TMA 4190 Introduction to Topology
Lecturer: Gereon Quick
Lecture 14[1]

14. Sard's Theorem and Morse functions

Now we are going to shift perspectives and ask:

Given a map f which does not have a property P. Is it possible to **bump f a little bit** such that it gets **property P**?

If this is possible for every map, P is a particularly nice property:

Generic properties

A property P of maps is called **generic** if, for any f_0, there is a homotopy F for f_0 and an $\epsilon > 0$ such that f_t has property P for all $t \in (0, \epsilon)$.

If we look back at the images we used to illustrate stable and unstable properties, we see that non-transversal intersections are rather the exception than the norm. Now we have a way to give this feeling a precise meaning: **Transversality is generic**.

We are not going to prove this statement for the moment, but content ourselves with looking at an important special case. Recall that transversality is a generalization of regularity:

$$f \pitchfork \{y\} \iff y \text{ is a regular value of } f.$$

An analog, though not equivalent, version of the above question is now: Given a smooth map $f\colon X \to Y$ and a critical value y. Is it possible to **bump y a little bit** such that it gets **regular**?

The answer is yes and is the content of a famous theorem:

Sard's Theorem

If $f\colon X \to Y$ is any smooth map of manifolds, then almost every point in Y is a regular value of f.

[1]Following the books of Guillemin and Pollack: Differential Topology, and Milnor: Topology from the differentiable viewpoint.

To say that "almost every point" is a regular value of f sounds sloppy, but is a well-defined term in measure theory. It means by definition that the complement of regular values in Y has **measure zero**. Since the complement of the regular values are the critical values, Sard's theorem says that the set of critical values of a smooth map of manifolds has measure zero.

Sard's Theorem for manifolds follows from Sard's Theorem in Calculus. We are not going to prove either of them, since the required techniques are not so interesting for this course.

Measure zero in a measure zero box

A rectangular solid in $\mathbb{R}^n$ is just a cartesian product of n intervals in $\mathbb{R}^n$, and its volume is the product of the lengths of the n intervals. An arbitrary set A in $\mathbb{R}^n$ is said to have **(Lebesgue) measure zero** if, for every $\epsilon > 0$, there exists a **countable** collection $\{S_1, S_2, \ldots\}$ of rectangular solids in $\mathbb{R}^n$, such that A is contained in the union of the S_i, and

$$\sum_{i=1}^{\infty} \operatorname{vol}(S_i) < \epsilon.$$

Then in a manifold X, an arbitrary subset $C \subset X$ has **measure zero** if, for every local parametrization ϕ of X, the preimage $\phi^{-1}(C)$ has measure zero in Euclidean space.
(Note that measure and volume depend on the ambient space.)
An example of a **measure zero** subset is given by the set of **rational numbers** in $\mathbb{R}$. Hence for measure theorists, "almost every" real number is irrational. This example illustrates that something that happens almost never, can still happen often enough to be noticed.

We learn from the previous box: By definition, no nonempty rectangular solid in $\mathbb{R}^n$ has measure zero. Hence it cannot be contained in a set of measure zero. Now, every nonempty open subset of $\mathbb{R}^n$ contains some nonempty rectangular solid. Thus, no nonempty open subset of $\mathbb{R}^n$ has measure zero. Hence, no nonempty open subset of a manifold Y has measure zero. In other words, no set of measure zero in a manifold Y can contain a nonempty open subset of Y.

In view of Sard's Theorem, this tells us that the set of critical values of a smooth map $f\colon X \to Y$ cannot contain any nonempty open subset of Y. Thus, its complement, the set of regular values, must have a **nonempty intersection with every nonempty open subset** of Y. A subset of a topological space with this property, i.e. having a nonempty intersection with every nonempty open subset, is called **dense**.

Hence we can rephrase Sard's Theorem in more topological terms by:

Sard's Theorem in dense form

The set of regular values of any smooth map $f\colon X \to Y$ is **dense** in Y. More generally, if $f_i\colon X_i \to Y$ are any countable number of smooth maps, then the points of Y that are simultaneously regular values for all of the f_i, are dense.

Morse Functions

Before we study a very interesting application of Sard's Theorem, we recall some terminology (we have already used these terms in the proof of the Fundamental Theorem of Algebra, but did not make a fuzz about it).

