provided that X has no isolated points. In $\mathbb{R}^N$, $N \geq 2$, there is a set of measure zero such that every real-valued locally Lipschitz function has points of differentiability in it. (See [7].) Not much more is known; for example the question of replacing real-valued by plane valued has not been answered yet.

In this paper we construct a large family of Cantor subsets X of $\mathbb{R}^N$ and <u>bi-Lipschitz</u> maps $f : X \to f(X) \subset \mathcal{H}$, where $\mathcal{H}$ is an infinite dimensional Hilbert space, such that f is <u>not</u> strongly differentiable at <u>any</u> point of X. (By a strong derivative D of f at $x \in X$ we mean a linear map $D : \mathbb{R}^N \to \mathcal{H}$ with the property that for any $\epsilon > 0$ there is $\delta > 0$ such that

1991 *Mathematics Subject Classification.* 26B05 58C20 26A27.

$$\|f(y) - f(x) - D(y-x)\| < \epsilon\|y-x\|$$

provided that $x, y \in X$ and $\|y-x\| < \delta$.) Furthermore, the image $M = f(X) \subset \mathcal{H}$ has the bizarre property that for any differentiable map $F : [0,1]^N \to \mathcal{H}$ with $dF(x)$ nonsingular for all $x \in [0,1]^N$, the set $F^{-1}(M)$ is a finite set. Hence, such a map f can agree with a nonsingular differentiable map on at + most a finite set.

Our constructions are general and mainly geometric in spirit. Specific examples follow directly from our general set up. For instance, let X be the Cantor ternary subset of $[0,1]$ with the metric $\rho(x,y) = 1/3^n$, where x, y are in the same interval at the n^{th} level and are in different intervals the $(n+1)^{\text{th}}$ level. Clearly, for any $x, y \in X$ we have

$$|x-y| \le \rho(x,y) \le 3|x-y|.$$

Let μ be the ($\log\ 2/\log\ 3$)-dimensional Hausdorff measure on X. (Throughout this paper, log means natural logarithm.) Then the map $f : (X,\rho) \hookrightarrow L^2(X,\mu)$ defined by

$$f(x) = \rho(x,\cdot)^{1-(\log\ 2/2\log\ 3)}$$

is a bi-Lipschitz map that is not strongly + differentiable at any $x \in X$. But, f is weakly differentiable at every $x \in X$ with weak derivative equal to zero.

2. Preliminaries. In this section we provide some of the basic machinery needed for the construction of our class examples.

Definition 2.1. Let (M,d) be a metric space and for $0 < r_2 < r_1$ let $\mathcal{N}(r_1, r_2)$ denote the maximum number of disjoint closed r_2-balls contained in a closed r_1-ball. Let

$$\Delta_{r,t}(M) = \sup\left\{ \frac{\log \mathcal{N}(r_1,r_2)}{\log(r_1/r_2)} | 0 < r_2 < r_1 < +r \text{ and } r_1 > tr_2 \right\}.$$

Then the metric dimension of M is defined by

$$\dim_m(M) = \lim_{r\to 0}\ {}_{t\to\infty} + lim\Delta_{r,t}(M).$$

These limits exist since they are monotonic, but may be infinite.

The metric dimension satisfies all the usual properties that a reasonable dimension should satisfy. Specifically,

(i) If (M_1, d_1) is a metric subspace of (M_2, d_2), then $\dim_m(M_+1) \le \dim_m(M_2)$.

(ii) If M is an open subset of $\mathbb{R}^N$ equipped with the inherited + metric, then $\dim_m(M) = N$.

(iii) If (M_1, d_1) and (M_2, d_2) are bi-Lipschitz homeomorphic, + then $\dim_m(M_1) = \dim_m(M_2)$.

The metric dimension is distinct from the Hausdorff dimension (as well as any of its approximations like limit capacity, etc.). Indeed, there exists a compact metric space whose Hausdorff dimension is 1 but whose metric dimension is infinite, [6].

