Differentialgeometrie

Wintersemester 21/22

Prof. Dr. E. Kuwert

Mathematisches Institut
Universität Freiburg

Inhaltsverzeichnis

Einleitung

Gegenstand der „klassischen“ Differentialgeometrie sind Flächen (oder Kurven) $M \subset \mathbb{R}^3$. Diese können lokal durch Parametrisierungen beschrieben werden. Ab dem 19. Jahrhundert wurden Situationen betrachtet, in denen Flächen nicht eingebettet sind, also nicht in einem umgebenden Raum liegen. Beispiele:

- die hyperbolische Ebene $\mathbb{H}^2$ (J. Bolyai/N. I. Lobatschewski um 1830) D. Hilbert hat 1901 gezeigt, es gibt keine isometrische Einbettung $\mathbb{H}^2 \subset \mathbb{R}^3$.

- das *Theorema egregium* von C. F. Gauß (1827) besagt, dass die Gaußkrümmung K einer Fläche invariant unter Verbiegungen ist, das heißt sie hängt nur von den Längenverhältnissen auf der Fläche ab und nicht von der Lage der Fläche im Raum. In seinem Habilitationsvortrag (1854) zum Thema *Über die Hypothesen, die der Geometrie zugrunde liegen* hat B. Riemann dann das Konzept eines abstrakten Raums mit Längen- und Winkelmessung entworfen; heute sprechen wir von einer Riemannschen Mannigfaltigkeit. Die Krümmung wird dabei durch den Riemannschen Krümmungstensor beschrieben.

- Grundlage der Allgemeinen Relativitätstheorie (um 1915) von A. Einstein ist der Begriff einer Raum-Zeit mit Längen- und Winkelmessung. Die Kopplung zwischen Materie und Geometrie der Raumzeit ist dabei durch die Einsteinschen Feldgleichungen gegeben.

- Das Problem der mehrdeutigen Lösung von algebraischen Gleichungen, wie zum Beispiel $w^2 - z = 0$, führte zur Theorie der Riemannschen Flächen. Der Begriff der Mannigfaltigkeit wurde erstmals in diesem Zusammenhang in moderner Weise formuliert, und zwar von H. Weyl in seinem Buch *Die Idee der Riemannschen Fläche* von 1913. Im Beispiel gibt es für $z \in \mathbb{C}\backslash\{0\}$ lokal zwei Lösungsfunktionen, die aber nicht auf ganz $\mathbb{C}$ stetig definiert werden können. Auf einer geeigneten abstrakten Fläche kann aber $\sqrt{z}$ global definiert werden.

1 Der Begriff der Mannigfaltigkeit

Definition 1.1 *Eine Topologie auf einer Menge X ist ein Mengensystem $\mathcal{O} \subset 2^X$ mit den folgenden Eigenschaften:*

(i) $\emptyset, X \in \mathcal{O}$,

(ii) $\bigcup_{\lambda \in \Lambda} U_\lambda \in \mathcal{O}$, *falls* $U_\lambda \in \mathcal{O}$ *für alle* $\lambda \in \Lambda$ *(Λ beliebige Indexmenge),*

(iii) $\bigcap_{i=1}^N U_i \in \mathcal{O}$, *falls* $U_i \in \mathcal{O}$ *für* $i = 1, \ldots, N$ *mit* $N < \infty$.

Die Menge X mit dem System $\mathcal{O}$ heißt topologischer Raum. Mit einem Vektorraum hat das nichts zu tun, man verwendet das Wort Raum nur wegen der Anschaulichkeit. Das griechische Wort *topos* kann mit *Ort* übersetzt werden. Die offenen Mengen in X sind genau die Mengen in $\mathcal{O}$. Eine Menge $A \subset X$ heißt abgeschlossen, wenn $X \backslash A$ offen ist, also $X \backslash A \in \mathcal{O}$. Für $x \in X$ heißt jede offene Menge U mit $x \in U$ offene Umgebung von X. Allgemein ist eine Umgebung von $x \in X$ eine beliebige Menge, die eine offene Umgebung von x enthält. Die folgende Eigenschaft ist nicht in jedem topologischen Raum erfüllt, aber in dieser Vorlesung wird sie immer verlangt.

Definition 1.2 *Eine Topologie auf X heißt Hausdorffsch bzw. X heißt Hausdorffraum*, wenn je zwei Punkte $x, y \in X$ mit $x \neq y$ offene Umgebungen U bzw. V besitzen mit $U \cap V = \emptyset$.*

Example 1.1 Jeder metrische Raum (X, d) ist ein topologischer Raum, und zwar ist $U \subset X$ nach Definition offen wenn gilt:

für alle $y \in U$ gibt es ein $\varrho > 0$ mit $B_\varrho(y) \subset U$.

Dabei ist $B_\varrho(y) = \{z \in X : d(y, z) < \varrho\}$ die offene Kugel um y mit Radius $\varrho > 0$. Es gilt:

(i) $B_r(x)$ ist offen für alle $x \in X$, $r > 0$.

(ii) Jedes offene U ist Vereinigung von offenen Kugeln.

(iii) Das System aller offenen Mengen ist eine Topologie.

(iv) X ist damit ein Hausdorffraum. //

Example 1.2 (Relativtopologie) Sei X ein topologischer Raum mit Topologie $\mathcal{O}$. Für eine Teilmenge $M \subset X$ ist

$$\mathcal{O}_M = \{M \cap U : U \in \mathcal{O}\}$$

eine Topologie auf M, die sogenannte Relativtopologie. Die Schnitte $M \cap A$ mit A abgeschlossen in X sind dann abgeschlossen bezüglich der Relativtopologie, denn

$$M \backslash (M \cap A) = M \cap (X \backslash A) = \text{ offen bzgl. der Relativtopologie.}$$

Alle abgeschlossenen Mengen bezüglich $\mathcal{O}_M$ sind von diesem Typ, denn für $U \subset X$ offen gilt

$$M \backslash (M \cap U) = M \cap (X \backslash U), \text{ und } X \backslash U \text{ ist abgeschlossen.}$$

*Felix Hausdorff, 1868-1942

Ist die Topologie auf X Hausdorffsch, so gilt das auch für die Relativtoplogie auf M: zu $x, y \in M$, $x \neq y$, gibt es disjunkte offene Umgebungen U, V in X, und dann sind $M \cap U$, $M \cap V$ ebenfalls disjunkt.

Statt *offen (abgeschlossen) bezüglich der Relativtopologie* sagt man auch kurz *relativ offen (abgeschlossen)*. Ist (X, d) ein metrischer Raum, so haben wir auf $M \subset X$ die induzierte Metrik $d_M(x, y) = d(x, y)$ $(x, y \in M)$. Die Relativtopologie ist die zugehörige Topologie.

Definition 1.3 *Ein System $\mathcal{B}$ von offenen Mengen in einem topologischen Raum X heißt Basis der Topologie, wenn jede offene Menge Vereinigung von Mengen aus $\mathcal{B}$ ist.*

Example 1.3 Sei (X, d) ein separabler metrischer Raum, das heißt es gibt eine abzählbare Teilmenge $\{p_i : i \in \mathbb{N}\}$, die dicht ist, zum Beispiel $\mathbb{Q} \subset \mathbb{R}$. Dann bilden die Kugeln $B_{\frac{1}{k}}(p_i)$ mit $i, k \in \mathbb{N}$ eine abzählbare Umgebungsbasis. Dazu ist zu zeigen: zu jeder Kugel $B_r(p)$ gibt es $i, k \in \mathbb{N}$ mit $p \in B_{\frac{1}{k}}(p_i)$ und $B_{\frac{1}{k}}(p_i) \subset B_r(p)$, bzw.

$$d(p, p_i) < \frac{1}{k} \quad \text{und} \quad d(p, p_i) + \frac{1}{k} \leq r.$$

Wir wählen erst k mit $\frac{1}{k} < r$, und dann p_i hinreichend nahe an p. Weiter zeigen wir: jede Teilmenge $M \subset X$, $M \neq \emptyset$, ist mit der induzierten Metrik ebenfalls separabel. Und zwar gibt es zu $i, k \in \mathbb{N}$ einen Punkt $q_{i,k} \in M$ mit

$$d(q_{i,k}, p_i) < \operatorname{dist}(p_i, M) + \frac{1}{k}.$$

Für einen gegebenen Punkt $q \in M$ und $i, k \in \mathbb{N}$ gilt die Abschätzung

$$\begin{aligned} d(q_{i,k}, q) &\leq d(q_{i,k}, p_i) + d(p_i, q) \\ &< \operatorname{dist}(p_i, M) + \frac{1}{k} + d(p_i, q) \\ &\leq 2d(p_i, q) + \frac{1}{k} < \varepsilon, \end{aligned}$$

für p_i hinreichend nahe bei q und k groß. Insbesondere ergibt sich: in einem separablen metrischen Raum hat die Relativtopologie auf einer Teilmenge M eine abzählbare Umgebungsbasis.

Wir wollen jetzt erklären, was eine stetige Abbildung zwischen topologischen Räumen ist. Im Gegensatz zur Analysis im $\mathbb{R}^n$ definieren wir nicht erst die Stetigkeit in einem festen Punkt, sondern gleich (bzw. nur) die Stetigkeit der gesamten Abbildung.

Definition 1.4 *Seien X, Y topologische Räume. Eine Abbildung $f : X \to Y$ heißt stetig, falls folgende Implikation gilt:*

$$V \text{ offen in } Y \quad \Rightarrow \quad f^{-1}(V) \text{ offen in } X.$$

Example 1.4 In metrischen Räumen X, Y ist das äquivalent zur Folgenstetigkeit in allen Punkten $x \in X$.

Example 1.5 Seien X, Y topologische Räume und $M \subset X$. Dann ist $f : M \to Y$ stetig bezüglich der Relativtopologie, wenn gilt: ist $V \subset Y$ offen, so gibt es $U \subset X$ offen mit

$$f^{-1}(V) = M \cap U.$$

Ist M eine offene Teilmenge von X, so ist das gleichbedeutend mit $f^{-1}(V)$ offen in X, dieser Fall tritt unten auf. Ist $f : X \to Y$ stetig, so ist die Einschränkung

$$f|_M : M \to Y,\ (f|_M)(x) := f(x),$$

stetig bezüglich der Relativtopologie auf M. Denn für $V \subset Y$ offen haben wir

$$(f|_M)^{-1}(V) = M \cap f^{-1}(V).$$

Definition 1.5 *Seien U, V offene Teilmengen der topologischen Räume X bzw. Y. Eine Abbildung $f : U \to V$ heißt Homeomorphismus, falls f bijektiv ist und f, f^{-1} beide stetig sind (als Abbildungen nach Y bzw. X).*

Die hier eingeführten topologischen Begriffe sind gewöhnungsbedürftig. Wir wollen aber nun zum Konzept der Mannigfaltigkeit kommen. Die zentrale Idee ist, dass lokal Koordinaten definiert werden können.

Definition 1.6 (C^0-Atlas) *Sei M ein topologischer Raum.*

(a) *Eine n-dimensionale Karte von M ist ein Homeomorphismus $\varphi : U \to \varphi(U)$ zwischen offenen Mengen $U \subset M$ und $\varphi(U) \subset \mathbb{R}^n$. U ist das Definitionsgebiet und $\varphi(U)$ das Parametergebiet der Karte φ.*

(b) *Ein n-dimensionaler C^0-Atlas von M ist ein System $\varphi_i : U_i \to \varphi_i(U_i)$, $i \in I$ Indexmenge, von n-dimensionalen Karten mit $\bigcup_{i \in I} U_i = M$.*

Wir können uns den Atlas tatsächlich wie ein Buch vorstellen, auf der Seite mit der Nummer $i \in I$ ist das Teilgebiet U_i von M parametrisiert.

Definition 1.7 (C^0-Mannigfaltigkeit) *Ein topologischer Raum M, der Hausdorffsch ist und eine abzählbare Umgebungsbasis hat, heißt n-dimensionale C^0-Mannigfaltigkeit, wenn es einen n-dimensionalen C^0-Atlas von M gibt.*

Example 1.6 Jede offene Menge $\Omega \subset \mathbb{R}^n$ mit der Relativtopologie ist eine n-dimensionale C^0-Mannigfaltigkeit, mit Atlas $\{\mathrm{id} : \Omega \to \Omega\}$.

Example 1.7 $M = \{(x, y) \in \mathbb{R}^2 : |x| + |y| = 1\}$ ist mit der Relativtopologie ein Hausdorffraum mit abzählbarer Basis, vgl. Beispiele 1.2 und 1.3. Schreiben wir M als Graph über den Koordinatenachsen, so ergibt sich der folgende ein-dimensionale Atlas:

$$\begin{array}{lll}
\varphi_1 : U_1 = M \cap \{x > 0\} \to (-1, 1), & \varphi_1(x, y) = y, & \varphi_1^{-1}(y) = (1 - |y|, y) \\
\varphi_2 : U_2 = M \cap \{y > 0\} \to (-1, 1), & \varphi_2(x, y) = x, & \varphi_2^{-1}(x) = (x, 1 - |x|) \\
\varphi_3 : U_3 = M \cap \{x < 0\} \to (-1, 1), & \varphi_3(x, y) = y, & \varphi_3^{-1}(y) = (|y| - 1, y) \\
\varphi_4 : U_4 = M \cap \{y < 0\} \to (-1, 1), & \varphi_4(x, y) = x, & \varphi_4^{-1}(x) = (x, |x| - 1).
\end{array}$$

Zur Stetigkeit der Karten siehe Beispiel 1.5.

Example 1.8 Das Kreuz $M = \{(x,y) \in \mathbb{R}^2 : |x| = |y|\}$, mit der induzierten Topologie, ist keine eindimensionale C^0-Mannigfaltigkeit (Übungsaufgabe).

Remark 1.1 Eine nichttriviale Bemerkung: zwei C^0-Atlanten $\mathcal{A}_1, \mathcal{A}_2$ auf demselben topologischen Raum M haben die gleiche Dimension. Wähle dazu Karten

$$\varphi_i : U_i \to \varphi_i(U_i) \subset \mathbb{R}^{n_i} \text{ in } \mathcal{A}_i \quad \text{ mit } U_1 \cap U_2 \neq \emptyset.$$

Dann ist $\varphi_2 \circ \varphi_1^{-1} : \varphi_1(U_1 \cap U_2) \to \varphi_2(U_1 \cap U_2)$ ein Homeomorphismus zwischen nichtleeren, offenen Teilmengen von $\mathbb{R}^{n_i}$. In der Topologie wird gezeigt, das ist nur für $n_1 = n_2$ möglich.

Die Bedingungen *Hausdorff* und *abzählbare Umgebungsbasis* sind technischer Natur, u.a. werden gewisse pathologische Fälle ausgeschlossen, wir wollen uns gar nicht damit befassen welche. Stattdessen stellen wir folgende wichtige Konsequenzen dieser Voraussetzungen fest.

Theorem 1.1 (Kompaktheitseigenschaften) *Für eine topologische Mannigfaltigkeit M gilt:*

(1) *M ist lokal kompakt.*

(2) *Es gibt eine Ausschöpfung $M = \bigcup_{k=1}^{\infty} G_k$, mit G_k offen und $G_k \subset\subset G_{k+1}$.*

(3) *M hat einen abzählbaren Atlas.*

Beweis: Für (1) sei $p \in M$ und V offene Umgebung von p. Wähle eine Karte $\varphi : U \to \varphi(U)$ mit $p \in U$, und setze $\varphi(p) = x$. Es gibt dann ein $\delta > 0$ mit $\overline{B_\delta(x)} \subset \varphi(U)$. Dann ist

$$K_\delta = \varphi^{-1}(\overline{B_\delta(x)})$$

eine kompakte Umgebung von p, und für $\delta > 0$ klein ist $K_\delta \subset V$.

Nun zu Aussage (2). Sei V_i, $i \in \mathbb{N}$, eine Umgebungsbasis von M. Dann bilden die V_i mit $\overline{V_i}$ kompakt auch eine Umgebungsbasis: sei $\Omega \subset M$ offen. Zu $p \in \Omega$ wähle eine offene Umgebung V mit $\overline{V}$ kompakt. Es gibt dann ein $i \in I$ mit $p \in V_i \subset \Omega \cap V$, also $\overline{V_i}$ kompakt. Also können wir von vornherein annehmen, dass die V_i relativ kompakt sind, das heißt $\overline{V_i}$ ist kompakt. Definiere nun $G_1 = V_1$ und induktiv

$$G_k = \bigcup_{i=1}^{i_k} V_i \quad \text{ wobei } i_k = \min\{i > i_{k-1} : \overline{G}_{k-1} \subset \bigcup_{j=1}^{i} V_j\}.$$

Nach Heine-Borel wird $\overline{G}_{k-1}$ durch endlich viele V_i übedeckt, daher bricht das Verfahren nicht ab und Behauptung (2) ist gezeigt. Behauptung (3) folgt, da jedes $\overline{G}_k$ durch endlich viele Kartengebiete überdeckt wird nach Heine-Borel. □

Eine Folge x_k in einem topologischen Raum X konvergiert gegen ein $x \in X$, wenn es zu jeder offenen Umgebung U von x ein $K \in \mathbb{N}$ gibt, so dass $x_k \in U$ für alle $k > K$. Die Hausdorff-Eigenschaft garantiert, dass eine Folge höchstens einen Grenzwert hat. In einer C^0-Mannigfaltigkeit M kann die Konvergenz wie folgt beschrieben werden:

> Ist $\varphi : U \to \varphi(U)$ Karte mit $x \in U$, so konvergiert eine Folge x_k genau dann gegen x, wenn $x_k \in U$ für alle bis auf endlich viele k, und $\varphi(x_k) \to \varphi(x)$ in $\mathbb{R}^n$.

Ist $f : X \to Y$ stetig, so gilt die Implikation

$$x_k \to x \quad \Rightarrow \quad f(x_k) \to f(x).$$

Denn ist V offene Umgebung von $f(x)$, so ist $f^{-1}(V)$ offene Umgebung von x und damit $x_k \in f^{-1}(V)$, also $f(x_k) \in V$, für k hinreichend groß. Umgekehrt: ist $f : X \to Y$ in jedem Punkt $x \in X$ folgenstetig, so ist f stetig im Sinne von Definition 1.4 (ÜA). Im Kontext der Mannigfaltigkeiten ist also das Konzept der Folgenstetigkeit ausreichend.

Wie kommt nun die Differenzierbarkeit ins Spiel?

Definition 1.8 *Sei M ein topologischer Raum. Zwei n-dimensionale Karten $\varphi_i : U_i \to \varphi_i(U_i)$, $i = 1, 2$, heißen C^∞-kompatibel, wenn der Koordinatenwechsel*

$$\varphi_2 \circ \varphi_1^{-1}|_{\varphi_1(U_1 \cap U_2)} : \varphi_1(U_1 \cap U_2) \to \varphi_2(U_1 \cap U_2)$$

ein C^∞-Diffeomorphismus ist. Ein n-dimensionaler C^0-Atlas $\mathcal{A}$ heißt C^∞-Atlas (oder differenzierbar), wenn alle Karten in $\mathcal{A}$ C^∞-kompatibel sind.

Remark 1.2 Für jeden Kartenwechsel $\phi = \varphi_2 \circ \varphi_1^{-1}$ ist $D\phi(x) \in \mathrm{GL}_n(\mathbb{R})$. Denn ϕ hat die Umkehrabbildung $\psi = \varphi_1 \circ \varphi_2^{-1}$, und mit der Kettenregel folgt

$$D\psi(\phi(x))D\phi(x) = D(\psi \circ \phi)(x) = \mathrm{E}_n \quad \text{für alle } x \in \varphi_1(U_1 \cap U_2).$$

Example 1.9 Die n-dimensionale Sphäre $\mathbb{S}^n = \{p \in \mathbb{R}^{n+1} : |p| = 1\}$ ist das absolute Standardbeispiel einer n-dimensionalen Mannigfaltigkeit (außer natürlich $\mathbb{R}^n$). Wir betrachten den Atlas, der durch stereographische Projektion vom Nord- bzw. Südpol gegeben ist:

$$U_1 = \{(x,t) \in \mathbb{S}^n \subset \mathbb{R}^n \times \mathbb{R} : t < 1\}, \quad U_2 = \{(x,t) \in \mathbb{S}^n \subset \mathbb{R}^n \times \mathbb{R} : t > -1\}.$$

In Formeln lauten die beiden Abbildungen $\varphi_i : U_i \to \mathbb{R}^n$, $i = 1, 2$, wie folgt:

$$\begin{aligned}
\varphi_1(x,t) &= \frac{x}{1-t}, \qquad \varphi_1^{-1}(y) = \left(\frac{2y}{1+|y|^2}, -\frac{1-|y|^2}{1+|y|^2}\right);\\
\varphi_2(x,t) &= \frac{x}{1+t}, \qquad \varphi_2^{-1}(y) = \left(\frac{2y}{1+|y|^2}, \frac{1-|y|^2}{1+|y|^2}\right).
\end{aligned}$$

Es gilt $U_1 \cap U_2 = \{(x,t) \in \mathbb{S}^n \subset \mathbb{R}^n \times \mathbb{R} : -1 < t < 1\}$ und somit folgt

$$\varphi_1(U_1 \cap U_2) = \mathbb{R}^n \backslash \{0\} = \varphi_2(U_1 \cap U_2).$$

Als Koordinatenwechsel ergibt sich

$$\varphi_2 \circ \varphi_1^{-1} : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}^n \backslash \{0\}, \; \varphi_2 \circ \varphi_1^{-1}(y) = \frac{\dfrac{2y}{1+|y|^2}}{1 - \dfrac{1-|y|^2}{1+|y|^2}} = \frac{y}{|y|^2}.$$

Somit ist $\{\varphi_1, \varphi_2\}$ ein C^∞-Atlas auf $\mathbb{S}^n$.

Example 1.10 Sei $M = \{(x, y) \in \mathbb{R}^2 : |x| + |y| = 1\}$ das Quadrat aus Beispiel 1.7. Für die dort betrachteten $\varphi_{1,2}$ gilt $U_1 \cap U_2 = \{(x, y) \in M : x, y > 0\}$, und

$$\varphi_2 \circ \varphi_1^{-1} : (0, 1) \to (0, 1),\ \varphi_2(\varphi_1^{-1}(y)) = 1 - y.$$

Analog sind die anderen Koordinatenwechsel differenzierbar, es handelt sich also, etwas überraschend, um einen C^∞-Atlas.

Allgemein hat man in der Wahl der Karten diverse Möglichkeiten. Sind zwei Karten (V_1, ψ_1) und (V_2, ψ_2) mit allen Karten (U_i, φ_i) aus einem C^∞-Atlas $\mathcal{A}_0$ kompatibel, so sind sie auch untereinander kompatibel. Denn $V_1 \cap V_2 = \bigcup_{i \in I} V_1 \cap V_2 \cap U_i$, und

$$\psi_2 \circ \psi_1^{-1} = \psi_2 \circ \varphi_i^{-1}|_{\varphi_i(V_2 \cap U_i)} \circ (\varphi_i \circ \psi_1)|_{\psi_1(V_1 \cap U_i)} \quad \text{auf } V_1 \cap V_2 \cap U_i.$$

Definition 1.9 *Ein n-dimensionaler C^∞-Atlas $\mathcal{A}$ auf M heißt maximal, wenn jede n-dimensionale Karte, die mit den Karten in $\mathcal{A}$ C^∞-kompatibel ist, selbst schon zu $\mathcal{A}$ gehört. Ein solcher Atlas heißt auch C^∞-Struktur oder differenzierbare Struktur auf M.*

Zu einem gegebenen C^∞-Atlas $\mathcal{A}_0$ gibt es genau einen maximalen C^∞-Atlas $\mathcal{A}$, der $\mathcal{A}_0$ enthält. Und zwar nehmen wir einfach alle Karten hinzu, die mit $\mathcal{A}_0$ kompatibel sind. Wie bemerkt sind diese Karten auch miteinander kompatibel, wir bekommen also so einen differenzerbaren Atlas. Dieser ist maximal, denn jede mit $\mathcal{A}_0$ kompatible Karte ist ja dabei. Ist $\mathcal{A}'$ ein anderer maximaler Atlas, der $\mathcal{A}_0$ enthält, so ist jede Karte in $\mathcal{A}'$ mit $\mathcal{A}_0$ kompatibel, also gilt $\mathcal{A}' \subset \mathcal{A}$, und sogar Gleichheit wegen der Maximalität. Man muss hier sagen, es hat noch niemand einen maximalen Atlas konkret gesehen, es handelt sich um einen theoretischen Begriff, der aber nützlich ist.

Remark 1.3 Natürlich können auch C^k-Atlanten für $k \in \mathbb{N}$ betrachtet werden. Aber jeder maximale C^k-Atlas $\mathcal{A}$ enthält einen C^∞-Atlas $\mathcal{A}'$ †.

Definition 1.10 *Eine n-dimensionale C^∞-Mannigfaltigkeit (oder: differenzierbare Mannigfaltigkeit) ist ein Hausdorffraum mit abzählbarer Basis, zusammen mit einem n-dimensionalen, maximalen C^∞-Atlas auf M.*

Um eine differenzierbare Mannigfaltigkeit anzugeben, reicht wie gesagt die Angabe eines differenzierbaren Atlas.

Example 1.11 Ist M eine C^0-Mannigfaltigkeit und besteht der zugehörige Atlas $\mathcal{A}$ nur aus einer Karte $\varphi : M \to \Omega$, so ist M automatisch eine C^∞-Mannigfaltigkeit, denn die Bedingung aus Definition 1.8 ist leer. Zum Beispiel ist jede offene Menge $\Omega \subset \mathbb{R}^n$ in kanonischer Weise eine C^∞-Mannigfaltigkeit, als Karte wählt man $\mathrm{id} : \Omega \to \Omega$. Insbesondere ist $\mathbb{R}^n$ mit dieser Struktur eine differenzierbare Mannigfaltigkeit.

Example 1.12 Sind M und N differenzierbare Mannigfaltigkeiten der Dimension m bzw. n, so ist $M \times N$ kanonisch eine differenzierbare Mannigfaltigkeit der Dimension $m + n$, die Produktmannigfaltigkeit von M und N (Übungsaufgabe).

Der folgende Satz liefert ein grundsätzliches Verfahren, um auf einer Menge ohne gegebene Topologie die Struktur einer Mannigfaltigkeit zu definieren.

† H. Whitney, Annals of Math. **37** (1936), 645–680.

Theorem 1.2 *Sei M eine Menge und $\varphi_\lambda : U_\lambda \to V_\lambda$, $\lambda \in \Lambda$, ein System von bijektiven Abbildungen von $U_\lambda \subset M$ auf offene Mengen $V_\lambda \subset \mathbb{R}^n$. Es gelte:*

(1) $\bigcup_{\lambda\in\Lambda} U_\lambda = M$.

(2) $\varphi_\lambda(U_\lambda \cap U_\mu)$ *ist offen für alle* $\lambda, \mu \in \Lambda$.

(3) $\varphi_\mu \circ \varphi_\lambda^{-1} : \varphi_\lambda(U_\lambda \cap U_\mu) \to \varphi_\mu(U_\lambda \cap U_\mu)$ *ist stetig für alle* $\lambda, \mu \in \Lambda$.

Dann gibt es genau eine Topologie $\mathcal{O}$ auf M, so dass $\{\varphi_\lambda : \lambda \in \Lambda\}$ ein C^0-Atlas ist. Wird M durch eine abzählbare Teilfamilie $\{U_\lambda : \lambda \in \Lambda'\}$ überdeckt, so hat $\mathcal{O}$ abzählbare Basis.

Bemerkung: Um zu zeigen, dass M eine C^0-Mannigfaltigkeit ist, ist im Einzelfall noch die Hausdorff-Eigenschaft nachzuweisen!

Beweis: (von Satz 1.2) Wir definieren die gesuchte Topologie durch

$$\mathcal{O} = \{U \subset M : \varphi_\lambda(U \cap U_\lambda) \subset \mathbb{R}^n \text{ ist offen für alle } \lambda \in \Lambda\}.$$

Dann ist $\mathcal{O}$ eine Topologie, wie sich aus folgenden Tatsachen ergibt:

(i) $\varphi_\lambda(M \cap U_\lambda) = V_\lambda \subset \mathbb{R}^n$ ist offen nach Voraussetzung.

(ii) Für $W_\alpha \subset M$, $\alpha \in A$, gilt $\varphi_\lambda((\bigcup_{\alpha\in A} W_\alpha) \cap U_\lambda) = \bigcup_{\alpha\in A} \varphi_\lambda(W_\alpha \cap U_\lambda)$.

(iii) Für $W_i \subset M$, $i = 1, \ldots, N$, gilt $\varphi_\lambda((\bigcap_{i=1}^N W_i) \cap U_\lambda) = \bigcap_{i=1}^N \varphi_\lambda(W_i \cap U_\lambda)$.

Bezüglich $\mathcal{O}$ sind die Mengen U_λ offen nach Voraussetzung (2). Weiter ist $\varphi_\lambda : U_\lambda \to V_\lambda$ stetig, denn für $V \subset V_\lambda$ offen und $\mu \in \Lambda$ ist

$$\varphi_\mu(\varphi_\lambda^{-1}(V) \cap U_\mu) = \underbrace{(\varphi_\mu \circ \varphi_\lambda^{-1})}_{(\varphi_\lambda\circ\varphi_\mu^{-1})^{-1}}(V \cap \underbrace{\varphi_\lambda(U_\lambda \cap U_\mu)}_{\text{offen}})$$

offen. Schließlich ist $\varphi_\lambda^{-1} : V_\lambda \to U_\lambda$ nach Definition stetig, denn für $U \in \mathcal{O}$ mit $U \subset U_\lambda$ ist $\varphi_\lambda(U) = \varphi(U \cap U_\lambda)$ offen. Damit sind die $\varphi_\lambda : U_\lambda \to V_\lambda$ Homeomorphismen und bilden einen C^0-Atlas auf M.

Sei $\mathcal{O}'$ eine Topologie, so dass die $\varphi_\lambda : U_\lambda \to V_\lambda$ homeomorph sind. Für $U \in \mathcal{O}$ folgt dann

$$U = \bigcup_{\lambda\in\Lambda} \varphi_\lambda^{-1}(\underbrace{\varphi_\lambda(U \cap U_\lambda)}_{\text{offen}}) \in \mathcal{O}' \quad \Rightarrow \quad \mathcal{O} \subset \mathcal{O}'.$$

Ist umgekehrt $U \in \mathcal{O}'$, so ist $U \cap U_\lambda \in \mathcal{O}'$ und $\varphi_\lambda(U \cap U_\lambda)$ ist offen, für alle $\lambda \in \Lambda$. Dies bedeutet $U \in \mathcal{O}$ bzw. $\mathcal{O}' \subset \mathcal{O}$.

Schließlich gelte $M = \bigcup_{\lambda\in\Lambda'} U_\lambda$ mit Λ' abzählbar. Jedes V_λ hat eine abzählbare Basis $\{V_{\lambda,i} : i \in \mathbb{N}\}$ bezüglich der Topologie auf $\mathbb{R}^n$. Dann ist $\{U_{\lambda,i} = \varphi_\lambda^{-1}(V_{\lambda,i}) : i \in \mathbb{N}\}$ eine abzählbare Basis für U_λ und somit $\{U_{\lambda,i} : \lambda \in \Lambda', i \in \mathbb{N}\}$ abzählbare Basis der Topologie $\mathcal{O}$ auf M. $\square$

Definition 1.11 *Ein topologischer Raum X heißt zusammenhängend, wenn für $A \subset X$ folgende Implikation gilt:*

$$A \text{ offen und abgeschlossen, } A \neq \emptyset \quad \Rightarrow \quad A = X.$$

Alternativ: jede lokal konstante Funktion auf X ist konstant.

Example 1.13 Das Intervall $[0,1] \subset \mathbb{R}$ ist zusammenhängend. Denn sei $A \subset [0,1]$ offen, abgeschlossen und nichtleer. Dann folgt für die charakteristische Funktion $\chi'_A \equiv 0$, also $\chi_A \equiv 1$ bzw. $A = X$. //

Definition 1.12 *Sei X topologischer Raum, und $x_0, x_1 \in X$. Ein Weg von x_0 nach x_1 ist eine Abbildung $\gamma \in C^0([0,1], X)$ mit $\gamma(0) = x_0$, $\gamma(1) = x_1$. Wenn es einen solchen Weg gibt, so heißen x_0, x_1 verbindbar. X heißt wegzusammenhängend, wenn je zwei Punkte $x_0, x_1 \in X$ verbindbar sind.*

Lemma 1.1 *Sei X ein topologischer Raum.*

(i) *„verbindbar“ ist Äquivalenzrelation. Die Äquivalenzklassen heißen Wegkomponenten.*

(ii) *X wegzusammenhängend $\Rightarrow$ X zusammenhängend.*

Beweis: Für (i) prüfen wir die drei Eigenschaften:

$x \sim x$: wähle $\gamma(t) \equiv x$ für alle $t \in [0,1]$.

$x_0 \sim x_1 \Rightarrow x_1 \sim x_0$: sei γ Weg von x_0 nach x_1; wähle $\sigma(t) = \gamma(1-t)$ für $t \in [0,1]$.

$x_0 \sim x_1$, $x_1 \sim x_2 \Rightarrow x_0 \sim x_2$: Seien γ_1, γ_2 Wege von x_0 nach x_1 bzw. x_1 nach x_2. Wähle

$$\gamma(t) = \begin{cases} \gamma_1(2t) & \text{für } 0 \leq t \leq \frac{1}{2} \\ \gamma_2(2t-1) & \text{für } \frac{1}{2} \leq t \leq 1. \end{cases}$$

Für (ii) sei $A \subset X$ offen und abgeschlossen, und $x_0 \in A$. Nach Vorassetzung gibt es zu $x \in X$ einen Weg γ von x_0 nach x. Dann ist $\gamma^{-1}(A) \subset [0,1]$ offen und abgeschlossen, sowie $0 \in \gamma^{-1}(A)$. Nach Beispiel 1.13 ist dann $x = \gamma(1) \in A$, also $A = X$. □

Theorem 1.3 *Sei M eine n-dimensionale topologische Mannigfaltigkeit. Dann gilt:*

$$M \quad \text{wegzusammenhängend} \quad \Leftrightarrow \quad M \quad \text{zusammenhängend.}$$

Beweis: Es ist nur noch die Aussage „$\Leftarrow$“ zu zeigen. Der Beweis beruht darauf, dass M *lokal* wegweise zusammenhängend ist. Wähle nämlich zu $x \in M$ eine Karte $\varphi : U \to V$ mit $x \in U$. Für $r > 0$ hinreichend klein gilt $B_r(\varphi(x)) \subset V$, und $U_x := \varphi^{-1}\big(B_r(\varphi(x))\big)$ ist wegweise zusammenhängend.

Sei nun $x_0 \in M$ beliebig und A die Wegkomponente von x_0. Für $x \in X$ gilt entweder $U_x \subset A$ oder $U_x \subset X \backslash A$. Somit sind A und $X \backslash A$ offen. Wegen $x_0 \in A$ folgt $A = M$, also ist M wegzusammenhängend. □

Example 1.14 Folgende Menge $M \subset \mathbb{R}^2$ ist zusammenhängend, aber nicht wegweise:

$$M = \{(x, \sin\frac{1}{x}) : 0 < x < 1\} \cup \{(0,y) : -1 \leq y \leq 1\}.$$

2 Abbildungen und Überlagerungen

Die Stetigkeit von Abbildungen zwischen differenzerbaren Mannigfaltigkeiten ist definiert, sie sind ja insbesondere topologische Räume. Jetzt geht es um C^k Abbildungen.

Definition 2.1 *Seien M, N differenzierbare Mannigfaltigkeiten der Dimension m bzw. n. Eine Abbildung $f : M \to N$ heißt k-mal stetig differenzierbar, wenn es zu jedem $p \in M$ Karten $\varphi : U \to \varphi(U)$ und $\psi : V \to \psi(V)$ von M bzw. N gibt mit*

(a) $p \in U$ *und* $f(U) \subset V$,

(b $\psi \circ f \circ \varphi^{-1} \in C^k(\varphi(U), \mathbb{R}^n)$.

Wir bezeichnen diese Abbildungen mit $C^k(M, N)$, im Fall $N = \mathbb{R}$ mit $C^k(M)$.

Hier zwei Anmerkungen zur Definition:

- Ist $f : M \to N$ stetig und $V \subset N$ offene Umgebung von $f(p)$, so hat p eine offene Umgebung U mit $f(U) \subset V$, zum Beispiel $U = f^{-1}(V)$.
- Für $k = 0$ gibt diese Definition eine äquivalente Fassung der Stetigkeit.
- Seien $\varphi' : U' \to \varphi(U')$ bzw. $\psi' : V' \to \psi(V')$ irgendwelche Karten von M bzw. N. mit $f(U') \subset V'$. Dann folgt

$$\psi' \circ f \circ (\varphi')^{-1} \in C^k(\varphi'(U'), \mathbb{R}^n).$$

  Zu $p \in U'$ wähle nämlich $\varphi : U \to \varphi(U)$, $\psi : V \to \psi(V)$ wie in Definition 2.1. Die Behauptung ergibt sich aus folgendem Diagramm:

  Statt *zu jedem p gibt es Karten mit ...* könnten wir somit in der Definition äquivalent schreiben: *für alle Karten mit ....*

Example 2.1 Betrachte $\mathbb{S}^n$ mit den Karten φ_1, φ_2 aus Beispiel 1.9 (stereographische Projektion von Nordpol bzw. Südpol). Für $\lambda > 0$ sei

$$\tau_\lambda : \mathbb{R}^n \to \mathbb{R}^n, \quad \tau_\lambda(x) = \lambda\, x.$$

Definiere nun $\sigma_\lambda : \mathbb{S}^n \to \mathbb{S}^n$ wie folgt:

$$\sigma_\lambda(p) = \begin{cases} \big(\varphi_1^{-1} \circ \tau_\lambda \circ \varphi_1\big)(p) & \text{falls } p \neq (0,1) \\ (0,1) & \text{falls } p = (0,1). \end{cases}$$

Wir zeigen, dass σ_λ eine C^∞-Abbildung ist: für $\varphi_1 : U_1 \to \mathbb{R}^n$ gilt $\sigma_\lambda(U_1) \subset U_1$ und

$$\varphi_1 \circ \sigma_\lambda \circ \varphi_1^{-1} = \tau_\lambda \in C^\infty(\mathbb{R}^n, \mathbb{R}^n).$$

Für $\varphi_2 : U_2 \to \mathbb{R}^n$ gilt $\sigma_\lambda(U_2) \subset U_2$ und

$$\varphi_2 \circ \sigma_\lambda \circ \varphi_2^{-1} = \begin{cases} 0 & x = 0 \\ (\varphi_2 \circ \varphi_1)^{-1} \circ \tau_\lambda \circ (\varphi_2 \circ \varphi_1)^{-1} & x \neq 0 \end{cases}$$

Nach Beispiel 1.9 ist $(\varphi_2 \circ \varphi_1^{-1})(x) = x/|x|^2$, also folgt für $x \neq 0$

$$(\varphi_2 \circ \sigma_\lambda \circ \varphi_2^{-1})(x) = \frac{\tau_\lambda\left(x/|x|^2\right)}{\left|\tau_\lambda\left(x/|x|^2\right)\right|^2} = \frac{1}{\lambda}x.$$

Insgesamt folgt $\varphi_2 \circ \sigma_\lambda \circ \varphi_2^{-1}(x) = \frac{1}{\lambda}x$ für alle $x \in \mathbb{R}^n$, insbesondere für $x = 0$, also ist $\varphi_2 \circ \sigma_\lambda \circ \varphi_2^{-1} \in C^\infty$.

Definition 2.2 *Eine Abbildung $f : M \to N$ zwischen differenzierbaren Mannigfaltigkeiten heißt Diffeomorphismus, falls f bijektiv ist, und f, f^{-1} von der Klasse C^∞ sind.*

Example 2.2 Die Abbildung $\sigma_\lambda : \mathbb{S}^n \to \mathbb{S}^n$ ist ein Diffeomorphismus, denn $\sigma_\lambda^{-1} = \sigma_{1/\lambda}$.

Example 2.3 Betrachte auf $\mathbb{R}$ die beiden Atlanten

$$\begin{aligned} \mathcal{A}_1 &= \{\varphi_1 : \mathbb{R} \to \mathbb{R}, \quad \varphi_1(x) = x\} \\ \mathcal{A}_2 &= \{\varphi_2 : \mathbb{R} \to \mathbb{R}, \quad \varphi_2(x) = x^3\}. \end{aligned}$$

$\mathcal{A}_1$ und $\mathcal{A}_2$ definieren verschiedene differenzierbare Strukturen auf $\mathbb{R}$, denn die Karten sind nicht einmal C^1-kompatibel (geschweige denn C^∞):

$$\begin{aligned} \varphi_2 \circ \varphi_1^{-1} : &\quad \mathbb{R} \to \mathbb{R}, \quad s \mapsto s^3 \\ \varphi_1 \circ \varphi_2^{-1} : &\quad \mathbb{R} \to \mathbb{R}, \quad t \mapsto \begin{cases} \sqrt[3]{t} & t \geq 0 \\ -\sqrt[3]{-t} & t < 0 \end{cases}. \end{aligned}$$

$\varphi_1 \circ \varphi_2^{-1}$ ist nicht differenzierbar in $t = 0$. Die Abbildung

$$f : (\mathbb{R}, \mathcal{A}_1) \to (\mathbb{R}, \mathcal{A}_2), \quad f(x) = \begin{cases} \sqrt[3]{x} & x \geq 0 \\ -\sqrt[3]{-x} & x < 0 \end{cases}$$

ist aber ein Diffeomorphismus! Denn

$$\begin{aligned} \varphi_2 \circ f \circ \varphi_1^{-1} &= \mathrm{id}_{\mathbb{R}} \in C^\infty(\mathbb{R}, \mathbb{R}), \\ \varphi_1 \circ f^{-1} \circ \varphi_2^{-1} &= \mathrm{id}_{\mathbb{R}} \in C^\infty(\mathbb{R}, \mathbb{R}). \end{aligned}$$

Auf der Menge der differenzierbaren Mannigfaltigkeiten (mit maximalen Atlanten) ist

$$M \sim N \quad \Leftrightarrow \quad M \text{ diffeomorph } N$$

eine Äquivalenzrelation; ihre Äquivalenzklassen heißen Diffeomorphietypen. Natürlich ist jede differenzierbare Mannigfaltigkeit insbesondere eine topologische Mannigfaltigkeit. Es ist eine sehr schwierige Frage, wieviele verschiedene Mannigfaltigkeiten – im Sinn von nicht homeomorph oder nicht diffeomorph - es in der jeweiligen Dimension gibt. Hier nur einige Ausblicke ohne Beweis, wobei wir stets von zusammenhängenden Mannigfaltigkeiten ausgehen:

- Es gibt topologische Mannigfaltigkeiten, die keinen differenzierbaren Atlas zulassen. Das erste Beispiel in Dimension 10 stammt von Kervaire (1960), durch Konstruktion einer topologischen Invariante. Donaldson fand 1984 Beispiele in Dimension 4, sein Beweis beruht auf der Analysis einer nichtlinearen elliptischen Differentialgleichung, der Yang-Mills Gleichung (Fields-Medaille 1986).

- Es gibt topologische Mannigfaltigkeiten, die mehrere nicht diffeomorphe differenzierbare Strukturen besitzen. So gibt es auf der 7-dimensionalen Sphäre 28 verschiedene differenzierbare Strukturen (Milnor 1962, Fields Medaille 1964). Und 1984 zeigte Gompf (aufbauend auf den Resultaten von Donaldson) die Existenz überabzählbar vieler, nicht diffeomorpher C^∞-Strukturen auf $\mathbb{R}^4$!

- Die Kreislinie $\mathbb{S}^1$ ist die einzige kompakte, eindimensionale Mannigfaltigkeit, die einzige nichtkompakte ist $\mathbb{R}$, das ist also relativ einfach. In Dimension 2 hat man eine vollständige Liste der kompakten orientierbaren Flächen: es sind die Sphäre, der Torus, die Brezel (2 Löcher), und so weiter. Die Brezel kann durch Zusammensetzen von 2 Tori konstruiert werden, man entfernt jeweils eine kleine Kreisscheibe und verbindet durch einen Zylinder. Analog kann eine Fläche mit p Löchern aus p Tori gebaut werden. Für die kompakten, nicht orientierbaren Flächen gibt es eine ganz analoge Beschreibung. Diese Klassifikation beruht auf komplexer Analysis, dem Uniformisierungssatz von Poincaré und Koebe (1883/1907). In Dimension 1 und 2 gibt es keine exotischen Beispiele wie oben, das heißt topologisch oder differenzierbar macht keinen Unterschied. Dasselbe gilt übrigens auch in Dimension 3.

- Die Poincaré-Vermutung besagt dass jede 3-dimensionale kompakte, einfach zusammenhängende Mannigfaltigkeit diffeomorph zur 3-Sphäre ist. Das wurde von Perelman 2003 bewiesen (Fields Medaille 2006, abgelehnt). Zum Beweis verwendet er eine Deformation der Mannigfaltigkeit durch den Ricci-Fluss, eine nichtlineare parabolische Differentialgleichung. Er zeigt auch eine allgemeine Klassifikation von kompakten 3-Mannigfaltigkeiten, die Vermutung von Thurston.

- In Dimension $n \geq 4$ gibt es eine verallgemeinerte Poincaré-Vermutung: grob gesagt soll eine Mannigfaltigkeit mit den Homotopieeigenschaften einer n-Sphäre bereits homeomorph zur n-Sphäre sein. Für $n \geq 5$ wurde das von Smale 1960 gezeigt (Fields-Medaille 1966), den Fall $n = 4$ hat Friedman 1982 erledigt (Fields-Medaille 1986). Diese Beweise sind topologischer Natur, während für $n = 3$ die Analysis entscheidend war.

Soweit die Highlights des 20. Jahrhuinderts, nun aber zurück zu den Niederungen der Vorlesung. Eine Verkettung von C^k-Abbildungen ist wieder eine C^k-Abbildung, das heißt

$$f \in C^k(M, N),\, g \in C^k(N, P) \quad \Rightarrow \quad g \circ f \in C^k(M, P). \tag{2.1}$$

Um das zu etwas genauer zu zeigen, wähle zu $p \in M$ Karten

$$\begin{array}{lll} \varphi : U \to \varphi(U) & \text{mit} & p \in U, \\ \psi : V \to \psi(V) & \text{mit} & f(p) \in V, \\ \chi : W \to \chi(W) & \text{mit} & g(f(p)) \in W. \end{array}$$

Wir können $f(U) \subset V$ und $g(V) \subset W$ annehmen, sonst ersetze erst V durch $V \cap g^{-1}(W)$ und dann U durch $U \cap f^{-1}(V)$. Unsere Behauptung (2.1) folgt nun aus der C^k-Kettenregel:

$$\chi \circ (g \circ f) \circ \varphi^{-1} = (\chi \circ g \circ \psi^{-1}) \circ (\psi \circ f \circ \varphi^{-1}) \in C^k(\varphi(U), \mathbb{R}^p) \quad \text{wobei } p = \dim P.$$

Theorem 2.1 *Die Menge* $\mathrm{Diff}(M)$ *der* C^∞*-Diffeomorphismen einer Mannigfaltigkeit* M *ist eine Untergruppe der Bijektionen von* M*.*

BEWEIS: Für $f, g \in \mathrm{Diff}(M)$ ist $g^{-1} \circ f \in \mathrm{Diff}(M)$ wegen (2.1). $\square$

Die Homeomorphismen und die C^k-Diffeomorphismen, $k \in \mathbb{N}$, bilden auch eine Gruppe.

Definition 2.3 *Sei* Γ *eine Untergruppe von* $\mathrm{Diff}(M)$*. Dann ist*

$$p \sim q \quad \Leftrightarrow \quad \exists\, g \in \Gamma \text{ mit } g(p) = q$$

eine Äquivalenzrelation auf M*. Sei* $\pi : M \to M/\!\sim$*,* $\pi(p) = [p]$*, die Projektion auf die Menge der Äquivalenzklassen. Wir betrachten auf* $M/\!\sim$ *die Quotiententopologie*

$$U \subset M/\!\sim \text{ offen} \quad \Leftrightarrow \quad \pi^{-1}(U) \subset M \text{ offen.}$$

Bezeichnungen: Die Äquivalenzklasse $[\,p\,] = \{g(p) : g \in \Gamma\}$ wird mit $\Gamma \cdot p$ bezeichnet, sie heißt Bahn von p. Der Quotient $M/\!\sim$ heißt Bahnenraum, die übliche Notation ist M/Γ.

Dass die Relation $\sim$ eine Äquivalenzrelation ist, sieht man leicht: es gilt $\mathrm{id_M}(p) = p$, aus $g(p) = q$ folgt $g^{-1}(q) = p$, und mit $g(p) = q$, $h(q) = r$ ist $(h \circ g)(p) = r$. Dabei ist $\mathrm{id_M} \in \Gamma$, und mit $g, h \in \Gamma$ sind auch $g^{-1}, h \circ g \in \Gamma$. Auch die Eigenschaften der Topologie folgen direkt, es gilt

$$\pi^{-1}(\emptyset) = \emptyset, \quad \pi^{-1}\Big(\bigcup_{i \in I} U_i\Big) = \bigcup_{i \in I} \pi^{-1}(U_i), \quad \pi^{-1}\Big(\bigcap_{i=1}^{N} U_i\Big) = \bigcap_{i=1}^{N} \pi^{-1}(U_i).$$

Nach Definition der Topologie ist die Abbildung $\pi : M \to M/\Gamma$ stetig. Sie bildet außerdem offene Mengen in offene Mengen ab, denn für $V \subset M$ offen gilt

$$\pi^{-1}(\pi(V)) = \bigcup_{g \in \Gamma} g(V) = \text{ offen in } M.$$

Sei V_i, $i \in I$, abzählbare Basis von M. Dann bilden die $U_i = \pi(V_i)$ eine abzählbare Basis von M/Γ. Denn für $U \subset M/\Gamma$ offen ist $\pi^{-1}(U)$ Vereinigung von Mengen V_i, $i \in I'$, und es folgt

$$U = \pi(\pi^{-1}(U)) = \pi\Big(\bigcup_{i \in I'} V_i\Big) = \bigcup_{i \in I'} \pi(V_i) = \bigcup_{i \in I'} U_i.$$

Die Frage ist nun, ob M/Γ eine Mannigfaltigkeit ist. Die folgenden zwei Beispiele zeigen, dass das ohne weiteres nicht stimmt.

Example 2.4 Die Gruppe $\Gamma = \pm 1$ operiert durch Multiplikation auf $\mathbb{R}$, aber $\mathbb{R}/\Gamma$ ist keine eindimensionale C^0-Mannigfaltigkeit. Denn sei $\varphi : U \to \varphi(U)$ eine beliebige Karte mit $[0] \in U$. Es gibt dann ein $\varepsilon > 0$ mit $(-\varepsilon, \varepsilon) \subset \pi^{-1}(U)$ bzw. $(-\varepsilon, \varepsilon)/\pm \subset U$. Durch Einschränkung haben wir den Homeomorphismus $\varphi : (-\varepsilon, \varepsilon)/\pm \to V$, für $V \subset \mathbb{R}$ offen. Setze $x = \varphi([0])$. Die Abbildung $\varphi \circ \pi : (0, \varepsilon) \to V \backslash \{x\}$ ist stetig und surjektiv, also ist $V \backslash \{x\}$ zusammenhängend, ein Widerspruch.

Example 2.5 Die Gruppe $\Gamma = (\mathbb{Q}, +)$ operiert durch Addition auf $\mathbb{R}$. Der topologische Raum $\mathbb{R}/\mathbb{Q}$ ist aber nicht Hausdorffsch: seien $U_i \subset \mathbb{R}/\mathbb{Q}$ offene Umgebungen von $\pi(x_i)$ für $i = 1, 2$. Dann sind $V_i = \pi^{-1}(U_i)$ offene Umgebungen von $x_{1,2}$. Da $\mathbb{Q}$ dicht, gibt es $q_k \in \mathbb{Q}$ mit $q_k \to x_2 - x_1$. Es folgt $U_1 \ni \pi(x_1) = \pi(x_1 + q_k) \in \pi(V_2) = U_2$ für k hinreichend groß.

Um diese Probleme auszuschliessen, führen wir folgende Bedingung ein.

Definition 2.4 *Die Untergruppe $\Gamma \subset \mathrm{Diff}(M)$ operiert eigentlich diskontinuierlich auf M, falls folgende zwei Bedingungen gelten:*

(i) *Zu $p \in M$ gibt es eine offene Umgebung U mit $g(U) \cap U = \emptyset$ für alle $g \in \Gamma \backslash \{\mathrm{id}_M\}$.*

(ii) *Zu $p, q \in M$ mit $q \notin \Gamma \cdot p$ gibt es offene Umgebungen U, V mit $V \cap \bigcup_{g \in \Gamma} g(U) = \emptyset$.*

Aus Bedingung (i) folgt insbesondere, dass Γ (Fixpunkt-)frei operiert, das heißt für alle $p \in M$ und alle $g \in \Gamma \backslash \{\mathrm{id}_M\}$ gilt $g(p) \neq p$. Der Begriff *eigentlich diskontinuierlich* ist in der Literatur nicht einheitlich, oft wird nur Bedingung (i) verlangt. Wir werden unten klar sehen, welche Rolle (i) und (ii) spielen. Vorher betrachten wir den Spezialfall, dass Γ bezüglich einer Metrik $d(\,\cdot\,,\,\cdot\,)$ isometrisch operiert, also

$$d\big(g(x), g(y)\big) = d(x, y) \quad \text{für alle } x, y \in M,\ g \in \Gamma.$$

Bedingung (i) ist dann erfüllt, wenn

$$\varrho(p) = \inf\{d(g(p), p) : g \in \Gamma \backslash \{\mathrm{id}_M\}\} > 0 \quad \text{für alle } p \in M. \tag{2.2}$$

Denn sei $U = B_{\frac{\varrho}{2}}(p)$ mit $\varrho = \varrho(p)$. Dann ist $g(U) = B_{\frac{\varrho}{2}}(g(p))$, und für $q \in g(U) \cap U$ folgt

$$d(g(p), p) \leq d(g(p), q) + d(q, p) < \frac{\varrho}{2} + \frac{\varrho}{2} = \varrho, \quad \text{also } g = \mathrm{id}_M.$$

Weiter ergibt sich Bedingung (ii) aus der Forderung

$$\delta(p, q) = \inf\{d\big(g(p), h(q)\big) : g, h \in \Gamma\} > 0 \quad \text{falls } \Gamma \cdot p \neq \Gamma \cdot q. \tag{2.3}$$

Man bezeichnet $\delta(p, q)$ auch als Bahnenabstand von p und q, offenbar gilt $\delta(p', q') = \delta(p, q)$ für alle $p' \in \Gamma \cdot p$, $q' \in \Gamma \cdot q$. Wir wählen als Umgebungen $U = B_{\delta/2}(p)$ und $V = B_{\delta/2}(q)$ mit $\delta = \delta(p, q) > 0$. Angenommen es gibt $x \in g(U) \cap h(V) = B_{\delta/2}(g(p)) \cap B_{\delta/2}(h(q))$, dann folgt

$$d\big(g(p), h(q)\big) \leq d(g(p), x) + d(x, h(q)) < \frac{\delta}{2} + \frac{\delta}{2} = \delta, \quad \text{Widerspruch.}$$

Als Anwendung betrachten wir zwei Beispiele.

Example 2.6 Die Gruppe $\Gamma = \{\pm\mathrm{id}\}$ operiert isometrisch auf $\mathbb{S}^n$ bezüglich des Euklidischen Abstands $d(p, q) = |p - q|$. Es gilt für alle $p, q \in \mathbb{S}^n$

$$\varrho(p) = d(-p, p) = 2 \quad \text{und} \quad \delta(p, q) = \min\big(|p - q|, |p - (-q)|\big).$$

Die Bedingungen (2.2) und (2.3) sind also erfüllt. Allgemein gilt das immer, wenn eine endliche Gruppe (Fixpunkt-)frei operiert.

Example 2.7 Betrachte auf $\mathbb{R}^n$ die Opeation von $\Gamma = \mathbb{Z}^n$ durch Translationen:

$$\tau_k(p) = p + k \quad \text{für } p \in \mathbb{R}^n,\ k \in \mathbb{Z}^n.$$

Es ist hier praktisch, auf $\mathbb{R}^n$ die Maximumsnorm $\|p\| = \max_{1 \leq i \leq n} |p^i|$ zu wählen. Die Operation ist bezüglich der zugehörigen Metrik isometrisch, denn es gilt

$$d(\tau_k(p), \tau_k(q)) = \|(p + k) - (q + k)\| = \|p - q\| = d(p, q).$$

Es gilt $\varrho(p) = \inf\{\|(p+k)-p\| : k \in \mathbb{Z}^n\backslash\{0\}\} = 1$. Weiter folgt für $q - p \notin \mathbb{Z}^n$ mit einer einfachen Kompaktheitsüberlegung

$$\delta(p,q) = \inf\{|q-p+k| : k \in \mathbb{Z}^n\} > 0.$$

Somit ist auch diese Operation eigentlich diskontinuierlich im Sinne von Definition 2.4.

Definition 2.5 *Eine stetige und surjektive Abbildung $\pi : M \to X$ zwischen topologischen Räumen heißt Überlagerung, falls es zu jedem $x \in X$ eine Umgebung U gibt, so dass*

$$\pi^{-1}(U) = \bigcup_{i \in I} V_i$$

mit $V_i \subset M$ offen, paarweise disjunkt und $\pi|_{V_i} : V_i \to U$ homeomorph.

Ein solches U heißt Umgebung mit Überlagerungseigenschaft. Finden Sie für die vorigen zwei Beispiele eine solche Umgebung.

Example 2.8 Standardbeispiel einer Überlagerung ist $c : \mathbb{R} \to \mathbb{S}^1$, $c(s) = e^{is}$. Nach Analysis 1 ist die Abbildung stetig und surjektiv. Zu $\omega = e^{it}$ wähle die offene Umgebung

$$U = \{z \in \mathbb{S}^1 : \langle z, \omega \rangle > 0\}.$$

Es gilt $\langle e^{is}, e^{it} \rangle = \cos(s-t)$, also ist $e^{is} \in U$ genau wenn $\cos(s-t) > 0$. Somit ist $c^{-1}(U)$ die disjunkte Vereinigung der offenen Intervalle I_k, $k \in \mathbb{Z}$, mit

$$I_k = \left\{2\pi k + t + \alpha : -\frac{\pi}{2} < \alpha < \frac{\pi}{2}\right\}.$$

Wir behaupten, dass $c|_{I_k}$ folgende Umkehrfunktion hat:

$$\varphi_k : U \to I_k,\ \varphi_k(z) = 2\pi k + t + \arcsin\langle z, i\omega \rangle \in I_k.$$

Denn für $z = e^{is}$, $s = 2\pi k + t + \alpha$ mit $-\frac{\pi}{2} < \alpha < \frac{\pi}{2}$, ist $\langle z, i\omega \rangle = \sin(s-t) = \sin\alpha$, also

$$\varphi_k(c(s)) = 2\pi k + t + \arcsin(\sin\alpha) = 2\pi k + t + \alpha = s.$$

φ_k ist stetig und damit $\pi|_{I_k}$ homeomorph.

Remark 2.1 Sei $\pi : M \to X$ Überlagerung und X zusammenhängend. Dann ist $\operatorname{card} \pi^{-1}\{x\} \in \mathbb{N} \cup \{\infty\}$ für alle $x \in X$ gleich, diese Zahl heißt Grad der Überlagerung. Denn ist U Umgebung von $x \in X$ mit $\pi^{-1}(U) = \bigcup_{i \in I} V_i$ wie in Definition 2.5, so gilt

$$\operatorname{card} \pi^{-1}\{x'\} = \operatorname{card} I = \operatorname{card} \pi^{-1}\{x\} \quad \text{für alle } x' \in U.$$

Für $k \in \mathbb{N} \cup \{\infty\}$ ist daher $X_k = \{x \in X : \operatorname{card} \pi^{-1}\{x\} = k\}$ offen. Wegen $X \backslash X_k = \bigcup_{\ell \neq k} X_\ell$ ist X_k auch abgeschlossen. Wählen wir k mit $X_k \neq \emptyset$ so folgt $X = X_k$.

Theorem 2.2 (diskrete Quotienten) *Sei M eine m-dimensionale Mannigfaltigkeit, auf der die Gruppe $\Gamma \subset \operatorname{Diff}(M)$ eigentlich diskontinuierlich operiert. Dann gelten folgende Aussagen:*

(i) *M/Γ ist Hausdorffsch mit abzählbarer Basis, und $\pi : M \to M/\Gamma$ ist Überlagerung.*

(ii) *Es gibt genau eine differenzierbare Struktur auf M/Γ, so dass π lokal diffeomorph ist.*

Beweis: Es ist schon klar, dass $X = M/\Gamma$ abzählbare Basis hat. Nach Voraussetzung hat jedes $p \in M$ eine offene Umgebung U mit $g(U) \cap U = \emptyset$ für alle $g \in \Gamma \backslash \{\mathrm{id}_M\}$. Es folgt

$$\pi^{-1}\big(\pi(U)\big) = \bigcup_{g \in \Gamma} g(U).$$

Aus $p \in g_1(U) \cap g_2(U)$ folgt $g_1^{-1}(p) \in U \cap (g_1^{-1} \circ g_2)(U)$, also $g_1^{-1} \circ g_2 = \mathrm{id}_M$ bzw. $g_2 = g_1$. Somit ist die Vereinigung disjunkt. Die Mengen $g(U)$ sind offen, und wie gezeigt ist $\pi : g(U) \to \pi(U)$ stetig und offen. Aus $\pi(q_1) = \pi(q_2)$ für $q_i \in g(U)$ folgt $\pi(p_1) = \pi(p_2)$ für $p_i = g^{-1}(q_i) \in U$. Es gibt dann $g' \in \Gamma$ mit $g'(p_1) = p_2$. Also folgt $g'(U) \cap U \neq \emptyset$, und damit $g' = \mathrm{id}_M$ und weiter $p_1 = p_2$ bzw. $q_1 = q_2$. Somit ist $\pi : g(U) \to \pi(U)$ bijektiv, und Behauptung (i) ist bewiesen.

Zum Beweis der Hausdorff-Eigenschaft seien $p, q \in M$ mit $\pi(p) \neq \pi(q)$. Für U, V wie in Definition 2.3(ii) sind $\pi(U)$, $\pi(V)$ offene Umgebungen der Punkte $\pi(p), \pi(q) \in X$, da π offen ist. Wäre $\pi(U) \cap \pi(V) \neq \emptyset$, so gibt es $p' \in U$, $q' \in V$ mit $\pi(p') = \pi(q')$, also $g(p') = q'$ für ein $g \in \Gamma$. Aber $g(U) \cap V = \emptyset$ nach Wahl von U, V, ein Widerspruch.

Als nächstes konstruieren wir die differenzierbare Struktur. Wie bereits gezeigt, gibt es zu $p \in M$ eine Umgebung U_p mit folgenden Eigenschaften:

- $g(U_p) \cap U_p = \emptyset$ für alle $g \in \Gamma \backslash \{\mathrm{id}_M\}$
- es gibt eine Karte $\psi_p : U_p \to \psi_p(U_p) \subset \mathbb{R}^m$.

Definiere nun auf X die Karten

$$\varphi_p = \psi_p \circ (\pi|_{U_p})^{-1} : \pi(U_p) \to \psi_p(U_p) \subset \mathbb{R}^m.$$

Die φ_p sind homeomorph, zu zeigen bleibt die Differenzierbarkeit der Kartenwechsel. Sei $x = \pi(p') = \pi(q')$ mit $p' \in U_p$, $q' \in U_q$; wir zeigen dass der Kartenwechsel auf einer Umgebung von $\varphi_p(x) = \psi_p(p')$ differenzierbar ist. Es gilt $q' = g(p')$ für ein $g \in \Gamma$, und

$$\pi|_{U_q} \circ g = \pi|_{U_p} \;\text{ bzw. }\; g = \big(\pi|_{U_q}\big)^{-1} \circ \pi|_{U_p} \qquad \text{auf } U_p \cap g^{-1}(U_q).$$

Daraus folgt auf der Umgebung $\psi_p\big(U_p \cap g^{-1}(U_q)\big) \ni \psi_p(p')$

$$\varphi_q \circ \varphi_p^{-1} = \big(\psi_q \circ (\pi|_{U_q})^{-1}\big) \circ \big(\psi_p \circ (\pi|_{U_p})^{-1}\big)^{-1} = \psi_q \circ g \circ \psi_p^{-1}.$$

Die rechte Seite ist die Koordinatendarstellung des Diffeomorphismus g, also in C^∞. Sei schließlich $\mathcal{A}$ ein Atlas auf X, so dass $\pi : M \to X$ lokal diffeomorph ist. Dann ist für jede Karte $\varphi : \Omega \to \varphi(\Omega)$ in $\mathcal{A}$ und jede Karte $\psi_p : U_p \to \psi_p(U_p)$ die zugehörige Koordinatendarstellung

$$\varphi \circ \pi \circ \psi_p^{-1} : \varphi_p\big(\pi(U_p) \cap \Omega\big) \to \varphi\big(\pi(U_p) \cap \Omega\big)$$

diffeomorph. Wegen $\pi \circ \psi_p^{-1} = \varphi_p^{-1}$ sind also die Kartenwechsel $\varphi \circ \varphi_p^{-1}$ differenzierbar, das heißt $\mathcal{A}$ definiert dieselbe differenzierbare Struktur. $\square$

Definition 2.6 *Sei X wegzusammenhängender topologischer Raum. Eine Schleife in X ist eine stetige Kurve $c : [0,1] \to X$, $c = c(s)$, mit $c(0) = c(1)$. Zwei Schleifen c_0, c_1 heißen homotop, wenn es ein $h \in C^0([0,1] \times [0,1], X)$, $h = h(s,t)$, gibt mit*

$$h(s,0) = c_0(s) \ \text{und} \ h(s,1) = c_1(s) \quad \text{für alle } s \in [0,1].$$

Ist dabei c_0 konstant, so heißt c_1 nullhomotop.

Definition 2.7 *Ein wegzusammenhängender topologischer Raum X heißt einfach zusammenhängend, wenn jede Schleife $c : [0,1] \to X$ nullhomotop ist.*

Example 2.9 Jede konvexe Teilmenge des $\mathbb{R}^n$ ist einfach zusammenhängend, und allgemeiner jede sternförmige Menge. Die Sphären $\mathbb{S}^n$ sind einfach zusammenhängend genau für $n \geq 2$.

Theorem 2.3 *Sei X eine zusammenhängende, m-dimensionale Mannigfaltigkeit. Dann existiert eine einfach zusammenhängende, m-dimensionale Mannigfaltigkeit M und eine Untergruppe $\Gamma \subset \mathrm{Diff}(M)$, die eigentlich diskontinuierlich operiert, so dass X diffeomorph ist zu M/Γ.*

Die Bedeutung dieses Satzes liegt darin, dass das Studium von Mannigfaltigkeiten in zwei Schritte zerlegt wird, nämlich erstens das Studium der einfach zusammenhängenden Mannigfaltigkeiten M und zweitens das Studium der eigentlich diskontinierlichen Gruppenoperationen auf solchen M. Zumindest ist das theoretisch so.

3 Tangentialvektoren und Differential

Die Differenzierbarkeit einer Abbildung $f : M \to N$ ist über ihre Koordinatendarstellungen definiert, unabhängig von der Wahl der der Koordinaten. Doch was ist die Ableitung von f? Hierzu müssen wir als erstes erklären, was der Tangentialraum von M im Punkt p sein soll. Es gibt drei verschiedene Zugänge:

a) geometrisch als Tangentenvektoren von Kurven

b) physikalisch mittels Koordinatenvektoren

c) algebraisch als Ableitungsoperatoren

Wir beginnen mit dem ersten Zugang. Die Idee ist, den Tangentialvektor durch Kurven zu repräsentieren. Dabei sind Kurven als äquivalent anzusehen, wenn ihre Koordinatendarstellung die gleiche Ableitung hat.

Definition 3.1 (Tangentialraum) *Sei M eine differenzierbare Mannigfaltigkeit. Ein Tangentialvektor im Punkt $p \in M$ ist eine Äquivalenzklasse $[c]$ von C^1-Kurven $c : (-\varepsilon, \varepsilon) \to M$ mit $c(0) = p$, unter folgender Äquivalenzrelation:* *

$$c_1 \sim c_2 \quad \Leftrightarrow \quad (\varphi \circ c_1)'(0) = (\varphi \circ c_2)'(0) \text{ für eine (und damit alle) Karte } \varphi : U \to \varphi(U) \text{ von } M.$$

Die Menge aller Tangentialvektoren in p heißt Tangentialraum T_pM.

Der Zusatz *und damit alle* ergibt sich wie folgt. Sei $\psi : V \to \psi(V)$ eine weitere Karte mit $p \in V$. Dann ist $c(t) \in U \cap V$ für $t \in (-\varepsilon, \varepsilon)$, und es folgt

$$(\psi \circ c)'(0) = (\psi \circ \varphi^{-1} \circ \varphi \circ c)'(0) = D(\psi \circ \varphi^{-1})(\varphi(p))(\varphi \circ c)'(0). \tag{3.1}$$

Example 3.1 Für eine offene Menge $U \subset \mathbb{R}^n$ und $x \in U$ ist T_xU kanonisch bijektiv zu $\mathbb{R}^n$ vermöge der Abbildung

$$T_xU \to \mathbb{R}^n,\ [c] \mapsto c'(0).$$

Die Injektivität folgt unmittelbar aus der Definition des Tangentialvektors. Für die Surjektivität betrachte für gegebenes $v \in \mathbb{R}^n$ die Kurve

$$c_{x,v} : (-\varepsilon, \varepsilon) \to U,\ c_{x,v}(t) = x + tv \quad \text{mit } \varepsilon > 0 \text{ hinreichend klein.} \tag{3.2}$$

Offenbar gilt dann $[c_{x,v}] \mapsto v$.

