NOTES ON SARD'S THEOREM

Definition. A set $A \subset R^n$ is a *null set* if, for all $\varepsilon > 0$, one may find a finite or countable covering of A by cubes (C_i), so that $\sum_i vol(C_i) < \varepsilon$.

Remark. Also called 'set of measure 0', a term we avoid, since we're not defining 'measure'. A cube is a cartesian product of n closed non-degenerate intervals, and the volume of a cube is the product of the lengths of those intervals.

Subsets of null sets and countable unions of null sets are null sets.

If nothing else is said, we mean 'closed cubes'. But we get an equivalent definition using coverings by open cubes, since given a closed cube Q, and any $\delta > 0$, we can find an open cube U containing Q, with the bound $vol(\overline{U}) \leq (1+\delta)vol(Q))$.

Proposition. $U \subset R^m$ open, $f : U \to R^n$ locally Lipschitz, $m \leq n$. Then if $X \subset U$ is a null set, $f(X)$ is a null set. (In particular, this holds if f is a C^1 map.)

Corollary 1. If $f : U \to R^n$ is of class C^1 with $U \subset R^m$ open and $m < n$, then $f(U)$ is a null set in R^n.

This allows us to define null subsets of differentiable manifolds, using local charts. (*Exercise:* explain exactly how this is done.)

Corollary 2. Let $f : M \to N$ be a C^1 map between manifolds. If $dim(M) < dim(N)$, then $f(M)$ is a null subset of N. If $dim(M) = dim(N)$ and $X \subset M$ is a null set, then $f(X)$ is a null subset of M.

We next observe that a cube $Q \subset R^n$ is not a null subset of R^n (this does indeed require proof!). First, if $Q \subset \sqcup_{i \geq 1} C_i$ is a cube contained in a (countable) union of essentially disjoint cubes (not intersecting on their interiors), then:

$$vol(Q) \leq \sum_i vol(C_i).$$

Proof: We may assume all the C_i intersect Q in its interior (discarding those that don't), and also that all $C_i \subset Q$ (replacing C_i by the cube $C_i \cap Q$ of smaller volume, if needed.) Then $Q = \sqcup_i C_i$: the C_i define a tiling [1] of Q,

[1] A *tiling* of a compact set $A \subset R^n$ is a covering of A by countably many cubes with disjoint and nonempty interiors. A countable partition of an interval $[a, b] \subset R$ is the same as a tiling of $[a, b]$. A *product tiling* of a cube $Q = [a_1, b_1] \times [a_2, b_2] \subset R^2$ is defined by countable partitions $\mathcal{P}$ of $[a_1, b_1]$, $\mathcal{Q}$ of $[a_2, b_2]$, taking all cartesian products of an atom of

and then the inequality claimed is in fact an equality. To see this, refine this tiling to a product tiling $(C'_j)_{j\geq 1}$, without changing the sum of volumes. And then with $Q=[a,b]\times[c,d]$ and $\mathcal{P}=\{I_k\}_k, \mathcal{Q}=\{J_l\}_l$ the countable partitions of $[a,b]$ (resp. $[c,d]$) defining the product tiling, we have:

$$vol(Q)=(b-a)(d-c)=(\sum_k vol(I_k))(\sum_l vol(J_l))$$

$$=\sum_{k,l} vol(I_k)vol(J_l)=\sum_j vol(C'_j)=\sum_i vol(C_i).$$

Next observe that, given a finite or countable collection of cubes $(Q_i)_{i\geq 1}$, we may find a second countable collection $(C_j)_{j\geq 1}$ of essentially disjoint cubes with the same union, without increasing the sum of volumes:

$$\bigsqcup_{j\geq 1} C_j=\bigcup_{i\geq 1} Q_i, \quad \sum_j vol(C_j)\leq \sum_i vol(Q_i).$$

Outline. Show this first for two intersecting cubes: for example, two such cubes in R^2 have the same union as at most seven essentially disjoint cubes. Then use induction on the number of cubes.

Lemma. A cube $Q\subset R^n$ is not a null set.

