Now we investigate the measure of the critical values of a map $f : M \longrightarrow N$ where $\dim M = \dim N$. Of course the set of critical points need not have measure zero, but we shall see that because the values of $f$ on the critical set do not vary much, the set of critical *values* will have measure zero.

**Theorem 1.38** (Equidimensional Sard)**.** *Let $f : M \longrightarrow N$ be a $C^1$ map of $n$-manifolds, and let $C \subset M$ be the set of critical points. Then $f(C)$ has measure zero.*

*Proof.* It suffices to show result for the unit cube. Let $f : I^n \longrightarrow \mathbb{R}^n$ a $C^1$ map and let $C \subset I^n$ be the set of critical points.

Let $a$ be the Lipschitz constant for $f, I^n$, obtained from the mean value equation

$$f(y) - f(x) = Df(z)(y - x), \tag{17}$$

and let $T_x$ be the affine map approximating $f$ at $x$, i.e.

$$T_x(y) = f(x) + Df(x)(y - x). \tag{18}$$

Then subtracting equations (17),(18), we obtain

$$f(y) - T_x(y) = (Df(z) - Df(x))(y - x). \tag{19}$$

Since $Df$ is continuous, there is a positive function $b(\epsilon)$ with $b \to 0$ as $\epsilon \to 0$ such that

$$||f(y) - T_x(y)|| \leq b(|y - x|)||y - x||.$$

If $x$ is a critical point, then $T_x$ has vanishing determinant, meaning that it maps $\mathbb{R}^n$ into a hyperplane $P_x \subset \mathbb{R}^n$ (i.e. of dimension $n - 1$). If $||y - x|| < \epsilon$, then $||f(y) - f(x)|| < a\epsilon$, and by (19), the distance of $f(y)$ from $P_x$ is less than $\epsilon b(\epsilon)$.

Therefore $f(y)$ lies in the cube centered at $f(x)$ of edge $a\epsilon$, but only $\epsilon b\epsilon$ in distance from the plane $P_x$. Choose the cube to have a face parallel to $P_x$, and we conclude $f(y)$ is in a region of volume $(a\epsilon)^{n-1}2\epsilon b(\epsilon)$.

Now partition $I^n$ into $h^n$ cubes each of edge $h^{-1}$. Any such cube containing a critical point $x$ is certainly contained in a ball around $x$ of radius $r = h^{-1}\sqrt{n}$. The image of this ball then has volume $\leq (ar)^{n-1}2rb(r) = Ar^n b(r)$ for $A = 2a^{n-1}$. The total volume of all the images is then less than

$$h^n A r^n b(r) = A n^{n/2} b(r).$$

Note that $A$ and $n$ are fixed, while $r = h^{-1}\sqrt{n}$ is determined by the number $h$ of cubes. By increasing the number of cubes, we may decrease their radius arbitrarily, and hence the above total volume, as required. ☐

The argument above will not work for $\dim N < \dim M$; we need more control on the function $f$. In particular, one can find a $C^1$ function from $I^2 \longrightarrow \mathbb{R}$ which fails to have critical values of measure zero (hint: $C + C = [0, 2]$ where $C$ is the Cantor set). As a result, Sard's theorem in general requires more differentiability of $f$.

**Theorem 1.39** (Big Sard's theorem)**.** *Let $f : M \longrightarrow N$ be a $C^k$ map of manifolds of dimension $m$, $n$, respectively. Let $C$ be the set of critical points, i.e. points $x \in U$ with*

$$\text{rank } Df(x) < n.$$

*Then $f(C)$ has measure zero if $k > \frac{m}{n} - 1$.*

*Proof.* As before, it suffices to show for $f : I^m \longrightarrow \mathbb{R}^n$.

Define $C_1 \subset C$ to be the set of points $x$ for which $Df(x) = 0$. Define $C_i \subset C_{i-1}$ to be the set of points $x$ for which $D^j f(x) = 0$ for all $j \leq i$. So we have a descending sequence of closed sets:

$$C \supset C_1 \supset C_2 \supset \cdots \supset C_k.$$

We will show that $f(C)$ has measure zero by showing

1. $f(C_k)$ has measure zero,

2. each successive difference $f(C_i \backslash C_{i+1})$ has measure zero for $i \geq 1$,

3. $f(C \backslash C_1)$ has measure zero.

