SARD'S THEOREM

ALEX WRIGHT

Abstract. A proof of Sard's Theorem is presented, and applications to the Whitney Embedding and Immersion Theorems, the existence of Morse functions, and the General Position Lemma are given.

Suppose $f : M^m \to N^n$ is a map from a m-dimensional manifold M to an n-dimensional manifold N. (All manifolds and maps are assumed to be smooth.) A *critical point* pf f is an $x \in M$ such that $(df_x)(T_xM) \neq T_{f(x)}N$. A *critical value* is the image of a critical point.

Theorem (Sard's Theorem)**.** *The set of critical values of f is null.*

We say that a set $S \subset N$ is null if its image in $\mathbb{R}^n$ under every chart is null. If $m < n$ there is a simple proof of Sard's Theorem, and if $n = m$ a relatively short proof can be found in M. Spivak's *Calculus on Manifolds* ([5], p.72). Here we make no assumption on m and n, and we benefit from this extra power in two of the three applications bellow. The following proof is from V. Guillemin and A. Pollack's *Differential Topology* ([1], p.205-207), which in turn cites [3] as its source.

Proof of Sard's Theorem. By passing to charts, and using the fact that there is a countable sub-collection of charts that cover M, we can assume $M = U \subset \mathbb{R}^m$, U open, and $N = \mathbb{R}^n$ (so $f : U \to \mathbb{R}^n$).

To begin, we break up C, the set of critical points of f, into a sequence of nested subsets $C \supset C_1 \supset C_2 \supset \cdots$, where C_1 is the set of all $x \in U$ such that $df_x = 0$, and C_i $(i \geq 1)$ is the set of all x such that all partial derivatives of order at most i vanish at x. We then proceed by induction on m and prove three lemmas. Lemmas 1 and 2 give that $f(C - C_1)$ and $f(C_i - C_{i+1})$ are null. These two lemmas use the inductive hypothesis and the fact that $\mathbb{R}^m$ is second countable (that is to say: if $\{U_\alpha\}$ is set of open sets in $\mathbb{R}^m$, there there is a countable sub-collection $\{U_{\alpha_k}\}$ so that $\cup_\alpha U_\alpha = \cup_k U_{\alpha_k}$). Lemma 1 makes use of Tonelli's Theorem, a variant of Fubini's Theorem. Lemma 3 uses Taylor's Theorem to show that if i is sufficiently big, then $f(C_i)$ is null. These lemmas clearly combine to give Sard's Theorem.

Date: February 2008.

We assume Sard's Theorem is true for $m-1$, and prove the three lemmas. The base case of $m=0$ is trivial, since $\mathbb{R}^0$ is a point.

Lemma (1). *$f(C-C_1)$ is null.*

Proof. Around each $x \in C - C_1$ we will find an open set V_x such that $f(V_x \cap C)$ is null. Since $\mathbb{R}^m$ is second countable, we will then be able to find a countable sub-collection $V_{x_1}, V_{x_2}, \cdots$, that covers $C - C_1$, and we will conclude

$$m(f(C-C_1)) \leq \sum_i m(f(V_{x_i} \cap (C-C_1))) \leq \sum_i m(f(V_{x_i} \cap C)) = 0,$$

where m is Lebesgue measure. So if we fix $x \in U$ it suffices to prove that we can find an open set V containing x with $f(V \cap C)$ null.

Since $x \notin C_1$, $f = (f_1, \cdots, f_m)$ has some partial, say $\frac{\partial f_1}{\partial x_1}$, which does not vanish at x. Define $h : U \to \mathbb{R}^m$ (recall $U \subset \mathbb{R}^m$) by

$$h(x) = (f_1(x), x_2, \cdots, x_m).$$

Now dh_x is non singular, so by the Inverse Function Theorem, h maps some neighbourhood V of x diffeomorphically onto an open set $V' \subset \mathbb{R}^m$. The composition $g = f \circ h^{-1`} : V' \to \mathbb{R}^n$ will then have the same critical values as $f|_V$ (f restricted to V). So we want to show that the set of critical values of g restricted to V' is null. Note that the first coordinates of h and f are the same, so $g = f \circ h^{-1}$ leaves the first coordinate unchanged. Therefore, for each t, g induces a map $g_t : (t \times \mathbb{R}^{m-1}) \cap V' \to \mathbb{R}^n$. Since dg has the form

$$\begin{pmatrix} 1 & 0 \\ * & (\frac{\partial g_i^t}{\partial x_j}) \end{pmatrix}$$

a point $(t, z) \in (t \times \mathbb{R}^{m-1}) \cap V'$ is a critical point of g if and only if z is a critical point for g^t. By induction, the set V^t of critical values of g^t is null for each t. The set of critical points of g is closed, so its image under g, the set V of critical values of g, is Borel. Thus χ_V (the indicator function of V), is measurable, and Tonelli's Theorem gives

$$m(V) = \int_{\mathbb{R}^n} \chi_V = \int_t \int_{\mathbb{R}^{n-1}} \chi_{V^t} = \int_{\mathbb{R}^{n-1}} 0 = 0.$$

