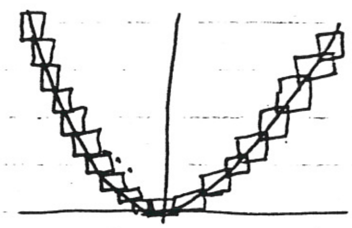
# Sard's Theorem

Def: A cube in $\mathbb{R}^n$ is a subset $I = [a_1, b_1] \times \cdots \times [a_n, b_n]$. The volume $\mathrm{vol}(I) = \prod_{i=1}^{n} (b_i - a_i)$.

A subset $X \subset \mathbb{R}^n$ has measure zero if, for every $\varepsilon > 0$, there exist cubes $I_1, I_2, \ldots$ containing $X$, $X \subset \bigcup_{i=1}^{\infty} I_i$, with total volume

$$\sum_{i=1}^{\infty} \mathrm{vol}(I_i) < \varepsilon.$$



Examples (1) $\{(x, x^2) \mid \in \mathbb{R}\}$ has measure 0

(2) The rationals $\mathbb{Q} \subset \mathbb{R}$ have measure 0.
More generally, any countable set in $\mathbb{R}^n$ has measure 0.

Lemma: (1) If $h : U \to V$ is a diffeomorphism, $U$ and $V$ open in $\mathbb{R}^n$, then $C \subset U$ has measure zero iff $h(C)$ has measure 0.

(2) A countable union of measure 0 sets has measure 0.

(3) If $C$ contains an open set then $C$ does not have measure 0.

(4) If $C \subset \mathbb{R}^n$ and $C \cap (\{x\} \times \mathbb{R}^{n-1})$ has measure 0 for each $x \in \mathbb{R}$, then $C$ has measure 0.

Theorem: Let $f : M \to N$ be a smooth map. Let $C \subset M$ be the set of critical points of $f$. Then $f(C)$ has measure zero.

Note: This was proved by A. Sard in 1942. The following corollary had been proved by A.B. Brown in 1935.

Corollary: The set of regular values of a smooth map $M \to N$ is everywhere dense in $N$.

Proof of Sard's Theorem: Since $M$ can be covered by a countable number of charts, it will suffice to prove the theorem for a smooth map

$$f : U \to \mathbb{R}^p, \quad U \text{ open in } \mathbb{R}^n.$$

Recall that $C = \{x \in \mathcal{U} \mid \operatorname{rank} T_x f < p\}$. We will prove the result by induction on $n$. Certainly it is true if $n=0$, since $\mathbb{R}^0 = \{0\}$. Let

$$C_i = \{x \in \mathcal{U} \mid \text{all partial derivatives of } f \text{ of order} \le i \text{ vanish}\}.$$

Then $C \supset C_1 \supset C_2 \supset C_3 \supset \cdots$ is a descending sequence of closed sets. We will show

1. $f(C \smallsetminus C_1)$ has measure $0$
2. $f(C_i \smallsetminus C_{i+1})$ has measure $0$ for $i > 0$.
3. $f(C_k)$ has measure $0$ if $k$ is sufficiently large.

We will prove steps 1 and 2 by using a nonzero partial derivative of $f$ to "straighten out" $f$ so that we may apply Fubini's theorem (part (4) of the Lemma).

Proof of step 1: We may assume $p \ge 2$ since $C = C_1$ when $p=1$. For each $\bar{x} \in C \smallsetminus C_1$ we will find a nhood $V$ of $\bar{x}$ such that $f(V \cap C)$ has measure $0$. Since $\mathcal{U}$ is second countable this is sufficient by part (2) of the Lemma.

Since $\bar{x} \notin C_1$ there is a partial derivative, say $\partial f_1 / \partial x_1$ which is nonzero at $\bar{x}$. Let

$$h : \mathcal{U} \longrightarrow \mathbb{R}^n$$

be $h(x_1, \dots, x_n) = (f_1(x), x_2, \dots, x_n)$. This is a diffeomorphism in a nhood of $\bar{x}$, so maps some neighborhood $V$ of $\bar{x}$ diffeomorphically onto an open set $V' \subset \mathbb{R}^n$. Let $g = f \circ h^{-1}$.