Das Differential kann nun als Abbildung zwischen den Tangentialräumen erklärt werden.

Definition 3.2 *Seien M, N differenzierbare Mannigfaltigkeiten und $f \in C^1(M, N)$. Die Ableitung von f im Punkt p ist die Abbildung*

$$Df(p) : T_pM \to T_{f(p)}N,\ [c] \mapsto [f \circ c].$$

Wir werden auf T_pM bzw. $T_{f(p)}N$ eine Vektorraumstruktur definieren, so dass $Df(p)$ eine lineare Abbildung ist.

*Dabei kann $\varepsilon > 0$ von c abhängen, und die Darstellung $\varphi \circ c$ muss nur nahe bei $t = 0$ definiert sein.

Lemma 3.1 *Die Ableitung $Df(p)$ von $f \in C^1(M,N)$ ist wohldefiniert.*

Beweis: Wir leiten dazu eine Transformationsformel her. Für $p \in M$ seien $\varphi : U \to \varphi(U)$ und $\psi : V \to \psi(V)$ Karten von M bzw. N mit $p \in U$ und $f(U) \subset V$. Sei $c : (-\varepsilon, \varepsilon) \to M$ eine C^1-Kurve mit $c(0) = p$. Dann folgt aus der Standard Kettenregel

$$(\psi \circ f \circ c)'(0) = (\psi \circ f \circ \varphi^{-1} \circ \varphi \circ c)'(0) = D(\psi \circ f \circ \varphi^{-1})(\varphi(p))(\varphi \circ c)'(0). \tag{3.3}$$

Also folgt aus $c_1 \sim c_2$ auch $f \circ c_1 \sim f \circ c_2$, das heißt $Df(p)([c])$ ist wohldefiniert. □

Example 3.2 Hier geht es um die Beziehung zur üblichen Ableitung im $\mathbb{R}^n$. Sei zuerst $f \in C^1(M,V)$ mit $V \subset \mathbb{R}^m$ offen. Wir fassen die $Df(p)$ aus Definition 3.2 als Abbildung nach $\mathbb{R}^m$ auf, indem wir $T_f(p)V$ mit $\mathbb{R}^m$ identifizieren nach Beispiel 3.1. Das ergibt

$$Df(p) : T_pM \to \mathbb{R}^n,\ [c] \mapsto [f \circ c] \mapsto (f \circ c)'(0).$$

Sei nun $f \in C^1(U,N)$ mit $U \subset \mathbb{R}^n$ offen. Wir fassen dann $Df(x)$, $x \in U$, als Abbildung von $\mathbb{R}^n$ nach $T_{f(x)}N$ auf, indem wir nun T_xU mit $\mathbb{R}^n$ identifizieren nach Beispiel 3.1:

$$Df(x) : \mathbb{R}^n \to T_{f(x)}N,\ v \mapsto [c_{x,v}] \mapsto= [f \circ c_{x,v}] \quad \text{wobei } c_{x,v}(t) = x + tv.$$

Sei schließlich $f \in C^1(U,V)$ mit offenen Mengen $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$. Mit obigen Identifikationen bildet $Df(x)$ wie folgt ab:

$$Df(x) : \mathbb{R}^m \to \mathbb{R}^n,\ v \mapsto [f \circ c_{x,v}] \mapsto (f \circ c_{x,v})'(0) = Df(x)v.$$

Dabei ist rechts die gewöhnliche mehrdimensionale Ableitung gemeint. Somit ist unsere Bezeichnung mit der gewohnten Notation im $\mathbb{R}^n$ konsistent.

Theorem 3.1 *Seien M, N, P differenzierbare Mannigfaltigkeiten. Ist $f \in C^1(M,N)$ und $g \in C^1(N,P)$, so folgt $g \circ f \in C^1(M,P)$ und es gilt*

$$D(g \circ f)(p) = Dg(f(p)) \circ Df(p) \quad \textit{für alle } p \in M.$$

Beweis: Die Aussage $g \circ f \in C^1(M,P)$ wurde bereits in (2.1) gezeigt. Ist $c : (-\varepsilon, \varepsilon) \to M$ eine C^1-Kurve mit $c(0) = p$, so gilt nach Definition

$$\begin{aligned} D(g \circ f)(p)([c]) &= [(g \circ f) \circ c] \\ &= [g \circ (f \circ c)] \\ &= Dg(f(p))\big([f \circ c]\big) \\ &= Dg(f(p))\big(Df(p)([\,c\,])\big). \end{aligned}$$

Damit ist der Satz bewiesen. □

Theorem 3.2 *Sei M eine m-dimensionale Mannigfaltigkeit, $\varphi : U \to \varphi(U)$ eine Karte und $p \in U$. Dann gelten folgende Aussagen:*

(i) *Die Abbildung $D\varphi(p) : T_pM \to \mathbb{R}^m$ ist definiert und bijektiv.*

(ii) *Ist $\psi : V \to \psi(V)$ eine weitere Karte mit $p \in V$, so gilt*

$$D\psi(p) = D(\psi \circ \varphi^{-1})(\varphi(p))\, D\varphi(p).$$

(iii) *Es gibt eine eindeutig bestimmte Vektorraumstruktur auf T_pM, so dass die Abbildung $D\varphi(p) : T_pM \to \mathbb{R}^m$ linear (und damit isomorph) ist.*

Beweis: Nach Beispiel 3.2 ist $D\varphi(p)$ gegeben durch

$$D\varphi(p) : T_pM \to \mathbb{R}^m,\ D\varphi(p)([c]) = (\varphi \circ c)'(0).$$

Genau genommen haben wir das Differential nur für C^1-Funktionen auf ganz M definiert. Aber die C^1-Eigenschaft mittels Kartendarstellungen lässt sich offensichtlich auf Abbildungen $f : \Omega \to N$ mit $\Omega \subset M$ offen verallgemeinern, man nimmt nur Karten $\varphi : U \to \varphi(U)$ mit $U \subset \Omega$; diese ergeben sich durch Einschränkung auf Ω. Insbesondere gilt hier $\varphi \in C^1(U, \mathbb{R}^m)$, denn $\varphi \circ \varphi^{-1} = \mathrm{id}_{\varphi(U)}$, vgl. Definition 2.1. Nach Definition des Tangentialvektors ist $D\varphi(p)$ injektiv. Für die Surjektivität setze $c_{x,v}(t) = x + tv$ mit $x = \varphi(p) \in \varphi(U) \subset \mathbb{R}^n$ und $v \in \mathbb{R}^n$, vgl. Beispiel 3.1. Dann ist $c = \varphi^{-1} \circ c_{x,v}$ eine C^1-Kurve wegen $\varphi \circ c = c_{x,v}$, und es gilt

$$D\varphi(p)([c]) = (\varphi \circ c)'(0) = c'_{x,v}(0) = v. \tag{3.4}$$

Damit ist Behauptung (i) gezeigt.

Behauptung (ii) wurde schon in (3.3) gezeigt, genauer hatten wir dort

$$(\psi \circ c)'(0) = D(\psi \circ \varphi^{-1})(\varphi(p))\,(\varphi \circ c)'(0).$$

Nach Beispiel 3.2 ist $(\psi \circ c)'(0) = D\psi(p)([c])$ und $(\varphi \circ c)'(0) = D\varphi(p)([c])$, und (ii) folgt.

Sei $\varphi : U \to \varphi(U)$ Karte mit $p \in U$. Für $X, Y \in T_pM$ und $\lambda \in \mathbb{R}$ sei eine Addition $X \oplus Y$ und eine Skalarmultiplikation $\lambda \odot X$ gegeben, so dass $D\varphi(p)$ linear ist. Dann folgt

$$\begin{aligned} D\varphi(p)(X \oplus Y) &= D\varphi(p)(X) + D\varphi(p)(Y), \\ D\varphi(p)(\lambda \odot X) &= \lambda\, D\varphi(p)(X). \end{aligned}$$

Da $D\varphi(p)$ injektiv ist, sind $X \oplus Y$ und $\lambda \odot X$ hierdurch eindeutig bestimmt. Da $D\varphi(p)$ auch surjektiv ist, können wir umgekehrt $X \oplus Y$ und $\lambda \odot X$ durch die Gleichungen definieren. Man prüft nach, dass sich die Vektorraumaxiome von $\mathbb{R}^m$ auf T_pM übertragen. Das neutrale Element der Addition ist $0_p = D\varphi(p)^{-1}(0)$, es gilt $0_p = [c_0]$ mit $c_0(t) \equiv p$, die konstante Kurve. Das zu $X \in T_pM$ inverse Element bzgl. $\oplus$ ist $-X := D\varphi(p)^{-1}\big(- D\varphi(p)(X)\big)$. Ist $X = [c]$ so gilt $-X = [\hat{c}]$ mit $\hat{c}(t) = c(-t)$. Die Linearität von $D\varphi(p)$ gilt nach Definition. Es bleibt zu zeigen, dass die Definition nicht von der Karte abhängt. Ist $\psi : V \to \psi(V)$ eine andere Karte mit $p \in V$, so gilt nach (ii)

$$D\psi(p) = D(\psi \circ \varphi^{-1})(\varphi(p))D\varphi(p),$$

das heißt $D\psi(p)$ ist linear bezüglich der durch φ definierten Vektorraumstruktur. Wegen der Eindeutigkeit muss diese gleich der durch ψ definierten Struktur sein. □

Definition 3.3 *Sei $\varphi : U \to \varphi(U)$, $x = \varphi(p)$, Karte auf der n-dimensionalen Mannigfaltigkeit M und $e_1, \ldots, e_n$ die Standardbasis des $\mathbb{R}^n$. Nach Satz 3.2 bilden dann die Vektoren*

$$\frac{\partial}{\partial x^j}(p) = D\varphi(p)^{-1} e_j \in T_pM \qquad \textit{mit } j = 1, \ldots, n$$

eine Basis von T_pM, die als Standardbasis bzgl. der Karte φ bezeichnet wird. Der Vektor $\frac{\partial}{\partial x^j}(p)$ wird durch die Kurve $\varphi^{-1} \circ c_{x,e_j}$ repräsentiert, wobei $c_{x,e_j}(t) = x + te_j$.

Die letzte Aussage gilt wegen

$$D\varphi(p)\frac{\partial}{\partial x^j}(p) = e_j \quad \text{und} \quad D\varphi(p)[\varphi^{-1} \circ c_{x,e_j}] = (\varphi \circ \varphi^{-1} \circ c_{x,e_j})'(0) = e_j.$$

Wir bezeichnen die Basisvektoren stets mit den Buchstaben der entsprechenden Koordinaten, zum Beispiel sind $\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}$ und $\frac{\partial}{\partial \varphi}$ die Basisvektoren zu den Polarkoordinaten r, θ, φ auf $\mathbb{R}^3$. Bei dieser Notation muss allerdings klar sein, bezüglich welcher Karte φ die Standardbasis gebildet wird. In Folgerung **??** wird erklärt, warum die Basisvektoren als partielle Ableitungen geschrieben werden.

Example 3.3 Ein C^1-Diffeomorphismus $\varphi : U \to V$ zwischen offenen Mengen U, V im $\mathbb{R}^n$ kann als Karte aufgefasst werden. Setzen wir $f = \varphi^{-1}$, so lautet die zugehörige Standardbasis im Punkt $p \in U$ gemäß Definition 3.3

$$\frac{\partial}{\partial x^j}(p) = Df(\varphi(p))e_j = \frac{\partial f}{\partial x^j}(\varphi(p)).$$

Insbesondere erhalten wir für $f = \mathrm{id}_U$ die Standardbasis $e_1, \ldots, e_n$ des $\mathbb{R}^n$. Im allgemeinen werden die Vektoren $\frac{\partial}{\partial x^j}(p)$ aber von p abhängen. Im Fall der Polarkoordinaten gilt

$$\begin{aligned} f : (0,\infty) \times (0,\pi) \times (0,2\pi) &\to \mathbb{R}^3 \backslash \{(x,y,z) : x \geq 0,\ y = 0\}, \\ f(r,\theta,\varphi) &= (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta). \end{aligned}$$

Die Standardbasis bezüglich φ im Punkt $p = f(r,\theta,\varphi)$ lautet somit

$$\begin{aligned} \frac{\partial}{\partial r}(p) &= (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta), \\ \frac{\partial}{\partial \theta}(p) &= (r\cos\theta\cos\varphi, r\cos\theta\sin\varphi, -r\sin\theta), \\ \frac{\partial}{\partial \varphi}(p) &= (-r\sin\theta\sin\varphi, r\sin\theta\cos\varphi, 0). \end{aligned}$$

Als Funktion von $p = (x,y,z)$ erhalten wir für die Basis

$$\begin{aligned} \frac{\partial}{\partial r}(p) &= \frac{\big(x,y,z\big)}{\sqrt{x^2+y^2+z^2}}, \\ \frac{\partial}{\partial \theta}(p) &= \frac{\big(zx, zy, -(x^2+y^2)\big)}{\sqrt{x^2+y^2}} \\ \frac{\partial}{\partial \varphi}(p) &= \big(-y, x, 0\big). \end{aligned}$$

Corollary 3.1 *Sei M eine n-dimensionale Mannigfaltigkeit und $X \in T_pM$. Bezüglich einer Karte $\varphi : U \to \varphi(U)$, mit $\varphi(p) = x$ hat X die Koordinatendarstellung*

$$X = \sum_{i=1}^{n} X^i \frac{\partial}{\partial x^i}(p) \quad \textit{mit } X^i = D\varphi^i(p)X. \tag{3.5}$$

Ist $\psi : V \to \psi(V)$, $\psi(p) = y$, eine weitere Karte, so transformieren sich die Koordinatenvektoren mit der Jacobimatrix des Kartenwechsels:

$$X_\psi = D(\psi \circ \varphi^{-1})(x)X_\varphi \quad \textit{und} \quad \frac{\partial}{\partial x^j}(p) = \sum_{i=1}^{n} \frac{\partial(\psi^i \circ \varphi^{-1})}{\partial x^j}(x)\,\frac{\partial}{\partial y^i}(p). \tag{3.6}$$

Beweis: Die Vektoren $\frac{\partial}{\partial x^i}(p)$, $1 \leq i \leq n$, bilden eine Basis von T_pM, also gilt die Darstellung für gewisse $X^i \in \mathbb{R}$. Wir berechnen mit $X = [c]$

$$\sum_{i=1}^n X^i e_i = D\varphi(p)X = (\varphi \circ c)'(0) = \sum_{i=1}^n (\varphi^i \circ c)'(0)\, e_i = \sum_{i=1}^n \big(D\varphi^i(p)X\big)\, e_i.$$

Die erste Behauptung folgt durch Koeffizientenvergleich. Weiter gilt

$$X_\psi = D\psi(p)X = D(\psi \circ \varphi^{-1})(x) D\varphi(p)X = D(\psi \circ \varphi^{-1})(x) X_\varphi.$$

Die Formel für die Basisvektoren ergibt sich mit $X = \frac{\partial}{\partial x^j}(p)$, den dann ist $X_\varphi = e_j$. □

Theorem 3.3 *Seien M, N differenzierbare Mannigfaltigkeiten der Dimension n bzw. m und $f \in C^1(M, N)$. Dann ist die Abbildung $Df(p) : T_pM \to T_{f(p)}N$ linear. Sind $\varphi : U \to \varphi(U) \subset \mathbb{R}^m$ und $\psi : V \to \psi(V) \subset \mathbb{R}^n$ Karten mit $p \in U$ und $f(U) \subset V$, so gilt*

$$D\psi(f(p)) Df(p) (D\varphi(p))^{-1} = D(\psi \circ f \circ \varphi^{-1})(\varphi(p)). \tag{3.7}$$

Bezüglich $\frac{\partial}{\partial x^j}(p)$, $j = 1, \ldots, n$, bzw. $\frac{\partial}{\partial y^i}(f(p))$, $i = 1, \ldots, m$ hat $Df(p)$ die Matrix

$$\left(\frac{\partial(\psi^i \circ f \circ \varphi^{-1})}{\partial x^j}\right)(\varphi(p)) \in \mathbb{R}^{m \times n}. \tag{3.8}$$

Beweis: Formel (3.7) folgt aus der Kettenregel, Satz 3.1. Beachte dabei

$$\varphi \circ \varphi^{-1} = \mathrm{id}_U \quad \Rightarrow \quad D\varphi(p) D(\varphi^{-1})(\varphi(p)) = \mathrm{Id}_{\mathbb{R}^m}.$$

$D\varphi(p)$ und $D\psi(f(p))$ sind Vektorraum-Isomorphismen, und $D(\psi \circ f \circ \varphi^{-1})(\varphi(p))$ ist die gewöhnliche Ableitung, vgl. Beispiel 3.2. Damit ist auch $Df(p)$ linear. Für die zugehörige Matrix berechnen wir

$$\begin{aligned}
Df(p)\frac{\partial}{\partial x^j}(p) &= D\psi(f(p))^{-1} D\psi(f(p)) Df(p) D\varphi(p)^{-1} e_j \\
&= D\psi(f(p))^{-1} D(\psi \circ f \circ \varphi^{-1})(\varphi(p))\, e_j \\
&= D\psi(f(p))^{-1} \sum_{i=1}^m \frac{\partial(\psi^i \circ f \circ \varphi^{-1})}{\partial x^j}(\varphi(p))\, e_i \\
&= \sum_{i=1}^m \frac{\partial(\psi^i \circ f \circ \varphi^{-1})}{\partial x^j}(\varphi(p))\, \frac{\partial}{\partial y^i}(f(p)).
\end{aligned}$$

□

In der Physik ist es üblich, eine Funktion in Koordinaten zu schreiben. Die Koordinaten und die zugehörigen Komponenten der Funktion werden mit demselben Buchstaben bezeichnet. In der Situation hier ergibt sich die Notation $f^i(p) := \psi^i \circ f(p)$. Die Matrix von $Df(p)$ hat dann die Einträge

$$\frac{\partial f^i}{\partial x^j}(p) := \frac{\partial(\psi^i \circ f \circ \varphi^{-1})}{\partial x^j}(\varphi(p)). \tag{3.9}$$

Es kann dann gerechnet werden wie es mit partiellen Ableitungen üblich ist. Das ist ziemlich praktisch, im Kapitel über die Lieklammer werden wir das zum Beispiel benutzen.

Example 3.4 Wir betrachten hier den Spezialfall einer C^1-Kurve $\gamma : I = (a,b) \to M$. Bezüglich der Karte $\varphi = \mathrm{id}_I$ haben wir den Basisvektor

$$\frac{\partial}{\partial t}(t_0) = 1 \in \mathbb{R}.$$

Der Vektor ist durch $c_{t_0,1}(t) = t_0 + t$ repräsentiert, denn $c'_{t_0,1}(t_0) = 1$. Für $f \in C^1(I, \mathbb{R}^m)$ berechnen wir nach Beispiel 3.2

$$Df(t_0)\frac{\partial}{\partial t}(t_0) = (f \circ c_{t_0,1})'(t_0) = \frac{d}{dt} f(t_0 + t)|_{t=t_0} = f'(t_0).$$

Wir definieren nun den Tangentialvektor von γ in t_0 durch

$$\gamma'(t_0) := D\gamma(t_0)\frac{\partial}{\partial t}(t_0).$$

Damit folgt in einer Karte $\psi : V \to \psi(V)$ von M mit der Kettenregel

$$D\psi(\gamma(t_0))\gamma'(t_0) = D\psi(\gamma(t_0))D\gamma(t_0)\frac{\partial}{\partial t}(t_0) = D(\psi \circ \gamma)(t_0)\frac{\partial}{\partial t}(t_0) = (\psi \circ \gamma)'(t_0).$$

Example 3.5 Sei N eine n-dimensionale Mannigfaltigkeit. $M \subset N$ heißt m-dimensionale Untermannigfaltigkeit der Klasse C^∞, wenn es zu jedem $p \in M$ eine M-adaptierte Karte $g : W \to g(W)$ von N gibt mit $p \in W$, das heißt

$$g(M \cap W) = (\mathbb{R}^m \times \{0\}) \cap g(W). \tag{3.10}$$

M ist mit der Relativtopologie ein Hausdorffraum mit abzählbarer Basis. Um einen Atlas zu definieren, führen wir folgende Bezeichnungen ein:

$$\begin{aligned} \iota &: \mathbb{R}^m \to \mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}, \quad \iota(x^1, \ldots, x^m) = (x^1, \ldots, x^m, 0, \ldots, 0), \\ \pi &: \mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}^m, \quad \pi(y^1, \ldots, y^n) = (y^1, \ldots, y^m). \end{aligned}$$

Für eine M-adaptierte Karte $g : W \to g(W)$ setzen wir nun

$$\varphi_g : U_g := M \cap W \to \varphi_g(U_g),\ \varphi_g = \pi \circ g|_{M \cap W}.$$

Dann gilt $\varphi_g^{-1} = f \circ \iota|_{\varphi_g(U_g)}$, wobei $f = g^{-1} : g(W) \to W$ die Umkehrabbildung der Karte g bezeichnet. Insbesondere ergeben sich die Kartenwechsel

$$\begin{aligned} \varphi_{g'} \circ \varphi_g^{-1}|_{\varphi_g(U_g \cap U_{g'})} : \quad & g(M \cap W \cap W') \to g'(M \cap W \cap W'), \\ \varphi_{g'} \circ \varphi_g^{-1} \quad & = \pi \circ (g' \circ g^{-1}) \circ \iota|_{g(M \cap W \cap W')}. \end{aligned}$$

Damit ist auf M ein differenzierbarer Atlas gegeben. Die Inklusionsabbildung $i : M \subset N$ ist bezüglich dieser Struktur differenzierbar, denn in den Karten gilt

$$g \circ i \circ \varphi_g^{-1} = \iota|_{(\pi \circ g)(M \cap W)} \in C^\infty(\varphi_g(U_g)).$$

Aus $i = f \circ \iota \circ \varphi_g$ auf $M \cap W$ folgt $Di(p) = Df(g(p)\,\iota\, D\varphi_g(p)$ für $p \in M \cap W$. Somit ist die lineare Abbildung

$$Di(p) : T_pM \to \text{Bild}\, Df(x,0)|_{\mathbb{R}^m \times \{0\}} \subset T_pN, \qquad \text{wobei } x = \varphi_g(p) = (\pi \circ g)(p),$$

ein Isomorphismus. Die Basisvektoren $\frac{\partial}{\partial x^j}(p) \in T_pM$ bzgl. der Karte φ_g werden mit den Vektoren $Df(x,0)(e_j, 0) \in T_pN$ mit $j = 1, \ldots, m$ identifiziert.

4 Tangentialbündel und Vektorfelder

Wir wollen nun die Tangentialräume T_pM, $p \in M$, zu einem Gesamtobjekt zusammenfassen; dieses ist das Tangentialbündel

$$TM = \bigcup_{p \in M} T_pM.$$

Natürlich soll TM auch wieder eine Mannigfaltigkeit sein. Allerdings ist TM eine Vereinigung von Vektorräumen, die sozusagen durch $p \in M$ parametrisiert sind. Es ist sinnvoll, hier folgenden Begriff einzuführen.

Definition 4.1 (Vektorbündel) *Seien E und M differenzierbare Mannigfaltigkeiten. Eine Abbildung $\pi \in C^\infty(E, M)$ heißt Vektorbündel vom Rang k über M, falls gilt:*

(i) *Für alle $p \in M$ trägt $E_p = \pi^{-1}\{p\}$ die Struktur eines k-dimensionalen $\mathbb{R}$-Vektorraums.*

(ii) *Für alle $p \in M$ existiert eine Umgebung $U \subset M$ und ein Diffeomorphismum der Form*

$$\phi : \pi^{-1}(U) \to U \times \mathbb{R}^k, \quad \phi(v) = \big(q, A(q)v\big) \text{ mit } q = \pi(v),$$

wobei $A(q) : E_q \to \mathbb{R}^k$ ein Vektorraumisomorphismus ist.

Hat man einmal einen Vektorraum V, so konstruiert die Lineare Algebra daraus weitere Vektorräume, etwa den Dualraum V^* oder den Raum $\mathrm{End}(V)$ der Endomorphismen von V. In unserer Situation ergeben sich so weitere Vektorbündel, das ist der Grund diesen Begriff allgemein einzuführen. Ich habe die Eigenschaft (ii) in die Definition aufgenommen (manche Autoren machen das nicht). Das Bündel $\pi : M \times \mathbb{R}^k$, $\pi(p, v) = p$, heißt das triviale Bündel. Eigenschaft (ii) besagt, dass das Bündel *lokal trivial* ist, also ein Produkt $U \times \mathbb{R}^k$. Allgemein heißt ϕ in (ii) Bündelkarte oder lokale Trivialisierung.

Lemma 4.1 *Sei M eine n-dimensionale Mannigfaltigkeit mit maximalem Atlas $\mathcal{A}$ und Projektion $\pi : TM \to M$, $\pi(X) = p$ für $X \in T_pM$. Dann gibt es eine eindeutig bestimmte Topologie auf TM, so dass die Abbildungen*

$$\varphi_\sharp : \pi^{-1}(U) \to \varphi(U) \times \mathbb{R}^n,\ \varphi_\sharp(X) = \big(\varphi(p), D\varphi(p)X\big) \text{ für } X \in T_pM \quad (\varphi \in \mathcal{A}) \tag{4.1}$$

einen C^∞-Atlas bilden. Damit ist TM eine differenzierbare Mannigfaltigkeit der Dimension $2n$, und die Projektion $\pi : TM \to M$ ist in $C^\infty(TM, M)$.

Beweis: Wir wollen Satz 1.2 anwenden. Die Abbildungen $\varphi_\sharp$ sind bijektiv, das Bild $\varphi(U) \times \mathbb{R}^n$ ist offen und die Mengen $\pi^{-1}(U)$, $\varphi \in \mathcal{A}$, überdecken TM. Für zwei Karten $\varphi : U \to \varphi(U)$ und $\psi : V \to \psi(V)$ ist die Menge

$$\varphi_\sharp(\pi^{-1}(U) \cap \pi^{-1}(V)) = \varphi(U \cap V) \times \mathbb{R}^n$$

offen in $\mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$. Weiter ist der Kartenwechsel nach Satz 3.2(ii) stetig:

$$\begin{aligned} \psi_\sharp \circ (\varphi_\sharp)^{-1} : \varphi(U \cap V) \times \mathbb{R}^n &\to \psi(U \cap V) \times \mathbb{R}^n, \\ (\psi_\sharp \circ (\varphi_\sharp)^{-1})(x, v) &= \big((\psi \circ \varphi^{-1})(x), D(\psi \circ \varphi^{-1})(x)v\big). \end{aligned}$$

Nach Satz 1.2 gibt es auf TM genau eine Topologie, so dass die $\varphi_\sharp$ homemorph sind und damit einen C^∞-Atlas von TM bilden, es gilt

$$W \subset TM \text{ offen} \quad \Leftrightarrow \quad \varphi_\sharp(W \cap \pi^{-1}(U)) \text{ offen für alle } \varphi \in \mathcal{A}.$$

Die Projektion $\pi : TM \to M$ ist C^∞, denn es gilt

$$\varphi \circ \pi \circ (\varphi_\sharp)^{-1} : \varphi(U) \times \mathbb{R}^n \to \varphi(U),\, \pi(x,v) = x.$$

Nach Satz 1.1(3) hat M einen abzählbaren Atlas $\mathcal{A}'$, und TM wird durch die Mengen $\pi^{-1}(U_\varphi)$, $\varphi \in \mathcal{A}'$, überdeckt. Somit hat TM eine abzählbare Basis nach Satz 1.2.

Zu zeigen bleibt die Hausdorff-Eigenschaft. Seien $X, Y \in TM$ mit $X \neq Y$. Ist $\pi(X) \neq \pi(Y)$, so gibt es offene Umgebungen $U, V \subset M$ von $\pi(X)$ bzw. $\pi(Y)$ mit $U \cap V = \emptyset$. Die Mengen $\pi^{-1}(U)$ und $\pi^{-1}(V)$ sind dann disjunkte, offene Umgebungen von X bzw. Y. Im Fall $\pi(X) = \pi(Y) =: p$ wähle eine Karte $\varphi : U \to \varphi(U)$ mit $p \in U$. Es gilt dann $D\varphi(p)X = v$ und $D\varphi(p)Y = w$ mit $v \neq w$. Wähle offene und disjunkte Umgebungen $V, W \subset \mathbb{R}^n$ von v, w. Dann sind $\varphi_\sharp^{-1}(U \times V)$ und $\varphi_\sharp^{-1}(U \times W)$ wie verlangt. □

Wir hatten nach Satz 1.1 die Konvergenz einer Folge von Punkten in einer Mannigfaltigkeit charakterisiert. Das besagt hier Folgendes:

> Seien $X_k \in T_{p_k}M$, $X \in T_pM$, und sei $\varphi : U \to \varphi(U)$ eine Karte mit $p \in U$. Dann konvergiert X_k genau dann gegen X, wenn $p_k \to p$, insbesondere $p_k \in U$ für k groß, und $(X_k)^i \to X^i$ mit $k \to \infty$ für alle $i = 1, \ldots, n$. Dabei sind $(X_k)^i$ und X^i die Koordinaten bzgl. der Basen $\frac{\partial}{\partial x^i}(p_k)$ bzw. $\frac{\partial}{\partial x^i}(p)$.

Theorem 4.1 (Tangentialbündel) *Für eine n-dimensionale differenzierbare Mannigfaltigkeit ist $\pi : TM \to M$ ein n-dimensionales Vektorbündel vom Rang n.*

Beweis: Sei $p \in M$ und $\varphi : U \to \varphi(U)$ eine Karte mit $p \in M$. Nach Satz 3.2 hat T_pM eine eindeutig bestimmte Vektorraumstruktur, so dass $D\varphi(p) : T_pM \to \mathbb{R}^n$ linear ist. Wir haben die Diffeomorphismen

$$\phi : \pi^{-1}(U) \to U \times \mathbb{R}^n,\, \phi(X) = (q, D\varphi(q)X) \quad \text{für } q = \pi(X) \in U.$$

□

Corollary 4.1 *Seien M, N differenzierbare Mannigfaltigkeiten der Dimension m bzw. n. Ist $f \in C^k(M,N)$ mit $k \geq 1$, so ist $Df \in C^{k-1}(TM, TN)$.*

Beweis: Seien $\varphi : U \to \varphi(U)$ bzw. $\psi : V \to \psi(V)$ Karten von M bzw. N mit $f(U) \subset V$. Dann gilt für $x \in \varphi(U)$ und $v \in \mathbb{R}^m$ nach Satz 3.3

$$\psi_\sharp \circ Df \circ (\varphi_\sharp)^{-1}(x,v) = \big((\psi \circ f \circ \varphi^{-1})(x), D(\psi \circ f \circ \varphi^{-1})(x)v\big) \in \psi(V) \times \mathbb{R}^n.$$

Die Behauptung folgt. □

Definition 4.2 *Ein Vektorfeld auf M ist eine Abbildung $X : M \to TM$ mit $\pi \circ X = \mathrm{id}_M$, also $X(p) \in T_pM$ für alle $p \in M$. Wir bezeichnen den Raum der C^k-Vektorfelder mit $C^k(TM)$.*

Die C^k-Eigenschaft ist durch die Darstellungen bezüglich Karten definiert. Ist $\varphi : U \to \varphi(U)$ eine Karte, so lauten die lokalen Darstellungen von X nach Lemma 4.1

$$\varphi_\sharp \circ X \circ \varphi^{-1}(x) = (\varphi(p), X^1(p), \ldots, X^n(p))|_{p=\varphi^{-1}(x)} \quad \text{mit } X^i(p) = D\varphi^i(p)X.$$

Damit ist $X \in C^k(M, TM)$ genau wenn für jede Karte die $X^i \circ \varphi^{-1}$ in $C^k(\varphi(U))$ sind.

Sei $\pi : E \to M$ ein Vektorbündel. Eine Funktion $X : M \to E$ mit $\pi \circ X = \mathrm{id}_M$, das heißt $X(p) \in E_p$ für alle $p \in M$, nennt man einen Schnitt von E. Die Differentialgeometer bezeichnen die C^∞-Schnitte mit $\Gamma(E)$. Diese Notation hat den Nachteil, dass sie andere Regularitätsstufen, zum Beispiel stetige oder C^1-Schnitte, nicht abdeckt. Ich bezeichne deshalb die C^k-Schnitte von E mit $C^k(E)$. Nachteil meiner Notation ist natürlich die Verwechselung mit dem Raum der C^k-Funktionen $f : E \to \mathbb{R}$. Ich hoffe dass sich aus dem Kontext ergibt was gemeint ist.

$C^k(TM)$ ist $\mathbb{R}$-Vektorraum mit der punktweisen Vektorraumstruktur, vgl. Satz 3.2,

$$(\lambda X + \mu Y)(p) = \lambda X(p) + \mu Y(p).$$

Ferner können Vektorfelder mit Funktionen multipliziert werden:

$$f \in C^k(M),\, X \in C^k(TM) \quad \Rightarrow \quad fX \in C^k(TM), \text{ wobei } (fX)(p) = f(p)X(p).$$

Definition 4.3 *Die Basisfelder einer Karte $\varphi : U \to \varphi(U)$, $x = \varphi(p)$, sind*

$$\frac{\partial}{\partial x^j} \in C^\infty(TU) : p \mapsto \frac{\partial}{\partial x^j}(p) = D\varphi(p)^{-1} e_j.$$

Ein Vektorfeld X auf M hat bezüglich der Basisfelder die Darstellung

$$X = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i} \quad \text{mit } X^i(p) = D\varphi^i(p)X.$$

Definition 4.4 *Sei $I \subset \mathbb{R}$ ein (evtl. uneigentliches) Intervall. $c \in C^1(I, M)$ heißt Integralkurve des Vektorfelds $X \in C^0(TM)$, falls gilt:*

$$c' = X \circ c \quad \Leftrightarrow \quad c'(t) = X(c(t)) \text{ für alle } t \in I. \tag{4.2}$$

c heißt maximal, wenn es keine Integralkurve $\tilde{c} : \tilde{I} \to M$ gibt mit $I \subsetneq \tilde{I}$ und $\tilde{c}|_I = c$.

Genauer könnte man auch die Begriffe rechts- bzw. linksseitig maximal erklären. Wir interessieren uns nun zuerst dafür, wie sich die Gleichung in lokale Koordinaten übersetzt.

Lemma 4.2 *Sei M eine n-dimensionale Mannigfaltigkeit und $X \in C^0(TM)$ ein Vektorfeld auf M, mit lokaler Darstellung $X = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i}$ bezüglich der Karte $\varphi : U \to \varphi(U)$. Definiere*

$$\xi \in C^0(U, \mathbb{R}^n),\, \xi(x) = D\varphi(p)X(p)|_{p=\varphi^{-1}(x)} = \sum_{i=1}^n X^i(\varphi^{-1}(x))\, e_i.$$

$c \in C^1(I, U)$ ist genau dann Integralkurve von X, wenn für $\gamma(t) = \varphi(c(t))$ gilt:

$$\gamma' = \xi \circ \gamma \quad \Leftrightarrow \quad \gamma'(t) = \xi(\gamma(t)) \text{ für alle } t \in I. \tag{4.3}$$

BEWEIS: Nach Beispiel 3.4 gilt, links mit der Standard Ableitung,

$$\gamma'(t) = (\varphi \circ c)'(t) = D\varphi(c(t))c'(t).$$

Andererseits gilt nach Definition wegen $\varphi^{-1}(\gamma(t)) = c(t)$

$$\xi(\gamma(t)) = D\varphi(c(t))X(c(t)).$$

Die Behauptung folgt. □

Gleichung (4.3) ist ein autonomes System von n gewöhnlichen Differentialgleichungen erster Ordnung. Die relevante Theorie im $\mathbb{R}^n$ liefert folgende Aussagen zur Eindeutigkeit und Existenz von Integralkurven, sowie zur Abhängigkeit der Integralkurve vom Anfangswert.

Theorem 4.2 *Sei $X \in C^1(TM)$. Dann gelten folgende Aussagen:*

(i) *Sind $c_k : I_k \to M$ Integralkurven von X mit $c_1(t_0) = c_2(t_0)$ für ein $t_0 \in I_1 \cap I_2$, so gilt $c_1 = c_2$ auf $I_1 \cap I_2$ und die zusammengesetzte Kurve*

$$c : I_1 \cup I_2 \to M,\ c(t) = \begin{cases} c_1(t) & \text{für } t \in I_1 \\ c_2(t) & \text{für } t \in I_2 \end{cases}$$

ist wieder eine Integralkurve von X.

(ii) *Zu jedem $p \in M$ gibt es ein $\delta > 0$, so dass das Anfangswertproblem $c'(t) = X(c(t))$, $c(0) = p$, eine Lösung $c_p : (-\delta, \delta) \to M$ hat.*

(iii) *Zu jedem $p \in M$ gibt es eine Umgebung U und ein $\delta > 0$, so dass die Abbildung*

$$\phi : U \times (-\delta, \delta) \to M,\ \phi(q,t) = c_q(t),$$

definiert und C^1 ist. Ist $X \in C^k(TM)$ mit $k \in \mathbb{N} \cup \{\infty\}$, so ist ϕ eine C^k-Abbildung.

BEWEIS: Für (i) betrachte die Menge $J = \{t \in I_1 \cap I_2 : c_1(t) = c_2(t)\}$. Nach Definition ist J abgeschlossen in $I_1 \cap I_2$ und nach Voraussetzung nicht leer, da $t_0 \in J$. Aber J ist auch offen in $I_1 \cap I_2$; dies folgt durch Anwendung des Eindeutigkeitssatzes im $\mathbb{R}^n$ in einer lokalen Karte gemäß Lemma 4.2. Also gilt $J = I_1 \cap I_2$, das heißt die Lösungen stimmen auf $I_1 \cap I_2$ überein. Die Lösungseigenschaft der zusammengesetzten Kurve ist offensichtlich. Die Behauptungen (ii) und (iii) sind direkte Konsequenzen der entsprechenden Aussagen im $\mathbb{R}^n$ zusammen mit Lemma 4.2. □

Ist nur $X \in C^0(TM)$, so gilt das Eindeutigkeitsresultat im allgemeinen nicht. Betrachte dazu die Gleichung $\gamma' = \sqrt{|\gamma|}$ für $\gamma : \mathbb{R} \to \mathbb{R}$. Es gibt eine Schar von C^1-Lösungen mit $\gamma(0) = 0$, die von zwei Parametern $t_1 \leq 0$, $t_2 \geq 0$ abhängt:

$$x(t) = \begin{cases} -\frac{1}{4}(t-t_1)^2 & t \leq t_1 \\ 0 & t_1 \leq t \leq t_2 \\ \frac{1}{4}(t-t_2)^2 & t \geq t_2. \end{cases}$$

Ein Beispiel dafür, dass die Integralkurven nicht immer auf ganz $\mathbb{R}$ definiert werden können, liefert die Gleichung $\gamma' = \gamma^2$. Die maximale Integralkurve mit $\gamma(0) = x_0 \neq 0$ lautet

$$\gamma(t) = \frac{1}{\frac{1}{x_0} - t} \text{ mit } \begin{cases} t \in (-\infty, \frac{1}{x_0}) & \text{falls } x_0 > 0, \\ t \in (\frac{1}{x_0}, \infty) & \text{falls } x_0 < 0. \end{cases}$$

Corollary 4.2 *Sei $X \in C^1(TM)$. Dann gibt es zu jedem $p \in M$ eine maximale Integralkurve $c_p : I_p \to M$ von X mit $c_p(0) = p$, und es gilt:*

(i) *$I_p = (\alpha_p, \omega_p)$ ist ein offenes Intervall mit $-\infty \leq \alpha_p < 0 < \omega_p \leq \infty$.*

(ii) *Ist $c : I \to M$ Integralkurve von X mit $c(0) = p$, so gilt $I \subset I_p$ und $c = c_p|_I$.*

(iii) *Zu $K \subset M$ kompakt gibt es ein $\delta > 0$ mit $(-\delta, \delta) \subset (\alpha_p, \omega_p)$ für alle $p \in K$.*

Beweis: Wir setzen

$$\begin{aligned}
\alpha_p &= \inf\{a < 0 : \text{ es gibt eine Integralkurve } c : (a, 0] \to M \text{ mit } c(0) = p\},\\
\omega_p &= \sup\{b > 0 : \text{ es gibt eine Integralkurve } c : [0, b) \to M \text{ mit } c(0) = p\}.
\end{aligned}$$

Nach Satz 4.2(ii) gilt $\alpha_p < 0 < \omega_p$, und mit Satz 4.2(i) erhalten wir eine wohldefinierte Lösung $c_p : (\alpha_p, \omega_p) \to M$ des Anfangswertproblems. Wäre c_p zum Beispiel auf $(\alpha_p, \omega_p]$ fortsetzbar, so gibt es nach Satz 4.2(ii) für geeignetes $\delta > 0$ eine Integralkurve $c : (\omega_p - \delta, \omega_p + \delta) \to M$ mit $c(\omega_p) = c_p(\omega_p)$, und wir erhalten eine Fortsetzung von c_p auf das Intervall $(\alpha_p, \omega_p + \delta)$ im Widerspruch zur Definition von ω_p. Also ist $c_p : I_p = (\alpha_p, \omega_p) \to M$ maximale Integralkurve. Die Eindeutigkeitsaussage ergibt sich aus Satz 4.2(i) und der Maximalität. Schließlich ergibt sich Behauptung (iii) aus Satz 4.2(iii) mit einem Überdeckungsargument. $\square$

Corollary 4.3 *Sei $c_p : (\alpha_p, \omega_p) \to M$ die maximale Integralkurve von $X \in C^1(TM)$ mit Anfangswert $c_p(0) = p$. Gilt dann $\omega_p < \infty$ (bzw. $\alpha_p > -\infty$), so gibt es zu $K \subset M$ kompakt ein $t^K < \omega_p$ (bzw. $t_K > \alpha_p$), so dass gilt:*

$$c_p(t) \notin K \text{ für alle } t > t^K \quad (\text{ bzw. } c_p(t) \notin K \text{ für alle } t < t_K).$$

Hat X kompakten Träger auf M, so gilt $(\alpha_p, \omega_p) = (-\infty, \infty)$.

Beweis: Sei $\omega_p < \infty$. Dann gilt $\omega_{c_p(t)} = \omega_p - t \searrow 0$ mit $t \nearrow \omega_p$. Folgerung 4.2(iii) liefert $c_p(t) \notin K$ für t hinreichend groß. Wäre $\omega_p < \infty$ im Fall $K := \operatorname{spt} X$ kompakt, so folgt $c_p(t) \notin K$ für $t > t_K$, also $c_p'(t) = X(c_p(t)) = 0$ für $t > t^K$. Aber dann ist $c_p(t)$ konstant für $t > t^K$, und wir können die Lösung konstant auf (t^K, ∞) fortsetzen, im Widerspruch zur Annahme. $\square$

Definition 4.5 *Für $X \in C^1(TM)$ sei $c_p : I_p \to M$ die maximale Integralkurve mit $c_p(0) = p$, und $M^X = \{(p, t) \in M \times \mathbb{R} : t \in I_p\}$. Der Fluss von X ist die Abbildung*

$$\phi : M^X \to M,\ \phi(p, t) = c_p(t).$$

Wir setzen $M_t^X = \{p \in M : t \in I_p\}$ und $\phi_t : M_t^X \to M$, $\phi_t(p) = c_p(t)$.

Lemma 4.3 *Sei ϕ der Fluss von $X \in C^1(TM)$. Sind $t, s + t \in I_p$, so folgt $s \in I_{\phi(p,t)}$ und*

$$(\phi_s \circ \phi_t)(p) = \phi_{s+t}(p). \tag{4.4}$$

Beweis: Da I_p Intervall, gilt:

$$\begin{aligned}
&\text{im Fall } s > 0 : \quad \sigma + t \in I_p \quad \text{für } \sigma \in [0, s] =: I,\\
&\text{im Fall } s < 0 : \quad \sigma + t \in I_p \quad \text{für } \sigma \in [s, 0] =: I.
\end{aligned}$$

Wir berechnen für $\sigma \in I$

$$\frac{d}{d\sigma} c_p(\sigma + t) = (c_p)'(\sigma + t) = X(c_p(\sigma + t)) \quad \text{und} \quad c_p(\sigma + t)|_{\sigma=0} = c_p(t) = \phi(p, t).$$

Es folgt $s \in I_{\phi(p,t)}$ und $(\phi_s \circ \phi_t)(p) = c_p(s + t) = \phi_{s+t}(p)$. $\square$

Corollary 4.4 *Sei $X \in C^k(TM)$. Der Fluss $\phi : M^X \to M$ hat folgende Eigenschaften:*

(i) *Die Mengen $M^X \subset M \times \mathbb{R}$ sowie $M_t^X \subset M$ sind offen.*

(ii) *$\phi : M^X \to M$ ist von der Klasse C^k.*

(iii) *$\phi_t : M_t^X \to M_{-t}^X$ ist ein C^k-Diffeomorphismus.*

Beweis: Sei G die Menge aller Punkte $(p, t) \in M^X$ mit folgender Eigenschaft: es gibt eine offene Umgebung $W \subset M \times \mathbb{R}$ von (p, t) mit $W \subset M^X$ und $\phi|_W \in C^k(W, M)$. Wir zeigen $G = M^X$, und damit die Behauptungen (i) und (ii). Beachte dabei: M_t^X offen folgt aus M^X offen, denn $M_t^X = i_t^{-1}(M^X)$ mit $i_t(p) = (p, t)$. Im folgenden sei $(p, T) \in M^X$ vorgegeben, ohne Einschränkung mit $T > 0$.

Schritt 1: Es gibt $U \supset c_p([0, T])$ offen und $\delta > 0$ mit $U \times (-\delta, \delta) \subset G$.
Nach Satz 4.2 gibt es zu jedem $t \in [0, T]$ eine Umgebung U_t von $c_p(t)$ und ein $\delta_t > 0$ mit $U_t \times (-\delta_t, \delta_t) \subset G$. Da $c_p([0, T])$ kompakt, wird es durch U_{t_i}, $i = 1, \dots, N$,überdeckt. Wir setzen $U = \bigcup_{i=1}^N U_{t_i}$ und $\delta = \min_{i=1,\dots,N} \delta_{t_i} > 0$. Es folgt

$$U \times (-\delta, \delta) \subset \bigcup_{i=1}^N U_{t_i} \times (-\delta_{t_i}, \delta_{t_i}) \subset G.$$

Schritt 2: Setze $\tau = \sup\{t \geq 0 : \{p\} \times [0, t] \subset G\}$. Dann ist $\tau > T$ und somit $(p, T) \in G$.
Nach Schritt 1 ist $p = c_p(0) \in U$, also $\{p\} \times [0, \delta) \subset G$ und damit $\tau \geq \delta > 0$. Angenommen es ist $\tau \leq T$. Für $t \in [0, \tau) \subset [0, T)$ gilt $(p, t) \in G$, also gibt es eine offene Umgebung V von p mit $V \times \{t\} \subset G$ und $\phi_t \in C^k(V, M)$. Es gilt $\phi_t(p) = c_p(t) \in U$. Nach Verkleinerung von V ist sogar $\phi_t(V) \subset U$. Definiere nun die C^k-Abbildung

$$\psi : V \times [t, t + \delta) \to M, \; \psi(q, s) = \phi(\phi_t(p), s - t).$$

$\psi(q, s)$ ist Integralkurve mit Anfangswert $\psi(q, t) = \phi_t(p)$, also haben wir eine C^k Fortsetzung von ϕ auf $V \times [0, t + \delta)$, und damit $\{p\} \times [0, t + \delta) \subset G$. Nach Definition von τ ist dann $t + \delta \leq \tau$, mit $t \nearrow \tau$ folgt $\delta \leq 0$, Widerspruch.

Beweis von (iii): Sei $p \in M_t^X$, also $t \in I_p$. Nach Lemma 4.3 gilt dann $-t \in I_{\phi_t(p)}$ bzw. $\phi_t(p) \in M_{-t}^X$, und $\phi_{-t}(\phi_t(p)) = p$. Da t und $-t$ vertauscht werden können, ist ϕ_t bijektiv. Außerdem gilt $\phi_t = \phi \circ i_t$ auf M_t^X, somit ist ϕ_t ein C^k-Diffeomorphismus. $\square$

Ist das Vektorfeld $X \in C^1(TM)$ vollständig integrierbar, das heißt alle Flusskurven sind auf ganz $\mathbb{R}$ definiert, so liefert der Fluss die Abbildung

$$\mathbb{R} \to \mathrm{Diff}\,(M), \; t \mapsto \phi_t, \quad \text{mit } \phi_s \circ \phi_t = \phi_{s+t};$$

das ist die von X erzeugte Einparametergruppe von Diffeomorphismen von M.

Lemma 4.4 *Sei $X \in T_pM$ mit $Df(p)X = 0$ für alle $f \in C^\infty(M)$. Dann folgt $X = 0$.*

Beweis: Wähle eine Karte $\varphi : U \to \varphi(U)$ mit $p \in U$. Für $i = 1, \ldots, n$ gibt es Funktionen $\widetilde{\varphi}^i \in C^\infty(M)$ mit $\widetilde{\varphi}^i = \varphi^i$ nahe bei p (verwende eine Abschneidefunkion). Nach Definition der Ableitung gilt dann $D\varphi^i(p) = D\widetilde{\varphi}^i(p)$, also folgt

$$X^i = D\varphi^i(p)X = D\widetilde{\varphi}^i(p)X = 0.$$

□

Vektorfelder $X \in C^0(TM)$ operieren auf Funktionen $f \in C^1(M)$ durch Richtungsableitung. In diesem Zusammenhang ist folgende Notation üblich:

$$Xf : M \to \mathbb{R},\ Xf(p) = Df(p)X(p).$$

Theorem 4.3 (Lieklammer) *Seien $X, Y \in C^k(TM)$ mit $k \in \mathbb{N} \cup \{\infty\}$. Dann gibt es ein eindeutig bestimmtes C^{k-1}-Vektorfeld, die Lieklammer $[X,Y]$ von X und Y, so dass gilt:*

$$[X,Y]f = X(Yf) - Y(Xf) \quad \text{für alle } f \in C^\infty(M). \tag{4.5}$$

Bezüglich einer Karte $\varphi : U \to \varphi(U)$ hat $[X,Y]$ die Koordinatendarstellung

$$\sum_{i,j=1}^n \big(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i}\big) \frac{\partial}{\partial x^j}. \tag{4.6}$$

Beweis: Wir berechnen bzgl. einer lokalen Karte $\varphi : U \to \varphi(U)$

$$X(Yf) = \sum_{i,j=1}^n X^i \frac{\partial}{\partial x^i}\big(Y^j \frac{\partial f}{\partial x^j}\big) = \sum_{i,j=1}^n X^i \frac{\partial Y^j}{\partial x^i}\frac{\partial f}{\partial x^j} + \sum_{i,j=1}^n X^i Y^j \frac{\partial^2 f}{\partial x^i \partial x^j}.$$

Nach Schwarz heben sich die zweiten Ableitungen beim Kommutator weg, es folgt

$$X(Yf) - Y(Xf) = \sum_{i,j=1}^n \big(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i}\big)\frac{\partial f}{\partial x^j}.$$

Wir definieren das Vektorfeld $[X,Y] \in C^{k-1}(TU)$ durch (4.6), auf U gilt dann die gewünschte Eigenschaft (4.5). Die Definitionen bezüglich zweier Karten stimmen überein, denn das Differenzvektorfeld ist Null nach Lemma 4.4. Also erhalten wir ein global definiertes Vektorfeld $[X,Y] \in C^{k-1}(TM)$. □

Lemma 4.5 *Für Vektorfelder $X, Y, Z \in C^k(TM)$ mit $k \geq 1$ gilt:*

(i) $[\lambda X + \mu Y, Z] = \lambda[X,Y] + \mu[Y,Z]$ *für alle* $\lambda, \mu \in \mathbb{R}$.

(ii) $[X,Y] = -[Y,X]$.

(iii) $\big[[X,Y],Z\big] + \big[[Y,Z],X\big] + \big[[Z,X],Y\big] = 0$, *falls* $k \geq 2$.

(iii) *Für die Basisfelder einer Karte ist* $[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$.

Beweis: Die Aussagen folgen aus der Definition. Für (iii) berechnen wir mit $f \in C^\infty(M)$:

$$\begin{aligned} \big[[X,Y],Z\big]f &= [X,Y]\,Zf - Z\,[X,Y]f \\ &= \underbrace{XYZf}_{\text{zyklisch}} - \underbrace{YXZf}_{\text{antizyklisch}} - \underbrace{ZXYf}_{\text{zyklisch}} + \underbrace{ZYXf}_{\text{antizyklisch}}\,. \end{aligned}$$

Also heben sich bei zyklischer Vertauschung und Addition alle Terme weg. Behauptung (iv) folgt aus dem Satz von Schwarz über die Vertauschbarkeit der partiellen Ableitungen. □

Bemerkung. Die Lieklammer ist nicht bilinear über $C^\infty(M)$, sondern es gilt

$$(4.7)\qquad [fX,Y] = f\,[X,Y] - (Yf)\,X \quad \text{bzw.} \quad [X,fY] = f[X,Y] + (Xf)Y.$$

Definition 4.6 *Seien M,N differenzierbare Mannigfaltigkeiten und $f \in C^1(M,N)$. Das Vektorfeld $Y : N \to TN$ heißt f-verwandt zu dem Vektorfeld $X : M \to TM$, wenn gilt:*

$$Y \circ f = Df \cdot X \quad \Leftrightarrow \quad Y(f(p)) = Df(p)X(p) \text{ für alle } p \in M.$$

Lemma 4.6 *Sei $f \in C^2(M,N)$. Sind $\widetilde{X},\widetilde{Y} \in C^1(TN)$ f-verwandt zu $X,Y \in C^1(TM)$, so ist $[\widetilde{X},\widetilde{Y}]$ auch f-verwandt zu $[X,Y]$.*

Beweis: In Koordinaten reduziert sich die Sache auf den Fall $f \in C^2(U,V)$ mit $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ offen. Es gilt $\tilde{X}(f(p)) = Df(p)X(p)$ und $\tilde{Y}(f(p)) = Df(p)Y(p)$. Wir berechnen

$$\begin{aligned} D(\tilde{Y} \circ f)(p)\,X(p) &= (D\tilde{Y})(f(p))\,Df(p)\,X(p) \\ &= (D\tilde{Y})(f(p))\,\tilde{X}(f(p)) \\ &= (D\tilde{Y} \cdot \tilde{X})(f(p)). \end{aligned}$$

Andererseits gilt

$$D(Df \cdot Y)(p)\,X(p) = Df(p)\,DY(p)\,X(p) + D^2f(p)(X(p),Y(p)).$$

Durch Vertauschung und Subtraktion folgt, da $D^2f(p)$ symmetrisch ist,

$$\begin{aligned} [\tilde{X},\tilde{Y}](f(p)) &= (D\tilde{Y} \cdot \tilde{X})(f(p)) - (D\tilde{X} \cdot \tilde{Y})(f(p)) \\ &= Df(p)\,\big(DY(p)X(p) - DX(p)Y(p)\big) \\ &= Df(p)\,[X,Y](p). \end{aligned}$$

□

Was ist nun die geometrische Bedeutung der Lieklammer? Wir wollen zeigen, dass das Verschwinden der Lieklammer von zwei Vektorfeldern X,Y – man spricht dann von kommutierenden Vektorfeldern – gleichbedeutend damit ist, dass die zugehörigen Flüsse der Vektorfelder kommutieren. Dazu brauchen wir etwas Vorbereitung.

Lemma 4.7 *Seien M und N differenzierbare Mannigfaltigkeiten, und $f \in C^1(M,N)$. Das Vektorfeld $Y : N \to TN$ sei f-verwandt zu $X : M \to TM$. Ist dann $c : I \to M$ Integralkurve von X, so ist $f \circ c : I \to N$ Integralkurve von Y.*

BEWEIS: $\frac{d}{dt}f(c(t)) = Df(c(t))c'(t) = Df(c(t))X(c(t)) = Y(f(c(t)))$. $\square$

Zu gegebenem X und f muss kein f-verwandtes Feld Y existieren, zum Beispiel kann es Punkte $p_{1,2} \in M$ geben mit $f(p_1) = f(p_2)$, aber $Df(p_1)X(p_1) \neq Df(p_2)X(p_2)$. Auch ist ein f-verwandtes Feld nicht eindeutig bestimmt, es sei denn, f ist surjektiv. Ist $f \in C^1(M,N)$ aber ein Diffeomorphismus, so gibt es genau ein f-verwandtes Feld, nämlich den pushforward von X unter f.

Definition 4.7 *Sei $f \in C^1(M,N)$ ein Diffeomorphismus. Dann gibt es zu jedem Vektorfeld $X : M \to TM$ genau ein f-verwandtes Feld $Y : N \to TN$, Bezeichnung $Y = f_*X$, nämlich*

$$Y(q) = Df(p)X(p)|_{p=f^{-1}(q)}.$$

Für $f \in C^{k+1}(M,N)$ und $X \in C^k(TM)$ ist f_*X in $C^k(TN)$.

Sind $f \in C^1(M,N)$ und $g \in C^1(N,P)$ diffeomorph, so gilt $(g \circ f)_*X = g_*(f_*X)$, denn

$$(g \circ f)_*X((g \circ f)(p) = D(g \circ f)(p)X(p) = Dg(f(p))Df(p)X(p) = g_*(f_*X)(g(f(p))). \tag{4.8}$$

Lemma 4.8 *Sei $f \in C^2(M,N)$ ein Diffeomorphismus, $X \in C^1(TM)$ und $Y = f_*X$. Dann gilt für die zugehörigen Flüsse $N_t^Y = f(M_t^X)$ und $\psi_t \circ f = f \circ \phi_t$.*

BEWEIS: Nach Lemma 4.7 gilt für die Existenzintervalle $I_{f(p)} = I_p$ und damit

$$f(M_t^X) = f\big(\{p \in M : t \in I_p\}\big) \subset \{q \in N : t \in I_q\} = N_t^Y.$$

Aus Symmetriegründen folgt Gleichheit. Für $p \in M$ ist $f \circ \phi_t(p)$, $t \in I_p$, Integralkurve von Y mit $f \circ \phi_t(p)|_{t=0} = f(p)$, also $f \circ \phi_t(p) = \psi_t(f(p))$ für alle $t \in I_p = I_{f(p)}$. $\square$

Theorem 4.4 *Seien $X, Y \in C^2(TM)$ und $\phi : M^X \to M$ der Fluss von X. Dann gilt*

$$\frac{\partial}{\partial t}\big((\phi_{-t})_*Y\big)(p)|_{t=0} = [X,Y](p). \tag{4.9}$$

BEWEIS: Nach Satz 4.2(iii) ist $\phi = \phi(x,t) \in C^2$ auf einer Umgebung von $(p,0)$. Wir berechnen, wobei D die x-Ableitung ist,

$$\begin{aligned}
\frac{\partial}{\partial t}\big(D\phi_{-t}(\phi_t(p))Y(p)|_{t=0} &= \frac{\partial}{\partial t}D\phi(\phi_t(p),-t)Y(p) \\
&= D^2\phi(p,0)(X(p),Y(p)) - \partial_t D\phi(p,0)Y(p).
\end{aligned}$$

Nun gilt $D^2\phi(p,0) = D^2\phi_0(p) = 0$, da $\phi_0(x) = x$. Aus $\partial_t\phi(\cdot,0) = X$ erhalten wir weiter $\partial_t D\phi(p,0) = D\partial_t\phi(p,0) = DX(p)$. Somit ergibt sich mit (4.6)

$$\begin{aligned}
\frac{\partial}{\partial t}\big((\phi_{-t})_*Y(p)\big)|_{t=0} &= \frac{\partial}{\partial t}\big(D\phi_{-t}(\phi_t(p))Y(\phi_t(p))\big)|_{t=0} \\
&= -DX(p)Y(p) + DY(p)X(p) \\
&= [X,Y](p).
\end{aligned}$$

In einer beliebigen differenzierbaren Mannigfaltigkeit M sieht die Sache zunächst kompliziert aus, denn die Ableitung einer Kurve in TM ist im allgemeinen ein Element von $T(TM)$.

Aber die Kurve in (4.9) verläuft in dem festen Vektorraum T_pM, denn Y wird in $\phi_t(p)$ ausgewertet und dann mit $D\phi_{-t}(\phi_t(p)) = D\phi_t(p)^{-1}$ nach T_pM zurück transportiert. Die Ableitung in (4.9) ist damit durch die Ableitung des Koordinatenvektors bezüglich einer Basis von T_pM gegeben.

Wähle eine Karte $\varphi : U \to V \subset \mathbb{R}^n$, und betrachte den Fluss ψ_t des Vektorfelds $\varphi_* X$ auf V. Nach Lemma 4.8 gilt $\psi_t \circ \varphi = \varphi \circ \phi_t$. Für ein Vektorfeld Z auf U und das zugehörigen Koordinatenfeld $Z^\varphi : U \to \mathbb{R}^n$ haben wir

$$\varphi_* Z(x) = D\varphi(p)Z(p)|_{p=\varphi^{-1}(x)} = X^\varphi(\varphi^{-1}(x)).$$

Mit (4.8) berechnen wir nun

$$\big((\phi_{-t})_* Y\big)^\varphi \circ \varphi^{-1} = \varphi_*\big((\phi_{-t})_* Y\big) = (\psi_{-t})_* \, \varphi_* \, Y.$$

Aus dem Fall $M = V \subset \mathbb{R}^n$ und Lemma 4.6 folgt

$$\begin{aligned}
\frac{\partial}{\partial t}\big((\phi_{-t})_* Y\big)^\varphi(p)|_{t=0} &= \frac{\partial}{\partial t}(\psi_{-t})_*(\varphi_* Y)(\varphi(p))|_{t=0} \\
&= [\varphi_* X, \varphi_* Y](\varphi(p)) \\
&= \varphi_*[X,Y](\varphi(p)) \\
&= [X,Y]^\varphi(p).
\end{aligned}$$

Die Behauptung ist allgemein bewiesen. □

Die folgende Aussage gilt auch wenn die Flüsse nicht aug ganz $M \times \mathbb{R}$ definiert sind, der Einfachheit halber setzen wir das aber voraus.

Theorem 4.5 *Die Vektorfelder $X, Y \in C^2(TM)$ seien vollständig integrierbar. Dann sind folgende Aussagen äquivalent:*

(i) $[X,Y] = 0$ *auf* M.

(ii) $\phi_s \circ \psi_t = \psi_t \circ \phi_s$, *wobei* ϕ_s, ψ_t *die Flüsse von* X *bzw.* Y *sind.*

Beweis: Wir zeigen zuerst (ii) ⇒ (i). Und zwar folgt für $p \in M$ durch sukzessive Anwendung von $\frac{\partial}{\partial s}|_{s=0}$ und $\frac{\partial}{\partial t}|_{t=0}$, wobei Satz 4.4 benutzt wird:

$$\begin{aligned}
(\phi_{-t} \circ \psi_s \circ \phi_t)(p) = \psi_s(p) \quad &\Rightarrow \quad D\phi_{-t}(\phi_t(p))Y(\phi_t(p)) = Y(p) \\
&\Rightarrow \quad [X,Y](p) = 0.
\end{aligned}$$

Umgekehrt berechnen wir $t = t_0 + s$, also $\phi_{-t} = \phi_{-t_0} \circ \phi_{-s}$,

$$\begin{aligned}
\frac{\partial}{\partial t}\big((\phi_{-t})_* Y)(p)\big)|_{t=t_0} &= \frac{\partial}{\partial s}\big((\phi_{-(t_0+s)})_* Y(p)\big)|_{s=0} \\
&= (\phi_{-t_0})_* \frac{\partial}{\partial s}\big(\phi_{-s})_* Y(p)\big)|_{s=0} \\
&= (\phi_{-t_0})_*[X,Y](p) \\
&= 0.
\end{aligned}$$

Also ist $(D\phi_{-t} \cdot Y)(\phi_t(p)) = Y(p)$ für alle (p,t). Betrachte nun für $t \in \mathbb{R}$ fest die Kurve $c(s) := (\phi_{-t} \circ \psi_s \circ \phi_t)(p)$. Es gilt $c(0) = p$ und

$$\begin{aligned} c'(s) &= (D\phi_{-t} \cdot Y)(\psi_s \circ \phi_t(p)) \\ &= (D\phi_{-t} \cdot Y)\big(\phi_t\left(\phi_{-t} \circ \psi_s \circ \phi_t)(p)\right)\big) \\ &= Y(c(s)). \end{aligned}$$

Der Eindeutigkeitssatz, siehe Satz 4.2(i)), liefert $c(s) = \psi_s(p)$ für alle $s \in \mathbb{R}$, und folglich $\phi_{-t} \circ \psi_s \circ \phi_t = \psi_s$ für alle $s,t \in \mathbb{R}$. $\square$

5 Alternierende Multilinearformen

Im folgenden sei V ein n-dimensionaler Vektorraum über $\mathbb{R}$, wir denken an den Modellfall $V = \mathbb{R}^n$ und an $V = T_pM$, den Tangentialraum einer n-dimensionalen Mannigfaltigkeit im Punkt p. Wir bezeichnen mit $\otimes^k V$ den Vektorraum der k-linearen Abbildungen

$$\omega : \underbrace{V \times \ldots \times V}_{k \text{ Faktoren}} \to \mathbb{R},\ (v_1, \ldots, v_k) \mapsto \omega(v_1, \ldots, v_k).$$

Die Addition und Skalarmultiplikation ist punktweise gegeben, also

$$\begin{aligned} (\omega + \eta)(v_1, \ldots, v_k) &= \omega(v_1, \ldots, v_k) + \eta(v_1, \ldots, v_k) \\ (\lambda\, \omega)(v_1, \ldots, v_k) &= \lambda\, \omega(v_1, \ldots, v_k). \end{aligned}$$

Definition 5.1 *$\omega \in \otimes^k V$ heißt alternierend, falls für alle Permutationen $\sigma \in S_k$ gilt:*

$$\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = \operatorname{sign}(\sigma)\, \omega(v_1, \ldots, v_k). \tag{5.1}$$

Wir bezeichnen den Raum der alternierenden k-Formen mit $\Lambda^k V$.

Äquivalent lassen sich alternierende Formen wie folgt charakterisieren: Transpositionen (Vertauschungen) haben Vorzeichen -1, also gilt

$$\omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -\omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k) \quad \text{für alle } 1 \leq i < j \leq k. \tag{5.2}$$

Da die Gruppe S_k durch Transpositionen erzeugt wird, folgt umgekehrt (5.1). Mit $v_i = v_j$ ergibt sich nun weiter

$$\omega(v_1, \ldots, \overset{i}{v}, \ldots, \overset{j}{v}, \ldots, v_k) = 0 \quad \text{für alle } 1 \leq i < j \leq k. \tag{5.3}$$

Umgekehrt folgt (5.2) durch Wahl von $v = v_i + v_j$ und Ausmultiplizieren. Schließlich

$$\omega(v_1, \ldots, v_k) = 0 \quad \text{falls } v_1, \ldots, v_k \text{ linear abhängig.} \tag{5.4}$$

Denn ein v_i kann als Linearkombination der anderen $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$ geschrieben werden, und diese treten dann doppelt auf. Die umgekehrte Implikation ist hier trivial.

Wir haben die folgenden, schon bekannten Spezialfälle:

$$\begin{array}{ll} \Lambda^0 V = \mathbb{R} & \text{(per Definition)}, \\ \Lambda^1 V = V^* & \text{(der Dualraum)}, \\ \Lambda^n V = \mathbb{R} \cdot \det & \text{(Lineare Algebra)}, \\ \Lambda^k V = \{0\} \text{ für } k > n & \text{(nach (5.4))}. \end{array}$$

Lemma 5.1 (Alternator) *Sei* $\mathrm{Alt} : \otimes^k V \to \otimes^k V$ *definiert durch*

$$\mathrm{Alt}(\omega)(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sign}(\sigma)\, \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}). \tag{5.5}$$

Dann ist Alt *eine lineare Projektion mit Bild $\Lambda^k V$.*

Beweis: Wir zeigen erst, dass Alt nach $\Lambda^k V$ abbildet. Für $\tau \in S_k$ berechnen wir, indem wir $v_{\tau(i)} =: w_i$ substituieren, also $w_{\sigma(i)} = v_{\tau\sigma(i)}$:

$$\begin{aligned}
\mathrm{Alt}(\omega)(v_{\tau(1)}, \ldots, v_{\tau(k)}) &= = \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sign}\,(\sigma)\omega(v_{\tau\sigma(1)}, \ldots, v_{\tau\sigma(k)}) \\
&= \frac{\mathrm{sign}\,(\tau)}{k!} \sum_{\sigma \in S_k} \mathrm{sign}\,(\tau\sigma)\omega(v_{\tau\sigma(1)}, \ldots, v_{\tau\sigma(k)}) \\
&= \mathrm{sign}\,(\tau)\mathrm{Alt}(\omega)(v_1, \ldots, v_k).
\end{aligned}$$

Im letzten Schritt wurde benutzt, dass die $\tau\sigma$ die Gruppe S_k genau einmal durchlaufen. Es ist klar dass Alt linear ist. Um zu sehen, dass Alt eine Projektion mit Bild $\Lambda^k V$ ist, zeigen wir nun $\mathrm{Alt}|_{\Lambda^k V} = \mathrm{Id}_{\Lambda^k V}$. Und zwar gilt für $\omega \in \Lambda^k V$

$$\begin{aligned}
\mathrm{Alt}(\omega)(v_1, \ldots, v_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sign}\,(\sigma)\, \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} \underbrace{\mathrm{sign}\,(\sigma)\mathrm{sign}\,(\sigma)}_{=1} \omega(v_1, \ldots, v_k) \\
&= \omega(v_1, \ldots, v_k).
\end{aligned}$$

$\square$

Für allgemeine Multilinearformen haben wir das Tensorprodukt

$$\otimes^k V \times \otimes^\ell V \to \otimes^{k+\ell} V, \quad (\omega \otimes \eta)(v_1, \ldots, v_{k+\ell}) = \omega(v_1, \ldots, v_k)\eta(v_{k+1}, \ldots, v_{k+\ell}).$$

Man sieht leicht, dass dieses Produkt assoziativ ist:

$$\begin{aligned}
(\omega \otimes \eta) \otimes \zeta(v_1, \ldots, v_{k+\ell+m}) &= (\omega \otimes \eta)(v_1, \ldots, v_{k+\ell})\, \zeta(v_{k+\ell+1}, \ldots, v_{k+\ell+m}) \\
&= \omega(v_1, \ldots, v_k)\, \eta(v_{k+1}, \ldots, v_{k+\ell})\, \zeta(v_{k+\ell+1}, \ldots, v_{k+\ell+m}) \\
&= \omega \otimes (\eta \otimes \zeta)(v_1, \ldots, v_{k+\ell+m}).
\end{aligned}$$

Wir definieren nun ein Produkt auf $\Lambda^k V$.

