Der Satz von Sard

Jonathan Gantner
0725250

5. Oktober 2010

Inhaltsverzeichnis

Kapitel 1

Definitionen

1.1 Differenzierbare Mannigfaltigkeiten

Gegenstand dieser Arbeit ist der Satz von Sard - um ihn zu verstehen benötigen wir einige Definitionen und Ergebnisse aus der Differentialgeometrie.

Definition 1.1 (Mannigfaltigkeit)**.** *Eine n-dimensionale Mannigfaltigkeit M ist ein topologischer Hausdorff-Raum mit abzählbarer Basis der Topologie, der lokal homöomorph zum $\mathbb{R}^n$ ist — das heißt zu jedem Punkt $p \in M$ gibt es eine offene Umgebung U von p und einen Homöomorphismus $h : U \to U'$, sodass $U' \subset \mathbb{R}^n$ und offen ist.*
So ein Homöomorphismus von einer offenen Menge $U \subset M$ auf eine offene Menge $U' \subset \mathbb{R}^n$ heißt Karte von M, U das zugehörige Kartengebiet.
Eine Menge von Karten $\{h_i \mid i \in I\}$ mit $\bigcup_{i\in I} U_i = M$ heißt Atlas von M.

Zu zwei verschiedenen Karten (U_i, h_i) und (U_j, h_j) lässt sich auf dem Durchschnitt $U_i \cap U_j$ der Kartengebiete ein *Kartenwechsel* $h_{ij} = h_j \circ h_i^{-1}$ von $h_i(U_i \cap U_j) \to h_j(U_i \cap U_j)$ definieren. Dieser ist als Zusammensetzung von zwei Homöomorphismen selbst wieder ein Homöomorphismus zwischen zwei offenen Teilmengen des $\mathbb{R}^n$.

Definition 1.2 (differenzierbare Mannigfaltigkeit)**.** *Ein Atlas einer Mannigfaltigkeit heißt differenzierbar, wenn alle seine Kartenwechsel differenzierbar sind.*
Eine differenzierbare Struktur ist ein bezüglich der Hinzunahme weiterer Karten maximaler differenzierbarer Atlas.
Ein Paar $(M, \mathfrak{F})$ bestehend aus einer Mannigfaltigkeit M und einer differenzierbaren Struktur $\mathfrak{F}$ heißt differenzierbare Mannigfaltigkeit.

Eine differenzierbare Struktur auf einer Mannigfaltigkeit ist also ein Atlas, dem keine weiteren Karten mehr hinzugefügt werden können, ohne dass Kartenwechsel entstehen, die nicht mehr differenzierbar sind. Es reicht deshalb einen weitaus kleineren Atlas anzugeben, um eine differenzierbare Struktur vollständig zu definieren. Offensichtlich gilt nämlich:

$$h_{ii} = \mathrm{Id}, \qquad h_{jk} \circ h_{ij} = h_{ik} \qquad \text{und damit auch} \qquad h_{ij}^{-1} = h_{ji}. \tag{1.1}$$

Hat man nun einen differenzierbaren Atlas $\mathfrak{A}$ gegeben, so sei $\mathfrak{D}(\mathfrak{A})$ die Menge aller Karten, für die jeder Kartenwechsel mit einer Karte aus $\mathfrak{A}$ differenzierbar ist. Dann ist $\mathfrak{D}(\mathfrak{A})$ wieder differenzierbar, denn jeder Kartenwechsel h_{ik} kann zumindest lokal dargestellt werden als Zusammensetzung $h_{ik} = h_{jk} \circ h_{ij}$ mit $h_j \in \mathfrak{A}$ — und offensichtlich ist $\mathfrak{D}(\mathfrak{A})$ auch maximal und kann nicht durch zusätzliche Karten erweitert werden. Die differenzierbare Struktur $\mathfrak{D}(\mathfrak{A})$ wird also durch $\mathfrak{A}$ eindeutig festgelegt.

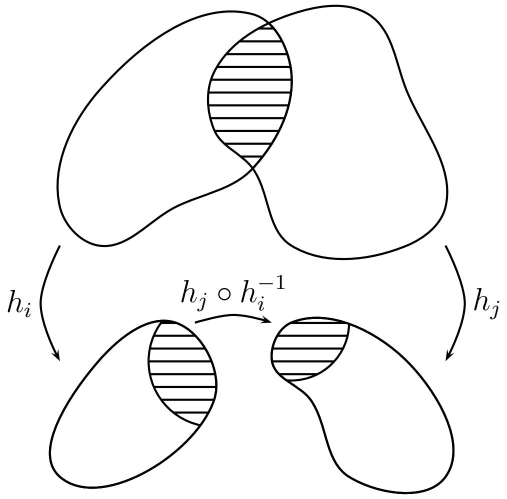

Abb. 1.1: Karten einer Mannigfaltigkeit

Außerdem sind wegen (1.1) die Inversen von Kartenwechseln eines differenzierbaren Atlas wieder Kartenwechsel und somit differenzierbar — Kartenwechsel eines differenzierbaren Atlas sind also Diffeomorphismen. Umgekehrt kann auch jeder Diffeomorphismus zwischen entsprechenden Teilmengen des $\mathbb{R}^n$ als Kartenwechsel interpretiert werden. Ist nämlich (U, h) eine Karte und f ein Diffeomorphismus $h(U) \to \widetilde{U} \subset \mathbb{R}^n$, dann ist $\tilde{h} = f \circ h$ ein Homöomorphismus $U \to \widetilde{U}$ — also eine Karte — und f bzw. f^{-1} sind genau die Kartenwechsel zwischen h und $\tilde{h}$.

Beispiele 1.3.

1. Jede offene Teilmenge des $\mathbb{R}^n$ ist offensichtlich eine n-dimensionale differenzierbare Mannigfaltigkeit. Die differenzierbare Struktur darauf wird durch den Atlas $\{\mathrm{Id}_U\}$ definiert, der nur eine einzige Karte enthält.

2. Die n-dimensionale Sphäre $S^n = \{x \in \mathbb{R}^{n+1} | \sqrt{x_1^2 + x_2^2 + \ldots + x_{n+1}^2} = 1\}$ ist eine n-dimensionale Mannigfaltigkeit. Auch auf ihr lässt sich durch folgende Karten eine differenzierbare Struktur definieren:

$$h_i^j \begin{cases} U_{ij} = \{x \in S^n |\ (-1)^j x_i > 0\} & \to \quad U_1^n(0) = \{x \in \mathbb{R}^n |\ \|x\|_2 < 1\} \\ \qquad (x_1, x_2, \ldots, x_{n+1}) & \mapsto \quad (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n+1}) \end{cases}$$

   wobei $i \in \{1, \ldots, n+1\}$ und $j \in \{1, 2\}$.

   Die Abbildungen h_i^j streichen also einfach die i-te Koordinate. Umgekehrt fügt die Inverse $(h_i^j)^{-1}$ einfach an der i-ten Stelle eine zusätzliche Koordinate mit dem Wert

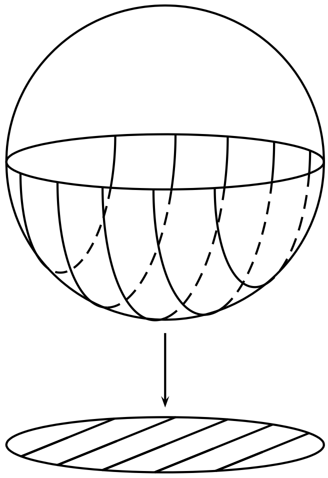

Abb. 1.2: Koordinatisierung der n-dimensionalen Sphäre

$(-1)^j(1-\sum_{k=1}^n x_k^2)^{\frac{1}{2}}$ ein. Die Kartenwechsel $h_k^l \circ (h_i^j)^{-1}$ haben dann folgende Form:

$$h_k^l \circ (h_i^j)^{-1} : \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ x_i \\ x_{i+1} \\ \vdots \\ x_{k-1} \\ x_k \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ (-1)^j(1-\sum_{l=1}^n x_l^2)^{\frac{1}{2}} \\ x_i \\ \vdots \\ x_{k-2} \\ x_{k-1} \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ (-1)^j(1-\sum_{l=1}^n x_l^2)^{\frac{1}{2}} \\ x_i \\ \vdots \\ x_{k-2} \\ x_k \\ \vdots \\ x_n \end{pmatrix}.$$

Sie sind damit auf ihren Definitionsbereichen offensichtlich differenzierbar.