It should be noted that our metric dimension is closely related to the one introduced by P. Assouad [2] (denoted by $\mathrm{Dim}(M,d)$) which, in turn, is a generalization of the dimensional order of Bouligand [4]. It is easy to see that for a metric space (M,d) we have $\dim_m(M) \leq \mathrm{Dim}(M,d)$. However, the reverse inequality is not clear.

Recall that a metric space (M,d) is an <u>ultrametric space</u> if the metric satisfies a stronger form of the triangle inequality: for all $x,y,z \in M$

$$d(x,y) \leq \max\{d(x,z), d(y,z)\}.$$

One special property of these spaces is the fact that any two closed balls in an ultrametric space are either disjoint or one is contained in the other. Consequently, every ultrametric space is totally disconnected. Indeed, it follows from a straight forward argument that every compact, perfect, ultrametric space is homeomorphic to a Cantor space; that is, to a metric space homeomorphic to the Cantor set. However, there exist Cantor spaces, e.g. a fat Cantor subset of the unit interval, which are not bi-Lipschitz homeomorphic to an ultrametric space; [6].

Theorem 2.2. *Let (M,d) be a compact ultrametric space. If $\dim_m(M) < N$, then (M,d) is bi-Lipschitz embeddable in $\mathbb{R}^N$. Conversely, if (M,d) is bi-Lipschitz embeddable in $\mathbb{R}^N$, then $\dim_m(M) \leq N$.*

For a proof of this theorem see [6].

Next, let $\mathcal{H}$ be a Hilbert space and $M \subset \mathcal{H}$ be closed. We say that a function $f : M \to \mathbb{R}$ is differentiable if and only if there is a differentiable function $F : \mathcal{H} \to \mathbb{R}$ such that $f = F|M$, where the differentiable functions on a Hilbert space are defined in the usual way. This is probably the most natural way to define a differentiable structure on a closed subset of a Hilbert space. (See [5]).

Definition 2.3. Let M be a subset of a Hilbert space $\mathcal{H}$ and let $x \in M$.

(i) The set of strong tangent vectors to M at x is defined as +
$S_x M = \{\xi | \xi =$(strong)$\lim\limits_{n\to\infty} \frac{y_n - z_n}{+\|y_n - z_n\|}$, where $\{y_n\}_{n\geq 1}$ and $\{z_n\}_{n\geq 1}$ are sequences + in
M, $y_n \neq z_n$, converging in <u>norm</u> to $x\}$.

(ii) The set of weak tangent vectors to M at x is defined as +
$W_x M = \{\xi | \xi =$ (weak) $\lim\limits_{n\to\infty} \frac{y_n - z_{+n}}{\|y_n - z_n\|}$, where $\{y_n\}_{n\geq 1}$ and $\{z_n\}_{n\geq 1}$ are + sequences in M, $y_n \neq z_n$, converging in <u>norm</u> to $x\}$.

Lemma 2.4. *Let M be a compact subset of a Hilbert space $\mathcal{H}$. Then for any $x \in M$ we have that S_xM and W_xM are invariant under C^1 diffeomorphisms.*

The proof is straight forward and is left to the reader. Observe that for a subset of a Euclidean space, the set of strong and weak tangent vectors are identical. Hence, if M is a compact subset of a Hilbert space and $W_xM \neq S_xM$ for some $x \in M$, then M is not C^1 diffeomorphic to a subset of $\mathbb{R}^N$ for any N, i.e., M is not C^1 embeddable in finite dimensions.

3. The Example. Let $\{q_n\}_{n\geq 1} \subset \mathbb{Z}$ with $q_n > 1$ for all $n \geq 1$, fix $1/2 < r_1 < 1$, and let $r_n = (q_1 q_2 \cdots q_{n-1})^{-1}$, $n > 1$. Let $\mathbb{Z}q_n = \mathbb{Z}/q_n\mathbb{Z}$ and set $G = \prod_{n\geq 1} \mathbb{Z}_{q_n}$. For $x = \{x_i\}_{i\geq 1}$ and $y = \{y_i\}_{i\geq 1}$ in G let $d(x,y) = r_n$ if $x_i = y_i$ for $1 \leq i \leq n-1$ and $x_n \neq y_n$. Clearly (G,d) is a compact perfect ultrametric space and so a Cantor space. In order to compute the metric dimension of (G,d) we note that for any $m,n \in \mathbb{N}$ with $n < m$ we have $\mathcal{N}(r_n, r_m) = q_n q_{n+1} \cdots q_{m-1}$ and $r_n/r_m = q_n q_{n+1} \cdots q_{m-1}$. Hence, we see that $\dim_m(G) = 1$.