**Step 1:** For $x \in C_k$, Taylor's theorem gives the estimate

$$f(x+t) = f(x) + R(x,t), \quad \text{with} \ \ ||R(x,t)|| \leq c||t||^{k+1},$$

where $c$ depends only on $I^m$ and $f$, and $t$ sufficiently small.

If we now subdivide $I^m$ into $h^m$ cubes with edge $h^{-1}$, suppose that $x$ sits in a specific cube $I_1$. Then any point in $I_1$ may be written as $x+t$ with $||t|| \leq h^{-1}\sqrt{m}$. As a result, $f(I_1)$ lies in a cube of edge $ah^{-(k+1)}$, where $a = 2cm^{(k+1)/2}$ is independent of the cube size. There are at most $h^m$ such cubes, with total volume less than

$$h^m(ah^{-(k+1)})^n = a^n h^{m-(k+1)n}.$$

Assuming that $k > \frac{m}{n} - 1$, this tends to 0 as we increase the number of cubes.
**Step 2:** For each $x \in C_i \backslash C_{i+1}$, $i \geq 1$, there is a $i+1^{th}$ partial $\partial^{i+1} f_j / \partial x_{s_1} \cdots \partial x_{s_{i+1}}$ which is nonzero at $x$. Therefore the function

$$w(x) = \partial^k f_j / \partial x_{s_2} \cdots \partial x_{s_{i+1}}$$

vanishes at $x$ but its partial derivative $\partial w / \partial x_{s_1}$ does not. WLOG suppose $s_1 = 1$, the first coordinate. Then the map

$$h(x) = (w(x), x_2, \ldots, x_m)$$

is a local diffeomorphism by the inverse function theorem (of class $C^k$) which sends a neighbourhood $V$ of $x$ to an open set $V'$. Note that $h(C_i \cap V) \subset \{0\} \times \mathbb{R}^{m-1}$. Now if we restrict $f \circ h^{-1}$ to $\{0\} \times \mathbb{R}^{m-1} \cap V'$, we obtain a map $g$ whose critical points include $h(C_i \cap V)$. Hence we may prove by induction on $m$ that $g(h(C_i \cap V)) = f(C_i \cap V)$ has measure zero. Cover by countably many such neighbourhoods $V$.
**Step 3:** Let $x \in C \backslash C_1$. Then there is some partial derivative, wlog $\partial f_1 / \partial x_1$, which is nonzero at $x$. the map

$$h(x) = (f_1(x), x_2, \ldots, x_m)$$

is a local diffeomorphism from a neighbourhood $V$ of $x$ to an open set $V'$ (of class $C^k$). Then $g = f \circ h^{-1}$ has critical points $h(V \cap C)$, and has critical values $f(V \cap C)$. The map $g$ sends hyperplanes $\{t\} \times \mathbb{R}^{m-1}$ to hyperplanes $\{t\} \times \mathbb{R}^{n-1}$, call the restriction map $g_t$. A point in $\{t\} \times \mathbb{R}^{m-1}$ is critical for $g_t$ if and only if it is critical for $g$, since the Jacobian of $g$ is

$$\begin{pmatrix} 1 & 0 \\ * & \frac{\partial g_t^i}{\partial x_j} \end{pmatrix}$$

By induction on $m$, the set of critical values for $g_t$ has measure zero in $\{t\} \times \mathbb{R}^{n-1}$. By Fubini, the whole set $g(C')$ (which is measurable, since it is the countable union of compact subsets (critical values not necessarily closed, but critical points are closed and hence a countable union of compact subsets, which implies the same of the critical values.) is then measure zero. To show this consequence of Fubini directly, use the following argument:

First note that for any covering of $[a,b]$ by intervals, we may extract a finite subcovering of intervals whose total length is $\leq 2|b-a|$. Why? First choose a minimal subcovering $\{I_1, \ldots, I_p\}$, numbered according to their left endpoints. Then the total overlap is at most the length of $[a,b]$. Therefore the total length is at most $2|b-a|$.

Now let $B \subset \mathbb{R}^n$ be compact, so that we may assume $B \subset \mathbb{R}^{n-1} \times [a,b]$. We prove that if $B \cap P_c$ has measure zero in the hyperplane $P_c = \{x^n = c\}$, for any constant $c \in [a,b]$, then it has measure zero in $\mathbb{R}^n$.