Definition 5.2 (Dachprodukt) *Das äußere Produkt (Dachprodukt) ist definiert durch*

$$\Lambda^k V \times \Lambda^\ell V \to \Lambda^{k+\ell} V,\ \omega \wedge \eta = \frac{(k+\ell)!}{k!\ell!} \mathrm{Alt}(\omega \otimes \eta). \tag{5.6}$$

Mit anderen Worten

$$\omega \wedge \eta(v_1, \ldots, v_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} \mathrm{sign}\,(\sigma)\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\eta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+\ell)}). \tag{5.7}$$

Example 5.1 Es ergibt sich

$$\begin{aligned}
&k = 1,\ \ell = 1 \qquad \omega \wedge \eta(v_1, v_2) = \omega(v_1)\eta(v_2) - \omega(v_2)\eta(v_1), \\
&k = 2,\ \ell = 1 \qquad \omega \wedge \eta(v_1, v_2, v_3) = \omega(v_1, v_2)\eta(v_3) + \omega(v_3, v_1)\eta(v_2) + \omega(v_2, v_3)\eta(v_1).
\end{aligned}$$

Prüfe nach, dass $\omega \wedge \eta$ tatsächlich alternierend ist.

Nach Definition ist das Dachprodukt in beiden beiden Faktoren linear, also zum Beispiel

$$(\lambda\omega + \mu\eta) \wedge \zeta = \lambda\,\omega \wedge \zeta + \mu\,\eta \wedge \zeta.$$

Theorem 5.1 (Eigenschaften des Dachprodukts) *Für alternierende Formen $\omega \in \Lambda^k V$, $\eta \in \Lambda^\ell V$, $\zeta \in \Lambda^m V$ gelten folgende Rechenregeln:*

$$\begin{aligned}
&(5.8) \qquad & \omega \wedge \eta &= (-1)^{k\ell}\eta \wedge \omega \quad \textit{Kommutativgesetz (graduiert)},\\
&(5.9) \qquad & (\omega \wedge \eta) \wedge \zeta &= \omega \wedge (\eta \wedge \zeta) \quad \textit{Assoziativgesetz}
\end{aligned}$$

Beweis: Um (5.8) zu zeigen, betrachten wir die Permutation

$$\tau = \begin{pmatrix} 1 & \dots & k & k+1 & \dots & k+\ell \\ \ell+1 & \dots & \ell+k & 1 & \dots & \ell \end{pmatrix}.$$

Damit berechnen wir

$$\begin{aligned}
(\eta \wedge \omega)(v_1, \dots, v_{k+\ell}) &= \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} \operatorname{sign}(\sigma)\eta(v_{\sigma(1)}, \dots, v_{\sigma(\ell)})\,\omega(v_{\sigma(\ell+1)}, \dots, v_{\sigma(\ell+k)})\\
&= \frac{\operatorname{sign}(\tau)}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} \operatorname{sign}(\sigma\tau)\eta(v_{\sigma\tau(k+1)}, \dots, v_{\sigma\tau(k+\ell)})\,\omega(v_{\sigma\tau(1)}, \dots, v_{\sigma\tau(k)})\\
&= (-1)^{kl}\omega \wedge \eta(v_1, \dots, v_{k+\ell}).
\end{aligned}$$

Die Assoziativität ist etwas schwieriger. Sei $I(k, k+\ell)$ die Menge der aufsteigenden Multiindizes $\alpha : \{1, \dots, k\} \to \{1, \dots, k+\ell\}$. Wir bezeichnen mit $\hat{\alpha} \in I(\ell, k+\ell)$ den dazu komplementären Multiindex, ebenfalls aufsteigend. Jede Permutation $\varrho \in S_{k+\ell}$ hat eine eindeutige Darstellung

$$\varrho = (\alpha \circ \sigma, \hat{\alpha} \circ \tau) \quad \text{mit } \alpha \in I(k, k+\ell),\ \sigma \in S_k,\ \tau \in S_\ell,$$

und es gilt $\operatorname{sign}(\varrho) = \operatorname{sign}(\alpha, \hat{\alpha})\operatorname{sign}(\sigma)\operatorname{sign}(\tau)$. Berechne nun für $\gamma \in \otimes^k V$ und $\zeta \in \otimes^\ell V$

$$\begin{aligned}
&\operatorname{Alt}(\gamma \otimes \zeta)(v_1, \dots, v_{k+\ell})\\
= &\frac{1}{(k+\ell)!} \sum_{\alpha \in I(k,k+\ell)} \operatorname{sign}(\alpha, \hat{\alpha}) \cdot\\
&\sum_{\sigma \in S_k, \tau \in S_\ell} \operatorname{sign}(\sigma)\operatorname{sign}(\tau)\,\gamma(v_{\alpha(\sigma(1))}, \dots, v_{\alpha(\sigma(k))})\,\zeta(v_{\hat{\alpha}(\tau(1))}, \dots, v_{\hat{\alpha}(\tau(\ell))})\\
= &\frac{k!\ell!}{(k+\ell)!} \sum_{\alpha \in I(k,k+\ell)} \operatorname{sign}(\alpha, \hat{\alpha})\operatorname{Alt}(\gamma)(v_{\alpha(1)}, \dots, v_{\alpha(k)})\,\operatorname{Alt}(\zeta)(v_{\hat{\alpha}(1)}, \dots, v_{\hat{\alpha}(\ell)}).
\end{aligned}$$

Im letzten Schritt substituiere $v_{\alpha(i)} =: w_i$, also $v_{\alpha(\sigma(i))} = w_{\sigma(i)}$. Die Rechnung zeigt

$$\operatorname{Alt}(\gamma) = 0 \quad \Rightarrow \quad \operatorname{Alt}(\gamma \otimes \zeta) = 0.$$

Diese Aussage brauchen wir: da Alt = Id auf alternierenden Formen, gilt $\operatorname{Alt}(\gamma) = 0$ für die $(k+\ell)$-lineare Form $\gamma = \operatorname{Alt}(\omega \otimes \eta) - \omega \otimes \eta$. Damit berechnen wir

$$\begin{aligned}
0 &= \operatorname{Alt}\big((\operatorname{Alt}(\omega \otimes \eta) - \omega \otimes \eta) \otimes \zeta\big)\\
&= \operatorname{Alt}\big(\operatorname{Alt}(\omega \otimes \eta) \otimes \zeta\big) - \operatorname{Alt}(\omega \otimes \eta \otimes \zeta)\\
&= \frac{k!\ell!m!}{(k+\ell+m)!}(\omega \wedge \eta) \wedge \zeta - \operatorname{Alt}(\omega \otimes \eta \otimes \zeta).
\end{aligned}$$

Bei anderer Klammerung folgt die analoge Formel, das Assoziativgesetzt ist gezeigt. $\square$

Aus dem Beweis ergibt sich für das Dachprodukt von $\gamma \in \Lambda^k V$ und $\zeta \in \Lambda^\ell V$ die Formel

$$\gamma \wedge \zeta(v_1, \ldots, v_{k+\ell}) = \sum_{\alpha \in I(k,k+\ell)} \operatorname{sign}(\alpha, \hat{\alpha}) \gamma\big(v_{\alpha(1)}, \ldots, v_{\alpha(k)}\big) \zeta\big(v_{\hat{\alpha}(1)}, \ldots, v_{\hat{\alpha}(\ell)}\big). \tag{5.10}$$

Die Zahl der Summanden in (5.7) ist $(k+\ell)!$, dagegen sind es hier nur $(k+\ell)!/k!\ell!$.

Theorem 5.2 (Dachprodukte von 1-Formen) *Für $\omega^1, \ldots, \omega^k \in \Lambda^1 V = V^*$ gilt:*

$$(\omega^1 \wedge \ldots \wedge \omega^k)(v_1, \ldots, v_k) = \det\big(\omega^i(v_j)\big). \tag{5.11}$$

$$\omega^1, \ldots, \omega^k \textit{ linear unabhängig} \quad \Leftrightarrow \quad \omega^1 \wedge \ldots \wedge \omega^k \neq 0. \tag{5.12}$$

Beweis: Wir zeigen (5.11) durch Induktion über k. Bezeichne mit $\hat{\jmath} = (1, \ldots, j-1, j+1, \ldots, k)$ den zu j komplementären, aufsteigenden $(k-1)$-index. Es gilt dann, vgl. vorangehender Beweis und Definition des Dachprodukts,

$$\begin{aligned} \omega^1 \wedge (\omega^2 \ldots \wedge \omega^k)(v_1, \ldots, v_k) &= \sum_{j=1}^k \operatorname{sign}(j, \hat{\jmath}) \omega^1(v_j) \omega^2 \wedge \ldots \wedge \omega^k(v_1, \ldots, \widehat{v_j}, \ldots, v_k) \\ &= \sum_{j=1}^k (-1)^{j-1} \omega^1(v_j) \det\big(\omega^i(v_\ell)\big)_{2 \leq i \leq k,\, \ell \neq j} \\ &= \det(\omega^i(v_j)). \end{aligned}$$

Die $\omega^1(v_j)$ sind die Einträge der ersten Zeile, diese werden mit den Determinanten multipliziert,die sich durch Streichen der ersten Zeile und j-ten Spalte ergeben. Der letzte Schritt gilt also nach dem Laplace-Entwicklungssatz.

In (5.12) ist $\Leftarrow$ einfach. Umgekehrt zeigen wir, dass zu $\omega^1, \ldots, \omega^k$ linear unabhängig $v_1, \ldots, v_k$ existieren mit $\omega^i(v_j) = \delta^i_j$, also

$$\omega^1 \wedge \ldots \wedge \omega^k(v_1, \ldots, v_k) = \det(\omega^i(v_j)) = 1.$$

Wir können dazu $k = n$ annehmen, sonst ergänze die ω^i zu einer Basis. Sei $\phi_1, \ldots, \phi_n$ die duale Basis in V^{**}, also $\phi_j(\omega^i) = \delta^i_j$. Die kanonische Abbildung $J : V \to V^{**}$ ist surjektiv, also gibt es $v_j \in V$ mit $Jv_j = \phi_j$, das heißt

$$\omega^i(v_j) = Jv_j(\omega^i) = \phi_j(\omega^i) = \delta^i_j.$$

$\square$

Theorem 5.3 (Basis von $\Lambda^k V$) *Sei $e_1, \ldots, e_n$ Basis von V mit dualer Basis $e^1, \ldots, e^n \in V^*$. Dann hat $\Lambda^k V$ die Basis $e^I = e^{i_1} \wedge \ldots \wedge e^{i_k}$ mit $1 \leq i_1 < \ldots < i_k \leq n$. Insbesondere gilt $\dim \Lambda^k V = \binom{n}{k}$, und jedes $\omega \in \Lambda^k V$ hat die Darstellung*

$$\omega = \sum_{I \in I(k,n)} \omega_I \, e^I \qquad \textit{mit } \omega_I = \omega(e_{i_1}, \ldots, e_{i_k}). \tag{5.13}$$

Beweis: Betrachte für aufsteigende Multiindizes $I, J \in I(k,n)$ die Matrix $\big(e^{i_\ell}(e_{j_m})\big) \in \mathbb{R}^{k\times k}$. Ist $I = J$, also $i_r = j_r$ für $r = 1,\ldots,k$, so ist $e^{i_\ell}(e_{j_m}) = \delta_{\ell m}$. Andernfalls sei $r \in \{1,\ldots,k\}$ der kleinste Index mit $i_r \neq j_r$. Im Fall $i_r > j_r$ ist $j_r \notin \{i_1,\ldots,i_k\}$, also ist die r-te Spalte der Matrix Null. Im Fall $i_r < j_r$ ist $i_r \notin \{j_1,\ldots,j_k\}$ und die r-te Zeile ist Null. Damit gilt

$$e^I(e_J) = \det\big(e^{i_\ell}(e_{j_m})\big)_{1\le \ell,m\le k} = \delta^I_J. \tag{5.14}$$

Für $\omega \in \Lambda^k V$ behaupten wir nun

$$\omega = \sum_{I\in I(k,n)} \omega_I e^I \quad \text{ mit } \omega_I = \omega(e_{i_1},\ldots,e_{i_k}).$$

Da beide Seiten multilinear und alternierend sind, müssen wir das nur auf $e_J = (e_{j_1},\ldots,e_{j_k})$, $J \in I(k,n)$, nachprüfen, dann folgt es aber direkt aus (5.14). Ist andererseits $\omega = \sum_{I\in I(k,n)} c_I e^I = 0$ mit beliebigen $c_I \in \mathbb{R}$, so folgt $c_J = \omega(e_J) = 0$ für alle $J \in I(k,n)$. □

Definition 5.3 (Pullback) *Seien V, W Vektorräume über $\mathbb{R}$ und $A \in L(V,W)$.*

(a) *Die transponierte (oder duale) Abbildung zu A ist*

$$A^* \in L(W^*, V^*),\ \ A^*\psi(v) = \psi(Av).$$

(b) *Der Pullback von k-Formen ω unter A ist*

$$A^* \in L(\Lambda^k W, \Lambda^k V),\ \ A^*\omega(v_1,\ldots,v_k) = \omega(Av_1,\ldots,Av_k).$$

Seien $v_1,\ldots,v_n$ bzw. $w_1,\ldots,w_m$ Basen von V und W, und $w^1,\ldots,w^m \in W^*$ die duale Basis. Für $A \in L(V,W)$ sei $A^i_j = w^i(Av_j)$. Für Multiindizes $I \in I(k,m)$ und $J \in I(k,n)$ folgt dann

$$A^*w^I(v_J) = w^I(Av_{j_1},\ldots,Av_{j_k}) = \det\big(w^{i_\ell}(Av_{j_m})\big) = \det\big(A^{i_\ell}_{j_m}\big).$$

Das ist die Determinante der $k\times k$ Submatrix von $A \in \mathbb{R}^{m\times n}$, die sich durch Auswahl der Zeilen $1 \le i_1 < \ldots < i_k \le m$ und der Spalten $1 \le j_1 < \ldots < j_k \le n$ ergibt. Diese Subdeterminante wird auch als $I\times J$-Minor der Matrix A bezeichnet.

Lemma 5.2 *Für den Pullback gelten folgende Rechenregeln:*

(i) *Sei $A \in L(V,W)$. Dann gilt für $\omega,\eta \in \Lambda^k W$ und $\lambda,\mu \in \mathbb{R}$*

$$A^*(\lambda\omega + \mu\eta) = \lambda\, A^*\omega + \mu\, A^*\eta.$$

(ii) *Sei $A \in L(V,W)$. Dann gilt für $\omega \in \Lambda^k W$ und $\eta \in \Lambda^\ell W$*

$$A^*(\omega\wedge\eta) = (A^*\omega)\wedge(A^*\eta).$$

(iii) *Sei $A \in L(V,W)$ und $B \in L(W,Z)$. Dann gilt $(BA)^* = A^*B^*$.*

Beweis: Regel (i) ist klar. Für (ii) berechnen wir mit der Definition des Dachprodukts

$$\begin{aligned}
A^*(\omega \wedge \eta)(v_1, \ldots, v_{k+\ell}) &= (\omega \wedge \eta)(Av_1, \ldots, Av_{k+\ell}) \\
&= \frac{1}{k!\,\ell!} \sum_{\sigma \in S_{k+\ell}} \operatorname{sign}(\sigma)\, \omega(Av_{\sigma(1)}, \ldots, Av_{\sigma(k)})\, \eta(Av_{\sigma(k+1)}, \ldots, Av_{\sigma(k+\ell)}) \\
&= \frac{1}{k!\,\ell!} \sum_{\sigma \in S_{k+\ell}} \operatorname{sign}(\sigma)\, A^*\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\, A^*\eta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+\ell)}) \\
&= (A^*\omega) \wedge (A^*\eta)(v_1, \ldots, v_{k+\ell}).
\end{aligned}$$

Wir zeigen schließlich (iii),und zwar gilt für $\zeta \in \Lambda^k Z$

$$(BA)^*\zeta(v_1, \ldots, v_k) = \zeta(BAv_1, \ldots, BAv_k) = B^*\zeta(Av_1, \ldots, Av_k) = A^*(B^*\zeta)(v_1, \ldots, v_k).$$

$\square$

6 Differentialformen

Definition 6.1 (Bündel $\Lambda^k TM$) *Sei M eine n-dimensionale differenzierbare Mannigfaltigkeit. Für $k = 0, 1, \ldots, n$ setzen wir*

$$\Lambda^k TM = \bigcup_{p \in M} \Lambda^k T_p M$$

$\Lambda^k TM$ ist ein differenzierbares Vektorbündel vom Rang $\binom{n}{k}$. Die Präzisierung dieser Aussage geschieht analog zum Tangentialbündel. Wir haben zunächst die Bündelprojektion

$$\pi : \Lambda^k TM \to M,\ \pi(\omega) = p \quad \text{für } \omega \in \Lambda^k T_p M.$$

Die Faser $\pi^{-1}\{p\} = \Lambda^k T_p M$ ist ein Vektorraum der Dimension $\binom{n}{k}$ nach Satz 5.3 . Bezüglich einer Karte $\varphi : U \to \varphi(U)$ mit $p \in U$ hat $\omega \in \Lambda^k T_p M$ die Koordinaten

$$\omega_I := \omega\Big(\frac{\partial}{\partial x^{i_1}}(p), \ldots, \frac{\partial}{\partial x^{i_k}}(p)\Big) \quad \text{mit } I = (i_1, \ldots, i_k) \in I(k, n).$$

Damit ergeben sich die Bündelkarten, mit $N = \binom{n}{k}$,

$$\phi : \pi^{-1}(U) \to U \times \mathbb{R}^N,\ \phi(\omega) = \big(\, p\, ,\, (\omega_I)_{I \in I(k,n)}\big)$$

Eine Menge $\Omega \subset \Lambda^k TM$ ist genau dann offen, wenn $\phi(\Omega \cap \pi^{-1}(U))$ offen in $U \times \mathbb{R}^N$ ist für alle Karten im maximalem Atlas $\mathcal{A}$. Eine Folge $\omega^\nu \in \Lambda^k TM$ konvergiert genau dann gegen $\omega \in \Lambda^k TM$, wenn Folgendes gilt:

- Für die Fußpunkte $p^\nu = \pi(\omega^\nu)$ und $p = \pi(\omega)$ gilt $p^\nu \to p$.
- Sei $\varphi : U \to \varphi(U)$ Karte mit $p \in U$. Dann gilt $\omega^\nu_I \to \omega_I$.

Definition 6.2 (Differentialform) *Einen Schnitt $\omega : M \to \Lambda^k TM$, also $\omega(p) \in \Lambda^k T_p M$ für alle p, bezeichnet man als alternierende Differentialform vom Grad k auf M (kurze Bezeichnung: k-Form).*

Wir haben folgende Spezialfälle:

$k = 0$: Wir hatten nach Definition $\Lambda^0 T_p M = \mathbb{R}$. Die Formen vom Grad $k = 0$ sind damit die Funktionen $f : M \to \mathbb{R}$.

$k = 1$: $\Lambda^1 T_p M$ ist der Raum der Linearformen $\omega : T_p M \to \mathbb{R}$, mit anderen Worten der Dualraum $(T_p M)^*$. Das Bündel $\Lambda^1 TM$ ist damit identisch mit dem Kotangentialbündel

$$T^*M = \bigcup_{p \in M} T^*_p M, \qquad \text{wobei } T^*_p M = (T_p M)^*.$$

Für eine Funktion $f \in C^1(M)$ ist das Differential $df(p) : T_p M \to \mathbb{R}$ linear, also gilt $df(p) \in (T_p M)^*$ und damit ist $df : M \to T^*M$ eine 1-Form.* Ist $\varphi : U \to \varphi(U)$ eine Karte, so haben wir auf U die 1-Formen

$$dx^i : U \to T^*U \subset T^*M,\ dx^i(p)v = (D\varphi(p)v)^i. \tag{6.1}$$

*Wir haben bisher $Df(p)$ geschrieben, für reelle Funktionen ist $df(p)$ aber üblich, siehe unten.

Die $dx^i(p)$ sind die duale Basis zu den $\frac{\partial}{\partial x^j}(p)$, denn

$$dx^i(p)\frac{\partial}{\partial x^j}(p) = \big(\underbrace{D\varphi(p)\frac{\partial}{\partial x^j}(p)}_{=e_j}\big)^i = \delta^i_j.$$

Somit hat jede 1-Form $\omega : M \to T^*M$ auf U die lokale Darstellung

$$\omega = \sum_{i=1}^n \omega_i\, dx^i \quad \text{ mit } \omega_i = \omega\Big(\frac{\partial}{\partial x^i}\Big) : U \to \mathbb{R}.$$

Nach Definition der Bündelkarten ist $\omega \in C^r(T^*M)$ genau wenn $\omega_i \in C^r(U)$ ist, für alle Karten φ und alle $i = 1, \ldots, n$.

Für k beliebig erhalten wir nun die Basisformen, bezüglich einer Karte $\varphi : U \to \varphi(U)$,

$$dx^I : U \to \Lambda^k(TU) \subset \Lambda^k(TM),\ dx^I = dx^{i_1} \wedge \ldots \wedge dx^{i_k}.$$

Jede k-Form ω hat auf U die Darstellung

$$\omega = \sum_{I \in I(k,n)} \omega_I\, dx^I \quad \text{ mit } \omega_I = \omega\Big(\frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_k}}\Big).$$

Definition 6.3 (Pullback von Differentialformen) *Sei $f \in C^1(M,N)$. Der Pullback der k-Form η auf N ist die k-Form $f^*\eta(p) = Df(p)^*\eta(f(p))$ auf M, bzw.*

$$f^*\eta(p)(v_1, \ldots, v_k) = \eta(f(p))(Df(p)v_1, \ldots, Df(p)v_k).$$

Lemma 6.1 *Für den Pullback von Differentialformen gelten folgende Rechenregeln, wobei jeweils $f \in C^1(M,N)$ ist:*

(a) $f^*(\lambda\omega + \mu\eta) = \lambda f^*\omega + \mu f^*\eta$ *(ω, η k-Formen auf N, $\lambda, \mu \in \mathbb{R}$).*

(b) $f^*(\omega \wedge \eta) = f^*\omega \wedge f^*\eta$ *(ω, η eine k-Form bzw. ℓ-Form auf N).*

(c) $(g \circ f)^*\zeta = f^*(g^*\zeta)$ *($g \in C^1(N,P)$ und ζ k-Form auf P).*

Wir wollen den Pullback auch in Koordinaten berechnen. Sei also $f \in C^1(M,N)$, mit $\dim M = m$ und $\dim N = n$, und seien $\varphi : U \to \varphi(U)$ sowie $\psi : V \to \psi(V)$ lokale Koordinaten auf M bzw. N, so dass $f(U) \subset V$. Sei $\eta = \sum_{J \in I(k,n)} \eta_J dy^J$ die Koordinatendarstellung einer k-Form auf N. Dann gilt nach Kettenregel

$$f^*dy^j = dy^j \circ Df = d(y^j \circ f) =: df^j.$$

Mit Aussage (b) von Lemma 6.1 folgt dann

$$f^*\eta = \sum_{J \in I(k,n)} \eta_J \circ f\, f^*dy^J = \sum_{J \in I(k,n)} \eta_J \circ f\, df^J, \tag{6.2}$$

Dabei ist $df^J = df^{j_1} \wedge \ldots \wedge df^{j_k}$ ein Produkt von 1-Formen. Wenden wir das auf die Vektoren $\frac{\partial}{\partial x^I}$ mit $I = (i_1, \ldots, i_k) \in I(k,m)$ an, so folgt mit Satz 5.2 die lokale Darstellung

$$f^*\eta = \sum_{I \in I(k,m)} \sum_{J \in I(k,n)} \eta_J \circ f\, \det\Big(\frac{\partial f^J}{\partial x^I}\Big)\, dx^I. \tag{6.3}$$

Theorem 6.1 (die äußere Ableitung d) *Sei M eine n-dimensionale differenzierbare Mannigfaltigkeit. Es gibt Operatoren $d : C^1(\Lambda^k TM) \to C^0(\Lambda^{k+1}TM)$, für $k = 0, 1, \ldots, n$, mit folgenden Eigenschaften:*

(a) *Sei $\omega = \sum_{I \in I(k,n)} \omega_I \, dx^I$ bezüglich einer Karte $\varphi : U \to \varphi(U)$. Dann gilt auf U*

$$d\omega = \sum_{i=1}^n dx^i \wedge \frac{\partial \omega}{\partial x^i} \quad \text{wobei} \quad \frac{\partial \omega}{\partial x^i} = \sum_{I \in I(k,n)} \frac{\partial \omega_I}{\partial x^i} \, dx^I. \tag{6.4}$$

Für $f \in C^1(M)$ ist df das Differential, also $df = \sum_{i=1}^n \frac{\partial f}{\partial x^j} \, dx^j$.

(b) *Es gilt $d^2 = 0$, genauer $d(d\omega) = 0$ für $\omega \in C^2(\Lambda^k TM)$.*

(c) *Es gilt die Produktregel*

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta \quad \text{für } \omega \in C^1(\Lambda^k TM),\ \eta \in C^1(\Lambda^\ell TM). \tag{6.5}$$

(d) *Ist $f \in C^2(M, N)$ mit $\dim N = m$, so gilt*

$$d(f^*\eta) = f^*(d\eta) \quad \text{für } \eta \in C^1(\Lambda^k TN). \tag{6.6}$$

Es ist unter anderem zu zeigen, dass d so wohldefiniert ist, also nicht von der Karte abhängt. Für n-Formen ω ist jedenfalls $d\omega = 0$, denn $\Lambda^{n+1}TM = 0$.
Beweis: *Schritt 1:* Für d^φ wie in (a) definiert gilt Regel (b).

$$\begin{aligned} d^\varphi(d^\varphi \omega) &= \sum_{i=1}^n dx^i \wedge \frac{\partial}{\partial x^i}\Big(\sum_{j=1}^n dx^j \wedge \frac{\partial \omega}{\partial x^j}\Big) \\ &= \sum_{i,j=1}^n dx^i \wedge dx^j \wedge \frac{\partial^2 \omega}{\partial x^i \partial x^j} \\ &= \frac{1}{2} \sum_{i,j=1}^n (dx^i \wedge dx^j + dx^j \wedge dx^i) \wedge \frac{\partial^2 \omega}{\partial x^i \partial x^j} \qquad \text{(Satz von Schwarz)} \\ &= 0. \end{aligned}$$

Schritt 2: Für d^φ wie in (a) definiert gilt Regel (c).

$$\begin{aligned} d^\varphi(\omega \wedge \eta) &= \sum_{i=1}^n dx^i \wedge \Big(\frac{\partial \omega}{\partial x^i} \wedge \eta + \omega \wedge \frac{\partial \eta}{\partial x^i}\Big) \\ &= \Big(\sum_{i=1}^n dx^i \wedge \frac{\partial \omega}{\partial x^i}\Big) \wedge \eta + (-1)^k \omega \wedge \Big(\sum_{i=1}^n dx^i \wedge \frac{\partial \eta}{\partial x^i}\Big) \\ &= d^\varphi \omega \wedge \eta + (-1)^k \omega \wedge d^\varphi \eta. \end{aligned}$$

Schritt 3: Seien $\varphi : U \to \varphi(U)$ und $\psi : V \to \psi(V)$ Karten von M bzw. N mit $f(U) \subset V$. Sind d^φ auf U sowie d^ψ auf V jeweils durch (a) definiert, so gilt $d^\varphi(f^*\eta) = f^*(d^\psi \eta)$.
Dazu berechnen wir für $\eta = \sum_{J \in I(k,n)} \eta_J \, dy^J$ zuerst

$$d^\psi \eta = \sum_{j=1}^n dy^j \wedge \sum_{J \in I(k,n)} \frac{\partial \eta_J}{\partial y^j} dy^J = \sum_{J \in I(k,n)} d\eta_J \wedge dy^J.$$

Jetzt verwenden wir Formel (6.2) für den Pullback, es gilt

$$\begin{aligned}
d^\varphi(f^*\eta) &= d^\varphi\Big(\sum_{J\in I(,n)} \eta_J\circ f\, d^\varphi f^{j_1}\wedge\ldots\wedge d^\varphi f^{j_k}\Big) \qquad (\text{beachte } df^j = d^\varphi f^j \text{ nach (a)})\\
&= \sum_{J\in I(k,n)} d^\varphi(\eta_J\circ f)\wedge d^\varphi f^{j_1}\wedge\ldots\wedge d^\varphi f^{j_k} \qquad (\text{Regeln (c) und (b)})\\
&= \sum_{J\in I(k,n)} d(\eta_J\circ f)\wedge df^{j_1}\wedge\ldots\wedge df^{j_k} \qquad (\text{wieder nach (a)})\\
&= \sum_{J\in I(k,n)} f^*(d\eta_J)\wedge f^*(dy^{j_1})\wedge\ldots\wedge f^*(dy^{j_k}) \qquad (\text{Kettenregel})\\
&= f^*(d^\psi\eta) \qquad (\text{Lemma 5.2(b)}).
\end{aligned}$$

Schritt 4: Wohldefiniertheit des Operators d
Seien $\varphi : U\to\varphi(U)$ und $\psi : V\to\psi(V)$ zwei Karten von M mit $U\cap V\neq\emptyset$. Wir wenden Schritt 3 an mit $f = \mathrm{id}_M$. Sind d^φ bzw. d^ψ die Operatoren nach (a), so folgt

$$d^\varphi\eta = d^\varphi(\mathrm{id}_M^*\eta) = \mathrm{id}_M^* d^\psi\eta = d^\psi\eta.$$

Damit ist der Satz bewiesen. □

In Schritt 3 trat eine alternative Formel für d auf, und zwar

$$(6.7)\qquad d\omega = \sum_{i=1}^n dx^i\wedge\sum_{I\in I(k,n)}\frac{\partial\omega_I}{\partial x^i}\,dx^I = \sum_{I\in I(k,n)}\sum_{i=1}^n\frac{\partial\omega_I}{\partial x^i}\,dx^i\wedge dx^I = \sum_{I\in I(k,n)} d\omega_I\wedge dx^I.$$

Example 6.1 Die Vektoranalysis im $\mathbb{R}^n$, insbesondere $n=3$, ist ein Spezialfall der äußeren Ableitung. Dazu führen wir folgende Notation ein (lies: v *eingesetzt in* ω):

$$(6.8)\qquad (v\llcorner\omega)(v_1,\ldots,v_{k-1}) = \omega(v,v_1,\ldots.v_{k-1}) \quad \text{für } \omega\in\Lambda^k V \text{ und } v,v_1,\ldots,v_k\in V.$$

Sei U offene Teilmenge des $\mathbb{R}^n$ mit Karte id_U. Für ein Vektorfeld $X\in C^1(U,\mathbb{R}^n)$ gilt

$$\begin{aligned}
X\llcorner(dx^1\wedge\ldots\wedge dx^n) &= \sum_{j=1}^n X^j e_j\llcorner(dx^1\wedge\ldots\wedge dx^n)\\
&= \sum_{j=1}^n (-1)^{j-1}X^j dx^1\wedge\ldots\wedge\widehat{dx}^j\wedge\ldots\wedge dx^n.
\end{aligned}$$

Es folgt für die äußere Ableitung

$$\begin{aligned}
d\big(X\llcorner(dx^1\wedge\ldots\wedge dx^n)\big) &= \sum_{i=1}^n dx^i\wedge\sum_{j=1}^n(-1)^{j-1}\frac{\partial X^j}{\partial x^i}\,dx^1\wedge\ldots\wedge\widehat{dx}^j\wedge\ldots\wedge dx^n\\
&= \sum_{i=1}^n\frac{\partial X^i}{\partial x^i}\,dx^1\wedge\ldots\wedge dx^n\\
&= \operatorname{div} X\,dx^1\wedge\ldots\wedge dx^n.
\end{aligned}$$

Im Fall $n = 3$ setze $\xi = \sum_{j=1}^{3} \xi_j dx^j$ mit $\xi_j = X^j$. Dann ergibt sich

$$\begin{aligned} d\xi &= \sum_{i=1}^{3} dx^i \wedge \sum_{j=1}^{3} \frac{\partial \xi_j}{\partial x^i}\, dx^j \\ &= \Big(\frac{\partial \xi_3}{\partial x^2} - \frac{\partial \xi_2}{\partial x^3}\Big)\, dx^2 \wedge dx^3 + \Big(\frac{\partial \xi_1}{\partial x^3} - \frac{\partial \xi_3}{\partial x^1}\Big)\, dx^3 \wedge dx^1 + \Big(\frac{\partial \xi_2}{\partial x^1} - \frac{\partial \xi_1}{\partial x^2}\Big)\, dx^1 \wedge dx^2 \\ &= \operatorname{rot} X \llcorner (dx^1 \wedge dx^2 \wedge dx^3). \end{aligned}$$

Insbesondere ergibt sich

$$\begin{aligned} 0 &= d(df) = \operatorname{rot} \operatorname{grad} f \llcorner dx^1 \wedge dx^2 \wedge dx^3, \\ 0 &= d(d\xi) = \operatorname{div} \operatorname{rot} X\, dx^1 \wedge dx^2 \wedge dx^3. \end{aligned}$$

Definition 6.4 (de-Rham-Kohomologie) *Für eine differenzierbare Mannigfaltigkeit ist der k-te de-Rham-Kohomologieraum*[†] *der Quotientenvektorraum*

$$H^k_{DR}(M) = Z^k(M)/B^k(M),$$

wobei $Z^k(M)$ der Kern von d und $B^k(M)$ das Bild von d ist, genauer

$$\begin{aligned} Z^k(M) &= \{\omega \in C^\infty(\Lambda^k TM) : d\omega = 0\}, \\ B^k(M) &= \{d\eta : \eta \in C^\infty(\Lambda^{k-1} TM)\}. \end{aligned}$$

Beachte, dass $B^k(M)$ ein Untervektorraum von $Z^k(M)$ ist wegen $d(d\eta) = 0$.

Man kann die Definition der Kohomologie durch das Problem motivieren, ob eine gegebene Form $\omega \in C^\infty(\Lambda^k TM)$ eine Stammform $\eta \in C^\infty(\Lambda^{k-1} TM)$ hat, also $d\eta = \omega$. Wegen $d(d\eta) = 0$ ist dazu die Bedingung $d\omega = 0$ offensichtlich notwendig. Wir werden unten zeigen, dass diese Bedingung lokal auch hinreichend ist (Lemma von Poincaré). Aber global gibt es eventuell zusätzliche Bedingungen, die mit der Topologie von M zu tun haben. Diese werden durch die Kohomologie beschrieben. Eine solche Situation ist schon aus der Theorie der Kurvenintegrale oder der komplexen Funktionentheorie bekannt. Übrigens nennt man ω geschlossen wenn $d\omega = 0$ bzw. $\omega \in Z^k(M)$, und exakt, wenn es eine $(k-1)$-Form η gibt mit $\omega = d\eta$, also $\omega \in B^k(M)$. Das folgende Lemma zeigt, die Kohomologie von M ist eine Diffeomorphieinvariante. Nach dem Satz von de Rham hängt $H^k_{DR}(M)$ tatsächlich nicht von der differenzierbaren Struktur ab, der Raum stimmt mit der singulären Kohomologie von M überein.

Lemma 6.2 *Eine Abbildung $f \in C^\infty(M,N)$ induziert die lineare Abbildung*

$$f^* : H^k_{DR}(N) \to H^k_{DR}(M),\ f^*[\omega] = [f^*\omega].$$

Ist $g \in C^\infty(N,P)$ so gilt $(g \circ f)^* = f^* \circ g^*$. Für einen Diffeomorphimus $f \in C^\infty(M,N)$ ist $f^* : H^k_{DR}(N) \to H^k_{DR}(M)$ ein Isomorphismus.

Beweis: Es gilt $f^* d\eta = df^*\eta$, also $[f^* d\eta] = 0$ in $H^k_{DR}(M)$. Somit ist f^* wohldefiniert und linear. Weiter gilt

$$(g \circ f)^*[\zeta] = [(g \circ f)^*\zeta] = [(g \circ f)^*\zeta] = [f^*(g^*\zeta)] = f^*(g^*[\zeta]).$$

[†]Georges de Rham, Sur l'analysis situs des varietés á n dimensions, Dissertation Paris 1931

Ist f Diffeomorphismus und $g = f^{-1}$, so folgt

$$\mathrm{id}_{H^k_{DR}(M)} = (f \circ g)^* = g^* \circ f^* \quad \text{und analog} \quad \mathrm{id}_{H^k_{DR}(N)} = f^* \circ g^*.$$

□

Wir werden später zeigen, dass $H^k_{DR}(M)$ endlichdimensional ist für kompakte Mannigfaltigkeiten M. Damit das Problem der Stammform lösbar ist, kommen also zur Gleichung $d\omega = 0$ nur endlich viele Bedingungen hinzu. Die Zahlen $b_k(M) = \dim H^k_{DR}(M)$ heißen Bettizahlen. [‡]. Der Satz von de Rham besagt, dass die Bettizahlen Homeomorphieinvarianten sind. Die Euler-Poincaré-Charakteristik ist die Wechselsumme

$$\chi(M) = \sum_{k=0}^{n} (-1)^k \, b_k(M).$$

Example 6.2 Für $k = 0$ gilt $B^0(M) = \{0\}$ per Definition, also ist $H^0_{DR}(M) = Z^0(M)$ der Raum der lokal konstanten Funktionen. Sei $C(M)$ die Menge der Zusammenhangskomponenten von M, und

$$\phi : H^0_{DR}(M) \to \mathrm{Abb}(C(M), \mathbb{R}), \quad \phi(f)(M_p) = f(p).$$

Dabei bezeichnet M_p die Komponente von p, die Abbildung ist wohldefiniert da f auf der Komponente konstant ist. Ist $\phi(f) = 0$, so ist f auf jeder Komponente gleich Null und damit die Nullfunktion. Zu gegebener Abbildung $\varphi : C(M) \to \mathbb{R}$ könen wir f lokal konstant wählen mit $f|_{M'} = \varphi(M')$ für jede Komponente M'. Damit ist ϕ ein Isomorphismus. Ist $\#C(M) = \ell \in \mathbb{N}$, so ist $H^0_{DR}(M)$ isomorph zu $\mathbb{R}^\ell$, also $b_0(M) = \ell$.

Example 6.3 Sei $I \subset \mathbb{R}$ offenes Intervall. Es gilt dann $Z^1(I) = C^\infty(I, \Lambda^1\mathbb{R})$, und jede solche Form γ hat die Darstellung $\gamma(x) = g(x)\,dx$ mit $g \in C^\infty(I)$. Die (glatten) Nullformen sind die Funktionen $f \in C^\infty(\mathbb{R})$, und $df(x) = f'(x)\,dx$. Es folgt $Z^1(I) = B^1(I)$ und $H^0_{DR}(I) = \{0\}$, denn für $x_0 \in I$ beliebig gilt

$$f(x) = \int_{x_0}^{x} g(x)\,dx \quad \Rightarrow \quad f'(x) = g(x), \text{ also } df = \gamma.$$

Example 6.4 Sei $c : [0, 2\pi] \to \mathbb{R}^2$, $c(t) = (\cos t, \sin t)$. Definiere mit dem Kurvenintegral längs c die Abbildung

$$\phi : H^1_{DR}(\mathbb{R}^2 \backslash \{0\}) \to \mathbb{R}, \, \phi[\omega] = \frac{1}{2\pi} \int_c \omega.$$

Dann ist ϕ ein Isomorphismus.

Wir wollen nun sehen, wie sich die Kohomologie unter Homotopien verhält. Betrachte dazu $[0,1] \times M$, für eine gegebene n-dimensionale Mannigfaltigkeit M. Ist $\varphi : U \to \varphi(U)$ Karte von M, so haben wir die induzierte Karte

$$\overline{\varphi} : [0,1] \times U \to [0,1] \times \varphi(U), \overline{\varphi}(t,p) = (t, \varphi(p)).$$

Bezeichnen wir die t-Koordinate mit dem Index $i = 0$, so lautet eine k-Form η

$$\eta = \sum_{I \in I(k-1,n)} \eta_{0I}(t,p)\,dt \wedge dx^I + \sum_{J \in I(k,n)} \eta_J(t,p)\,dx^J =: dt \wedge \alpha + \beta.$$

[‡]Enrico Betti, Sopra gli spazi di un numero qualunque di dimensioni, Annali di Matematica 1870

Diese Zerlegung hängt nicht von der Karte φ ab, und zwar gilt

$$\alpha = \frac{\partial}{\partial t} \llcorner \eta, \qquad \text{und dann } \beta = \eta - dt \wedge \alpha.$$

Sei $d_x = \sum_{j=1}^n dx^j \wedge \frac{\partial}{\partial x^j}$ die *partielle* äußere Ableitung. Wir berechnen

$$d\eta = dt \wedge \frac{\partial \eta}{\partial t} + d_x \eta = -dt \wedge d_x \alpha + dt \wedge \frac{\partial \beta}{\partial t} + d_x \beta.$$

Es folgt

$$\frac{\partial}{\partial t} \llcorner d\eta = -d_x \alpha + \frac{\partial \beta}{\partial t} \qquad \text{und} \qquad d\big(\frac{\partial}{\partial t} \llcorner \eta\big) = d\alpha = dt \wedge \frac{\partial \alpha}{\partial t} + d_x \alpha.$$

Damit haben wir die Formel

$$\frac{\partial \eta}{\partial t} = dt \wedge \frac{\partial \alpha}{\partial t} + \frac{\partial \beta}{\partial t} = \frac{\partial}{\partial t} \llcorner d\eta + d\big(\frac{\partial}{\partial t} \llcorner \eta\big). \tag{6.9}$$

Wir wollen nun die t-Variable ausintegrieren. Sei $i_t : M \to [0,1] \times M$, $i_t(p) = (t,p)$. Definiere die Abbildung $I : C^\infty(\Lambda^k T([0,1] \times M)) \to C^\infty(\Lambda^{k-1} TM)$ durch

$$I\eta(p)(v_1, \ldots, v_{k-1}) = \int_0^1 \eta(t,p)\big(\frac{\partial}{\partial t}, v_1, \ldots, v_{k-1}\big)\, dt. \tag{6.10}$$

Dabei identifizieren wir $v_i \in T_pM$ mit $(0, v_i) \in T_{(t,p)}([0,1] \times M)$. Alternativ gilt die Darstellung

$$I\eta(p) = \int_0^1 \underbrace{(i_t)^*\big(\frac{\partial}{\partial t} \llcorner \eta\big)(p)}_{\in \Lambda^{k-1} T_pM}\, dt.$$

Mit Parameterdifferentiation und Vertauschung von d und i_t^* ergibt sich

$$\begin{aligned} d(I\eta) &= d \int_0^1 (i_t)^*\big(\frac{\partial}{\partial t} \llcorner \eta\big)\, dt = \int_0^1 (i_t)^* d\big(\frac{\partial}{\partial t} \llcorner \eta\big)\, dt \\ I(d\eta) &= \int_0^1 (i_t)^*\big(\frac{\partial}{\partial t} \llcorner d\eta\big)\, dt. \end{aligned}$$

Aus (6.9) folgt

$$d(I\eta) + I(d\eta) = \int_0^1 (i_t)^* \frac{\partial \eta}{\partial t}\, dt.$$

Berechne nun für $t_0 \in [0,1]$ und $v_1, \ldots, v_k \in T_pM$

$$(i_{t_0})^* \frac{\partial \eta}{\partial t}(p)(v_1, \ldots, v_k) = \frac{\partial \eta}{\partial t}(t_0, p)(v_1, \ldots, v_k) = \frac{\partial}{\partial t} \eta(t,p)(v_1, \ldots, v_k)|_{t=t_0}.$$

Somit folgt

$$d(I\eta) + I(d\eta) = (i_1)^*\eta - (i_0)^*\eta. \tag{6.11}$$

Theorem 6.2 (Lemma von Poincaré) [§] *Gegeben sei eine Homotopie $f \in C^\infty([0,1] \times M, N)$. Dann gilt für jede k-Form $\omega \in C^\infty(\Lambda^k TN)$ die Homotopieformel*

$$(f_1)^*\omega - (f_0)^*\omega = d(If^*\omega) + I(f^*d\omega) \qquad \textit{wobei } f_t(p) = f(t,p). \tag{6.12}$$

[§] Henri Poincaré 1895, Vito Volterra 1889

Beweis: Wende (6.11) an mit $\eta = f^*\omega$, also $(i_t)^*\eta = (f \circ i_t)^*\omega = (f_t)^*\omega$. Mit $d\eta = df^*\omega = f^*d\omega$ folgt die Behauptung. $\square$

Um Formel (6.12) anschaulich zu verstehen, ist es nützlich den Satz von Stokes zur Verfügung zu haben, wir kommen dann darauf zurück.

Corollary 6.1 (Homotopieinvarianz) *Sind $f_0, f_1 \in C^\infty(M,N)$ glatt homotop, so sind die induzierten Abbildungen $f_0^*, f_1^* : H^k_{DR}(N) \to H^k_{DR}(M)$ gleich.*

Beweis: Sei $\omega \in Z^k(M)$, also $d\omega = 0$. Dann folgt $[f_1^*\omega] = [f_0^*\omega + dIf^*\omega] = [f_0^*\omega]$. $\square$

Corollary 6.2 *Ist $U \subset \mathbb{R}^n$ offen und sternförmig, so gilt $H^k(U) = \{0\}$ für $k \geq 1$.*

Beweis: Sei U sternförmig bezüglich des Punkts $x_0 \in U$. Dann ist $f : [0,1] \times U \to U$, $f(t,x) = tx + (1-t)x_0$, eine Homotopie zwischen $f_0 \equiv x_0$ und $f_1 = \mathrm{id}_U$. Es folgt

$$H^k(U) = \text{Bild}\, f_1^* = \text{Bild}\, f_0^* = \begin{cases} \mathbb{R} & \text{für } k = 0 \\ \{0\} & \text{für } k \geq 1. \end{cases}$$

Wir wollen die Stammform hier explizit bestimmen, und zwar ist

$$\begin{aligned} If^*\omega(x)(v_1,\ldots,v_{k-1}) &= \int_0^1 (f^*\omega)(t,x)\big(\frac{\partial}{\partial t}, v_1,\ldots,v_{k-1}\big)\,dt \\ &= \int_0^1 \omega(tx+(1-t)x_0)(x-x_0, tv_1,\ldots,tv_{k-1})\,dt, \end{aligned}$$

das heißt die Stammform von ω ist

$$I(f^*\omega) = \int_0^1 t^{k-1}(x-x_0)\llcorner\omega(tx+(1-t)x_0)\,dt. \tag{6.13}$$

$\square$

Für die Definition des Integrals einer n-Form auf einer n-dimensionalen Mannigfaltigkeit brauchen wir den Begriff der Orientierung. Zunächst erinnern wir an dieses Konzept in der Linearen Algebra. Seien $\mathcal{A} = \{a_1,\ldots,a_n\}$, $\mathcal{B} = \{b_1,\ldots,b_n\}$ zwei Basen des $\mathbb{R}$-Vektorraums V. Man hat dann die Transformationsmatrix $T = \mathrm{id}_{\mathcal{BA}}$, also $a_j = \sum_{i=1}^n T_{ij} b_i$ für $j = 1,\ldots,n$. Die Basen heißen gleich orientiert, Symbol $\mathcal{A} \sim \mathcal{B}$, wenn $\det T > 0$. Bekanntlich gilt

$$\mathrm{id}_{\mathcal{AA}} = \mathrm{E}_n, \quad \mathrm{id}_{\mathcal{AB}} = \mathrm{id}_{\mathcal{BA}}^{-1} \quad \text{und} \quad \mathrm{id}_{\mathcal{CA}} = \mathrm{id}_{\mathcal{CB}}\,\mathrm{id}_{\mathcal{BA}}.$$

Somit ist $\mathcal{A} \sim \mathcal{B}$ eine Äquivalenzrelation mit genau zwei Äquivalenzklassen. Ist σ eine dieser Klassen, so heißt σ Orientierung von V und (V,σ) ist ein orientierter $\mathbb{R}$-Vektorraum. Eine Basis $\mathcal{A}$ heißt dann positiv (bzw. negativ) orientiert, wenn $\mathcal{A} \in \sigma$ (bzw. $\mathcal{A} \notin \sigma$). Ein Isomorphismus zwischen orientierten Vektorräumen, also $A : (V,\sigma) \to (W,\tau)$, heißt orientierungstreu, falls gilt:

$$[v_1,\ldots,v_n] = \sigma \quad \Rightarrow \quad [Av_1,\ldots,Av_n] = \tau.$$

Insbesondere ist ein Automorphismus $A : (V,\sigma) \to (V,\sigma)$ genau dann orientierungstreu, wenn $\det A > 0$, und A heißt dann auch orientierungserhaltend.

Definition 6.5 (Orientierung einer Mannigfaltigkeit) *Sei M eine n-dimensionale differenzierbare Mannigfaltigkeit. Eine Orientierung von M ist eine Wahl $\sigma(p)$ der Orientierungen von T_pM, für alle $p \in M$, so dass gilt: zu $p \in M$ gibt es eine Karte (U, φ) mit*

$$\sigma|_U = \Big[\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\Big].$$

M heißt orientierbar, wenn eine Orientierung von M existiert.

Die Bedingung garantiert, dass die Orientierung nicht lokal beliebig springen kann.

Lemma 6.3 *M ist genau dann orientierbar, wenn es einen Atlas von M gibt, so dass alle Kartenwechsel positive Jacobideterminante haben.*

Beweis: Seien (U, φ) und (V, ψ) Karten auf M mit Basisfeldern $\mathcal{A} = \frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}$ und $\mathcal{B} = \frac{\partial}{\partial y^1}, \ldots, \frac{\partial}{\partial y^n}$. Nach (3.6) gilt

$$\frac{\partial}{\partial x^j}(p) = \sum_{i=1}^{n} \frac{\partial(\psi^i \circ \varphi^{-1})}{\partial x^j}(\varphi(p))\, \frac{\partial}{\partial y^i}(p).$$

Ist σ eine Orientierung von M, so bilden die Karten aus Definition 6.5 einen Atlas mit der gewünschten Eigenschaft: die Orientierung der Basisfelder ist durch σ gegeben und damit gleich auf den overlaps, also haben die Kartenwechsel positive Jacobideterminante. Ist umgekehrt ein solcher Atlas $\mathcal{A}$ gegeben, so setzen wir für jede Karte $(U, \varphi) \in \mathcal{A}$

$$\sigma|_U := \Big[\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\Big].$$

Auf Overlaps stimmen die Definitionen überein, also ist σ eine Orientierung von M. $\square$

Lemma 6.4 *Sei M zusammenhängend und orientierbar. Dann gibt es genau zwei Orientierungen von M.*

Beweis: Seien σ, τ Orientierungen von M. Wir zeigen, dass $G = \{p \in M : \sigma(p) = \tau(p)\}$ offen ist. Zu $p \in G$ wähle Karten $\varphi : U \to \varphi(U)$ und $\psi : V \to \psi(V)$ mit $p \in U \cap V$, und

$$\sigma|_U = \Big[\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\Big] \quad \text{sowie } \tau|_V = \Big[\frac{\partial}{\partial y^1}, \ldots, \frac{\partial}{\partial y^n}\Big].$$

Da $\sigma(p) = \tau(p)$ ist $\det D(\psi \circ \varphi^{-1})(\varphi(p)) > 0$. Aber dann ist $\det D(\psi \circ \varphi^{-1}) \circ \varphi > 0$, und damit $\sigma = \tau$, auf einer Umgebung von p. Analog sieht man, dass auch $M \backslash A = \{p \in M : \sigma(p) \neq \tau(p)\}$ offen ist. Somit ist entweder $A = \emptyset$ oder $A = M$, und es gibt höchstens zwei Orientierungen. Ist nun σ eine Orientierung von M, so können wir in jedem $p \in M$ die entgegengesetzte Orientierung $\tau(p)$ wählen. Ist $\mathcal{A}$ ein Atlas wie in Definition 6.5, so bekommen wir einen entsprechenden Atlas für τ, indem wir jede Karte (U, φ) durch $(U, S \circ \varphi)$ ersetzen, wobei $S \in \mathrm{GL}(\mathbb{R}^n)$ mit $\det S < 0$. $\square$

Sei jetzt $f : M \to N$ ein lokaler Diffeomorphismus, und τ sei eine Orientierung von N. Wir definieren dann die induzierte Orientierung $f^*\tau$ auf M durch

$$f^*\tau(p) = \big[Df(p)^{-1}w_1, \ldots, Df(p)^{-1}w_n\big] \quad \text{wobei } \tau(f(p)) = [w_1, \ldots, w_n].$$

Für zwei Basen $\mathcal{B}_{1,2}$ von $T_{f(p)}N$ und $\mathcal{A}_{1,2} = Df(p)^{-1}\mathcal{B}_{1,2}$ gilt $\mathrm{Id}_{\mathcal{A}_1\mathcal{A}_2} = \mathrm{Id}_{\mathcal{B}_1\mathcal{B}_2}$; also ist $f^*\tau(p)$ wohldefiniert. Wir zeigen jetzt, dass es zu $p \in M$ eine Karte wie in Definition 6.5 gibt. Wähl eine Karte (V, ψ) auf N bei $f(p)$ mit Basisfeldern $\frac{\partial}{\partial y^i}$ in der Orientierung τ. Es gibt eine Umgebung U von p, so dass $f|_U$ diffeomorph auf $f(U) \subset V$ abbildet. Dann ist $(U, \psi \circ f|_U)$ eine Karte von M, und für die zugehörigen Basisfelder $\frac{\partial}{\partial x^i}$ gilt für alle $q \in U$

$$\frac{\partial}{\partial x^i}(q) = D(\psi \circ f|_U)(q)^{-1}e_i = Df(q)^{-1}D\psi(f(q))^{-1}e_i = Df(q)^{-1}\frac{\partial}{\partial y^i}(f(q)).$$

Nach Definition ist also $f^*\tau = \left[\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right]$ auf U.

Definition 6.6 *Ein lokaler Diffeomorphismus $f : (M, \sigma) \to (N, \tau)$ heißt orientierungstreu (bzw. orientierungsumkehrend), falls $f^*\tau = \sigma$ (bzw. $f^*\tau \neq \sigma$).*

Example 6.5 $\mathbb{S}^n = \{p \in \mathbb{R}^{n+1} : |p| = 1\} \subset \mathbb{R}^{n+1}$ ist orientierbar, und zwar sei

$$\sigma(p) = [v_1, \ldots, v_n] \quad \text{wobei } v_i \in T_pM \text{ mit } \det(p, v_1, \ldots, v_n) > 0.$$

Offenbar ist $\sigma(p)$ wohldefiniert. Sei (U, φ) eine Karte bei p mit U zusammenhängend. Dann ist die Funktion $\det\big(p, \frac{\partial}{\partial x^1}(p), \ldots, \frac{\partial}{\partial x^n}(p)\big)$ entweder positiv auf U oder negativ auf U. Im zweiten Fall schalten wir die Spiegelung an der Hyperebene $x^n = 0$ dahinter; damit haben wir in beiden Fällen eine Karte wie in Definiton 6.5 verlangt. Wir betrachten nun die Antipodenabbildung

$$f : \mathbb{S}^n \to \mathbb{S}^n,\ f(p) = -p.$$

Seien $v_i \in T_{-p}\mathbb{S}^n$ mit $\sigma(-p) = [v_1, \ldots, v_n]$, also $\det(-p, v_1, \ldots, v_n) > 0$. Dann gilt $Df(p)^{-1}v_i = -v_i$ und $\det(p, -v_1, \ldots, -v_n) = (-1)^{n+1}\det(-p, v_1, \ldots, v_n)$. Also ist $f^*\sigma = \sigma$ genau wenn n ungerade ist.

Der folgende Satz gibt eine alternative Definition der Orientierbarkeit.

Theorem 6.3 *Für eine n-dimensionale differenzierbare Mannigfaltigkeit M sind äquivalent:*

(i) *M ist orientierbar*

(ii) *Es gibt eine Form $\omega \in C^\infty(\Lambda^n TM)$ mit $\omega(p) \neq 0$ für alle $p \in M$.*

Beweis: (ii) $\Rightarrow$ (i): Definiere für jedes $q \in M$ auf T_qM die Orientierung

$$\sigma(q) = [v_1, \ldots, v_n] \quad \text{wobei } v_i \in T_qM \text{ mit } \omega(q)(v_1, \ldots, v_n) > 0.$$

Sei (U, φ) eine Karte bei $p \in M$ mit U zusammenhängend. Dann gilt $\omega|_U = f\, dx^1 \wedge \ldots \wedge dx^n$ mit $f \neq 0$. Wir können $f > 0$ auf U annehmen, sonst gehe über zu $S \circ \varphi$, wobei S die Spiegelung an $\{x^n = 0\}$ ist. Es folgt $\omega(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}) = f > 0$, also $\sigma|_U = [\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}]$ auf U.

(i) $\Rightarrow$ (ii): Wähle einen orientierten Atlas $(U_\lambda, \varphi_\lambda)$, $\lambda \in \Lambda$. Es gibt dann eine untergeordnete, lokal endliche Teilung der Eins $\eta_\lambda \in C^\infty(M)$, das heißt $\eta_\lambda \geq 0$ und

(a) $\operatorname{spt} \eta_\lambda \subset U_\lambda$ für alle $\lambda \in \Lambda$,

(b) jedes $p \in M$ hat eine Umgebung U mit $\{\lambda \in \Lambda : \operatorname{spt} \eta_\lambda \cap U \neq \emptyset\}$ endlich.

(c) $\sum_{\lambda\in\Lambda} \eta_\lambda = 1$ auf ganz M.

Die Existenz einer solchen Teilung der Eins wird in Satz 18.1 gezeigt. Definiere nun

$$\omega = \sum_{\lambda\in\Lambda} \eta_\lambda\, \omega_\lambda \qquad \text{wobei } \omega_\lambda = d\varphi^1_\lambda \wedge \ldots \wedge d\varphi^n_\lambda.$$

Für jedes $\lambda \in \Lambda$ ist M Vereinigung der offenen Mengen U_λ und $M \backslash \operatorname{spt} \eta_\lambda$, also ist $\eta_\lambda\, \omega_\lambda$ glatt. Wegen (b) ist dann $\omega \in C^\infty(\Lambda^n TM)$, und es gilt für $p \in U_\lambda$

$$\omega(p)\Big(\frac{\partial}{\partial x^1}(p), \ldots, \frac{\partial}{\partial x^n}(p)\Big) = \sum_{\mu:\, p\in U_\mu} \underbrace{\eta_\mu(p)}_{\geq 0}\, \underbrace{\det D(\varphi_\mu \circ \varphi_\lambda^{-1})(\varphi_\lambda(p))}_{>0} > 0.$$

□

7 Integration von n-Formen

Jetzt kommen wir endlich zur Integration von n-Formen. Der folgende Satz spielt dabei eine wesentliche Rolle.

Transformationssatz Sei $\phi \in C^1(U,V)$ ein Diffeomorphismus der offenen Mengen $U, V \subset \mathbb{R}^n$. Ist $f : V \to [-\infty, \infty]$ messbar bezüglich $\mathcal{L}^n$, so gilt

$$\int_V f(y)\,dy = \int_U f(\phi(x))|\det D\phi(x)|\,dx,$$

wenn eines der Integrale existiert.

Die Symbole dx und dy in diesem Satz sind keine Differentialformen. sie stehen für Integration bezüglich des Lebesguemaßes, sind also nur Abkürzungen für $d\mathcal{L}^n(x)$ bzw. $d\mathcal{L}^n(y)$. Auf der Mannigfaltigkeit M haben wir kein kanonisches Maß gegeben, deshalb führen wir den Begriff der Messbarkeit mittels Karten auf $\mathbb{R}^n$ zurück.

Definition 7.1 *Sei M eine n-dimensonale differenzierbare Mannigfaltigkeit. $E \subset M$ heißt messbar (bzw. Nullmenge), wenn $\varphi(E \cap U) \subset \mathbb{R}^n$ messbar (bzw. Nullmenge) ist bezüglich $\mathcal{L}^n$, für alle Karten $\varphi : U \to \varphi(U)$.*

Es genügt hier, die Eigenschaft für einen Atlas von M nachzuweisen, für jede andere Karte folgt sie dann aus dem Transformationssatz, angewandt auf die Kartenwechsel. Ähnlich gehen wir vor, um die Messbarkeit von n-Formen zu definieren. In lokalen Koordinaten (U, φ) hat eine n-Form $\omega : M \to \Lambda^k(TM)$ die Darstellung

$$\omega|_U = \omega^\varphi\, dx^1 \wedge \ldots \wedge dx^n.$$

Wir nennen ω messbar, wenn die Funktion $\omega^\varphi \circ \varphi^{-1} : \varphi(U) \to \mathbb{R}$ Lebesguemessbar ist, für jede Karte (U, φ). Ist (V, ψ) eine andere Karte so gilt auf dem Overlap $U \cap V$

$$\omega^\varphi = \omega^\psi \det D(\psi \circ \varphi^{-1}) \circ \varphi.$$

Insbesondere reicht es auch hier aus, wenn die Messbarkeit für einen festen Atlas erfüllt ist.

Sei nun M eine orientierte Mannigfaltigkeit, und $E \subset M$ sei messbar und im Gebiet einer orientierten Karte (U, φ) enthalten. Eine messbare n-Form auf M nennen wir auf E integrierbar, falls

$$\|\omega\|_{L^1(E)} := \int_{\varphi(E)} |\omega^\varphi \circ \varphi^{-1}(x)|\,dx < \infty.$$

Ist das der Fall, so definieren wir das Integral

$$\int_E \omega := \int_{\varphi(E)} \omega^\varphi \circ \varphi^{-1}(x)\,dx \in \mathbb{R}.$$

Diese Definitionen hängen nicht von der Wahl der Karte ab: ist (V, ψ) eine andere orientierte Karte mit $E \subset V$, so gilt auf $U \cap V \supset E$

$$\omega^\psi = \omega^\varphi \det D(\psi \circ \varphi^{-1}) \circ \varphi.$$

Mit der Substitution $y = \psi \circ \varphi^{-1}(x)$ folgt aus dem Transformationssatz zunächst

$$\begin{aligned}\int_{\psi(E)} |\omega^{\psi} \circ \psi^{-1}(y)|\, dy &= \int_{\varphi(E)} |\omega^{\psi} \circ \varphi_i^{-1}(x)|\, |\det D(\psi \circ \varphi_i^{-1})(x)|\, dx \\ &= \int_{\varphi(E)} |\omega^{\varphi} \circ \varphi_i^{-1}(x)|\, dx < \infty.\end{aligned}$$

Da die Karten gleich orientiert sind, also $\det D(\psi \circ \varphi_i^{-1}) > 0$, folgt weiter

$$\begin{aligned}\int_{\psi(E)} \omega^{\psi} \circ \psi^{-1}(y)\, dy &= \int_{\varphi(E)} \omega^{\psi} \circ \varphi_i^{-1}(x)\, |\det D(\psi \circ \varphi_i^{-1})(x)|\, dx \\ &= \int_{\varphi(E)} \omega^{\varphi} \circ \varphi_i^{-1}(x)\, dx.\end{aligned}$$

Folgender Hinweis ist hier angebracht: die Integrierbarkeit und das Integral einer Funktion kann nicht auf diese Weise mittels Karten erklärt werden. Zum Beispiel gilt für die konstante Funktion $f(x) = 1$ auf $\mathbb{R}$ und die Karte $\varphi(x) = \arctan x$

$$\int_{\varphi(\mathbb{R})} f \circ \varphi^{-1}(y)\, dy = \pi \neq \infty = \int_{\mathbb{R}} f(x)\, dx.$$

Es ist wesentlich, dass n-Formen ein anderes Transformationsgesetz haben als Funktionen.

Um nun das Integral einer messbaren n-Form ω auf ganz M zu definieren, wählen wir eine höchstens abzählbare messbare Zerlegung $M = \bigcup_{i=1}^{\infty} E_i$, so dass $E_i \subset U_i$ für eine orientierte Karte (U_i, φ_i). Das ist immer möglich: nach Satz 1.1 ist M σ-kompakt, also existiert ein orientierter, abzählbarer Atlas (U_i, φ_i). Wir können dann $E_i = U_i \backslash (U_1 \cup \ldots \cup U_{i-1})$ wählen. Wir bezeichnen mit $L^1(\Lambda^n TM)$ den Raum der messbaren n-Formen mit

$$\|\omega\|_{L^1(M)} = \sum_{i \in I} \|\omega\|_{L^1(E_i)} < \infty. \tag{7.1}$$

Das Integral von $\omega \in L^1(\Lambda^n TM)$ definieren wir durch

$$\int_M \omega = \sum_{i \in I} \int_{E_i} \omega. \tag{7.2}$$

Diese Definitionen hängen wieder nicht von der Wahl der Zerlegung und der Karten ab: sei $M = \bigcup_{j \in J} F_j$ eine andere abzählbare, messbare Zerlegung mit $F_j \subset V_j$ für orientierte Karten $\psi_j : V_j \to \psi_j(V_j)$. Mit dem Satz über monotone Konvergenz folgt

$$\begin{aligned}\sum_{j \in J} \int_{\psi_j(F_j)} |\omega^{\psi_j} \circ \psi_j^{-1}(y)|\, dy &= \sum_{j \in J} \sum_{i \in I} \int_{\psi_j(E_i \cap F_j)} |\omega^{\psi_j} \circ \psi_j^{-1}(y)|\, dy \\ &= \sum_{i \in I} \sum_{j \in J} \int_{\varphi_i(E_i \cap F_j)} |\omega^{\varphi_i} \circ \varphi_i^{-1}(x)|\, dx \\ &= \sum_{i \in I} \int_{\varphi_i(E_i)} |\omega^{\varphi_i} \circ \varphi_i^{-1}(x)|\, dx < \infty.\end{aligned}$$

Somit ist der Raum $L^1(\Lambda^n TM)$ und die L^1-Norm unabhängig von der Wahl der Zerlegung. Für $\omega \in L^1(\Lambda^n TM)$ folgt analog, aber jetzt mit dem Konvergenzsatz von Lebesgue,

$$\sum_{j \in J} \int_{\psi_j(F_j)} \omega^{\psi_j} \circ \psi_j^{-1}(y)\, dy = \sum_{i \in I} \int_{\varphi_i(E_i)} \omega^{\varphi_i} \circ \varphi_i^{-1}(x)\, dx.$$

Also ist auch das Integral von ω unabhängig von der Zerlegung definiert. Hier noch drei Bemerkungen zur Definition des Integrals:

(a) Für eine Teilung der Eins η_i, $i \in I$, gilt

$$\int_M \omega = \sum_{i \in I} \int_M \eta_i \omega \qquad \text{für } \omega \in L^1(\Lambda^n TM).$$

(b) Es gilt $C^0_c(\Lambda^n TM) \subset L^1(\Lambda^n TM)$.

(c) Seien M, N orientierte, n-dimensionale Mannigfaltigkeiten. Ist $\phi \in C^1(M,N)$ orientierungstreuer (bzw. orientierungsumkehrender) Diffeomorphismus und $\omega \in L^1(\Lambda^n TN)$, so ist $\phi^*\omega \in L^1(\Lambda^n TM)$ und es gilt

$$\int_M \phi^*\omega = \pm \int_N \omega.$$

Die letzte Aussage wollen wir noch zeigen. Seien (U,φ) und (V,ψ) orientierte Karten von M bzw. N, so dass $\phi(U) \subset V$. Auf U gilt dann

$$\phi^*\omega = \phi^*(\omega^\psi\, dy^1 \wedge \ldots \wedge dy^n) = \omega^\psi \circ \phi \, \det D(\psi \circ \phi \circ \varphi^{-1}) \circ \varphi \, dx^1 \wedge \ldots \wedge dx^n.$$

Für $E \subset U$ messbar folgt aus dem Transformationssatz mit der Substitution $y = \psi \circ \phi \circ \varphi^{-1}(x)$

$$\begin{aligned}
\|\omega\|_{L^1(\phi(E))} &= \int_{\psi \circ \phi(E)} |\omega^\psi \circ \psi^{-1}(y)|\, dy \\
&= \int_{\varphi(E)} |\omega^\psi \circ \phi \circ \varphi^{-1}(x)|\, |\det D(\psi \circ \phi \circ \varphi^{-1})(x)|\, dx \\
&= \|\phi^*\omega\|_{L^1(E)}.
\end{aligned}$$

Weiter berechnen wir, wieder mit dem Transformationssatz,

$$\int_{\phi(E)} \omega = \int_{\varphi(E)} \omega^\psi \circ \phi \circ \varphi^{-1}(x)\, |\det D(\psi \circ \phi \circ \varphi^{-1})(x)|\, dx = \pm \int_E \phi^*\omega.$$

Das Vorzeichen ist je nach Orientierung von ϕ. Sei nun F_j, $j \in J$, eine messbare Zerlegung von N mit $F_j \subset V_j$ für Karten (V_j, ψ_j). Dann ist $E_j = \phi^{-1}(F_j)$ eine messbare Zerlegung von M, und es gilt $E_j \subset U_j$ für die Karten $(U_j = \phi^{-1}(V_j), \varphi_j = \psi_j \circ \phi)$. Da $\phi(U_j) = V_j$ folgt Aussage (c) aus der Rechnung oben durch Summation über $j \in J$.