Thus V is null and the proof of Lemma 1 is complete. ☐

Lemma (2). *$f(C_k - C_{k+1})$ is null if $k \geq 1$.*

Proof. This is a similar argument, but easier. For each $x \in C_k - C_{k+1}$, there is some $(k+t)$st partial of f that is not zero at x. Thus we can find a kth partial of f, say ρ, that has a first partial, say $\frac{\partial \rho}{\partial x_1}$, that is non-zero at x. Then the map $h : U \to \mathbb{R}^m$ defined by

$$h(x) = (\rho(x), x_2, \cdots, x_m)$$

maps a neighbourhood V of x diffeomorphically onto an open set $V' \subset \mathbb{R}^m$. Since all kth partials vanish on C_k, and ρ is a kth partial, h carries $C_k \cap V$ into the hyperplane $0 \times \mathbb{R}^{m-1}$.

Define $g = f \circ h^{-1'} : V' \to \mathbb{R}^n$. Of course $f|_V$ and $g|_{V'}$ have the same critical values. As in Lemma 1, it suffices to show that the set of critical values of $g|_{V'}$ is null. But these values all come from points in $0 \times \mathbb{R}^{m-1}$. Let $\tilde{g} : (0 \times \mathbb{R}^{m-1}) \cap V' \to \mathbb{R}^n$ be the restriction of g. If x is a critical point of g, then $(d\tilde{g})_x(T_x\mathbb{R}^{m-1}) \subset (dg)_x(T_x\mathbb{R}^m) \neq T_{g(x)}\mathbb{R}^n$, so x is also a critical value for $\tilde{g}$. By induction, the set of critical values of $\tilde{g}$ is null, so Lemma 2 is proved. □

Lemma (3)**.** *For $k > m/n - 1$, $f(C_k)$ is null.*

Proof. Fix such a k. Let $S \subset U$ be a cube with sides of length δ. We will show that $f(C_k \cap S)$ is null. Since U is covered by a countable number of such cubes, this will prove that $f(C_k)$ is null. From Taylor's Theorem, the compactness of S, and the definition of C_k, we see that

$$f(x+h) = f(x) + R(x,h)$$

where $|R(x,h)| < a|h|^{k+1}$ for $x \in C_k \cap S$. Here a is a constant that depends only on f and S. Now subdivide S into r^m cubes whose sides are of length δ/m. Let S_1 be a cube of the subdivision that contains a point x of C_k. Then any point of S_1 can be written as $x + h$ with $|h| < \sqrt{m}(\frac{\delta}{m})$.. Now if $x + h \in S_1$, then

$$|f(x+h) - f(x)| = |R(x+h)| < a(\sqrt{m}\frac{\delta}{m})^{k+1} = b/r^{k+1}$$

where b is a constant. So $f(S_1)$ lies in a cube of side length at most b'/r^{k+1} centered at $f(x)$ (b' a new constant). Hence $f(C_k \cap S)$ is contained in the union of at most r^m cubes having total volume at most

$$r^m (b')^m r^{m-(k+1)n}.$$

If $m - (k+1)n < 0$, (that is $k > m/n - 1$) then letting $r \to 0$ gives that $f(C_k \cap S)$ is null. □

This completes the proof of Sard's Theorem. □

We proceed to our first application.

Theorem. *Every manifold M^m admits an injective immersion into $\mathbb{R}^{2m+1}$.*

Note that if M is compact, then an injective immersion is an embedding, so this theorem comes very close to the Whitney Embedding Theorem, which says: Every manifold M^m can be embedded into R^{2m}. An injective immersion can be turned into an embedding with extra work (see [1], p.53), but the reduction from $2m+1$ to $2m$ is very difficult, and the author knows of no friendly exposition of this $2m$ Whitney Embedding Theorem.

Proof. We assume M can be embedded into some $\mathbb{R}^n$. For compact manifolds, this can be proved using partitions of unity ([2], p.23). If $n = 2m+1$, we're done, so we assume $n > 2m+1$. For $a \in \mathbb{R}^n$, $a \neq 0$, let π_a be the projection of $\mathbb{R}^n$ onto the perp space of a. By iteration, it suffices to show that $\pi_a : M \to \mathbb{R}^{n-1}$ is an injective immersion for at least one a. We will in fact use Sard's Theorem to show that it is true for a.e. a! Define

$$
\begin{aligned}
&g : M \times M \times \mathbb{R} \to \mathbb{R}^n \\
&g(x, y, t) = t(x - y) \\
&h : TM \to \mathbb{R}^n \\
&h((p, v)) = v
\end{aligned}
$$

where $(p, v) \in TM$ represents the tangent vector $v \in \mathbb{R}^n$ at the point $p \in M$. (Note immediately that the domain of g has dimension $2m+1$.) Now, if $\pi_a : M \to \mathbb{R}^{n-1}$ is not injective, then we have some $x, y \in M, t \in \mathbb{R}$ so that $x \neq y$ and $x - y = ta$. That is to say, $g(x, y, 1/t) = a$. Furthermore, if π_a is not an immersion, then there is some $(p, v) \in TM$ such that $v = sa$ for some a. Since M is immersed into $\mathbb{R}^n$, we must have $s \neq 0$, so $h(v/s) = a$.

