Лекция 23. Теорема Сарда

Пусть $f\colon \mathbb{R}^m \to \mathbb{R}^n$ — гладкое отображение. Что можно сказать о его слоях $f^{-1}(y) \subset \mathbb{R}^n$ для произвольного $y \in \mathbb{R}^n$?

При такой постановке ответ прост и безрадостен: помимо того, что эти слои обязаны быть замкнутыми подмножествами в $\mathbb{R}^m$, ничего сказать нельзя, так кака *всякое* замкнутое множество является множествуом нулей некоторой бесконечно гладкой функции:

Предложение 23.1. *Для всякого замкнтуого множества $F \subset \mathbb{R}^n$ существует C^∞-функция $\varphi\colon \mathbb{R}^n \to \mathbb{R}$, обладающая тем свойством, что $\varphi(x) \geqslant 0$ для всех $x \in \mathbb{R}^n$ и $\varphi(x) = 0$ тогда и только тогда, когда $x \in F$.*

Доказательство. Начнем с построения так называемых колоколообразных функций (они пригодятся нам и во многих других вопросах). Как известно, функция

$$f(x) = \begin{cases} 0, & x \leqslant 0, \\ e^{-1/x^2}, & x > 0 \end{cases}$$

является C^∞-гладкой на $\mathbb{R}$. Для всякого $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ положим $\psi_0(x) = f(1 - x_1^2 - \ldots - x_n^2)$. Тогда $\psi_0\colon \mathbb{R}^n \to \mathbb{R}$ — функция класса C^∞, принимающая только неотрицательные значения, положителяная всюду на внутренности единичного шара с центром в начале координат и равная нулю в остальных точках. Если $B_r(a)$ — открытый шар радиуса r с центром в точке a, то функция $\psi(x) = \psi_0(r^{-1}(x - a))$ — также неотрицательная функция класса C^∞, положительная на шаре $B_r(a)$ и раная нулюд вне его; будем называть всякую такую функцию колоколообразной функцией для этого шара.

Заметим теперь, что открытое множество $\mathbb{R}^n \setminus F$ представляется в виде счетного объединения $\mathbb{R}^n \setminus F = \bigcup_{j=1}^\infty B_j$ открытых шаров. Для каждого шара B_j можно выьбрать соответствующую ему колоколообразную функцию ψ_j, обладающую следующим свойством:

$$\sup_{\substack{x \in \mathbb{R}^n \\ |\alpha| \leqslant j}} |D^\alpha \psi_j(x)| \leqslant \frac{1}{2^j} \tag{23.1}$$

(в самом деле, поскольку (23.1) задает условия только на конечное число функций, для этого достаточно взять произвольную колоколообразную функцию для B_j и умножить ее на подходящую положительную константу).

Положим теперь $\varphi(x) = \sum_{j=1}^{\infty} \psi_j(x)$. Ввиду (23.1) ряд в правой части равномерно сходится, так что φ — неотрицательная функция, множество нулей которой, очевидно, совпадает с F. Осталось показать, что она бесконечно дифференцируема.

Это доказывается с помощью следующей леммы о функциях одного переменного.

Лемма 23.2. *Пусть $\{f_n\}$ — последовательность функций класса C^1 на интервале $(a;b) \subset \mathbb{R}$; предположим, что последовательность $\{f_n\}$ равномерно сходится к функции f, а последовательность $\{f'_n\}$ равномерно сходится к функции $\tilde{f}$. Тогда f — также функция класса C^1, причем $f' = \tilde{f}$.*

Доказательство леммы. Ввиду равномерной сходимости функции f и $\tilde{f}$ непрерывны. Зафиксируем точку $x_0 \in (a;b)$; по формуле Ньютона–Лейбница достаточно проверить, что

$$f(x) = f(x_0) + \int_{x_0}^{x} \tilde{f}(t)dt$$

для всех $x \in (a;b)$; это же равенство вытекает из того, что

$$\int_{x_0}^{x} \tilde{f}(t)dt = \lim_{n\to\infty} \int_{x_0}^{x} f'_n(t)dt$$

ввиду равномерной сходимости, вкупе с очевидным равенством $f(x_0) = \lim f_n(x_0)$. $\square$