Um den Satz von Stokes zu formulieren, benötigen wir noch den Begriff der Mannigfaltigkeit mit Rand. Das Standardmodell ist der abgeschlossene Halbraum $H^n = \{x \in \mathbb{R}^n : x^1 \leq 0\}$, versehen mit der von $\mathbb{R}^n$ induzierten Topologie. Die offenen Mengen in H^n sind also genau die Durchschnitte $H^n \cap W$ mit W offen in $\mathbb{R}^n$.

Definition 7.2 (Mannigfaltigkeit mit Rand) *Ein topologischer Raum (Hausdorffsch mit abzählbarer Umgebungsbasis) heißt n-dimensonale topologische Mannigfaltigkeit mit Rand, falls es einen C^0-Atlas $\varphi_i : U_i \to \varphi_i(U_i)$, $i \in I$, gibt mit $\varphi_i(U_i)$ offen in H^n. M heißt differenzierbar, wenn der Atlas C^∞-Kartenwechsel hat.*

Hier ist zuerst zu erklären, wie die Differenzierbarkeit der Kartenwechsel aufzufassen ist. Für $V \subset H^n$ offen bezeichnen wir mit $C^k(V)$ die Menge aller Funktionen $f \in C^k(\operatorname{int} V)$, deren Ableitungen $D^\alpha f$, $0 \leq |\alpha| \leq k$, stetig auf $V = \operatorname{int} V \cup (\partial H \cap V)$ fortsetzbar sind. Da jeder Punkt in V Häufungspunkt von $\operatorname{int} V$ ist, sind die Fortsetzungen eindeutig bestimmt, wir bezeichnen sie wieder mit $D^\alpha f$. Es gilt mit $\overline{D} = (\partial_2, \ldots, \partial_n)$

$$f|_{\partial H \cap V} \in C^k(\partial H \cap V) \text{ mit } \overline{D}^\alpha(f|_{\partial H \cap V}) = \big(\overline{D}^\alpha f\big)|_{\partial H \cap V} \quad \text{für } 0 \leq |\alpha| \leq k. \tag{7.3}$$

Denn für $\varepsilon \searrow 0$ konvergieren die Funktionen $f_\varepsilon(y) = f(y - \varepsilon e_1)$ lokal auf $\partial H \cap V$ gegen $f(y)$, und alle Ableitungen $\overline{D}^\alpha f_\varepsilon$ konvergieren lokal gleichmäßig gegen $\big(\overline{D}^\alpha f\big)|_{\partial H \cap V}$.

Lemma 7.1 *Seien (U_i, φ_i), $i = 1, 2$, Karten bei $p \in M$. Dann gelten für den Kartenwechsel $\phi = \varphi_2 \circ \varphi_1^{-1} : V_1 \to V_2$, wobei $V_i = \varphi_i(U_1 \cap U_2)$, folgende Aussagen:*

(a) $\phi(\operatorname{int} H^n \cap V_1) = \operatorname{int} H^n \cap V_2$,

(b) $\phi(\partial H^n \cap V_1) = \partial H^n \cap V_2$,

(c) $D\phi(x) \in \mathrm{GL}_n(\mathbb{R})$ *für alle* $x \in V_1$.

Beweis: Sei $x \in \operatorname{int} H^n \cap V_1$. Angenommen es gibt $B = B_\delta(x) \subset \operatorname{int} H^n$ mit $\phi(B) \subset \partial H^n$. Mit $\psi = \phi^{-1}|_{\partial H^n \cap V_2}$ folgt

$$\mathrm{id}_B = \phi^{-1} \circ \phi|_B = \psi \circ \phi|_B.$$

Nach (7.3) ist ψ differenzierbar, also folgt mit der Kettenregel $\mathrm{E}_n = D\psi(\phi(x))D\phi(x)$. Die Ableitung von ψ hat aber höchstens Rang $n-1$, Widerspruch. Also existiert eine Folge $x_k \to x$ mit $\phi(x_k) \in \operatorname{int} H^n \cap V_2$, und nach Kettenregel gilt

$$\mathrm{E}_n = D\phi^{-1}(\phi(x_k))\, D\phi(x_k).$$

Alle beteiligten Funktionen sind stetig, also gilt auch $\mathrm{E}_n = D\phi^{-1}(\phi(x))D\phi(x)$, und $D\phi(x)$ ist invertierbar. Nach dem Satz über inverse Funktionen ist $\phi(B_\delta(x))$, für $\delta > 0$ klein, offen im $\mathbb{R}^n$. Also ist $\phi(x) \in \operatorname{int} H^n$ bzw. $\phi(\operatorname{int} H^n \cap V_1) \subset \operatorname{int} H^n \cap V_2$. Durch Anwendung auf ϕ^{-1} folgt Aussage (a), und (b) ergibt sich direkt. Für (c) verwenden wir, dass jedes $x \in V_1$ Grenzwert einer Folge $x_k \in V_1 \cap \operatorname{int} H^n$ ist, und argumentieren wie oben. □

Für Punkte $p \in M$ haben wir damit folgende Dichotomie:

- entweder gilt $\varphi(p) \in \operatorname{int} H^n$ für alle Karten bei p, diese Punkte nennen wir innere Punkte (Bezeichnung: $p \in \operatorname{int} M$),
- oder es gilt $\varphi(p) \in \partial H^n$ für alle Karten bei p, diese Punkte heißen Randpunkte (Bezeichnung: $p \in \partial M$).

Die Mengen $\operatorname{int} H^n$ und ∂H^n sind bezüglich der Topologie auf $\mathbb{R}^n$ wie üblich definiert. Für die Mengen $\operatorname{int} M$ und ∂M ist das anders, denn M ist a priori gar nicht Teilmenge irgendeines topologischen Raums.

Man sieht leicht, dass jede Teilmenge einer Mannigfaltigkeit mit der induzierten Topologie Hausdorffsch ist und eine abzählbare Basis hat, siehe Beispiel 1.3 und Satz 1.1. Indem wir das Gebiet der Karte gegebenenfalls verkleinern, haben wir zu jedem $p \in \operatorname{int} M$ eine Karte (U, φ) bei p mit $\varphi(U) \subset \operatorname{int} H^n$, das heißt $\varphi(U)$ ist offen im $\mathbb{R}^n$. Damit ist $\operatorname{int} M$ eine n-dimensionale differenzierbare Mannigfaltigkeit (ohne Rand). Auf ∂M haben wir den C^0-Atlas

$$\varphi|_{\partial M \cap U} : \partial M \cap U \to \partial H^n \cap \varphi(U).$$

Hier ist $\varphi(U)$ offen in H^n, das heißt es gibt $W \subset \mathbb{R}^n$ offen mit $\varphi(U) = H^n \cap W$, also ist $\partial H^n \cap \varphi(U) = \partial H^n \cap W$ offen in $\partial H^n \cong \mathbb{R}^{n-1}$. Nach (7.3) sind die Kartenwechsel differenzierbar, also ist ∂M eine differenzierbare Mannigfaltigkeit der Dimension $n-1$.

Die Definition des Tangentialbündels, siehe Satz 4.1, ist auch im Fall einer n-dimensionalen, differenzierbaren Mannigfaltigkeit M mit Rand anwendbar. Dabei ist wesentlich, dass nach Lemma 7.1 die Kartenwechsel invertierbare Ableitung haben, auch in Randpunkten. Ist (U, φ) eine Karte, so haben wir insbesondere wieder die Basisfelder

$$\frac{\partial}{\partial x^j}(p) = D\varphi(p)^{-1} e_j \quad \text{für } p \in U.$$

Sei nun $p \in \partial M$, und $v \in T_pM$ habe die Basisdarstellung

$$v = \sum_{j=1}^n v_\varphi^j \frac{\partial}{\partial x^j}(p).$$

Ist $v_\varphi^1 > 0$, so nennen wir v nach außen weisend. Dieser Begriff hängt nicht von der Karte ab. Setze dazu $x = \varphi(p)$ und betrachte eine andere Karte (V, ψ) bei p. Nach (3.6), mit $f = \mathrm{id}_M$ und $\phi = \psi \circ \varphi^{-1}$, gilt dann

$$\frac{\partial}{\partial x^j}(p) = \sum_{i=1}^n \frac{\partial \phi^i}{\partial x^j}(x) \frac{\partial}{\partial y^i}(p).$$

Wir haben $\phi^1 \leq 0$ auf $\varphi(U)$ und $\phi^1 = 0$ auf $\varphi(U) \cap \partial H$. Daraus folgt

$$\frac{\partial \phi^1}{\partial x^j}(x) = \begin{cases} 0 & \text{für } j = 2, \ldots, n \\ \geq 0 & \text{für } j = 1. \end{cases}$$

Für $j \geq 2$ beachte (7.3), für $j = 1$ wende den Mittelwertsatz an auf $\phi^1(x + te_1)$ mit $t < 0$:

$$0 \leq \frac{\phi^1(x) - \phi^1(x + te_1)}{0 - t} = \frac{\partial \phi^1}{\partial x^1}(x + \tau e_1) \overset{t \nearrow 0}{\longrightarrow} \frac{\partial \phi^1}{\partial x^1}(x).$$

Tatsächlich gilt für $j = 1$ die strikte Ungleichung, sonst wäre $D\phi(x)$ nicht surjektiv im Widerspruch zu Lemma 7.1(c). Für die Basisdarstellung $v = \sum_{i=1}^n v_\psi^i \frac{\partial}{\partial y^i}(p)$ folgt

$$v_\psi^1 = \sum_{j=1}^n v_\varphi^j \frac{\partial \phi^1}{\partial x^j}(x) = v_\varphi^1 \frac{\partial \phi^1}{\partial x^1}(x) > 0.$$

Sei nun σ eine Orientierung auf M. Wir definieren auf ∂M eine induzierte Orientierung $\sigma_{\partial M}$. Für $p \in \partial M$ sei $v \in T_pM$ nach außen weisend. Dann setzen wir

$$\sigma_{\partial M}(p) = [v_2, \ldots, v_n] \quad \text{falls} \quad \sigma(p) = [v, v_2, \ldots, v_n]. \tag{7.4}$$

Ist (U, φ) eine Randkarte von M, so sind die Vektoren $\frac{\partial}{\partial x^2}, \ldots, \frac{\partial}{\partial x^n}$ positiv orientiert, denn $\frac{\partial}{\partial x^1}$ weist nach außen. Also ist der induzierte Atlas von ∂M positiv orientiert.

Example 7.1 Sei $D = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 \leq 1\}$. Dann ist D eine orientierte 2-dimensionale Mannigfaltigkeit mit Rand ∂D. Für $(x,y) \in \partial D$ ist der Vektor $(-y,x)$ positiv orientiert, denn der Vektor (x,y) weist nach außen, und $(x,y),(-y,x)$ ist eine positiv orientierte Basis von $\mathbb{R}^2$.

Der folgende Satz wird üblicherweise als allgemeiner Satz von Stokes bezeichnet. Tatsächlich betrachtete Stokes ein Vektorfeld X auf $\mathbb{R}^3$, mit zugehörigem Rotationsfeld $\operatorname{rot} X$. Für eine parametrisierte Fläche M gilt dann (1854): das Kurvenintegral von X längs ∂M ist gleich dem Fluss des Felds $\operatorname{rot} X$ durch die Fläche M, also in klassischer Notation

$$\int_{\partial M} X \cdot \overrightarrow{ds} = \int_M \operatorname{rot} X \cdot \overrightarrow{dA}.$$

Die sogenannte Vektoranalysis mit den Operatoren div und rot hat übrigens Maxwell eingeführt (1873). Die erste höherdimensionale Version des Satzes von Stokes stammt von Volterra (1889). Der Kalkül der Differentialformen wurde von Poincaré und Élie Cartan um 1900 entwickelt. Dieser hat dann in seinen Vorlesungen 1936-37 die moderne Fassung des Satzes von Stokes formuliert und diese später auch publiziert.*

Theorem 7.1 (Stokes-Cartan) *Sei M eine n-dimensionale, orientierte und kompakte Mannigfaltigkeit mit Rand ∂M. Dann gilt für $\omega \in C^1(\Lambda^{n-1}TM)$*

$$\int_M d\omega = \int_{\partial M} i^*\omega. \tag{7.5}$$

Dabei ist $i : \partial M \to M$ die Inklusionsabbildung, und ∂M hat die induzierte Orientierung.

Beweis: Wir betrachten zunächst den Fall, dass ω eine Form mit kompaktem Träger auf H^n ist. Wir schreiben dann

$$\omega = \sum_{i=1}^n (-1)^{i-1} \omega_i \, dx^1 \wedge \ldots \wedge dx^{i-1} \wedge \widehat{dx^i} \wedge dx^{i+1} \wedge \ldots \wedge dx^n.$$

Es folgt $d\omega = \sum_{i=1}^n \frac{\partial \omega_i}{\partial x^i} \, dx^1 \wedge \ldots \wedge dx^n$, und damit

$$\begin{aligned}
\int_{H^n} \omega &= = \int_{H^n} \sum_{i=1}^n \frac{\partial \omega_i}{\partial x^i}(x)\, dx \\
&= \int_{\mathbb{R}^{n-1}} \int_{-\infty}^0 \frac{\partial \omega_1}{\partial x^1}(x^1, y)\, dx^1\, dy + \sum_{i=2}^n \int_{-\infty}^0 \underbrace{\int_{\mathbb{R}^{n-1}} \frac{\partial \omega_i}{\partial x^i}(x^1,y)\, dy}_{=0}\, dx^1 \\
&= \int_{\mathbb{R}^{n-1}} \omega_1(0,y)\, dy.
\end{aligned}$$

Für die Karte $\varphi : \partial H \to \mathbb{R}^{n-1}$, $\varphi(0,y) = y$, gilt $\frac{\partial}{\partial y^i}(0,y) = D\varphi^{-1}(y) e_i = e_{i+1}$, also ist φ positiv orientiert. Wir berechnen $(i^*\omega)(0,y) = \omega_1(0,y)\, dy^1 \wedge \ldots dy^{n-1}$, also folgt

$$\int_{\partial H^n} i^*\omega = \int_{\mathbb{R}^{n-1}} \omega_1(0,y)\, dy.$$

*Elie Cartan: Les systèmes différentielles extérieurs et leurs applications géometriques, Hermann Paris 1945.

Im Fall dass ω kompakten Träger auf $\mathbb{R}^n$ hat, verschwindet das Integral. Als nächstes betrachte den Fall dass ω kompakten Träger im Gebiet einer orientierten Karte $\varphi : U \to \varphi(U)$ mit $\varphi(U)$ offen in H^n hat. Dann gilt

$$\begin{aligned}
\int_M d\omega &= \int_{H^n} (\varphi^{-1})^* d\omega = \int_{H^n} d(\varphi^{-1})^* \omega = \int_{\partial H^n} i^*_{\partial H^n} (\varphi^{-1})^* \omega,\\
\int_{\partial M} i^*_{\partial M} \omega &= \int_{\partial H^n} (\varphi|^{-1}_{\partial M})^* i^*_{\partial M} \omega.
\end{aligned}$$

Da $\varphi^{-1} \circ i_{\partial H^n} = i_{\partial M} \circ \varphi|^{-1}_{\partial M}$, folgt die Behauptung. Wir kommen schließlich zum allgemeinen Fall. Da M kompakt ist, gibt es eine endliche Überdeckung U_λ, $\lambda \in \Lambda$, mit Koordinatenumgebungen. Sei η_j, $j \in J$, eine lokal endliche untergeordnete Teilung der Eins, das heißt zu $j \in J$ gibt es ein $\lambda \in \Lambda$ mit $\operatorname{spt} \eta_j \subset\subset U_\lambda$. Es folgt

$$\int_M d\omega = \int_M d\Big(\sum_{j \in J} \eta_j \omega\Big) = \sum_{j \in J} \int_M d(\eta_j \omega) = \sum_{j \in J} \int_{\partial M} i^*(\eta_j \omega) = \int_{\partial M} i^* \omega.$$

Der Satz ist bewiesen. $\square$

Wenn man an Ästhetik interessiert ist, kann man den pullback im Satz auch unterdrücken. Bei aller Begeisterung muss gesagt werden, dass die wesentliche Arbeit für diese Aussage im Kalkül der Differentialformen und in der Definition der Begriffe (Mannigfaltigkeit, Orientierung, Rand,...) steckt. Am Ende ist alles so arrangiert, dass sich der Satz auf den Hauptsatz der Differential- und Integralrechnung reduziert. Ich denke, dass die Aussage oft als black box benutzt wird.

8 Definition der Riemannschen Mannigfaltigkeit

Sei M eine n-dimensionale differenzierbare Mannigfaltigkeit. Für $p \in M$ bezeichnen wir mit $T^{2,0}_p M$ den Vektorraum der Bilinearformen auf $T_p M$, und betrachten das Bündel

$$T^{2,0} M = \bigcup_{p \in M} T^{2,0}_p M.$$

Allgemeiner bezeichnet $T^{r,0}_p M$ den Raum der r-linearen Formen auf TM, und $T^{r,1}_p M$ den Raum der r-linearen Abbildungen mit Werten in $T_p M$, zum Beispiel $T^{1,1}_p M = \operatorname{End}(TM)$. Aber hier geht es nur um Bilinearformen.

Die Bündelprojektion $\pi : T^{2,0} M \to M$ ist natürlich $\pi(b) = p$ für $b \in T^{2,0}_p M$. Bezüglich einer Karte $\varphi : U \to \varphi(U)$ mit Feldern $\frac{\partial}{\partial x^i}$ haben wir die Basisformen

$$dx^i \otimes dx^j \in T^{2,0} U, \ \ (dx^i \otimes dx^j)(v, w) = v^i w^j \quad \text{ wobei } 1 \leq i, j \leq n.$$

Ein Schnitt $b : M \to T^{2,0} M$, also $b(p) \in T^{2,0}_p M$, hat die lokale Darstellung

$$b = \sum_{i,j=1}^n b_{ij}\, dx^i \otimes dx^j, \qquad \text{mit } b_{ij} = b\Big(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\Big) : U \to \mathbb{R}.$$

Definition 8.1 (Riemannsche Metrik) *Ein Schnitt $g \in C^\infty(T^{2,0} M)$ heißt Riemannsche Metrik, wenn für jedes $p \in M$ die Bilinearform $g(p) : T_p M \times T_p M \to \mathbb{R}$ ein Skalarprodukt ist, also symmetrisch und positiv definit.*

Bezüglich einer Karte $\varphi : U \to \varphi(U)$ haben wir die Matrixfunktion

$$G : U \to \mathbb{R}^{n\times n},\ G(p) = \big(g_{ij}(p)\big)_{1\leq i,j\leq n}.$$

Der Schnitt $g \in C^\infty(T^{2,0}M)$ ist genau dann eine Riemannsche Metrik, wenn für jede Karte $\varphi : U \to \varphi(U)$ und jedes $p \in U$ die Matrix $G(p)$ symmetrisch und positiv definit ist, also

$$g_{ij}(p) = g_{ji}(p) \text{ für } 1 \leq i,j \leq n \quad \text{ und } \quad \langle G(p)v,v\rangle = \sum_{i,j=1}^n g_{ij}(p)v^i v^j > 0 \quad \text{ für } v \neq 0.$$

Ist $\psi : V \to \psi(V)$, $y = \psi(p)$, eine weitere Karte, so gilt auf dem Overlap $U \cap V$ nach (3.6)

$$\frac{\partial}{\partial x^i} = \sum_{k=1}^n \frac{\partial \psi^k}{\partial x^i}\frac{\partial}{\partial y^k} \quad \text{ mit } \frac{\partial \psi^k}{\partial x^i}(p) = \frac{\partial(\psi^k \circ \varphi^{-1})}{\partial x^i}|_{x=\varphi(p)}.$$

Daraus folgt für eine Riemannsche Metrik g (bzw. jede Bilinearform), vgl. (3.9),

$$g^\varphi_{ij} = g\big(\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j}\big) = \sum_{k,l=1}^n \frac{\partial \psi^k}{\partial x^i}\frac{\partial \psi^l}{\partial x^j} g^\psi_{kl},$$

oder in Matrixschreibweise

$$G^\varphi = DF^{\mathrm{T}} \circ \varphi\ G^\psi\ DF \circ \varphi \quad \text{ mit } F = \psi \circ \varphi^{-1}|_{\varphi(U\cap V)}.$$

Theorem 8.1 (Riemannsches Volumen) *Sei (M,g) Riemannsche Mannigfaltigkeit. Dann gibt es ein eindeutig bestimmtes Maß μ_g, das Volumenmaß von g, mit*

$$\mu_g(E) = \int_{\varphi(E)} \sqrt{\det G}\,(\varphi^{-1}(x))\,dx,$$

falls $\varphi : U \to \varphi(U)$ Karte und $E \subset U$ Lebesgue-messbar.

Beweis: Sei $\psi : V \to \psi(V)$ andere Karte mit $E \subset V$, und $F = \psi \circ \varphi^{-1} : \varphi(U\cap V) \to \psi(U \cap V)$. Die Transformationsformel der Metrik ergibt

$$(\det G^\varphi) \circ \varphi^{-1} = \det\big(DF^{\mathrm{T}} G^\psi \circ \varphi^{-1} \det DF\big) = (\det G^\psi) \circ \varphi^{-1} |\det DF|^2.$$

Mit Substitution $y = F(x) = \psi \circ \varphi^{-1}(x)$ folgt aus dem Transformationssatz

$$\begin{aligned}\int_{\varphi(E)} \sqrt{\det G}(\varphi^{-1}(x))\,dx &= \int_{\varphi(E)} \sqrt{\det G^\psi}(\varphi^{-1}(x))|\det DF(x)|\,dx \\ &= \int_{F(\varphi(E))=\psi(E)} \sqrt{\det G^\psi}(\psi^{-1}(y))\,dy.\end{aligned}$$

Wähle nun eine messbare Zerlegung $M = \bigcup_{i\in I} E_i$, wobei jedes E_i in einem Kartengebiet liegt, und setze für $E \subset M$ messbar

$$\mu_g(E) = \sum_{i\in I} \mu_g(E \cap E_i).$$

Es ist leicht zu sehen, dass μ_g ein wohldefiniertes Maß auf M ist. □

Weitere assoziierte Größen sind

- die Länge $\|v\| = \sqrt{g(p)(v,v)}$ eines Vektors $v \in T_pM$

- der Winkel $\angle(v,w) = \arccos g(p)(v,w)$ zwischen $v, w \in T_pM$ mit $\|v\| = \|w\| = 1$.

Lemma 8.1 *Sei (M,g) Riemannsche Mannigfaltigkeit.*

(a) *Sei $f \in C^1(M)$, und $\operatorname{grad}_g f(p)$ bezeichne den eindeutigen Vektor in T_pM mit*

$$g(p)(\operatorname{grad}_g f(p), v) = Df(p)v \quad \text{für alle } v \in T_pM.$$

Dann gilt bezüglich einer Karte $\varphi : U \to \varphi(U)$

$$\operatorname{grad}_g f = \sum_{i,j=1}^n g^{ij} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^j} \qquad \text{wobei } g^{ij} := (G^{-1})_{ij}.$$

Insbesondere ist $\operatorname{grad}_g f \in C^{k-1}(TM)$ falls $f \in C^k(M)$, $k \geq 1$.

(b) *Zu $X \in C^1(TM)$ gibt es eine eindeutig bestimmte Funktion $\operatorname{div}_g X \in C^0(M)$ mit*

$$\int_M g(\operatorname{grad}_g \eta, X)\, d\mu_g = -\int_M \eta \operatorname{div}_g X \, d\mu_g \quad \text{für alle } \eta \in C_c^\infty(M).$$

Bezüglich einer Karte $\varphi : U \to \varphi(U)$ gilt

$$(\operatorname{div}_g X) \circ \varphi^{-1} = \frac{1}{\sqrt{\det G}} \sum_{i=1}^n \frac{\partial}{\partial x^i} \big(\sqrt{\det G}\, X^i \circ \varphi^{-1}\big).$$

Insbesondere ist $\operatorname{div}_g X \in C^{k-1}(M)$ falls $X \in C^k(TM)$.

Beweis: Für (a) berechnen wir

$$g\Big(\sum_{i,j=1}^n g^{ij} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^j}\,, \frac{\partial}{\partial x^k}\Big) = \sum_{i=1}^n \frac{\partial f}{\partial x^i} \underbrace{\sum_{j=1}^n g^{ij} g_{jk}}_{=\delta^i_k} = \frac{\partial f}{\partial x^k} = df(\frac{\partial}{\partial x^k}).$$

Nach dem Fundamentallemma der Variationsrechung gilt die Implikation

$$\int_M \eta f \, d\mu_g = 0 \text{ für alle } \eta \in C_c^\infty(M) \quad \Rightarrow \quad f = 0.$$

Somit ist die gesuchte Funktion in (b) eindeutig bestimmt; wir definieren $\operatorname{div}_g X$ auf dem Gebiet einer Karte $\varphi : U \to \varphi(U)$ durch die Formel. Dann gilt für $\eta \in C_c^\infty(U)$

$$g(\operatorname{grad}_g \eta, X) = D\eta \cdot X = \sum_{i=1}^n \frac{\partial \eta}{\partial x^i} X^i.$$

Es folgt mit partieller Integration

$$\begin{aligned}
\int_M g(\operatorname{grad}_g \eta, X)\, d\mu_g &= \int_{\varphi(U)} \sum_{i=1}^n \frac{\partial (\eta \circ \varphi^{-1})}{\partial x^i}\, X^i \circ \varphi^{-1} \sqrt{\det G}\, dx \\
&= -\int_{\varphi(U)} \eta \circ \varphi^{-1} \sum_{i=1}^n \frac{\partial}{\partial x^i} \big(X^i \circ \varphi^{-1} \sqrt{\det G}\big)\, dx \\
&= -\int_U \eta \operatorname{div}_g X \, d\mu_g.
\end{aligned}$$

Wegen der Eindeutigkeit stimmen die Formeln für zwei Karten auf den Overlaps überein. Für $\eta \in C_c^\infty(M)$ überdecke den Träger mit Kartengebieten und verwende eine untergeordnete Teilung der Eins. $\square$

Lemma 8.2 *Sei (M,g) Riemannsche Mannigfaltigkeit mit Rand ∂M. Dann gibt es auf ∂M ein eindeutig bestimmtes Normalenfeld $\nu : \partial M \to TM$ mit folgenden Eigenschaften:*

(1) $\nu(p) \perp T_p(\partial M)$ *bezüglich* g,

(2) $\|\nu(p)\| = 1$,

(3) $\nu(p)$ *weist nach außen.*

Beweis: Sei $\varphi : U \to \varphi(U) \subset H^n$ Randkarte von M. Betrachte $N : \partial M \cap U \to TM$, $N = \sum_{j=1}^n g^{j1} \frac{\partial}{\partial x^j}$. Wir berechnen

$$\begin{aligned} g\big(\frac{\partial}{\partial x^i}, N\big) &= \sum_{j=1}^n g_{ij} g^{j1} = 0 \qquad \text{für } i = 2, \ldots, n, \\ g(N,N) &= \sum_{i,j=1}^n g_{ij} g^{i1} g^{j1} = g^{11}, \\ \langle D\varphi(p) N, e_1 \rangle &= g^{11} > 0. \end{aligned}$$

Somit hat das Vektorfeld $\nu : \partial M \cap U \to TM$, $\nu = N/\sqrt{g^{11}}$, alle gewünschten Eigenschaften. Wegen Eindeutigkeit stimmen diese Definitionen auf den Overlaps überein. $\square$

Theorem 8.2 (Riemannscher Satz von Gauß) *Sei (M,g) kompakte Riemannsche Mannighfaltigkeit mit Rand ∂M. Dann gilt für $X \in C^1(TM)$*

$$\int_M \mathrm{div}_g X \, d\mu_g = \int_{\partial M} g(X,\nu)\, d\sigma_g. \tag{8.6}$$

Dabei ist σ_g das Riemannsche Oberflächenmaß auf ∂M, und ν die Riemannsche äußere Normale.

Beweis: Mit dem Satz von Fubini und dem Hauptsatz der Differential- und Integralrechnung.

Fall 1: spt X liegt im Gebiet einer inneren Karte $\varphi : U \to \varphi(U)$

$$\int_M \mathrm{div}_g X \, d\mu_g = \int_{\mathbb{R}^n} \frac{1}{\sqrt{\det G}} \sum_{i=1}^n \frac{\partial}{\partial x^i} \big(\sqrt{\det G} X^i\big) \sqrt{\det G}\, dx = 0.$$

Fall 2: spt X liegt im Gebiet einer Randkarte $\varphi : U \to \varphi(U)$

$$\begin{aligned} \int_M \mathrm{div}_g X \, d\mu_g &= \int_{\partial H^n} \int_{-\infty}^0 \frac{\partial}{\partial x^1} \big(\sqrt{\det G}\big)(x^1, y)\, dx^1\, dy \\ &\quad + \sum_{i=2}^n \int_{-\infty}^0 \int_{\partial H^n} \frac{\partial}{\partial x^i} \big(\sqrt{\det G}\, X^i\big)(x^1, y)\, dy\, dx^1 \\ &= \int_{\partial H^n} \big(\sqrt{\det G}\, X^1\big)(0, y)\, dy. \end{aligned}$$

Für das Randintegral berechnen wir mit Lemma 8.2

$$g(X,\nu) = \sum_{i,j=1}^{n} g_{ij} X^i \frac{1}{\sqrt{g^{11}}} g^{1j} = \frac{1}{\sqrt{g^{11}}} X^1.$$

Das Oberflächenmaß ist $d\sigma_g = \sqrt{\det \widehat{G}_{11}}$, wobei sich $\widehat{G}_{11} \in \mathbb{R}^{(n-1)\times(n-1)}$ aus G durch Streichen der ersten Zeile und Spalte ergibt. Mit der Cramerschen Regel folgt

$$g(X,\nu)\, d\sigma_g = \frac{1}{\sqrt{g^{11}}} \sqrt{\det \widehat{G}_{11}}\, X^1\, dy = \sqrt{\det G}\, dy.$$

Fall 3: $X \in C^1(TM)$ beliebig
Überdeckung von M mit endlich vielen Kartengebiete und untergeordnete Teilung der Eins.

□

Ist M orientierbar, so können wir die Riemannsche Volumenform $\omega_g \in C^\infty(\Lambda^n TM)$ definieren, bezüglich einer orientierten Karte $\varphi : U \to \varphi(U)$ lautet sie

$$\omega_g = \sqrt{\det G^\varphi}\, dx^1 \wedge \ldots \wedge dx^n.$$

Ist $\psi : V \to \psi(V)$, $y = \psi(p)$, eine andere Karte, so gilt mit $F = \psi \circ \varphi^{-1}$

$$\begin{aligned} G^\varphi &= DF^{\mathrm{T}} \circ \varphi\, G^\psi\, DF \circ \varphi \\ dy^1 \wedge \ldots \wedge dy^n &= \det DF \circ \varphi\, dx^1 \wedge \ldots \wedge dx^n, \\ \sqrt{\det G^\varphi}\, dx^1 \wedge \ldots \wedge dx^n &= \operatorname{sign}(\det DF \circ \varphi) \sqrt{\det G^\psi}\, dy^1 \wedge \ldots \wedge dy^n. \end{aligned}$$

Indem wir einen orientierten Atlas verwenden, ist ω_g wohldefiniert, und es gilt

$$\int_M f\, \omega_g = \int_M f\, d\mu_g.$$

Betrachte nun für $X \in C^1(TM)$ die $(n-1)$-Form

$$X \llcorner \omega_g(v_1, \ldots, v_{n-1}) = \omega_g(X, v_1, \ldots, v_{n-1}).$$

In lokalen Koordinaten gilt

$$X \llcorner \omega_g = \sum_{i=1}^{n} \sqrt{\det G}\, X^i (-1)^{i-1} dx^1 \wedge \ldots \wedge \widehat{dx^i} \wedge \ldots \wedge dx^n.$$

Wir berechnen damit

$$d(X \llcorner \omega_g) = (\operatorname{div}_g X)\, \omega_g. \tag{8.7}$$

Weiter gilt bezüglich einer Randkarte $X^\top \llcorner \omega_g = 0$, also

$$i^*_{\partial M}(X \llcorner \omega_g) = g(X,\nu)\; i^*_{\partial M}(\nu \llcorner \omega_g).$$

$i^*_{\partial M}(\nu \llcorner \omega_g)$ ist aber die Oberflächenform auf ∂M, denn

$$\begin{aligned} i^*_{\partial M}(\nu \llcorner \omega_g) &= \sqrt{\det G}\, dx^1 \wedge \ldots \wedge dx^n \Big(\sum_{j=1}^{n} \frac{g^{1j}}{\sqrt{g^{11}}} \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^2}, \ldots, \frac{\partial}{\partial x^n} \Big) \\ &= \sqrt{\det G}\, \sqrt{g^{11}} = \sqrt{\det \widehat{G}_{11}}, \end{aligned}$$

wieder nach der Cramerschen Regel. Also haben wir gezeigt

$$i_{\partial M}^*(X \llcorner \omega_g) = g(X,\nu)\,\omega_g^{\partial M}. \tag{8.8}$$

Insgesamt gilt also

$$\begin{aligned}
\int_M d(X \llcorner \omega_g) &= \int_M \mathrm{div}_g X\,\omega_g = \int_M \mathrm{div}_g X\,d\mu_g,\\
\int_{\partial M} i_{\partial M}^*(X \llcorner \omega_g) &= \int_{\partial M} g(X,\nu)\,\omega_g^{\partial M} = \int_{\partial M} g(X,\nu)\,d\sigma_g.
\end{aligned}$$

Zur Relation der beiden Sätze: aus Gauß folgt Stokes. Denn eine Riemannsche Metrik gibt es immer, siehe unten, und eine gegebene $(n-1)$-Form ξ kann als $X \llcorner \omega_g$ geschrieben werden, wähle dazu in Koordinaten

$$X^i = \frac{1}{\sqrt{\det G}}\,\xi_i \qquad \text{wobei } \xi = \sum_{i=1}^n \xi_i\,dx^1 \wedge \ldots \wedge \widehat{dx^i} \wedge \ldots \wedge dx^n.$$

Umgekehrt gilt auch: aus Stokes folgt Gauss, allerdings mit der kleinen Einschränkung dass M orientierbar sein muss, was im Satz von Gauß a priori nicht der Fall ist.

Theorem 8.3 (Existenz Riemannscher Metriken) *Auf jeder n-dimensionalen differenzierbaren Mannigfaltigkeit gibt es eine (glatte) Riemannsche Metrik.*

Beweis: Sei $\varphi_\lambda : U_\lambda \to \varphi_\lambda(U_\lambda)$, $\lambda \in \Lambda$, ein lokal endlicher Atlas von M, vgl. Satz 1.1. Wähle auf U_λ die glatte Riemannsche Metrik

$$g_\lambda = \sum_{i=1}^n dx^i \otimes dx^i, \qquad \text{wobei } x^i = \varphi_\lambda^i.$$

Wähle eine Teilung der Eins $\eta_\lambda \in C^\infty(M)$, $0 \leq \eta_\lambda \leq 1$, mit $\operatorname{spt} \eta_\lambda \subset U_\lambda$, und setze

$$g = \sum_{\lambda \in \Lambda} \eta_\lambda\, g_\lambda.$$

Dann ist $g \in C^\infty(T^{2,0}M)$ symmetrisch, und es gilt für $X \in TM$, mit $p = \pi(X)$,

$$g(X,X) = \sum_{\lambda \in \Lambda : p \in U_\lambda} \underbrace{\eta_\lambda(p)}_{\geq 0}\, \underbrace{g_\lambda(p)(X,X)}_{>0} \geq 0.$$

Es ist $\eta_\lambda(p) > 0$ für wenigstens ein λ, daraus folgt die strikte Ungleichung. □

Definition 8.2 *Sei (N,h) Riemannsche Mannigfaltigkeit und $f : M \to N$ eine Immersion, also* $\operatorname{rang} Df = \dim M$ *und insbesondere* $\dim M \leq \dim N$.

(a) *Dann ist $g(X,Y) = h(Df \cdot X, Df \cdot Y)$ eine Riemannsche Metrik auf M, sie heißt Pullback von h unter f (Bezeichnung: $g = f^*h$).*

(b) *Ist g gegebene Riemannsche Metrilk auf M und ist $g = f^*h$, so heißt $f : (M,g) \to (N,h)$ isoimetrische Immersion.*

(c) *Ein Diffeomorphismus $f : (M,g) \to (N,h)$ mit $g = f^*h$ heißt Isometrie.*

Die Klassifikation Riemannscher Mannigfaltigkeiten ist, im Vergleich zur Klassifikation der differenzierbaren Mannigfaltigkeiten, ein viel umfangreicheres Problem, denn es gibt auf einer differenzierbaren Mannifaltigkeiten viele Metriken, die nicht isometrisch sind. Trotzdem können Riemannsche Metriken helfen, wenn es um die differenzierbare Klassifikation geht. Die Idee ist, dass gewisse Normalformen von Metriken konstruiert werden, etwa Metriken konstanter Krümmung. Das war jedenfalls für 2-dimensionale Mannigfaltigkeiten sehr erfolgreich, und ist grundsätzlich auch der Ansatz für 3-dimensionale Mannigfaltigkeiten.

Wir kommen jetzt zu einem weiteren Aspekt, nämlich der Struktur einer Riemannschen Mannigfaltigkeit als metrischer Raum. Sei $c : [a,b] \to M$ eine stückweise C^1-Kurve, das heißt $c \in C^0([a,b], M)$ und es gibt eine Unterteilung $a = t_0 < \ldots < t_N \leq b$, so dass c auf jedem Teilintervall $[t_{i-1}, t_i]$, $1 \leq i \leq N$, von der Klasse C^1 ist. Insbesondere existieren in den t_i die links- und rechtsseitigen Ableitungen, sie müssen aber nicht übereinstimmen. Wir bezeichnen den Raum dieser Kurven mit $PC^1([a,b], M)$.

Definition 8.3 (Bogenlänge und Abstand) *Sei (M,g) eine Riemannsche Mannigfaltigkeit. Die Länge einer Kurve $c \in PC^1([a,b], M)$ ist*

$$L_g(c) = \int_a^b \|c'(t)\|_g \, dt. \tag{8.9}$$

Der Riemannsche Abstand von zwei Punkten $p, q \in M$ ist

$$d(p,q) = \inf \big\{ L_g(c) : c \in PC^1([a,b], M),\ c(a) = p,\ c(b) = q \big\}. \tag{8.10}$$

Ist $c(t)$ im Gebiet einer Karte $\varphi : U \to \varphi(U) \subset \mathbb{R}^n$, so lautet das Bogenlängenintegral

$$L_g(c) = \int_a^b \Big(\sum_{i,j=1}^n g_{ij}(c(t))\, (c^i)'(t) (c^j)'(t) \Big)^{\frac{1}{2}} dt.$$

Example 8.1 Im $\mathbb{R}^n$ gilt $d(p,q) = |p-q|$. Im Fall $p = q$ ist das klar, für $p \neq q$ betrachte dazu

$$e = \frac{q-p}{|q-p|}, \qquad \text{also } |e| = 1.$$

Sei $c \in PC^1([a,b], \mathbb{R}^n)$ mit $c(a) = p$, $c(b) = q$. Mit Cauchy-Schwarz und dem Hauptsatz folgt

$$L(c) = \int_a^b |c'(t)| \, dt \geq \int_a^b \langle c'(t), e \rangle \, dt = [\langle c(t), e \rangle]_{t=a}^{t=b} = |q-p|.$$

Andererseits können wir die Stecke $c(t) = (1-t)q + tp$, $t \in [0,1]$, als Verbindungskurve wählen, diese hat die Länge $|p-q|$.

Example 8.2 In $M = \mathbb{R}^n \backslash \{0\}$ gilt ebenfalls $d(p,q) = |p-q|$. Die untere Schranke ist klar, und für $p, q \in M$ können wir wieder die Strecke als Verbindungskurve wählen, es sei denn der Nullpunkt liegt auf der Strecke. In diesem Fall können wir aber den Nullpunkt umgehen, indem wir ein kleines Stück durch einen Halbkreis ersetzen. Wir erhalten so eine Folge von Kurven, deren Länge gegen $|p-q|$ konvergiert. Man kann sich überlegen, dass das Infimum $|p-q|$ in diesem Fall nicht angenommen wird.

Example 8.3 In $\mathbb{S}^n$, versehen mit der induzierten Euklidischen Metrik, gilt

$$d(p,q) = \arccos\langle p, q\rangle \in [0,\pi].$$

Ein Großkreis ist nach Definition ein Schnitt der Sphäre mit einer 2-Ebene durch den Nullpunkt. Im Fall $q \neq \pm p$ spannen p, q eine solche Ebene auf, und sie zerlegen den Großkreis in zwei Bögen. Der Abstand ist die Länge des kürzeren Bogens (so die Behauptung). Um das zu beweisen, sei $c : [a,b] \to \mathbb{S}^n$ von p nach q gegeben. Die tangentiale Komponente von p im Punkt $c(t)$ ist

$$p^\top = p - \langle p, c(t)\rangle c(t), \qquad \text{insbesondere } |p^\top| = \sqrt{1 - \langle p, c(t)\rangle^2}.$$

Es gilt wieder mit Cauchy-Schwarz und Hauptsatz

$$\begin{aligned} L(c) &= \int_a^b |c'(t)|\,dt \\ &\geq -\int_a^b \Big\langle \frac{p^\top}{|p^\top|}, c'(t) \Big\rangle\,dt \\ &= -\int_a^b \frac{\langle p, c'(t)\rangle}{\sqrt{1-\langle p, c(t)\rangle^2}}\,dt \qquad \text{(da } c'(t) \text{ tangential)} \\ &= \int_a^b \frac{d}{dt} \arccos\langle p, c(t)\rangle\,dt = \arccos\langle p, q\rangle. \end{aligned}$$

Die Abschätzung gilt so nur wenn $c(t) \neq \pm p$ für alle $t \in (a,b)$, denn für $c(t) = \pm p$ wird das Integral singulär. Das ist aber keine Einschränkung, wie man leicht sieht.

Theorem 8.4 (Riemannscher Abstand) *Sei (M,g) eine zusamenhängende Riemannsche Mannigfaltigkeit. Dann ist*

$$d(p,q) = \inf \big\{ L_g(c) : c \in PC^1([a,b], M),\ c(a) = p,\ c(b) = q \big\}$$

eine Metrik auf M, und die zugehörige Topologie ist gleich der gegebenen Topologie auf M.

Beweis: Es ist klar, dass $d(p,q)$ nichtnegativ und symmetrisch ist, und dass die Dreicksungleichung gilt. Es bleibt zu zeigen, dass $d(p,q) < \infty$ und $d(p,q) > 0$ für $p \neq q$.

Für das erste sei $p \in M$ fest, und $A \subset M$ sei die Menge aller $q \in M$, die mit p durch eine stückweise C^1-Kurve verbindbar sind. Zu $q \in M$ gibt es eine Karte $\varphi : U_q \to B_1(0)$ mit $\varphi(q) = 0$. Aus $q \in A$ folgt dann $U_q \subset A$, also ist A offen. Andererseits folgt aus $q \in M \backslash A$ auch $U_q \subset M \backslash A$, somit ist $M \backslash A$ auch offen. Da $p \in A$ folgt $A = M$, und hieraus $d(p,q) < \infty$.

Für die zweite Aussage wähle zu $p \in M$ eine Karte $\varphi : U \to \varphi(U)$ mit $\varphi(p) = 0$ und $\overline{B_1(0)} \subset \varphi(U)$. Die Funktion

$$f : \overline{B_1(0)} \times \mathbb{S}^{n-1} \to \mathbb{R},\ f(x,v) = \Big(\sum_{i,j=1}^n g_{ij}(x) v^i v^j \Big)^{\frac{1}{2}} =: \|v\|_{g(x)},$$

ist stetig und strikt positiv. Da $\overline{B_1(0)} \times \mathbb{S}^{n-1}$ kompakt ist, gibt es $\lambda, \Lambda \in (0,\infty)$ mit

$$\lambda |v| \leq \|v\|_{g(x)} \leq \Lambda |v| \qquad \text{für alle } (x,v) \in \overline{B_1(0)} \times \mathbb{R}^n.$$

Das folgt direkt für $|v| = 1$, und dann für $v \in \mathbb{R}^n$ beliebig durch Skalierung. Sei $c \in PC^1([a,b], M)$ Kurve von p nach q, und $c(t) \in \varphi^{-1}(\overline{B_1(0)})$ für $t \in [0,\tau]$. Dann gilt

$$L_g(c) \geq L_g(c|_{[0,\tau]}) = \int_0^\tau \|(\varphi \circ c)'(t)\|_{g(\varphi \circ c(t))}\, dt \geq \lambda \int_0^\tau |(\varphi \circ c)'(t)|\, dt \geq \lambda\, |\varphi(c(\tau))|.$$

Also gilt die Abschätzung

$$d(p,q) \geq \begin{cases} \lambda|\varphi(q)| & \text{falls } q \in \varphi^{-1}(B_1(0)), \\ \lambda & \text{sonst.} \end{cases}$$

Für alle $\varrho \in (0,\lambda]$ folgen die Implikationen

$$q \in \varphi^{-1}(B_{\frac{\varrho}{\Lambda}}(0)) \quad \Rightarrow \quad d(p,q) < \varrho \quad \Rightarrow \quad q \in \varphi^{-1}(B_{\frac{\varrho}{\lambda}}(0)).$$

Mengen dieses Typs erzeugen jeweils die Topologie, also stimmen die metrische und die gegebene Topologie überein. $\square$

In Satz 2.2 haben wir die Operation einer Gruppe $\Gamma \subset \mathrm{Diff}(M)$ auf einer Mannigfaltigkeit M betrachtet. Ist die Operation eigentlich diskontinuierlich, so ist M/Γ eine Mannigfaltigkeit und $\pi : M \to M/\Gamma$ eine Überlagerung. Jetzt ergänzen wir den Satz im Fall, dass die Diffeomorphismen Isometrien sind.

Theorem 8.5 (Riemannsche Überlagerung) *Sei (M,g) eine Riemannsche Mannigfaltigkeit, auf der die Gruppe $\Gamma \subset \mathrm{Isom}(g)$ eigentlich diskontinuierlich operiert. Dann gibt es auf M/Γ genau eine Riemannsche Metrik $\overline{g}$, so dass $\pi : (M,g) \to (M/\Gamma, \overline{g})$ isometrisch ist.*

Beweis: Die Projektion $\pi : (M,g) \to (M/\Gamma, \overline{g})$ ist genau dann isometrisch, wenn für alle $x \in M/\Gamma$ und alle $p \in M$ mit $\pi(p) = x$ gilt:

$$\overline{g}(x)(X_1,X_2) = g(p)(\tilde{X}_1,\tilde{X}_2) \quad \text{falls } D\pi(p)\tilde{X}_i = X_i \text{ für } i = 1,2. \tag{8.11}$$

Insbesondere ist $\overline{g}$ eindeutig bestimmt. Umgekehrt wähle zu $x \in M/\Gamma$ ein $p \in M$ mit $\pi(p) = x$, und definiere $\overline{g}(x)$ durch (8.11). Das Ergebnis hängt nicht von der Wahl von p ab: zu $q \in M$ mit $\pi(q) = x$ gibt es ein $\gamma \in \Gamma$ mit $\gamma(p) = q$, und es gilt

$$D\pi(q)D\gamma(p)\tilde{X}_i = D(\pi \circ \gamma)(p)\tilde{X}_i = D\pi(p)\tilde{X}_i = X_i.$$

Ist nun $D\pi(q)\tilde{Y}_i = X_i$, so folgt $\tilde{Y}_i = D\gamma(p)\tilde{X}_i$ und weiter $g(q)(\tilde{Y}_1,\tilde{Y}_2) = g(p)(\tilde{X}_1,\tilde{X}_2)$, da γ eine Isometrie ist. Es bleibt zu zeigen, dass $\overline{g}$ glatt ist. Nach Satz 2.2 gibt es zu $p \in M$ eine Karte $\varphi : U \to \varphi(U)$, so dass $\pi|_U : U \to \pi(U)$ diffeomorph ist. Es folgt

$$\pi^*\overline{g}\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = \overline{g} \circ \pi \left(D\pi \cdot \frac{\partial}{\partial x^i}, D\pi \cdot \frac{\partial}{\partial x^j}\right) = g_{ij} \in C^\infty(U).$$

$\square$

Zum Schluss des Kapitels wollen wir einen Spezialfall genauer studieren. Für eine Basis v_1, v_2 von $\mathbb{R}^2$ betrachten wir das Gitter $\Gamma = \mathbb{Z}v_1 \oplus \mathbb{Z}v_2$, das auf $\mathbb{R}^2$ durch Translationen operiert; wir haben dann die Riemannsche Überlagerung $\pi_\Gamma : \mathbb{R}^2 \to \mathbb{R}^2/\Gamma$. Ist $A \in \mathrm{GL}(\mathbb{R}^2)$, so ist $\Gamma' = A\Gamma$ wieder ein Gitter und A induziert eine Abbildung

$$\varphi_A : \mathbb{R}^2/\Gamma \to \mathbb{R}^2/\Gamma',\ \varphi_A([z]_\Gamma) = [Az]_{\Gamma'} \quad \text{wobei } \Gamma' = A\Gamma.$$

φ_A ist wohldefiniert, denn für $\gamma \in \Gamma$ gilt $[A(z+\gamma)]_{\Gamma'} = [Az + A\gamma]_{\Gamma'} = [Az]_{\Gamma'}$. Ist auch $B \in \mathrm{GL}(\mathbb{R}^2)$ und $\Gamma'' = B\Gamma'$, so folgt

$$\varphi_B \circ \varphi_A([z]_\Gamma) = \varphi_B([Az]_{\Gamma'}) = [BAz]_{\Gamma''} = \varphi_{BA}([z]).$$

Also ist φ_A bijektiv mit $\varphi_A^{-1} = \varphi_{A^{-1}}$. Nach Definition gilt $\varphi_A \circ \pi_\Gamma = \pi_{\Gamma'} \circ A$ und π_Γ ist lokal diffeomorph, also sind die φ_A glatt. Die $\mathbb{R}^2/\Gamma$ sind damit diffeomorph zu $\mathbb{R}^2/\mathbb{Z}^2$, man bezeichnet sie als flache Tori, weil ihre Metrik lokal Euklidisch ist. Es gilt $\pi_\Gamma^* \varphi_A^* g_{\Gamma'} = A^* \pi_{\Gamma'}^* g_{\Gamma'} = A^* g_{\mathbb{R}^2}$, beziehungweise bezüglich der Standardbasis

$$\varphi_A^* g_{\Gamma'}(D\pi_\Gamma \cdot e_i, D\pi_\Gamma \cdot e_j) = g_{\mathbb{R}^2}(Ae_i, Ae_j) \quad \text{für } 1 \leq i,j \leq 2.$$

Für $A = \lambda S$ mit $\lambda > 0$ und $S \in \mathbb{O}(2)$ folgt $\varphi_A^* g_{\Gamma'} = \lambda^2 g_\Gamma$.

Definition 8.4 (konforme Äquivalenz) *$\varphi : (M,g) \to (N,h)$ heißt konform, wenn es eine Funktion $\lambda : M \to (0,\infty)$ gibt mit $\varphi^* h = \lambda^2 g$. Ist φ zusätzlich diffeomorph, so heißt φ konforme Äquivalenz.*

Werden g, h mit positiven Funktionen multipliziert, so ändert sich der sogenannte Konformfaktor λ, aber die Klasse der konformen Abbildungen bzw. Äquivalenzen bleibt gleich. Die Verkettung konformer Abbildungen ist wieder konform. Ist φ konforme Äquivalenz so auch φ^{-1}, wir berechnen explizit

$$(\varphi^{-1})^* g = (\varphi^{-1})^* \Big(\frac{1}{\lambda^2} \varphi^* h \Big) = \frac{1}{(\lambda \circ \varphi^{-1})^2} (\varphi \circ \varphi^{-1})^* h = \frac{1}{(\lambda \circ \varphi^{-1})^2} h.$$

Für $A = \lambda S$ mit $\lambda > 0$ und $S \in \mathbb{O}(2)$ ist $\varphi_A : \mathbb{R}^2/\Gamma \to \mathbb{R}^2/\Gamma'$ mit $\Gamma' = A\Gamma$ konforme Äquivalenz. Allerdings ist φ_A speziell, zum Beispiel ist der Konformfaktor $\lambda > 0$ konstant. Im allgemeinen ist das nicht so, zum Beispiel gilt

$$\varphi^* g_{\mathbb{R}^n}(x) = \frac{1}{|x|^4} g_{\mathbb{R}^n} \quad \text{für } \varphi : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}^n \backslash \{0\},\ \varphi(x) = \frac{x}{|x|^2}.$$

Wir wollen untersuchen, welche der flachen Tori zueinander konform äquivalent sind. Das folgende Lemma besagt, dass eine konforme Äquivalenz solcher Tori speziell ist.

Lemma 8.3 *Sei $\varphi : \mathbb{R}^2/\Gamma \to \mathbb{R}^2/\Gamma'$ konforme Äquivalenz. Dann gibt es $S \in \mathbb{O}(2)$, $\lambda > 0$ und $b \in \mathbb{R}^2$, so dass gilt:*

$$\Gamma' = A\Gamma \text{ mit } A = \lambda S \quad \text{und} \quad \varphi([x]_\Gamma) = [Ax + b]_{\Gamma'} \text{ für alle } x \in \mathbb{R}^2.$$

Beweis: Wir benötigen den folgenden Liftungssatz aus der Überlagerungstheorie.

Theorem 8.6 (Liftung) *Sei $\pi : M \to Y$ eine Überlagerung und $f \in C^0(X,Y)$ mit $f(x_0) = y_0$. Ist X einfach zusammenhängend, so gibt es zu jedem $p_0 \in M$ mit $\pi(p_0) = y_0$ genau ein $F \in C^0(X,M)$ mit $\pi \circ F = f$ und $F(x_0) = p_0$.*

Wir wenden wir den Satz an auf $\pi_{\Gamma'} : \mathbb{R}^2 \to \mathbb{R}^2/\Gamma'$ und $f = \varphi \circ \pi_\Gamma : \mathbb{R}^2 \to \mathbb{R}^2/\Gamma'$. Wähle $b \in \mathbb{R}^2$ mit $[b]_{\Gamma'} = \varphi([0]_\Gamma)$. Es gibt dann genau ein $\phi \in C^0(\mathbb{R}^2, \mathbb{R}^2)$ mit

$$\pi_{\Gamma'} \circ \phi = \varphi \circ \pi_\Gamma \quad \text{und} \quad \phi(0) = b.$$

Hier das Ganze in einem (nicht besonders schönen) Diagramm:

$$\begin{array}{ccccccc}
 & \mathbb{R}^2, 0 & \cdots & \xrightarrow{\phi} & \cdots & \mathbb{R}^2, b & \\
 & \vdots & \ddots & & & \vdots & \\
\pi_\Gamma & \downarrow & \varphi \circ \pi_\Gamma & \searrow & & \downarrow & \pi_{\Gamma'} \\
 & \vdots & & & \ddots & \vdots & \\
 & \mathbb{R}^2/\Gamma, [0]_\Gamma & \cdots & \xrightarrow{\varphi} & \cdots & \mathbb{R}^2/\Gamma', [b]_{\Gamma'} &
\end{array}$$

Wir zeigen nun ϕ differenzierbar. Zu $x \in \mathbb{R}^2$ wählen wir Umgebung V von $\phi(x)$ mit $\pi_{\Gamma'} : V \to \pi_{\Gamma'}(V)$ diffeomorph. Da ϕ stetig, gibt es Umgebung U von x mit $\phi(U) \subset V$. Es folgt

$$\phi|_U = (\pi_{\Gamma'}|_V)^{-1} \circ \varphi \circ \pi_\Gamma,$$

also ist ϕ glatt. Sei nun $\lambda : \mathbb{R}^2/\Gamma \to (0,\infty)$ der Konformfaktor von φ. Dann gilt

$$\phi^* g_{\mathbb{R}^2} = \phi^* \pi_{\Gamma'}^* g_{\Gamma'} = \pi_\Gamma^* \varphi^* g_{\Gamma'} = (\lambda \circ \pi_\Gamma)^2 g_{\mathbb{R}^2}.$$

Ist $\det D\phi > 0$, so ist $\phi : \mathbb{C} \to \mathbb{C}$ holomorph. Aber $|\phi'(z)| = \lambda(\pi_\Gamma(z))$ ist beschränkt, also ist $\phi'(z)$ konstant nach Liouville und es folgt $\phi(z) = az + b$. Im Fall $\det D\phi < 0$ argumentieren wir analog für $\phi(\overline{z})$, damit ist der zweite Teil der Behauptung gezeigt. Für $\gamma \in \Gamma$ folgt weiter

$$[b]_{\Gamma'} = \varphi([0]_\Gamma) = \varphi([\gamma]_\Gamma) = [A\gamma + b]_{\Gamma'},$$

also $A\gamma + b \in \Gamma' + b$ bzw. $A\Gamma \subset \Gamma'$. Aber $\varphi^{-1}([y]_{\Gamma'}) = [A^{-1}(y-b)]_\Gamma$. Es gilt dann analog $A^{-1}\Gamma' \subset \Gamma$, und somit $A\Gamma = \Gamma'$. $\square$

Behauptung: jedes Gitter $\Gamma = \mathbb{Z}v_1 \oplus \mathbb{Z}v_2$ ist unter $\mathbb{R}^+\mathbb{O}(2)$ äquivalent zu einem Gitter

$$\mathbb{Z}e_1 + \mathbb{Z}v \quad \text{mit } v \in M = \Big\{v \in \mathbb{R}^2 : |v| \geq 1,\ 0 \leq v^1 \leq \frac{1}{2},\ v^2 > 0\Big\}.$$

Das sehen wir wie folgt: wähle einen kürzesten Vektor e in $\Gamma \backslash \{0\}$, und dann einen kürzesten Vektor v in $\Gamma \backslash \mathbb{Z}e$. Beachte dabei, dass in jeder Kreisscheibe $B_R(0)$ nur endlich viele Gitterpunkte liegen. Denn ist A die lineare Abbildung mit $Av_i = e_i$, so gilt

$$\big(m_1^2 + m_2^2\big)^{\frac{1}{2}} = |m_1 e_1 + m_2 e_2| = |A(m_1 v_1 + m_2 v_2)| \leq C\,|m_1 v_1 + m_2 v_2|.$$

e, v sind linear unabhängig, denn sonst ist $v = \alpha e$ mit $k < \alpha < k+1$ für ein $k \in \mathbb{Z}$, und

$$0 \neq (\alpha - k)e = v - ke \in \Gamma \text{ mit } |(\alpha-k)e| < |e|, \ \text{ Widerspruch.}$$

Wir können annehmen dass $\langle v, e\rangle \geq 0$, sonst ersetze v durch $-v$. Nach Drehung und Streckung des Gitters gilt nun $e = e_1$. Durch evtl. Spiegelung des Gitters an der x-Achse wird $v^2 > 0$, dabei bleibt $v^1 \geq 0$. Wegen $v - e_1 \in \Gamma \backslash \mathbb{Z}e_1$ folgt nun

$$|v|^2 \leq |v - e_1|^2 = |v|^2 - 2v^1 + 1, \quad \text{ also } v^1 \leq \frac{1}{2} \text{ und damit } v \in M.$$

Wir zeigen nun, dass e, v das Gitter Γ erzeugen. Sei dazu $w = \lambda e + \mu v \in \Gamma$ gegeben, ohne Einschränkung $-\frac{1}{2} \leq \lambda, \mu \leq \frac{1}{2}$. Wir zeigen $\lambda = \mu = 0$. Es gilt

$$|w| \leq |\lambda| + |\mu|\,|v| \leq (|\lambda| + |\mu|)|v| \leq |v|.$$

Angenommen es ist $w \neq 0$. Dann folgt $w \notin \mathbb{Z}e_1$ und somit $|w| \geq |v|$, wir hätten also Gleichheit. Dann sind λe_1 und μv linear abhängig, das geht nur für $\lambda = 0$ oder $\mu = 0$. Aber dann ist $(|\lambda| + |\mu|)|v| \leq \frac{1}{2}|v| < |v|$, Widerspruch.

Theorem 8.7 (Klassifikation flacher Tori) *Jeder Torus $\mathbb{R}^2/\Gamma$ ist konform äquivalent zu genau einem Torus $\mathbb{R}^2/(\mathbb{Z}e_1 + \mathbb{Z}v)$ mit $v \in M$.*

Beweis: Es ist nur noch die Eindeutigkeit zu zeigen. Seien $\Gamma_i = \mathbb{Z}e_1 \oplus \mathbb{Z}v_i$, $i = 1,2$, mit $v_i \in M$. Es gelte $\Gamma_2 = \lambda S \Gamma_1$ für $\lambda > 0$ und $S \in \mathbb{O}(2)$. Dann folgt

$$1 = \min_{y \in \Gamma_2 \backslash \{0\}} |y| = \min_{x \in \Gamma_1 \backslash \{0\}} |\lambda S x| = \lambda \min_{x \in \Gamma_1 \backslash \{0\}} |x| = \lambda.$$

S induziert eine Isometrie zwischen den $\mathbb{R}^2/\Gamma_i$, deren Flächeninhalt ist $\langle v_i, e_2 \rangle$. Also folgt $\langle v_1, e_2 \rangle = \langle v_2, e_2 \rangle$. Wir zeigen nun $|v_1| = |v_2|$; daraus folgt dann $v_1 = v_2$ bzw. $\Gamma_1 = \Gamma_2$.

Fall 1: $Se_1 = \pm e_1$. Dann gilt $S\mathbb{Z}e_1 = \mathbb{Z}e_1$, und es folgt

$$|v_2| = \min_{y \in \Gamma_2 \backslash \mathbb{Z}e_1} |y| = \min_{x \in \Gamma_1 \backslash \mathbb{Z}e_1} |Sx| = |v_1|.$$

Fall 2: $Se_1 \neq \pm e_1$. Dann gilt $Se_1 \in \Gamma \backslash \mathbb{Z}e_1$, also

$$1 \leq |v_2| = \min_{y \in \Gamma_2 \backslash \mathbb{Z}e_1} |y| \leq |Se_1| = 1.$$

Es gilt auch $S^{-1}e_1 \neq \pm e_1$, also folgt analog $|v_1| = 1$. □

Der Satz hat eine wesentliche Verallgemeinerung, nämlich den *Uniformisierungssatz im Fall von Geschlecht Eins:* jeder Riemannsche Torus (T, g), also T diffeomorph zu $\mathbb{R}^2/\mathbb{Z}^2$, ist konform äquivalent zu genau einem flachen Torus $\mathbb{R}^2/\Gamma$ mit $\Gamma = \mathbb{Z}e_1 + \mathbb{Z}v$, $v \in M$. Die Eindeutigkeit haben wir schon bewiesen. Für die Existenz bestimmt man eine konforme Metrik $g_0 = e^{2u}g$ mit Krümmung $K_{g_0} = 0$. Die Formel für die Krümmung (vgl. das Theorema Egregium) lautet

$$-\Delta_g u - K_g = K_{g_0} e^{2u}.$$

Es ist also die Laplacegleichung $-\Delta_g u = K_g$ zu lösen. Nach der Fredholmschen Alternative ist das genau dann möglich, wenn die rechte Seite in $L^2(g)$ senkrecht zum Kern des Operators Δ_g ist. Der Kern besteht nur aus den Konstanten, denn mit partieller Integration für $\Delta_g = \mathrm{div}_g \mathrm{grad}_g$ folgt

$$0 = -\int_T f \Delta_g f \, d\mu_g = \int_T \|\mathrm{gradf}\|_g^2 \, d\mu_g.$$

Die Fredholmsche Alternative folgt mit Gauß-Bonnet, es gilt

$$\int_T K_g \, d\mu_g = 2\pi \chi(T) = 0.$$

Der Torus (T, g_0) ist isometrisch zu einem der flachen Tori $\mathbb{R}^2/\Gamma$ (Satz von Minding).

9 Kovariante Ableitungen

In diesem Kapitel führen wir den Begriff der kovarianten Ableitung in einem Vektorbündel ein. Für eine vektorwertige C^1-Funktion $\xi : M \to \mathbb{R}^k$ kennen wir die übliche Ableitung

$$D\xi \in C^0(T^*M \otimes \mathbb{R}^k), \; D\xi = \sum_{i=1}^k d\xi^i \otimes e_i.$$

Der Operator $D : C^1(M, \mathbb{R}^k) \to C^0(T^*M \otimes \mathbb{R}^k)$ ist linear in ξ, und es gilt

$$D(f\xi) = \sum_{i=1}^k d(f\xi^i) \otimes e_i = df \otimes \xi + f\, D\xi \quad \text{für } f \in C^1(M).$$

Sei nun $\pi : E \to M$ ein Vektorbündel vom Rang k. Für einen C^1-Schnitt $\xi : M \to E$, also $\pi \circ \xi = \mathrm{id}_M$, könnten wir auch die übliche Ableitung $D\xi$ betrachten. Diese bildet aber nach TE ab, mehrfache Ableitungen würden auf komplizierte Bündel führen. Außerdem enthält $D\xi$ überflüssige Information, denn es gilt ja $D\pi(\xi(p))D\xi(p) = D(\pi \circ \xi)(p) = \mathrm{Id}_{T_pM}$.

Definition 9.1 (Kovariante Ableitung) *Sei $\pi : E \to M$ ein Vektorbündel vom Rang k. Eine kovariante Ableitung oder Zusammenhang auf E ist eine lineare Abbildung $\nabla : C^1(E) \to C^0(T^*M \otimes E)$, die folgende Produtregel erfüllt:*

$$\nabla(f\xi) = df \otimes \xi + f\, \nabla\xi. \tag{9.12}$$

Wir zeigen unten, dass es kovariante Ableitungen auf E gibt. Erstmal betrachten wir die Sache aber in einer lokalen Trivialisierung $\phi : \pi^{-1}(V) \to V \times \mathbb{R}^k$. Wir haben dann die Basisfelder

$$v_i \in C^1(E|_V),\ v_i(p) = \phi^{-1}(p, e_i) \quad \text{für } i = 1, \ldots, k.$$

Bezüglich der Basis $\mathcal{A} = \{v_1, \ldots, v_k\}$ wird ∇ durch eine 1-Form A auf V mit Werten in $\mathbb{R}^{k\times k}$ dargestellt. Und zwar definieren wir

$$\nabla_X v_j = \sum_{i=1}^k A^i_j(X)\, v_i \quad \text{für } X \in C^0(T^*M|_V) \text{ und } j = 1, \ldots, k. \tag{9.13}$$

Für einen beliebigen C^1-Schnitt $\xi = \sum_{j=1}^k \xi^j v_j$ folgt mit der Produktregel

$$\nabla_X \xi = \sum_{j=1}^k d\xi^j(X) v_j + \sum_{i,j=1}^k \xi^j A^i_j(X) v_i = \sum_{i=1}^k \big(d\xi^i(X) + \sum_{j=1}^k A^i_j(X)\xi^j\big)\, v_i.$$

Sei $\xi_{\mathcal{A}}$ der Koordinatenvektor von ξ bezüglich der Basis $\mathcal{A}$. Dann lautet die Formel

$$(\nabla\xi)_{\mathcal{A}} = d\xi_{\mathcal{A}} + A\, \xi_{\mathcal{A}}, \quad \text{oder kurz } \nabla = d + A. \tag{9.14}$$

Sei jetzt $\psi : \pi^{-1}(W) \to W \times \mathbb{R}^k$ eine andere Trivialisierung, mit Basis $\mathcal{B} = \{w_1, \ldots, w_k\}$ auf W, also $w_j \in C^1(E|_W)$, $w_j(p) = \psi^{-1}(p, e_j)$. Setze

$$S = \mathrm{Id}_{\mathcal{AB}} : V \cap W \to \mathrm{GL}_k(\mathbb{R}),\ w_j = \sum_{i=1}^k S^i_j\, v_i.$$

S ist die Matrix der Übergangsfunktion zwischen ϕ und ψ, es gilt

$$\phi \circ \psi^{-1}(p, e_j) = \phi(p, w_j) = \sum_{i=1}^k S^i_j\, \phi(p, v_i) = \sum_{i=1}^k S^i_j\, e_i.$$

Für die Koordinatenvektoren $\xi_{\mathcal{A}}$ und $\xi_{\mathcal{B}}$ gilt

$$\xi = \sum_{j=1}^{k} \xi_{\mathcal{B}}^j\, w_j = \sum_{i=1}^{k}\sum_{j=1}^{k} S_j^i\, \xi_{\mathcal{B}}^j\, v_i, \qquad \text{also } \xi_{\mathcal{A}} = S\xi_{\mathcal{B}}.$$

Für die Koordinatendarstellung $(\nabla\xi)_{\mathcal{A}}$ ergibt sich die Relation

$$\begin{aligned}(\nabla\xi)_{\mathcal{A}} &= d\xi_{\mathcal{A}} + A\xi_{\mathcal{A}}\\ &= d(S\xi_{\mathcal{B}}) + AS\xi_{\mathcal{B}}\\ &= S\, d\xi_{\mathcal{B}} + (dS)\,\xi_{\mathcal{B}} + AS\xi_{\mathcal{B}}.\end{aligned}$$

Andererseits ist $(\nabla\xi)_{\mathcal{A}} = S\,(\nabla\xi)_{\mathcal{B}} = S\big(d\xi_{\mathcal{B}} + B\,\xi_{\mathcal{B}}\big)$. Durch Vergleich ergibt sich, wenn wir noch $T = \mathrm{Id}_{\mathcal{B}\mathcal{A}} = S^{-1}$ setzen,

$$B = S^{-1}AS + S^{-1}dS \quad \text{bzw.} \quad B = TAT^{-1} - (dT)T^{-1}. \tag{9.15}$$

Beachte dabei $0 = d(ST) = (dS)T + S(dT) = S(S^{-1}dS + (dT)T^{-1})T$.

Sei $G \subset \mathrm{GL}_k(\mathbb{R})$ eine Untergruppe. Man sagt, die Strukturgruppe des Bündels E ist nach G reduzierbar, wenn es einen Bündelatlas gibt mit Übergangsfunktionen in G.