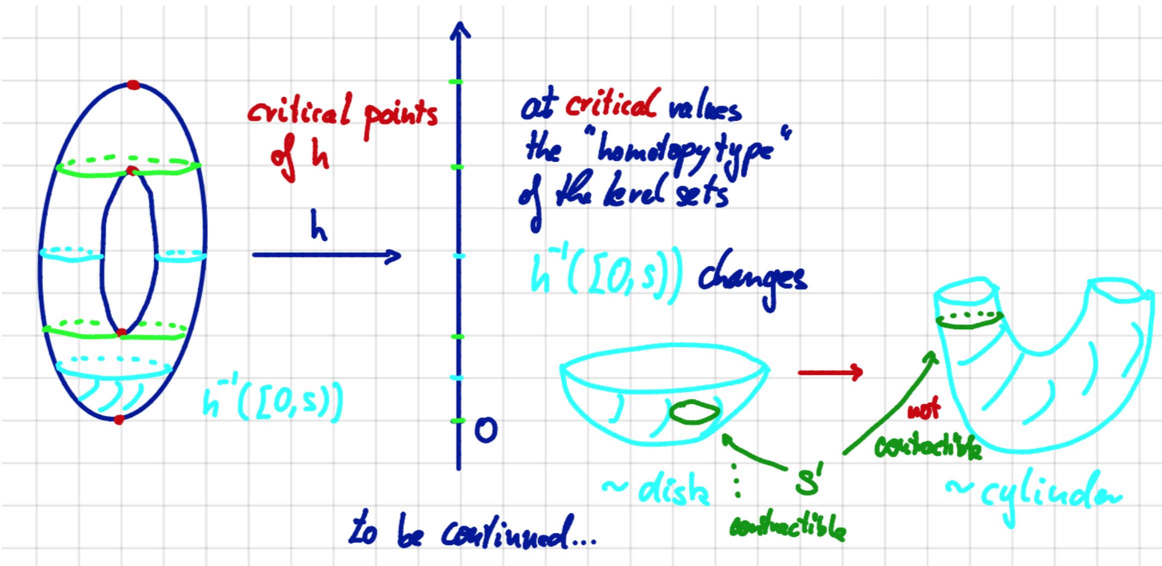
If $f\colon X \to Y$ is a smooth map, a regular value of f is a point $y \in Y$ such that df_x is surjective for every $x \in X$ with $f(x) = y$. We call such an $x \in X$ also a regular point of f. Note that this is the same as to say that f is regular at x. Hence y is a **regular value** of f if **every** $x \in f^{-1}(y)$ is a **regular point**.

On the other hand, if df_x is not surjective, we call x a critical point of f. Hence $y \in Y$ is a **critical value** if **at least one** of the points $x \in f^{-1}(y)$ is a **critical point**.

We understand the local behavior of smooth maps at regular points by the Local Submersion Theorem (up to diffeomorphism look like the canonical submersion). But what about the local behavior at critical points? In fact, it is often at critical points that the interesting stuff happens. It is often at critical points that the topology of a manifold can change.

For example, for a smooth map $f\colon X \to \mathbb{R}$, if X is **compact**, then we know that f must have a **maximum** and a **minimum**. At a point $x \in X$ where $f(x)$ is either a maximal or a minimal value, f cannot change in any direction in X. In other words, the derivative df_x must vanish (recall $df_x(h)$ is a measure for the change of f in direction h). Hence x is a critical point in our terminology.

A standard example is given by the height function on a torus:



So let us stick to smooth functions, i.e. smooth maps to $\mathbb{R}$. We want to understand how critical points look like locally. Let us look a smooth function $f\colon \mathbb{R}^k \to \mathbb{R}$. Locally around a point $c \in X$, we can describe f by

$$f(x) = f(c) + \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}(c) \cdot (x_i - c_i) + \frac{1}{2} \sum_{i,j=1}^{k} \frac{\partial^2 f}{\partial x_i \partial x_j}(c) \cdot (x_i - c_i)(x_j - c_j) + o(|x|^3).$$

If c is a critical point, then by definition

$$df_c = (\partial f/\partial x_1(c), \ldots, \partial f/\partial x_k(c)) = 0$$

(otherwise df_c was surjective as a linear map $\mathbb{R}^k \to \mathbb{R}$). Hence the best possible measure for the local behavior of f at c is the Hessian matrix of the second partial derivatives. Critical points where the Hessian matrix is invertible is the best we can hope for.