The set of critical values of $g$ is the set $f(C \cap V)$. The crucial property of $g$ is that

$$g(t, x_2, \dots, x_n) \in \{t\} \times \mathbb{R}^{p-1}$$

$$\begin{tikzcd} V \arrow[r, "h", "\cong"'] \arrow[d, "f"'] & V' \arrow[dl, "g"] \\ \mathbb{R}^p & \end{tikzcd}$$

so that the derivative of $g$ is

$$Dg(t, x_2, \ldots, x_n) = \begin{bmatrix} 1 & 0 \\ * & Dg_t(x_2, \ldots, x_n) \end{bmatrix}$$

where $g_t$ is $g$ restricted to $\{t\} \times \mathbb{R}^{n-1}$. Therefore a point in $\{t\} \times \mathbb{R}^{n-1}$ is a critical point of $g$ if and only if it is a critical point of $g_t$. By induction, the critical values of $g_t$ have measure zero in $\{t\} \times \mathbb{R}^{n-1}$. The measure zero Fubini's Theorem (Lemma (4)) then implies that $f(C \cap V)$ has measure zero. This completes Step 1.

Proof of step 2: If $\bar{x} \in C_K - C_{K+1}$ there is a $(K+1)^{st}$ derivative

$$\frac{\partial^{K+1} f_i}{\partial x_{j_1} \cdots \partial x_{j_{K+1}}}(\bar{x}) \neq 0$$

Then

$$w(x) = \frac{\partial^K f_i}{\partial x_{j_2} \cdots \partial x_{j_{K+1}}}$$

vanishes at $\bar{x}$ but $\partial w / \partial x_{j_1}$ does not. For simplicity we may assume that $j_1 = 1$. The map

$$h : U \longrightarrow \mathbb{R}^n$$

given by $h(x) = (w(x), x_2, \ldots, x_n)$ is a diffeomorphism in a nhood of $\bar{x}$, say $h : V \longrightarrow V'$, $V$ and $V'$ open in $\mathbb{R}^n$. Again let $g = f \circ h^{-1}$, so that

$$f(C_K \cap V) = gh(C_K \cap V).$$

Note that $h(C_K \cap V) \subset 0 \times \mathbb{R}^{n-1}$ so that $gh(C_K \cap V) = \bar{g}(V' \cap (0 \times \mathbb{R}^{n-1}))$, where $\bar{g}$ is the restriction of $g$ to $V' \cap (0 \times \mathbb{R}^{n-1})$. Clearly $h(C_K \cap V)$ lies in the critical points of $\bar{g}$, so by induction, $gh(C_K \cap V)$ has measure $0$. Since $C_K$ can be covered by a countable number of such nhoods $V$, we have finished Step 2.

Proof of Step 3: Since $C_K$ can be covered by countably many cubes, it is sufficient to show $f(C_K \cap I)$ has measure zero, where $I$ is a cube of side $\delta$. Now, for $x \in C_K \cap I$, Taylor's theorem gives

$$f(x+h) = f(x) + R(x,h), \qquad |R(x,h)| \leq c|h|^{K+1},$$

where $c$ is a constant depending only on $f$ and $I$.

If we subdivide $I$ into cubes of side $\delta/r$ and let $I_1$ be one of these which contains a point $x$ of $C_k$, then for $x+h \in I_1$ we have

$$|h| \leq \sqrt{n}\,(\delta/r)$$

from which it follows that $f(C_k \cap I_1)$ is contained in a cube centered on $f(x)$ of edge

$$2c\,(\sqrt{n}\,\delta)^{k+1}/r^{k+1} = a/r^{k+1}.$$

and hence volume $a^p/r^{p(k+1)}$. Therefore, $f(C_k \cap I)$ is contained in $r^n$ cubes of total volume

$$a^p r^{n-p(k+1)}.$$

For $k+1 > n/p$ this goes to 0 as $r \to \infty$, finishing step 3. //

This proof is taken from Milnor, Topology from the Differentiable Viewpoint, where it is attributed to Pontrjagin.

Corollary: For any countable set of smooth functions $M \to N$, the set of points of $N$ which are regular values for all of them is dense in $N$.

Example: It is important to note the difference between regular values and regular points. The set of regular values is dense, but the set of regular points need not be. For example, if $f: M \to N$ is a constant function, then every point of $N$ but one is a regular value, while the set of regular points is empty.