Lemma 9.1 *Jedes Vektorbündel E vom Rang k über M besitzt eine Riemannsche Metrik. Insbesondere ist die Strukturgruppe immer nach $\mathbb{O}(k)$ reduzierbar.*

Beweis: Sei $\phi_\lambda : \pi^{-1}(U_\lambda) \to U_\lambda \times \mathbb{R}^k$. Definiere auf U_λ die Riemannsche Metrik

$$g_\lambda(p)(v,w) = \langle \phi_\lambda(p,v), \phi_\lambda(p,w)\rangle_{\mathbb{R}^k}.$$

Überdecke M mit Gebieten U_λ und verwende eine untergeordnete Teilung der Eins. Mittels Gram-Schmidt kann jede lokale Basis in eine Orthonormalbasis überführt werden, daraus folgt die zweite Aussage. □

Definition 9.2 (metrischer Zusammenhang) *Sei g Riemannsche Metrik auf dem Vektorbündel E. Eine kovariante Ableitung auf E heißt metrisch, wenn gilt:*

$$d[g(v,w)] = g(\nabla v, w) + g(v, \nabla w) \quad \textit{für alle } v,w \in C^1(E).$$

Lemma 9.2 *Sei g Riemannsche Metrik und ∇ Zusammenhang auf E. Dann gilt:*

(1) *Ist ∇ metrisch, so gilt $A + A^{\mathrm{T}} = 0$ für jede Orthonormalbasis $\mathcal{E}$.*

(2) *Ist $\mathcal{E}$ Orthonormalbasis auf U mit $A + A^{\mathrm{T}} = 0$, so ist ∇ metrisch auf $E|_U$.*

Beweis: Bezüglich einer Orthonormalbasis $\mathcal{E} = \{e_1, \ldots, e_k\}$ gilt

$$d\,[g(e_i,e_j)] = 0 \quad \text{und} \quad g(\nabla e_i, e_j) + g(e_i, \nabla e_j) = A_i^j + A_j^i.$$

Ist ∇ metrisch so folgt direkt $A + A^{\mathrm{T}} = 0$. Umgekehrt gilt, da A schiefsymmetrisch,

$$\begin{aligned} g(\nabla v, w) + g(v, \nabla w) &= \langle dv_{\mathcal{E}} + Av_{\mathcal{E}}, w_{\mathcal{E}}\rangle_{\mathbb{R}^k} + \langle v_{\mathcal{E}}, dw_{\mathcal{E}} + Aw_{\mathcal{E}}\rangle_{\mathbb{R}^k}\\ &= d\,\langle v_{\mathcal{E}}, w_{\mathcal{E}}\rangle_{\mathbb{R}^k}\\ &= d\,[g(v,w)].\end{aligned}$$

□

Lemma 9.3 *Die Strukturgruppe eines Vektorbündels E vom Rang k über M ist genau dann nach $\mathbb{S}O(k)$ reduzierbar, wenn das Bündel orientierbar ist.*

Beweis: Es gebe einen Bündelatlas mit Übergangsfunktionen in $\mathbb{S}O(k)$. Für zugehörige lokale Trivialisierung definieren wir eine Orientierung durch die Basis $\phi^{-1}(e_j)$, $1 \leq j \leq k$. Das liefert eine wohldefinierte Orientierung. Ist umgekehrt eine Orientierung gegeben, so definiere diesbezüglich positiv orientierte Karten (ggf. durch Verkettung mit einer Spiegelung). $\square$

Theorem 9.1 (Raum der Zusammenhänge) *Auf jedem Vektorbündel $\pi : E \to M$ existiert eine kovariante Ableitung. Ist ∇_0 eine feste kovariante Ableitung, so ist die Menge aller kovarianten Ableitungen gegeben durch*

$$\nabla = \nabla_0 + \Omega \quad \text{mit } \Omega \in C^0\big(T^*M \otimes \mathrm{End}(E)\big).$$

Beweis: Wähle lokale Trivialisierungen $\phi_\lambda : \pi^{-1}(V_\lambda) \to V_\lambda \times \mathbb{R}^k$, $\lambda \in \Lambda$, so dass M durch die V_λ überdeckt wird. Seien $v_{\lambda,i}$ die Basisfelder von ϕ_λ, und betrachte auf $E|_{V_\lambda}$

$$\nabla_\lambda \xi = \sum_{i=1}^{k} d\xi^i_\lambda \otimes v_{\lambda,i}.$$

Wähle eine untergeordnete Teilung der Eins η_λ und setze

$$\nabla \xi = \sum_{\lambda \in \Lambda} \eta_\lambda \nabla_\lambda \xi.$$

Dann ist $\nabla : C^1(E) \to C^0(T^*M \otimes E)$ wohldefiniert und linear. Da ∇_λ eine kovariante Ableitung auf $E|_{V_\lambda}$ ist, gilt für $f \in C^1(M)$

$$\nabla(f\xi) = \sum_{\lambda \in \Lambda} \eta_\lambda \nabla_\lambda (f\xi) = \sum_{\lambda \in \Lambda} \eta_\lambda \big(df \otimes \xi + f\,\nabla_\lambda \xi\big) = df \otimes \xi + f\,\nabla\xi.$$

Damit ist die Existenz einer kovarianten Ableitung gezeigt. Seien nun ∇ und ∇_0 zwei kovariante Ableitungen. Bezüglich einer lokalen Basis $\mathcal{A} = \{v_1, \ldots, v_k\}$ auf U haben ∇, ∇_0 Darstellungen $d + A$ bzw. $d + A_0$. Definiere auf Ω auf U durch

$$\Omega\,\xi = \sum_{i,j=1}^{k} (A - A_0)^i_j \xi^j v_i = \nabla\xi - \nabla_0 \xi.$$

Wegen der rechten Darstellung hängt Ω nicht von der Wahl der Trivialisierung ab, und ist damit die gesuchte 1-Form mit Werten in $\mathrm{End}(E)$. $\square$

Remark 9.1 Hat E eine Riemannsche Metrik, so gibt es einen metrischen Zusammenhang. Wir können dazu annehmen, dass die Basen $\mathcal{A}_\lambda$ der Trivialisierungen orthonormal sind. Nach Definition hat ∇_λ bezüglich $\mathcal{A}_\lambda$ die Zusammenhangsform Null, und ist damit metrisch auf $E|_{V_\lambda}$ nach Lemma 9.2(2). An der Definition von $\nabla\xi$ sind lokal nur endlich viele ∇_λ beteiligt. Nach Lemma 9.2(1) hat jeder dieser Zusammenhänge bezüglich einer fest gewählten Orthonormalbasis eine schiefsymmetrische Zusammenhangsform, also auch ∇. Nach Lemma 9.2(2) ist ∇ dann ein metrischer Zusammenhang.

Sei ∇ gegebener Zusammenhang auf E. Wir definieren jetzt die kovariante Ableitung eines Feldes $\xi \in C^1([a,b],E)$ längs $c \in C^1([a,b],M)$, also $\pi \circ \xi = c$. Man nennt c die Fußpunktkurve von ξ. Wir nehmen zuerst an dass $c(t) \in V$ für alle $t \in [a,b]$, wobei auf V eine Basis $\mathcal{A}$ existiert. Dann setzen wir

$$\left(\frac{\nabla\xi}{dt}\right)_{\mathcal{A}} = \frac{d\xi_{\mathcal{A}}}{dt} + A(c')\,\xi_{\mathcal{A}}. \tag{9.16}$$

Genauer sollten wir $(A \circ c)(c')$ schreiben, es ist aber üblich den Fußpunkt da wegzulassen. Wir zeigen jetzt, dass die Definition nicht von der Basis abhängt. Sei auch $c(t) \in W$ für alle $t \in [a,b]$, und auf W gebe es eine lokale Basis $\mathcal{B}$. Es gilt dann $\xi_{\mathcal{A}}(t) = S(c(t))\xi_{\mathcal{B}}(t)$ mit $S = \mathrm{Id}_{\mathcal{A}\mathcal{B}}$. Es folgt für die Zusammenhangsformen A, B

$$\begin{aligned}
\frac{d\xi_{\mathcal{A}}}{dt} + A(c')\,\xi_{\mathcal{A}} &= \frac{d((S\circ c)\xi_{\mathcal{B}})}{dt} + A(c')(S\circ c)\xi_{\mathcal{B}} \\
&= (S\circ c)\Big(\frac{d\xi_{\mathcal{B}}}{dt} + (S\circ c)^{-1} dS(c')\,\xi_{\mathcal{B}} + (S\circ c)^{-1}A(c')(S\circ c)\,\xi_{\mathcal{B}}\Big) \\
&= (S\circ c)\Big(\frac{d\xi_{\mathcal{B}}}{dt} + B(c')\,\xi_{\mathcal{B}}\Big).
\end{aligned}$$

Somit gilt $(\frac{\nabla\xi}{dt})_{\mathcal{A}} = \mathrm{Id}_{\mathcal{A}\mathcal{B}}\,(\frac{\nabla\xi}{dt})_{\mathcal{B}}$, das heißt $\frac{\nabla\xi}{dt}$ ist unabhängig von der Wahl der Basis definiert.

Lemma 9.4 *Die kovariante Ableitung $\frac{\nabla\xi}{dt}$ ist wohldefiniert und hat folgende Eigenschaften:*

(a) *Linearität: für C^1-Felder ξ, η längs c und $\lambda, \mu \in \mathbb{R}$ gilt*

$$\frac{\nabla(\lambda\xi + \mu\eta)}{dt} = \lambda\,\frac{\nabla\xi}{dt} + \mu\,\frac{\nabla\eta}{dt}. \tag{9.17}$$

(b) *Produktregel: für ein C^1-Feld ξ längs c auf $[a,b]$ und $f \in C^1([a,b])$ gilt*

$$\frac{\nabla(f\xi)}{dt} = \frac{df}{dt}\,\xi + f\,\frac{\nabla\xi}{dt}. \tag{9.18}$$

(c) *Ist ∇ metrisch bezüglich g, so gilt für C^1-Felder ξ, η längs c*

$$\frac{d}{dt}\,g(\xi,\eta) = g\Big(\frac{\nabla\xi}{dt},\eta\Big) + g\Big(\xi,\frac{\nabla\eta}{dt}\Big). \tag{9.19}$$

Beweis: Dies geschieht durch Nachrechnen in einer lokalen Basis, für (c) in einer orthonormalen lokalen Basis. $\square$

Definition 9.3 (paralleles Feld) *Sei ∇ kovariante Ableitung auf dem Vektorbündel E. Ein Feld $\xi \in C^1([a,b],E)$ heißt parallel (längs $c = \pi \circ \xi$), falls gilt:*

$$\frac{\nabla\xi}{dt} = 0 \qquad \textit{auf } [a,b]. \tag{9.20}$$

Ein Feld $\xi(t)$ im Produktbündel $M \times \mathbb{R}^k$ hat die Form $(c(t), v(t))$ mit $v(t) \in \mathbb{R}^k$. Der Produktzusammenhang ist

$$\frac{\nabla\xi}{dt}(t) = \big(c(t), \frac{dv}{dt}(t)\big).$$

Das Feld ξ ist parallel längs c, wenn $\frac{dv}{dt} = 0$ ist, also $v(t)$ konstant. Wir können uns vorstellen, dass sich $v(t)$ durch Parallelverschiebung längs der Kurve $c(t)$ ergibt. Der Begriff des parallelen (oder kovariant konstanten) Felds im Vektorbündel E verallgemeinert die Situation. Bezüglich einer lokalen Basis $\mathcal{A}$ sieht die Bedingung wie folgt aus:

$$\left(\frac{\nabla\xi}{dt}\right)_{\mathcal{A}} = \frac{d\xi_{\mathcal{A}}}{dt} + A(c')\,\xi_{\mathcal{A}} = 0. \tag{9.21}$$

Es handelt sich um ein lineares, homogenes System von k Differentialgleichungen erster Ordung für die Funktion $\xi_{\mathcal{A}} : I \to \mathbb{R}^k$.

Lemma 9.5 (parallele Felder) *Sei $\pi : E \to M$ Vektorbündel mit kovarianter Ableitung ∇, und $c \in C^1([a,b], M)$ gegeben mit $c(t_0) = x_0$. Dann gibt es zu jedem $v \in E_{x_0}$ genau ein paralleles Feld $\xi_v \in C^1([a,b], E)$ längs c mit $\xi_v(t_0) = v$. Die Abbildung $v \mapsto \xi_v$ ist ein Vektorraumisomorphismus.*

Beweis: Sei zunächst $c : [a,b] \to V$, wobei auf V eine Basis $\mathcal{A} = \{v_1, \ldots, v_k\}$ existiert. Die Funktion $(A \circ c)c'$ ist stetig auf $[a,b]$. Die Lösungstheorie für lineare Systeme besagt, dass (9.21) zu gegebenem Anfangswert eine eindeutige Lösung auf dem ganzen Intervall $[a,b]$ hat. Genauer gilt die lokale Lösbarkeit nach Picard-Lindelöf, und die Existenz auf ganz $[a,b]$ verwendet eine Wachstumsschranke nach dem Lemma von Gronwall. Im allgemeinen wähle eine Unterteilung $a = t_0 < \ldots < t_N = b$, so dass $c([t_{i-1}, t_i])$ im Bereich einer lokalen Basis liegt, und konstruiere stückweise die Lösung. $\square$

Definition 9.4 (Parallelverschiebung) *Sei $\pi : E \to M$ Vektorbündel mit kovarianter Ableitung ∇. Für eine Kurve $c \in C^1([a,b], M)$ definieren wir den Vektorraumisomorphismus*

$$P_c : E_{c(a)} \to E_{c(b)},\ P_c(v) = \xi_v(b), \qquad \textit{wobei } \xi_v \textit{ parallel längs } c \textit{ mit } \xi_v(a) = v.$$

Hierzu drei Bemerkungen.

- Sei ∇ metrisch bezüglich g. Dann gilt für parallele Felder ξ, η längs c
$$\frac{d}{dt}\, g(\xi, \eta) = g\Big(\frac{\nabla\xi}{dt}, \eta\Big) + g\Big(\xi, \frac{\nabla\eta}{dt}\Big) = 0.$$
Also ist die Parallelverschiebung eine Isometrie, denn
$$g(P_c v, P_c w) = g(\xi_v(b), \xi_w(b)) = g(\xi_v(a), \xi_w(a)) = g(v, w).$$

- Ist E orientiertes Vektorbündel, so ist P_c orientierungstreu. Sei erst $c : [a,b] \to V$, wobei auf V eine positiv orientierte Basis $\mathcal{A}$ existiert. Sei $P\mathcal{A}(t)$ die Parallelverschiebung von $\mathcal{A}(c(a))$ längs $c(t)$. Nach dem Zwischenwertsatz hat die Transformationsmatrix zwischen $P\mathcal{A}(t)$ und $\mathcal{A}(c(t))$ immer positive Determinante, also erhält die Parallelverschiebung die Orientierung. Für c beliebig verwende das stückweise auf den Intervallen einer Unterteilung von $[a,b]$.

- Die Parallelverschiebung ändert sich nicht bei Umparametrisierung von c. Denn sei $t = \varphi(s)$ mit $s \in [\alpha, \beta]$. Dann ist $\tilde{\xi}(s) = \xi(\varphi(s))$ ein Feld längs $\tilde{c}(s) = c(\varphi(s))$, und in einer lokalen Basis gilt

$$\begin{aligned}\frac{\nabla\tilde{\xi}}{ds}(s) &= \frac{d\tilde{\xi}}{ds}(s) + A(\tilde{c}(s))(\tilde{c}'(s))\,\tilde{\xi}(s)\\ &= \frac{d\xi}{dt}(\varphi(s))\varphi'(s) + (A(c(t))(c'(t))\,\xi(t)|_{t=\varphi(s)}\,\varphi'(s)\\ &= \frac{\nabla\xi}{dt}(\varphi(s))\,\varphi'(s).\end{aligned}$$

Ist $\xi(t)$ parallel längs $c(t)$, so ist $\tilde{\xi}(s) = \xi(\varphi(s))$ parallel längs $\tilde{c}(s) = c(\varphi(s))$. Insbesondere gilt $P_{c\circ\varphi} = P_c$.

Im allgemeinen hängt die Parallverschiebung aber sehr wohl von der Kurve c ab.

Example 9.1 Betrachte auf $T\mathbb{S}^2$ die kovariante Ableitung $\nabla X = (DX)^\top$. Offenbar ist ∇ linear und für $f : \mathbb{S}^2 \to \mathbb{R}$ gilt $\nabla(fX) = df \otimes X + f\nabla X$. Sei $X \in C^1([a,b], T\mathbb{S}^2)$ ein Vektorfeld längs $c \in C^1([a,b], \mathbb{S}^2)$. Dann gilt

$$\frac{\nabla X}{dt} = \Big(\frac{dX}{dt}\Big)^\top.$$

Betrachte speziell einen Großkreis

$$c(t) = (\cos t)p + (\sin t)q \qquad \text{mit } |p| = |q| = 1,\ \langle p, q\rangle = 0.$$

Dann ist $v_1(t) = c'(t)$ parallel längs $c(t)$, denn es gilt $(c''(t))^\top = -(c(t))^\top = 0$. Ein zweites paralleles Feld längs $c(t)$ ist $v_2(t) = p \times q$ (Kreuzprodukt), das ist die Normale zur Ebene des Großkreises. Jedes andere parallele Feld längs c ergibt sich als Linearkombination von v_1, v_2 mit konstanten Koeffizienten. Es ist leicht zusehen, dass zum Beispiel die Parallelverschiebung längs des Randes eines Oktanten nicht Null ist.

Definition 9.5 *Sei ∇ ein C^1-Zusammenhang auf dem Vektorbündel $\pi : E \to M$. Die Krümmung von ∇ ist, für $X, Y \in C^1(TM)$ und $\xi \in C^1(E)$,*

$$F(X,Y)\xi = \nabla_X(\nabla_Y\xi) - \nabla_Y(\nabla_X\xi) - \nabla_{[X,Y]}\xi.$$

Hier sind X, Y Vektorfelder auf M und ξ ist ein Feld in E, genauer $F \in C^0(\Lambda^2 TM \otimes \mathrm{End}(E))$.

Das ist natürlich ein kompliziertes Objekt, das die Differentialgeometrie seit Riemann beschäftigt. Wir nähern uns der Sache wie üblich, nämlich durch Übersetzung in lokale Koordinaten. In einer Karte $\varphi : U \to \varphi(U)$ von M haben wir die kovarianten Ableitungen in Richtung der Koordinatenfelder, also

$$\nabla_\alpha\xi = \nabla_{\frac{\partial}{\partial x^\alpha}}\,\xi \in C^0(E|_U).$$

Bekanntlich gilt $\big[\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\big] = 0$, die Definition der Krümmung lautet dann

$$F\Big(\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\Big)\,\xi = (\nabla_\alpha\nabla_\beta - \nabla_\beta\nabla_\alpha)\,\xi.$$

Das macht auf jeden Fall Sinn, die Krümmung misst die Abweichung bei Vertauschung der Operatoren ∇_α und ∇_β, also die Asymmetrie der zweiten kovarianten Ableitung. Wir wollen nun die Formel für die Krümmung bezüglich einer lokalen Basis von E herleiten. Betrachte dazu auf $V \subset M$ zwei 1-Formen A, B mit Werten in $\mathbb{R}^{k\times k}$, also lokal

$$A = \sum_{\alpha=1}^{n} A_\alpha dx^\alpha, \quad B = \sum_{\alpha=1}^{n} B_\alpha dx^\alpha \quad \text{ mit } A_\alpha, B_\alpha : V \to \mathbb{R}^{k\times k}.$$

Wir defineren dann

$$\begin{aligned} dA &= \sum_{\alpha,\beta=1}^{n} \partial_\alpha A_\beta \, dx^\alpha \wedge dx^\beta, \\ A \wedge B &= \sum_{\alpha,\beta=1}^{n} A_\alpha B_\beta \, dx^\alpha \wedge dx^\beta. \end{aligned}$$

Im Unterschied zu reellen 1-Formen gilt hier im allgemeinen nicht $A \wedge A = 0$, sondern

$$A \wedge A = \frac{1}{2} \sum_{\alpha,\beta=1}^{n} [A_\alpha, A_\beta] \, dx^\alpha \wedge dx^\beta = \sum_{1\le\alpha<\beta\le n} [A_\alpha, A_\beta] \, dx^\alpha \wedge dx^\beta,$$

wobei $[A_\alpha, A_\beta]$ der Kommutator der Matrizen ist.

Lemma 9.6 (lokale Formel der Krümmung) *Sei ∇ kovariante Ableitung auf E. Dann gilt bezüglich einer lokalen Basis $\mathcal{E}$ mit Zusammenhangsform A*

$$F_{\mathcal{A}} = dA + A \wedge A.$$

Beweis: Wir berechnen bezüglich der Basis $\mathcal{A}$ und einer lokalen Karte auf M

$$\begin{aligned} \Big(F\big(\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\big)\,\xi\Big)_{\mathcal{A}} &= \big(\nabla_\alpha(\nabla_\beta \xi) - \nabla_\beta(\nabla_\alpha \xi)\big)_{\mathcal{A}} \\ &= (\partial_\alpha + A_\alpha)\,(\nabla_\beta \xi)_{\mathcal{A}} - (\partial_\beta + A_\beta)\,(\nabla_\alpha \xi)_{\mathcal{A}} \\ &= (\partial_\alpha + A_\alpha)(\partial_\beta + A_\beta)\,\xi_{\mathcal{A}} - (\partial_\beta + A_\beta)(\partial_\alpha + A_\alpha)\,\xi_{\mathcal{A}} \\ &= (\partial_\alpha A_\beta - \partial_\beta A_\alpha)\xi_{\mathcal{A}} + (A_\alpha A_\beta - A_\beta A_\alpha)\xi_{\mathcal{A}} \\ &= dA\big(\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\big)\,\xi_{\mathcal{A}} + A \wedge A\big(\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\big)\,\xi_{\mathcal{A}}. \end{aligned}$$

□

Lemma 9.7 (Krümmung und Basiswechsel) *Seien $\mathcal{A}$, $\mathcal{B}$ lokale Basen von E. Dann gilt*

$$F_{\mathcal{B}} = T\,F_{\mathcal{A}}\,T^{-1} \quad \textit{mit } T = \mathrm{Id}_{\mathcal{B}\mathcal{A}}.$$

Beweis: Es gilt für eine 1-Form A und eine Funktion B, beide mit Werten in $\mathbb{R}^{k\times k}$,

$$\begin{aligned} d(AB) &= \sum_{\alpha=1}^{n} dx^\alpha \wedge \partial_\alpha \sum_{\beta=1}^{n} A_\beta B \, dx^\beta \\ &= \sum_{\alpha,\beta=1}^{n} \big((\partial_\alpha A_\beta)\,B + A_\beta\,\partial_\alpha B\big)\,dx^\alpha \wedge dx^\beta \\ &= (dA)\,B - A \wedge dB. \end{aligned}$$

Bei Vertauschung von A und B ergibt sich

$$\begin{aligned} d(BA) &= \sum_{\alpha=1}^{n} dx^\alpha \wedge \partial_\alpha \sum_{\beta=1}^{n} BA_\beta \, dx^\beta \\ &= \sum_{\alpha,\beta=1}^{n} \big((\partial_\alpha B)\, A_\beta + B\, \partial_\alpha A_\beta\big)\, dx^\alpha \wedge dx^\beta \\ &= dB \wedge A + B(dA). \end{aligned}$$

Damit berechnen wir

$$\begin{aligned} F_{\mathcal{B}} &= dB + B \wedge B \\ &= d\,(TAT^{-1} - (dT)T^{-1}) + (TAT^{-1} - (dT)T^{-1}) \wedge (TAT^{-1} - (dT)T^{-1}) \\ &= dT \wedge AT^{-1} + T(dA)T^{-1} - TA \wedge d(T^{-1}) + dT \wedge d(T^{-1}) \\ &\quad + T(A \wedge A)T^{-1} - TAT^{-1} \wedge (dT)T^{-1} - (dT) \wedge AT^{-1} + (dT)T^{-1} \wedge (dT)T^{-1}. \end{aligned}$$

Verwende nun $d(TT^{-1}) = 0$, also $d(T^{-1}) = -T^{-1}(dT)T^{-1}$. Das ergibt

$$\begin{aligned} dT \wedge AT^{-1} - (dT) \wedge AT^{-1} &= 0 \\ -TA \wedge d(T^{-1}) - TAT^{-1} \wedge (dT)T^{-1} &= 0, \\ dT \wedge d(T^{-1}) + (dT)T^{-1} \wedge (dT)T^{-1} &= 0. \end{aligned}$$

Insgesamt folgt $F_{\mathcal{B}} = T\,(dA + A \wedge A)\,T^{-1} = T\,F_{\mathcal{A}}\,T^{-1}$. □

Theorem 9.2 (Flache Zusammenhänge) *Sei ∇ Zusammenhang auf E. Äquivalent sind:*

(a) *∇ hat Krümmung $F^\nabla = 0$.*

(b) *Sind $c_0, c_1 : [a,b] \to M$ C^2-homotop mit festen Endpunkten, so folgt $P_{c_0} = P_{c_1}$.*

Beweis: Sei erst D Zusammenhang auf dem trivialen Bündel $Q \times \mathbb{R}^k$, mit $Q = [a,b] \times [0,1]$. Durch Parallelverschiebung erhalten wir eine Basis $\mathcal{B} = \{v_1(s,t), \ldots, v_k(s,t)\}$, so dass

$$\begin{aligned} v_j(a,0) &= e_j && \text{(die Standardbasis)}, \\ \frac{Dv_j}{dt}(a,t) &= 0 && \text{für alle } t \in [0,1], \\ \frac{Dv_j}{ds}(s,t) &= 0 && \text{für alle } s \in [a,b],\ t \in [0,1]. \end{aligned}$$

Sei $\omega : Q \to \Lambda^1(\mathbb{R}^2) \otimes \mathbb{R}^{k \times k}$ die Zusammenhangsform dieser Basis, also

$$(D\xi)_{\mathcal{B}} = d\xi_{\mathcal{B}} + \omega \xi_{\mathcal{B}} \quad \text{für } \xi \in C^1(Q, \mathbb{R}^k).$$

Es gilt $(v_j)_{\mathcal{B}} = e_j$, also $(Dv_j)_{\mathcal{B}} = \omega e_j$, und nach Konstruktion

$$\omega_2(a,t) = \omega(a,t)\Big(\frac{\partial}{\partial t}\Big) = 0 \quad \text{und} \quad \omega_1(s,t) = \omega(s,t)\Big(\frac{\partial}{\partial s}\Big) = 0.$$

Damit berechnen wir für alle $(s,t) \in Q$

$$\omega_2(s,t) = \int_a^s \big(\partial_1 \omega_2 - \partial_2 \omega_1 + [\omega_1, \omega_2]\big)(\sigma, t)\, d\sigma = \int_a^s (F^D_{12})_{\mathcal{B}}(\sigma, t)\, d\sigma. \tag{9.22}$$

Außerdem gilt nach Definition

$$\left(\frac{Dv_j}{dt}\right)_{\mathcal{B}}(s,t) = \omega_2(s,t)e_j. \tag{9.23}$$

Ist nun $F^D = 0$, so folgt $\omega_2 = 0$ aus (9.22), also ist $\omega = 0$ und $Dv_j = 0$. Ein Feld ξ ist damit genau dann parallel längs einer Kurve in Q, wenn es konstante Koordinaten bezüglich der Basis $\mathcal{B}$ hat. Insbesondere liefert Verschiebung rund um ∂Q die Identität.

Es gelte umgekehrt die Wegunabhängigkeit. Verschieben wir $v_j(s,0)$ parallel im Rechteck hintereinander nach $(a,0)$, (a,t) und (s,t), so ergibt sich $v_j(s,t)$. Damit ist $v_j(s,t)$ auch die Parallelverschiebung von $v_j(s,0)$ längs der vierten Seite $\{s\} \times [0,t]$. Aus (9.23) folgt $\omega_2 e_j = 0$ und damit $\omega = 0$, also $F^D_{\mathcal{B}} = d\omega + \omega \wedge \omega = 0$.

Um dieses Argument auf ein Bündel über M anzuwenden, konstruieren wir einen Pullback. Wir nehmen dazu an, dass die Homotopie $c \in C^2([a,b] \times [0,1], M)$, $c = c(s,t)$, in eine offene Menge V abbildet, auf der eine Basis $\mathcal{A} = \{a_1, \ldots, a_k\}$ existiert. Sei $A \in C^1(T^*V \otimes \mathbb{R}^{k \times k})$ die Zusammenhangsform von $\mathcal{A}$. Definiere auf $Q \times \mathbb{R}^k$ den Zusammenhang

$$D\xi(s,t) = d\xi(s,t) + (c^*A)(s,t)\,\xi(s,t) \qquad \text{für } \xi : Q \to \mathbb{R}^k.$$

Sei $\xi \in C^1(I, \mathbb{R}^k)$ ein Feld längs der Kurve $\gamma \in C^1(I,Q)$. Dann haben wir längs $\tilde{\gamma} = c \circ \gamma$ den Lift $\tilde{\xi} = \sum_{j=1}^k \xi^j\, a_j \circ \tilde{\gamma}$. Es gilt

$$\begin{aligned}
\frac{\nabla \tilde{\xi}}{du} &= \sum_{j=1}^k \frac{d\xi^j}{du}\, a_j \circ \tilde{\gamma} + \sum_{i,j=1}^k \xi^j A^i_j \circ \tilde{\gamma}(\tilde{\gamma}')\, a_i \circ \tilde{\gamma} \\
&= \sum_{i=1}^k \Big(\frac{d\xi^i}{du} + \sum_{j=1}^k (c^*A^i_j) \circ \gamma(\gamma')\,\xi^j\Big)\, a_i \circ \tilde{\gamma} \\
&= \sum_{i=1}^n \Big(\frac{D\xi}{du}\Big)^i a_i \circ \tilde{\gamma}.
\end{aligned}$$

$\tilde{\xi}$ ist genau dann parallel längs $\tilde{\gamma}$, wenn ξ parallel längs γ ist. Liefert Parallelverschiebung rund um ∂Q im Pullbackbündel die Identität, so gilt das auch längs c in $E|_V$, und vice versa. Für die Krümmungen folgt, bezüglich der Standardbasis bzw. der Basis $\mathcal{A}$,

$$F^D = d(c^*A) + (c^*A) \wedge (c^*A) = c^*(dA + A \wedge A) = c^* F^{\nabla}_{\mathcal{A}}. \tag{9.24}$$

Aus $F^{\nabla} = 0$ folgt also $F^D = 0$, und damit die Wegunabhängigkeit der Parallelverschiebung wie in (b) behauptet, zunächst nur in $E|_V$. Umgekehrt können wir zu $p \in M$ und $v, w \in T_pM$ eine Homotopie $c(s,t)$ wählen mit $c(0,0) = p$, $\frac{\partial c}{\partial s}(0,0) = v$ und $\frac{\partial c}{\partial t}(0,0) = w$. Aus der Wegunabhängigkeit folgt dann

$$0 = F^D_{12}(a,0) = F^{\nabla}_{\mathcal{A}}\Big(\frac{\partial c}{\partial s}, \frac{\partial c}{\partial t}\Big)|_{s,t=0} = F^{\nabla}_{\mathcal{A}}(p)(v,w).$$

Es bleibt $(a) \Rightarrow (b)$ zu zeigen, wenn das Bild von $c : [a,b] \times [0,1] \to M$ nicht im Bereich einer Trivialisierung liegt. Betrachte $Q_{\mu\nu} = [s_{\mu-1}, s_\mu] \times [t_{\nu-1}, t_\nu]$ für $a = s_0 < \ldots < s_M = b$

und $0 = t_0 < \ldots < t_N = 1$. Die Unterteilung sei so gewählt, dass jedes $c(Q_{\mu\nu})$ im Gebiet einer Basis liegt. Wie bewiesen ist dann die Parallelverschiebung längs jeder Masche $\partial Q_{\mu\nu}$ die Identität. Ausgehend von der Randkurve ∂Q kann man durch Weglassen von Maschen zu der konstanten Kurve $c(a,0)$ gelangen, ohne dabei den Wert der Parallelverschiebung zu ändern. Also folgt allgemein $P_{c_0} = P_{c_1}$. $\square$

Es ist naheliegend zu fragen, ob und wie die Parallelverschiebung abgeschätzt werden kann, wenn die Krümmung kontrolliert ist.

Theorem 9.3 (Abschätzung des Paralleltransports) *Sei (M,g) Riemannsche Mannigfaltigkeit, und $\pi : E \to M$ ein Vektorbündel mit Metrik $\langle \cdot\,, \cdot \rangle$ und metrischem Zusammenhang ∇. Sei $c \in C^2(Q,M)$, $Q = [a,b] \times [0,1]$, eine Homotopie. Dann gilt für den Paralleltransport rund um ∂Q die Abschätzung*

$$|P_{\partial Q} - \mathrm{Id}| \leq \Lambda e^{\Lambda} \qquad \textit{mit } \Lambda = \int_Q \|F^\nabla\|_g \, d\mu_g.$$

Beweis: Wir gehen wie im Beweis von Satz 9.2 vor. Allerdings nehmen wir an, dass D metrisch ist bezüglich des Standardskalarprodukts. Die Basis $\mathcal{B}$ ist dann orthonormal, und die Zusammenhangsform ω ist schiefsymmetrisch. Wir verwenden in den Abschätzungen die Hilbert-Schmidt Norm für Matrizen. Ist ξ parallel längs ∂Q, so ist $\xi_{\mathcal{B}}$ konstant außer auf $b \times [0,1]$. Dort gilt

$$0 = \Big(\frac{D\xi}{dt}\Big)_{\mathcal{B}}(t) = \frac{d\xi_{\mathcal{B}}}{dt}(t) + \omega_2(b,t)\xi_{\mathcal{B}}(t).$$

Die Ungleichung von Gronwall (bzw. Integration) liefert, da $\mathcal{B}$ Orthonormalbasis,

$$|\xi(b,t)| \leq |\xi(b,0)|\, e^{\kappa} \qquad \text{mit } \kappa = \int_0^1 |\omega_2(b,s)|\, ds.$$

Rückeinsetzen in die Differentialgleichung oben ergibt

$$|\xi(b,1) - \xi(b,0)| \leq \int_0^1 |\omega_2(b,t)|\, |\xi(b,t)|\, dt \leq |\xi(b,0)|\, \kappa e^{\kappa}.$$

Nach (9.22) ist

$$\kappa = \int_0^1 \Big| \int_a^b F_{12}^D(s,t)\, ds \Big|\, dt \leq \int_0^1 \int_a^b |F_{12}^D(s,t)|\, ds\, dt.$$

Sei $J^g c$ die Jacobische von c bzgl. g. Ist τ_1, τ_2 Orthonormalbasis von Bild $Dc(s,t)$, so gilt

$$\Big| F_{\mathcal{A}}^\nabla \Big(\frac{\partial c}{\partial s}, \frac{\partial c}{\partial t}\Big)\Big| = |F_{\mathcal{A}}^\nabla(\tau_1,\tau_2)|\, J^g c.$$

Dies sieht man leicht für geeignete τ_1, τ_2. Mit $d\mu_g(s,t) = J^g c(s,t)\, ds\, dt$ folgt

$$\int_Q |F_{12}^D|\, ds\, dt = \int_Q \Big| F_{\mathcal{A}}^\nabla \Big(\frac{\partial c}{\partial s}, \frac{\partial c}{\partial t}\Big)\Big|\, ds\, dt \leq \int_Q \|F^\nabla\|_g\, d\mu_g.$$

Wir haben benutzt, dass die Hilbert-Schmidt Norm von $F_{\mathcal{A}}^\nabla$ nicht von der Orthonormalbasis $\mathcal{A}$ abhängt. Dies beweist die Abschätzung, wenn das Bild von c im Gebiet einer Trivialisierung liegt. Allgemein wählen wir wieder eine Unterteilung von Q wie im Beweis von Satz 9.2. Wir haben dann für jede Masche eine Fehlerabschätzung

$$|P_{\partial Q_{\mu\nu}} - \mathrm{Id}| \leq \Lambda_{\mu\nu} e^{\Lambda_{\mu\nu}} \quad \text{mit } \Lambda_{\mu\nu} = \int_{Q_{\mu\nu}} \|F^\nabla\|_g \, d\mu_g.$$

Bei sukzessive Streichung der Maschen folgt mit der Dreiecksungleichung

$$|P_{\partial Q} - \mathrm{Id}| \leq \sum_{\mu,\nu} |P_{\partial Q_{\mu\nu}} - \mathrm{Id}| \leq \sum_{\mu,\nu} \Lambda_{\mu\nu} e^{\Lambda_{\mu\nu}} \leq e^\Lambda \sum_{\mu,\nu} \Lambda_{\mu\nu} = \Lambda e^\Lambda.$$

□

10 Shortest connections

Let $g \in C^0(T^{2,0}M)$ be a Riemannian metric on the n-dimensional connected manifold M. In Definition 8.3 the Riemannian distance of points $p, q \in M$ was introduced as

$$d(p,q) = \inf \Big\{ L(c) : c \in PC^1([a,b], M),\, c(a) = p,\, c(b) = q \Big\}.$$

Here $L(c)$ is the Riemannian length of the curve $c(t)$ given by

$$L(c) = \int_a^b \|c'(t)\| \, dt \quad \text{where } \|c'(t)\| = g(c(t))\big(c'(t), c'(t)\big)^{\frac{1}{2}}.$$

In general the infimum need not be attained. For instance, existence fails for $M = \mathbb{R}^n \backslash \{0\}$ and $q = -p \neq 0$, see Example 8.2. However, we will prove that for appropriate (M, g) the distance is realized by some connecting curve. For this we employ the direct method in the calculus of variations, that is we study the convergence of a minimizing sequence. The existence of shortest connections was probably the first variational problem where the direct method was applied (Hilbert 1900). Our approach involves two basic steps:

(a) the choice of a convergence (or topology), which allows to extract a limit c of the minimizing sequence c_k, after passing to a subsequence.

(b) the verification of the lower semicontinuity $L(c) \leq \liminf_{k \to \infty} L(c_k)$, for the limit as constructed in step (a).

Actually, step (a) cannot be carried out as it stands. To obtain a limit c in the class $PC^1([a,b], M)$ one needs strong convergence of the c_k, essentially convergence in the C^1 norm. However, all we know is a bound on the length of the c_k, this does not give any pointwise control on c_k'. To find a limit we are forced to enlarge the class of admissible curves, and to extend the definition of length to this enlarged class. This situation is quite typical for variational problems, we proceed as follows.

Definition 10.1 (length functional) *The length of $c \in C^0([a,b], M)$ with respect to g is*

$$L(c) = \sup_Z L_Z(c) \quad \textit{where } L_Z(c) = \sum_{i=1}^N d(c(t_{i-1}, c(t_i)).$$

Here the supremum is over all partitions $Z : a = t_0 \leq \ldots \leq t_N = b$ of the interval $[a,b]$.

For curves $c \in PC^1([a,b],\mathbb{R}^n)$ we have now two competing definitions of $L(c)$; below we show that these coincide. From now on $L(c)$ is always as in Definition 10.1 unless stated otherwise. The partition Z consisting of the two endpoints a, b yields $L_Z(c) = d(c(a), c(b))$. Therefore we have the following basic estimate:

$$d(c(a), c(b)) \leq L(c). \tag{10.1}$$

We now collect some facts about the definition of the length.

Lemma 10.1 (additivity)

$$L(c) = L(c|_{[a,\tau]}) + L(c|_{[\tau,b]}) \quad \textit{for any } \tau \in [a,b].$$

Proof. Put $I_1 = [a,\tau]$ and $I_2 = [\tau, b]$. If $Z_{1,2}$ are partitions of $I_{1,2}$, then $Z = (Z_1, Z_2)$ is a partition of $I = [a,b]$, and

$$L_{Z_1}(c|_{I_1}) + L_{Z_2}(c|_{I_2}) = L_Z(c) \leq L(c).$$

Taking the supremum over all Z_1 and Z_2 yields

$$L(c|_{I_1}) + L(c|_{I_2}) \leq L(c).$$

Reversely, let $Z : a = t_0 \leq \ldots \leq t_N = b$ be a partition of I. Choose $r \in \{1, \ldots, N\}$ with $t_{r-1} \leq \tau \leq t_r$, and consider the partitions $Z_1 : t_0 \leq \ldots \leq t_{r-1} \leq \tau$ and $Z_2 : \tau \leq t_r \leq \ldots \leq t_N$ of $I_{1,2}$. Then we estimate by the triangle inequality

$$\begin{aligned}
L_Z(c) &= \sum_{i=1}^{N} d(c(t_{i-1}), c(t_i)) \\
&\leq \sum_{i=1}^{r-1} d(c(t_{i-1}), c(t_i)) + d(c(t_{r-1}), c(\tau)) + d(c(\tau), c(t_r)) + \sum_{i=r+1}^{N} d(c(t_{i-1}), c(t_i)) \\
&\leq L(c|_{I_1}) + L(c|_{I_2}).
\end{aligned}$$

The result follows by taking the supremum over Z. □

The following fact is analogous to a well-known property of the Riemann integral.

Lemma 10.2 (approximation) *Let $c \in C^0([a,b], M)$. Then for any $L < L(c)$ there exists a $\delta > 0$ such that*

$$L_Z(c) > L \quad \textit{for any partition } Z \textit{ with } \Delta(Z) < \delta.$$

Proof. Chose a partition $Z_0 : a = s_0 \leq \ldots \leq s_M = b$ such that $L_{Z_0}(c) > L$, and take $\varepsilon > 0$ with $L + \varepsilon < L_{Z_0}(c)$. By uniform contituity we can choose $\delta > 0$ with

$$d(c(t), c(t')) < \frac{\varepsilon}{2M} \quad \text{whenever } |t - t'| < \delta.$$

Now let $Z : a = t_0 \leq \ldots \leq t_N = b$ be any partition with $\Delta(Z) < \delta$. Assume that the open interval (t_{j-1}, t_j) contains $m_j > 0$ of the points s_i. Then

$$L_{Z_0}(c|_{[t_{j-1},t_j]}) \leq (m_j + 1)\frac{\varepsilon}{2M} \leq m_j \frac{\varepsilon}{M}.$$

Now consider the partition $Z \cup Z_0$ of $[a, b]$. We estimate by the triangle inequality

$$L_{Z_0}(c) \leq L_{Z \cup Z_0}(c) \leq \sum_{m_j = 0} d(c(t_{j-1}, t_j)) + \sum_{m_j > 0} m_j \frac{\varepsilon}{M} \leq L_Z(c) + \varepsilon.$$

Recalling $L + \varepsilon < L_{Z_0}(c)$ we get $L_Z(c) > L$ as claimed. □

Lemma 10.3 (absolute continuity) *Let $c \in C^0([a, b], M)$ be a curve with $L(c) < \infty$. Then for any $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$L(c|_J) < \varepsilon \quad \text{for any } J = [\alpha, \beta] \subset [a, b] \text{ with } |J| < \delta.$$

Proof. By Lemma 10.2 we can choose $\delta > 0$ such that

$$\begin{aligned} d(c(t), c(t')) &< \frac{\varepsilon}{2} && \text{for } |t - t'| < \delta, \\ L(c) &< L_Z(c) + \frac{\varepsilon}{2} && \text{for any partition } Z \text{ of } [a, b] \text{ with } \Delta(Z) < \delta. \end{aligned}$$

Let Z be a partition of $[a, b]$ with $\Delta(Z) < \delta$, and such that J is a subinterval. Then for any partition Z' of J we compute

$$L_Z(c) + L_{Z'}(c|_J) = L_{Z \cup Z'}(c) + d(c(\alpha), c(\beta)) < L(c) + \frac{\varepsilon}{2}.$$

Thus $L_{Z'}(c|_J) < L(c) - L_Z(c) + \frac{\varepsilon}{2} < \varepsilon$. □

We now show that the notions of length are the same for PC^1-curves.

Theorem 10.1 (arclength formula) *For any $c \in PC^1([a, b], M)$ we have*

$$L(c) = \int_a^b \|c'(t)\| \, dt.$$

Proof. By Lemma 10.1 both sides are additive with respect to subdivision, hence we can assume $c \in C^1([a, b], \mathbb{R}^n)$. We have for any partition Z

$$\sum_{i=1}^{N} d(c(t_i), c(t_{i-1})) \leq \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} \|c'(t\| \, dt = \int_a^b \|c'(t)\| \, dt.$$

Taking the supremum we see that

$$L(c) \leq \int_a^b \|c'(t)\| \, dt.$$

Now consider the arclength function $\ell(t) = L(c|_{[a,t]}$. For $a \leq t < t + h \leq b$ we have

$$\frac{d(c(t+h), c(t))}{h} \leq \frac{\ell(t+h) - \ell(t)}{h} \leq \frac{1}{h} \int_t^{t+h} \|c'(t)\| \, dt \xrightarrow{h \searrow 0} \|c'(t)\|.$$

Here we used the additivity of length with respect to the 3-point partition $a \leq t < t + h$. We claim that

$$\liminf_{h \searrow 0} \frac{d(c(t+h), c(t))}{h} \geq \|c'(t)\|.$$

Then we conclude $\ell'(t) = \|c'(t)\|$, in particular $\ell \in C^1([a,b]$, and the claim follows by the fundamental theorem of calculus.

Now for $p = c(t)$ we select a chart $\varphi : U \to B_\varrho(0)$ with $\varphi(p) = 0$ und $g_{ij}(p) = \delta_{ij}$. By restricting to a smaller $\varrho > 0$, we can assume that

$$\|v\|_{g(x)} \geq (1-\varepsilon)|v| \text{ for all } x \in B_\varrho(0).$$

Let $\gamma : I \to M$ be a piecewise C^1-curve going from p to $q \in U$. If $\varphi(I) \subset U$ then

$$L(\gamma) \geq (1-\varepsilon) L^{\mathbb{R}^n}(\varphi \circ \gamma) \geq (1-\varepsilon)|\varphi(q)|.$$

But otherwise, the argument shows $L(\gamma) \geq (1-\varepsilon)\varrho$. Therefore we conclude $d(q,p) \geq (1-\varepsilon)|\varphi(q)|$. For $h > 0$ we have $c(t+h) \in U$. Applying our estimate for $q = c(t+h)$, and using $\varphi(c(t)) = 0$, we have

$$d(c(t+h), c(t)) \geq (1-\varepsilon)|\varphi(c(t+h)) - \varphi(c(t))|.$$

Dividing by $h > 0$ end letting $h \searrow 0$ we get

$$\begin{aligned} \liminf_{h \searrow 0} \frac{d(c(t+h), c(t))}{h} & \geq (1-\varepsilon) \liminf_{h \searrow 0} \Big| \frac{\varphi(c(t+h)) - \varphi(c(t))}{h} \Big| \\ & = (1-\varepsilon)|D\varphi(c(t))c'(t)| \\ & = (1-\varepsilon)\|c'(t)\|. \end{aligned}$$

The claim follows by letting $\varepsilon \searrow 0$. □

Theorem 10.2 (Lower semicontinuity of length) *Let $c_k, c \in C^0([a,b], M)$ such that $c_k \to c$ pointwise on $[a,b]$. Then*

$$L(c) \leq \liminf_{k \to \infty} L(c_k).$$

Proof. Let $Z : a = t_0 \leq \ldots \leq t_N = b$ be a partition of $[a,b]$. Since d is continuous on $M \times M$, we have

$$L_Z(c) = \sum_{i=1}^N d(c(t_i), c(t_{i-1})) = \lim_{k \to \infty} \sum_{i=1}^N d(c_k(t_i), c_k(t_{i-1})) \leq \liminf_{k \to \infty} L(c_k).$$

Now take the supremum over Z. □

When enlarging the class of admissible functions for a variational problem, it may happen that the infimum of the functional on the larger class becomes strictly smaller, this is called the Lavrentiev gap phenomenon. For instance, consider the walking distance in New York City, minimizing over paths following the grid. On the larger class of all PC^1 paths, the distance becomes smaller by a factor $1/\sqrt{2}$. Luckily such a gap does not come up in our situation, namely the basic etimate implies

$$\begin{aligned} d(p,q) & \leq \inf \big\{ L(c) : c \in C^0([a,b], M),\, c(a) = p,\, c(b) = q \big\} \\ & \leq \inf \big\{ L(c) : c \in PC^1([a,b], M),\, c(a) = p,\, c(b) = q \big\} \\ & = d(p,q). \end{aligned}$$

We now take up the problem of convergence for a minimizing sequence. A first difficulty is the invariance of the length under reparametrizations. For instance if $c : [0,1] \to M$ is a given shortest connection of its endpoints p, q, then the same ist true for the curves $c_k(t) = c(t^k)$, $t \in [0,1]$, where $k \in \mathbb{N}$. But then

$$c_k(t) = c(t^k) \stackrel{k\to\infty}{\longrightarrow} \begin{cases} p = c(0) & \text{für } 0 \leq t < 1, \\ q = c(1) & \text{für } t = 1. \end{cases}$$

To rule out this loss of compactness, we need to fix some parametrization.

Definition 10.2 *$c \in C^0(I, M)$ is length-preserving (or isometric) if*

$$L(c|_{[s_1,s_2]}) = |s_1 - s_2| \quad \textit{for all } [s_1, s_2] \in I.$$

We note that a length-preserving curve $c(s)$ is Lipschitz continuous with constant $L = 1$, in fact the basic estimate shows

$$d(c(s_1), c(s_2)) \leq L(c|_{[s_1,s_2]}) = |s_1 - s_2|. \tag{10.2}$$

Now let $c \in c \in C^0([a,b], M)$ be a curve with $L = L(c) < \infty$, and consider the associated arclength function

$$\ell : [a,b] \to [0, L],\ \ell(t) = L(c|_{[a,t]}).$$

By Lemma 10.1 and Lemma 10.3 $\ell(t)$ is monotonically increasing (not necessarily strictly) and continuous with $\ell(a) = 0$ and $\ell(b) = L$. Put

$$\tilde{c} : [0, L] \to M,\ \tilde{c}(s) = c(t) \quad \text{where } \ell(t) = s.$$

The existence of some $t \in [a,b]$ with $\ell(t) = s$ follows by the intermediate value theorem. If $\ell(t_1) = \ell(t_2) = s$, then again by the basic estimate

$$d(c(t_1), c(t_2)) = L(c|_{[t_1,t_2]}) = \ell(t_2) - \ell(t_1) = 0.$$

Theorem 10.3 (Reparametrization by arclength) *The above curve $\tilde{c}(s)$ is well-defined and length-preserving, and it has the same endpoints as the given curve $c(t)$.*

Proof. We need to show that $L(\tilde{c}|_{[s_1,s_2]}) = |s_1 - s_2|$. Choose $t_{1,2} \in [a,b]$ with $\ell(t_{1,2}) = s_{1,2}$, and let Z_j be a sequence of partitions of $[t_1, t_2]$ with mesh $\Delta(Z_j) \to 0$. Then $\ell(Z_j)$ is a sequence of partitions of $[s_1, s_2]$ with $\Delta(\ell(Z_j)) \to 0$ by Lemma 10.3. Using Lemma 10.2 we obtain

$$\begin{aligned} L(\tilde{c}|_{[s_1,s_2]}) &= \lim_{j\to\infty} L_{\ell(Z_j)}(\tilde{c}|_{[s_1,s_2]}) \\ &= \lim_{j\to\infty} L_{Z_j}(c|_{[t_1,t_2]}) \\ &= L(c|_{[t_1,t_2]}) \\ &= |\ell(t_1) - \ell(t_2)| = |s_1 - s_2|. \end{aligned}$$

The existence of a convergent subsequence is an application of the following principle.

Theorem 10.4 (Arzelà-Ascoli 1895) *Let $f_k : (X, d) \to (Y, d)$ be a sequence of maps, where X is compact. Assume that:*

(1) *f_k is uniformly bounded:* $\exists\, B_R(y_0) \subset Y$ *with* $f_k(X) \subset B_R(y_0)$ *for all* k*;*

(2) *f_k is equicontinuous, i.e.*

$$\sup_{k \in \mathbb{N}} \sup_{d(x,x')<\delta} d(f(x), f(x')) \to 0 \quad \text{as } \delta \searrow 0.$$

(3) *For each $x \in X$ the sequence $f_k(x)$ has a convergent subsequence.*

Then there exists an $f \in C^0(X,Y)$ such that $f_k \to f$ in $C^0(X,Y)$ (uniformly) as $k \to \infty$, after passing to a subsequence.

Our goal is to apply this result in the following situation:

- $c_k \in C^0([0,1], M)$ is a sequence of curves from $p = c_k(0)$ to $q = c_k(1)$, with length $L_k \to L := d(p,q) < \infty$. In other words, c_k is a minimizing sequence.
- c_k is parametrized proportional to arclength, i.e. $L(c_k|_{[s_1,s_2]}) = L_k|s_1 - s_2|$.

The second condition is achieved by first passing to a parametrization on $[0, L_k]$ by arclength, and then scaling the interval to $[0,1]$. The effect of the scaling is that all c_k are defined on the same interval $[0,1]$. Now we check the conditions of Theorem 10.4.

(1) For any $k \in \mathbb{N}$, $s \in [0,1]$ we have

$$d(p, c_k(s)) = d(c_k(0), c_k(s)) \leq L(c_k) \leq \Lambda < \infty.$$

Here Λ is a common bound for $L_k = L(c_k)$. Thus we $c_k([0,1]) \subset \overline{B_\Lambda(p)}$.

(2) For $|s_1 - s_2| < \delta$ we have the estimate

$$d(c_k(s_1), c_k(s_2)) \leq L(c_k|_{[s_1,s_2]}) = L_k\, \delta \leq \Lambda\, \delta.$$

The bound is independent of k, therefore the sequence c_k is equicontinuous.

(3) This condition is not satisfied a priori, as the example $M = \mathbb{R}^n \backslash \{0\}$ shows. We ensure (3) by imposing a condition on M.

Theorem 10.5 (Existence of shortest connections) *Let M be a connected differentiable manifold with Riemannian metric $g \in C^0(T^{2,0}M)$. Assume that distance balls $B_r(p) = \{q \in M : d(p,q) < R\}$ are relatively compact. Then $d(p,q) < \infty$ for all $p, q \in M$, and there exists a curve $c \in C^0([0,1], M)$ from p to q with length $L(c) = d(p,q)$.*

Proof. Let $c_k \in C^0([0,1], M)$ be the minimizing sequence considered above. As $d(p, c_k(s)) \leq \Lambda$, condition (3) follows by assumption. We conclude that after passing to a subsequence, we have $c_k \to c \in C^0([0,1], M)$ uniformly, in particular $c(0) = p$ and $c(1) = q$. By lower semicontinuity, see Theorem 10.2, we have $L(c) \leq \liminf_{k \to \infty} L(c_k) = d(p,q)$. The theorem is proved. □

A final remark: it is of interest to modify the minimization problem by requiring that all c_k belong to a fixed homotopy class. This is easily possible. In particular, the reparametrization by arclength preserves the homotopy class. For this let $c : [a,b] \to M$ be given; by a linear substitution we can assume that $[a,b] = [0,L]$ for $L = L(c)$. Define

$$\lambda : [0,L] \times [0,1] \to [0,L],\ \lambda(t,u) = (1-u)t + u\ell(t),$$

and further

$$h :: [0, L] \times [0, 1] \to M,\ h(s, u) = c(t) \text{ if } \lambda(t, u) = s.$$

We have $\lambda(t, 0) = t$, so $h(s, 0) = c(s)$. Also $\lambda(t, 1) = \ell(t)$, which implies $h(s, 1) = \tilde{c}(s)$. We need to show that h is continuous, then it is a homotopy between c and $\tilde{c}$ (exercise).

Lemma 10.4 (Regularity I) *Let g be a Riemannian metric of class C^r on M where $r \geq 1$. If $c \in C^1(I, M)$, $I = (a, b)$, is a locally shortest curve parametrized by arclength, then c belongs to $C^{r+1}(I, M)$.*

Proof. The statement is local, hence we can assume that $c \in C^1([a, b], U)$ is a shortest connection in the open set $U \subset \mathbb{R}^n$ with respect to the metric $g = (g_{ij}) \in C^r(U, \mathbb{R}^{n \times n})$.

Step 1. *First variation of length*
For any $\phi \in C^1_c(I, \mathbb{R}^n)$ the curve $c + t\phi$ is in U for $t \in (\varepsilon, \varepsilon)$ small. Therefore the function $t \mapsto L(c + t\phi)$ has a minimum at $t = 0$. We have

$$L(c + t\phi) = \int_0^L g(c + t\phi)(c' + t\phi', c' + t\phi')^{\frac{1}{2}}\, ds \qquad \text{where } c = c(s),\ \phi = \phi(s).$$

Here we write $g(c)$ instead of $g \circ c$ for convenience. As $g(c)(c', c') = 1$ by assumption, differentiation under the integral is justified and we get

$$\int_0^L \Big(\sum_{i,k=1}^n g_{ik}(c)(c^i)'(\phi^k)' + \frac{1}{2} \sum_{k=1}^n \partial_k g(c)(c', c')\phi^k \Big)\, ds = 0. \tag{10.3}$$

Step 2. *Higher regularity: $r \geq 2$*
Let $g \in C^r(U)$ for $r \geq 2$, and assume we know already $c \in C^r(I, U)$. We can then integrate by parts in (10.3). Applying the fundamental lemma of the calculus of variations, we get the Euler-Lagrange system

$$\frac{d}{ds} \sum_{i=1}^n g_{ik}(c)(c^i)' = \frac{1}{2} \partial_k g(c)(c', c') \in C^{r-1}(I) \qquad \text{for } k = 1, \ldots, n.$$

The functions $f_k = \sum_{i=1}^n g_{ik}(c)(c^i)'$ thus belong to $C^r(I)$. Multiplying by $g^{jk}(c)$ and summing for $k = 1, \ldots, n$ gives

$$(c^j)' = \sum_{k=1}^n g^{jk}(c) f_k \in C^r(I),$$

which proves $c \in C^{r+1}(I, U)$.

Step 3. *Initial regularity: $r = 1$*
Here we assume that g and c are of class C^1. Then the partial integration as in step 2 cannot be done, the argument must be modified. For $k = 1, \ldots, n$ we choose $f_k \in C^1(I)$ with

$$-f_k' = \frac{1}{2} \partial_k g(c)(c', c') \in C^0(I).$$

Multiplying by ϕ^k and integrating by parts gives

$$\int_0^L \sum_{k=1}^n f_k\, (\phi^k)'\, ds = \frac{1}{2} \int_0^L \sum_{k=1}^n \partial_k g(c)(c', c)\, \phi^k\, ds.$$

Subtracting this from (10.3) yields

$$\int_0^L \sum_{k=1}^n \Big(\sum_{i=1}^n g_{ik}(c)(c^i)' - f_k \Big) (\phi^k)' \, ds = 0.$$

For each standard vector e_k and $\varphi \in C^1_c(I)$, consider the variation

$$\phi \in C^1(I, \mathbb{R}^n), \; \phi(s) = \Big(\int_a^s \varphi(s') \, ds' \Big) \, e_k.$$

If $\int_I \varphi(s) \, ds = 0$ then ϕ has compact support in I and we obtain

$$\int_I \Big(\sum_{i=1}^n g_{ik}(c)(c^i)' - f_k \Big) \, \varphi \, ds = 0 \quad \text{for } k = 1, \ldots, n. \tag{10.4}$$

Fix some $\varphi_0 \in C^0_c(I)$ with $\int_I \varphi_0(s) \, ds = 1$. By adding a constant to f_k, we get

$$\int_I \Big(\sum_{i=1}^n g_{ik}(c)(c^i)' - f_k \Big) \varphi_0 \, ds = 0 \quad \text{for } k = 1, \ldots, n.$$

But any function $\varphi \in C^0_c(I)$ decomposes as $\varphi = \alpha\varphi_0 + (\varphi - \alpha\varphi_0)$ where

$$\int_I (\varphi - \alpha\varphi_0) \, ds = 0 \quad \text{for } \alpha = \int_I \varphi \, ds.$$

Thus equation (10.4) holds for all $\varphi \in C^0_c(I)$, and we conclude as in step 2

$$(c^j)' = \sum_{k=1}^n g^{jk}(c) f_k \in C^1(I).$$

This proves $c \in C^2(I, U)$ which is the start of the induction. $\square$

The trick in step 3 appears in J.K. Whittemore: Lagrange's equation in the calculus of variations, and the extension of a theorem of Erdmann, Ann. Math. **2** (1900), 130-136.

11 Some functional analysis

In the following Ω always denotes an open subset of $\mathbb{R}^n$.

Definition 11.1 (weak derivative) *We say that $u \in L^1_{\mathrm{loc}}(\Omega)$ has weak derivative $\partial_i u = g$, where also $g \in L^1_{\mathrm{loc}}(\Omega)$, if*

$$\int_\Omega u \, \partial_i \eta = - \int_\Omega g \, \eta \quad \textit{for all } \eta \in C^\infty_c(\Omega). \tag{11.1}$$

The function g is unique, in fact if (11.1) holds for $g_{1,2}$ then by subtracting we get

$$0 = \int_\Omega (g_1 - g_2)\eta \quad \text{for all } \eta \in C^\infty_c(\Omega),$$

and hence $g_1 = g_2$ by the fundamental lemma of the calculus of variations. The weak derivative operation is linear, and higher weak derivatives can be defined. A good exercise is to check which of the functions $u(x) = |x|^\alpha$, $\alpha \in \mathbb{R}$, has weak derivatives in $L^p_{\mathrm{loc}}(\mathbb{R}^n)$.

Definition 11.2 (Sobolev spaces) *For $1 \leq p \leq \infty$ we define $W^{1,p}(\Omega)$ as the space of functions $u \in L^p(\Omega)$ having weak derivatives $\partial_i u \in L^p(\Omega)$ for $i = 1, \ldots, n$, with norm*

$$\|u\|_{W^{1,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \sum_{i=1}^{n} \|\partial_i u\|_{L^p(\Omega)}.$$

Theorem 11.1 *$W^{1,p}(\Omega)$ is a Banach space.*

BEWEIS: Let u_k be a Cauchy sequence in $W^{1,p}(\Omega)$. Then u_k and also the $\partial_i u_k$ are Cauchy sequences in $L^p(\Omega)$. By Fischer-Riesz we have $u_k \to u$, $\partial_i u_k \to g_i$ in $L^p(\Omega)$. For any $\eta \in C^\infty_c(\Omega)$ we get

$$\int u\, \partial_i \eta = \lim_{k\to\infty} \int u_k\, \partial_i \eta = -\lim_{k\to\infty} \int (\partial_i\, u_k)\eta = -\int g_i \eta.$$

This shows $\partial_i u = g_i$ weakly, thus $u \in W^{1,p}(\Omega)$ and $u_k \to u$ in $W^{1,p}(\Omega)$. □

A useful technique in connection with Sobolev functions is the approximation by classical functions. A standard tool is the convolution, depending on a radius $\varrho > 0$. At each $x \in \mathbb{R}^n$ one takes a certain average over values on $B_\varrho(x)$ (see below). When dealing with functions on a domain $\Omega \subset \mathbb{R}^n$, one has the little disadvantage that the convolution is defined only on the smaller domain

$$\Omega^\varrho = \{x \in \Omega : \mathrm{dist}\,(x, \mathbb{R}^n \backslash \Omega) > \varrho\}.$$

Theorem 11.2 (Mollification on $L^p_{\mathrm{loc}}(\Omega)$) *Let $\eta \in C^\infty_c(\mathbb{R}^n)$, $\eta \geq 0$, have support in $B_1(0)$ and $\int_{\mathbb{R}^n} \eta(x)\, dx = 1$. For $u \in L^1_{\mathrm{loc}}(\Omega)$ consider the convolution*

$$\eta_\varrho * u : \Omega^\varrho \to \mathbb{R},\ \eta_\varrho * u(x) = \int \eta_\varrho(x-y)\, u(y)\, dy \qquad \text{where } \eta_\varrho(x) = \varrho^{-n}\, \eta\Big(\frac{x}{\varrho}\Big).$$

The following statements hold:

1. *$\eta_\varrho * u \in C^\infty(\Omega^\varrho)$ and $D^\alpha(\eta_\varrho * u) = \varrho^{-|\alpha|}(D^\alpha \eta)_\varrho * u$.*
2. *$\|\eta_\varrho * u\|_{L^p(\Omega^\varrho)} \leq \|u\|_{L^p(\Omega)}$ for $1 \leq p \leq \infty$.*
3. *If $p < \infty$ and $u \in L^p_{\mathrm{loc}}(\Omega)$ then $\eta_\varrho * u \to u$ in $L^p(\Omega')$ as $\varrho \searrow 0$, for any $\Omega' \subset\subset \Omega$.*
4. *If $u \in L^\infty_{\mathrm{loc}}(\Omega)$, then $\eta_\varrho * u \to u$ weak* in $L^\infty(\Omega')$, for any $\Omega' \subset\subset \Omega$.*

Remark: weak* convergence in $L^\infty(\Omega')$ means that

$$\int_{\Omega'} (\eta_\varrho * u)\varphi \to \int_{\Omega'} u\varphi \qquad \text{for all } \varphi \in L^1(\Omega').$$

For continuous functions the L^∞ norm is just the C^0 norm. Therefore if $u_k \to u$ in $L^\infty_{\mathrm{loc}}(\Omega)$ where $u_k \in C^0(\Omega)$, then u_k converges locally uniformly and thus u has a continuous representative. Therefore statement (4) cannot be improved to L^∞-convergence in general.

Lemma 11.1 (Mollification on $W^{1,p}_{\mathrm{loc}}(\Omega)$) *For $u \in W^{1,p}_{\mathrm{loc}}(\Omega)$ we have*

$$\partial_i(\eta_\varrho * u) = \eta_\varrho * \partial_i u \qquad \text{on } \Omega^\varrho.$$

Moreover for $p < \infty$ we have $\eta_\varrho * u \to u$ in $W^{1,p}_{\mathrm{loc}}(\Omega')$), for any $\Omega' \subset\subset \Omega$.

BEWEIS: We differentiate under the integral and then use the definition of weak derivative:

$$\begin{aligned}
\partial_i(\eta_\varrho * u)(x) &= \int \Big(\frac{\partial}{\partial x_i}\eta_\varrho(x-y)\Big)\, u(y)\, dy \\
&= -\int \Big(\frac{\partial}{\partial y_i}\eta_\varrho(x-y)\Big)\, u(y)\, dy \\
&= \int \eta_\varrho(x-y)\, \partial_i u(y)\, dy \\
&= (\eta_\varrho * \partial_i u)(x).
\end{aligned}$$

The second statement now follows using Theorem 11.2(3). □

Lemma 11.2 (product rule) *If $u, v \in (W^{1,p} \cap L^\infty)_{\mathrm{loc}}(\Omega)$ then $uv \in (W^{1,p} \cap L^\infty)_{\mathrm{loc}}(\Omega)$ and*

$$\partial_i(uv) = (\partial_i u)v + u(\partial_i v).$$

BEWEIS: For any $\varphi \in C^\infty_c(\Omega)$ we calculate

$$\begin{aligned}
\int_\Omega uv\, \partial_i\varphi &= \lim_{\varrho \searrow 0} \int u(\eta_\varrho * v)\, \partial_i\varphi \\
&= \lim_{\varrho \searrow 0} \int u\big[\partial_i\big((\eta_\varrho * v)\varphi\big) - \partial_i(\eta_\varrho * v)\varphi\big] \\
&= \lim_{\varrho \searrow 0} \Big(- \int \big[(\partial_i u)\, \eta_\varrho * v + u(\eta_\varrho * \partial_i v)\big]\varphi \Big) \\
&= -\int \big((\partial_i u)v + u(\partial_i v)\big)\varphi.
\end{aligned}$$

□

Lemma 11.3 *For $1 \leq p < \infty$ the space $C^\infty_c(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$.*

BEWEIS: It suffices to approximate $u \in W^{1,p}(\mathbb{R}^n)$ by functions in $W^{1,p}(\mathbb{R}^n)$ having compact support, then mollification yields the desired approximation. Consider the cutoff function

$$\varphi(x) = \begin{cases} 1 & \text{for } |x| \leq \frac{1}{2}, \\ 0 & \text{for } |x| \geq 1. \end{cases}$$

Put $\varphi_R(x) = \varphi\big(\frac{x}{R}\big)$, hence $\partial_i\varphi_R(x) = \frac{1}{R}\partial_i\varphi\big(\frac{x}{R}\big)$. For $u_R = \varphi_R\, u$ the product rule yields

$$\partial_i u_R(x) = \varphi_R(x)\, \partial_i u(x) + \frac{1}{R}\partial_i\varphi\big(\frac{x}{R}\big)\, u(x).$$

Taking norms we obtain by dominated convergence for $R \nearrow \infty$

$$\begin{aligned}
\|u - u_R\|_{L^p(\mathbb{R}^n)} &\leq \|(1-\varphi_R)u\|_{L^p(\mathbb{R}^n)} \to 0, \\
\|\partial_i u - \partial_i u_R\|_{L^p(\mathbb{R}^n)} &\leq \|(1-\varphi_R)\, \partial_i u\|_{L^p(\mathbb{R}^n)} + \frac{C}{R}\, \|u\|_{L^p(\mathbb{R}^n)} \to 0.
\end{aligned}$$

□

For a bounded domain $\Omega \subset \mathbb{R}^n$ the space $C^1_c(\Omega)$ is not dense in $W^{1,p}(\Omega)$. The closure is the space $W^{1,p}_0(\Omega)$ of $W^{1,p}$-functions which have boundary value zero, in an appropriate sense.