Differenzierbare Strukturen ermöglichen es uns nun die Differentialrechnung auf Mannigfaltigkeiten einzuführen:

Definition 1.4 (Differenzierbare Abbildung)**.** *Seien M und N differenzierbare Mannigfaltigkeiten und $\varphi : M \to N$ eine stetige Abbildung. Dann heißt φ differenzierbar im Punkt $p \in M$, wenn es zwei Karten (U, g) und (V, h) von M bzw. N gibt, sodass $p \in U$ und $\varphi(p) \in V$ ist, und sodass $h \circ \varphi \circ g^{-1}$ differenzierbar im Punkt $g(p)$ ist.*

Die Abbildung φ heißt differenzierbar, wenn sie in jedem Punkt $p \in M$ differenzierbar ist.

Diese Definition macht Sinn, da ja die Kartenwechsel differenzierbarer Mannigfaltigkeiten ebenfalls differenzierbar sind, und die Differenzierbarkeit einer Abbildung im

Punkt p deshalb unabhängig von der Wahl der Karten erhalten bleibt.
Analog definiert man Strukturen und Mannigfaltigkeiten der Klasse C^k, indem man fordert, dass alle Kartenwechsel k-mal stetig differenzierbar sind. Eine Funktion φ zwischen C^k-Mannigfaltigkeiten heißt k-mal stetig differenzierbar, wenn für alle Karten g und h die Darstellung $h \circ \varphi \circ g^{-1}$ k-mal stetig differenzierbar ist. Man beachte allerdings, dass eine Funktion zwischen C^k-Mannigfaltigkeiten höchstens k-mal stetig differenzierbar sein kann.

Definition 1.5. *Seien M und N differenzierbare Mannigfaltigkeiten und $\varphi : M \to N$ differenzierbar und seien (U, g) und (V, h) Karten von M bzw. N mit $p \in U$ und $\varphi(p) \in V$. Weiters sei $f = h \circ \varphi \circ g^{-1}$ die Darstellung von φ bezüglich g und h. Dann ist der Rang von φ im Punkt p definiert als der Rang der Jacobimatrix df im Punkt $g(p)$.*

Auch der Rang von φ ist unabhängig von der Wahl der Karten g und h, wie man mit Hilfe der Kettenregel leicht sieht.

Lemma 1.6. *Seien M und N zwei $m-$ bzw. $n-$dimensionale C^1–Mannigfaltigkeiten und φ eine stetig differenzierbare Abbildung $M \to N$. Hat φ bei $p \in M$ Rang k, dann existieren Karten $(\widetilde{U}, \widetilde{g})$ und $(\widetilde{V}, \widetilde{h})$ mit $p \in \widetilde{U}$ und $\varphi(p) \in \widetilde{V}$, auf die bezogen φ die Darstellung $\psi = \widetilde{h} \circ \varphi \circ \widetilde{g}^{-1}$ hat mit*

$$\psi_i(x_1, \ldots, x_m) = x_i \qquad \text{für } i = 1 \ldots k \tag{1.2}$$

Gilt rang $\varphi = k$ auf einer Umgebung von p, kann man $\widetilde{g}$ und $\widetilde{h}$ sogar so wählen, dass

$$\psi_i(x_1, \ldots, x_m) = 0 \qquad \text{für } i > k \tag{1.3}$$

Beweis. Seien (U, g) und (V, h) zwei Karten mit $p \in U$ und $\varphi(p) \in V$ und sei $f = h \circ \varphi \circ g^{-1}$ die Darstellung von φ bezüglich dieser Karten. Dann hat die Jacobimatrix von f bei $g(p)$ rang k und o.B.d.A. können wir annehmen, dass

$$\det \left(\frac{\partial f_i}{\partial x_j} \right) \neq 0 \qquad i, j = 1 \ldots k \tag{1.4}$$

— andernfalls lässt sich die Reihenfolge der Koordinaten mit Hilfe eines Kartenwechsels ändern. Betrachtet man nun die Abbildung u:

$$\begin{aligned} u_i(x_1, \ldots, x_m) &= f_i(x_1, \ldots, x_m) \qquad & i &= 1 \ldots, k \\ u_i(x_1 \ldots, x_m) &= x_i & i &= k+1 \ldots m \end{aligned}$$

dann ist die Jacobi-Matrix du im Punkt $g(p)$ invertierbar - also existiert nach dem Satz über die inverse Funktion eine Umgebung von $g(p)$, auf der u ein Diffeomorphismus ist und deshalb als Kartenwechsel interpretiert werden kann. Bezüglich der Karten h und $\widetilde{g} = u \circ g$ hat φ nun die gewünschte Darstellung ψ die (1.2) erfüllt.

Ist zusätzlich rang $\varphi = k$ auf einer Umgebung W von p, dann gilt dort auch $\frac{\partial \psi_i}{\partial x_j} = 0$ für $i, j > k$. Die Jacobi-Matrix von ψ hat nämlich die Form

$$\begin{pmatrix} I_{k \times k} & 0 \\ \frac{\partial \psi_i}{\partial x_j}_{i \leq k < j} & \frac{\partial \psi_i}{\partial x_j}_{i,j > k} \end{pmatrix}$$

Wäre also $\frac{\partial \psi_i}{\partial x_j} \neq 0$ mit $i, j > k$, so wäre rang $\psi > k$ — deshalb hängt ψ auf W nur von $x_1, \ldots, x_k$ ab und wir können hier schreiben $\psi = \psi(x_1, \ldots, x_k)$. Betrachtet man nun die Abbildung z:

$$\begin{aligned} z_i(y_1, \ldots, y_{n_2}) &= y_i && i = 1, \ldots, k \\ z_i(y_1, \ldots, y_{n_2}) &= y_i - \psi_i(y_1, \ldots, y_k) && i = k+1, \ldots, n_2 \end{aligned}$$

so sieht man leicht, dass dz im Punkt $h \circ \varphi(p)$ invertierbar ist — und mit dem Satz über die inverse Funktion folgt wieder die Existenz einer Umgebung von $h \circ \varphi(p)$, sodass z auf dieser Umgebung ein Diffeomorphismus ist. Die Darstellung $\widetilde{\psi}$ von φ bezüglich der Karten $\widetilde{g}$ und $\widetilde{h} = z \circ \widetilde{h}$ — also $\widetilde{\psi} = \widetilde{h} \circ \varphi \circ \widetilde{g}^{-1} = z \circ \psi$ — erfüllt nun zusätzlich (1.3).

□

1.2 Mengen vom Maß null

Um den Satz von Sard formulieren zu können müssen wir noch erklären, was es bedeutet, dass eine Teilmenge einer differenzierbaren Mannigfaltigkeit Maß null hat. Dazu überlegen wir zuerst, wann das Lebesgue-Maß λ_d einer Teilmenge des $\mathbb{R}^d$ null ist. Betrachtet man das Lebesgue-Maß als Vervollständigung der Einschränkung des nach der Borel-Lebesgue-Carathéodory-Methode konstruierten äußeren Maßes auf die Borel-Mengen[1], so ergibt sich die folgende

Definition 1.7. *Eine Teilmenge $M \subset \mathbb{R}^d$ hat Lebesgue-Maß null, wenn für alle $\varepsilon > 0$ eine Folge von Würfeln W_n^ε existiert, sodass gilt*

$$M \subset \bigcup_{n=1}^{\infty} W_n^\varepsilon \qquad \text{und} \qquad \sum_{n=1}^{\infty} vol(W_n^\varepsilon) < \varepsilon \tag{1.5}$$

wobei $vol(W)$ den elementaren Inhalt a^d (bzw. das Lebesgue-Maß) eines d-dimensionalen Würfels mit Seitenlänge a bezeichnet.

Eine Menge hat also Maß null, wenn sie sich von höchstens abzählbar vielen „beliebig kleinen“ Mengen überdecken lässt.

Insbesondere lässt sich mit dieser Definition auch die bekannten Tatsache zeigen, dass jede Teilmenge einer Nullmenge und jede Vereinigung von abzählbar vielen Nullmengen wieder Maß null hat. Ist nämlich $A = \bigcup_{n \in \mathbb{N}} A_n$ mit $\lambda_d(A_n) = 0$, dann gibt es zu $\varepsilon > 0$ für jedes $n \in \mathbb{N}$ eine Überdeckung $A_{nm}, m \in \mathbb{N}$ von A_n mit $\sum_{m \in \mathbb{N}} \mathrm{vol}(A_{nm}) < 2^{-n}\varepsilon$. Es ist $A_{nm}, n, m \in \mathbb{N}$ dann eine abzählbare Überdeckung von A mit $\sum_{n,m \in \mathbb{N}} \mathrm{vol}(A_{nm}) < \varepsilon$.

Lemma 1.8. *Eine Teilmenge $M \subset \mathbb{R}^d$ hat genau dann Lebesgue-Maß null, wenn für alle $\varepsilon > 0$ eine Folge von Kugeln B_n^ε existiert, sodass gilt*

$$M \subset \bigcup_{n=1}^{\infty} B_n^\varepsilon \qquad \text{und} \qquad \sum_{n=1}^{\infty} vol(B_n^\varepsilon) < \varepsilon \tag{1.6}$$

[1]Vgl. [Wer07, Kapitel 5 - 8]

wobei vol(B) den elementaren Inhalt $r^d\omega_1$ einer d–dimensionalen Kugel mit Radius r bezeichnet, wenn ω_1 das Volumen der d-dimensionalen Einheitskugel ist.