Для доказательства бесконечной дифференцируемости функции φ достаточно установить, что все ее частные производный существуют и непрерывны на $\mathbb{R}^n$; ввиду леммы это доказывается индукцией по порядку производной, так как ввиду (23.1) все ряды вида $\sum D^\alpha \psi_j$ равномерно сходятся. $\square$

(Заметим в скобках, что с аналитическими отображениями все обстоит куда лучше: их слои вовсе не произвольны и вполне поддаются изучению. Мы, однако же, занимаемся анализом гладким.)

Потерпев фиаско в исследовании произвольных слоев, попробуем все же сказать что-нибудь позитивное. Именно, пусть для всех $x \in f^{-1}(y)$ ранг производной $f'(x)$ равен n. Тогда из предложения 20.2 следует, что $f^{-1}(y)$ является гладким подмногообразием в $\mathbb{R}^m$, и это множество не произвольное, а очень хорошее. Таким образом, есть точки, слои над которыми изучать можно. Для таких точек существует специальный термин.

Определение 23.3. Пусть $f\colon U \to \mathbb{R}^n$ — гладкое отображение, где $U \subset \mathbb{R}^m$ — открытое подмножество. Точка $y \in \mathbb{R}^n$ называется *критическим значением отображения* f, если существует точка $x \in f^{-1}(y)$, для которой $\operatorname{rank} f'(x) < n$. При этом точка x, для которой $\operatorname{rank} f'(x) < n$, называется *критической точкой* отображения f.

Точка $y \in \mathbb{R}^n$ называется *регулярным значением отображения* f, если она не является критическим значением.

Два замечания к этому определению.

1. Если $y \notin f(U)$, то точка y автоматически является регулярным значением (хотя и не является значением отображения f!).

2. Определение не исключает случая $m < n$. В этом случае всякое значение отображения f является критическим.

В силу обсуждения перед определнием 23.3, прообраз всякого регулярного значения является (если он непуст) гладким подмногообразием в $\mathbb{R}^m$. Основной результат этой лекции утверждает, что если отображение f достаточно гладко, то регулярных значений у него очень много.

Теорема 23.4 (теорема Сарда). *Пусть $U \subset \mathbb{R}^m$ — открытое подмножество, и пусть $f\colon U \to \mathbb{R}^n$ — отображение класса C^r, где $r > \max(m-n, 0)$. Тогда множество критических значений отображения f имеет меру нуль.*

Мы докажем теорему Сарда при упрощающем предположении, считая, что f бесконечно дифференцируемо.

Начнем с простого случая $m \leqslant n$; для него мы докажем теорему Сарда в полном объеме.

Предложение 23.5. *Теорема 23.4 верна при $m \leqslant n$.*

Доказательство. При $m \leqslant n$ теорема гласит, что множество критических значений отображения f имеет меру нуль, как только f является C^1-гладким.

Пусть сначала $m = n$. Достаточно доказать, что меру нуль имеет множество критических значений ограничения $f|_Q$ для всякого замкнутого куба $Q \subset U$. Зафиксируем такой куб; пусть $f = (f_1, \ldots, f_n)$, где $f_i\colon U \to \mathbb{R}$ — функции класса C^1 на U. Из формулы Тейлора, примененной ко всем f_i, следует, что существует такая возрастающая функция $\eta\colon [0;\tau) \to [0;+\infty)$ (для некоторого $\tau > 0$), обладающая свойством

$\lim_{t\to 0}\eta(t)=0$, что

$$\|f(x+h)-f(x)-f'(x)h\|\leqslant \eta(\|h\|)\|h\|, \tag{23.2}$$

как только $x\in Q$ и $\|h\|<\tau$ (под нормой здесь и ниже понимаем обычную евклидову норму). Кроме того, из той же формулы Тейлора вытекает существование константы C с тем свойством, что $\|f(x)-f(x')\|\leqslant C\|x-x'\|$, как только $x,x'\in Q$ и $\|x-x'\|$ достаточно мало; так как куб — выпуклая фигура, отсюда следует, что $\|f(x)-f(x')\|\leqslant C\|x-x'\|$ для любых $x,x'\in Q$ (соедините x и x' отрезком).