Theorem 11.3 (chain rule) *Let $u \in W^{1,1}_{\mathrm{loc}}(\Omega)$ and $f \in C^1(\mathbb{R})$ with f' bounded. Then*

$$D(f \circ u) = (f' \circ u) Du \in L^1_{\mathrm{loc}}(\Omega).$$

Beweis: Let $\varphi \in C^\infty_c(\Omega)$. The mollification $u^\varrho = \eta_\varrho * u$ is defined and smooth on the support of φ for small $\varrho > 0$, and we have

$$-\int (f \circ u^\varrho)\, \partial_i \varphi = \int (f' \circ u^\varrho)\, (\partial_i u^\varrho)\, \varphi.$$

We want to take the limit $\varrho \searrow 0$ on both sides. Clearly $u^\varrho \to u$ in $W^{1,1}_{\mathrm{loc}}(\Omega)$. Assuming $|f'| \leq L$ and hence $\mathrm{Lip}(f) \leq L$ we estimate for the left hand side

$$\Big| \int (f \circ u^\varrho) \partial_i \varphi - \int (f \circ u) \partial_i \varphi \Big| \leq L \, \|D\varphi\|_{C^0(\Omega)} \int_{\mathrm{spt}\, \varphi} |u^\varrho - u| \to 0.$$

For the right hand side we use the triangle inequality to get

$$\begin{aligned} \Big| \int (f' \circ u^\varrho)(\partial_i u^\varrho)\, \varphi - \int (f' \circ u)(\partial_i u)\, \varphi \Big| \;&\leq \\ \int \underbrace{|f' \circ u^\varrho - f' \circ u|}_{|\ldots| \leq 2L} \, \underbrace{|Du|\, |\varphi|}_{\in L^1} \;&+\; \int \underbrace{|f' \circ u^\varrho|}_{\leq L} |Du - Du^\varrho|\, \varphi. \end{aligned}$$

The first integral goes to zero by dominated convergence, at least for a subsequence $\varrho_i \searrow 0$. The chain rule follows. $\square$

Theorem 11.4 (transformation rule) *Let $\phi \in C^1(U, V)$ be a diffeomorphism. Then for $v \in W^{1,1}_{\mathrm{loc}}(V)$ we have $v \circ \phi \in W^{1,1}_{\mathrm{loc}}(U)$ and*

$$D(v \circ \phi) = (Dv) \circ \phi \, D\phi.$$

Beweis: We first recall that ϕ maps null sets to null sets, therefore $v \circ \phi$ is well-defined almost everywhere, likewise $(Dv) \circ \phi$. If $K \subset U$ is compact, then $\phi(K) \subset V$ is also compact and

$$\int_K |(v \circ \phi)(x)|\, dx = \int_{\phi(K)} |v(y)|\, |\det D\psi(y)|\, dy < \infty \qquad \text{where } \psi = \phi^{-1} \in C^1(V, U).$$

Thus $v \circ \phi$ and $(Dv) \circ \phi$ belong to $L^1_{\mathrm{loc}}(U)$. Let $\eta \in C^\infty_c(U)$. Then for v^ϱ with $\varrho > 0$ small

$$\int_U (v^\varrho \circ \phi)\, \partial_i \eta = - \int_U (Dv^\varrho) \circ \phi \, \partial_i \phi \, \eta.$$

Using transformation we estimate as $\varrho \searrow 0$

$$\begin{aligned} \int_U \big| (v^\varrho \circ \phi - v \circ \phi) \partial_i \eta \big|\, dx \;&\leq\; \int_{\phi(\mathrm{spt}\, \eta)} |v^\varrho - v|\, \underbrace{|(\partial_i \eta) \circ \psi|\, |\det D\psi|}_{\leq C}\, dy \to 0, \\ \int_U \big| (Dv^\varrho) \circ \phi - Dv \circ \phi)(\partial_i \phi)\, \eta \big|\, dx \;&\leq\; \int_{\phi(\mathrm{spt}\, \eta)} |Dv^\varrho - Dv|\, \underbrace{|(\partial_i \phi) \circ \psi|\, |\eta \circ \psi|\, |\det D\psi|}_{\leq C}\, dy \to 0. \end{aligned}$$

We conclude

$$\int_U (v \circ \phi)\, \partial_i \eta \, dx = -\int_U Dv \circ \phi \, \partial_i \phi \, \eta.$$

□

We now come to embedding theorems, which are stated for functions in $W^{1,p}(\mathbb{R}^n)$. Sobolev embeddings for $W^{1,p}(\Omega)$ follow by suitably extending functions on Ω to all of $\mathbb{R}^n$. We start with the very simple one-dimensional case.

Lemma 11.4 *Any $u \in W^{1,1}(\mathbb{R})$ has a continuous representative satisfying*

$$\|u\|_{C^0(\mathbb{R})} \leq \|u'\|_{L^1(\mathbb{R})}. \tag{11.2}$$

Moreover the fundamental theorem of calculus holds, that is

$$u(y) = \int_{-\infty}^{y} u'(x)\, dx \quad \text{for all } y \in \mathbb{R}.$$

Proof. For $u \in C^1_c(\mathbb{R})$ we have

$$|u(x)| = \Big| \int_{-\infty}^{x} u'(\xi)\, d\xi \Big| \leq \|u'\|_{L^1(\mathbb{R})}.$$

By Lemma 11.3 any $u \in W^{1,1}(\mathbb{R})$ can be approximated by functions $u_k \in C^1_c(\mathbb{R})$, and then

$$\|u_k - u_\ell\|_{C^0(\mathbb{R})} \leq \|u_k' - u_\ell'\|_{L^1(\mathbb{R})} \to 0 \quad \text{as } k, \ell \to \infty.$$

Therefore u_k converges uniformly to a representative $u \in C^0(\mathbb{R})$, which satisfies estimate (11.2). The second claim also follows by passing to the limit. □

We now turn to the case $n \geq 2$. Our goal is an estimate

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C\, \|Du\|_{L^p(\mathbb{R}^n)} \text{ for all } u \in W^{1,p}(\mathbb{R}^n) \quad \text{where } C = C(n,p). \tag{11.3}$$

Using scaling, one finds that q is is uniquely determined. Namely let $u_\lambda(x) = u(\lambda x)$ for $\lambda > 0$. Then $Du_\lambda(x) = \lambda Du(\lambda x)$, and we compute

$$\|u_\lambda\|_{L^q(\mathbb{R}^n)} = \lambda^{-\frac{n}{q}} \|u\|_{L^q(\mathbb{R}^n)} \quad \text{and} \quad \|Du_\lambda\|_{L^p(\mathbb{R}^n)} = \lambda^{1-\frac{n}{p}} \|Du\|_{L^p(\mathbb{R}^n)}.$$

To have (11.3) we need the same scaling. This is only possible for $p < n$, then

$$-\frac{n}{q} = 1 - \frac{n}{p} \quad \Leftrightarrow \quad q = \frac{np}{n-p}.$$

From the viewpoint of scaling, the critical case $p = n$ would be allowed with $q = \infty$. However, the estimate (11.3) fails unless when $n = 1$.

Theorem 11.5 (Sobolev) *Let $1 \leq p < n$. Then for any $u \in W^{1,p}(\mathbb{R}^n)$ we have*

$$\|u\|_{L^q(\mathbb{R}^n)} \leq \frac{(n-1)p}{n-p} \|Du\|_{L^p(\mathbb{R}^n)} \quad \text{where } q = \frac{np}{n-p}. \tag{11.4}$$

Beweis: We follow Gagliardo-Nirenberg, assuming $u \in C_c^1(\mathbb{R}^n)$ without loss of generality.

Step 1. *Reduction to the case $p = 1$.*
For $1 < p < n$ we estimate, using the case $p = 1$ and Theorem 11.3

$$\begin{aligned}
\Big[\int |u|^{\frac{np}{n-p}}\Big]^{\frac{n-1}{n}} &= \Big[\int \big(|u|^{\frac{(n-1)p}{n-p}}\big)^{\frac{n}{n-1}}\Big]^{\frac{n-1}{n}} \\
&\leq \int \big|D(|u|^{\frac{(n-1)p}{n-p}})\big| \\
&= \frac{(n-1)p}{n-p}\int |u|^{\frac{n(p-1)}{n-p}}\,|Du| \\
&\leq \frac{(n-1)p}{n-p}\Big(\int |u|^{\frac{np}{n-p}}\Big)^{\frac{p-1}{p}}\Big(\int |Du|^p\Big)^{\frac{1}{p}}.
\end{aligned}$$

By cancelling we obtain the claimed inequality (14.48).

Step 2. *Proof for $p = 1$ by induction on n.*
Write $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$, $x = (\xi, z)$, and define

$$f(\xi) = \int_{-\infty}^{\infty} |u(\xi, z)|\,dz \quad \text{and} \quad g(\xi) = \int_{-\infty}^{\infty} |Du(\xi, z)|\,dz.$$

We claim that f has a weak derivative satiyfying $|Df(\xi)| \leq g(\xi)$ almost everywhere. Let us first show that then the estimate follows. By Lemma 11.4, which may be viewed formally as the start of the induction with $n = p = 1$, we obtain

$$\begin{aligned}
\int_{\mathbb{R}^n} |u(x)|^{\frac{n}{n-1}}\,dx &\leq \int_{\mathbb{R}^{n-1}}\int_{\mathbb{R}} |u(\xi,z)|\,\Big(\underbrace{\int_{\mathbb{R}} \Big|\frac{\partial u}{\partial z}(\xi,s)\Big|\,ds}_{\leq g(\xi)}\Big)^{\frac{1}{n-1}}\,dz\,d\xi \\
&\leq \int_{\mathbb{R}^{n-1}} f(\xi)\,g(\xi)^{\frac{1}{n-1}}\,d\xi \\
&\leq \begin{cases} \Big(\int_{\mathbb{R}^{n-1}} f(\xi)^{\frac{n-1}{n-2}}\,d\xi\Big)^{\frac{n-2}{n-1}}\Big(\int_{\mathbb{R}^{n-1}} g(\xi)\,d\xi\Big)^{\frac{1}{n-1}} & \text{for } n \geq 3, \\ \|f\|_{C^0(\mathbb{R})}\int_{\mathbb{R}} g(\xi)\,d\xi & \text{for } n = 2 \end{cases} \\
&\leq \Big(\int_{\mathbb{R}^{n-1}} \underbrace{|Df(\xi)|}_{\leq g(\xi)}\,d\xi\Big)\cdot\Big(\int_{\mathbb{R}^{n-1}} g(\xi)\,d\xi\Big)^{\frac{1}{n-1}} \\
&\leq \Big(\int_{\mathbb{R}^n} |Du(x)|\,dx\Big)^{\frac{n}{n-1}}.
\end{aligned}$$

It remains to check the claim. For $\eta \in C_c^\infty(\mathbb{R}^{n-1})$ we compute for $i = 1, \ldots, n-1$

$$\begin{aligned}
\int_{\mathbb{R}^{n-1}} f(\xi)\,\partial_i\eta(\xi)\,d\xi &= \int_{\mathbb{R}}\int_{R^{n-1}} |u(x,\xi)|\,\partial_i\eta(\xi)\,d\xi\,dz \\
&= -\int_{R^{n-1}}\int_{\mathbb{R}} \operatorname{sign} u(\xi,z)\,\partial_i u(\xi,z)\,\eta(\xi)\,d\xi\,dz.
\end{aligned}$$

This shows that f has weak derivatives

$$\partial_i f(\xi) = \int_{\mathbb{R}} \operatorname{sign} u(\xi,z)\,\partial_i u(\xi,z)\,dz.$$

An analogous formula holds when $\partial_i f$ is replaced by the derivative in direction of any unit vector $v \in \mathbb{R}^{n-1}$. We conclude

$$|Df(\xi)| \leq \int_{\mathbb{R}} |Du(\xi, z)|\, dz.$$

This completes the proof of the theorem. □

Remark. Let $\Omega \subset \mathbb{R}^n$ be a bounded C^2 domain. We apply the Sobolev inequality with

$$u(x) = \big(1 - \frac{1}{\varepsilon}\mathrm{dist}\,(x, \Omega)\big)_+, \quad \text{thus } |Du(x)| = \frac{1}{\varepsilon}\,\chi_{\{0 < \mathrm{dist}(x,\Omega) < \varepsilon\}}.$$

This yields

$$|\Omega|^{\frac{n-1}{n}} \leq \Big(\int_{\mathbb{R}^n} |u(x)|^{\frac{n}{n-1}}\, dx\Big)^{\frac{n-1}{n}} \leq \int_{\mathbb{R}^n} |Du(x)|\, dx = \frac{1}{\varepsilon}\mathcal{L}^n\big(\{0 < \mathrm{dist}\,(x, \Omega) < \varepsilon\}\big).$$

Letting $\varepsilon \searrow 0$ we get the isoperimetric inequality

$$|\Omega|^{\frac{n-1}{n}} \leq \mathcal{H}^{n-1}(\partial\Omega).$$

Reversely, the Sobolev inequality can be derived from the isoperimetric inequality. This allows to obtain the optimal constant in the Sobolev estimate, see for instance the book of Schoen-Yau (Lectures on Differential geometry, International Press 2010).

Lemma 11.5 *Let $u \in W^{1,p}(\mathbb{R}^n)$ where $1 \leq p < \infty$. Then*

$$\begin{array}{rcll} \|u \circ \tau_h - u\|_{L^p(\mathbb{R}^n)} & \leq & |h|\, \|Du\|_{L^p(\mathbb{R}^n)} & \text{for } h \in \mathbb{R}^n,\ \tau_h(x) = x + h, \\ \|\eta_\varrho * u - u\|_{L^p(\mathbb{R}^n)} & \leq & \varrho\, \|Du\|_{L^p(\mathbb{R}^n)} & \text{for } \varrho > 0. \end{array}$$

Beweis: We can assume $u \in C_c^\infty(\mathbb{R}^n)$, see Lemma 11.3. For the first statement we compute

$$\begin{aligned} \int_{\mathbb{R}^n} |u(x+h) - u(x)|^p\, dx &= \int_{\mathbb{R}^n} \Big| \int_0^1 \frac{d}{dt} u(x + th)\, dt\Big|^p\, dx \\ &\leq \int_0^1 \int_{\mathbb{R}^n} |Du(x + th)|^p\, |h|^p\, dx\, dt \\ &= \|Du\|^p_{L^p(\mathbb{R}^n)}\, |h|^p. \end{aligned}$$

Using this we deduce further

$$\begin{aligned} \int_{\mathbb{R}^n} |\eta_\varrho * u(x) - u(x)|^p\, dx &= \int_{\mathbb{R}^n} \Big| \int_{\mathbb{R}^n} \eta_\varrho(x - y)\big(u(y) - u(x)\big)\, dy\Big|^p\, dx \\ &= \int_{\mathbb{R}^n} \Big| \int_{\mathbb{R}^n} \eta(z)\big(u(x - \varrho z) - u(x)\big)\, dz\Big|^p\, dx \\ &\leq \int_{\mathbb{R}^n} \eta(z) \int_{\mathbb{R}^n} |u(x - \varrho z) - u(x)|^p\, dx\, dz \\ &\leq \varrho^p\, \|Du\|^p_{L^p(\mathbb{R}^n)}. \end{aligned}$$

□

Any continuous function on a bounded, closed subset of $\mathbb{R}^n$ attains its minimum. The proof of this fact relies on the Bolzano-Weierstraß theorem, which guarantees that a minimizing sequence has a convergent subsequence. Unfortunately, this argument fails in Banach spaces of infinite dimension. Inspecting the statement, one realizes that the norm of the space enters twice, first in the word *bounded* and second in the word *convergent*. One comes to the idea that the statement might become true if one uses a weaker notion of convergence in the second.

Theorem 11.6 (Rellich) *Let $u_k \in W^{1,p}(\mathbb{R}^n)$, $1 \leq p < \infty$, be a sequence with*

$$\|u_k\|_{W^{1,p}(\mathbb{R}^n)} \leq M < \infty \quad \text{for all } k.$$

Then there exists a subsequence u_{k_j} with $u_{k_j} \to u$ locally in $L^p(\mathbb{R}^n)$. Moreover then

$$\|u\|_{L^p(\mathbb{R}^n)} \leq \liminf_{j\to\infty} \|u_{k_j}\|_{L^p(\mathbb{R}^n)}.$$

In the statement the word *locally* cannot be dropped, as seen by a sequence of translations of a fixed function (*the wandering hill*).

Beweis: We prove that for $\varrho > 0$ the sequence $\eta_\varrho * u_k$ is uniformly bounded and equicontinuous. We may then use the Arzelá-Ascoli theorem to obtain a sublimit u^ϱ, for any fixed $\varrho > 0$. Recalling Theorem 11.2(1) we have

$$D^\alpha(\eta_\varrho * u_k)(x) = \varrho^{-|\alpha|} \int_{\mathbb{R}^n} \varrho^{-n} D^\alpha \eta\Big(\frac{x-y}{\varrho}\Big)\, u_k(y)\, dy.$$

We estimate

$$\begin{aligned} |D^\alpha(\eta_\varrho * u_k)(x)| &\leq C(\alpha)\varrho^{-n-|\alpha|}\, \|u_k\|_{L^1(B_\varrho(x))} \\ &\leq C(n,p,\alpha)\varrho^{-|\alpha|-\frac{n}{p}}\, \|u_k\|_{L^p(\mathbb{R}^n)} \\ &\leq C(n,p,\alpha)\varrho^{-|\alpha|-\frac{n}{p}} M. \end{aligned}$$

For fixed $\varrho > 0$ Arzelá-Ascoli yields a subsequence u_{k_j} with $\eta_\varrho * u_{k_j} \to u^\varrho$ locally in $C^1(\mathbb{R}^n)$. For a sequence $\varrho_i \searrow 0$ choose successive, locally C^1-convergent subsequences

$$\{k_j^{(1)}\} \supset \{k_j^{(2)}\} \supset \ldots \quad \text{for } \varrho_1, \varrho_2, \ldots.$$

Then for the diagonal sequence $\{k_j^{(j)}\}$ we have, after renumbering,

$$\eta_\varrho * u_k \to u^\varrho \quad \text{for all } \varrho = \varrho_1, \varrho_2, \ldots.$$

Furthermore

$$\begin{aligned} \|u^\varrho\|_{L^p(\mathbb{R}^n)} &\leq \liminf_{k\to\infty} \|\eta_\varrho * u_k\|_{L^p(\mathbb{R}^n)} \leq \|u_k\|_{L^p(\mathbb{R}^n)} \leq M, \\ \|Du^\varrho\|_{L^p(\mathbb{R}^n)} &\leq \liminf_{k\to\infty} \|D(\eta_\varrho * u_k)\|_{L^p(\mathbb{R}^n)} \leq \|Du_k\|_{L^p(\mathbb{R}^n)} \leq M. \end{aligned}$$

Now let $0 < \varrho' \leq \varrho$ where $\varrho, \varrho' \in \{\varrho_1, \varrho_2, \ldots\}$. Then

$$\begin{aligned} \|u^\varrho - u^{\varrho'}\|_{L^p(\mathbb{R}^n)} &\leq \liminf_{k\to\infty} \|\eta_\varrho * u_k - \eta_{\varrho'} * u_k\|_{L^p(\mathbb{R}^n)} \\ &\leq \liminf_{k\to\infty} \Big(\|\eta_\varrho * u_k - u_k\|_{L^p(\mathbb{R}^n)} + \|\eta_{\varrho'} * u_k - u_k\|_{L^p(\mathbb{R}^n)}\Big) \\ &\leq 2\varrho \liminf_{k\to\infty} \|Du_k\|_{L^p(\mathbb{R}^n)} \\ &\leq 2M\varrho. \end{aligned}$$

Thus for $\varrho_i \searrow 0$, the sequence u^{ϱ_i} converges by Fischer-Riesz in $L^p(\mathbb{R}^n)$ to some function $u \in L^p(\mathbb{R}^n)$, moreover by letting $\varrho' \searrow 0$

$$\|u^\varrho - u\|_{L^p(\mathbb{R}^n)} \leq 2M\varrho.$$

For any compact set $K \subset \mathbb{R}^n$ we have, again for $\varrho \in \{\varrho_1, \varrho_2, \ldots\}$,

$$\begin{aligned}
\|u - u_k\|_{L^p(K)} &\leq \|u - u^\varrho\|_{L^p(K)} + \|u^\varrho - \eta_\varrho * u_k\|_{L^p(K)} + \|\eta_\varrho * u_k - u_k\|_{L^p(K)} \\
&\leq 2M\varrho + \|u^\varrho - \eta_\varrho * u_k\|_{L^p(K)} + M\varrho.
\end{aligned}$$

Letting $k \to \infty$ we conclude

$$\liminf_{k \to \infty} \|u - u_k\|_{L^p(K)} \leq 3M\varrho \searrow 0 \text{ as } \varrho = \varrho_i \searrow 0.$$

The theorem is proved. $\Box$

Let $1 \leq p < n$. In view of the Sobolev theorem, one may ask whether the sequence u_k in Theorem 11.6 converges also (locally) in $L^r(\mathbb{R}^n)$ for $p \leq r \leq \frac{np}{n-p}$. For this we need the following lemma.

Lemma 11.6 (interpolation) *Let $1 \leq p < r < q \leq \infty$. Then for $f \in (L^p \cap L^q)(\mu)$*

$$\|f\|_{L^r(\mu)} \leq \|f\|^{1-\alpha}_{L^p(\mu)} \|f\|^\alpha_{L^q(\mu)} \qquad \textit{where } \alpha = \frac{\frac{1}{p} - \frac{1}{r}}{\frac{1}{p} - \frac{1}{q}} \in (0,1).$$

BEWEIS: For $\alpha \in (0,1)$ we apply Hölder with exponents $\frac{p}{(1-\alpha)r}$ and $\frac{q}{\alpha r}$. This gives

$$\int |f|^r \, d\mu = \int |f|^{(1-\alpha)r} |f|^{\alpha r} \, d\mu \leq \Big(\int |f|^p \, d\mu \Big)^{\frac{(1-\alpha)r}{p}} \Big(\int |f|^q \, d\mu \Big)^{\frac{\alpha r}{q}} = \|f\|^{(1-\alpha)r}_{L^p(\mu)} \|f\|^{\alpha r}_{L^q(\mu)}.$$

We need that the two exponents are conjugate, in fact by the choice of α we have

$$\frac{(1-\alpha)r}{p} + \frac{\alpha r}{q} = r\Big(\frac{1}{p} - \alpha\Big(\frac{1}{p} - \frac{1}{q}\Big)\Big) = 1.$$

$\Box$

In the situation above we have $p < r < q = \frac{np}{n-p}$. For K compact we estimate using Sobolev

$$\begin{aligned}
\|u - u_k\|_{L^r(K)} &\leq \|u - u_k\|^{(1-\alpha)}_{L^p(K)} \|u - u_k\|^\alpha_{L^q(\mathbb{R}^n} \\
&\leq C(n,p,r,M) \, \|u - u_k\|^{(1-\alpha)}_{L^p(K)} \to 0 \quad \text{ for } k \to \infty.
\end{aligned}$$

Note that $\alpha = n\big(\frac{1}{p} - \frac{1}{r}\big)$ goes to one if $r \nearrow q$. In fact, the functions u_k do not converge locally in the critical space $L^q(\mathbb{R}^n)$ in general (exercise).

One of the basic concepts of functional analysis is weak convergence, it appears in many compactness arguments. As an example, consider the sequence space $L^p(\mathbb{N})$ with norm

$$\|x\|_{L^p(\mathbb{N})} = \Big(\sum_{i=1}^\infty |x^i|^p \Big)^{\frac{1}{p}}.$$

In general, bounded sequences in a Banach space do not have convergent subsequences, e.g.

$$x_k = (0,\ldots,0,\overset{k}{1},0,\ldots) \quad \text{where } \|x_k - x_m\|_{L^p(\mathbb{N})} = 2^{\frac{1}{p}} \text{ for } k \neq m.$$

But for a sequence $x_k \in L^p(\mathbb{N})$ with $\|x_k\|_{L^p(\mathbb{N})} \leq C$ we have $|x_k^i| \leq \|x_k\|_{L^p(\mathbb{N})} \leq C$. Therefore by Bolzano-Weierstraß and a diagonal argument we can pass to a subsequence such that

$$x_k^i \to x^i \text{ as } k \to \infty, \quad \text{for any } i \in \mathbb{N}.$$

In other words the sequence converges componentwise. Moreover

$$\Big(\sum_{i=1}^N |x^i|^p\Big)^{\frac{1}{p}} = \lim_{k\to\infty} \Big(\sum_{i=1}^N |x_k^i|^p\Big)^{\frac{1}{p}} \leq \liminf_{k\to\infty} \|x_k\|_{L^p(\mathbb{N})} \leq C.$$

Letting $N \to \infty$ we infer the lower semicontinuity

$$\|x\|_{L^p(\mathbb{N})} \leq \liminf_{k\to\infty} \|x_k\|_{L^p(\mathbb{N})},$$

in particular $x \in L^p(\mathbb{N})$. In the following definition of weak convergence, the role of the coordinates is taken by the linear functionals.

Definition 11.3 (weak convergence) *Let X be a Banach space. Then*

$$\begin{aligned} x_k \to x \text{ weak in } X \quad &\Leftrightarrow \quad \varphi(x_k) \to \varphi(x) \quad \text{for all } \varphi \in X', \\ \varphi_k \to \varphi \text{ weak}^* \text{ in } X' \quad &\Leftrightarrow \quad \varphi_k(x) \to \varphi(x) \quad \text{for all } x \in X. \end{aligned}$$

The key compactness result for weak convergence is the following.

Theorem 11.7 (weak*-compactness) *Let X be a separable Banach space. Then any bounded sequence $(\varphi_k)_{k\in\mathbb{N}}$ in X' has a subsequence which converges weak* to some $\varphi \in X'$.*

Beweis: We have $C := \sup_{k\in\mathbb{N}} \|\varphi_k\| < \infty$, and there is a dense set $\{x_n : n \in \mathbb{N}\} \subset X$. Thus

$$|\varphi_k(x_n)| \leq C\|x_n\| \quad \text{for all } k, n \in \mathbb{N}.$$

Using a diagonal process, we have after renumbering

$$\exists \lim_{k\to\infty} \varphi_k(x_n) \quad \text{for all } n \in \mathbb{N}.$$

Now let X_0 be the span of the set $\{x_n : n \in \mathbb{N}\}$. Each $x \in X_0$ has a representation

$$x = \sum_{n=1}^\infty \alpha_n x_n \quad \text{where } \alpha_n = 0 \text{ for all but finitely many } n \in \mathbb{N}.$$

Define the map $\varphi : X_0 \to \mathbb{R}$, $\varphi_0(x) = \lim_{k\to\infty} \varphi_k(x)$. The limit exists and φ is linear. Moreover

$$|\varphi(x)| = \lim_{k\to\infty} |\varphi_k(x)| \leq C\|x\| \quad \text{for any } x \in X_0.$$

In particular φ is Lipschitz on the dense subspace X_0, and hence extends to a functional $\varphi \in X'$ with $\|\varphi\| \leq C$. Let $x \in X$ be given. Then for any $x_0 \in X_0$

$$|\varphi_k(x) - \varphi(x)| \leq |\varphi_k(x) - \varphi_k(x_0)| + |\varphi_k(x_0) - \varphi(x_0)| + |\varphi(x_0) - \varphi(x)|.$$

Letting $k \to \infty$ we infer

$$\limsup_{k\to\infty} |\varphi_k(x) - \varphi(x)| \leq 2C\,\|x - x_0\|.$$

By density of X_0 we conclude $\varphi_k(x) \to \varphi(x)$, which shows $\varphi_k \to \varphi$ weak* in X'. $\square$

The Banach-Alaoglu theorem asserts that for any Banach space X the unit ball in X' is compact for the weak* topology, in the sense of the Heine-Borel covering property. The proof of this result is non-constructive, one needs Tychonov's theorem and hence the axiom of choice. But the two notions of compactness coincide in the case when X is separable. It is the above simple sequential version we are interested in.

Now assume we are given a bounded sequence in a Banach space Y. To apply the theorem we need to identify Y as the dual space X' of some separable Banach space X, or at least Y being isomorphic to that dual space. We address this question first in the fundamental case $Y = L^p(\mu)$. In the following we always assume that $1 \leq p, q \leq \infty$ are conjugate, i.e. $\frac{1}{p} + \frac{1}{q} = 1$. Consider

$$J : L^p(\mu) \to L^q(\mu)',\; Ju(v) = \int uv\,d\mu.$$

The map is well-defined by Hölder's inequality, in fact we have

$$\Big|\int uv\,d\mu\Big| \leq \|u\|_{L^p(\mu)}\|v\|_{L^q(\mu)}, \qquad \text{thus } \|Ju\|_{L^q(\mu)'} \leq \|u\|_{L^p(\mu)}.$$

Lemma 11.7 *J is an isometric embedding (here for $p = \infty$ we assume that μ is σ-finite).*

Beweis: We assume $1 \leq p < \infty$. For a given $u \in L^p(\mu)$ with $\|u\|_{L^p(\mu)} = 1$, consider

$$u^* = \begin{cases} |u|^{p-2}u & \text{if } u \neq 0,\\ 0 & \text{if } u = 0.\end{cases}$$

We compute

$$\|u^*\|_{L^q(\mu)} = \begin{cases} \|u\|^{p-1}_{L^p(\mu)} & = 1 \quad \text{for } 1 < p < \infty,\\ \|\chi_{\{u \neq 0\}}\|_{L^\infty(\mu)} & = 1 \quad \text{for } p = 1.\end{cases}$$

Moreover we have

$$Ju(u^*) = \int |u|^p\,d\mu = 1.$$

Hence $\|Ju\|_{L^q(\mu)'} \geq 1 = \|u\|_{L^p(\mu)}$. The case $p = \infty$ is left to the reader. $\square$

Theorem 11.8 (Riesz L^p-duality) *Let $1 < p \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then the following hold:*

(1) *$J : L^p(\mu) \to L^q(\mu)'$ is an isometry, i.e. surjective (for $p = \infty$, we assume μ σ-finite).*

(2) *If μ is a Radon measure on a compact metric space, then $L^q(\mu)$ is separable.*

BEWEIS: We indicate the proof of (1) in the case $p < \infty$. Let $\Lambda \in L^q(\mu)'$ be given, w.l.o.g. with norm $\|\Lambda\| = 1$. Assume there exists $v \in L^q(\mu)$ with

$$\|v\|_{L^q(\mu)} = 1 \quad \text{and} \quad \Lambda(v) = 1.$$

Let $v_* \in L^p(\mu)$, $\|v_*\|_{L^p(\mu)} = 1$, be defined by

$$v_* = \begin{cases} |v|^{q-2}v & \text{for } v \neq 0, \\ 0 & \text{for } v = 0. \end{cases}$$

We claim that $J(v_*) = \Lambda$. For any $\varphi \in L^q(\mu)$ and $t \in \mathbb{R}$ small we know that

$$\Lambda\Big(\frac{v + t\varphi}{\|v + t\varphi\|_{L^q(\mu)}}\Big) \leq 1, \quad \text{with equality for } t = 0.$$

We compute the derivative at $t = 0$. First we have for all t

$$\frac{\partial}{\partial t}|v + t\varphi|^q = \begin{cases} q|v + t\varphi|^{q-2}(v + t\varphi)\varphi & \text{if } v + t\varphi \neq 0, \\ 0 & \text{if } v + t\varphi = 0. \end{cases}$$

In particular for $|t| \leq 1$ we can estimate

$$\Big\|\frac{\partial}{\partial t}|v + t\varphi|^q\Big\| \leq C(|v|^q + |\varphi|^q) \in L^1(\mu).$$

Therefore we can differentiate under the integral to get

$$\frac{d}{dt}\|v + t\varphi\|_{L^q(\mu)}|_{t=0} = \frac{1}{q}\frac{d}{dt}\int |v + t\varphi|^q \, d\mu|_{t=0} = \int |v|^{q-2}v\varphi \, d\mu = Jv_*(\varphi).$$

Now by the maximizing property we conclude

$$0 = \frac{d}{dt}\Lambda\Big(\frac{v + t\varphi}{\|v + t\varphi\|_{L^q(\mu)}}\Big)|_{t=0} = \Lambda(\varphi) - Jv_*(\varphi)\Lambda(v) = \Lambda(\varphi) - Jv_*(\varphi).$$

This proves our claim. It remains to show the existence of a ϕ maximizer in the $L^q(\mu)$ unit sphere. Unfortunately this is technically involved, one proves uniform convexity of the L^q unit ball, for $1 < q < \infty$, using an inequality of Clarkson (see for instance Hirzebruch-Scharlau, Einführung in die Funktionalanalysis, B.I. Hochschultaschenbuch, 1971). We remark that the case $p = \infty$ then follows by a limit consideration (exercise). $\square$

Corollary 11.1 (weak*-compactness in $L^p(\mu)$) *In the situation of Theorem 11.8, let u_k be a bounded sequence in $L^p(\mu)$. Then there exists a function $u \in L^p(\mu)$ such that after passing to a subsequence*

$$\int uv \, d\mu = \lim_{k \to \infty} \int u_k v \, d\mu \quad \textit{for all } v \in L^q(\mu).$$

Moreover $\|u\|_{L^p(\mu)} \leq \liminf_{k \to \infty} \|u_k\|_{L^p(\mu)}$.

The second statement follows from the first, namely by Hölder we have for $\|v\|_{L^q(\mu)} \leq 1$

$$Ju(v) = \int uv \, d\mu = \lim_{k \to \infty} \int u_k v \, d\mu \leq \liminf_{k \to \infty} \|u_k\|_{L^p(\mu)}.$$

Taking the supremum with respect to such v we get by Lemma 11.7

$$\|u\|_{L^p(\mu)} = \|Ju\|_{L^q(\mu)'} \leq \liminf_{k \to \infty} \|u_k\|_{L^p(\mu)}.$$

Theorem 11.9 (Riesz-Markow) *Let (X,d) be a compact metric space. Then for any non-negative functional $\Lambda \in C^0(X)'$ there exists a unique Radon measure μ such that*

$$\Lambda(\phi) = \int_X \phi \, d\mu \quad \text{for all } \phi \in C^0(X).$$

Here nonnegative means that $\Lambda(\phi) \geq 0$ for all $\phi \geq 0$. The identification of μ with Λ is isometric if we put $\|\mu\| = \mu(X)$.

Corollary 11.2 (weak*-compactness for Radon measures) *In the situation of Theorem 11.9, let μ_k be a sequence of Radon measures with $\mu_k(X)$ bounded. Then there exists a Radon measure μ such that after passing to a subsequence*

$$\int \phi \, d\mu = \lim_{k\to\infty} \int \phi \, d\mu_k \quad \text{for all } \phi \in C^0(X).$$

Note that we can take $\phi \equiv 1$ to get $\mu(X) = \lim_{k\to\infty} \mu_k(X)$.

Theorem 11.10 (Riesz representation in Hilbert space) *Let X be a Hilbert space on $\mathbb{R}$. Then for any $\phi \in X'$ there exists a unique $x_0 \in X$ such that*

$$\phi(x) = \langle x, x_0 \rangle \quad \text{for all } x \in X. \tag{11.5}$$

Moroever $\|x_0\| = \|\phi\|$, hence the Riesz map $\mathcal{R} : X \to X'$, $\mathcal{R}x_0(x) = \langle x, x_0\rangle$, is an isometry.

Remark. *x_0 is the unique minimizer of the functional $Q_\phi(x) = \frac{1}{2}\|x\|^2 - \phi(x)$.*

Beweis: We prove the following statements:

(a) Any minimizer x_0 of Q_ϕ solves (11.5).

(b) There exists a minimizer x_0 of Q_ϕ.

(c) Any solution x_0 of (11.5) satisfies $\|x_0\| = \|\phi\|$.

(a): Let x_0 be a minimizer of Q_ϕ. For any $x \in X$ and $t \in \mathbb{R}$ we compute

$$Q_\phi(x_0 + tx) = Q_\phi(x_0) + t\big(\langle x, x_0\rangle - \phi(x)\big) + \frac{t^2}{2}\|x\|^2.$$

By assumption, the function has a minimum at $t = 0$.

(b): Q_ϕ is bounded below, more precisely

$$Q_\phi(x) \geq \frac{1}{2}\|x\|^2 - \|\phi\|\,\|x\| = \frac{1}{2}(\|x\| - \|\phi\|)^2 - \frac{1}{2}\|\phi\|^2 \geq -\frac{1}{2}\|\phi\|^2.$$

Put $\lambda = \inf_{x\in X} Q_\phi(x) > -\infty$. For a minimizing sequence $x_k \in X$ we write

$$\begin{aligned}
Q_\phi(x_k) &= \frac{1}{2}\|x_k\|^2 - \phi(x_k),\\
Q_\phi(x_l) &= \frac{1}{2}\|x_l\|^2 - \phi(x_l),\\
2\,Q_\phi\Big(\frac{x_k + x_l}{2}\Big) &= \frac{1}{2}\|x_k + x_l\|^2 - \phi(x_k) - \phi(x_l).
\end{aligned}$$

By the parallelogram identity we obtain

$$\begin{aligned}\frac{1}{4}\|x_k - x_l\|^2 &= \frac{1}{2}\|x_k\|^2 + \frac{1}{2}\|x_l\|^2 - \frac{1}{4}\|x_k + x_l\|^2 \\ &= Q_\phi(x_k) + Q_\phi(x_l) - 2\,Q_\phi\Big(\frac{x_k + x_l}{2}\Big) \\ &\leq Q_\phi(x_k) + Q_\phi(x_l) - 2\lambda \\ &< \varepsilon \quad \text{for } k,l \text{ sufficiently large.}\end{aligned}$$

Thus $x_0 = \lim_{k\to\infty} x_k$ exists. Moreover $Q_\phi(x_0) = \lim_{k\to\infty} Q_\phi(x_k) = \lambda$, by continuity of Q_ϕ.

(c): Cauchy-Schwarz yields

$$\|\phi\| = \sup_{\|x\|\leq 1} \phi(x) = \sup_{\|x\|\leq 1} \langle x, x_0\rangle \leq \|x_0\|.$$

For the reverse direction we use x_0 as test function in (11.5), this is a standard argument:

$$\|x_0\|^2 = \langle x_0, x_0\rangle = \phi(x_0) \leq \|\phi\|\,\|x_0\|.$$

In particular $\phi = 0$ implies $x_0 = 0$, and hence the uniqueness. □

Corollary 11.3 (weak*-compactness in Hilbert space) *Let x_k be a bounded sequence in a separable Hilbert space. Then there exists an $x_0 \in X$ such that for a subsequence*

$$\langle x_k, x\rangle \to \langle x_0, x\rangle \quad \textit{for all } x \in X.$$

The separability condition can be dropped in Corollary 11.3, and likewise in Corollary 11.1 when $1 < p < \infty$. The simple argument uses that these spaces are reflexive, which follows directly from Theorem 11.8 and 11.10. A space X is reflexive if the canonical map into its bidual X'' is surjective.

12 Hodge theory

We start with the definition of Sobolev differential forms on an n-dimensional manifold M. For a local chart $\varphi : U \to V$ the bundle $\Lambda^p TM|_U$ is trivial with local basis $dx^I = dx^{i_1} \wedge \ldots \wedge dx^{i_p}$ $I = (i_1, \ldots, i_p) \in I(n,p)$. The associated coordinates of a form ω are denoted by ω^φ.

Definition 12.1 (local Sobolev space) *Let M be a differentiable manifold.*

(1) *$L^2_{\mathrm{loc}}(\Lambda^* TM)$ is the space of forms α with $\alpha^\varphi \in L^2_{\mathrm{loc}}(V)$ for any chart $\varphi : U \to V$.*

(2) *$W^{1,2}_{\mathrm{loc}}(\Lambda^* TM)$ is the subspace with $\alpha^\varphi \in W^{1,2}_{\mathrm{loc}}(V)$ for any chart $\varphi : U \to V$.*

If $\varphi_i : U_i \to V_i$, $i = 1,2$, are charts and $\phi : \varphi_1(U_1 \cap U_2) \to \varphi_2(U_1 \cap U_2)$ is the coordinate change, then by (6.3) the representations of a p-form ω are related by

$$\omega_I^{\varphi_1} = \sum_{J \in I(n,p)} \omega_J^{\varphi_2} \circ \phi \,\det\Big(\frac{\partial \phi^J}{\partial x^I}\Big). \tag{12.6}$$

By the transformation and product rules, Theorem 11.4 and Lemma 11.2, we see that if ω^{φ_2} is locally in $W^{1,2}$ on $\varphi_2(U_1 \cap U_2)$, then also $\omega^{\varphi_1} \in W^{1,2}$ locally on $\varphi_1(U_1 \cap U_2)$. In particular, it is sufficient in Definition 12.1 to require this property for a fixed atlas. In order to define the Sobolev norm we need to choose an atlas with a little care.

Definition 12.2 (good atlas) *Let M be a compact manifold. We say that a finite atlas $\mathcal{A}$ with charts $\varphi_i : U_i \to V_i$, $i = 1, \ldots, N$, is good, if there exist charts $\tilde{\varphi}_i : \tilde{U}_i \to \tilde{V}_i$ such that*

$$U_i \subset\subset \tilde{U}_i \quad \textit{and} \quad \varphi_i = \tilde{\varphi}_i|_{U_i}.$$

For such a good atlas $\mathcal{A}$ we define the norm

$$\|\alpha\|_{W^{1,2}_{\mathcal{A}}} = \sum_{i=1}^{N} \|\alpha^{\varphi_i}\|_{W^{1,2}(V_i)}.$$

The norm is finite for any $\alpha \in W^{1,2}_{\mathrm{loc}}(\Lambda^* TM)$. Assume that each of the charts $\varphi_{1,2}$ in (12.6) belongs to some good atlas $\mathcal{A}_{1,2}$. Then $\varphi_1(U_1 \cap U_2) \subset\subset \varphi_1(\tilde{U}_1 \cap \tilde{U}_2)$, in particular ϕ has bounded C^2-norm on $\varphi_1(U_1 \cap U_2)$ and hence, again by the transformation and product rule,

$$\|\omega^{\varphi_1}\|_{W^{1,2}(\varphi_1(U_1 \cap U_2))} \leq C(\|\phi\|_{C^2})\, \|\omega^{\varphi_2}\|_{W^{1,2}(\varphi_2(U_1 \cap U_2))}.$$

This shows that the norms associated to $\mathcal{A}_{1,2}$ are equivalent. We denote the resulting Banach space by $W^{1,2}(\Lambda^* TM)$, and the norm by $\|\cdot\|_{W^{1,2}(M)}$.

Lemma 12.1 *For M compact the space $C^\infty(\Lambda^* TM)$ is dense in $W^{1,2}(\Lambda^* TM)$.*

Proof. Let $\omega \in W^{1,2}(\Lambda^* TM)$ be given. Assume first that $\operatorname{spt} \omega$ is contained in U where $U \subset\subset \tilde{U}$ for a chart $\varphi : \tilde{U} \to \tilde{V}$. By Lemma 11.1 there is a sequence $\omega_k^\varphi \in C^\infty_c(U)$ such that

$$\omega_k^\varphi \to \omega^\varphi \in W^{1,2}(U).$$

Put $\omega_k = (\omega_k^\varphi)_I \circ \varphi\, dx^I \in C^\infty(\Lambda^* TM)$. We can assume that $\operatorname{spt} \omega_k \subset K$ for all k where $K \subset U$ is compact. In fact, the ω_k^φ are obtained by mollification, hence their support has maximal distance to $\operatorname{spt} \omega^\varphi$ going to zero. Now in a chart $\varphi_i : U_i \to V_i$ from a good atlas, we have

$$\omega^{\varphi_i} = \omega_k^{\varphi_i} = 0 \quad \text{ on } \varphi_i(U_i \backslash K),$$

while on $\varphi_i(U \cap U_i)$ we have from (12.6) for $\phi = \varphi \circ \varphi_i^{-1}|_{\varphi_i(U_i \cap U)}$

$$\begin{aligned}
(\omega^{\varphi_i})_I &= \sum_{J \in I(n,p)} (\omega^\varphi)_J \circ \phi \, \det\Big(\frac{\partial \phi^J}{\partial x^I}\Big) \quad \text{ on } \varphi_i(U_i \cap U),\\
(\omega_k^{\varphi_i})_I &= \sum_{J \in I(n,p)} (\omega_k^\varphi)_J \circ \phi \, \det\Big(\frac{\partial \phi^J}{\partial x^I}\Big) \quad \text{ on } \varphi_i(U_i \cap U).
\end{aligned}$$

By the transformation and product rules, and recalling that ϕ is C^2-bounded, we conclude that $\omega_k^{\varphi_i} \to \omega^{\varphi_i}$ in $W^{1,2}$ on $\varphi_i(U_i) = \varphi_i(U_i \backslash K) \cup \varphi_i(U_i \cap U)$. To get the approximation for general $\omega \in W^{1,2}(\Lambda^* TM)$, one uses a partition of unity subordinate to a good atlas in the obvious way. $\square$

Next let $g = g_{ij}\, dx^i \otimes dx^j$ be a Riemannian metric on M. We have the isomorphism

$$TM \to T^*M,\ X \mapsto g(X, \cdot\,) = g_{ij}\, X^i dx^j =: X^\flat.$$

The inverse isomorphism is, writing g^{ij} for the elements of the inverse Gram matrix,

$$T^*M \to TM,\ \xi \mapsto g^{ij}\, \xi_i \frac{\partial}{\partial x^j} = \xi^\sharp.$$

In fact we check that

$$(X^\flat)^\sharp = g^{jk}(X^\flat)_j \frac{\partial}{\partial x^k} = g^{jk} g_{ij} X^i \frac{\partial}{\partial x^k} = X^i \frac{\partial}{\partial x^i} = X.$$

On the bundle of 1-forms we get the induced metric

$$g(\xi,\eta) = g(\xi^\sharp,\eta^\sharp) = g_{kl}\, g^{ik}\xi_i\, g^{jl}\eta_j = g^{ij}\xi_i\eta_j.$$

Further on the bundle $\bigotimes^p(TM)$ of p-linear forms we get

$$g(\xi,\eta) = g^{i_1j_1}\cdot\ldots\cdot g^{i_pj_p}\xi_{i_1\ldots i_p}\eta_{j_1\ldots j_p}.$$

The definition is independent of the coordinates, in fact we have

$$\begin{aligned} g(\xi_1\otimes\ldots\otimes\xi_p,\eta_1\otimes\ldots\otimes\eta_p) &= g^{i_1j_1}\cdot\ldots\cdot g^{i_pj_p}(\xi_1)_{i_1}\cdot\ldots\cdot(\xi_p)_{i_p}\,(\eta_1)_{j_1}\cdot\ldots\cdot(\eta_p)_{j_p} \\ &= g(\xi_1,\eta_1)\cdot\ldots\cdot g(\xi_p,\eta_p). \end{aligned}$$

The space $\Lambda^p(TM)$ is a subspace of $\bigotimes^p(TM)$, thus we get a metric also there. However, it is common to normalize by putting

$$\langle\xi,\eta\rangle_g = p!\,g(\xi,\eta) \qquad \text{for } \xi,\eta\in\Lambda^p(TM):$$

With this definition we obtain, for $\xi_i,\,\eta_j\in T^*M$,

$$\langle\xi_1\wedge\ldots\wedge\xi_p,\eta_1\wedge\ldots\wedge\eta_p\rangle_g = \det g(\xi_i,\eta_j).$$

Namely by definition of the wedge product

$$\xi_1\wedge\ldots\wedge\xi_p = \frac{1}{p!}\sum_{\sigma\in S_p}\operatorname{sign}(\sigma)\,\xi_{\sigma(1)}\otimes\ldots\otimes\xi_{\sigma(p)}.$$

We compute, using that $\sigma(i)=k$ implies $\tau(i)=\tau\sigma^{-1}(k)$,

$$\begin{aligned} \langle\xi_1\wedge\ldots\wedge\xi_p,\eta_1\wedge\ldots\wedge\eta_p\rangle_g &= \frac{1}{p!}\sum_{\sigma,\tau\in S_p}\operatorname{sign}(\sigma\tau)\,g(\xi_{\sigma(1)},\eta_{\tau(1)})\cdot\ldots\cdot g(\xi_{\sigma(p)},\eta_{\tau(p)}) \\ &= \frac{1}{p!}\sum_{\sigma\in S_p}\sum_{\tau\in S_p}\operatorname{sign}(\tau\sigma^{-1})g(\xi_1,\eta_{\tau\sigma^{-1}(1)})\cdot\ldots\cdot g(\xi_p,\eta_{\tau\sigma^{-1}(p)}) \\ &= \det g(\xi_i,\eta_j). \end{aligned}$$

It follows that if $\omega^1,\ldots,\omega^n$ is an orthonormal basis of T^*M, then $\omega^I=\omega^{i_1}\wedge\ldots\wedge\omega^{i_p}$, $I\in I(n,p)$, is an orthonormal basis for Λ^pTM.

Lemma 12.2 *For $\alpha\in\Lambda^p(TM)$, $\beta\in\Lambda^{p+1}(TM)$ and $\xi\in T^*M$ we have*

$$\langle\xi\wedge\alpha,\beta\rangle_g = \langle\alpha,X\llcorner\beta\rangle_g. \qquad \textit{where } X=\xi^\sharp.$$

Proof. It suffices to check the claim in $\mathbb{R}^n$ for $\alpha=dx^I$, $\beta=dx^J$ and $\xi=dx^i$, hence $X=e_i$. Using the determinant formula, one checks that then both sides are equal to $(-1)^k$ in the case $J=(i_1,\ldots,i_k,i,i_{k+1},\ldots,i_p)$ for $0\le k\le n$, and zero otherwise. □

Theorem 12.1 (codifferential) *Let (M,g) be a Riemannian manifold of dimension n. Then there is a unique operator $d^*: C^1(\Lambda^*TM) \to C^0(\Lambda^{*-1}TM)$ with the property*

$$\int_M \langle d\alpha, \beta\rangle_g \, d\mu_g = \int_M \langle \alpha, d^*\beta\rangle_g \, d\mu_g \quad \text{for all } \alpha \in C^1_c(\Lambda^*TM). \tag{12.7}$$

In local coordinates, for $g^{IJ} = \langle dx^I, dx^J\rangle_g$ and (g_{IJ}) the matrix inverse, d^* is given by

$$d^*\beta = -\frac{1}{\sqrt{\det g}}\, g_{IJ}\, \partial_i \left(g^{ij} \sqrt{\det g}\, \langle dx^I, \frac{\partial}{\partial x^j} \llcorner \beta\rangle_g\right) dx^J. \tag{12.8}$$

Proof. We first show uniqueness. For linear maps $d^*_{1,2}$ with (12.7) we conclude

$$\int_M \langle \alpha, d^*_1\beta - d^*_2\beta\rangle_g \, d\mu_g = 0 \quad \text{for all } \alpha \in C^1_c(\Lambda^*TM).$$

The calculus of variations fundamental lemma yields $d^*_1\beta = d^*_2\beta$. For existence we calculate

$$\begin{aligned}
\langle d\alpha, \beta\rangle_g \sqrt{\det g} &= \partial_i \alpha_I \langle dx^i \wedge dx^I, \beta\rangle_g \sqrt{\det g} \\
&= \partial_i \alpha_I \langle dx^I, g^{ij} \frac{\partial}{\partial x^j} \llcorner \beta\rangle_g \sqrt{\det g} \\
&= \partial_i \big(\alpha_I \langle dx^I, g^{ij} \frac{\partial}{\partial x^j} \llcorner \beta\rangle_g \sqrt{\det g}\big) - \alpha_I\, \partial_i \big(g^{ij} \sqrt{\det g}\, \langle dx^I, \frac{\partial}{\partial x^j} \llcorner \beta\rangle_g\big)
\end{aligned}$$

We want to write

$$-\partial_i \big(g^{ij} \sqrt{\det g}\, \langle dx^I, \frac{\partial}{\partial x^j} \llcorner \beta\rangle_g\big) = g^{IK} \gamma_K \sqrt{\det g},$$

since then we obtain

$$-\alpha_I\, \partial_i \big(g^{ij} \sqrt{\det g}\, \langle dx^I, \frac{\partial}{\partial x^j} \llcorner \beta\rangle_g\big) = g^{IK} \alpha_I \gamma_K \sqrt{\det g} = \langle \alpha, \gamma\rangle_g \sqrt{\det g},$$

and the claim of the lemma follows by putting $d^*\beta = \gamma$. To determine γ, we multiply by g_{IL}, then sum over I and divide by $\sqrt{\det g}$ to get

$$\gamma_L = g^{IK} g_{IL} \gamma_K = -\frac{1}{\sqrt{\det g}}\, g_{IL}\, \partial_i \left(g^{ij} \sqrt{\det g}\, \langle dx^I, \frac{\partial}{\partial x^j} \llcorner \beta\rangle_g\right).$$

□

Definition 12.3 (Hodge Laplacian) *Let (M,g) be a Riemannian manifold. The Hodge Laplace operator on forms $\alpha \in C^2(\Lambda^*TM)$ is*

$$\Delta\alpha = (d^*d + dd^*)\alpha.$$

Lemma 12.3 *For the standard metric on $\mathbb{R}^n$, the Hodge Laplacian is given by*

$$\Delta\alpha = -\sum_{i=1}^n \partial_i^2 \alpha. \tag{12.9}$$

Proof. In the case of the the standard metric on $\mathbb{R}^n$ equation (12.8) yields

$$d\alpha = \sum_{i=1}^{n} dx^i \wedge \partial_i \alpha \quad \text{and} \quad d^*\alpha = -\sum_{i=1}^{n} e_i \llcorner \partial_i \alpha.$$

It follows that

$$d^* d\alpha = -\sum_{i,j=1}^{n} e_i \llcorner (dx^j \wedge \partial_{ij}^2 \alpha) \quad \text{and} \quad dd^*\alpha = -\sum_{i,j=1}^{n} dx^i \wedge (e_j \llcorner \partial_{ij}^2 \alpha).$$

We claim that for any $v \in \mathbb{R}^n$, $\omega = \langle v, \cdot \rangle$ and any $\alpha \in \Lambda^*(\mathbb{R}^n)$ we have

$$v \llcorner (\omega \wedge \alpha) + \omega \wedge (v \llcorner \alpha) = |v|^2 \alpha.$$

For this we may assume $v = v_1$ where $v_1, \ldots, v_n$ is an orthonormal basis. Put $\omega^i = \langle v_i, \cdot \rangle$ and write $\alpha = \omega \wedge \beta + \gamma$ where β, γ do not have $\omega = \omega^1$ in their basis representation. Then $\omega \wedge \alpha = \omega \wedge \gamma$ and $v \llcorner \alpha = \beta$. This yields further $v \llcorner (\omega \wedge \alpha) = \gamma$ and $\omega \wedge (v \llcorner \alpha) = \omega \wedge \beta$, and the claim follows. Next for $B(e_i, e_j)\alpha = e_i \llcorner (dx^j \wedge \alpha) + dx^i \wedge (e_j \llcorner \alpha)$ we compute

$$\begin{aligned}
\frac{1}{2}\big(B(e_i, e_j) + B(e_j, e_i)\big)\,\alpha &= \frac{1}{2}\big(B(e_i + e_j, e_i + e_j) - B(e_i, e_i) - B(e_j, e_j)\big)\,\alpha \\
&= \frac{1}{2}\big(|e_i + e_j|^2 - |e_i|^2 - |e_j|^2\big)\alpha \\
&= \delta_{ij}\,\alpha.
\end{aligned}$$

We conclude using the symmetry of $\partial_{ij}^2 \alpha$

$$d^* d\alpha + dd^*\alpha = -\sum_{i,j=1}^{n} \frac{1}{2}\big(B(e_i, e_j) + B(e_j, e_i)\big)\,\partial_{ij}^2 \alpha = -\sum_{i=1}^{n} \partial_i^2 \alpha.$$

□

Unfortunately Hodge decided to take the sign opposite to Laplace, which is a constant source of confusion. We now come to a basic estimate for the Hodge Laplace operator.

Theorem 12.2 (Gaffney's inequality) *Let g be a C^1 Riemannian metric on a compact manifold M. There exists a constant $C = C(M, g) < \infty$ such that for all $\omega \in W^{1,2}(\Lambda^* TM)$*

$$\|\omega\|_{W^{1,2}(M)}^2 \leq C \int_M \big(\|d\omega\|_g^2 + \|d^*\omega\|_g^2 + \|\omega\|_g^2\big)\, d\mu_g. \tag{12.10}$$

Our strategy is to first prove a local version, and then globalize it using a partition of unity. In the local situation we want to employ a special coordinate system.

Lemma 12.4 (almost Euclidean coordinates) *Let g be a C^1 metric on a manifold M of dimensio n. Then at any $p \in M$ there is a chart $\varphi : U \to \varphi(U) \subset \mathbb{R}^n$ such that*

$$g_{ij}(p) = \delta_{ij} \quad \text{and} \quad \partial_i g_{jk}(p) = 0 \quad \text{for } 1 \leq i, j, k \leq n. \tag{12.11}$$

Proof. We may assume that $g = (g_{ij})$ is given on a neighborhood of $p = 0 \in \mathbb{R}^n$, and that we have already $g_{ij}(0) = \delta_{ij}$. This is achieved by composing with a suitable linear map. We now consider a local diffeomorphism ϕ of the form

$$\phi(x) = \Big(x^k + \frac{1}{2} C^k_{ij} x^i x^j\Big) e_k \quad \text{where } C^k_{ij} = C^k_{ji}.$$

We have $\phi(0) = 0$, $\partial_j \phi(0) = e_j$ and $\partial^2_{ij}\phi(0) = C^k_{ij} e_k$. Using this we compute

$$\begin{aligned}
(\phi^* g)_{jk}(0) &= g(\phi(0))(\partial_j \phi(0), \partial_k \phi(0)) = \delta_{jk},\\
\partial_i (\phi^* g)_{jk}(0) &= \partial_i [g(\phi(x))(\partial_j \phi(x), \partial_k \phi(x))]_{x=0}\\
&= g(0)(C^l_{ij} e_l, e_k) + g(0)(e_j, C^l_{ik} e_l) + \partial_i g(0)(e_j, e_k)\\
&= C^k_{ij} + C^j_{ik} + \partial_i g_{jk}(0).
\end{aligned}$$

We choose $C^k_{ij} = -\frac{1}{2}\big(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\big)(0)$. Then $C^k_{ij} = C^k_{ji}$, and moreover

$$\begin{aligned}
C^k_{ij} + C^j_{ik} &= -\frac{1}{2}\big(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij}\big)(0)\\
&\quad -\frac{1}{2}\big(\partial_i g_{kj} + \partial_k g_{ij} - \partial_j g_{ik}\big)(0)\\
&= -\partial_i g_{jk}(0).
\end{aligned}$$

Therefore the metric $\phi^* g$ has the desired properties. $\square$

As an alternative, Riemann normal coordinates also have the property (12.11). There are defined via the Riemannian exponential map. However, the differentiability of that map is more involved, and the simple construction here suffices for our purposes.

Lemma 12.5 *Let $G = (g_{ij})$ be a C^1 metric on $U \subset \mathbb{R}^n$ which satisfies, for all $1 \leq i, j \leq n$,*

$$\|g_{ij} - \delta_{ij}\|_{C^0(U)} \leq \delta \quad \text{and} \quad \|Dg_{ij}\|_{C^0(U)} \leq \Lambda. \tag{12.12}$$

If $\delta < \delta_0 = \delta_0(n,p)$ then we have the inequality, for a constant $C = C(n,p) < \infty$,

$$\int_U |D\omega|^2 \leq 2 \int_M \big(\|d\omega\|_g^2 + \|d_g^*\omega\|_g^2\big)\, d\mu_g + C\Lambda^2 \int_U \|\omega\|_g^2\, d\mu_g \quad \textit{for all } \omega \in C^1_c(U, \Lambda^p \mathbb{R}^n).$$

Proof. For the Euclidean metric, partial integration and Lemma 12.3 imply the identity

$$\begin{aligned}
\int_U |D\omega|^2 &= \sum_{i=1}^n \int_U \langle \partial_i \omega, \partial_i \omega\rangle\\
&= -\sum_{i=1}^n \int_U \langle \partial_i^2 \omega, \omega\rangle\\
&= \int_U \big\langle (d^*d + dd^*)\omega, \omega\big\rangle\\
&= \int_U \big(|d\omega|^2 + |d^*\omega|^2\big).
\end{aligned}$$

We emphasize that in general $D\omega$ is not estimated pointwise by $d\omega$ and $d^*\omega$, in fact $D\omega$ can be nonzero while both $d\omega$ and $d^*\omega$ vanish. In the above calculation we needed

$\omega \in C^2_c(U, \Lambda^p \mathbb{R}^n)$, however this is removed afterwards by approximation. In fact, the equation follows for all $\omega \in W^{1,2}_0(U, \Lambda^p \mathbb{R}^n)$, i.e. the $W^{1,2}$-closure of $C^\infty_c(U, \Lambda^p \mathbb{R}^n)$. Our goal now is to apply this Euclidean identity, treating the metric g as a pertubation.

Step 1: There exist smooth functions A^{ij}, B^{ijklK}_J on $\mathrm{GL}_n(\mathbb{R})$ such that

$$(d^*_g \omega)_J = (d^*\omega)_J + A^{ij}(G)\, \partial_i \omega_{j,J} + B^{ijklK}_J(G)\, \partial_i g_{kl}\, \omega_{j,K}. \tag{12.13}$$

In fact, starting with formula (12.8) we compute

$$\begin{aligned}
(d^*_g \omega)_J &= -\frac{1}{\sqrt{\det g}} g_{IJ}\, \partial_i \big(g^{ij} \sqrt{\det g}\, \langle dx^I, \frac{\partial}{\partial x^j} \llcorner \omega \rangle_g\big) \\
&= -\frac{1}{\sqrt{\det g}} g_{IJ}\, \partial_i \big(g^{ij} \sqrt{\det g}\, g^{IK} \omega_{j,K}\big) \\
&= -\partial_i \omega_{i,J} + \underbrace{\big(\delta^{ij} - g^{ij}\big)}_{=A^{ij}(G)} \partial_i \omega_{j,J} \underbrace{- \frac{1}{\sqrt{\det g}}\, g_{IJ}\, \frac{\partial}{\partial g_{kl}} \big(g^{ij} \sqrt{\det g}\, g^{IK}\big)}_{=B^{ijklK}_J(G)}\, \partial_i g_{kl}\, \omega_{j,K}.
\end{aligned}$$

In the last step we used the chain rule. Recall that by Cramer's rule, g^{ij} is a function of $G = (g_{ij})$, and so is $g^{IK} = \det g^{i_j k_l}$ where $1 \leq j, l \leq p$. For $f \in C^1(\mathrm{GL}_n(\mathbb{R}))$ let

$$C(f) = \|f\|_{C^0(K_0)} \quad \text{where } K_0 = \big\{G \in \mathbb{R}^{n \times n} : |G - \mathrm{E}_n| \leq \frac{1}{2}\big\}.$$

Note that K_0 is a compact subset of $\mathrm{GL}_n(\mathbb{R})$, hence $C(f) < \infty$. From now on we assume that $\delta_0 > 0$ is chosen such that $G(x) \in K_0$ for all $x \in U$. The collection of functions B^{ijklK}_J, for $1 \leq i, j, k, l \leq n$ and $J, K \in I(n,p)$, depends only on n, p, therefore we have a bound $|B(G)| \leq C(B) = C(n,p)$. Furthermore as $A^{ij}(\mathrm{E}_n) = 0$ we can estimate

$$|A^{ij}(G)| = \Big| \int_0^1 DA^{ij}\big((1-t)\mathrm{E}_n + tG\big)(G - \mathrm{E}_n)\, dt \Big| \leq C(DA^{ij})\, |G - \mathrm{E}_n| \leq C(n,p)\delta. \tag{12.14}$$

Step 2: We claim that for any p-form ω one has the inequality

$$\int_U |\omega|^2 \leq (1 + C\delta) \int_U \|\omega\|^2_g\, d\mu_g \quad \text{where } C = C(n,p) < \infty. \tag{12.15}$$

Namely in coordinates we can estimate, arguing as for (12.14),

$$|\omega|^2 = \|\omega\|^2_g \sqrt{\det g} + \big(\delta^{IJ} - g^{IJ} \sqrt{\det g}\big) \omega_I \omega_J \leq \|\omega\|^2_g \sqrt{\det g} + C\delta |\omega|^2.$$

Rearranging gives (12.15), for an appropriate constant $C = C(n,p) < \infty$.

Step 3: We now put things together. Using step 1 we have, for any $\varepsilon > 0$,

$$\begin{aligned}
|d^*\omega|^2 &= |d^*_g \omega - A(G) D\omega - B(G)(DG)\omega|^2 \\
&\leq (1+\varepsilon)\, |d^*_g \omega^2| + C(\varepsilon)\, \delta^2 |D\omega|^2 + C(\varepsilon) \Lambda^2 |\omega|^2.
\end{aligned}$$

By the Euclidean identity we obtain, estimating with (12.15),

$$
\begin{aligned}
\int_U |D\omega|^2 &= \int_U (|d\omega|^2 + |d^*\omega|^2) \\
&\leq \int_U |d\omega|^2 + (1+\varepsilon) \int_U |d_g^*\omega^2| + C(\varepsilon)\,\delta^2 \int_U |D\omega|^2 + C(\varepsilon)\Lambda^2 \int_U |\omega|^2 \\
&\leq (1+C\delta) \int_U \|d\omega\|_g^2 \, d\mu_g + (1+\varepsilon)(1+C\delta) \int_U \|d_g^*\omega\|_g^2 \, d\mu_g \\
&\quad + C(\varepsilon)\,\delta^2 \int_U |D\omega|^2 + C(\varepsilon)\Lambda^2 (1+C\delta) \int_U \|\omega\|_g^2 \, d\mu_g.
\end{aligned}
$$

Choosing $\varepsilon = \frac{1}{2}$ this simplifies to

$$
(1 - C\delta^2) \int_U |D\omega|^2 \leq \frac{3}{2}(1+C\delta) \int_U \big(\|d\omega\|_g^2 + \|d_g^*\omega\|_g^2\big)\, d\mu_g + C\Lambda^2 \int_U \|\omega\|_g^2 \, d\mu_g.
$$

The claim of the lemma follows by taking $\delta > 0$ sufficiently small. $\square$

Proof of Theorem 12.2: We employ a covering argument to globalize the estimate. We will need the following product rules, where $\omega \in C^1(\Lambda^* TM)$ and $\eta \in C^1(M)$,

$$
d(\eta\omega) = \eta\, d\omega + d\eta \wedge \omega \quad \text{ and } \quad d_g^*(\eta\omega) = \eta\, d_g^*\omega - \mathrm{grad_g}\, \eta \llcorner \omega.
$$

The first is well-kown, the second then follows from Theorem 12.1 and Lemma 12.2, namely

$$
\int_M \langle d_g^*(\eta\omega), \varphi \rangle_g \, d\mu_g = \int_M \langle \omega, d(\eta\varphi) - d\eta \wedge \varphi \rangle_g \, d\mu_g = \int_M \langle d_g^*\omega - \mathrm{grad_g}\, \eta \llcorner \omega, \varphi \rangle \, d\mu_g.
$$

Let (U, g) be as in Lemma 12.5. Then it is easy to see that

$$
\begin{aligned}
\|d(\eta\omega)\|_g &\leq |\eta|\, \|d\omega\|_g + C|d\eta|\, \|\omega\|_g, \\
\|d_g^*(\eta\omega)\|_g &\leq |\eta|\, \|d_g^*\omega\|_g + C|d\eta|\, \|\omega\|_g.
\end{aligned}
$$

We now assume that $\eta \in C_c^1(U)$. Then Lemma 12.5 applies to $\eta\omega$, and we obtain

$$
\begin{aligned}
\int_U |D(\eta\omega)|^2 &\leq 2 \int_U \big(\|d(\eta\omega)\|_g^2 + \|d_g^*(\eta\omega)\|_g^2\big)\, d\mu_g + C\Lambda^2 \int_U \|\eta\omega\|_g^2 \, d\mu_g \\
&\leq 4 \int_M \big(\|d\omega\|_g^2 + \|d_g^*\omega\|_g^2\big)\, d\mu_g + C(\|\eta\|_{C^1(U)}^2 + \Lambda^2) \int_M \|\omega_g\|^2 \, d\mu_g.
\end{aligned}
$$

By Lemma 12.4, for any $p \in M$ there is a local chart (U, φ) which satisfies the conditions (12.12) with $\Lambda = 1$. Chose $\varrho > 0$ such that $\varphi(U)$ contains the closed ball around $x = \varphi(p)$ of radius 2ϱ, and let $\eta \in C_c^1(B_{2\varrho}(x))$ be a standard test function with $\eta \equiv 1$ on $B_\varrho(x)$. We then have, also recalling (12.15), on $U_p = \varphi^{-1}(B_\varrho(x))$,

$$
\|\omega\|_{W^{1,2}(U_p)}^2 \leq 4 \int_M \big(\|d\omega\|_g^2 + \|d_g^*\omega\|_g^2\big)\, d\mu_g + C(\varrho) \int_M \|\omega_g\|^2 \, d\mu_g.
$$

Since M is compact, we can select points p_i for $i = 1, \ldots, N < \infty$ such that the $U_i := U_{p_i}$ are a covering of M. Denoting the corresponding charts by φ_i, we conclude that $(U_i, \varphi_i|_{U_i})$ is a good atlas of M. This finishes the proof for forms of class $C^1(\Lambda^* TM)$. For $\omega \in W^{1,2}(\Lambda^* TM)$ the result follows by approximation using Lemma 12.1, since all terms in the statement pass to the limit under $W^{1,2}$ convergence. Alternatively, it is easy to check that the proof of Lemma 12.5 applies to forms in $W^{1,2}(\Lambda^*TM)$. $\square$

We now employ standard Hilbert space theory to analyse the analyse the linear operator Δ_g. Our approach is to define Δ_g in the weak sense. For M compact, which is now always assumed, we let

$$\Delta_g : W^{1,2}(\Lambda^*TM) \to W^{1,2}(\Lambda^*TM)',\ (\Delta_g\omega)(\eta) = \int_M \big(\langle d\omega, d\eta\rangle_g + \langle d_g^*\omega, d_g^*\eta\rangle_g\big)\, d\mu_g. \tag{12.16}$$

By Cauchy-Schwarz and the coordinate representations, we have

$$|\Delta_g\omega(\eta)| \leq \|d\omega\|_{L^2(g)}\|d\eta\|_{L^2(g)} + \|d_g^*\omega\|_{L^2(g)}\|d_g^*\eta\|_{L^2(g)} \leq C\,\|\omega\|_{W^{1,2}(M)}\|\eta\|_{W^{1,2}(M)}.$$

Thus Δ_g as in (12.16) is well-defined and continuous. For $\omega \in W^{2,2}(\Lambda^*TM)$ one can alternatively define $\Delta_g\omega \in L^2(\Lambda^*TM)$ by the pointwise formula. This yields the strong (or classical) definition of the operator, we write here $\Delta_g^{strong} : W^{2,2}(\Lambda^*TM) \to L^2(\Lambda^*TM)$ for distinction. To explain that the two notions are consistent we introduce the embedding

$$I : L^2(\Lambda^*TM) \to W^{1,2}(\Lambda^*TM)',\ I\phi(\eta) = \int_M \langle \phi, \eta\rangle_g\, d\mu_g.$$

We claim that for all $\omega \in W^{2,2}(\Lambda^*TM)$, $\phi \in L^2(\Lambda^*TM)$ we have the equivalence

$$\Delta_g^{strong}\omega = \phi \quad \Leftrightarrow \quad \Delta_g\omega = I\phi. \tag{12.17}$$

In fact partial integration yields, for any $\eta \in W^{1,2}(\Lambda^*TM)$,

$$\begin{aligned}
\Delta_g\omega(\eta) &= \int_M \big(\langle d\omega, d\eta\rangle_g + \langle d_g^*\omega, d_g^*\eta\rangle_g\big)\, d\mu_g \\
&= \int_M \langle (d_g^*d + dd_g^*)\omega, \eta\rangle\, d\mu_g \\
&= \int_M \langle \Delta_g^{strong}\omega, \eta\rangle\, d\mu_g \\
&= I(\Delta_g^{strong}\omega)(\eta).
\end{aligned}$$

Thus $\Delta_g^{strong}\omega = \phi$ implies $\Delta_g\omega = I\phi$. Reversely, the calculation implies

$$I\phi(\eta) = \Delta_g\omega(\eta) = I(\Delta_g^{strong}\omega)(\eta) \quad \text{for all } \eta \in W^{1,2}(\Lambda^*TM).$$

As $W^{1,2}(\Lambda^*TM)$ is dense in $L^2(\Lambda^*TM)$, we conclude $\Delta_g^{strong}\omega = \phi$.