If $B \cap P_c$ has measure zero, we can find a covering by open sets $R_c^i \subset P_c$ with total volume $< \epsilon$. For sufficiently small $\alpha_c$, the sets $R_c^i \times [c - \alpha_c, c + \alpha_c]$ cover $B \cap \bigcup_{z \in [c-\alpha_c, c+\alpha_c]} P_z$ (since $B$ is compact). As we vary $c$, the sets $[c - \alpha_c, c + \alpha_c]$ form a covering of $[a, b]$, and we extract a finite subcover $\{I_j\}$ of total length $\leq 2|b - a|$.

Let $R^i_j$ be the set $R^i_c$ for $I_j = [c - \alpha_c, c + \alpha_c]$. Then the sets $R^i_j \times I_j$ form a cover of $B$ with total volume $\leq 2\epsilon|b - a|$. We can make this arbitrarily small, so that $B$ has measure zero.

□

**Corollary 1.40.** *Let $M$ be a compact manifold with boundary. There is no smooth map $f : M \longrightarrow \partial M$ leaving $\partial M$ pointwise fixed. Such a map is called a smooth retraction of $M$ onto its boundary.*

*Proof.* Such a map $f$ must have a regular value by Sard's theorem, let this value be $y \in \partial M$. Then $y$ is obviously a regular value for $f|_{\partial M} = \mathrm{Id}$ as well, so that $f^{-1}(y)$ must be a compact 1-manifold with boundary given by $f^{-1}(y) \cap \partial M$, which is simply the point $y$ itself. Since there is no compact 1-manifold with a single boundary point, we have a contradiction. □

For example, this shows that the identity map $S^n \longrightarrow S^n$ may not be extended to a smooth map $f : \overline{B(0,1)} \longrightarrow S^n$.

**Lemma 1.41.** *Every smooth map of the closed $n$-ball to itself has a fixed point.*

*Proof.* Let $D^n = \overline{B(0,1)}$. If $g : D^n \longrightarrow D^n$ had no fixed points, then define the function $f : D^n \longrightarrow S^{n-1}$ as follows: let $f(x)$ be the point nearer to $x$ on the line joining $x$ and $g(x)$.

This map is smooth, since $f(x) = x + tu$, where

$$u = ||x - g(x)||^{-1}(x - g(x)),$$

and $t$ is the positive solution to the quadratic equation $(x+tu)\cdot(x+tu) = 1$, which has positive discriminant $b^2 - 4ac = 4(1 - |x|^2 + (x \cdot u)^2)$. Such a smooth map is therefore impossible by the previous corollary. □

**Theorem 1.42** (Brouwer fixed point theorem)**.** *Any continuous self-map of $D^n$ has a fixed point.*

*Proof.* The Weierstrass approximation theorem says that any continuous function on $[0,1]$ can be uniformly approximated by a polynomial function in the supremum norm $||f||_\infty = \sup_{x\in[0,1]} |f(x)|$. In other words, the polynomials are dense in the continuous functions with respect to the supremum norm. The Stone-Weierstrass is a generalization, stating that for any compact Hausdorff space $X$, if $A$ is a subalgebra of $C^0(X, \mathbb{R})$ such that $A$ separates points $(\forall x, y, \exists f \in A \;: f(x) \neq f(y))$ and contains a nonzero constant function, then $A$ is dense in $C^0$.

Given this result, approximate a given continuous self-map $g$ of $D^n$ by a polynomial function $p'$ so that $||p' - g||_\infty < \epsilon$ on $D^n$. To ensure $p'$ sends $D^n$ into itself, rescale it via

$$p = (1 + \epsilon)^{-1}p'.$$

Then clearly $p$ is a $D^n$ self-map while $||p - g||_\infty < 2\epsilon$. If $g$ had no fixed point, then $|g(x) - x|$ must have a minimum value $\mu$ on $D^n$, and by choosing $2\epsilon = \mu$ we guarantee that for each $x$,

$$|p(x) - x| \geq |g(x) - x| - |g(x) - p(x)| > \mu - \mu = 0.$$

Hence $p$ has no fixed point. Such a smooth function can't exist and hence we obtain the result. □