Beweis. Sei $\varepsilon > 0$ und $M \subset \mathbb{R}^d$ habe Lebesgue-Maß null. Sei weiters W_n eine Folge von Würfeln, sodass gilt $\sum_{n\in\mathbb{N}} \mathrm{vol}(W_n) < (\frac{2}{\sqrt{d}})^d \omega_1^{-1}\varepsilon$. Ist m_n der Mittelpunkt und a_n die Seitenlänge von W_n, so definiere man B_n als die abgeschlossene Kugel mit Mittelpunkt m_n und Radius $\frac{\sqrt{d}}{2}a_n$. Dann gilt $M \subset \bigcup_{n\in\mathbb{N}} B_n$ und

$$\sum_{n\in\mathbb{N}} \mathrm{vol}(B_n) = \sum_{n\in\mathbb{N}} \left(\frac{\sqrt{d}}{2}a_n\right)^d \omega_1 = \left(\frac{\sqrt{d}}{2}\right)^d \omega_1 \sum_{n\in\mathbb{N}} a_n^d = \left(\frac{\sqrt{d}}{2}\right)^d \omega_1 \sum_{n\in\mathbb{N}} \mathrm{vol}(W_n) < \varepsilon.$$

Existiert umgekehrt zu jedem $\varepsilon > 0$ eine Folge von Kugeln, die (1.5) erfüllt, so betrachte man zu einer Folge von Kugeln B_n die Folge von Würfeln W_n, sodass W_n den gleichen Mittelpunkt wie B_n hat und die Seitenlänge von W_n dem doppelten Radius von B_n entspricht. Dann gilt $M \subset \bigcup_{n\in\mathbb{N}} W_n$ und

$$\sum_{n\in\mathbb{N}} \mathrm{vol}(W_n) = \sum_{n\in\mathbb{N}} (2r_n)^d = 2^d\omega_1^{-1} \sum_{n\in\mathbb{N}} r_n^d\omega_1 = 2^d\omega_1^{-1} \sum_{n\in\mathbb{N}} \mathrm{vol}(B_n)$$

Zu beliebigem $\varepsilon > 0$ wähle man eine Überdeckung B_n^ε von M mit $\sum_{n\in\mathbb{N}} \mathrm{vol}(B_n) < 2^{-d}\omega_1\varepsilon$. Die entsprechende Folge von Würfeln W_n^ε erfüllt dann (1.5). Also hat M Maß null.

□

Mit ähnlichen Argumenten kann man zeigen, dass es egal ist, ob in Definition 1.7 Würfel, Quader oder Kugeln verwendet werden und ob diese offen oder abgeschlossen sind.

Lemma 1.9. *Sei $O \subset \mathbb{R}^d$ offen, $f : O \to \mathbb{R}^d$ stetig differenzierbar und $A \subset O$ mit $\lambda_d(A) = 0$. Dann hat auch $f(A)$ Maß null.*

Beweis. Nehmen wir zuerst an, dass A ganz in einer kompakten Menge enthalten ist. In diesem Fall genügt f auf A einer Lipschitzbedingung

$$\|f(x) - f(y)\| \leq L\|x - y\| \qquad \text{für } x, y \in A.$$

Für beliebiges $\varepsilon > 0$ sei dann $A_n, n \in \mathbb{N}$ eine Überdeckung von A aus Kugeln mit $\sum_{n\in\mathbb{N}} \mathrm{vol}(A_n) < L^{-d}\varepsilon$ und r_n sei der Radius von A_n. Wegen der Lipschitzbedingung ist $f(A_n)$ ganz in einer Kugel C_n mit Radius $L \cdot r_n$ enthalten. Somit gilt $f(A) \subset \bigcup_{n\in\mathbb{N}} C_n$ und $\sum_{n\in\mathbb{N}} \mathrm{vol}(C_n) < \varepsilon$. Also hat $f(A)$ Lebesgue-Maß null.

Ist A nun nicht ganz in einer kompakten Menge enthalten, dann existieren kompakte Mengen $K_n, n \in \mathbb{N}$ mit $K_n \uparrow O$, sodass $f(A)$ sich als die abzählbare Vereinigung der Nullmengen $f(B_n)$ mit $B_n = K_n \cap A$ darstellen lässt und somit selbst Maß null hat.

□

Dieses Lemma macht es nun sinnvoll von Nullmengen einer differenzierbaren Mannigfaltigkeit zu sprechen.

Definition 1.10. *Eine Teilmenge A einer n-dimensionalen differenzierbaren Mannigfaltigkeit M heißt vom Maß null, wenn sie darstellbar ist als $A = \bigcup_{i\in\mathbb{N}} A_i$, sodass Karten (U_i, h_i) existieren mit $A_i \subset U_i$ und $h_i(A_i)$ hat Maß null in $\mathbb{R}^n$.*

Wegen Lemma 1.9 ist diese Definition unabhängig von der Wahl der Mengen A_i und der Karten U_i. Tatsächlich hat A genau dann Maß null, wenn für jede beliebige Karte (V, g) die Menge $g(V \cap A) \subset \mathbb{R}^n$ Maß null hat. Es gilt nämlich

$$g\left(V \cap A\right) = \bigcup_{i \in \mathbb{N}} g\left(V \cap A_i\right) = \bigcup_{i \in \mathbb{N}} g \circ h_i^{-1} \circ h_i(V \cap A_i).$$

Da aber $h_i(V \cap A_i) \subset \mathbb{R}^n$ Maß null hat und der Kartenwechsel $g \circ h_i^{-1}$ stetig differenzierbar ist, hat mit Lemma 1.9 auch $g(V \cap A)$ Maß null.

Offensichtlich gilt auch, dass jede Teilmenge einer abzählbaren Vereinigung von Nullmengen wieder Maß null hat. Da die Topologie einer Mannigfaltigkeit eine abzählbare Basis besitzt, folgt unmittelbar

Korollar 1.11. *Sei A Teilmenge einer n-dimensionalen Mannigfaltigkeit M. Hat jeder Punkt $p \in M$ eine Umgebung U_p, sodass $U_p \cap A$ Maß null hat, so hat auch A Maß null.*

Beweis. Da die Topologie von M eine abzählbare Basis hat, können wir abzählbar viele U_i auswählen, sodass gilt $A = \bigcup_{i \in \mathbb{N}} (U_i \cap A)$. Als abzählbare Vereinigung von Nullmengen hat dann auch A Maß null.

□

Zu beachten ist allerdings, dass wir auf einer Mannigfaltigkeit kein Maß im eigentlichen Sinn definiert haben, sondern nur eine Klasse von „kleinen“ Mengen, die wir als Nullmengen oder „Mengen mit Maß null“ bezeichnen.

Kapitel 2

Der Satz von Sard

Definition 2.1. *Seien M_1 und M_2 n_1- bzw. n_2-dimensionale differenzierbare Mannigfaltigkeiten und φ eine differenzierbare Abbildung $M_1 \to M_2$. Jeder Punkt $p \in M_1$ mit $\operatorname{rang} \varphi < n_2$ heißt kritischer Punkt von φ.*
Ein Punkt $q \in M_2$ heißt kritischer Wert von φ, wenn $\varphi^{-1}(q)$ mindestens einen kritischen Punkt enthält.

Insbesondere ist also für $n_1 < n_2$ jeder Punkt aus M_1 kritisch. Der Satz von Sard lautet nun wie folgt:

Satz 2.2 (von Sard)**.** *Seien M_1 und M_2 C^k-Mannigfaltigkeiten der Dimension n_1 bzw. n_2 und φ eine C^k-Abbildung $M_1 \to M_2$. Gilt*

$$k - 1 \geq \max\{n_1 - n_2, 0\}, \tag{2.1}$$

so hat die Menge der kritischen Werte von φ Maß null.