Theorem 12.3 *Let g be a C^1 metric on a compact n-dimensional manifold M. Then*

$$\Delta_g\omega = 0 \quad \Leftrightarrow \quad d\omega = 0,\ d_g^*\omega = 0. \tag{12.18}$$

The space of harmonic forms $H^p(M) = \{\omega \in W^{1,2}(\Lambda^*TM) : \Delta_g\omega = 0\}$ is finite dimensional.

Remark 12.1 To define harmonic forms on a general, possibly noncompact manifold M one allows that $\omega \in W^{1,2}_{\mathrm{loc}}(\Lambda^*TM)$, and uses only test functions with compact support:

$$\Delta_g\omega(\eta) = \int_M \big(\langle d\omega, d\eta\rangle_g + \langle d_g^*\omega, d_g^*\eta\rangle_g\big)\, d\mu_g = 0 \quad \text{for all } \eta \in W^{1,2}_c(\Lambda^*TM),$$

However, Theorem 12.3 does not extend to the noncompact case. For instance, a 1-form $\omega = adx + bdy$ on the unit disk $D \subset \mathbb{R}^2$ has derivatives

$$d\omega = (b_x - a_y)\, dx \wedge dy \quad \text{and} \quad d^*\omega = -(a_x + b_y).$$

Now ω harmonic just means that a, b are harmonic functions, while the equations $d\omega = 0$, $d^*\omega = 0$ are equivalent to $f = a - ib$ being holomorphic, a stronger condition. Moreover, none of the two spaces has finite dimension. One may try to select a finite-dimensional subspace by adding boundary or asymptotic conditions, and enlarging the space of test functions in an appropriate way.

Proof of Theorem 12.3. By definition (12.16) the equations $d\omega = 0$, $d_g^*\omega = 0$ trivially imply $\Delta_g \omega = 0$. Reversely, we get by testing with $\omega \in W^{1,2}(\Lambda^* TM)$

$$0 = (\Delta_g \omega)(\omega) = \int_M \left(\|d\omega\|_g^2 + \|d_g^*\omega\|_g^2\right) d\mu_g,$$

and hence $d\omega = 0$, $d_g^*\omega = 0$. Now assume that $\dim H^p(M) = \infty$, so that there is a linearly independent sequence $\omega_k \in H^p(M)$. We equip $H^p(M)$ with the $L^2(g)$ scalar product. Applying the Gram-Schmidt process we can assume that

$$\langle \omega_k, \omega_l \rangle_{L^2(g)} = \int_M \langle \omega_k, \omega_l \rangle_g \, d\mu_g = \delta_{kl} \quad \text{for } k, l \in \mathbb{N}.$$

Now as $d\omega_k = 0$, $d_g^*\omega_k = 0$ we obtain by Theorem 12.2

$$\|\omega_k\|_{W^{1,2}(M)} \leq C \, \|\omega_k\|_{L^2(g)} \leq C.$$

By Rellich, see Theorem 11.6, there exists an $\omega \in L^2(g)$ such that $\omega_k \to \omega$ in $L^2(g)$, after passing to a subsequence. But this contradicts

$$\|\omega_k - \omega_l\|_{L^2(g)} = \sqrt{2} \quad \text{whenever } k \neq l.$$

$\square$

Lemma 12.6 *Let g be a C^1 metric on a compact manifold M. There exists a constant $\lambda = \lambda(M, g) > 0$ such that for $\omega \in W^{1,2}(\Lambda^* TM)$*

$$\omega \perp_{L^2(g)} H^*(M) \quad \Rightarrow \quad \int_M \left(\|d\omega\|_g^2 + \|d_g^*\omega\|_g^2\right) d\mu_g \geq \lambda \, \|\omega\|_{L^2(g)}^2. \tag{12.19}$$

Proof. Assume that the claim fails for $\lambda = \frac{1}{k}$, for any $k \in \mathbb{N}$. Then there exist $\omega_k \in W^{1,2}(\Lambda^* TM)$ with $\omega_k \perp_{L^2(g)} H^*(M)$, such that

$$\int_M \left(\|d\omega_k\|_g^2 + \|d_g^*\omega_k\|_g^2\right) d\mu_g < \frac{1}{k} \, \|\omega\|_{L^2(g)}^2.$$

We normalize to $\|\omega_k\|_{L^2(g)} = 1$ and conclude from Theorem 12.2

$$\|\omega_k\|_{W^{1,2}(M)} \leq C\big(\|d\omega_k\|_{L^2(g)} + \|d_g^*\omega_k\|_{L^2(g)} + \|\omega_k\|_{L^2(g)}\big) \leq C.$$

Again by Rellich, Theorem 11.6, there exists an $\omega \in L^2(g)$ such that $\omega_k \to \omega$ in $L^2(g)$ for a subsequence, in particular

$$\|\omega\|_{L^2(g)} = 1 \quad \text{and} \quad \omega \perp_{L^2(g)} H^*(M). \tag{12.20}$$

Furthermore for $\eta \in W^{1,2}(\Lambda^*TM)$ we compute

$$\begin{aligned}
\langle \omega, d_g^*\eta \rangle_{L^2(g)} &= \lim_{k\to\infty} \langle \omega_k, d_g^*\eta \rangle_{L^2(g)} = \langle d\omega_k, \eta \rangle_{L^2(g)} = 0,\\
\langle \omega, d\eta \rangle_{L^2(g)} &= \lim_{k\to\infty} \langle \omega_k, d\eta \rangle_{L^2(g)} = \langle d_g^*\omega_k, \eta \rangle_{L^2(g)} = 0.
\end{aligned}$$

But now $\omega_k \to \omega$ weakly in $W^{1,2}(\Lambda^*TM)$ by Corollary 11.3, and we conclude $d\omega = 0$, $d_g^*\omega = 0$ by partial integration. This means $\omega \in H^*(M)$, contradicting (12.20). $\square$

Theorem 12.4 *Let g be a C^1 metric on a compact manifold M. Then the following hold:*

(a) $\operatorname{im} \Delta_g = \{\Lambda \in W^{1,2}(\Lambda^*TM)' : \Lambda|_{H^*(M)} = 0\}$.

(b) *Let $H^*(M)^\perp = \{\omega \in W^{1,2}(\Lambda^*TM) : \omega \perp_{L^2(g)} H^*(M)\}$. Then $\Delta_g : H^*(M)^\perp \to \operatorname{im}(\Delta_g)$ is an isomorphism.*

Proof. The condition $\Lambda|_{H^*(M)} = 0$ in (a) is clearly necessary, in fact for $\eta \in H^*(M)$ we have $d\eta = 0$, $d_g^*\eta = 0$ and hence

$$\Delta_g\omega(\eta) = \int_M \big(\langle d\omega, d\eta\rangle_g + \langle d_g^*\omega, d_g^*\eta\rangle_g \, d\mu_g = 0.$$

To prove that the condition is sufficient, we need to solve the equation $\Delta_g\omega = \Lambda$. The space $H^*(M)^\perp$ is a closed subspace of $W^{1,2}(\Lambda^*TM)$, and is therefore a Hilbert space with the $W^{1,2}$ scalar product. We consider on $H^*(M)^\perp$ the bilinear form

$$(\omega, \eta) = \int_M \big(\langle d\omega, d\eta\rangle_g + \langle d_g^*\omega, d_g^*\eta\rangle_g\big) \, d\mu_g.$$

The associated norm clearly has a bound $\|\omega\| \leq C\, \|\omega\|_{W^{1,2}(M)}$. On the other hand by Gaffney's inequality, Theorem 12.2, and Lemma 12.6 we have

$$\|\omega\|_{W^{1,2}(M)} \leq C\big(\|\omega\| + \|\omega\|_{L^2(g)}\big) \leq C\big(1 + \frac{1}{\sqrt{\lambda}}\big)\, \|\omega\|.$$

Thus the Riesz representation theorem, Theorem 11.10, applies on $H^*(M)^\perp$ with scalar product $(\cdot,\cdot)$. Given any $\Lambda \in W^{1,2}(\Lambda^*TM)'$, there exists an $\omega \in H^*(M)^\perp$ such that

$$\int_M \big(\langle d\omega, d\eta\rangle_g + \langle d_g^*\omega, d\eta\rangle_g\big)\, d\mu_g = \Lambda(\eta) \quad \text{for all } \eta \in H^*(M)^\perp.$$

Now if Λ vanishes on $H^*(M)$, then the equation holds also for $\eta \in H^*(M)$, and hence for all $\eta \in W^{1,2}(\Lambda^*TM)$. By definition this means $\Delta_g\omega = \Lambda$.

As $H^*(M)$ is the kernel of Δ_g, the operator in (b) is clearly bijective. Let $\Lambda = \Delta_g\omega$ for $\omega \in H^*(M)^\perp$. By testing with ω we infer

$$\int_M \big(\|d\omega\|_g^2 + \|d_g^*\omega\|_g^2\big)\, d\mu_g = (\Delta_g\omega)(\omega) \leq \|\Lambda\|\, \|\omega\|_{W^{1,2}(M)}.$$

Combining Theorem 12.2 and Lemma 12.6 we get further

$$\begin{aligned}
\|\omega\|^2_{W^{1,2}(M)} &\leq C\int_M \big(\|d\omega\|_g^2 + \|d_g^*\omega\|_g^2 + \|\omega\|_g^2\big)\,d\mu_g \\
&\leq C\Big(1+\frac{1}{\lambda}\Big)\int_M \big(\|d\omega\|_g^2 + \|d_g^*\omega\|_g^2\big)\,d\mu_g \\
&\leq C\Big(1+\frac{1}{\lambda}\Big)\,\|\Lambda\|\,\|\omega\|_{W^{1,2}(M)}.
\end{aligned}$$

After cancelling we obtain

$$\|\omega\|_{W^{1,2}(M)} \leq C\Big(1+\frac{1}{\lambda}\Big)\,\|\Lambda\|. \tag{12.21}$$

It follows that the operator in (b) has inverse bounded by $C\big(1+\frac{1}{\lambda}\big)$. □

By definition, an operator $T \in L(X,Y)$ between Banach spaces is Fredholm, if $\operatorname{im} T$ is a closed subspace of Y and both $\ker T$ and $\operatorname{coker} T = T/\operatorname{im} T$ have finite dimension. The difference $\operatorname{ind} T = \dim\ker T - \dim\operatorname{coker} T \in \mathbb{Z}$ is called the Fredholm index. The set of Fredholm operators is open in $L(X,Y)$ and the index is locally constant, see Theorem 25.2 in [**?**]. Theorem 12.4 asserts that $\Delta_g : W^{1,2}(\Lambda^*TM) \to W^{1,2}(\Lambda^*TM)'$ is a Fredholm operator of index zero. The fact that the index is zero is expected from the symmetry of the operator. Namely, if $T \in L(X,Y)$ has closed range then $T' \in L(Y',X')$, $T'\psi(x) = \psi(Tx)$, has also closed range and one has the canonical isomorphism, see Theorem 25.4 in [**?**],

$$F : \ker T' \to (Y/\mathrm{im\,T})',\ F(\psi)[y] = \psi(y).$$

In particular $\dim\ker T' = \dim\operatorname{coker} T$ if T is Fredholm. In our case we have $\Delta_g' \circ J = \Delta_g$, where $J : W^{1,2}(\Lambda^*TM) \to W^{1,2}(\Lambda^*TM)''$ denotes the canonical isomorphism. Namely using the symmetry in definition (12.16) we compute

$$\Delta_g'(J\omega)(\eta) = (J\omega)(\Delta_g\eta) = \Delta_g\eta(\omega) = \Delta_g\omega(\eta).$$

We conclude $\dim\operatorname{coker}\Delta_g = \dim\ker\Delta_g' = \dim\ker\Delta_g$. Before we can proceed to the main result of Hodge, we need a fundamental regularity result.

Theorem 12.5 ($W^{2,2}$-regularity) *Let g be a C^2 metric on a compact manifold M, and assume that $\omega \in W^{1,2}(\Lambda^*TM)$ solves $\Delta_g\omega = I\phi$ where $\phi \in L^2(\Lambda^*TM)$, that is*

$$\int_M \big(\langle d\omega, d\eta\rangle_g + \langle d_g^*\omega, d_g^*\eta\rangle_g\big)\,d\mu_g = \int_M \langle\phi,\eta\rangle_g\,d\mu_g \quad \textit{for all } \eta \in W^{1,2}(\Lambda^*TM).$$

Then $\omega \in W^{2,2}(\Lambda^*TM)$, and

$$\|\omega\|_{W^{2,2}(M)} \leq C\,\big(\|\phi\|_{L^2(g)} + \|\omega\|_{L^2(g)}\big) \quad \textit{where } C = C(M,g). \tag{12.22}$$

In the case of the Laplace operator on functions, this result is standard in any PDE lecture. We postpone a discussion of the proof.

Theorem 12.6 (Hodge) *Let g be a C^2 metric on a compact manifold M. Then one has an $L^2(g)$-orthogonal decomposition*

$$L^2(\Lambda^pTM) = dW^{1,2}(\Lambda^{p-1}TM) \oplus d_g^*W^{1,2}(\Lambda^{p+1}TM) \oplus H^p(M), \tag{12.23}$$

where $dW^{1,2}(\Lambda^{p-1}TM)$ and $d_g^*W^{1,2}(\Lambda^{p+1}TM)$ are closed subspaces.

Proof. We first show that the spaces are orthogonal. For $\alpha \in W^{1,2}(\Lambda^{p-1}TM)$, $\beta \in W^{1,2}(\Lambda^{p+1}TM)$ and $\eta \in H^p(M)$ we have by partial integration and Theorem 12.3

$$\begin{aligned}\langle d\alpha, \eta\rangle_{L^2(g)} &= \langle \alpha, d_g^*\eta\rangle_{L^2(g)} = 0,\\ \langle d_g^*\beta, \eta\rangle_{L^2(g)} &= \langle \alpha, d\eta\rangle_{L^2(g)} = 0.\end{aligned}$$

Furthermore, assuming that $\alpha \in W^{2,2}(\Lambda^{p-1}TM)$, partial integration yields

$$\langle d\alpha, d_g^*\beta\rangle_{L^2(g)} = \langle d^2\alpha, \beta\rangle_{L^2(g)} = 0.$$

The last orthogonality extends to $\alpha \in W^{1,2}(\Lambda^{p-1}TM)$ by approximation using Lemma 12.1. Now let $\phi \in L^2(\Lambda^pTM)$ be given. By Theorem 12.3 the space $H^p(M)$ has a finite basis $\eta_1, \ldots, \eta_N$ which is orthonormal in $L^2(g)$. We put

$$\eta = \sum_{i=1}^N \langle \phi, \eta_i\rangle_g \, \eta_i \in H^p(M).$$

By Theorem 12.4 there exists $\omega \in W^{1,2}(\Lambda^pTM)$, $\omega \perp_{L^2(g)} H^p(M)$, solving

$$\Delta_g \omega = I(\phi - \eta).$$

By Theorem 12.5 we get $\omega \in W^{2,2}(\Lambda^pTM)$. As discussed in (12.17) the equation then holds in the strong sense, that is we have

$$d_g^* d\omega + dd_g^*\omega = \phi - \eta \quad \text{ in } L^2(g).$$

Now put $\alpha = d_g^*\omega \in W^{1,2}(\Lambda^{p-1}TM)$ and $\beta = d\omega \in W^{1,2}(\Lambda^{p+1}TM)$, so that $\omega = d\alpha + d_g^*\beta + \eta$. Moreover from (12.24) and (12.21) we get

$$\begin{aligned}\|\alpha\|_{W^{1,2}(M)} + \|\beta\|_{W^{1,2}(M)} &\leq C\,\|\omega\|_{W^{2,2}(M)}\\ &\leq C\|I(\phi-\eta)\|\\ &\leq C\,\|\phi - \eta\|_{L^2(g)} \leq C\,\|\phi\|_{L^2(g)}.\end{aligned}$$

In the last step we used $\|\eta\|_{L^2(g)} \leq \|\phi\|_{L^2(g)}$, as η is defined by $L^2(g)$ projection. Now let $\gamma_k \in W^{1,2}(\Lambda^{p-1}TM)$ such that $d\gamma_k \to \xi$ in $L^2(\Lambda^pTM)$. We decompose

$$d\gamma_k = d\alpha_k + d_g^*\beta_k + \eta_k.$$

By orthogonality $d_g^*\beta_k = 0$ and $\eta_k = 0$. Furthermore we can estimate

$$\|\alpha_k\|_{W^{1,2}(M)} \leq C\,\|d\gamma_k\|_{L^2(g)} \leq C.$$

By Theorem 11.3 there exists $\alpha \in W^{1,2}(\Lambda^{p-1}TM)$ such that for a subsequence $\alpha_k \to \alpha$ weakly in $W^{1,2}(\Lambda^{p-1}TM)$. By Rellich 11.6 the convergence is also in $L^2(g)$. We conclude that for any form $\varphi \in C^\infty(\Lambda^pTM)$

$$\langle \xi, \varphi\rangle_{L^2(g)} = \lim_{k\to\infty} \langle d\gamma_k, \varphi\rangle_{L^2(g)} = \lim_{k\to\infty} \langle d\alpha_k, \varphi\rangle_{L^2(g)} = \langle d\alpha, \varphi\rangle_{L^2(g)}.$$

We conclude $\xi = d\alpha \in dW^{1,2}(\Lambda^{p-1}TM)$. The fact that $d_g^*W^{1,2}(\Lambda^{p+1}TM)$ is L^2 closed goes along the same lines. $\square$

Corollary 12.1 *For any $\phi \in W^{1,2}(\Lambda^p TM)$ with $d\phi = 0$ there is a unique $\eta \in H^p(M)$ with*

$$\phi - \eta \in dW^{1,2}(\Lambda^{p-1} TM).$$

Thus η is a harmonic representative in the cohomology class $[\phi]$ of ϕ.

Proof. Consider the Hodge decomposition

$$\phi = d\alpha + d_g^* \beta + \eta.$$

Testing with $d_g^*\beta$ yields, by orthogonality and since ϕ is closed,

$$0 = \langle \phi, d_g^*\beta \rangle_{L^2(g)} = \langle d_g^*\beta, d_g^*\beta \rangle_{L^2(g)}.$$

This shows $\phi - \eta = d\alpha$ as desired. For uniqueness assume that $\phi = d\alpha_1 + \eta_1 = d\alpha_2 + \eta_2$. Then $d(\alpha_1 - \alpha_2) + \eta_1 - \eta_2 = 0$, thus by orthogonality $\eta_1 = \eta_2$. □

We finally address the Hodge decomposition of forms $\phi \in C^\infty(\Lambda^* TM)$. We need the following generalization of the $W^{2,2}$-regularity.

Theorem 12.7 (higher regularity) *Let g be a C^{k+2} metric, $k \in \mathbb{N}_0$, on a compact manifold M. Let $\omega \in W^{1,2}(\Lambda^* TM)$ be a solution of*

$$\Delta_g \omega = I\phi \quad \text{where } \phi \in W^{k,2}(\Lambda^* TM).$$

Then $\omega \in W^{k+2,2}(\Lambda^* TM)$, and

$$\|\omega\|_{W^{k+2,2}(M)} \le C \left(\|\phi\|_{W^{k,2}(g)} + \|\omega\|_{L^2(g)} \right) \quad \text{where } C = C(M, g, k). \tag{12.24}$$

If g is a C^∞ metric and $\phi \in C^\infty(\Lambda^* TM)$, then also $\omega \in C^\infty(\Lambda^* TM)$.

The Sobolev embedding theorem implies that any $\omega \in W^{k,2}(\Lambda^* TM)$ has a representative in $C^j(\Lambda^* TM)$ if $k > j + \frac{n}{2}$, and the induced embedding of $W^{k,2}(\Lambda^* TM)$ into $C^j(\Lambda^* TM)$ is continuous. In particular the last statement regarding smoothness of ω follows form the $W^{k,2}$ regularity.

Corollary 12.2 (smooth Hodge decompostion) *Let g be a C^∞ metric on a compact manifold M. Then the following hold.*

(a) *Harmonic forms are smooth, i.e. $H^p(M) \subset C^\infty(\Lambda^p TM)$.*

(b) *Let $\phi = d\alpha + d_g^*\beta + \eta$ be the Hodge decomposition of $\phi \in C^\infty(\Lambda^p TM)$. Then $\alpha \in C^\infty(\Lambda^{p-1} TM)$ and $\beta \in C^\infty(\Lambda^{p+1} TM)$.*

Proof. For a harmonic form $\eta \in W^{1,2}(\Lambda^p TM)$ we have $\Delta_g \eta = 0$, thus $\eta \in C^\infty(\Lambda^p TM)$ is immediate from Theorem 12.7. Moreover by construction of the Hodge decomposition

$$\alpha = d_g^*\omega,\ \beta = d\omega \quad \text{where } \Delta_g \omega = I(\phi - \eta).$$

Thus $\omega \in C^\infty(\Lambda^p TM)$ again by Theorem 12.7, and claim (b) follows. □

13 Riemannian connection and curvature

The first goal in this section is to construct a canonical covariant derivative D, the Levi-Civita connection, acting on vector fields on any Riemannian manifold (M, g). We will ask that D is a metric connection, i.e. for $f \in C^\infty(M)$ and $X, Y, Z \in C^\infty(TM)$ we have the product rules

$$\begin{aligned} D_X(fY) &= (D_X f)\, Y + f\, D_X Y, \\ D_X[g(Y,Z)] &= g(D_X Y, Z) + g(Y, D_X Z). \end{aligned}$$

Here we write $D_X f = df(X)$ for convenience. As discussed in Theorem 9.1 and subsequent remark, such a connection exists always, however it is not unique. If D is one choice then all such connections are given by $D + A$ where $A \in C^\infty(T^*M \otimes \mathrm{End}\,(TM))$ is skewsymmetric. To make D unique a further property will be needed. We say that D is symmetric if

$$D_X Y - D_Y X = [X, Y] \quad \text{for all } X, Y \in C^\infty(TM). \tag{13.25}$$

Clearly this makes only sense in the case of connections on TM. The standard derivative $D_X Y = \sum_{i=1}^n X^i \partial_i Y^j\, e_j$ on $\mathbb{R}^n$ has this property, cf. Theorem 4.3.

Theorem 13.1 (Levi-Civita) *Let M be a differentiable manifold with smooth Riemannian metric g. Then on TM there is a unique connection D which is metric and symmetric. If we denote its coordinate representation by*

$$D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \sum_{k=1}^n \Gamma_{ij}^k \frac{\partial}{\partial x^k}, \tag{13.26}$$

then the coefficients are given by Christoffel's formula

$$\Gamma_{ij}^k = \frac{1}{2} \sum_{l=1}^n g^{kl} \big(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij} \big). \tag{13.27}$$

Proof. For a metric connection we have the identities,

$$\begin{aligned} \partial_X g(Y,Z) &= g(D_X Y, Z) + g(Y, D_X Z) \\ \partial_Y g(Z,X) &= g(D_Y Z, X) + g(Z, D_Y X) \\ -\,\partial_Z g(X,Y) &= -\,g(D_Z X, Y) - g(X, D_Z Y). \end{aligned}$$

Adding the equations and using that D is symmetric yields

$$\begin{aligned} 2g(D_X Y, Z) &= \partial_X g(Y,Z) + \partial_Y g(Z,X) - \partial_Z g(X,Y) \\ &\quad + g([X,Y], Z) + g([Z,X], Y) - g([[Y,Z], X). \end{aligned}$$

The right hand side is independent of D which shows uniqueness. Now for a given coordinate patch define a local connection, with Γ_{ij}^k given by (13.27),

$$D_X Y = \sum_{i=1}^n X^i D_{\frac{\partial}{\partial x^i}} \Big(\sum_{j=1}^n Y^j \frac{\partial}{\partial x^j} \Big) = \sum_{i,j=1}^n X^i \partial_i Y^j \frac{\partial}{\partial x^j} + \sum_{i,j,k=1}^n X^i Y^j \Gamma_{ij}^k \frac{\partial}{\partial x^k}.$$

D is symmetric since $\Gamma^k_{ij} = \Gamma^k_{ji}$. Furthermore

$$\begin{aligned} g\Big(D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^l}\Big) &= \sum_{k=1}^n g_{kl}\,\Gamma^k_{ij} \\ &= \frac{1}{2}\sum_{k,m=1}^n g_{kl}\,g^{km}(\partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}) \\ &= \frac{1}{2}(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}). \end{aligned}$$

Interchanging j, l and adding yields

$$\partial_i g_{jl} = g\Big(D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^l}\Big) + g\Big(\frac{\partial}{\partial x^j}, D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^l}\Big).$$

This is the metric property of D with respect to the coordinate fields, the case of general vector fields follows easily by expanding. We have shown that (13.27) defines a local connection which is symmetric and metric. Then by uniqueness, the definitions coincide on the overlaps, so that D is globally well-defined. □

Christoffel's formula (13.27) can be deduced by taking in the uniqueness argument coordinate vector fields $X = \frac{\partial}{\partial x^i}$, $Y = \frac{\partial}{\partial x^j}$ and $Z = \frac{\partial}{\partial x^l}$. This yields

$$\sum_{m=1}^n g_{ml}\,\Gamma^l_{ij} = \frac{1}{2}(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}).$$

Multiplication with g^{kl} and summation for $l = 1, \ldots, n$ gives (13.27). The definition of the Christoffel symbols and the covariant derivative was probably inspired by the first variation formula of the arclength functional, where the term $D_{c'}c'$ occurs as Euler-Lagrange operator.

The Levi-Civita connection induces metric conections on all bundles derived from TM; these are obtained by imposing a product rule. Instead of developing this in general, we consider only the case of one-forms $\omega \in C^\infty(T^*M)$. We put

$$(D_X\omega)(Y) = D_X(\omega(Y)) - \omega(D_X Y) \quad \text{for } \omega \in C^\infty(T^*M).$$

The product rule $D_X(f\omega) = (D_X f)\omega + f D_X\omega$ is easily checked. Moreover, $D\omega$ is linear with respect to functions in both entries X, Y, therefore it is a section of $T^{2,0}M$. To show the metric property, we choose almost Euclidean coordinates at center $p \in M$, see Lemma 12.4. At p we then have $g_{ij} = \delta_{ij}$, $\partial_i g_{jk} = 0$ and in particular $\Gamma^k_{ij} = 0$. Now by definition

$$\big(D_{\frac{\partial}{\partial x^i}} dx^j\big)\big(\frac{\partial}{\partial x^k}\big) = D_{\frac{\partial}{\partial x^i}}\Big(dx^j\big(\frac{\partial}{\partial x^k}\big)\Big) - dx^j\big(D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^k}\big) = 0.$$

We conclude, again at p,

$$D_{\frac{\partial}{\partial x^i}} g\big(dx^j, dx^k\big) = D_{\frac{\partial}{\partial x^i}} g^{jk} = 0 = g\big(D_{\frac{\partial}{\partial x^i}} dx^j, dx^k\big) + g\big(dx^j, D_{\frac{\partial}{\partial x^i}} dx^k\big).$$

The general case follows. Taking $\omega = Df$ where $f \in C^2(M)$ leads to the Hessian

$$D^2 f(X, Y) = (D_X(Df))(Y) = D_X(D_Y f) - D_{D_X Y} f,$$

in coordinates $D^2_{ij} f = \partial^2_{ij} f - \Gamma^k_{ij}\,\partial_k f$. We see that the Hessian is symmetric in X, Y. Moreover

$$\Delta_g f = \mathrm{tr}_g D^2 f = \sum_{i,j=1}^n g^{ij} \partial^2_{ij} f - \sum_{i,j,k=1}^n g^{ij} \Gamma^k_{ij}\,\partial_k f.$$

To show the identity, we use again almost Euclidean coordinates. We have at the center p

$$\Delta_g f = \frac{1}{\sqrt{\det g}} \sum_{i,j=1}^n \partial_i\big(\sqrt{\det g}\, g^{ij} \partial_j f\big) = \sum_{i=1}^n \partial^2_i f = \mathrm{tr}_g D^2 f.$$

Next we consider the second covariant derivative of vector fields, defined by $D^2_{X,Y} Z = D_X(D_Y Z) - D_{D_X Y}$. In contrast to the case of functions, this derivative is in general not symmetric with respect to X, Y. In fact we have the following

Definition 13.1 (Riemann curvature tensor) *Let g be a C^2 Riemannian metric on the differentiable manifold M. The Riemann curvature tensor of g is defined by*

$$R(X,Y)Z = D^2_{X,Y} Z - D^2_{Y,X} Z \quad \textit{where } X, Y, Z \in C^\infty(TM). \tag{13.28}$$

The curvature measures the asymmetry of the second derivative, which is quite natural from the viewpoint of analysis. However the geometric meaning of the curvature is a priori mysterious, its understanding is the central object of Riemnnian geometry. A few simple remarks:

(0) The Riemann curvature vanishes in the case $\dim M = 1$.

(1) Using $D^2_{X,Y} Z = D_X(D_Y Z) - D_{D_X Y} Z$ and recalling $D_X Y - D_Y X = [X,Y]$ by symmetry, we obtain the alternative formula

$$R(X,Y)Z = D_X(D_Y Z) - D_Y(D_X Z) - D_{[X,Y]} Z. \tag{13.29}$$

(2) Definition 13.1 yields in fact a tensor, more precisely a section of the bundle $T^{3,1}M$. For this one needs to show that R is linear with respect to the three entries X, Y, Z. As regards X, Y this is clear. The property for Z follows by straightforward calculation, assuming without loss of genrality that $D_X Y = D_Y X = 0$ at the given point:

$$\begin{aligned} R(X,Y)(fZ) &= D_X(D_Y(fZ)) - D_Y(D_X(fZ)) \\ &= D_X((D_Y f)Z + f D_Y Z) - D_Y((D_X f)Z + f D_X Z) \\ &= (D^2_{X,Y} f)Z - (D^2_{Y,X} f)Z + f(D^2_{X,Y} Z - D^2_{Y,X} Z) \\ &= f\, R(X,Y)Z. \end{aligned}$$

Here the symmetry of the Hessian of f was used.

(3) With respect to local coordinates we obtain the representation

$$R\Big(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\Big)\frac{\partial}{\partial x^k} = \sum_{l=1}^l R^l_{ijk} \frac{\partial}{\partial x^l}, \quad \text{where}$$

$$R^l_{ijk} = \partial_i \Gamma^l_{jk} - \partial_j \Gamma^l_{ik} + \sum_{m=1}^n \big(\Gamma^l_{im}\Gamma^m_{jk} - \Gamma^l_{jm}\Gamma^m_{ik}\big). \tag{13.30}$$

Here the Γ^k_{ij} are the Christoffel symbols, see (13.27).

(4) In Chapter 9 the curvature of a general vector bundle was introduced. In that context R is regarded as a 2-form taking values in $\mathrm{End}(TM)$, i.e. a section of $\Lambda^2 TM \otimes \mathrm{End}(TM)$. The connection one-form with respect to the coordinate basis is given by

$$A_j^k = \sum_{i=1}^n \Gamma_{ij}^k \, dx^i, \qquad \text{since } D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \sum_{k=1}^n \Gamma_{ij}^k \frac{\partial}{\partial x^k}.$$

The local representation of the curvature derived in (3) follows alternatively with the formula $R = dA + A \wedge A$. If R vanishes identically then parallel translation in TM is a homotopy invariant, a first relevant geometric information. Although D is metric the connection matrix A_j^k is in general not skew-symmetric. This is of course due to the fact that the coordinate vectorfields are not orthonormal. In can be useful to introduce a local orthonormal frame, this is always possible by Gram-Schmidt. This approach was systematically developed by S.S. Chern.

(5) The curvature tensor may also be regarded as a section of $T^{4,0}M$ by putting

$$R(X,Y,Z,W) = g(R(X,Y)Z,W). \tag{13.31}$$

Lemma 13.1 (Symmetries of the curvature tensor) *The following relations hold.*

(a) $R(X,Y,Z,W) = -R(Y,X,Z,W) = -R(X,Y,W,Z)$,

(b) $R(X,Y,Z,W) = R(Z,W,X,Y)$,

(c) $R(X,Y)Z + R(Z,X)Y + R(Y,Z)X = 0$ *(1. Bianchi identity).*

Proof. The first equation in (a) is clear. For the second identity we show that $R(X,Y)$ is skew-symmetric. Assuming that X, Y, Z have covariant derivative zero at $p \in M$, we compute

$$\begin{aligned} g(D_X(D_Y Z), Z) - g(D_{D_X Y} Z, Z) &= D_X[g(D_Y Z, Z)] \\ &= \frac{1}{2} D_X\big(D_Y[g(Z,Z)]\big) \\ &= \frac{1}{2} D^2_{X,Y}[g(Z,Z)]. \end{aligned}$$

Interchanging X, Y and subtracting proves $g(R(X,Y)Z,Z) = 0$. For (c) we note

$$\begin{aligned} R\Big(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\Big)\frac{\partial}{\partial x^k} &= D_{\frac{\partial}{\partial x^i}}\Big(D_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k}\Big) - D_{\frac{\partial}{\partial x^j}}\Big(D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^k}\Big) \\ &= D_{\frac{\partial}{\partial x^i}}\Big(D_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k}\Big) - D_{\frac{\partial}{\partial x^j}}\Big(D_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^i}\Big). \end{aligned}$$

Under cyclic permutation all terms cancel. Finally (b) follows from (a) and (c):

$$\begin{aligned} 2\,R(X,Y,Z,W) &= R(X,Y,Z,W) + R(Y,X,W,Z) \\ &= -R(Y,Z,X,W) - R(Z,X,Y,W) - R(X,W,Y,Z) - R(W,Y,X,Z) \\ &= -R(Z,Y,W,X) - R(X,Z,W,Y) - R(W,X,Z,Y) - R(Y,W,Z,X) \\ &= R(W,Z,Y,X) + R(Z,W,X,Y) \\ &= 2\,R(Z,W,X,Y). \end{aligned}$$

□

To quantify the curvature one has to extract scalar information. Let X, Y be a basis of a two-dimensional subspace $E \subset T_pM$, and denote by $\|X \wedge Y\|_g$ the area of the parallelogram spanned by X, Y. If $Z = aX + cY$, $W = bX + dY$ is another basis, then $\|Z \wedge W\|_g = |ad - bc|\, \|X \wedge Y\|_g$, and we see by expanding

$$\frac{R(Z,W,W,Z)}{\|Z \wedge W\|_g^2} = \frac{R(X,Y,Y,X)}{\|X \wedge Y\|_g^2}.$$

Thus the following definition makes sense.

Definition 13.2 *The sectional curvature of a two-dimensional subspace E of T_pM is*

$$K(E) = \frac{R(X,Y,Y,X)}{\|X \wedge Y\|_g^2} \qquad \text{where } X, Y \text{ is any basis of } E.$$

Clearly $\|v_1 \wedge v_2\|_g = 1$ for an orthonormal basis v_1, v_2 of E. Writing a general basis as $X = av_1 + cv_2$, $Y = bv_1 + dv_2$ it is easy to check that

$$\|X \wedge Y\|_g^2 = g(X,X)g(Y,Y) - g(X,Y)^2.$$

The sectional curvature determines the full Riemann curvature tensor, in fact any tensor $T(X,Y,Z,W)$ having the symmetries as in Lemma 13.1 is determined by the values of $T(X,Y,Y,X)$. To see this, denote by $c(X,Y,Z,W)$ the coefficient of $\alpha\beta$ in the polynomial $T(X + \alpha Z, Y + \beta W, Y + \beta W, X + \alpha Z)$. We compute

$$\begin{aligned} c(X,Y,Z,W) &= T(Z,W,Y,X) + T(Z,Y,W,X) + T(X,W,Y,Z) + T(X,Y,W,Z) \\ &= 2\,T(Y,Z,X,W) - 2\,T(X,Y,Z,W). \end{aligned}$$

Antisymmetrizing the first term with respect to Z, W, we get using the Bianchi identity

$$T(Y,Z,X,W) - T(Y,W,X,Z) = T(Y,Z,X,W) + T(Z,X,Y,W) = -T(X,Y,Z,W)$$

Therefore we obtain

$$\begin{aligned} c(X,Y,Z,W) - c(X,Y,W,Z) &= 2\,(T(Y,Z,X,W) - T(Y,W,X,Z)) - 4T(X,Y,Z,W) \\ &= -6\,T(X,Y,Z,W). \end{aligned}$$

Example 13.1 Assume that the sectional curvature of (M,g) is pointwise konstant, i.e. there exists a function $\varkappa : M \to \mathbb{R}$ such that $K(E) = \varkappa(p)$ for any 2-plane $E \subset T_pM$. We claim

$$R(p)(X,Y,Z,W) = \varkappa(p)\,\big(g(X,W)g(Y,Z) - g(X,Z)g(Y,W)\big) \quad \text{for any } X,Y,Z,W \in T_p(M).$$

By definition of the sectional curvature, we know that

$$R(p)(X,Y,Y,X) = \varkappa(p)\,\big(g(X,X)g(Y,Y) - g(X,Y)g(Y,X)\big).$$

One checks that the tensor on the right has the symmetries as in Lemma 13.1, hence the claim follows. In two dimensions there is generally just one sectional curvature function, the formula applies. If the metric g is induced by an immersion into $\mathbb{R}^3$ then the function $\varkappa$ is the Gauß curvature, i.e. the product of the principal curvatures. This is the content of the Theorema egregium of Gauß.

Example 13.2 A Riemannian manifold (M,g) with $K(E) = \varkappa \in \mathbb{R}$, for all 2-planes $E \subset T_pM$ and all $p \in M$, is called a space form. By the above we have

$$R(X,Y)Z = \varkappa\big(g(Y,Z)X - g(X,Z)Y\big).$$

The model cases are $\mathbb{R}^n$ with $\varkappa = 0$, $\mathbb{S}^n$ with $\varkappa = 1$ and $\mathbb{H}^n$ with $\varkappa = -1$.

As we have seen, the sectional curvature function determines the full Riemann curvature tensor. Instead, one can take averages to produce more coarse curvature invariants.

Definition 13.3 *Let (M,g) be a Riemannian manifold. The Riccitensor $\mathrm{Ric} \in C^\infty(\mathrm{End}\, TM)$ and the scalar curvature $S \in C^\infty(M)$ are defined by*

$$\mathrm{Ric}(p)X = \sum_{j=1}^{n} R(X,v_j)v_j, \tag{13.32}$$

$$S(p) = \mathrm{tr}\,\mathrm{Ric}(p) = \sum_{i,j=1}^{n} R(v_i,v_j,v_j,v_i). \tag{13.33}$$

Here $v_1,\ldots,v_n$ is an orthonormal basis of T_pM. One introduces also the quadratic form

$$\mathrm{ric}(p)(X,X) = g(\mathrm{Ric}(p)X,X).$$

If the curvature is pointwise equal to a constant $\varkappa$, then

$$\mathrm{Ric} = (n-1)\varkappa\,\mathrm{Id} \quad \text{and} \quad S = n(n-1)\varkappa.$$

Example 13.3 It useful to check how the curvature quantities change under a scaling $g_\lambda = \lambda^2 g$ for $\lambda > 0$. We see from (13.27) that the Christoffel symbols remain unchanged and hence $D_\lambda = D$. It follows that $R_\lambda(X,Y)Z = R(X,Y)Z$. This implies

$$K_\lambda = \lambda^{-2}K, \quad \mathrm{Ric}_\lambda = \lambda^{-2}\mathrm{Ric}, \quad \mathrm{ric}_\lambda = \mathrm{ric}, \quad S_\lambda = \lambda^{-2}S.$$

14 The Yamabe problem

Two Riemannian metrics $g, \tilde{g}$ on M are conformal if there is a function $u : M \to \mathbb{R}$ such that $\tilde{g} = e^{2u}g$. The factor is written as an exponential just for convenience. We denote by $[g]$ the corresponding equivalence class of conformal metrics. A classic theorem states that any compact, two-dimensional manifold (M,g) has a conformal metric $\tilde{g}$ of constant curvature. We briefly explain how this follows from complex analysis, assuming that M is oriented. Locally M has conformal parametrizations, hence it becomes a one-dimensional complex manifold. Then by the uniformization theorem the universal covering is biholomorphic to either $\mathbb{C}$, $\mathbb{D}$ or $\mathbb{S}^2$, and a constant curvature metric is obtained as a quotient. A special case of the uniformization theorem is the Riemann mapping theorem.

Here we consider a compact manifold of dimension $n \geq 3$. Since the Riemann curvature tensor has many components, we cannot expect to find a constant curvature metric by only a conformal change. However, it seems reasonable to ask whether there is a conformal metric which has constant scalar curvature, this is the Yamabe problem. To find the analytic formulation of the problem, we first need to calculate how the geometric quantities are related for two conformal metrics. Recall our convention to write $D_X u$ instead of $du(X)$.

Lemma 14.1 *Let $\tilde{g} = e^{2u}g$ be conformal metrics on an n-dimensional manifold. The following relations hold:*

$$
\begin{aligned}
&(14.34) \qquad & d\mu_{\tilde{g}} &= e^{nu}\, d\mu_g, \\
&(14.35) \qquad & \mathrm{grad}_{\tilde{g}}\varphi &= e^{-2u}\mathrm{grad}_g\varphi, \\
&(14.36) \qquad & \mathrm{div}_{\tilde{g}}X &= \mathrm{div}_g X + nD_X u, \\
&(14.37) \qquad & \Delta_{\tilde{g}}\varphi &= e^{-2u}\big(\Delta_g\varphi + (n-2)\langle Du, D\varphi\rangle_g\big), \\
&(14.38) \qquad & D^{\tilde{g}}_X Y &= D_X Y + (D_X u)Y + (D_Y u)X - g(X,Y)\mathrm{grad}_g u.
\end{aligned}
$$

Next we collect the formulae for the curvature, starting with the full curvature tensor of type $(4,0)$. See below for the definition of $g \wedge T$, the Kulkarni-Nomizu product.

$$
\begin{aligned}
&(14.39) \qquad & R_{\tilde{g}} &= e^{2u}(R_g + g \wedge T) \qquad \textit{where } T = D^2u - du \otimes du + \frac{1}{2}|Du|_g^2\, g, \\
&(14.40) \qquad & \mathrm{ric}_{\tilde{g}} &= \mathrm{ric}_{\tilde{g}} - (n-2)D^2u + (n-2)Du \otimes Du - (\Delta_g u + (n-2)|Du|_g^2)\, g, \\
&(14.41) \qquad & S_{\tilde{g}} &= e^{-2u}\big(S_g - 2(n-1)\Delta_g u - (n-1)(n-2)|Du|_g^2\big).
\end{aligned}
$$

Proof. We have $\tilde{g}^{ij} = e^{-2u}g^{ij}$ and $\sqrt{\det \tilde{g}} = e^{nu}\sqrt{\det g}$. Using this we get easily

$$
\begin{aligned}
d\mu_{\tilde{g}} &= \sqrt{\det \tilde{g}}\, dx = e^{nu}\, d\mu_g, \\
\mathrm{grad}_{\tilde{g}}\varphi &= \tilde{g}^{ij}\partial_i\varphi \frac{\partial}{\partial x^j} = e^{-2u}\mathrm{grad}_{\,g}\varphi, \\
\mathrm{div}_{\tilde{g}}X &= \frac{1}{\sqrt{\det \tilde{g}}}\,\partial_i(\sqrt{\det \tilde{g}}\, X^i) = \mathrm{div}_g X + nD_X u.
\end{aligned}
$$

Formula (14.37) follows from (14.35) and (14.36). Next, for the Christoffel symbols

$$
\begin{aligned}
(\Gamma_{\tilde{g}})^k_{ij} &= \frac{1}{2}\tilde{g}^{kl}(\partial_i\tilde{g}_{jl} + \partial_j\tilde{g}_{il} - \partial_l\tilde{g}_{ij}) \\
&= \Gamma^k_{ij} + (\partial_i u)\delta_{jk} + (\partial_j u)\delta_{ik} - (\mathrm{grad}_g u)^k g_{ij}.
\end{aligned}
$$

Formula (14.38) follows. For the curvature identities we start with the second derivative

$$
\begin{aligned}
\tilde{D}_X(\tilde{D}_Y Z) &= D_X\big(D_Y Z + (D_Y u)Z + (D_Z u)Y - g(Y,Z)\mathrm{grad}_g u\big) \\
&\quad + (D_X u)\big(D_Y Z + (D_Y u)Z + (D_Z u)Y - g(Y,Z)\,\mathrm{grad}_g u\big) \\
&\quad + Du \cdot \big(D_Y Z + (D_Y u)Z + (D_Z u)Y - g(Y,Z)\mathrm{grad}_g u\big)\, X \\
&\quad - g\big(X, D_Y Z + (D_Y u)Z + (D_Z u)Y - g(Y,Z)\mathrm{grad}_g u\big)\,\mathrm{grad}_g u.
\end{aligned}
$$

We simplify this a little by assuming that X, Y, Z have zero derivative for D at a given point.

$$
\begin{aligned}
\tilde{D}_X(\tilde{D}_Y Z) &= D^2_{X,Y}Z + (D^2_{X,Y}u)Z + (D^2_{X,Z}u)Y - g(Y,Z)D_X\mathrm{grad}_g u \\
&\quad + (D_X u)(D_Y u)Z + (D_X u)(D_Z u)Y - g(Y,Z)(D_X u)\mathrm{grad}_g u \\
&\quad + 2(D_Y u)(D_Z u)X - g(Y,Z)|Du|_g^2\, X \\
&\quad - g(X,Z)(D_Y u)\mathrm{grad}_g u - g(X,Y)(D_Z u)\mathrm{grad}_g u + g(Y,Z)(D_X u)\mathrm{grad}_g u.
\end{aligned}
$$

Now we antisymmetrize with respect to X, Y. As $[X,Y]=0$ at the given point we obtain

$$\begin{aligned} R_{\tilde g}(X,Y)Z &= R(X,Y)Z + (D^2_{X,Z}u)Y - (D^2_{Y,Z}u)X - g(Y,Z)D_X\mathrm{grad}_g u + g(X,Z)D_Y\mathrm{grad}_g u \\ &\quad +(D_Xu)(D_Zu)Y - (D_Yu)(D_Zu)X - g(Y,Z)(D_Xu)\mathrm{grad}_g u + g(X,Z)(D_Yu)\mathrm{grad}_g u \\ &\quad +2(D_Yu)(D_Zu)X - 2(D_Xu)(D_Zu)Y - g(Y,Z)|Du|_g^2 X + g(X,Z)|Du|_g^2 Y \\ &\quad -g(X,Z)(D_Yu)\mathrm{grad}_g u + g(Y,Z)(D_Xu)\mathrm{grad}_g u + g(Y,Z)(D_Xu)\mathrm{grad}_g u - g(X,Z)(D_Yu)\mathrm{gra} \\ &= R(X,Y)Z + (D^2_{X,Z}u)Y - (D^2_{Y,Z}u)X - g(Y,Z)D_X\mathrm{grad}_g u + g(X,Z)D_Y\mathrm{grad}_g u \\ &\quad -(D_Xu)(D_Zu)Y + (D_Yu)(D_Zu)X + g(Y,Z)(D_Xu)\mathrm{grad}_g u - g(X,Z)(D_Yu)\mathrm{grad}_g u \\ &\quad -g(Y,Z)|Du|_g^2 X + g(X,Z)|Du|_g^2 Y. \end{aligned}$$

Choosing $Z = Y$ and taking the g-scalar product with X we get

$$\begin{aligned} e^{-2u}R_{\tilde g}(X,Y,Y,X) &= R(X,Y,Y,X) + g(X,Y)(D^2_{X,Y}u) - g(X,X)(D^2_{Y,Y}u) \\ &\quad -g(Y,Y)D^2_{X,X}u + g(X,Y)D^2_{X,Y}u - g(X,Y)(D_Xu)(D_Yu) + g(X,X)(D_Yu)^2 \\ &\quad +g(Y,Y)(D_Xu)^2 - g(X,Y)(D_Xu)(D_Yu) \\ &\quad -g(X,X)g(Y,Y)|Du|_g^2 + g(X,Y)^2|Du|_g^2 \\ &= R(X,Y,Y,X) \\ &\quad +2g(X,Y)\big(D^2_{X,Y}u - (D_Xu)(D_Yu) + 1/2\, g(X,Y)|Du|_g^2\big) \\ &\quad -g(X,X)\big(D^2_{Y,Y}u - (D_Yu)^2 + 1/2\, g(Y,Y)|Du|_g^2\big) \\ &\quad -g(Y,Y)\big(D^2_{X,X}u - (D_Xu)^2 + 1/2\, g(X,X)|Du|_g^2\big). \end{aligned}$$

Following Kulkarni-Nomizu one introduces for symmetric bilinear forms a, b the product

$$\begin{aligned} (a\wedge b)(X,Y,Z,W) &= a(X,Z)b(Y,W) - a(X,W)b(Y,Z) \\ &\quad + b(X,Z)a(Y,W) - b(X,W)a(Y,Z). \end{aligned}$$

In particular $(a\wedge b)(X,Y,Y,X) = 2a(X,Y)b(X,Y) - a(X,X)b(Y,Y) - a(Y,Y)b(X,X)$, hence

(14.42) $e^{-2u}R_{\tilde g} = R_g(X,Y,Y,X) + (g\wedge T)(X,Y,Y,X), \quad$ for $T = D^2u - du\otimes du + \frac{1}{2}|Du|^2 g$.

As a symmetric bilinear form on $\Lambda_2(TM)$ By definition, $a\wedge b$ has the symmetries of type (a) and (b) in Lemma 13.1. Therefore (14.42) implies the formula for the full curvature $(4,0)$-tensor. Next we put $Y = e_j$ where $e_1, \dots, e_n$ is an orthonormal basis for g. By summation we have

$$\begin{aligned} \mathrm{ric}_{\tilde g}(X,X) &= \mathrm{ric}_g(X,X) + (g\wedge T)(X,e_j,e_j,X) \\ &= \mathrm{ric}_g(X,X) - (n-2)T(X,X) - g(X,X)\,\mathrm{tr}\,T \\ &= \mathrm{ric}_g(X,X) - (n-2)D^2u(X,X) + (n-2)Du(X)^2 \\ &\quad -g(X,X)\big(\Delta_g u + (n-2)|Du|^2\big). \end{aligned}$$

Finally we take $X = e_i$ and sum over i to obtain

$$\begin{aligned} e^{2u}S_{\tilde g} &= S_g - (n-2)\Delta_g u + (n-2)|Du|_g^2 - n\big(\Delta_g u + (n-2)|Du|^2\big) \\ &= S_g - 2(n-1)\Delta_g u - (n-1)(n-2)|Du|_g^2. \end{aligned}$$

The lemma is proved. $\square$

To make equation (14.41) a little nicer we substitute $e^{2u} = \varphi^{2\alpha}$ with $\varphi > 0$. Then

$$\begin{aligned}\varphi^{2\alpha} S_{\tilde g} &= S_g - 2(n-1)\Delta_g u - (n-1)(n-2)|Du|_g^2 \\ &= S_g - 2(n-1)\Big(\frac{\alpha}{\varphi}\Delta_g\varphi - \frac{\alpha}{\varphi^2}|D\varphi|^2\Big) - (n-1)(n-2)\frac{\alpha^2}{\varphi^2}|D\varphi|^2.\end{aligned}$$

By choosing $\alpha = \frac{2}{n-2}$ the gradient terms cancel, and we arrive at

$$L_g\varphi := -\frac{4(n-1)}{n-2}\Delta_g\varphi + S_g\varphi = S_{\tilde g}\varphi^{\frac{n+2}{n-2}} \quad \text{for } \tilde g = \varphi^{\frac{4}{n-2}} g. \tag{14.43}$$

The second order elliptic operator L_g is called the conformal Laplacian. Renaming φ into u the Yamabe problem takes the following form:

Find a constant $\lambda \in \mathbb{R}$ and a function $u > 0$ on M such that $L_g u = \lambda u^{\frac{n+2}{n-2}}$.

Remarkably, this has a variational interpretation. We introduce the functionals

$$Q_g(u) = \int_M \Big(\frac{4(n-1)}{n-2}|Du|_g^2 + S_g u^2\Big)\, d\mu_g, \tag{14.44}$$

$$V_g(u) = \int_M |u|^{\frac{2n}{n-2}}\, d\mu_g \quad \Big(\text{note that } \frac{2n}{n-2} = \frac{n+2}{n-2} + 1\Big). \tag{14.45}$$

The corresponding Euler-Lagrange operators are as follows:

$$\begin{aligned} DQ_g(u)\varphi &= 2\int_M (L_g u)\varphi\, d\mu_g, \\ DV_g(u)\varphi &= \frac{2n}{n-2}\int_M |u|^{\frac{4}{n-2}} u\,\varphi\, d\mu_g. \end{aligned}$$

Lemma 14.2 *Assume that $u \in C^\infty(M)$, $u > 0$, has the minimizing property*

$$Q_g(u) \le Q_g(u+v) \quad \textit{for all } v \in C^\infty(M),\ V_g(u+v) = V_g(u),\ \|v\|_{C^1(M)} \ll 1.$$

Then $L_g u = \lambda u^{\frac{n+2}{n-2}}$ for some $\lambda \in \mathbb{R}$.

Proof. Choose a smooth function φ with

$$DV_g(u)\varphi = \frac{2n}{n-2}\int_M u^{\frac{n+2}{n-2}}\varphi\, d\mu_g = 1.$$

For given v we consider $G(s,t) = V_g(u + sv + t\varphi)$. Then $G(0,0) = V_g(u)$ and $\partial_t G(0,0) = DV_g(u)\varphi = 1$. By the implicit function theorem, there exists locally a function $t = \tau(s)$, $\tau(0) = 0$, with

$$G(s,t) = V_g(u) \quad \Leftrightarrow \quad t = \tau(s).$$

The derivative of $\tau(s)$ is computed by the chain rule, we have

$$0 = \frac{d}{ds} G(s,\tau(s))_{s=0} = DV_g(u)(v + \tau'(0)\varphi) = DV_g(u)v + \tau'(0).$$

From the minimizing property we now conclude

$$\begin{aligned}
0 &= \frac{d}{ds} Q_g(u + sv + \tau(s)\varphi)|_{s=0} \\
&= DQ_g(u)v - DQ_g(u)\varphi(DV_g(u)v) \\
&= DQ_g(u)v - \lambda\, DV_g(u)v \quad \text{ where } \lambda = DQ_g(u)\varphi \in \mathbb{R}.
\end{aligned}$$

The claim follows by the fundamental lemma in the calculus of variations. □

Next we observe that the functionals Q_g and V_g have a geometric interpretation. Namely for $\tilde{g} = u^{\frac{4}{n-2}} g$ we have recalling (14.43)

$$\begin{aligned}
V_g(u) &= \int_M (u^{\frac{2}{n-2}})^n \, d\mu_g = \int_M d\mu_{\tilde{g}}, \\
Q_g(u) &= \int_M (L_g u) u \, d\mu_g = \int_M S_{\tilde{g}} \, u^{\frac{n+2}{n-2}+1} \, d\mu_g = \int_M S_{\tilde{g}} \, d\mu_{\tilde{g}}.
\end{aligned}$$

Hence Yamabe metrics are characterized as critical points of the total scalar curvature functional with precribed volume, however only in the class of metrics which are conformal. In relativity, the total scalar curvature is called the Einstein-Hilbert action. The foundations of general relativity were published by Einstein in 1915. Almost at the same time, Hilbert found that the vacuum Einstein equations are the Euler-Lagrange equations of the total scalar functional, however in the full space of metrics. Apparently Einstein was also looking for such a variational characterization, in connection with the coordinate invariance of his field equations.

Definition 14.1 (Yamabe invariant) *For a smooth metric g on M compact we define*

$$\begin{aligned}
\lambda([g]) &= \inf \big\{ Q_g(u); u \in C^\infty(M),\, V_g(u) = 1,\, u > 0 \big\} \\
&= \inf \Big\{ \int_M S_{\tilde{g}} \, d\mu_{\tilde{g}} : \tilde{g} \in [g],\, \mathrm{vol}_{\tilde{g}}(M) = 1 \Big\}.
\end{aligned}$$

The second equality shows that $\lambda([g])$ is a conformal invariant, i.e. it depends only on the conformal class of g. We note that

$$Q_g(u) \geq -\int_M S_- u^2 \, d\mu_g \geq -\Big(\int_M S_-^{\frac{n}{2}} \, d\mu_g \Big)^{\frac{2}{n}} \Big(\int_M u^{\frac{2n}{n-2}} \, d\mu_g \Big)^{\frac{n-2}{n}},$$

thus $\lambda([g]) \geq \|S_-\|_{L^{n/2}(\mu_g)} > -\infty$. We will see that $\lambda([g])$ plays a key role in the solution of the Yamabe problem.

Lemma 14.3 *We have*

$$\lambda([g]) = \inf \Big\{ \frac{Q_g(u)}{\|u\|_{L^p}^2} : u \in W^{1,2}(M), u \neq 0 \Big\}.$$

Proof. We must show that the right hand side is not smaller than the left hand side. Let $u \in W^{1,2}(M)$, $u \neq 0$, be given. In a first step we approximate u by functions $u_k \in C^\infty(M)$ in $W^{1,2}(M)$, see Lemma 12.1. The functional $Q_g(u)$ is continuous with respect to $W^{1,2}$-convergence. The same holds for the L^p-norm, by Sobolev embedding. Thus

$$\frac{Q_g(u)}{\|u\|_{L^p}^2} = \lim_{k \to \infty} \frac{Q_g(u_k)}{\|u_k\|_{L^p}^2}.$$

In the second step, let $u \in C^\infty(M)$, $u \neq 0$, be given. For $\varepsilon > 0$ consider the smooth functions $u^\varepsilon = \sqrt{u^2+\varepsilon^2} \geq \varepsilon > 0$. We have $u^\varepsilon \to |u|$ in $L^2(M)$, moreover

$$Du^\varepsilon = \frac{u}{\sqrt{u^2+\varepsilon^2}} Du \to \operatorname{sign}(u) Du = D|u| \quad \text{ in } L^2(M).$$

This implies

$$\frac{Q_g(u)}{\|u\|_{L^p}^2} = \frac{Q_g(|u|)}{\|u\|_{L^p}^2} = \lim_{\varepsilon \searrow 0} \frac{Q_g(u^\varepsilon)}{\|u^\varepsilon\|_{L^p}^2} \geq \lambda([g]).$$

Note that $\|u^\varepsilon\|_{L^p(M)} = 1$ can be achieved by scaling. Combining the steps yields the result. $\square$

The case $M = \mathbb{S}^n$ will be of particular relevance. Putting $\omega_n = \operatorname{vol}(\mathbb{S}^n)$ we have

$$\int_{\mathbb{S}^n} S_{g_{\mathbb{S}^n}} \, d\mu_{\mathbb{S}^n} = n(n-1)\omega_n, \qquad \text{hence } \lambda(\mathbb{S}^n) \leq n(n-1)\omega_n^{\frac{2}{n}}.$$

Consider the stereographic projections from the north and south pole

$$\varphi_\pm : \mathbb{S}^n \backslash \{(0, \pm 1)\} \to \mathbb{R}^n, \; \varphi_\pm(x,z) = \frac{x}{1 \mp z},$$

with corresponding inverse maps

$$\psi_\pm : \mathbb{R}^n \to \mathbb{S}^n \backslash \{(0, \pm 1)\}, \; \psi_\pm(x) = \Big(\frac{2x}{|x|^2+1}, \pm \frac{|x|^2-1}{|x|^2+1} \Big).$$

These are conformal diffeomorphisms, one computes that

$$\psi_\pm^* g_{\mathbb{S}^n} = u_1^{\frac{4}{n-2}} g_{\mathbb{R}^n} \qquad \text{where } u_1(x) = \Big(\frac{2}{|x|^2+1} \Big)^{\frac{n-2}{2}}.$$

On $\mathbb{R}^n$ there is the one-parameter family of dilations $\tau_\alpha(x) = \alpha x$ where $\alpha > 0$. These induce conformal diffeomorphisms $\sigma_\alpha : \mathbb{S}^n \to \mathbb{S}^n$ by

$$\sigma_\alpha(x,z) = \begin{cases} (\psi_+ \circ \tau_\alpha \circ \varphi_+)(x,z) & \text{for } (x,z) \neq (0,1), \\ (0,1) & \text{for } (x,z) = (0,1). \end{cases}$$

σ_α is smooth at the north pole. Namely for $x \neq 0$ we have $(\varphi_\pm \circ \psi_\mp)(x) = \frac{x}{|x|^2}$, and hence

$$\varphi_- \circ \sigma_\alpha \circ \psi_-(x) = (\varphi_- \circ \psi_+ \circ \tau_\alpha \circ \varphi_+ \circ \psi_-)(x) = \frac{1}{\alpha} x.$$

A short calculation yields $\sigma_\alpha^* g_{\mathbb{S}^n} = v_\alpha^{\frac{4}{n-2}} g_{\mathbb{S}^n}$ where

$$v_\alpha(x,z) = \frac{u_\alpha(y)}{u_1(y)}\Big|_{y=\varphi_+(x,z)} = \Big(\frac{2}{\alpha^{-1}(1-z) + \alpha(1+z)} \Big)^{\frac{n-2}{2}}.$$

As the metric is a pullback of the standard metric, we have

$$\int_{\mathbb{S}^n} v_\alpha^{\frac{2n}{n-2}} \, d\mu_{\mathbb{S}^n} = \omega_n \qquad \text{and} \qquad Q_{g_{\mathbb{S}^n}}(v_\alpha) = n(n-1)\omega_n,$$

in particular

$$\frac{Q_{g_{\mathbb{S}^n}}(v_\alpha)}{\|v_\alpha\|_{L^p}^2} = n(n-1)\omega_n^{\frac{2}{n}} \quad \text{for all } \alpha > 0.$$

However the family v_α, $\alpha > 0$, is not compact in $W^{1,2}(\mathbb{S}^n)$, in fact

$$v_\alpha^{\frac{2n}{n-2}} \mu_{\mathbb{S}^n} \to \omega_n \, \delta_{(0,\pm 1)} \quad \text{as } \alpha \searrow 0 \text{ resp. } \alpha \nearrow \infty.$$

Aubin and Talenti proved in 1976 that $v_1 \equiv 1$ is minimizing on $\mathbb{S}^n$, so that indeed

$$\lambda(\mathbb{S}^n) = n(n-1)\omega_n^{2/n}. \tag{14.46}$$

This is not unexpected, from the viewpoint that a minimizer should have maximal symmetry. However, by the above all the functions v_α are also minimizers, and they are less symmetric. Moreover, we see that minimizing sequences need not (sub-)converge, as even the actual minimizers v_α diverge as α goes to zero or infinity. This could be attributed to the fact that the group of conformal diffeomorphisms of $\mathbb{S}^n$ is not compact, as we have seen.

The key to the Yamabe problem is to show that this divergent scenario happens only in the single case when the manifold is globally conformally equivalent to $\mathbb{S}^n$. For the analysis it will be useful to relate the constant $\lambda(\mathbb{S}^n)$ to the corresponding constant on $\mathbb{R}^n$.

Definition 14.2 (Sobolev-Yamabe constant) *We put*

$$\lambda_n = \inf\Big\{\frac{Q_{\mathbb{R}^n}(u)}{\|u\|_{L^p(\mathbb{R}^n)}} : u \in W^{1,2}(\mathbb{R}^n),\, u \neq 0\Big\}.$$

Using stereographic projection it is easy to see that actually $\lambda_n = \lambda(\mathbb{S}^n)$. Moreover by the Gagliardo-Nirenberg inequality, see Theorem 11.5, we know that

$$\lambda_n \geq \frac{n-2}{n-1} > 0. \tag{14.47}$$

Of course this bound should not be expected to be optimal.

Lemma 14.4 *Let (M,g) be a compact Riemannian manifold. For any $\varepsilon > 0$ there exists a constant $C_\varepsilon = C_\varepsilon(M,g)$ such that*

$$\lambda_n \|u\|_{L^p(\mu_g)}^2 \leq (1+\varepsilon) \int_M \frac{4(n-1)}{n-2} |Du|_g^2 \, d\mu_g + C_\varepsilon \int_M u^2 \, d\mu_g. \tag{14.48}$$

Proof. For each $x \in M$ there exist almost Euclidean coordinates (U_x, φ_x) such that

$$\frac{1}{1+\varepsilon} \leq (g_{ij}) \leq (1+\varepsilon) \quad \text{and} \quad \frac{1}{1+\varepsilon} \leq \sqrt{\det g} \leq (1+\varepsilon).$$

Let U_i, $i \in I$, be a finite subcovering and η_i^2 a subordinate partition of unity. Using the definition of λ_n we can estimate

$$\begin{aligned}
\lambda_n \|u\|^2_{L^p(\mu_g)} &= \lambda_n \| \sum_{i=1}^n \eta_i^2 u^2 \|_{L^{\frac{p}{2}}(\mu_g)} \\
&\leq \lambda_n \sum_{i \in I} \Big(\int_{U_i} |\eta_i u|^p \, d\mu_g \Big)^{\frac{2}{p}} \\
&\leq (1 + C\varepsilon) \lambda_n \sum_{i \in I} \Big(\int_{U_i} |\eta_i u|^p \, dx \Big)^{\frac{2}{p}} \\
&\leq (1 + C\varepsilon) \sum_{i \in I} \int_{U_i} \frac{4(n-1)}{n-2} |D(\eta_i u)|^2 \, dx \\
&\leq (1 + C\varepsilon) \sum_{i \in I} \int_{U_i} \frac{4(n-1)}{n-2} |D(\eta_i u)|^2_g \, d\mu_g.
\end{aligned}$$

Now by the product rule

$$\begin{aligned}
|D(\eta_i u)|^2_g &= \eta_i^2 |Du|^2_g + |D\eta_i|^2_g u^2 + 2\eta_i u \langle D\eta_i, Du \rangle_g \\
&\leq (1 + \varepsilon) \eta_i^2 |Du|^2_g + C_\varepsilon u^2.
\end{aligned}$$

Inserting yields

$$\lambda_n \|u\|^2_{L^p(\mu_g)} \leq (1 + C\varepsilon) \int_M \frac{4(n-1)}{n-2} |Du|^2_g \, d\mu_g + C_\varepsilon(M, g) \int_M u^2 \, d\mu_g.$$

□

The following measure-theoretic lemma proves useful in the analysis of minimizing sequences. The lemma holds for any $p > 0$, but this is not needed here.

Lemma 14.5 (Brézis-Lieb 1983) *Let μ be a measure on M and $p \geq 2$. Assume that u_j is a bounded sequence in $L^p(\mu)$ which converges pointwise to $u \in L^p(\mu)$ μ-almost-everywhere. Writing $u_j = u + w_j$ we then have*

(14.49)
$$\lim_{j \to \infty} \int_M \big| |u_j|^p - (|u|^p + |w_j|^p) \big| \, d\mu = 0.$$

Moreover as a consequence

(14.50)
$$\limsup_{j \to \infty} \big(\|u_j\|^2_{L^p} - (\|u\|^2_{L^p} + \|w_j\|^2_{L^p}) \big) \leq 0.$$

Proof. By elementary calculus, for any $\varepsilon > 0$ there exists a constant $C_\varepsilon = C_\varepsilon(p) < \infty$ with

$$\big| |a + b|^p - |a|^p \big| \leq \varepsilon |a|^p + C_\varepsilon |b|^p \quad \text{for all } a, b \in \mathbb{R}.$$

For fixed $\varepsilon > 0$ consider now the functions

$$f_{\varepsilon,j} = \big(|u_j|^p - (|u|^p + |w_j|^p) - \varepsilon |w_j|^p \big)_+.$$

Clearly $f_{\varepsilon,j} \to 0$ pointwise almost everywhere. Moreover we have

$$\big| |u_j|^p - (|u|^p + |w_j|^p) \big| \leq \big| |w_j + u|^p - |w_j|^p \big| + |u|^p \leq \varepsilon |w_j|^p + C_\varepsilon |u|^p.$$

It follows that $0 \le f_{\varepsilon,j} \le C_\varepsilon |u|^p \in L^1(\mu)$, and hence by dominated convergence

$$\lim_{j\to\infty} \int f_{\varepsilon,j}\,d\mu = 0.$$

Now $\big||u_j|^p - (|u|^p + |w_j|^p)\big| \le f_{\varepsilon,j} + \varepsilon |w_j|^p$, and we conclude, sind w_j is bounded in $L^p(\mu)$,

$$\limsup_{j\to\infty} \int \big||u_j|^p - (|u|^p + |w_j|^p)\big|\,d\mu \le \limsup_{j\to\infty} \int \big(f_{\varepsilon,j} + \varepsilon|w_j|^p\big)\,d\mu \le C\varepsilon.$$

Letting $\varepsilon \searrow 0$ shows (14.49). This implies further, for a null sequence ε_j,

$$\|u_j\|_{L^p}^2 \le \Big(\int |u|^p\,d\mu + \int |w_j|^p\,d\mu + \varepsilon_j\Big)^{\frac{2}{p}} \le \Big(\int |u|^p\,d\mu\Big)^{\frac{2}{p}} + \Big(\int |w_j|^p\,d\mu\Big)^{\frac{2}{p}} + \varepsilon_j^{\frac{2}{p}},$$

which proves (14.50). $\square$

Theorem 14.1 (Aubin) *Let (M,g) be a compact Riemannian manifold. Recall that*

$$\lambda([g]) = \inf\big\{Q_g(u) : u \in W^{1,2}(M),\ \|u\|_{L^p} = 1,\ u \ge 0\big\}.$$

If $\lambda([g]) < \lambda_n$ then this infimum is attained. In fact, any minimzing sequence (sub-)converges strongly in $W^{1,2}(M)$ to a minimizer.