Next, let μ be the probability measure on G defined by $\mu = \mu_1 \times \mu_2 \times \ldots$ where μ_n is the normalized counting measure on $\mathbb{Z}q_n$. Observe that (G,d) is a commutative compact topological group with multiplication

$$\{x_i\}_{i\geq 1} \cdot \{y_i\}_{i\geq 1} = \{x_i + y_i (\mod q_i)\}_{i\geq 1}$$

and μ is its unique invariant Haar measure. For $x \in G$ and $n \geq 1$ let $B(x, r_n) = \{y \in G | d(x,y) \leq r_n\}$. Then $\mu(B(x,r_n)) = (q_1 q_2 \ldots q_{n-1})^{-1} = r_n$ for $n > 1$. It follows that for each $x \in G$ and $n > 1$ we have $\log \mu(B(x,r_n))/\log r_n = 1$. Therefore, the Hausdorff dimension of G is also 1. (See [3]).

Now, let $0 < \epsilon < 9/16$ be fixed and let ρ be the ultrametric on G defined by the sequence $\{\epsilon^n\}_{n\geq 1}$. Specifically, for $x, y \in (G,d)$, let $\rho(x,y) = \epsilon^n$ if x and y are in the same ϵ^{n-1}-ball (in metric d) but in distinct ϵ^n-balls. Then $\epsilon^n < d(x,y) \leq \epsilon^{n-1}$ and so

$$\epsilon d(x,y) \leq \rho(x,y) \leq d(x,y).$$

Therefore, (G,d) and (G,ρ) are bi-Lipschitz homeomorphic under the identity map. Furthermore, for each ϵ^n-ball $B(x,\epsilon^n) \subset (G,\rho)$ we have $\epsilon^{n+1} \leq \mu(B(x,\epsilon^n)) \leq \epsilon^n$, $n > 1$.

The next step is to look for a bi-Lipschitz embedding of (G,d) in some Hilbert space. To this end, let $\varphi : (G,d) \hookrightarrow L^2(G,\mu)$ by setting $\varphi(x) = \rho(x,\cdot)^{1/2}$.

Lemma 3.1. *The map φ is bi-Lipschitz.*

Proof. Let $x, y \in G$ with $\rho(x,y) = \epsilon^n$ for some $n \geq 1$. Denote $B_n(x) = \{z \in G | \rho(x,z) \leq \epsilon^n\}$, $n \geq 1$. Then

$$\|\varphi(x) - \varphi(y)\|_2^2 = \int_G |\rho(x,z)^{1/2} - \rho(y,z)^{1/2}|^2 d\mu(z) +$$

$$= \int_{B_n(x)} |\rho(x,z)^{1/2} - \rho(y,z)^{1/2}|^2 d\mu(z).$$

The last equality follows from the fact that the integrand is identically zero on $G\backslash B_n(x)$. Note that if $z \in B_{n+1}(x)$, then $\rho(y,z) = \epsilon^n$ and if $z \in B_{n+1}(y)$, then $\rho(x,z) = \epsilon^n$. Furthermore, for $z \in B_n(x) \setminus (B_{n+1}(x) \cup B_{n+1}(y))$ we have $\rho(x,z) = \rho(y,z) = \epsilon^n$. Hence,

$$\begin{aligned}\|\varphi(x) - \varphi(y)\|_2^2 &= \int_{B_{n+1}(x)} \left(\rho(x,z)^{1/2} - \epsilon^{+n/2}\right)^2 d\mu(z) \\ &\quad + \int_{B_{n+1}(y)} \left(\epsilon^{n/2} - \rho(y,z)^{1/2}\right)^2 d\mu(z) \\ &= 2\int_{B_{n+1}(x)} \left(\epsilon^{n/2} - \rho(x,z)^{1/2}\right)^2 d\mu(z).\end{aligned}$$