Пусть теперь $R=\{x\in Q\mid \operatorname{rank} f'(x)<n\}$.

Лемма 23.6. *Пусть $E\subset U$ — подмножество, диаметр которого не превосходит $\varepsilon>0$. Если $E\cap R\neq\varnothing$, то $\bar\mu(f(E))\leqslant 2^nC^{n-1}\varepsilon^n\eta(\varepsilon)$ (в левой части стоит внешняя мера).*

Доказательство леммы. Пусть $x\in E\cap R$; обозначим через L аффинное подпространство в $\mathbb{R}^n$, проходящее через $f(x)$ и параллельное векторному подпространству $f'(x)(\mathbb{R}^n)\subset\mathbb{R}^n$. По условию $\dim L\leqslant n-1$; если неравенство строгое, дополним L произвольным образхом до гиперплоскости, содержащей $f(x)$, и будем далее обозначать через L эту гиперплоскость. Из неравенства (23.2) следует, что для всякой $x'\in E$ точка $f(x')$ отстоит от гиперплоскости L не более чем на $\varepsilon\eta(\varepsilon)$; кроме того, из неравенства $\|f(x)-f(x')\|\leqslant C\|x-x'\|$ следует, что для всякого $x'\in E$ имеем $\|f(x)-f(x')\|\leqslant C\varepsilon$; стало быть, проекция $f(x')$ на гиперплоскость L также отстоит от $f(x)$ не более чем на $C\varepsilon$. Тем самым $f(E)$ содержится в прямоугольном параллелепипеде с центром в $f(x)$, у которого $n-1$ ребер параллельны L и имеет длину $2C\varepsilon$, а последнее ребро, ортогональное к L, имеет длину $2\varepsilon\eta(\varepsilon)$. Внешняя мера $\mu(f(E))$ не превосходит объема этого параллелепипеда. □

Теперь для произвольного натурального $k>1$ разобьем куб Q на k^n кубиков с ребром d/k, где d — ребро куба Q. Если какой-то из этих кубиков имеет непустое пересечение с R, то, согласно лемме, мера его образа не превосходит $2^nC^{n-1}(d\sqrt{n}/k)^n\eta(d\sqrt{n}/k)$. Поскольку общее количество таких кубиков не превосходит k^n, имеем

$$\mu(f(R))\leqslant \underbrace{2^nC^{n-1}(d\sqrt{n})^n}_{\text{const}}\cdot\eta(d\sqrt{n}/k);$$

так как правая часть стремится к нулю при $k\to\infty$, имеем $\mu(f(R))=0$, и теорема Сарда для случая $m=n$ доказана.

Если $m < n$, то теорема равносильна утверждению, что $\mu(f(U)) = 0$; это немедленно следует из уже доказанного и того факта, что $f(U)$ совпадает с множеством критических значений отображения $f \circ \pi \colon U \times \mathbb{R}^{n-m} \to \mathbb{R}^n$, где π — проекция на первый сомножитель. $\square$

Предложение 23.7. *Теорема 23.4 верна, если $m > n$ и f бесконечно гладка.*

Доказательство. Начнем с простой леммы, с помощью которой мы установим, что некоторые интересующие нас множества измеримы.

Лемма 23.8. *Пусть $E \subset \mathbb{R}^m$ — подмножество, представимое в виде разности двух замкнутых множеств (такие подмножества называют «локально замкнутыми») и $f \colon E \to X$ — непрерывное отображение в хаусдорфово пространство. Тогда $f(E)$ — борелевское подмножество в X.*

Доказательство леммы. Можно считать, что $E = F \setminus G$, где F и G замкнуты и $G \subset F$. Поскольку F, очевидно, представимо в виде счетного объединения компактных подмножеств, можно считать, что F и G компактны. Далее, имеем $G = \bigcap_{i=1}^{\infty} U_i$, где все U_i открыты (см. доказательство теоремы 16.15); Поэтому $E = \bigcup_{i=1}^{\infty}(F \setminus U_i)$, т. е. E — объединение счетного семейства компактных множеств; следовательно, этим же свойством обладает и $f(E)$, а поскольку компактные подмножества в X замкнуты, множество $f(E)$ является борелевским. $\square$

Переходя к доказательству предложения, проведем индукцию по m (базу индукции обеспечивает предложение 23.5).