Beweis. Ist $n_1 < n_2$ so ist der Beweis einfach — in diesem Fall ist jeder Punkt aus M_1 kritisch. Zu zeigen ist also, dass $\varphi(M_1)$ eine Nullmenge ist. Um das einzusehen, betrachte man einen abzählbaren Atlas $(U_i, g_i)_{i \in \mathbb{N}}$ von M_1 und Karten (V_i, h_i) von M_2 mit $\varphi(U_i) \subset V_i$. Sei $U_i' = g_i(U_i) \subset \mathbb{R}^{n_1}$ und sei $f_i = h_i \circ \varphi \circ g_i^{-1}$ die Darstellung von φ bezüglich der Karten (U_i, g_i) und (V_i, h_i). Betrachtet man nun für jedes $i \in \mathbb{N}$ die Abbildung

$$\tilde{f}_i : \begin{cases} U_i' \times \mathbb{R}^{n_2 - n_1} \subset \mathbb{R}^{n_2} & \to \quad \mathbb{R}^{n_2} \\ \qquad (x, y) & \mapsto \quad f_i(x) \end{cases}$$

dann ist $\tilde{f}_i$ stetig differenzierbar und es gilt $f_i(U_i') = \tilde{f}_i(U_i' \times \{0\})$. Da aber bekanntermaßen $\mathbb{R}^{n_1} \times \{0\}$ und somit sicher auch $U_i' \times \{0\}$ in $\mathbb{R}^{n_2}$ Lebesgue-Maß null haben, ist mit Lemma 1.9 auch $f_i(U_i')$ eine Nullmenge. Also hat für jedes $i \in \mathbb{N}$ die Menge $\varphi(U_i) \subset M_2$ Maß null und deshalb auch $\varphi(M_1) = \bigcup_{i \in \mathbb{N}} \varphi(U_i)$.

Sei nun $n_1 \geq n_2$. Wegen Definition 1.10 lässt sich das Problem auf den Fall reduzieren, dass $M_2 = \mathbb{R}^{n_2}$, M_1 Teilmenge des n_1-dimensionalen abgeschlossenen Einheitswürfels E^{n_1} und f eine C^k-Abbildung von einer Umgebung von E^{n_1} in den $\mathbb{R}^{n_2}$ ist. Ist nämlich $(U_i, g_i)_{i \in \mathbb{N}}$ ein abzählbarer Atlas von M_1 und $(V_i, h_i)_{i \in \mathbb{N}}$ ein Atlas von M_2 mit $\varphi(U_i) \subset V_i$ und ist für alle U_i das Bild $U_i' = g_i(U_i) \subset \mathbb{R}^{n_1}$ beschränkt, dann existiert zu jedem $i \in \mathbb{N}$ eine affine Abbildung t_i, sodass $t_i(U_i) \subset \mathrm{E}^{n_1}$ ist. Die Darstellung f_i von φ bezüglich der Karten $\widetilde{g}_i = t_i \circ g_i$ und h_i bildet (wie verlangt) eine Teilmenge des E^{n_1} in den $\mathbb{R}^{n_2}$ ab.

Ist nun C die Menge der kritischen Punkte von φ und $C_i = U_i \cap C$, so ist $\widetilde{C}_i = \widetilde{g}_i(C_i)$ die Menge der kritischen Punkte von f_i. Gilt dann für f_i der Satz von Sard, sodass $f_i(\widetilde{C}_i)$ Maß null hat, so ist $\varphi(C_i) = h_i^{-1}(f_i(\widetilde{C}_i))$ eine Nullmenge. Also hat auch $\varphi(C) = \bigcup_{i\in\mathbb{N}} \varphi(C_i)$ als abzählbare Vereinigung von Nullmengen Maß null.

Falls der Atlas $(U_i, g_i)_{i\in\mathbb{N}}$ aber Karten enthält, für die U_i' nicht beschränkt ist, so lässt sich aus ihm ein Atlas konstruieren, der diese Bedingung erfüllt. Jede unbeschränkte Menge U_i' lässt sich nämlich von abzählbar vielen beschränkten offenen Mengen $U_{ij}', j \in \mathbb{N}$ überdecken. Ersetzt man jede solche Karte (U_i, g_i) durch die Karten $(U_{ij}, g|_{U_{ij}})$ mit $U_{ij} = g^{-1}(U_{ij}')$, so erhält man den gesuchten Atlas.

Betrachten wir nun den Fall $n_1 = n_2$. Zu zeigen ist

Lemma 2.3. *Sei $\varphi = (\varphi_1, \ldots, \varphi_n)^T$ eine Abbildung von E^n nach $\mathbb{R}^n$, wobei jede der Abbildungen φ_i auf einer Umgebung von E^n definiert und stetig differenzierbar ist. Dann hat die Menge der kritischen Werte von φ Maß null.*

Beweis. Zu zwei Punkten $x = (x_1, \ldots, x_n)^T$ und $y = (y_1, \ldots, y_n)^T$ aus E^n existiert nach dem Mittelwertsatz der Differentialrechnung für jedes $i \in \{1, \ldots, n\}$ ein Punkt $z_i = (z_{i\,1}, \ldots, z_{i\,n})^T$ aus der Verbindungsstrecke von x und y, sodass gilt

$$\varphi_i(y) - \varphi_i(x) = \sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial x_j}(z_i)(y_j - x_j). \tag{2.2}$$

Also gilt

$$|\varphi_i(y) - \varphi_i(x)| \leq \sum_{j=1}^{n} \left| \frac{\partial \varphi_i}{\partial x_j}(z_i) \right| |y_j - x_j| \leq \|y - x\| \sum_{j=1}^{n} \left| \frac{\partial \varphi_i}{\partial x_j}(z_i) \right|$$

Und weiter

$$\begin{aligned} \|\varphi(y) - \varphi(x)\|^2 &= \sum_{i=1}^{n} |\varphi_i(y) - \varphi_i(x)|^2 \leq \\ &\leq \sum_{i=1}^{n} \left(\|y - x\| \sum_{j=1}^{n} \left| \frac{\partial \varphi_i}{\partial x_j}(z_i) \right| \right)^2 = \|y - x\|^2 \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial x_j}(z_i) \right)^2 \end{aligned}$$

Da die partiellen Ableitungen $\frac{\partial \varphi_i}{\partial x_j}$ stetig und somit auf E^n beschränkt sind, lässt sich eine Lipschitzkonstante L finden, sodass gilt

$$\|\varphi(y) - \varphi(x)\| < L\ \|y - x\|. \tag{2.3}$$

Sei weiters $T_x = (T_{x\,1}, \ldots, T_{x\,n})^T$ die bei x zu φ tangentiale Abbildung $T_x(y) = \varphi(x) + \mathrm{d}\varphi(x)(y - x)$. Die einzelnen Komponentenfunktionen haben dann die Form

$$T_{x\,i}(y) = \varphi_i(x) + \sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial x_j}(x)(y_j - x_j)$$

Mit (2.2) folgt

$$\varphi_i(y) - T_{x\,i}(y) = \sum_{j=1}^{n} \left(\frac{\partial \varphi_i}{\partial x_j}(z_i) - \frac{\partial \varphi_i}{\partial x_j}(x) \right) (y_j - x_j)$$

Da φ eine C^1-Funktion ist, sind die partiellen Ableitungen $\frac{\partial \varphi_i}{\partial x_j}$ stetig und damit gleichmäßig stetig auf der kompakten Menge E^n. Also existieren für alle $i, j \in \{1, \ldots, n\}$ Funktionen $b_{ij}(\varepsilon)$, sodass gilt: $b_{ij}(\varepsilon) \to 0$ für $\varepsilon \to 0$ und $|\frac{\partial \varphi_i}{\partial x_j}(y) - \frac{\partial \varphi_i}{\partial x_j}(x)| \leq b_{ij}(\|x - y\|)$. Außerdem ist sicher $\|z_i - x\| \leq \|y - x\|$ und damit folgt

$$\begin{aligned} \|\varphi(y) - T_x(y)\|^2 &\leq \sum_{i=1}^{n} \left| \sum_{j=1}^{n} \left(\frac{\partial \varphi_i}{\partial x_j}(z_i) - \frac{\partial \varphi_i}{\partial x_j}(x) \right) (y_i - x_i) \right|^2 \leq \\ &\leq \sum_{i=1}^{n} \left(\sum_{j=1}^{n} b_{ij}(\|x - y\|) \|y - x\| \right)^2 = \|x - y\|^2 \sum_{i=1}^{n} \left(\sum_{j=1}^{n} b_{ij}(\|x - y\|) \right)^2 \end{aligned}$$

Insbesondere existiert also eine Funktion $b(\varepsilon)$, sodass $b(\varepsilon) \to 0$ für $\varepsilon \to 0$ und

$$\|\varphi(y) - T_x(y)\| \leq b(\|x - y\|)\|x - y\| \tag{2.4}$$

Ist x ein kritischer Punkt von φ, dann hat die Jacobi-Matrix $\mathrm{d}\varphi(x)$ nicht vollen Rang und das Bild von $\mathbb{R}^n$ unter T_x ist in einer $n-1$ – dimensionalen Ebene P_x enthalten.

Sei nun ε so klein, dass $b(\varepsilon) \leq L$ ist, und es gelte $\|x - y\| \leq \varepsilon$. Wegen (2.4) gilt dann $\|\varphi(y) - T_x(y)\| \leq b(\varepsilon)\varepsilon$. Es ist also $\varphi(y)$ in jener Region Q enthalten, die durch zwei im Abstand von $b(\varepsilon)\,\varepsilon$ zu P_x parallel liegende $n-1$–dimensionale Ebenen begrenzt wird.