Indeed, for any (finite) covering by open cubes Q_i, we have, from the remarks above: $vol(Q)\leq \sum_i vol(Q_i)$.

Exercise. Use the facts described above (in particular, the lemma) to show that null subsets of R^n (or of a manifold) have empty interior.

SARD'S THEOREM

Let $f: M^m\to N^n$ be differentiable. The *critical set* of f is:

$$Crit(f)=C_f=\{x\in M; df(x) \text{ is not onto }\}$$

(Recall $df(x)\in\mathcal{L}(T_xM, T_{f(x)}N)$.)

Set of critical values: $f(C_f)\subset N$.

Sard's theorem. Let $U\subset R^m$ open, $f:U\to R^n$ of class C^k, $k> m-n+1$. Then $f(C_f)$ is a null set in R^n.

Comment on differentiability. The value of k in this statement is not quite optimal, but works for the proof given. Note it is invariant under

---

$\mathcal{P}$ and one of $\mathcal{Q}$.

reducing both domain and range dimension by one (used in Step 3); and also that if $m \geq n$, we have $m - n + 1 \geq m/n$ (check it.) $k > m/n$ is used in step 1. As seen above (Corollary 2), we only need to prove the result when $m \geq n$.

Corollary: The same is true for $f : M \to N$ of class C^k, where M and N are manifolds: $f(C_f)$ is a null set in N.

Proof. Exercise: reduce to Sard's theorem in euclidean space, using local charts on M and N.

Proof of Sard's theorem in euclidean space. The proof is by induction on the domain dimension m, using the corollary (for domain manifolds of dimension $m - 1$) as induction hypothesis. (The induction starts with $m = 0$, where the result is trivial. But we can also check it directly for C^2 functions $f : R \to R$, using the argument in step 1. *Exercise.*) Consider for each $i \geq 1$ the subsets of C_f:

$$\Sigma_i = \{x \in U;\ \text{all partial derivatives of order} \leq i \text{ of each component of } f \text{ vanish at } x\}.$$

Then:

$$C_f = (C_f \setminus \Sigma_1) \cup (\Sigma_1 \setminus \Sigma_2) \cup \ldots \cup (\Sigma_{k-1} \setminus \Sigma_k) \cup \Sigma_k,$$

where we choose a $k > m/n$.

Step 1: $f(\Sigma_k)$ is a null set in R^n.

Proof. Let $C \subset U$ be a cube of edge length l containing a point $x \in \Sigma_k$. By Taylor's theorem and compactness, for $x \in \Sigma_k \cap C, y \in C$ we have the order k Lipschitz condition:

$$||f(x) - f(y)|| \leq L||x - y||^k,$$

where L depends on f, k and C. Now fix an $s > 0$, and subdivide C into s^m cubes C_i of edge length l/s, hence each C_i contained in a ball in R^n of diameter $\sqrt{m}(l/s)$. Thus if C_i contains a point of Σ_k, the above Lipschitz condition implies its image $f(C_i)$ is contained in a cube in $Q_i \subset R^n$ of edge length:

$$(const.)s^{-k}, \quad const = L(\sqrt{m}l)^k.$$

So $f(C \cap \Sigma_k)$ may be covered by s^m cubes with total volume bounded above by: $s^m.(const)^n s^{-nk}$. Since $m - nk < 0$, this can be made arbitrarily small, by taking s sufficiently large. Thus $f(\Sigma_k \cap C)$ is a null set for any such cube C, and covering Σ_k by countably many of them we see $f(\Sigma_k)$ is a null set.

Step 2: $f(\Sigma_i \setminus \Sigma_{i+1})$ is a null set, if $i \geq 1$.

Let $x_0 \in \Sigma_i \setminus \Sigma_{i+1}$. Let $g : U \to R$ be a partial derivative of order i of some component of f so that $dg(x_0)[e_j] \neq 0$ for some e_j in the standard basis of R^m. (Such a g exists since $x_0 \notin \Sigma_{i+1}$.) In particular $dg(x_0)$ is onto R, so g is a submersion at (and near) x_0, and for a neighborhood $V \subset U$ of x_0 in R^m, $S = V \cap g^{-1}(0)$ is a submanifold of R^m, of dimension $m-1$. And $\Sigma_i \cap V \subset S$, since g is a partial derivative of order i.