Пусть $f = (f_1, \ldots, f_n)$, и пусть $R = R_0 = \{x \in U \mid \operatorname{rank} f'(x) < n\}$. Нам надо доказать, что $\mu(f(R)) = 0$. Обозначим через $R_k \subset R$ подмножество, состоящее из тех точек $x \in R$, в которых все частные производные порядка $\leqslant k$ от всех функций f_i обращаются в нуль. Достаточно доказать следующие три утверждения:

(1) $\mu(f(R \setminus R_1)) = 0$; (2) $\mu(f(R_k \setminus R_{k+1})) = 0$ при всех $k > 0$; (3) $\mu(f(R_k)) = 0$, если $k \geqslant m/n$. Сразу же заметим, что все R_k замкнуты, так что ввиду леммы 23.8 все множества вида $f(R_j)$ и $f(R_j \setminus R_{j+1})$ измеримы.

Докажем утверждение (1). Нам достаточно проверить, что у всякой точки $x \in R \setminus R_1$ существеут такая окрестность V, что $\mu(f(R \cap V)) = 0$. Пусть теперь $x \in R \setminus R_1$; тогда $\frac{\partial f_i}{\partial x_j}(x) \neq 0$ для каких-то i и j. Рассмотрим отображение $g\colon U \to \mathbb{R}^m$, действующее по формуле

$$(x_1, \ldots, x_m) \mapsto (x_1, \ldots, x_{j-1}, f_i(x_1, \ldots, x_m), x_{j+1}, \ldots, x_m);$$

согласно теореме об обратной функции, существует такая окрестность $V \ni x$, что ограничение $g|_V$ является C^∞-диффеоморфизмом на свой образ V_1. Критические значения отображения $\tilde{f} = f|_V \circ g|_{V_1}{}^{-1}$ — такие же, как у отображения $f|_V$. Теперь заметим, что $\tilde{f}$ действет по правилу

$$(x_1, \ldots, x_m) = x \mapsto (\tilde{f}_1(x), \ldots, \tilde{f}_{i-1}(x), x_j, \tilde{f}_{i+1}(x), \ldots, f_n(x). \tag{23.3}$$

Поэтому $\tilde{f}$ переводит пересечение $H_c = V_1 \cap \{x_j = c\}$ в гиперплоскость $L_c = \{(y_1, \ldots, y_n) \in \mathbb{R}^n \mid y_i = c\}$, причем из (23.3) явствует, что $y \in L_c$ является критическим значением для $\tilde{f}$ тогда и только тогда, когда y является критическим значением для отображения $\tilde{f}|_{H_c}\colon H_c \to L_c$. Согласно предположению индукции, множество критических значений каждого из $\tilde{f}|_{H_c}$ имеет меру нуль в L_c; стало быть, по теореме Фубини множество критических значений $\tilde{f}$ имеет меру нуль, а это множество совпадает, как мы уже отмечали, с $f(R \cap V)$.

Докажем утверждение (2). По-прежнему достаточно проверить, что у всякой точки $x \in R_k \setminus R_{k+1}$ существеут такая окрестность V, что $\mu(f(R_k \cap V)) = 0$. Пусть теперь $x \in R \setminus R_1$; тогда существуют мультииндекс α, $|\alpha| = k$, и индексы i и j, для которых $(\partial D_\alpha f_i / \partial x_j)(x) \neq 0$. Рассмотрим отображение $g\colon U \to \mathbb{R}^m$, действующее по формуле

$$(x_1, \ldots, x_m) \mapsto (x_1, \ldots, x_{j-1}, D_\alpha f_i(x_1, \ldots, x_m), x_{j+1}, \ldots, x_m);$$