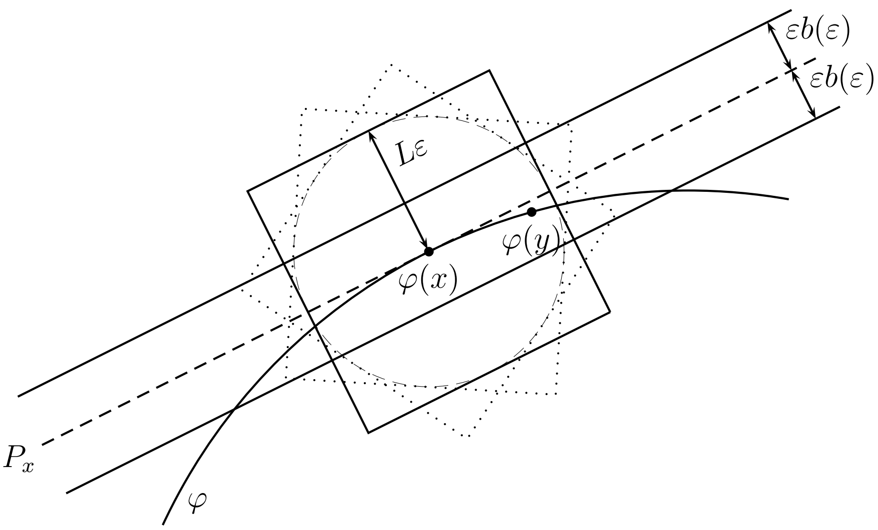

Abb. 2.1: Zwei zu P_x parallele Hyperebenen und Würfel um $\varphi(x)$ mit Seitenlänge $2L\varepsilon$ schränken die Position von $\varphi(y)$ ein

Wegen (2.3) ist $\varphi(y)$ außerdem in jedem Würfel, dem die Kugel mit Mittelpunkt in $\varphi(x)$ und Radius $L\varepsilon$ einbeschrieben werden kann, — also jedem Würfel mit Mittelpunkt in $\varphi(x)$ und Seitenlänge $2L\varepsilon$ — enthalten (vgl. Abb. 2.1). Sei W nun ein solcher Würfel, der 2 zu P_x parallele Seiten hat. Dann ist $W \cap Q$ ein n–dimensionaler Quader mit Volumen $2^n L^{n-1} \varepsilon^n b(\varepsilon)$ und es gilt offensichtlich auch $\varphi(y) \in W \cap Q$.

Teilt man nun E^n in p^n Würfel mit Seitenlänge $1/p$ auf und betrachtet alle jene Würfel, die einen kritischen Punkt x enthalten, dann ist jeder solcher Würfel in einer Kugel um x mit Radius $\sqrt{n}/p$ enthalten. Wählt man p groß genug, so ist nach der obigen Ausführung also das Bild eines solchen Würfels in einem n–dimensionalen Quader mit Maß $2^n L^{n-1} (\frac{\sqrt{n}}{p})^n b(\frac{\sqrt{n}}{p})$ enthalten.

Die Menge der kritischen Werte von φ ist nun aber sicher Teilmenge des Bilds aller Würfel, die zumindest einen kritischen Punkt enthalten. Da es höchstens p^n solche Würfel gibt, ist das Maß der Menge der kritischen Werte sicher kleiner gleich

$$p^n 2^n L^{n-1} \left(\frac{\sqrt{n}}{p}\right)^n b\left(\frac{\sqrt{n}}{p}\right) = 2^n L^{n-1} \left(\sqrt{n}\right)^n b\left(\frac{\sqrt{n}}{p}\right). \tag{2.5}$$

Für $p \to \infty$ geht dieser Ausdruck gegen null — also hat die Menge der kritischen Werte Maß null.

$\square$

Sei nun $n_1 > n_2$ — dann muss (2.5) ersetzt werden durch:

$$p^{n_1} 2^{n_2} L^{n_2} \left(\frac{\sqrt{n_1}}{p}\right)^{n_2} b\left(\frac{\sqrt{n_1}}{p}\right) = 2^{n_2} L^{n_2} \left(\sqrt{n_1}\right)^{n_2} p^{n_1-n_2} b\left(\frac{\sqrt{n_1}}{p}\right) \tag{2.6}$$

Dieser Ausdruck konvergiert für $p \to \infty$ im Allgemeinen allerdings nicht — der Beweis für den Fall $n_1 > n_2$ ist komplexer und erfordert eine genauere Untersuchung des Verhaltens von φ bei kritischen Punkten. Was wir zeigen wollen ist, dass $\|\varphi(x) - \varphi(y)\|$ bei kritischen Punkten „schnell genug“ gegen 0 konvergiert, sodass auch (2.6) für $p \to \infty$ verschwindet. Lemma 2.6 liefert die benötigte Abschätzung — zunächst benötigen wir jedoch noch die beiden folgenden Resultate:

Lemma 2.4. *Sei $\varphi : B_\varepsilon^m = \{x \in \mathbb{R}^m \mid \|x\| < \varepsilon\} \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung, sodass alle partiellen Ableitungen von φ beschränkt sind. Sei weiters $f : O \subset \mathbb{R}^n \to \mathbb{R}$ mit $\varphi(B_\varepsilon^m) \subset O$ stetig differenzierbar und seien $y \in B_\varepsilon^m$, $k \in \mathbb{N}_0$ und $b(\epsilon)$ eine monotone Funktion mit $\lim_{\epsilon \to 0} b(\epsilon) = 0$, sodass für alle $x \in B_\varepsilon^m$ gilt:*

$$\left| \frac{\partial f}{\partial x_j}(\varphi(x)) \right| \leq b(\|x-y\|)\|x-y\|^k \qquad j = 1, \ldots, n.$$

Dann existiert eine Konstante K, die nur von φ abhängt, sodass

$$|f(\varphi(x)) - f(\varphi(y))| \leq K b(\|x-y\|)\|x-y\|^{k+1}$$

für alle $x \in B_\varepsilon^m$ gilt.

Beweis. Sei $x \in B_\varepsilon^m$ und sei $F(t) = f\left(\varphi(y + (x-y)t)\right)$. Dann ist $F(t)$ stetig differenzierbar und es gilt $f(\varphi(x)) = F(1)$ und $f(\varphi(y)) = F(0)$. Mit dem Mittelwertsatz der

Differentialrechnung folgt $f(\varphi(x)) - f(\varphi(y)) = F'(t)$ für ein $t \in (0,1)$. Es gilt aber nach der Kettenregel

$$\begin{aligned} F'(t) &= df(\varphi(y+(x-y)t)) \cdot d\varphi(y+(x-y)t) \cdot (x-y) = \\ &= \sum_{j=1}^{n} \sum_{i=1}^{m} \frac{\partial f}{\partial x_j}(\varphi(y+(x-y)t)) \cdot \frac{\partial \varphi_j}{\partial x_i}(y+(x-y)t) \cdot (x_i - y_i) \end{aligned}$$

und mit $K_1 = m \sup \left| \frac{\partial \varphi_i}{\partial x_j} \right|$ folgt weiter

$$\begin{aligned} |F'(t)| &\leq \sum_{j=1}^{n} \left| \frac{\partial f}{\partial x_j}(\varphi(y+(x-y)t)) \sum_{i=1}^{m} \frac{\partial \varphi_j}{\partial x_i}(y+(x-y)t) \cdot (x_i - y_i) \right| \leq \\ &\leq K_1 \|x-y\| \sum_{j=1}^{n} \left| \frac{\partial f}{\partial x_j} (\varphi(y+(x-y)t)) \right| < \\ &< nK_1 b(\|x-y\|)\|x-y\|^k \end{aligned}$$

Setzt man nun $K = nK_1$, so ist das Lemma bewiesen.

□

Lemma 2.5. *Sei A eine Teilmenge des $\mathbb{R}^n$, $k \in \mathbb{N}$ und $O \subset \mathbb{R}^n$ offen mit $A \subset O$. Dann existieren eine Folge von Mengen $A_i, i \in \mathbb{N}_0$ und Abbildungen $\varphi_i, i \in \mathbb{N}$, sodass $A \subset \bigcup_{i=0}^{\infty} A_i$ ist, A_0 abzählbar ist und für $i \geq 1$ gilt:*

1. *φ_i ist eine Einbettung[1] $B_{\varepsilon_i}^{m_i} = \{x \in \mathbb{R}^{m_i} |\ \|x\| < \varepsilon_i\} \to \mathbb{R}^n$, mit beschränkten partiellen Ableitungen, sodass $A_i \subset \varphi_i(B_{\varepsilon_i}^{m_i})$ ist und für $x, y \in B_{\varepsilon_i}^{m_i}$ gilt*

$$\|\varphi_i(x) - \varphi_i(y)\| \geq \|x - y\| \tag{2.7}$$

2. *Für jede k-mal stetig differenzierbare Funktion $f : O \to \mathbb{R}$, die auf A verschwindet, existieren monotone Funktionen $b_i(\epsilon)$ mit $\lim_{\epsilon \to 0} b_i(\epsilon) = 0$, sodass*

$$|f(\varphi_i(x))| \leq b_i(\|x-y\|)\|x-y\|^k \tag{2.8}$$

   für alle $x, y \in B_{\varepsilon_i}^{m_i}$ mit $\varphi_i(y) \in A_i$ gilt.