At points $x \in S$ we have $T_x S = ker(dg(x))$ is an $(m-1)$-dimensional subspace of R^m, so g is a submersion also at x and $x \notin \Sigma_{i+1}$. This shows $\Sigma_i \cap V \subset (\Sigma_i \setminus \Sigma_{i+1}) \cap V$, so in fact: $\Sigma_i \cap V = (\Sigma_i \setminus \Sigma_{i+1}) \cap V$. And since $i \geq 1$:

$$(\Sigma_i \setminus \Sigma_{i+1}) \cap V = \Sigma_i \cap V \subset Crit(f) \cap V \subset Crit(f_{|S}).$$

So we may apply the induction hypothesis (in the form of the corollary) to $f_{|S} : S \to R^n$ and conclude:

$$f((\Sigma_i \setminus \Sigma_{i+1}) \cap V) \subset f_{|S}(Crit(f_{|S}))$$

is a null set in R^n. Now cover $\Sigma_i \setminus \Sigma_{i+1}$ by countably many such V to finish this step.

Step 3: (cp. [Milnor]). $f(C_f \setminus \Sigma_1)$ is a null set in R^n.

Let $x_0 \in C_f \setminus \Sigma_1$. Then some partial derivative of some component of f does not vanish at x_0, say $\frac{\partial f_1}{\partial x_1}(x_0) \neq 0$. Consider the map $h : U \to R^m$, $h(x) = (f_1(x), x_2, \ldots, x_m)$. $dh(x_0)$ is an isomorphism, so there exist neighborhoods $V \subset U$ of x_0, $V' \subset R^m$ of $h(x_0)$ so that h is a diffeomorphism from V to V'. Let $g = f \circ h^{-1} : V' \to R^n$. The critical set of g is $C_g = h(C_f \cap V) \subset V'$, and it is enough to show that the set $g(C_g) = f(C_f \cap V)$ of critical values of g is a null set in R^n.

It is easy to see that g preserves the first component, that is:

$$g(t,y) = (t, g_t(y)), \quad (t,y) \in V' \subset R \times R^{m-1}, y \in R^{m-1}, \quad g_t : V'_t \to R^{n-1},$$

where $V'_t \subset R^{m-1}$ is the open set of $y \in R^{m-1}$ such that $(t,y) \in V'$, and g_t is of class C^k. From the form of the differential of g, one sees that:

$$C_g = \bigcup_{t \in I} \{t\} \times C_{g_t}, \quad C_{g_t} \subset V'_t \subset R^{m-1},$$

where the interval $I \subset R$ is the projection of V' onto R. For the set of critical values:

$$g(C_g) = \bigcup_{t \in I} \{t\} \times g_t(C_{g_t}) \subset R^n.$$

Now use the induction hypothesis: Sard's theorem is true for C^k maps from open sets in R^{m-1} to R^{n-1} (note that the requirement on k doesn't change.) This allows us to conclude that $g_t(C_{g_t})$ is a null set in R^{n-1}, for all $t \in I$. Since $g_t(C_{g_t})$ is measurable in R^n (countable union of compact sets), FUBINI'S Theorem (see below) allows us to conclude $g(C_g)$ is a null set in R^n.

Remark/problems. Only a small part of the full statement of Fubini's theorem is needed, and in fact what is needed is contained in problem 2 below.

Problem 1. (Easy.) Suppose $A \subset R^n$ is a null set. Then for any $B \subset R^p$, $A \times B$ is a null set in $R^n \times R^p$.

Problem 2. (Not as easy.) Let $Q = I \times Q' \subset R^n$ be a cube ($I = [0,1] \subset R$, $Q' \subset R^{n-1}$ a cube.) Let $K \subset int(Q)$ be a compact subset of its interior. Suppose that for each $t \in I$, the set $K_t = \{y \in Q'; (t,y) \in K\}$ is a null set in Q'. Then K is a null set.