But if $z \in B_{n+1}(x)$, then

$$\epsilon^{n/2} - \epsilon^{(n+1)/2} \leq \epsilon^{n/2} - \rho(x,z)^{1/2} \leq \epsilon^{+}n/2.$$

Consequently, for $\epsilon < 9/16$ we have

$$\frac{\epsilon^2}{2} \cdot \epsilon^{2n} \leq \int_{B_{n+1}(x)} \left(\epsilon^{n/2} - +\rho(x,z)^{1/2}\right)^2 d\mu(z) \leq \epsilon^{2n+1}.$$

Therefore,

$$\epsilon^2 d(x,y) \leq \epsilon\rho(x,y) \leq \|\varphi(x) - \varphi(y)\|_2 \leq \sqrt{2\epsilon}\rho(x,y) \leq \sqrt{2\epsilon}\, d(x,y). \quad \blacksquare$$

Let $M = \varphi(G) \subset L^2(G,\mu)$. Since φ is bi-Lipschitz we se that $\dim_m(M) = 1$ and so by Theorem 2.2 M is bi-Lipschitz embeddable in $\mathbb{R}^N$ for any $N \geq 2$. We show, however, that M is not C^1 embeddable in $\mathbb{R}^N$ for <u>any</u> N. This fact is a consequence of the following lemma.

Lemma 3.2. *For any $x \in M$ we have $S_xM = \phi$ but $W_xM = \{0\}$.*

Proof. We first show that $S_xM = \phi$. To this end, assume that there are sequences $\{x_n\}_{n\geq 1}$ and $\{y_n\}_{n\geq 1}$ in G converging to $x \in G$ such that the sequence

$$\{(\varphi(x_n) - \varphi(y_n))/\|\varphi(x_n) - \varphi(y_n)\|_2\}_{n\geq 1}$$

converges strongly (in L^2 norm) to some $\xi \in S_xM$. Then $\|\xi\|_2 = 1$. Given $r > 0$ there is $n_0 \in \mathbb{N}$ such that if $n_0 < n$, then $x_n \in B(x,r)$ and $y_n \in B(x,r)$. Because G is an ultrametric space, if $n_0 < n$, then

$$\rho(x_n, y_n) \leq \max\{\rho(x_n,x), \rho(y_n,x)\} \leq r.$$

Therefore, $B(x_n, \rho(x_n, y_n)) = B(y_n, \rho(x_n, y_n)) \subset B(x, r)$ for $n_0 < n$.

On the other hand, suppose $z \in G$ with $\rho(z, x_n) > \rho(x_n, y_n)$ and $\rho(z, y_n) > \rho(x_n, y_n)$. Then we must have

$$\rho(z, y_n) \leq \max\{\rho(z, x_n), \rho(x_n, y_n)\} = \rho(z, x_n).$$

Similarly, we get $\rho(z, x_n) \leq \rho(z, y_n)$ and so $\rho(z, x_n) = \rho(z, y_n) +$. Hence, if $z \notin B(x_n, \rho(x_n, y_n)) = B(y_n, \rho(x_n, y_n))$, then $\varphi(x_n)(z) - \varphi(y_n)(z) = 0$ for $n_0 < n$. In other words, if $n_0 < n$, then

$$\frac{\varphi(x_n) - \varphi(y_n)}{\|\varphi(x_n) - \varphi(y_n)\|_2}(z) = 0 \quad + \text{ for all } z \notin B(x, r).$$

This means that $\text{supp}(\xi) \subset B(x, r)$ for all $r > 0$. Therefore, $\text{diam}(\text{supp}(\xi)) = 0$ and so $\xi = 0$ almost everywhere with respect to μ. This contradiction proves that $S_x M = \phi$.