Nondegenerate critical points and Morse functions

For a smooth function $f\colon \mathbb{R}^k \to \mathbb{R}$, a point $c \in \mathbb{R}^k$ where df_c vanishes, but the Hessian matrix $H(f)_c = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(c)\right)$ is invertible at c, is called a **nondegenerate critical point**.
A smooth function $f\colon \mathbb{R}^k \to \mathbb{R}$ for which all critical points are nondegenerate is called a **Morse function**.

Nondegenerate critical points are much easier to study than arbitrary critical points, since they are **isolated from the other critical points**, i.e. there is an open neighborhood which does not contain any other critical points. Hence Morse functions are easier to understand than arbitrary smooth functions.

To see that nondegenerate critical points are isolated, we define a map $g\colon \mathbb{R}^k \to \mathbb{R}^k$ by the formula

$$g = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_k}\right). \tag{1}$$

$$\text{Then } df_x = 0 \iff g(x) = 0.$$

Moreover, the matrix representing the derivative dg_x is the Hessian of f at x. So if x is nondegenerate, then not only is $g(x) = 0$, but **g maps a neighborhood of x diffeomorphically onto a neighborhood of 0** as well. In particular, g is injective in that neighborhood of x. Thus g can be zero at no other points in this neighborhood, and f has no other critical point in this neighborhood.

Another reason to be interested in Morse functions is the fact that there are a lot of them.

Morse functions on $\mathbb{R}^k$ are generic

Let $f\colon U \to \mathbb{R}$ be a smooth function defined on some open $U \subseteq \mathbb{R}^k$ and $a \in \mathbb{R}^k$, define

$$f_a(x) = f(x) + a \cdot x.$$

Then, **for almost all** $a \in \mathbb{R}^k$, f_a is a Morse function.

Proof: We us again the function g from (1).The derivative of f_a at a point $p \in U$ then satisfies

$$(df_a)_p = \left(\frac{\partial f_a}{\partial x_1}(p), \dots, \frac{\partial f_a}{\partial x_k}(p)\right) = g(p) + a.$$

Hence the **critical points of f_a** are the points $p \in U$ with $g(p)+a = 0$. Moreover, the **Hessian of f_a** at p is the matrix dg_p, i.e.

$$H(f_a)_p = H(f)_p = dg_p.$$

Hence

$$\begin{aligned} f_a \text{ is Morse} &\iff \det(H(f_a)_p) \neq 0 \text{ at all critical points } p \\ &\iff \det(dg_p) \neq 0 \text{ at all } p \text{ with } g(p) + a = 0 \\ &\iff -a \text{ is a regular value of } g. \end{aligned}$$

By **Sard's Theorem**, $-a$ is a regular value of g for almost all $a \in \mathbb{R}^k$. Therefore almost every f_a is a Morse function. **QED**

Now we would like to transport the concept of nondegenerate critical points to **manifolds**. So let X be a smooth manifold. Suppose that $f\colon X \to \mathbb{R}$ has a **critical point at x** and that $\phi\colon U \to X$ is a local parametrization with $\phi(0) = x$. Then

$$d(f \circ \phi)_0 = df_x \circ d\phi_0$$

and hence 0 is a critical point for the function $f \circ \phi$. We call x a **nondegenerate critical point for f** if 0 is a nondegenerate critical point for $f \circ \phi$.

Independence of choice

Since we made a choice of a local parametrization for this definition, we need to make sure that the criterion is independent of the choice.
So let $\psi\colon V \to X$ be another local parametrization with $\psi(0) = x$. We define $\theta := \psi^{-1} \circ \phi\colon U \to V$. Since θ is a diffeomorphism, the critical points of $f \circ \phi$ and $f \circ \psi \circ \theta$ are the same.
Assuming that x is a critical point of f, i.e. $df_x = 0$, the chain rule implies for the two Hessian matrices at 0:

$$H(f \circ \phi)_0 = (d\theta_0)^t H(f \circ \psi)_0 d\theta_0.$$

Since $d\theta_0$ is invertible, we see

$$H(f \circ \phi)_0 \text{ is invertible} \iff H(f \circ \psi)_0 \text{ is invertible.}$$

An important result on Morse functions is that they can be described in some sort of canonical form. It extends our understanding of the local behavior of smooth maps.