Now it is clear that if a is in neither the range of g or the range of h, then π_a is the desired injective immersion. Since the dimensions of the domains of g and h are $2m+1$ and $2m$ respectively, and $n > 2m+1$, every point in the range of these functions is a critical value! Thus we can pick almost any $a \in \mathbb{R}^n$ and get that $\pi_a : M \to \mathbb{R}^{n-1}$ is an injective immersion. □

We leave it as an exercise to the reader to modify this proof to get that every M^m can be immersed into $\mathbb{R}^{2m}$ (with the same starting assumption that it can be immersed into some R^n). This essentially comes from the fact that we can drop g, and the domain of h has dimension $2m$ instead of $2m+1$.

Our next application of Sard's Theorem will be the existence of Morse functions. Given a function $f : M \to \mathbb{R}$, a critical point $x \in M$ is called *non-degenerate* if the Hessian of f at x, $\mathrm{Hess}(f)_x = (\frac{\partial^2 f}{\partial x_i \partial x_j})$

is non-singular in local coordinates. See [1] p.42 or compute using the chain rule to see that this does not depend on local coordinates. Such critical points turn out to be very important because f is locally quadratic at these points. (This is known as the Morse Lemma.) If all of f's critical points are non-degenerate, f is called a Morse function: such functions say a great deal about the topology of M. If $M \in \mathbb{R}^3$, and $f(x, y, z) = z$ is a Morse function, we think of filling up $\mathbb{R}^3$ with water up to the level z. The the topology of the part underwater, $f^{-1}((-\infty, z))$, changes only with the water covers a mountain top (of M), fills a valley (saddle point), or meets a bowl (local minimum).:These events correspond to the water level reaching a non-degenerate critical point, and this intuitive picture is used to think of all Morse functions.

Theorem. *There are lots of Morse functions: Given $M \subset \mathbb{R}^n$, and $f : M \to \mathbb{R}$, then $f_a = f + a_1x_1 + \cdots + a_nx_n$ is a Morse function for almost every $a \in \mathbb{R}^n$.*

Proof. Define $g = df = (\frac{\partial f}{\partial x_1} + \cdots + \frac{\partial f}{\partial x_n})$ on M. Note that $df_a = g + a$, and $\text{Hess}(f_a) = \text{Hess}(f) = dg$. Pick any a so that $-a$ is a regular value for g. Then if x is a critical point of f_a, $g(x) = -a$ so $\text{Hess}(f_a)_x = dg_x$ is non-singular. Thus f_a is a Morse function. □

The reader who wants to learn more about Morse theory is urged to consult J. Milnor's *Morse Theory* ([4]). Our final application is the General Position Lemma. Recall that manifolds $M, N \subset \mathbb{R}^n$ are said to be transverse (written $M \pitchfork N$) if $T_pM + T_pN = T_p\mathbb{R}^n$ for all $p \in M \cap N$. Tranverse manifolds are said to be in general position.

Theorem (General Position Lemma)**.** *For almost every $a \in \mathbb{R}^n$, $(M + a) \pitchfork N$.*

Note that if $\dim M + \dim N < n$, and $p \in M \cap N$, then $T_pM + T_pN \neq T_p\mathbb{R}^n$ (the dimension of the left hand side is too small). So in this case M and N are transverse if and only if they are disjoint, and the General Position Lemma has a marvelous consequence: We can budge M a bit so that it is disjoint from N.

Proof. Consider $g : M \times N \to \mathbb{R}^n$ defined by $g(x, y) = x - y$. Pick any $a \in \mathbb{R}^n$ that is a regular value of g. We claim $(M + a) \pitchfork N$. If not, then there would be an $x \in M, y \in N$ such that $y = x + a$ and $T_xM + T_yN \neq \mathbb{R}^n$ ($\mathbb{R}^n$ is the tangent space $T_y\mathbb{R}^n$). Then $g(x, y) = a$ and $dg_{(x,y)}(T_{(x,y)}M \times N) = T_xM + T_yN$, which since $T_xM + T_yN \neq \mathbb{R}^n$ contradicts the fact that a is a regular value. Thus, it must be that $(M + a) \pitchfork N$. □

References

[1] V. Guillemin and A. Pollack, *Differential Topology.* Prentice Hall, 1974.
[2] M. Hirsch, *Differential Topology.* Springer, GTM No. 33, 1976.
[3] J. Milnor, *Topology from a Differential Viewpoint.* University of Virginia Press, 1965.
[4] J. Milnor, *Morse Theory.* Princeton, N.J.: Princeton University Press, No. 51, 1963.
[5] M. Spivak, *Calculus on Manifolds.* Addison-Wesley, 1965.

E-mail address: alexonlinemail@gmail.com