согласно теореме об обратной функции, существует такая окрестность $V \ni x$, что ограничение $g|_V$ является C^∞-диффеоморфизмом на свой образ V_1. Критические значения отображения $\tilde{f} = f|_V \circ g|_{V_1}{}^{-1}$ — такие же, как у отображения $f|_V$. Заметим теперь, что условие обращение в нуль в точке x всех частных производных порядка $\leqslant k$ у отображения f класса C^k (и тем более класса C^∞) равносильно, ввиду формулы Тейлора, тому, что $\|f(x+h) - f(x)\| = o(\|h\|^k)$. Это последнее условие, очевидно, сохраняется при композиции с диффеоморфизмом. Поэтому у отображения $\tilde{f} = f|_V \circ g|_{V_1}{}^{-1}$ в каждой точке $x' \in g(R_k \cap V)$ все частные производные порядка $\leqslant k$ обращаются в нуль. С другой стороны, из определения множества R_k следует, что $f(R_k \cap V) \subset V_1 \cap \{x_j = 0\}$. Стало быть, множество $g(R_k \cap V)$ содержится в множестве критических точек отображения $\tilde{f}|_{V_1 \cap \{x_j = 0\}}$. По предположению индукции имеем $\mu(\tilde{f}(g(R_k \cap V))) = 0$, так что $\mu(f(R_k \cap V)) = 0$.

Докажем утверждение (3). Будем действовать так же, как в доказательстве предложения 23.5. Именно, пусть $k \geqslant m/n$. Достаточно показать, что $\mu(f(R_k \cap Q)) = 0$ для всякого замкнутого куба $Q \subset U$. Из формулы Тейлора следует, что существует такая возрастающая функция $\eta\colon [0;\tau) \to [0;+\infty)$ (для некоторого $\tau > 0$), обладающая свойством $\lim_{t\to 0} \eta(t) = 0$, что

$$\|f(x+h) - f(x)\| \leqslant \eta(\|h\|)\|h\|^k, \tag{23.4}$$

как только $x \in Q \cap R_k$ и $\|h\| < \tau$ (под нормой здесь и ниже понимаем обычную евклидову норму).

Следовательно, для всех достаточно малых $\varepsilon > 0$ верно такое утверждение: если $E \subset Q$ — множество диаметра $\leqslant \varepsilon$, для которого $E \cap R_k \neq \varnothing$, то $f(E)$ содержится в кубе с ребром $2\eta(\varepsilon)\varepsilon^k$. Если теперь разбить куб Q на r^m кубиков с ребром d/r, где d — ребро куба Q, то мера образа такого кубика, пересекающегося с R_k, не превосходит $2^n\eta(d\sqrt{m}/r)(d\sqrt{m})^{kn}/r^{kn}$. Поскольку таких кубиков не более чем r^m штук, получаем, что

$$\mu(f(R_k)) \leqslant \underbrace{2^n(d\sqrt{m})^{kn}}_{\text{const}} \cdot \eta(d\sqrt{m}/r) r^{m-kn}.$$

Поскольку $m - kn \leqslant 0$, то устремляя r к бесконечности, получаем, что $\mu(f(R_k)) = 0$. $\square$

Стоит разобраться, почему приведенное рассуждение не дает доказательства теоремы Сарда в полной общности. Легко видеть, что доказательство утверждения (1) ($\mu(f(R \setminus R_1)) = 0$) проходит и в случае, когда f — всего лишь отображение класса C^s, где $s > m - n$, а также что по-прежнему верно утверждение (3) ($\mu(f(R_s)) = 0$, так как $m - n + 1 \geqslant m/n$ при m/n). Трудности возникают только с утверждением (2), гласящим, что $\mu(f(R_k \setminus R_{k+1})) = 0$: если k велико, то диффеоморфизм, «выпрямляющий» гиперповерхность $D^\alpha f_i = 0$, имеет слишком низкий класс гладкости, вследствие чего не удается опереться на теорему Сарда для меньшей размерности.