Sard's Theorem

An extremely important notion in differential topology is that that of general position or genercity. A particular map may have some horrible pathologies but often a nearby map has much nicer properties.

For example the map

$$f(\theta) = ((\cos(2\theta)\cos(\theta), \cos(2\theta)\sin(\theta), 0).$$

maps the unit circle in the plain to the a figure 8 lying in a plane in $\mathbb{R}^3$ while the near by map

$$f_\epsilon(\theta) = (\cos(2\theta)\cos(\theta), \cos(2\theta)\sin(\theta), \epsilon\cos(\theta)).$$

is an embedding. We will develop a general setting in which we can decide when a nearby map will have some nice property. These ideas have been central in topology since early days of Lagrange, Poincaré and where put into a modern efficient setting by Thom and Smale.

The most basic result we will need is Sard's Theorem. A subset of a manifold is said to have measure zero if its intersection with every chart has measure zero with respect to the Lebesque measure on $\mathbb{R}^n$. We will need an easy version of Fubini's theorem.

Theorem 12.1. *Suppose a measureable $C \subset \mathbb{R}^n$ has the property that for all $t \in \mathbb{R}$ $C \cap \{t\} \times \mathbb{R}^{n-1}$ has measure zero. Then C has measure zero.*

We will also use the following lemma.

Lemma 12.2. *If $C \subset \mathbb{R}^m$ is measureable and $f : \mathbb{R}^m \to \mathbb{R}^n$ is continuous then $f(C)$ is measureable.*

Theorem 12.3. *Let $f : M \to N$ be a smooth map of finite dimensional manifolds. Then the set of critical values has measure zero in N.*

Proof. (Copied from Milnor's little blue book *Topology from the differentiable viewpoint*, this proof does not give the sharp result that a C^k map with $k \geq \max\{1, m-n+1\}$ also satisifies the conclusion.) The definition of measure zero is local so it suffices to prove the result in case $M \subset \mathbb{R}^m$ and $N \subset \mathbb{R}^n$ are open subsets.

The proof is by induction on m the dimension of the domain. The case $m = 0$ is trivial. Let $C = Crit(f)$ denote the critical set of f. It suffices to prove that for every point $y \in f(C)$ there is neighborhood of y whose intersection with $f(C)$ has measure zero. Now set

$$C_s = \{x \in M | d_x^j f = 0, \text{for all } 1 \leq j \leq k\}$$

Then $C \supset C_1 \supset C_2 \supset \dots$ is a desceding sequence of closed sets and hence measureable sets. Futhermore the sets $f(C_s \setminus C_{s+1})$ are all measureable.

The proof has three steps. If $m \leq n$ then you can skip directly to step 3.

Step 1. $f(C \setminus C_1)$ has measure zero. If $x \in C \subset C_1$ then there is some first partial which doesn't vanish so assume that

$$\frac{\partial f^1}{\partial x_1}(x) \neq 0.$$

Then we consider the map $g \colon \mathbb{R}^m \to \mathbb{R}^m$.

$$g(x^1, \dots, x^m) = (f^1(x^1, \dots, x^m), x^2, \dots, x^m)$$

Notice that from our assumption

$$d_x g = \begin{bmatrix} \frac{\partial f^1}{\partial x_1}(x) & \frac{\partial f^1}{\partial x_2}(x) & \dots & \frac{\partial f^1}{\partial x_m}(x) & \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots 1 & \end{bmatrix}$$

which is clearly invertible. The inverse function theorem then provides an inverse, $h\colon V \to \mathbb{R}^m$,on small neighborhood of x Then consider the map $f \circ h$ we have

$$f \circ h(x^1, \ldots, x^m) = (x^1,\ f^2 \circ h(x^1, \ldots, x^m), \ldots,\ f^n \circ h(x^1, \ldots, x^m)).$$

So $f(C \cap h(V)) = f \circ h(h^{-1}(C) \cap V)$. The inverse image of the set critical $h^{-1}(C) \cap V$ are simply the critical points of $f \circ h$. If we set

$$k_t(x^2, x^3, \ldots, x^m) = (f^2 \circ h(t, \ldots, x^m), \ldots,\ f^n \circ h(t, \ldots, x^m))$$

then

$$h^{-1}(C) \cap V = \cup_t \{t\} \times Crit(k_t).$$

By the induction hypothesis we have

$$k_t(Crit(k_t))$$

has measure zero in $\mathbb{R}^{m-1}$ and hence by Fubini

$$f(C \cap h(V)) = \cup_t \{t\} \times k_t(Crit(k_t))$$

has measure zero in $\mathbb{R}^m$.

Step 2. Suppose $x \in C_s \backslash C_{s+1}$. Then without loss of generality we can assume that there is some s-th order mixed partial derivative so that if we set

$$w = \frac{\partial^{i_1+\ldots+i_m} f}{\partial(x^1)^{i_1} \ldots \partial(x^m)^{i_m}}$$

so that

$$\frac{\partial w}{\partial x^1}(x) \neq 0.$$

Define

$$g(x^1, \ldots, x^m) = (w(x^1, \ldots, x^m), x^2, \ldots, x^m).$$

Again this map is a diffeomorphism with inverse $h\colon V \to \mathbb{R}^m$ for some neighborhood V of $g(x)$. Let

$$k = f \circ h$$

and let

$$\bar{k} = k|_{\{0\}\times\mathbb{R}^{m-1}\cap V}.$$

Clearly $g(C_k \cap h(V)) \subset \{0\} \times \mathbb{R}^{m-1} \cap V$ and the critical set of $\bar{k}$ contains $g(C_k \cap h(V))$ since it contains $g(C \cap h(V))$. Thus

$$f(C_k \cap h(V)) \subset \bar{k}(Crit(\bar{k}))$$

which has measure zero by the induction hypothesis.

Step 3. Suppose that $x \in C_k$ where $k+1 > \frac{m}{n}$. Choose a little cube I of side length δ. We have from Taylors theorem and the compactness of I that there is a constant $M > 0$ so that for all $y \in I$ and all $x \in C_k \cap I$

$$\|f(x) - f(y)\| \leq M\|x-y\|^{k+1}$$

Subdivide I into l^m subcubes of side length δ/l. By the above estimate if I' is such a subcube containing a point of C_k then $f(I')$ is contained in a cube of side length at most

$$2M\sqrt{m}(\delta/l)^{k+1}$$

Thus the $f(C_k \cap I)$ is contained in set of total volume bounded above

$$(2M\sqrt{m}(\delta/l)^{k+1})^n l^m = Cl^{m-n(k+1)}.$$

By our assumption this goes to zero as l goes to infinity. □