Morse Lemma

Let X be a smooth manifold and $f\colon X \to \mathbb{R}$. Suppose that $a \in X$ is a **nondegenerate critical point** of f. Then there is a local parametrization $\phi\colon U \to X$ with $\phi(0) = a$ and local coordinate functions $\phi^{-1} = (x_1, \ldots, x_k)$ around a such that

$$f(x) = f(a) + \sum_{ij} h_{ij} x_i x_j$$

for all $x \in \phi(U)$ where the h_{ij} are the entries of the Hessian of f at a:

$$h_{ij} = (H(f \circ \phi))_0)_{ij} = \frac{\partial^2 (f \circ \phi)}{\partial x_i \partial x_j}(0).$$

(Note that the h_{ij} depend on the chosen coordinate system.)

We are not going to discuss the proof of this classical result. However, we are going to show that it applies to many functions.

In fact, we can generalize the fact that "almost all" functions are Morse to the level of **manifolds**: Suppose $X \subset \mathbb{R}^N$, and let $x_1, \ldots, x_N \in \mathbb{R}^N$ be the usual coordinate functions on R^N. If $f\colon X \to \mathbb{R}$ is a smooth function on X and $a = (a_1, \ldots, a_N)$ is an N-tuple of numbers, we define again a new function $f_a\colon X \to \mathbb{R}$ by

$$f_a := f + a_1 x_1 + \cdots + a_N x_N.$$

Morse functions on any manifold are generic

For every smooth function $f\colon X \to \mathbb{R}$ and for **almost every** $a \in \mathbb{R}^N$, f_a is a Morse function on X, i.e. all its critical points are nondegenerate.

Proof: We would like to use the above result for $U \subset \mathbb{R}^k$ open. Since $X \subset \mathbb{R}^N$ is in general **not open** (in fact, it is never open if $\dim X < N$), the **strategy is to cover X** by open subsets and then try to lift the k-dimensional result to open sets in $\mathbb{R}^N$.

So let x be any point in X. First we are going to choose a suitable local coordinate system around x. Let $v_1, \ldots, v_k \in \mathbb{R}^N$ be a basis of $T_x(X)$ (for $k = \dim X$). Then the matrix $[v_1 \cdots v_k]$, having the v_i's as columns, has rank k. Hence it has k linearly independent rows, say $i_1, \ldots, i_k$. Let $\pi\colon \mathbb{R}^N \to \mathbb{R}^k$ be projection defined by $(x_1, \ldots, x_N) \mapsto (x_{i_1}, \ldots, x_{i_k})$ where the $x_1, \ldots, x_N$ denote the standard coordinates on $\mathbb{R}^N$. Then

$$(d\pi_x)_{|T_x(X)}\colon T_x(X) \to \mathbb{R}^k \text{ is an isomorphism}$$

by construction. Hence, by the Inverse Function Theorem,

$$\pi_{|X}\colon X \to \mathbb{R}^k \text{ is a local diffeomorphism.}$$

Hence we can take the k-tuple of functions $(x_{i_1}, \ldots, x_{i_k})\colon X \to \mathbb{R}^k$ to define a **local coordinate system around x**.

Therefore we can cover X with open subsets $U_\alpha \subseteq \mathbb{R}^N$ such that on each U_α **some k-tuple** of the functions $x_1, \dots, x_N$ on $\mathbb{R}^N$ form a coordinate system. Moreover, it is always possible to choose a **countable** subfamily of the U_α's. Hence we may assume there are only **countably many U_α**.

Let $S \subset \mathbb{R}^N$ be the subset of a such that f_a is **not Morse**. Since the countable union of sets with measure zero has measure zero, it suffices to show that **for each U_α** the set S_α of a's such that $f_a \colon U_\alpha \to \mathbb{R}$ is **not Morse**, has **measure zero**.

So let us look at one of the U_α's. We want to show that S_α has measure zero in $\mathbb{R}^N$.

For simplicity, assume $x_1, \dots, x_k$ form a coordinate system around x on U_α. We can write any $a \in \mathbb{R}^N$ as $a = (b,c)$, where b denotes the **first k** coordinates and c denotes the **last $N-k$** coordinates. Around a given point x, we can thus write

$$f_a(x) = f(x) + c \cdot (x_{k+1}, \dots, x_N) + b \cdot (x_1, \dots, x_k).$$

The function $x \mapsto f(x) + c \cdot (x_{k+1}, \dots, x_N)$ is smooth. Hence we can apply our previous result on genericity of Morse functions on open subsets in $\mathbb{R}^k$ to this function and get that **f_a is a Morse function for almost every $b \in \mathbb{R}^k$**.