Beweis. Im Fall $n = 1$ ist der Beweis einfach. A_0 sei die Menge der isolierten Punkte von A. Für jeden Punkt $x \in A \backslash A_0$ existiert dann eine Folge von Punkten $x_i \in A$ mit $x = \lim_{i \to \infty} x_i$. Verschwindet $f \in C^k(O)$ auf A, so gilt $f(x_i) = 0$ und es folgt $f'(x) = f^{(2)}(x) = \ldots = f^k(x) = 0$. Klarerweise ist nämlich $f'(x) = \lim_{n \to \infty} \frac{f(x_n) - f(x)}{x_n - x} = 0$. Induktiv kann man nun für $m \leq k$ zeigen, dass auch $f^{(m)}(x) = 0$ ist. Gilt $f^{(j)}(x) = 0$ für

[1] Eine Einbettung ist ein stetig differenzierbarer Homöomorphismus $f : O \subset \mathbb{R}^n \to f(O) \subset \mathbb{R}^m$ mit $n \leq m$, wobei $O \neq \emptyset$ und in $\mathbb{R}^n$ offen und $f(O)$ mit der Spurtopologie versehen ist, sodass $\mathrm{d}f(x)$ für alle $x \in O$ maximalen Rang n hat.

$1 \leq j < k$, so existiert nämlich nach dem Satz von Taylor für alle $i \in \mathbb{N}$ ein ξ_i zwischen x_i und x, sodass gilt:

$$0 = f(x_i) = \sum_{j=0}^{m-1} \frac{f^{(j)}(x)}{j!}(x_i - x)^j + \frac{f^{(m)}(\xi_i)}{m!}(x_i - x)^m = \frac{f^{(m)}(\xi_i)}{m!}(x_i - x)^m$$

Also gilt $f^{(m)}(\xi_i) = 0$ — und da mit $(x_i)_{i\in\mathbb{N}}$ auch $(\xi_i)_{i\in\mathbb{N}}$ gegen x konvergiert, folgt $f^{(m)}(x) = \lim_{i\to\infty} f^{(m)}(\xi_i) = 0$.

$A\backslash A_0$ lässt sich sicher durch abzählbar viele Intervalle (a_i, b_i) mit $[a_i, b_i] \subset O$ überdecken. Wählt man nun φ_i als Translation von $(-\frac{b_i-a_i}{2}, \frac{b_i-a_i}{2})$ nach (a_i, b_i) und $A_i = (A\backslash A_0) \cap (a_i, b_i)$, so ist das Lemma erfüllt. Für $z_1 = \varphi_i(x)$ und $z_2 = \varphi_i(y) \in A_i$ gilt nämlich wegen dem gerade Besprochenen nach dem Satz von Taylor:

$$f(z_1) = \sum_{j=0}^{k-1} \frac{f^{(j)}(z_2)}{j!}(z_1 - z_2)^j + \frac{f^{(k)}(\xi)}{k!}(z_1 - z_2)^k = \frac{f^{(k)}(\xi)}{k!}(z_1 - z_2)^k$$

für ein ξ zwischen z_1 und z_2. Da aber $f^{(k)}$ stetig und deshalb auf $[a_i, b_i]$ gleichmäßig stetig ist, existiert eine Funktion $b_i(\epsilon)$ mit $\lim_{\epsilon\to 0} b_i(\epsilon) = 0$, sodass für $x_1, x_2 \in [a_i, b_i]$ gilt $|f^{(k)}(x_1) - f^{(k)}(x_2)| \leq b(|x_1 - x_2|)$. Wegen $f(z_2) = 0$ folgt:

$$|f(\varphi(x))| = |f(z_1)| = \frac{|f^{(k)}(\xi)|}{k!}|z_1 - z_2|^k \leq b_i(|z_1 - z_2|)|z_1 - z_2|^k = b_i(|x - y|)|x - y|^k$$

Auch der Fall $k = 0$ ist leicht — wobei die Dimension n beliebig sein kann. Da O offen ist, findet man dann zu jedem Punkt $p \in A$ eine Umgebung U_p von p für deren Abschluss gilt: $\overline{U_p} \subset O$ — wir können U_p sogar aus der abzählbaren Basis aus ε - Kugeln der Topologie von O wählen. In Summe erhält man dadurch eine abzählbare Überdeckung $U_i, i \in \mathbb{N}$ von A aus ε–Kugeln. Wählt man nun $A_i = (A \cap U_i)$ und φ_i als die Translation von U_i zum Ursprung, dann ist f als stetige Funktion auf der kompakten Menge $\overline{U_i}$ — und damit auch auf U_i selbst — gleichmäßig stetig. Also existieren Funktionen b_i die (2.8) erfüllen.

Der Beweis funktioniert nun durch vollständige Induktion nach $p = n + k$ — die Induktionsannahme ist, dass Lemma 2.5 für $n+k < p$ wahr ist. Sei also $n+k = p$, $n > 1$, $k > 0$ und $A \subset \mathbb{R}^n$. A^1 sei dann definiert als die Menge aller Punkte $x \in A$, bei denen für alle $f \in C^k(O)$, die auf A verschwinden, auch deren partielle Ableitungen verschwinden, wo also gilt:

$$\frac{\partial f}{\partial x_j}(x) = 0 \quad \text{für alle } j = 1, \dots, n$$

und sei $A^2 = A\backslash A^1$.

Laut Induktionsvoraussetzung existieren nun eine abzählbare Menge A_0^1 sowie Mengen A_i^1 und Abbildungen φ_i^1 von $B_{\varepsilon_i}^{m_i} \to \mathbb{R}^n$, $i \in \mathbb{N}$, sodass gilt: $A_i^1 \subset \varphi_i^1(B_{\varepsilon_i}^{m_i})$, die φ_i^1 erfüllen (2.7) und für jede Funktion $g \in C^{k-1}$, die auf A^1 verschwindet, existieren Funktionen $b_i^1(\epsilon)$, mit

$$|g(\varphi_i^1(x))| \leq b_i^1(\|x - y\|)\|x - y\|^{k-1} \tag{2.9}$$

wenn $x \in A_i^1$ ist.

Wegen der Definition von A^1 gilt (2.9) insbesondere für $g = \frac{\partial f}{\partial x_j}$, wenn $f \in C^k$ auf A verschwindet. Wir wollen nun zeigen, dass die Mengen A_r^1 und φ_r^1 (2.8) erfüllen. Das folgt aber unmittelbar aus Lemma 2.4, wenn wir entsprechende Funktionen $\bar{b}_i(\epsilon)$ finden können, die

$$|\frac{\partial f}{\partial x_j}(\varphi_i^1(x))| \leq \bar{b}_i(\|x-y\|)\|x-y\|^{k-1}$$

für alle $j \in \{1, \ldots, n\}$ erfüllen. Die Summen $\bar{b}_i(\epsilon) = \sum_{j=1}^n b_{ij}^1(\epsilon)$ aller Schrankenfunktionen $b_{ij}^1(\epsilon)$ aus (2.9) für die jeweiligen partiellen Ableitungen $\frac{\partial f}{\partial x_j}$ erfüllen dieses Kriterium. Also haben wir Mengen A_r^1 und Abbildungen φ_r^1 gefunden, die ganz A^1 überdecken und den geforderten Bedingungen genügen. Das Gleiche wollen wir nun für A^2 tun.

Sei $p \in A^2$ — dann gibt es also eine Funktion $g \in C^k$, die auf A verschwindet, sodass gilt $\frac{\partial g}{\partial x_i}(p) \neq 0$ für zumindest ein $i \in \{1, \ldots, n\}$. O.b.d.A. können wir annehmen, dass $i = n$. Nach dem Satz über die implizite Funktion existieren eine Umgebung N von p und eine k–mal stetig differenzierbare Funktion $\tau_n(x_1, \ldots, x_{n-1})$, die auf einer Umgebung M von $(p_1, \ldots, p_{n-1})^T$ definiert ist, sodass die vollständige Lösung der Gleichung

$$g(x_1, \ldots, x_n) = 0$$

auf N durch die Punkte $(x_1, \ldots, x_{n-1}, \tau_n(x_1, \ldots, x_{n-1}))^T$ gegeben ist. Betrachtet man nun die Abbildung τ von $M \to \mathbb{R}^n$, die definiert ist durch:

$$\tau((x_1, \ldots, x_{n-1})^T) = (x_1, \ldots, x_{n-1}, \tau_n(x_1, \ldots, x_{n-1}))^T$$

dann gilt: $\tau \in C^k$, $\|\tau(x) - \tau(y)\| \geq \|x-y\|$ und $N \cap A \subset \tau(M)$.