An analogous argument using weak limits establishes that $W_x M = \{0\}$ and the lemma is proved. ■

Corollary 3.3. *M is not C^1 embeddable in $\mathbb{R}^N$ for any N.*

Theorem 3.4. *There exists a Cantor subspace M of a Hilbert space $\mathcal{H}$ such that for any differentiable map $F : [0, 1]^N \to \mathcal{H}$ with $dF(x)$ nonsingular for all $x \in [0, 1]^N$, the set $F^{-1}(M)$ is a finite set.*

Proof. Let $M = \varphi(G)$ above and $\mathcal{H} = L^2(G, \mu)$. Suppose that $F : [0, 1]^N \to \mathcal{H}$ is differentiable with $dF(x)$ nonsingular for all $x \in [0, 1]^N$ and that the set $K = F^{-1}(M)$ is infinite. Since K is compact, there is a sequence $\{z_n\}_{n \geq 1} \subset K$ converging to some $z \in K$ such that the sequence $\{(z_n - z)/|z_n - z|\}_{n \geq 1}$ converges to a unit vector $\sigma \in S_z K \subset \mathbb{R}^N$. Then $dF(z)\sigma \neq 0$ and $dF(z)\sigma/\|dF(z)\sigma\|_2 \in S_{F(z)} M$. But $S_{F(z)} M = \phi$. This contradiction proves the theorem. ■

Theorem 3.5. *There exist Cantor subspaces X of $\mathbb{R}^N$ and bi-Lipschitz maps $f : X \to f(X) \subset \mathcal{H}$, where $\mathcal{H}$ is an infinite dimensional Hilbert space, such that f is not strongly + differentiable at <u>any</u> point of X.*

Proof. Let $M = \varphi(G)$ above and recall that $\dim_m(M) = 1$. Therefore M is bi-Lipschitz embeddable in $\mathbb{R}^N$, $N \geq 2$. Let $j : M \hookrightarrow \mathbb{R}^N$, $N \geq 2$, be any bi-Lipschitz embedding and let $X = j(M)$ inherit the C^1 structure from $\mathbb{R}^N$. Because $S_x M = \phi$ for all $x \in M$, the map

$$f = j^{-1} : j(M) = X \longrightarrow M \subset L^2(G, \mu)$$

is a bi-Lipschitz map which is not strongly differentiable at <u>any</u> point of X. ■

Corollary 3.6. *The map f above can agree with a nonsingular differentiable map at most on a finite set.*

Observe that the map f above is weakly differentiable at every point of X with weak derivative equal to zero.

We conclude by remarking that the above construction represents a family of examples. Indeed, any compact ultrametric space $(\prod_{n\geq 1} \mathbb{Z}q_n,\ \{r_n\}_{n\geq 1})$ yields one such example provided that the sequences $\{q_n\}_{n\geq 1} \subset \mathbb{N}$ and $\{r_n\}_{n\geq 1} \subset (0,1)$ are chosen so that $\dim_m(\prod_{n\geq 1} \mathbb{Z}q_n) < \infty$ and $a r_n^D \leq \mu(B(x, r_n)) \leq A r_n^D$ for some constants $0 < D < 2$ and $0 < a \leq A$. In such a case, the map φ in Lemma 3.1 becomes $\varphi(x) = \rho(x, \cdot)^{(1-D/2)}$.

References

1. Aronszajn, N., *"Differentiability of Lipschitz mapping between Banach spaces"*, Studia Mathematica, T. LVII. (1976), 147–190.
2. Assouad, P., *"Plongements Lipschitziens dans $\mathbb{R}^N$"*, Bull. Soc. Math. France **111** (1983), 429–448.
3. Billingsley, P., *Ergodic Theory and Information*, Wiley, New York, 1965.
4. Bouligand, G., *"Ensembles impropres et ordre dimensionnel"*, Bull. Sci. Math. **52** (1928), 320–344 and 361–376.
5. Milnor, J., *Topology from a Differentiable Viewpoint*, Univ. of Virginia Press, Charlottesville, 1965.
6. Movahedi-Lankarani, H., *Minimal Lipschitz Embeddings*, Ph.D. Thesis, Pennsylvania State University, 1990.
7. Preiss, J., J. Funct. Anal. 91 (1990), pp. 312–345.
8. Zahorski, Bull. Soc. Math., France 74 (1946), pp. 147–178.

Received October 28, 1991