Thus, for a fixed c, the subset of all $b \in \mathbb{R}^k$ where f_a is not Morse, has measure zero in $\mathbb{R}^k$. Hence $S_\alpha \cap (\mathbb{R}^k \times \{0\})$ has measure zero in $\mathbb{R}^N$. It is a calssical result in Measure Theory, called **Fubini's Theorem**, which then implies that the set S_α of all $a = (b,c)$ where a does not yield a Morse function has measure zero in $\mathbb{R}^N$. Hence f_a is a Morse function for almost every a. **QED**

Finally, we can also show that being a Morse function is a stable property. In order to prove stability, we start with a little lemma:

First Lemma

Let f be a smooth function on an open set $U \subset \mathbb{R}^k$. For each $x \in U$, let $H(f)_x$ be the Hessian matrix of f at x. Then f is a Morse function if and only if

$$(\det(H(f)_x))^2 + \sum_{i=1}^{k} \left(\frac{\partial f}{\partial x_i}(x) \right)^2 > 0 \text{ for all } x \in U. \tag{2}$$

Proof: A point x is regular if $df_x = (\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_k}(x)) \neq 0$, and x is a nondegenerate critical point if $df_x = 0$ and $\det(H(f)_x) \neq 0$. Hence f is Morse if and only if (2) is satisfied. **QED**

Second Lemma

Suppose that f_t is a homotopic family of functions on $\mathbb{R}^k$. If f_0 is a Morse function on some oepn subset $U \subset \mathbb{R}^k$ containing a **compact** set $K \subset \mathbb{R}^k$, then so is every f_t for t sufficiently small.

Proof: We define the map

$$F\colon U \times [0,1] \to \mathbb{R},\ (x,t) \mapsto (\det(H(f_t)_x))^2 + \sum_{i=1}^{k} \left(\frac{\partial f}{\partial x_i}(x) \right)^2.$$

Since f is smooth, F depends smoothly on both variables. By the First Lemma and the assumption, we know $F(x,0) > 0$ for all $x \in U \times \{0\}$. Since $K \subset U$ is compact, F has a minimum on $K \times \{0\}$, i.e. there is a $\delta > 0$ such that $F(x,0) \geq 2\delta$ for all $x \in K$. Since F is continuous, there an open neighborhood $W \subset U \times [0,1]$ containing $K \times \{0\}$ such that $F(x,t) > \delta$ for all $(x,t) \in W$. In fact, we can cover $K \times \{0\}$ by open subsets $W_i \subset U \times [0,1]$ such that $F(x,t) > \delta$ for all $(x,t) \in W_i$. Each such open subset W_i has the form $V_i \times [0,\epsilon_i)$ for some open $V_i \subset U$ and $\epsilon_i > 0$. Since K is compact, **finitely many** such open W_i suffice to cover $K \times \{0\}$. Let ϵ be the **minimum** of the finitely many ϵ_i. Then we have $F(x,t) > \delta$ for all $(x,t) \in K \times [0,\epsilon)$. Since F is continuous, for any fixed $t \in [0,\epsilon)$, there is again an open subset $V\mathbb{R}^k$ containing K such that $F(x,t) > 0$ for all $(x,t) \in V \times \{t\}$. Thus f_t is Morse in a neighborhood of K for all sufficiently small t. **QED**

Finally, we are ready to prove stability of Morse functions.

Stability of Morse functions

Let X be a **compact** smooth manifold, let $f_0\colon X \to \mathbb{R}$ be a smooth function and f_t be a homotopy of f_0. If f_0 is Morse, then there is an $\epsilon > 0$ such that f_t is a Morse function for all $t \in [0,\epsilon)$.

Proof: For $x \in X$, let $\phi_x\colon U_x \to X$ be a local parametrization around x. Then $f_0 \circ \phi_x$ is a Morse function on U. Since $\{0\}$ is a compact subset of U, the Second Lemma above implies that there is an open subet $V_x \subset U_x$ containing $\{0\}$ and an $\epsilon(x) > 0$ such that f_t is Morse on V_x for all $t \in [0,\epsilon(x))$. The images $\phi_x(V_x)$ are open in X and cover X. Since X is compact, finitely many suffice to cover

X, say

$$X = \phi_{x_1}(V_{x_1}) \cup \cdots \cup \phi_{x_n}(V_{x_n})$$

Then we can set $\epsilon :=$ minimum of $\epsilon(x_1), \ldots, \epsilon(x_n)$. Then $f_t \colon X \to \mathbb{R}$ is a Morse function for all $t \in [0, \epsilon)$. **QED**