Wendet man nun die Induktionsvoraussetzung auf $D = \tau^{-1}(N \cap A)$ an, so erhält man Mengen $D_i, i \in \mathbb{N}_0$, sodass $D \subset \bigcup_{i \in \mathbb{N}_0} D_i$ und D_0 abzählbar ist, und Abbildungen ψ_i, die (2.7) und (2.8) erfüllen, sodass für jede C^k-Funktion h, die auf D verschwindet, Funktionen $b_i(\epsilon)$ existieren, die

$$|h(\psi_i(x))| < b_i(\|x-y\|)\|x-y\|^k$$

erfüllen, wenn $\psi_i(y) \in D_i$ ist. Verschwindet aber $f \in C^k$ auf A, dann verschwindet $h = f \circ \tau \in C^k$ auf D. Setzt man also $\varphi_i = \tau \circ \psi_i$ und $N_i = \tau(D_i)$, so erhält man eine Überdeckung von $N \cap A$, die allen geforderten Kriterien genügt.

Da zu jedem Punkt $p \in A^2$ eine solche Umgebung N_p existiert, lässt sich ganz A^2 durch solche überdecken — und da eine abzählbare Basis der Topologie des $\mathbb{R}^n$ existiert können wir sogar eine abzählbare Teilüberdeckung $N_q, q \in \mathbb{N}$ auswählen. Verfährt man nun mit jeder dieser Mengen N_q wie oben dargelegt so erhält man in Summe wieder eine abzählbare Überdeckung von A^2, die die geforderten Kriterien erfüllt. Vereinigt man diese Überdeckung mit der von A^1 und benennt die Vereinigung aller darin enthaltenen abzählbaren Mengen mit A_0, so erhält man die gesuchte Überdeckung von A.

□

Mit Lemma 2.4 und Lemma 2.5 erhält man nun leicht das folgende Lemma 2.6, auf dem der Beweis des Satz von Sard im Fall $n_1 > n_2$ beruht:

Lemma 2.6. *Sei $A \subset \mathbb{R}^n$, sei $O \subset \mathbb{R}^n$ offen mit $A \subset O$ und sei $q \in \mathbb{N}$. Dann existieren Mengen $A_i, i \in \mathbb{N}_0$ mit $A \subset \bigcup_{i=0}^{\infty} A_i$, sodass A_0 abzählbar ist und für die $A_i, i \geq 1$ gilt: Für jede q mal stetig differenzierbare Abbildung $f : O \to \mathbb{R}$, deren kritische Punkte eine Obermenge von A sind, existieren Funktionen $b_i, i \in \mathbb{N}$ für die gilt: $b_i(\epsilon)$ ist monoton fallend, $b_i(\epsilon) \to 0$ für $\epsilon \to 0$ und*

$$|f(x) - f(y)| < b_i(\|x - y\|)\|x - y\|^q \tag{2.10}$$

für alle $x, y \in A_i$.

Beweis. Da jeder Punkt $x \in A$ kritischer Punkt von f ist, gilt rang $\mathrm{d}f(x) = 0$ — also verschwinden alle partiellen Ableitungen von f auf A. Wenden wir Lemma 2.5 an, so erhalten wir Mengen $A_i, i \in \mathbb{N}_0$, und Abbildungen $\varphi_i, i \in \mathbb{N}$, sodass A_0 abzählbar ist und alle φ_i (2.7) erfüllen. Außerdem existieren für jede partielle Ableitung $\frac{\partial f}{\partial x_j}, j = 1, \ldots, n$ monotone Funktionen $b_{ji}(\epsilon), i \in \mathbb{N}$, mit $\lim_{\epsilon \to 0} b_{ji}(\epsilon) = 0$, sodass gilt

$$\left|\frac{\partial f}{\partial x_j}(\varphi(x))\right| < b_{ji}(\|x - y\|)\|x - y\|^{q-1}$$

falls $\varphi_i(y) \in A_i$ ist. Insbesondere gibt es deshalb auch für jedes $i \in \mathbb{N}$ eine monotone Funktion $b_i(\epsilon)$ mit $\lim_{\epsilon \to 0} b_i(\epsilon) = 0$, sodass dann für alle $j \in \{1, \ldots, n\}$ gilt:

$$\left|\frac{\partial f}{\partial x_j}(\varphi(x))\right| < b_i(\|x - y\|)\|x - y\|^{q-1}$$

Aus Lemma 2.4 folgt nun weiter die Existenz von Funktionen $\hat{b}_i(\varepsilon)$ mit $\lim_{\varepsilon \to 0} \hat{b}_i(\varepsilon) = 0$, sodass gilt

$$|f(\varphi(x)) - f(\varphi(y))| < \hat{b}_i(\|x - y\|)\|x - y\|^q$$

falls $\varphi(y) \in A_i$ — wegen $A_i \subset \varphi(B_{\epsilon_i}^{m_i})$ und (2.7) gilt damit insbesondere für $x, y \in A_i$

$$|f(x) - f(y)| < \hat{b}_i(\|x - y\|)\|x - y\|^q$$

□

Der Beweis des Satz von Sard ist nun nicht mehr schwer — zuerst beweisen wir:

Lemma 2.7. *Sei $\varphi = (\varphi_1, \ldots, \varphi_{n_2})$ eine Abbildung des n_1-dimensionalen Einheitswürfel $E^{n_1} \to \mathbb{R}^{n_2}$, wobei jede Funktion φ_i eine auf einer Umgebung von E^{n_1} definierte C^q-Funktion ist, und sei A die Menge der Punkte von E^{n_1}, wo gilt rang $\varphi = 0$. Gilt $q \geq \frac{n_1}{n_2}$, dann hat $\varphi(A)$ Maß null.*

Beweis. Wir wenden Lemma 2.6 an und erhalten Mengen A_r, $r \in \mathbb{N}_0$, sodass $A \subset \bigcup_{r \in \mathbb{N}} A_r$ ist. Zu zeigen ist nun, dass $\varphi(A_r)$ für alle $r \in \mathbb{N}_0$ Maß null hat. Im Fall $r = 0$ ist das trivial, da A_0 abzählbar ist — betrachten wir also A_r mit $r \geq 1$. Bei einem Punkt x gilt rang $\varphi = 0$ genau dann, wenn x ein kritischer Punkt von allen φ_i ist. Mit (2.10) existieren nun Funktionen $b_{ir}(\epsilon)$, sodass für $x, y \in A_r$ mit $\|x - y\| < \varepsilon$ gilt

$$|\varphi_i(x) - \varphi_i(y)| < b_{ir}(\varepsilon)\varepsilon^q$$

und damit existiert auch eine Funktion $b_r(\epsilon)$ mit $\lim_{\varepsilon\to 0} b_r(\varepsilon) = 0$, sodass gilt

$$\|\varphi(x) - \varphi(y)\| < b_r(\varepsilon)\varepsilon^q. \tag{2.11}$$

Teilt man nun, wie im Beweis von Lemma 2.3, E^{n_1} in p^{n_1} Würfel W_α mit Seitenlänge$\frac{1}{p}$, dann ist für jeden dieser Würfel $\varphi(A_r \cap W_\alpha)$ wegen (2.11) in einer Kugel mit Radius $b\left(\frac{\sqrt{n_1}}{p}\right)\left(\frac{\sqrt{n_1}}{p}\right)^q$ enthalten. Da es höchstens p^{n_1} solcher Kugeln gibt, und $\varphi(A_r)$ sicher ganz in diesen enthalten ist, ist das Maß von $\varphi(A_r)$ ist auf jeden Fall kleiner gleich

$$p^{n_1}\left(b\left(\frac{\sqrt{n_1}}{p}\right)\left(\frac{\sqrt{n_1}}{p}\right)^q\right)^{n_2}\omega_{n_2} = \sqrt{n_1}^{qn_2}\omega_{n_2} b\left(\frac{\sqrt{n_1}}{p}\right)^{n_2} p^{n_1-qn_2}$$

wobei ω_{n_2} das Volumen der n_2-dimensionalen Einheitskugel bezeichnet. Da $q \geq \frac{n_1}{n_2}$ ist, konvergiert dieser Ausdruck für $p \to \infty$ gegen null — also ist $\varphi(A_r)$ eine Nullmenge.

□

Lemma 2.8. *Sei φ wie in Lemma 2.7 und sei A die Menge der Punkte mit rang $\varphi = r$ $(0 \leq r < n_2)$. Dann hat $\varphi(A)$ Maß null, wenn gilt $q \geq \frac{n_1-r}{n_2-r}$.*

Beweis. Den Fall $r = 0$ haben wir schon in Lemma 2.7 betrachtet — sei also $r \geq 1$. Wegen Lemma 1.11 reicht es zu zeigen, dass für jeden Punkt aus A eine Umgebung U existiert, sodass $\varphi(U \cap A)$ eine Nullmenge ist. Sei also $x_0 \in A$. Wegen Lemma 1.2 existieren Karten (U, g) um x_0 und (V, h) um $\varphi(x)$, auf die bezogen φ die Darstellung $\psi = h \circ \varphi \circ g^{-1}$ hat mit

$$\begin{aligned}
\psi_i(x_1,\ldots,x_{n_1}) &= x_i & i &= 1,\ldots,r\\
\psi_i(x_1,\ldots,x_{n_1}) &= F_i(x_1,\ldots,x_n) & i &= r+1,\ldots,n_2
\end{aligned}$$

Hält man nun $x_1,\ldots,x_r$ fest, dann hat die Abbildung[2]

$$\eta : \left\{\begin{array}{ccc} U_{(x_1,\ldots,x_r)} & \to & \mathbb{R}^{n_2-r} \\ (x_{r+1},\ldots,x_{n_1}) & \mapsto & F(x_1,\ldots,x_r,x_{r+1},\ldots,x_{n_1}) \end{array}\right.$$

genau dann Rang null, wenn $(x_1,\ldots,x_{n_1}) \in U \cap A$ ist — nach Lemma 2.7 hat also die Menge $\eta(\,(U \cap A)_{(x_1,\ldots,x_r)}) = (\varphi(U \cap A))_{(x_1,\ldots,x_r)}$ Maß null in $\mathbb{R}^{n_2-r}$.

Da nach dem Satz von Fubini für jede messbare Menge $M \subset \mathbb{R}^d$ gilt:

$$\lambda_d(M) = \int_{\mathbb{R}^{d_1}} \lambda_{d_2}(M_x)\, d\lambda_{d_1}(x)$$

falls $d = d_1 + d_2$ ist, folgt damit unmittelbar, dass ganz $\varphi(U \cap A)$ eine Nullmenge ist.

□

Der Beweis des Satz von Sard für den Fall, dass $n_1 > n_2$ ist nun einfach:

Sei A die Menge der kritischen Punkte von φ, dann gilt $A = \bigcup_{r=0}^{n_2-1} A_r$, wobei A_r die Menge der Punkte ist, wo φ Rang r hat. Die Menge der kritischen Werte ist dann sicher enthalten in $\bigcup_{r=0}^{n_2-1} f(A_r)$.

[2] $U_{(x_1,\ldots,x_n)}$ bezeichnet den $(x_1,\ldots,x_n)$–Schnitt von U: $\{(x_{r+1},\ldots,x_{n_1}) \in \mathbb{R}^{n_1-r} \mid (x_1,\ldots,x_{n_1}) \in U\}$.

Wegen Voraussetzung (2.1) ist $k \geq n_1 - n_2 + 1$. Also ist $k \geq \frac{n_1 - r}{n_2 - r}$ für $r \in \{0, \ldots, n_2 - 1\}$ und nach Lemma 2.8 hat deshalb dann jede der Mengen $f(A_r)$ Maß null — und somit auch $f(A)$ selbst.

$\square$

Aus dem Satz von Sard erhält man insbesondere folgendes

Korollar 2.9. *Sei $O \subset \mathbb{R}^n$ offen, $f : O \to \mathbb{R}^n$ stetig differenzierbar und C die Menge der kritischen Punkte von f. Dann hat $f(C)$ Lebesgue-Maß null.*

Kapitel 3

Verallgemeinerte Transformationsformel

Seien $X, Y \subset \mathbb{R}^p$ offen und $t : X \to Y$ ein C^1–Diffeomorphismus. Dann lautet die bekannte Transformationsformel für Integrale

$$\int_Y f \; d\lambda_p = \int_X f \circ T \cdot |\det dt| \; d\lambda_p \tag{3.1}$$

Mit dem Satz von Sard ist es nun möglich, die Transformationsformel für Integrale zu verallgemeinern:

Satz 3.1 (Verallgemeinerte Transformationsformel)**.** *Seien $X \subset \mathbb{R}^p$ offen, $t : X \to \mathbb{R}^p$ stetig differenzierbar, $Y = t(X)$ und C die Menge der kritischen Punkte von t. Für $y \in Y$ sei $N(y) \in [0, +\infty]$ die Anzahl der $x \in X \backslash C$ mit $t(x) = y$. Dann ist $N(y) \in \mathcal{M}^+(Y, \mathfrak{L}^p_Y)$ und für alle $f \in \mathcal{M}^+(Y, \mathfrak{L}^p_Y)$ gilt:*[1]

$$\int_Y Nf \; d\lambda_p = \int_X f \circ t \cdot |\det dt| \; d\lambda_p \tag{3.2}$$

Für Lebesuge-meßbare Funktionen $f : Y \to \mathbb{R}$ oder $f : Y \to \mathbb{C}$ ist Nf genau dann λ_p-integrierbar über Y, wenn $f \circ t \cdot |\det dt|$ über X λ^p-integrierbar ist. In diesem Fall gilt (3.2)*.*

Beweis. Sei zunächst $\det dt \neq 0$ und $K \subset X$ kompakt. Zu jedem $x \in K$ existieren dann offene Umgebungen U_x von x und V_x von $t(x)$, sodass $t|_{U_x} : U_x \to V_x$ ein C^1-Diffeomorphismus ist. $\{U_x \mid x \in K\}$ bildet eine offene Überdeckung von K und da K kompakt ist, existieren endlich viele $x_1, \dots, x_m \in K$, sodass gilt $K \subset \bigcup_{j=1}^m U_{x_j}$.
Die Mengen $A_1 = U_{x_1} \cap K$, $A_2 = (U_{x_2} \cap K) \backslash A_1$, …, $A_m = (U_{x_m} \cap K) \backslash \bigcup_{j=1}^{m-1} A_j$ sind dann disjunkte Lebesgue-Mengen mit $\bigcup_{j=1}^m A_j = K$. Für jede dieser Mengen A_j gilt nach der Transformationsformel außerdem:

$$\int_{t(A_j)} f \; d\lambda_p = \int_{A_j} f \circ t \cdot |\det dt| \; d\lambda_p$$

[1]Anmerkung: $\mathcal{M}^+(Y, \mathfrak{L}^p_Y)$ bezeichnet die Menge der Lebesgue-messbaren Funktionen von Y nach $[0, +\infty]$.

Addiert man diese Gleichungen, so erhält man

$$\int_{t(K)} N_K f \; d\lambda_p = \int_K f \circ t \cdot |\det dt| \; d\lambda_p \tag{3.3}$$

wobei $N_K(y) = \sum_{j=1}^m \mathbb{1}_{t(A_j)}(y)$ gleich der (endlichen) Anzahl der $x \in K$ mit $t(x) = y$ ist. Offensichtlich gilt $N_K \in \mathcal{M}^+(Y, \mathfrak{L}_Y^p)$.

Wählt man nun eine Folge kompakter Mengen $K_n \subset X$ mit $K_n \uparrow X$, dann gilt: $t(K_n) \uparrow Y$ und $N_{K_n} \uparrow N$. Also ist $N \in \mathcal{M}^+(Y, \mathfrak{L}_Y^p)$ und für $n \to \infty$ liefert die Gleichung (3.3) mit K_n statt K schließlich die verallgemeinerte Transformationsformel (3.2).
Wir haben die Gültigkeit der verallgmeinerten Transformationsformel also gezeigt, falls $\det dt \neq 0$ ist.

Sei nun $\det dt$ nicht notwendigerweise ungleich null. Dann ist C als Urbild von $\{0\}$ unter der stetigen Funktion $\det dt$ abgeschlossen und $X \backslash C$ deshalb offen - also gilt (3.2) wie oben gezeigt für $X \backslash C$ und $t(X \backslash C)$ an Stelle von X und Y. Da aber $Y \backslash t(X \backslash C) \subset t(C)$ nach dem Satz von Sard bzw. Korollar 2.9 eine Nullmenge ist, folgt die Behauptung in vollem Umfang.

$\square$

Literaturverzeichnis

[BJ90] Bröcker, Theodor und Klaus Jänich: *Einführung in die Differentialtopologie.* Springer, 1990.

[Els05] Elstrodt, Jürgen: *Maß- und Integrationstheorie.* Springer, 4. Auflage, 2005.

[Ste65] Sternberg, Shlomo: *Lectures on Differential Geometry.* Prentice-Hall, 2. Auflage, 1965.

[Wer07] Wertz, Wolfgang: *Maß- und Integrationstheorie.* TU Wien, 2007.