Remark 14.1 Recall that on $\mathbb{R}^n$ there are the metrics $(\psi_+ \circ \tau_\alpha)^* g_{\mathbb{S}^n} = u_\alpha^{4/(n-2)} g_{\mathbb{R}^n}$, where

$$u_\alpha(x) = \Big(\frac{1}{\alpha|x|^2 + \alpha^{-1}}\Big)^{\frac{n-2}{2}}.$$

The functions concentrate at the origin as $\alpha \nearrow 0$, i.e. $u_\alpha^p\,dx \to \omega_n \delta_0$. Using an almost Euclidean coordinate system, one can transplant the u_α to any manifold, showing that (exercise)

$$\lambda([g]) \le n(n-1)\omega_n^{\frac{2}{n}} \quad \text{for any compact } (M,g).$$

By the quoted Aubin-Talenti result the constant on the right is λ_n, thus we always have $\lambda([g]) \le \lambda_n$. In Theorem 14.1 it is crucial that the strict inequality is assumed.

Proof of Theorem 14.1. Let $u_j \ge 0$ be a minimizing sequence. We first note that u_j is bounded in $W^{1,2}(M)$, namely we have since $\|u_j\|_{L^p} = 1$

$$\begin{aligned}
\|u_j\|_{L^2} &\le \operatorname{vol}(M)^{\frac{2}{n}}\\
\frac{4(n-1)}{n-2}\int_M |Du_j|_g^2\,d\mu_g &\le Q_g(u_j) - \int_M S_g\,u_j^2\,d\mu_g \le Q_g(u_j) + \|(S_g)_-\|_{L^{\frac{n}{2}}(M)}.
\end{aligned}$$

The standard compactness theorems yield a subsequence such that $u_j \to u$ in $L^q(M)$ for $q < p$, and also pointwise almost everywhere, and further $Du_j \to Du$ weakly in $L^2(TM)$. Putting $u_j = u + w_j$ we now estimate, writing again ε_j for any null sequence,

$$\begin{aligned}
\lambda([g]) + \varepsilon_j \ge Q_g(u_j) &= Q_g(u) + \frac{4(n-1)}{n-2}\int_M |Dw_j|_g^2\,d\mu_g\\
&\quad + \frac{8(n-1)}{n-2}\int_M \langle Du, Dw_j\rangle_g\,d\mu_g + 2\int_M S_g u w_j\,d\mu_g + \int_M S w_j^2\,d\mu_g.
\end{aligned}$$

The last three integrals vanish as $j \to \infty$. By definition of $\lambda([g])$ and Lemma 14.4

$$Q_g(u) + \frac{4(n-1)}{n-2} \int_M |Dw_j|_g^2 \, d\mu_g \quad \geq \quad \lambda([g]) \, \|u\|_{L^p}^2 + \frac{\lambda_n}{1+\varepsilon} \|w_j\|_{L^p}^2 - C_\varepsilon \|w_j\|_{L^2}^2.$$

The last term also goes to zero. Now $\|u\|_{L^p} \leq 1$ by Fatou, thus for $\lambda([g]) \leq 0$ we have

$$\frac{\lambda_n}{1+\varepsilon} \|w_j\|_{L^p}^2 \leq \lambda([g]) \left(1 - \|u\|_{L^p}^2\right) + \varepsilon_j \leq \varepsilon_j \to 0.$$

As $\lambda_n > 0$ by (14.47) we get $w_j \to 0$ in $L^p(M)$. For $\lambda([g]) > 0$ we use $\|u_j\|_{L^p} = 1$ to write

$$\Big(\frac{\lambda_n}{1+\varepsilon} - \lambda([g])\Big) \, \|w_j\|_{L^p}^2 \leq \lambda([g]) \left(\|u_j\|_{L^p}^2 - (\|u\|_{L^p}^2 + \|w_j\|_{L^p}^2)\right) + \varepsilon_j.$$

By (14.50) the $\limsup$ on the right is non-positive. For $\varepsilon > 0$ small we get again $w_j \to 0$ and hence $u_j \to u$ in $L^p(M)$. As u is admissible for the infimum, we have further

$$\begin{aligned}
\limsup_{j \to \infty} 4\frac{n-1}{n-2} \int_M |Du_j|_g^2 \, d\mu_g \quad &= \quad \limsup_{j \to \infty} \Big(Q_g(u_j) - \int_M S_g u_j^2 \, d\mu_g\Big) \\
&= \quad \lambda([g]) - \int_M S_g u^2 \, d\mu_g \\
&\leq \quad 4\frac{n-1}{n-2} \int_M |Du|_g^2 \, d\mu_g.
\end{aligned}$$

Therefore the convergence is also strong in $W^{1,2}(M)$. □

Example 14.1 A simple estimate for $\lambda([g])$ is obtained by taking $u \equiv 1$, then we have

$$\begin{aligned}
\|u\|_{L^p}^2 \quad &= \quad \mathrm{vol}(M)^{1-\frac{2}{n}}, \\
Q_g(u) \quad &= \quad \int_M S_g \, d\mu_g \leq \|(S_g)_+\|_{L^{\frac{n}{2}}} \, \mathrm{vol}(M)^{1-\frac{2}{n}}.
\end{aligned}$$

Therefore if $\|(S_g)_+\|_{L^{\frac{n}{2}}} < \lambda_n$, then the hypothesis of the theorem is satisfied.

Aubin's theorem shifts the question about the Yamabe problem to the construction of a test function with Yamabe quotient below λ_n. Of course this can only be done if (M, g) is not globally conformally equivalent to $\mathbb{S}^n$. At this point we give a short historical overview.

- In 1960 Yamabe discovered the variational structure of the problem. He proposed a subcritical approach to the problem, by first minimizing under the constraint $\|u\|_{L^s} = 1$ for $s < p$ and then studying the limit $s \nearrow p$. The minimization step poses no difficulties since the Rellich theorem is available for $s < p$. However the limit $s \nearrow p$ is nontrivial, and Yamabe's arguments were not sufficient.

- In 1968 Trudinger took up Yamabe's approach and justified the argument, in the case $\lambda([g]) \leq 0$ or when $\lambda([g])$ is small.

- In 1976 Aubin solved the problem under the optimal assumption $\lambda([g]) < \lambda_n$. As mentioned before, he and also Talenti determined the value $\lambda_n = n(n-1)\omega_n^{2/n}$.

- In the same year, Aubin verified the assumption $\lambda([g]) < \lambda_n$ for manifolds of dimension $\dim M \geq 6$ which are not locally conformally flat. Here (M, g) is called locally conformally flat. if it is locally conformal to $\mathbb{R}^n$, i.e. there are charts in which $g = e^{2u} g_{\mathbb{R}^n}$. The local flatness implies the vanishing of the Weyl curvature tensor, which is the conformally invariant part of the Riemann curvature tensor. The test functions used by Aubin are local in nature, exploiting that the Weyl tensor is nonzero. Aubin's argument can not be applied in the case when M is locally conformally flat, because the desired inequality fails fo the round sphere. Instead, a global construction of a test function is needed. This is also true, but less obvious, in low dimensions.

- In 1984 Schoen proved the inequality $\lambda([g]) < \lambda_n$ in the cases when $\dim M \leq 7$. For the construction he applied the positive mass theorem of general relativity, which he had proved before in 1979 together with Yau.

- In 1988 Schoen and Yau finally showed the inequality in the case when M is locally conformally flat, which was the remaining open case. The proof does not employ the positive mass theorem, which is still unknown in high dimensions.

We close this section by proving the following regularity theorem.

Theorem 14.2 *Let $u \in W^{1,2}(M)$ be a weak solution of the equation $L_g u = \lambda u^{\frac{n+2}{n-2}}$, that is*

$$\int_M \Big(\frac{4(n-1)}{n-2}\langle Du, D\phi\rangle_g + S_g u\phi\Big)\, d\mu_g = \lambda \int_M u^{\frac{n+2}{n-2}}\phi\, d\mu_g \quad \text{for all } \phi \in C^\infty(M).$$

Then $u \in C^\infty(M)$ and $u > 0$, hence $\tilde{g} = u^{\frac{4}{n-2}} g$ is a smooth Riemannian metric with scalar curvature $S_{\tilde{g}} \equiv \lambda$.

The result applies to the minimizers constructed in Theorem 14.1. In fact, as noted earlier, the condition $u \geq 0$ in the definition of $\lambda([g])$ can be dropped without changing the infimum. Therefore the functions minimize under the single condition $\|u\|_{L^p} = 1$, and the weak equation can be derived as in Lemma 14.2 using the implicit function theorem.

Our approach to Theorem 14.2 uses L^p-estimates, which go back to Calderon-Zygmund (1952). They apply to linear elliptic equations with a right hand side locally in L^p, where $1 < p < \infty$, and assert that the solution is locally in $W^{2,p}$. The case $p = 2$ is the standard L^2 theory. Let us explain on a heuristic level how L^p theory can be used to push up the regularity of the solution to the Yamabe equation. Assume we know already that $u \in W^{1,r}$ locally for $2 \leq r < n$. By the Sobolev embedding theorem

$$u \in L^q_{\text{loc}} \text{ where } \frac{1}{q} = \frac{1}{r} - \frac{1}{n}.$$

Then for the right hand side of Yamabe's equation

$$u^{\frac{n+2}{n-2}} \in L^p \quad \text{ where } p = \frac{n-2}{n+2}\, q.$$

We note that

$$0 < \frac{1}{p} = \frac{n+2}{n-2}\Big(\frac{1}{r} - \frac{1}{n}\Big) \leq \frac{n+2}{n-2}\Big(\frac{1}{2} - \frac{1}{n}\Big) = \frac{n+2}{2n} < 1.$$

Thus by L^p regularity we obtain $u \in W^{2,p}$ locally, and again by Sobolev

$$u \in W^{1,s}_{\mathrm{loc}} \quad \text{where } \frac{1}{s} = \frac{1}{p} - \frac{1}{n}.$$

We calculate

$$\frac{1}{s} = \frac{n+2}{n-2}\Big(\frac{1}{r} - \frac{1}{n}\Big) - \frac{1}{n} = \frac{1}{r} + \frac{4}{n-2}\Big(\frac{1}{r} - \frac{1}{2}\Big).$$

We arrive at the following conclusions:

- if initially $r > 2$, then $s > r$ and hence the regularity is improved. In fact, the argument can be iterated to show that $u \in W^{1,s}_{\mathrm{loc}}$ for all $s < \infty$. In particular the solution u is Hölder continuous, by the corresponding Sobolev embedding. Then the Schauder regularity theorem shows $u \in C^{2,\alpha}$, and u is a classical solution.
- if we have initially only $r = 2$, then $s = r$ and we get no improvement.

In the initial step the following precise result from Calderon-Zygmund will be used.

Lemma 14.6 *Let $B = B_1(0) \subset \mathbb{R}^n$ and $1 < r < \infty$. Then the classical Laplacian*

$$\Delta : W^{2,r}_0(B) \to L^r(B), \quad \text{where } W^{2,r}_0(B) := W^{2,r}(B) \cap W^{1,r}_0(B),$$

is an isomorphism, that is it has a bounded inverse.

Proof of Theorem 14.2. Due to scale invariance it is not possible to get absolute a priori estimates. Instead we are going for bounds under the assumption that the solution is not too much concentrated. This will be measured by the local L^p integral, where $p = \frac{2n}{n-2}$.

By chosing coordinates, we can assume that g_{ij} is given near the origin, satisfying

$$g_{ij}(0) = \delta_{ij} \quad \text{and} \quad \partial_i g_{jk}(0) = 0.$$

On $B = B_1(0)$ we consider the rescaled metric $g^\varrho = \varrho^{-2}\tau_\varrho^* g$, where $\tau_\varrho : B \to B_\varrho(0)$, $\tau_\varrho(x) = \varrho x$. Then we have $g^\varrho_{ij}(x) = g_{ij}(\varrho x)$ and

$$\|g^\varrho - g_{\mathbb{R}^n}\|_{C^1(B)} \to 0 \quad \text{as } \varrho \searrow 0. \tag{14.51}$$

Now we transform the situation back to B. In the following we write $\varepsilon(\varrho)$ for a quantity which goes to zero as $\varrho \searrow 0$. We have

$$v = \varrho^{\frac{n-2}{2}} u \circ \tau_\varrho \quad \Rightarrow \quad \int_B v^p \, d\mu_{g^\varrho} = \int_{B_\varrho(0)} u^p \, d\mu_g = \varepsilon(\varrho).$$

The weak Yamabe equation becomes, for test functions related by $\varphi = \varrho^{\frac{n-2}{2}} \phi \circ \varrho$,

$$\int_B \Big(\langle Dv, D\varphi\rangle_{g^\varrho} + \frac{1}{c_n} S_{g^\varrho} v\varphi\Big)\, d\mu_{g^\varrho} = \frac{\lambda}{c_n}\int_B v^{\frac{n+2}{n-2}}\varphi \, d\mu_{g^\varrho} \quad \text{where } c_n = \frac{4(n-1)}{n-2}.$$

For $\eta \in C^\infty_c(B)$ with $\eta = 1$ on $B_{\frac{1}{2}}(0)$ we now compute the equation satisfied by ηv.

$$\begin{aligned}
&\int_B \Big(\langle D(\eta v), D\varphi\rangle_{g^\varrho} + \frac{1}{c_n} S_{g^\varrho}\, \eta v \varphi\Big)\, d\mu_{g^\varrho} \\
= \;& \int_B \Big(\langle Dv, D(\eta\varphi)\rangle_{g^\varrho} + \frac{1}{c_n} S_{g^\varrho}\, v(\eta\varphi)\Big)\, d\mu_{g^\varrho} + \int_B \big(v\langle D\eta, D\varphi\rangle_{g^\varrho} - \langle D\eta, Dv\rangle_g \varphi\big)\, d\mu_{g^\varrho} \\
= \;& \frac{\lambda}{c_n}\int_B v^{\frac{n+2}{n-2}}\eta\,\varphi\, d\mu_{g^\varrho} - 2\int_B \langle D\eta, Dv\rangle_{g^\varrho}\varphi\, d\mu_{g^\varrho} - \int_B (\Delta_{g^\varrho}\eta) v\,\varphi\, d\mu_{g^\varrho}.
\end{aligned}$$

We abbreviate $f = \eta v$ and rearrange to

$$\int_B \Big(\langle Df, D\varphi \rangle_{g^\varrho} - \underbrace{\frac{\lambda}{c_n} v^{\frac{4}{n-2}}}_{=:q} f\varphi \Big)\, d\mu_{g^\varrho} = - \int_B \underbrace{\frac{1}{c_n} \big(2\langle D\eta, Dv\rangle_{g^\varrho} + (\Delta_{g^\varrho}\eta)v + S_{g^\varrho}\eta v\big)}_{=:h} \varphi\, d\mu_{g^\varrho}.$$

Now we have the bound

$$\int_B |q|^{\frac{n}{2}}\, d\mu_{g^\varrho} \le C|\lambda|^{\frac{n}{2}} \int_B v^p \, d\mu_{g^\varrho} = C\,|\lambda|^{\frac{n}{2}}\, \varepsilon(\varrho).$$

This implies for $1 \le r < \frac{n}{2}$, using Hölder and Sobolev,

$$\|qf\|_{L^r(B)} \le C\, \|q\|_{L^{\frac{n}{2}}}\, \|f\|_{L^{\frac{nr}{n-2r}}(B)} \le C\,|\lambda|\, \varepsilon(\varrho)^{\frac{2}{n}}\, \|f\|_{W^{2,r}(B)}.$$

Furthermore we check that

$$\begin{aligned} \|(\Delta_{g^\varrho} - \Delta_{\mathbb{R}^n})f\|_{L^r(B)} &\le \|\big((g^\varrho)^{ij} - \delta^{ij}\big)\partial^2_{ij} f\|_{L^r(B)} + \|\frac{\partial_i(\sqrt{\det g^\varrho}(g^\varrho)^{ij})}{\sqrt{\det g^\varrho}}\,\partial_j f\|_{L^r(B)} \\ &\le \varepsilon(\varrho)\, \|f\|_{W^{2,r}(B)}. \end{aligned}$$

The estimates show that the operator $\Delta_{g^\varrho} + q : W^{2,r}_0(B) \to L^r(B)$ is close to $\Delta_{\mathbb{R}^n}$ in operator norm, and hence also invertible. Now η, $\Delta_{g^\varrho}\eta$ and $S_{g^\varrho} = \varrho^2 S_g \circ \tau_\varrho$ are all bounded functions, hence $\|h\|_{L^r(B)} \le C\,\|v\|_{W^{1,r}(B)}$, in particular

$$\|h\|_{L^r(B)} \le C\,\|v\|_{W^{1,2}(B)} \quad \text{for } r \le 2.$$

Now for $1 < r < \frac{n}{2}$, $r \le 2$, let $w \in W^{2,r}_0(B)$ be the solution of the equation $\Delta_{g^\varrho} w + qw = h$. By Sobolev embedding we get $w \in W^{1,s}(B)$ where

$$\frac{1}{s} = \frac{1}{r} - \frac{1}{n} = \frac{1}{2} + \frac{1}{r} - \frac{n+2}{2n}.$$

In particular $s > 2$ by choosing $r > \frac{2n}{n+2}$. We now claim that $w = \eta v$, which completes the intial regularity step. Integrating by parts, we have for all $\varphi \in C^\infty_c(B)$

$$0 = -\int_B (\Delta_{g^\varrho} w + qw - h)\varphi\, d\mu_{g^\varrho} = \int_B \big(\langle Dw, D\varphi\rangle_{g^\varrho} - (qw - h)\varphi\big)\, d\mu_{g^\varrho}.$$

Subtracting the equations we get

$$0 = \int_B \big(\langle D(w - \eta v), D\varphi\rangle_{g^\varrho} - q(w - \eta v)\varphi\big)\, d\mu_{g^\varrho}.$$

Approximating we can take $\varphi = w - \eta v$ as test fucntion, and conclude

$$\begin{aligned} \int_B |D(w - \eta v)|^2\, d\mu_{g_\varrho} &\le \int_B |q|\, |w - \eta v|^2 \\ &\le C\, \|q\|_{L^{\frac{n}{2}}}\, \|w - \eta v\|^2_{L^p} \\ &\le C\,|\lambda|\, \varepsilon(\varrho)^{\frac{2}{n}}\, \|D(w - \eta v)\|^2_{L^2}. \end{aligned}$$

Thus $w = \eta v$ for $\varrho > 0$ suffciently small. This finishes the proof. □

15 Remarks on the Einstein-Hilbert action

This section discusses some aspects of the total scalar curvature functional for Riemannian metrics. In the physical literature the functional is called the Einstein-Hilbert action.

Definition 15.1 *Let (M, g) be a Riemannian manifold. For a tensor T of type (r, s) with $r \geq 1$ we define the tensor D^*T of type $(r-1, s)$ as*

$$D^*T = -\mathrm{tr}_g DT = -\sum_{i,j=1}^{n} g^{ij} D_{\frac{\partial}{\partial x^i}} T\Big(\frac{\partial}{\partial x^j}, \,\cdot\,\Big).$$

Lemma 15.1 (partial integration) *Let S, T be tensors of type $(r-1, s)$ and (r, s), where $r \geq 1$, on the compact Riemannian manifold (M, g), possibly with boundary ∂M. Then*

$$\int_M \big(\langle DS, T\rangle_g - \langle S, D^*T\rangle_g\big)\, d\mu_g = \int_{\partial M} \langle S, T(\nu, \,\cdot\,)\rangle_g \, d\omega_g. \tag{15.52}$$

Here ν denotes the exterior unit normal.

Proof. In a first step, we let $\xi \in C^1(T^*M)$ be a 1-form with associated vector field $X \in C^(TM)$, that is $\xi = g(\,\cdot\, X)$. We have in coordinates or $X^i = \sum_{j=1}^n g^{ij}\xi_j$. We have

$$\begin{aligned} D^*\xi &= -\sum_{i,j=1}^{n} g^{ij} D_{\frac{\partial}{\partial x^i}} \xi\Big(\frac{\partial}{\partial x^j}\Big) = -\sum_{i,j=1}^{n} g^{ij}\partial_i \xi_j + \sum_{i,j=1}^{n} \xi\Big(D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}\Big), \\ \mathrm{div}_g X &= \frac{1}{\sqrt{\det g}} \sum_{i=1}^{n} \partial_i(\sqrt{\det g} X^i). \end{aligned}$$

Now $X^i = \sum_{j=1}^n g^{ij}\xi_j$. Using almost Euclidean coordinates at a given point $p \in M$, i.e. $g_{ij}(p) = \delta_{ij}$ and $\partial_i g_{jk}(p) = 0$, we check easily that

$$D^*\xi = -\mathrm{div}_g X.$$

The theorem of Gauß now yields

$$-\int_M D^*\xi \, d\mu_g = \int_M \mathrm{div}_g X \, d\mu_g = \int_{\partial M} g(X, \nu)\, d\omega_g = \int_{\partial M} \xi(\nu)\, d\omega_g.$$

For S, T as in the claim, consider now $\xi = \langle S, T\rangle_g$, in coordinates (summation convention)

$$\xi = \, g_{j_1\ell_1} \cdot \ldots \cdot g_{j_s\ell_s} \, g^{i_1k_1} \cdot \ldots \cdot g^{i_{r-1}k_{r-1}} \, S^{j_1\ldots j_s}_{i_1\ldots i_{r-1}} \, T^{\ell_1\ldots\ell_s}_{i\,k_1\ldots k_{r-1}} \, dx^i.$$

Using again almost Euclidean coordinates we compute

$$D^*\xi = -\big(\partial_i S^{j_1\ldots j_s}_{i_1\ldots i_{r-1}}\big)\, T^{j_1\ldots j_s}_{i\,i_1\ldots i_{r-1}} - S^{j_1\ldots j_s}_{i_1\ldots i_{r-1}} \,\big(\partial_i T^{j_1\ldots j_s}_{i\,i_1\ldots i_{r-1}}\big) = -\langle DS, T\rangle_g + \langle S, D^*T\rangle_g.$$

We conclude

$$\int_M \big(\langle DS, T\rangle_g - \langle S, D^*T\rangle_g\big)\, d\mu_g = -\int_M D^*\xi \, d\mu_g = \int_{\partial M} \xi(\nu)\, d\omega_g = \int_{\partial M} \langle S, T(\nu), \,\cdot\,)\rangle_g.$$

$\square$

Theorem 15.1 (Hilbert) *Let $g(t)$ be a smooth family of Riemannian metrics on the n-dimensional manifold M. Then for $g = g(0)$ and $h = \partial_t g|_{t=0}$ we have*

$$\frac{\partial}{\partial t}\big(S_{g(t)}\,d\mu_{g(t)}\big)|_{t=0} = \big(-\langle G, h\rangle_g + D^*\alpha\big)\,d\mu_g, \tag{15.53}$$

where $G = \mathrm{ric}_g - \frac{1}{2}S_g\,g$ is the Einstein tensor and α is the 1-form $D^*h + d(\mathrm{tr}_g h)$.

Note that for $n = 2$ one has always $G = 0$, hence the first variation vanishes if h has compact support. This relates to the fact that the Einstein-Hilbert functional is just twice the total Gauß curvature, which is deformation invariant due to the Gauß-Bonnet theorem. Even in higher dimensions there is some cancellation: the scalar curvature depends on second derivatives of the metric, so in general one would expect the Euler-Lagrange operator to be of order four in the metric. By contrast G contains only second order derivatives.

The proof of the theorem is completely straightforward, but unfortunately the calculations are long. We postpone them and first discuss an application. For $X \in C^2(TM)$ and $\Omega \subset\subset M$ we have the flow $\phi \in C^2(\Omega \times (-\varepsilon, \varepsilon), M)$ defined by

$$\frac{\partial \phi}{\partial t}(p,t) = X(\phi(p,t)), \quad \phi(p,0) = 0.$$

We want to apply the variational formula with $g(t) = \phi_t^* g$. To compute the derivative we conisider a curve $\gamma(s)$ with $\gamma(0) = p$ and $\gamma'(0) = Y \in T_pM$. Then

$$\begin{aligned}
\frac{D}{\partial t} D\phi_t(p)Y &= \frac{D}{\partial t}\frac{\partial}{\partial s}\phi_t(\gamma(s))|_{s,t=0}\\
&= \frac{D}{\partial s}\frac{\partial}{\partial t}\phi_t(\gamma(s))|_{t,s=0}\\
&= \frac{D}{\partial s}X(\gamma(s))|_{s=0}\\
&= D_Y X(p).
\end{aligned}$$

Using this we calculate

$$h_{ij} := \frac{\partial g_{ij}}{\partial t}|_{t=0} = \frac{\partial}{\partial t} g\big(D\phi_t \frac{\partial}{\partial x^i}, D\phi_t \frac{\partial}{\partial x^j}\big) = g\big(D_{\frac{\partial}{\partial x^i}} X, \frac{\partial}{\partial x^j}\big) + g\big(\frac{\partial}{\partial x^i}, D_{\frac{\partial}{\partial x^j}} X\big).$$

For simplicity we assume X has compact support in Ω. As $\phi_t : (\Omega, \phi_t^* g) \to (\Omega, g)$ is by definition an isometry, we have

$$\int_\Omega S_{\phi_t^* g}\,d\mu_{\phi_t^* g} = \int_\Omega S_g\,d\mu_g \quad \text{for all } t \in \mathbb{R}.$$

The integral of $D^*\alpha$ vanishes. Using the symmetry of G we conclude

$$0 = \frac{d}{dt}\int_\Omega S_{\phi_t^* g}\,d\mu_{\phi_t^* g}|_{t=0} = -\int_\Omega \langle G, DX\rangle_g\,d\mu_g = -\int_\Omega \langle D^*G, X\rangle_g.$$

This shows that D^*G is identically zero. One can rewrite this by splitting

$$G = B - \frac{n-2}{2n} S_g\,g,$$

i.e. B is the tracefree component of the Ricci tensor. Now $D^*(S_g\, g) = -dS_g$, thus

$$D^*B = -\frac{n-2}{2n} dS_g.$$

For $n \geq 3$ this formula has the following standard consequence: if $B = 0$ then S_g is constant and hence g is an Einstein metric.

Let's return to conformal geometry. On a compact manifold M consider the metrics

$$\tilde{g} = \varphi^{-2} g \quad \text{ where } \varphi > 0.$$

Writing the factor in this form is just more comvenient here. Namely one obtains

$$\tilde{B} = B + \frac{n-2}{\varphi} \big(D^2\varphi - \frac{1}{n}(\Delta\varphi)\, g\big).$$

Taking the scalar product with φB and integrating yields

$$\int_M \varphi\, \langle \tilde{B}, B\rangle_g \, d\mu_g = \int_M \varphi\, |B|_g^2 \, d\mu_g + (n-2) \int_M \langle D^2\varphi - \frac{1}{n}(\Delta\varphi)\, g,\, B\rangle_g \, d\mu_g.$$

Clearly $\langle g, B\rangle_g = 0$ sine B is trace-free. Moroever

$$\int_M \langle D^2\varphi, B\rangle_g \, d\mu_g = \int_M \langle D\varphi, D^*B\rangle_g \, d\mu_g = -\frac{n-2}{2n} \int_M \langle d\varphi, dS_g\rangle_g \, d\mu_g.$$

Now we asume that S_g is constant. Then by Cauchy-Schwarz and rearranging

$$\int_M \varphi\, |B|_g^2 \, d\mu_g \leq \int_M \varphi\, |\tilde{B}|_g^2 \, d\mu_g.$$

Here is an application of this inequality.

Theorem 15.2 (Obata) *Let (M, g) be a compact Riemannian manifold with constant scalar curvature. If there is a metric $\tilde{g}$ conformal to g which is Einstein, then g is also Einstein.*

In particular, consider a metric g on $\mathbb{S}^n$ with constant scalar curvature, which is conformal to the standard metric g_0. Then g is Einstein, moreover the conformally invariant part of the Riemann curvature tensor (the Weyl tensor) is also zero. Together this implies that g is a metric of constant sectional curvature, and hence there exists an isometry $\phi : (\mathbb{S}^n, g) \to (\mathbb{S}^n, g_0)$. It follows that ϕ is a conformal diffeomorphism of $\mathbb{S}^n$. This shows that any constant scalar curvature metric on $\mathbb{S}^n$ which is conformal to the standard metric g_0 is actually the pullback under a conformal diffeomorphism. These diffeomorphisms are classified, they come from Möbius transformations of $\mathbb{R}^{n+1}$.

16 Yang-Mills Theory: Background Material

Definition 16.1 (vector bundle) *A surjective map $\pi \in C^\infty(E, M)$ between manifolds is called a complex vector bundle of rank r, if the following holds for every $p \in M$:*

(i) *$E_p = \pi^{-1}\{p\}$ is a complex vector space of dimension r.*

(ii) *There is a smooth basis $\mathcal{A} = \{a_1, \dots, a_r\}$ on a neighborhood U of p.*

More precisely, $a_i \in C^\infty(U, E)$ and $\{a_1(q), \dots, a_r(q)\}$ is a basis of E_q for all $q \in U$. The basis induces the bundle chart (or local trivialization)

$$\varphi_{\mathcal{A}} : \pi^{-1}(U) \to U \times \mathbb{C}^r,\ \varphi_{\mathcal{A}}(v) = (\pi(v), v_{\mathcal{A}}).$$

Here $v_{\mathcal{A}}$ is the coordinate vector with respect to $\mathcal{A}$. The inverse of the chart is

$$\varphi_{\mathcal{A}}^{-1} : U \times \mathbb{C}^r \to \pi^{-1}(U),\ \varphi_{\mathcal{A}}^{-1}(q, z) = \sum_{i=1}^r z^i a_i(q).$$

Given two bases $\mathcal{A}$,$\mathcal{B}$ on open sets U,V, we have the transition function

$$\varphi_{\mathcal{B}\mathcal{A}} : U \cap V \to \mathrm{GL}_r(\mathbb{C}),\ a_j = \sum_{i=1}^r (\varphi_{\mathcal{B}\mathcal{A}})^i_j\, b_i.$$

It follows that $\varphi_{\mathcal{B}} \circ \varphi_{\mathcal{A}}^{-1}(q, z) = (q, \varphi_{\mathcal{B}\mathcal{A}}(q)z)$, and we have the cocycle relations

$$\varphi_{\mathcal{C}\mathcal{A}} = \varphi_{\mathcal{C}\mathcal{B}} \cdot \varphi_{\mathcal{B}\mathcal{A}} \quad \text{and} \quad \varphi_{\mathcal{A}\mathcal{A}} = \mathrm{E}_k.$$

In the following we write G instead of $\mathrm{GL}_r(\mathbb{C})$ and $\mathfrak{g}$ instead of $\mathbb{C}^{r \times r}$. Similarly, we put $\mathfrak{g}_E = \bigcup_{p \in M} \mathrm{End}(E_p)$. This notation potentially allows to cover cases when the bundle has some extra structure, with transition functions taking values in a subgroup of $\mathrm{Gl}_r(\mathbb{C})$. A typical case is when E carries a hermitian metric. By Gram-Schmidt one can then construct local bases which are orthonormal, and hence the transition functions are in $\mathrm{U}(r)$. Moreover, one then restricts to endomorphisms in $\mathfrak{u}_E$, i.e. skew-hermitian ($B^* + B = 0$) with respect to the metric h. The set $\mathfrak{g}_E$ is again a vector bundle over M, with induced charts

$$\phi_{\mathcal{A}} : \pi^{-1}(U) \to U \times \mathfrak{g},\ \phi_{\mathcal{A}}(f) = (\pi(f), f_{\mathcal{A}}).$$

Of course, $f_{\mathcal{A}}$ denotes the matrix of f with respect to the basis $\mathcal{A}$. We have

$$\phi_{\mathcal{A}}^{-1} : U \times \mathfrak{g} \to \pi^{-1}(U),\ \phi_{\mathcal{A}}^{-1}(q, A)\, a_j = \sum_{i=1}^r A^i_j\, a_i.$$

Under a base change with matrix $g = \mathrm{Id}_{\mathcal{B}\mathcal{A}}$, the matrix of $f \in \mathrm{End}(E)$ transforms by

$$f_{\mathcal{B}} = \varphi_{\mathcal{B}\mathcal{A}}\, f_{\mathcal{A}}\, \varphi_{\mathcal{A}\mathcal{B}} = g\, f_{\mathcal{A}}\, g^{-1}.$$

Definition 16.2 *Let $\pi : E \to M$ be a (complex) vector bundle of rank r. A covariant derivative on E is a linear map $D : C^1(E) \to C^0(T^*M \otimes E)$ satisfying the product rule*

$$D(fv) = df \otimes v + f\, D\xi.$$

With respect to a basis $\mathcal{A} = \{a_1, \ldots, a_r\}$ on U, the connection is represented by a matrix-valued one-form A, in other words $A \in C^\infty(T^*U \otimes \mathfrak{g})$. Namely we have

$$D_X a_j = \sum_{i=1}^{r} A^i_j(X)\, a_i \quad \text{or} \quad (Dv)_{\mathcal{A}} = dv_{\mathcal{A}} + Av_{\mathcal{A}}.$$

To verify the second formula we compute

$$\begin{aligned} Dv = D\Big(\sum_{j=1}^{r} v^j_{\mathcal{A}}\, a_j\Big) &= \sum_{j=1}^{r} \big(dv^j_{\mathcal{A}} \otimes a_j + v^j_{\mathcal{A}} Da_j\big) \\ &= \sum_{i=1}^{r} \big(dv^i_{\mathcal{A}} + A^i_j v^j_{\mathcal{A}}\big) \otimes a_i \\ &= \sum_{i=1}^{r} \big(dv_{\mathcal{A}} + A\, v_{\mathcal{A}}\big)^i \otimes a_i. \end{aligned}$$

In the case when E carries a hermitian metric one can require that D is a metric connection, see Definition 9.2. With respect to a local orthonormal basis, A then takes values in $\mathfrak{u}(r)$.

Lemma 16.1 *Let A, B be the connection one-forms of D with respect to the bases $\mathcal{A}$, $\mathcal{B}$ on U and E. Then on $U \cap V$ we have*

$$B = gAg^{-1} + g\, d(g^{-1}) = gAg^{-1} - (dg)\, g^{-1} \qquad \textit{where } g = \varphi_{\mathcal{B}\mathcal{A}}.$$

Proof. We compute, using $v_{\mathcal{A}} = \varphi_{\mathcal{A}\mathcal{B}} v_{\mathcal{B}} = g^{-1} v_{\mathcal{B}}$,

$$\begin{aligned} (Dv)_{\mathcal{A}} &= dv_{\mathcal{A}} + Av_{\mathcal{A}} \\ &= d(g^{-1}v_{\mathcal{B}}) + Ag^{-1}v_{\mathcal{B}} \\ &= g^{-1}dv_{\mathcal{B}} + d(g^{-1})v_{\mathcal{B}} + Ag^{-1}v_{\mathcal{B}}. \end{aligned}$$

On the other hand $(Dv)_{\mathcal{A}} = \varphi_{\mathcal{A}\mathcal{B}}(Dv)_{\mathcal{B}} = g^{-1}(dv_{\mathcal{B}} + Bv_{\mathcal{B}})$. Comparing we obtain

$$Bv_{\mathcal{B}} = gAg^{-1}v_{\mathcal{B}} + g\, d(g^{-1})\, v_{\mathcal{B}}.$$

□

In Theorem 9.1 we showed that connections always exist. Moreover, if D_0 is some connection then all connections are given by $D = D_0 + \omega$, where ω is a 1-form with values in $\mathfrak{g}_E$. that is $\omega \in C^\infty\big(T^*M \otimes \mathfrak{g}_E\big)$. It is common practice to write D_A instead of D. Perhaps this notation is motivated by the goal to distinguish the connection from the standard derivative. It may also come from physics, where one always works in a local situation.

We now introduce an operator d_A on sections of $\Lambda^p TM \otimes E$ which couples the exterior derivative with the connection. Using a local basis $\mathcal{A}$, we define $d_A\omega$ by

$$(d_A\omega)_{\mathcal{A}} = d\omega_{\mathcal{A}} + A \wedge \omega_{\mathcal{A}} \quad \text{where } (A \wedge \omega_{\mathcal{A}})^i = \sum_{j=1}^{r} A^i_j \wedge \omega^j_{\mathcal{A}}. \tag{16.54}$$

This is independent of $\mathcal{A}$, for a basis $\mathcal{B}$ and $g = \varphi_{\mathcal{B}\mathcal{A}}$ we have on the overlap

$$d\omega_{\mathcal{B}} + B \wedge \omega_{\mathcal{B}} = d(g\omega_{\mathcal{A}}) + \big(g\, A\, g^{-1} - (dg)\, g^{-1}\big) \wedge g\omega_{\mathcal{A}} = g\,(d\omega_{\mathcal{A}} + A \wedge \omega_{\mathcal{A}}).$$

The standard exterior derivative satisfies $d^2 = 0$, so we are interested in d_A^2. We compute for any section v of E, using again a local basis $\mathcal{A}$,

$$\begin{aligned}(d_A^2 v)_{\mathcal{A}} &= d\big((d_A v)_{\mathcal{A}}\big) + A \wedge (d_A v)_{\mathcal{A}} \\ &= d\big(dv_{\mathcal{A}} + Av_{\mathcal{A}}\big) + A \wedge \big(dv_{\mathcal{A}} + Av_{\mathcal{A}}\big) \\ &= (dA + A \wedge A)\, v_{\mathcal{A}}.\end{aligned}$$

We now define a section F_A of $\Lambda^2 TM \otimes \mathfrak{g}_E$ as follows: for given $v \in E_p$ we chose a section $\tilde{v}$ of E with $\tilde{v}(p) = v$, and put $F_A(p)v = d_A^2 \tilde{v}(p)$. By the above local formula, $F_A(p)v$ does not depend on the choice of $\tilde{v}$ and $F_A(p)$ acts linearly on E_p. Moreover the map $F_A : M \to \Lambda^2 TM \otimes \mathfrak{g}_E$ is a smooth section. Under a base change the matrix of F_A transforms by

$$(F_A)_{\mathcal{B}} = g\,(F_A)_{\mathcal{A}}\, g^{-1} \qquad \text{where } g = \varphi_{\mathcal{B}\mathcal{A}} : U \cap V \to G. \tag{16.55}$$

This just holds for any section of $\mathfrak{g}_E$, explicitely we have

$$(F_A v)_{\mathcal{B}} = g(F_A v)_{\mathcal{A}} = g\,(F_A)_{\mathcal{A}}\, v_{\mathcal{A}} = g\,(F_A)_{\mathcal{A}}\, g^{-1} v_{\mathcal{B}}.$$

We summarize the calculations as follows.

Definition 16.3 (curvature) *The operator $F_A = d_A^2$ is a section of $\Lambda^2 TM \otimes \mathfrak{g}_E$, it is the curvature of D_A. With respect to a local basis, we have*

$$F_A = dA + A \wedge A = dA + \frac{1}{2}[A, A].$$

In the above we used the wedge product for differential forms with values in $\mathfrak{g}$. If A, B are forms of degree p, q then $A \wedge B$ is a $(p+q)$-form also with values in $\mathfrak{g}$, given by

$$(A \wedge B)^i_k = \sum_{j=1}^{k} A^i_j \wedge B^j_k. \tag{16.56}$$

From the standard definition of the wedge product one obtains

$$(A \wedge B)(X_1, \ldots, X_{p+q}) = \frac{1}{p!\,q!} \sum_{\sigma \in S_{p+q}} \operatorname{sign}(\sigma) A(X_{\sigma(1)}, \ldots, X_{\sigma(p)}) B(X_{\sigma(p+1)}, \ldots, X_{\sigma(p+q)}).$$

This formula may be used to define also the commutator, namely

$$[A, B](X_1, \ldots, X_{p+q}) = \frac{1}{p!\,q!} \sum_{\sigma \in S_{p+q}} \operatorname{sign}(\sigma) [A(X_{\sigma(1)}, \ldots, X_{\sigma(p)}), B(X_{\sigma(p+1)}, \ldots, X_{\sigma(p+q)})].$$

Computations with these definitions need some care. For instance, one has $a \wedge a = 0$ for any 1-form with values in $\mathbb{R}$ or $\mathbb{C}$, but the term $A \wedge A$ in the definition of F_A is in general nonzero. The following identities are easily checked.

$$\begin{aligned}[A, B] &= A \wedge B - (-1)^{pq} B \wedge A, \\ [B, A] &= (-1)^{pq+1} [A, B].\end{aligned}$$

In particular $A \wedge A = \frac{1}{2}[A, A]$ in Definition 16.3. We also see easily that

$$\begin{aligned} d(A \wedge B) &= (dA) \wedge B + (-1)^p A \wedge B, \\ d[A, B] &= [dA, B] + (-1)^p [A, dB]. \end{aligned}$$

Unfortunately, it is getting still more complicated, namely we need the operator d_A also for differential forms with values in $\mathfrak{g}_E$. First for sections of $\mathfrak{g}$ we have the induced connection, defined by the product rule,

$$(D_A f)v = D_A(fv) - f D_A v.$$

This is again coupled to the exterior derivative on forms to get d_A.

Lemma 16.2 *Let $\mathcal{A} = \{a_1, \ldots, a_r\}$ be a local basis. For $\phi \in C^\infty(\Lambda^p TM \otimes \mathfrak{g}_E)$ we have*

$$(d_A \phi)_{\mathcal{A}} = d\phi_{\mathcal{A}} + [A, \phi_{\mathcal{A}}]. \tag{16.57}$$

Proof. We first assume $p = 0$, so ϕ is just a section of $\mathfrak{g}_E$. Then

$$\begin{aligned} ((D_A \phi)v)_{\mathcal{A}} &= (D_A(\phi v) - \phi\, D_A v)_{\mathcal{A}} \\ &= d((\phi v)_{\mathcal{A}}) + A\,(\phi v)_{\mathcal{A}} - \phi_{\mathcal{A}}(D_A v)_{\mathcal{A}} \\ &= d(\phi_{\mathcal{A}} v_{\mathcal{A}}) + A\, \phi_{\mathcal{A}} v_{\mathcal{A}} - \phi_{\mathcal{A}}\,(dv_{\mathcal{A}} + A\, v_{\mathcal{A}}) \\ &= d\phi_{\mathcal{A}} v_{\mathcal{A}} + [A, \phi_{\mathcal{A}}] v_{\mathcal{A}}. \end{aligned}$$

For general p the formula now holds by definition analogous to equation (16.54). $\square$

Theorem 16.1 (Bianchi's identity) *The curvature F_A of a connection D_A satisfies*

$$d_A F_A = 0. \tag{16.58}$$

Proof. We compute in a local basis $\mathcal{A}$

$$\begin{aligned} (d_A F_A)_{\mathcal{A}} &= d(F_A)_{\mathcal{A}} + [A, (F_A)_{\mathcal{A}}] \\ &= d\big(dA + \frac{1}{2}[A, A]\big) + \big[A, dA + \frac{1}{2}[A, A]\big] \\ &= \frac{1}{2}[dA, A] - \frac{1}{2}[A, dA] + [A, dA] + \frac{1}{2}\big[A, [A, A]\big] \\ &= \frac{1}{2}\big[A, [A, A]\big]. \end{aligned}$$

We used that $[dA, A] = (-1)^{1 \cdot 2 + 1}[A, dA] = -[A, dA]$. Now the last term vanishes by the Jacobi identity, or more explicitely

$$\frac{1}{2}\big[A, [A, A]\big] = \frac{1}{2} A \wedge [A, A] - (-1)^{1 \cdot 2}[A, A] \wedge A = A \wedge (A \wedge A) - (A \wedge A) \wedge A = 0.$$

$\square$

Let E, E' be vector bundles over M. A differentiable map $\phi : E \to E'$ is an isomorphism (over id_M), if $\phi|_{E_p} : E_p \to E'_p$ is an isomorphism of vector spaces for all $p \in M$. By definition any bundle is locally trivial, in the sense that any $p \in M$ has a neighborhood U such that

$\pi^{-1}(U)$ is isomorphic to the product bundle $U \times \mathbb{C}^r$. But globally E need not be isomorphic to the product bundle $M \times \mathbb{C}^r$. Following S.S. Chern (1946), we now introduce cohomology classes on M which quantify the nontriviality of the bundle.

A symmetric k-linear form $p : \mathfrak{g} \times \ldots \times \mathfrak{g} \to \mathbb{C}$ is called invariant, if it satisfies:

$$p(gB_1g^{-1}, \ldots, gB_kg^{-1}) = p(B_1, \ldots, B_k) \quad \text{for all } B_1, \ldots, B_k \in \mathfrak{g},\ g \in G.$$

Taking $g(t) = \exp(tX)$ and differentiating at $t = 0$, we get

$$\sum_{j=1}^{k} p(B_1, \ldots, [X, B_j], \ldots, X_k) = 0 \quad \text{for all } B_1, \ldots, B_k \in \mathfrak{g},\ X \in \mathfrak{g}.$$

Putting $B_j = B$ for $j = 1 \ldots, k$ we obtain a polynomial $P(B)$ which is homogeneous of degree k and invariant under conjugation with $g \in G$. The symmetric k-linear form $p(B_1, \ldots, B_k)$ can be reconstructed from $P(B)$ by polarization. For $k = 1, \ldots, r$ we have the examples

$$P_k(B) = \Big(\frac{i}{2\pi}\Big)^k \operatorname{tr} \Lambda_k(B).$$

For $B = \operatorname{diag}(\lambda_1, \ldots, \lambda_r)$ the function $P_k(B)$ becomes the k-th elementary symmetric function of the eigenvalues $\lambda_1, \ldots, \lambda_r$, that is

$$P_k(B) = \sum_{1 \le i_1 < \ldots < i_k \le r} \lambda_{i_1} \cdot \ldots \cdot \lambda_{i_k}.$$

Using the Jordan normal form it is easy to see that the matrices which are conjugate to a diagonal matrix are dense in $\mathfrak{g}$.

Next we consider matrix-valued differential forms $B_i \in C^\infty(\Lambda^{m_i} TU \otimes \mathfrak{g})$ for $i = 1, \ldots, k$. Using the definition of the wedge product given in (16.56), we obtain a differential form

$$p(B_1, \ldots, B_k) \in C^\infty(\Lambda^m TU) \quad \text{where } m = m_1 + \ldots + m_k.$$

By the product rule of exterior differentiation, we have

$$d\big(p(B_1, \ldots, B_k)\big) = \sum_{i=1}^{k} (-1)^{m_1 + \ldots + m_{i-1}}\, p(B_1, \ldots, dB_i, \ldots, B_k). \tag{16.59}$$

We want to extend this definition to differential forms $\phi_i \in C^\infty(\Lambda^{m_i} TM \otimes \mathfrak{g}_E)$, i.e. the product bundle $M \times \mathfrak{g}$ is replaced by the bundle $\mathfrak{g}_E$. However we can reduce to matrix-valued forms by choosing a local basis $\mathcal{A}$ on U, and put

$$p(\phi_1, \ldots, \phi_k) = p\big((\phi_1)_{\mathcal{A}}, \ldots, (\phi_k)_{\mathcal{A}}\big) \in C^\infty(\Lambda^m TU).$$

For another basis $\mathcal{B}$ we have $(\phi_i)_{\mathcal{B}} = g\,(\phi_i)_{\mathcal{A}}\,g^{-1}$ on the overlap $U \cap V$, where $g = \varphi_{\mathcal{BA}}$. As the form p is assumed to be invariant, the definitions coincide on $U \cap V$ and hence we obtain a well-defined form $p(\phi_1, \ldots, \phi_k)$ in $C^\infty(\Lambda^m TM)$. In particular we can define

$$P(F_A) = p(F_A, \ldots, F_A) \in C^\infty(\Lambda^{2k} TM). \tag{16.60}$$

Lemma 16.3 *For an invariant k-linear form p and sections ϕ_i of $\Lambda^{m_i}TM \otimes \mathfrak{g}_E$ we have*

$$d\big(p(\phi_1,\dots,\phi_k)\big) = \sum_{i=1}^{k}(-1)^{m_1+\dots+m_{i-1}}\, p(\phi_1,\dots,d_A\phi_i,\dots,\phi_k). \tag{16.61}$$

Proof. Using a local basis on U we may assume that the ϕ_i are differential forms of degree m_i with values in $\mathfrak{g}$. By expanding into a basis we can in fact assume that $\phi_i = \theta_i B_i$ where θ_i is a standard m_i-form and $B_i : U \to \mathfrak{g}$. For A the connection 1-form, we compute

$$[A,\phi_i] = \sum_{\alpha=1}^{m}[dx^\alpha A_\alpha, \theta_i B_i] = \sum_{\alpha=1}^{m} dx^\alpha \wedge \theta_i [A_\alpha, B_i].$$

Interchanging the dx^α with $\theta_1,\dots,\theta_{i-1}$ we get

$$\begin{aligned} &\sum_{i=1}^{k}(-1)^{m_1+\dots+m_{i-1}} p(\phi_1,\dots,[A,\phi_i],\dots,\phi_k) \\ = \;& \sum_{\alpha=1}^{m} dx^\alpha \wedge \theta_1 \wedge \dots \wedge \theta_k \sum_{i=1}^{k} p(B_1,\dots,[A_\alpha,B_i],\dots,B_k) = 0, \end{aligned}$$

recalling the (infinitesimal) invariance of p. Combining with (16.59) we conclude

$$d\big(p(\phi_1,\dots,\phi_k)\big) = \sum_{i=1}^{k}(-1)^{m_1+\dots+m_{i-1}} p(\phi_1,\dots,d\phi_i + [A,\phi_i],\dots,\phi_k).$$

As $d_A\phi_i = d\phi_i + [A,\phi_i]$ in the local basis, the lemma follows. □

Theorem 16.2 (Chern-Weil) *Let D_A be a connection on the complex vector bundle E over M. For any invariant k-linear form p and associated polynomial P the following holds:*

(1) *The differential form $P(F_A) \in C^\infty(\Lambda^{2k}TM)$ is closed.*

(2) *The cohomology class $[P(F_A)] \in H^{2k}_{\mathrm{dR}}(M;\mathbb{C})$ is independent of D_A.*

Proof. The first statement follows by Lemma 16.3 and the Bianchi identity, namely

$$d\big(P(F_A)\big) = \sum_{i=1}^{k}(-1)^{2(i-1)} p(F_A,\dots,d_AF_A,\dots,F_A) = 0. \tag{16.62}$$

Now let D_0, D_1 be given connections. We put $\omega = D_1 - D_0 \in C^\infty(T^*M \otimes \mathfrak{g}_E)$ and consider $D_t = D_0 + t\omega$. In a local basis we have

$$\begin{aligned} F_{A_t} = dA_t + \frac{1}{2}[A_t,A_t] &= d(A_0 + t\omega) + \frac{1}{2}[A_0 + t\omega, A_0 + t\omega] \\ &= F_{A_0} + t(d\omega + [A_0,\omega]) + \frac{t^2}{2}[\omega,\omega]. \end{aligned}$$

Thus we have $\frac{\partial}{\partial t}F_{A_t} = d_{A_0+t\omega}\omega = d_{A_t}\omega$. This implies

$$\begin{aligned}\partial_t\big(p(F_{A_t},\dots,F_{A_t})\big) &= \sum_{i=1}^k p(F_{A_t},\dots,d_{A_t}\omega,\dots,F_{A_t})\\ &= d\sum_{i=1}^k p(F_{A_t},\dots,\omega,\dots,F_{A_t}).\end{aligned}$$

In the last step we use Lemma 16.3. Recalling that $d_{A_t}F_t = 0$ by Bianchi we verify

$$d\big(p(F_{A_t},\dots,\overset{i}{\omega},\dots,F_{A_t}) = \underbrace{(-1)^{2(i-1)}}_{=1}p(F_{A_t},\dots,d_{A_t}\overset{i}{\omega},\dots,F_{A_t}).$$

Now integrating with respect to $t\in[0,1]$ we conclude

$$P(F_{A_1}) - P(F_{A_0}) = d\int_0^1\sum_{i=1}^k p(F_{A_t},\dots,\overset{i}{\omega},\dots,F_{A_t})\,dt.$$

□

Chosing $p(F) = \frac{i}{2\pi}\operatorname{tr}F$ yields the first Chern class

$$c_1(E) = \Big[\frac{i}{2\pi}\operatorname{tr}F_A\Big] \in H^2(M;\mathbb{C}).$$

For structure group $U(r)$ one has $F_A\in\mathfrak{u}(r)$, which means $F_A^* = -F_A$ and hence $\operatorname{tr}F_A\in i\,\mathbb{R}$. It follows that $c_1(E)$ is a real cohomology class.

Let V be a real vector space of dimension m, equipped with a scalar product $\langle\cdot,\cdot\rangle$. To define the Hodge star operator, we select an orthonormal basis $e_1,\dots,e_m$. For $k=1,\dots,m$ the space $\Lambda_k(V)$ then has the orthonormal basis

$$e_\alpha = e_{\alpha_1}\wedge\cdots\wedge e_{\alpha_k}\quad\text{ where } \alpha\in I(m,k),\text{ i.e. } 1\le\alpha_1<\ldots<\alpha_k\le m.$$

By definition, the Hodge star is the linear map $*:\Lambda_k(V)\to\Lambda_{m-k}(V)$ such that

$$*e_\alpha = \operatorname{sign}(\alpha,\widehat{\alpha})\,e_{\widehat{\alpha}}\quad\text{ for } \alpha\in I(m,k).$$

Here $\widehat{\alpha}\in I(m,m-k)$ is the complementary increasing multi-index. We also define

$$*1 = e_1\wedge\ldots\wedge e_m\quad\text{ and }\quad *\,e_1\wedge\ldots\wedge e_m = 1.$$

A simple example is, for $1\le i\le m$,

$$*e_i = \operatorname{sign}(i,1,\dots,\widehat{i},\dots,m)\,e_1\wedge\ldots\wedge\widehat{e_i}\wedge\ldots\wedge e_m = (-1)^{i-1}e_1\wedge\ldots\wedge\widehat{e_i}\wedge\ldots\wedge e_m.$$

As a further example, we calculate

$$\begin{aligned}**e_\alpha &= *\big(\operatorname{sign}(\alpha,\widehat{\alpha})\,e_{\widehat{\alpha}}\big)\\ &= \operatorname{sign}(\alpha,\widehat{\alpha})\operatorname{sign}(\widehat{\alpha},\alpha)\,e_\alpha\\ &= (-1)^{k(m-k)}\operatorname{sign}(\alpha,\widehat{\alpha})^2\,e_\alpha\\ &= (-1)^{k(m-k)}\,e_\alpha.\end{aligned}$$

Lemma 16.4 *For $v, w \in \Lambda_k(V)$ we have*

$$v \wedge *w = \langle v, w\rangle \, e_1 \wedge \ldots \wedge e_m. \tag{16.63}$$

Proof. We compute

$$v \wedge *w = \sum_{\alpha,\beta \in I(m,k)} v^\alpha w^\beta \, e_\alpha \wedge *e_\beta = \sum_{\alpha,\beta \in I(m,k)} v^\alpha w^\beta \operatorname{sign}(\beta, \widehat{\beta}) \, e_\alpha \wedge e_{\widehat{\beta}}.$$

Now we see easily that

$$e_\alpha \wedge e_{\widehat{\beta}} = \begin{cases} 0 & \text{for } \alpha \neq \beta, \\ \operatorname{sign}(\beta, \widehat{\beta}) \, e_1 \wedge \ldots \wedge e_m & \text{for } \alpha = \beta. \end{cases}$$

Inserting we obtain

$$v \wedge *w = \Big(\sum_{\beta \in I(m,k)} v^\beta w^\beta \Big) \, e_1 \wedge \ldots \wedge e_m.$$

The lemma is proved. □

Clearly, the operator $** = (-1)^{k(m-k)}$ does not depend on the choice of the orthonormal basis. We now show that almost the same is true for $*$ itself.

Theorem 16.3 (Hodge star) *The Hodge star with respect to the orthonormal bases $e_1, \ldots, e_m$ and $e'_1, \ldots, e'_m$ coincides, if they have the same orientation.*

Proof. For $w \in \Lambda_k(V)$ let $\xi = *w - *'w$. The lemma implies, for any $v \in \Lambda_k V$,

$$v \wedge \xi = v \wedge *w - v \wedge *'w = \langle v, w\rangle \, (e_1 \wedge \ldots \wedge e_m - e'_1 \wedge \ldots \wedge e'_m) = 0.$$

The assumption on the orientation was used in the last equality. We conclude

$$0 = \sum_{\alpha \in I(m,k)} \sum_{\beta \in I(m,m-k)} v^\alpha \xi^\beta e_\alpha \wedge e_\beta.$$

For given $\beta \in I(m, m-k)$ we choose $v = e_\alpha$ where $\alpha = \widehat{\beta}$, and conclude

$$0 = \xi^\beta \, e_\alpha \wedge e_\beta = \operatorname{sign}(\alpha, \beta) \, \xi^\beta \, e_1 \wedge \ldots \wedge e_m.$$

This shows $\xi^\beta = 0$ and hence $\xi = 0$ as claimed. □

We now focus on the case $m = 4$, $k = 2$, hence $** = \mathrm{Id}$. Any $v \in \Lambda^2(V)$ decomposes as

$$v = \frac{1}{2}(v + *v) + \frac{1}{2}(v - *v) =: v^+ + v^-.$$

The decomposition $\Lambda^2(V) = \Lambda^+ V \oplus \Lambda^- V$ has as summands the eigenspaces of $*$ for the eigenvalues ± 1. From Lemma 16.4 we see that

$$v \wedge v = (|v^+|^2 - |v^-|^2) \, e_1 \wedge \ldots \wedge e_4.$$

In the Yang-Mills context the 2-vectors in $\Lambda^+ V$ are called selfdual (SD), the 2-vectors in $\Lambda^- V$ are called antiselfdual (ASD). An explicit basis of these spaces is as follows:

selfdual	antiselfdual
$e_1 \wedge e_2 + e_3 \wedge e_4$	$e_1 \wedge e_2 - e_3 \wedge e_4$
$e_2 \wedge e_3 + e_1 \wedge e_4$	$e_2 \wedge e_3 - e_1 \wedge e_4$
$e_3 \wedge e_1 + e_2 \wedge e_4$	$e_3 \wedge e_1 - e_2 \wedge e_4$.

To get an orthonormal basis, these 2-vectors should be multiplied by the factor $1/2$.

17 The Yang-Mills functional

In this section we consider a complex vector bundle $\pi : E \to M$, where M is a compact, 4-dimensional manifold with Riemannian metric g. We put $\dim E = r$ and assume that E carries a hermitian metric $\langle \cdot\,, \cdot \rangle$. We let

$$\begin{aligned} G_E &= \bigcup_{p \in M} \{g \in \mathrm{Aut}(E_p) : g^* g = \mathrm{Id}_{E_p}\}, \\ \mathfrak{g}_E &= \bigcup_{p \in M} \{a \in \mathrm{End}(E_p) : a^* + a = 0\}. \end{aligned}$$

By Remark 9.1 the set $\mathcal{C}$ of metric connections on E is nonempty. If $D_0 \in \mathcal{C}$ is fixed, then all connections $D \in \mathcal{C}$ are given by $D = D_0 + \omega$, where ω is a 1-form with values in $\mathfrak{g}_E$, that is a section of the bundle $\Lambda^1 TM \otimes \mathfrak{g}_E$. On $\mathfrak{g}_E$ we have the Hilbert-Schmidt norm

$$|a|^2 = \mathrm{tr}\,(a^* a) = -\mathrm{tr}\,(a^2).$$

For a connection $D \in \mathcal{C}$, the curvature F_D is a 2-form with values in $\mathfrak{g}_E$, i.e. a section of $\Lambda^2 TM \otimes \mathfrak{g}_E$. Combining the norms yields a curvature energy density at any p

$$|F_D|^2 = \frac{1}{2} \sum_{\alpha,\beta=1}^{4} |F_D(e_\alpha, e_\beta)|^2 \quad \text{where } e_\alpha \text{ is an orthonormal basis of } TM.$$

The Yang-Mills energy is obtained by integrating that density on M:

$$YM : \mathcal{C} \to \mathbb{R},\ YM(D) = \int_M |F_D|^2 \, d\mu_g.$$

A fundamental property of the functional is the gauge invariance. We define the group of gauge transformations of the bundle by

$$\mathcal{G} = C^\infty(G_E) = \{g \in C^\infty(\mathfrak{g}_E) : g^* g = \mathrm{Id}\}.$$

Any $g \in \mathcal{G}$ acts on $\mathcal{C}$ by

$$g(D)v = g\, D(g^{-1} v) = Dv - (Dg)\, g^{-1} v.$$

Recall that D is defined on sections of $\mathfrak{g}_E$ using the product rule.

Lemma 17.1 (gauge invariance) *For all $D \in \mathcal{C}$, $g \in \mathcal{G}$ we have*

$$F_{g(D)} = g\, F_D\, g^{-1}, \quad \text{hence } |F_{g(D)}|^2 = |F_D|^2.$$

In particular $YM(g(D)) = YM(D)$.

Proof. Let D and $g(D)$ have connection 1-forms A, B in a local basis $\mathcal{A}$. We compute

$$\begin{aligned} B &= A - ((Dg) g^{-1})_{\mathcal{A}} \\ &= A - \big(dg_{\mathcal{A}} + [A, g_{\mathcal{A}}]\big) g_{\mathcal{A}}^{-1} \\ &= A - dg_{\mathcal{A}}\, g_{\mathcal{A}}^{-1} - A + g_{\mathcal{A}} A\, g_{\mathcal{A}}^{-1} \\ &= g_{\mathcal{A}} A\, g_{\mathcal{A}}^{-1} - dg_{\mathcal{A}}\, g_{\mathcal{A}}^{-1}. \end{aligned}$$

Thus B is exactly the connection 1-form of D with respect to the basis $\mathcal{B} = g^{-1}\mathcal{A}$, since then $\varphi_{\mathcal{B}\mathcal{A}} = g_{\mathcal{A}}$. Now in the basis $\mathcal{B}$ the connection D has the curvature

$$(F_{g(D)})_{\mathcal{A}} = dB + B \wedge B = (F_D)_{\mathcal{B}} = \varphi_{\mathcal{B}\mathcal{A}}\,(F_D)_{\mathcal{A}}\,\varphi_{\mathcal{B}\mathcal{A}}^{-1} = (g\,F_D\,g^{-1})_{\mathcal{A}}.$$

This proves the formula. The invariance of the Hilbert-Schmidt norm follows, using $g^*g = \mathrm{Id}$,

$$\operatorname{tr}(gF_Dg^{-1})^*gF_Dg^{-1} = \operatorname{tr} g(F_D)^*F_Dg^{-1} = \operatorname{tr}(F_D)^*F_D.$$

□

We note that the group $\mathcal{G}$ of gauge transformations is huge, formally its tangent space at the identity is $C^\infty(\mathfrak{g}_E)$.

Lemma 17.2 (Conformal invaraince) *For* $\dim M = 4$ *we have*

$$|F_D|^2_{\tilde{g}}\,d\mu_{\tilde{g}} = |F_D|^2_g\,d\mu_g \quad \textit{for } \tilde{g} = e^{2u}g.$$

In particular $YM(D,\tilde{g}) = YM(D,g)$.

Proof. If e_i is an orthonormal basis for g, then $e^{-u}e_i$ is an orthonormal basis for $\tilde{g}$. Since the curvature is a 2-form, the energy density gets a factor e^{-4u} which cancels with the factor e^{4u} in the 4-dimensional volume element. □

From the viewpoint of geometry these invariances are kind of beautiful. However from the viewpoint of analysis, they pose serious difficulties. A sequence of connections with bounded Yang-Mills energy need not subconverge. On the one hand, there is the possibility of concentrations of curvature density, similar to the Yamabe problem. This is due to conformal or scaling invariance, we will see an example below. Secondly, there is the possibility of divergence due to bad gauge transformations, which is also a central issue for Yang-Mills.

We now compute the Euler-Lagrange equation of the functional.

Lemma 17.3 (Yang-Mills equation) *If D is a critical point of the Yang.Mills energy with respect to compactly supported variations, then it satisfies the Yang-Mills equation*

$$d_A^* F_A = 0.$$

Proof. From Theorem 16.2 we know that the connections $D_t = D + t\omega$ have curvature

$$F_{A_t} = F_A + t\,d_A\omega + \frac{t^2}{2}\,[\omega,\omega].$$

In particular $\frac{\partial}{\partial t}F_{A_t}|_{t=0} = d_A\omega$. The claim of the lemma follows. □

For a Yang-Mills connection we thus have $d_A^*F_A = 0$ and $d_AF_A = 0$ (Bianchi).

Definition 17.1 *A connection D_A on an oriented Riemannian 4-manifold is called*

$$\begin{aligned} \textit{selfdual (SD)} \quad &\Leftrightarrow \quad F_A^- = 0,\\ \textit{antiselfdual (ASD)} \quad &\Leftrightarrow \quad F_A^+ = 0. \end{aligned}$$

Theorem 17.1 *Let D_A be a metric connection on a hermitian vector bundle over an oriented 4-manifold M. If D_A is self-dual (or antiselfdual), then it minimizes the Yang-Mills energy on any open set $\Omega \subset\subset M$ subject to Dirichlet boundary conditions.*

Proof. We consider the invariant symmetric bilinear form

$$p : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C},\ p(B_1, B_2) = \mathrm{tr}\,(B_1 B_2).$$

Let $\omega_i^{\pm}$, $i = 1, 2, 3$, be an orthonormal basis of $\Lambda^{\pm}(TM)$. Decompose

$$F_A = \sum_{i=1}^{3} F_i^+ \otimes \omega_i^+ + \sum_{i=1}^{3} F_i^- \otimes \omega_i^-, \quad \Rightarrow \quad F_A \wedge F_A = \sum_{i=1}^{3} \big((F_i^+)^2 - (F_i^-)^2\big)\, d\mu_g.$$

For $B \in U(r)$ we have $\mathrm{tr}\, B^2 = -\mathrm{tr}\, B^* B = -|B|^2$. Thus we obtain

$$P(F) = (|F_A^-|^2 - |F_A^+|^2)\, d\mu_g.$$

Now let ω be a smooth $\mathfrak{g}_E$-valued 1-form on $\overline{\Omega}$ with $\omega|_{\partial\Omega} = 0$. By Chern-Weil theory, see Theorem 16.2, we have for $D + \omega$ by Gauß-Stokes

$$\int_\Omega (P(F_{D+\omega}) - P(F_D)) = 2 \int_0^1 \int_\Omega d\,\mathrm{tr}\,(F_{D+t\omega}\, \omega) = 0.$$

If D_A is antiselfdual, then we conclude

$$\begin{aligned} YM(D+\omega) &= \int_\Omega (|F^-_{D+\omega}|^2 + |F^+_{D+\omega}|^2)\, d\mu_g \\ &\geq \int_\Omega (|F^-_{D+\omega}|^2 - |F^+_{D+\omega}|^2)\, d\mu_g \\ &= \int_\Omega (|F^-_D|^2 - |F^+_D|^2)\, d\mu_g \\ &= YM(D). \end{aligned}$$

□

The argument applies of course also on the whole manifold M, when there is no boundary term at all, and hence SD or ASD connections are absolute minimizers. We point out an analogy to the Dirichlet energy for maps from a surface into $\mathbb{S}^2$. The role of the Chern class is then taken by the degree of the map, and the (anti-)holomorphic maps minimize with given degree. We now briefly mention an example of an ASD connection, discovered by Belavin, Polyakov, Schwarz and Tyupkin (1975).

On $M = \mathbb{R}^4 = \mathbb{C} \times \mathbb{C}$, with standard basis e_0, e_1, e_2, e_3 one lets

$$\gamma(a, b) = \begin{pmatrix} a & b \\ -\bar{b} & a \end{pmatrix}.$$

One checks $\gamma(a, b)^* \gamma(a, b) = (|a|^2 + |b|^2)\mathrm{E}_2$, and notes the values

$$\gamma(e_0) = \mathrm{E}_2, \quad \gamma(e_1) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad \gamma(e_2) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \gamma(e_3) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

One now checks that the following defines a $\mathfrak{su}(2)$-connection:

$$A(x)v = \frac{1}{1+|x|^2}\big(\langle x, v\rangle \mathrm{E}_2 + \gamma(x)^*\gamma(v)\big).$$

One calculates further

$$F_A(x) = \frac{2}{(1+|x|^2)^2}\sum_{i=1}^{3}\omega_i \otimes \gamma(e_i),$$

where the forms ω_i are a basis of $\Lambda^-(\mathbb{R}^4)$, explicitly

$$\begin{aligned}\omega_1 &= dx^0\wedge dx^1 - dx^2 \wedge dx^3,\\ \omega_2 &= dx^0\wedge dx^2 - dx^3 \wedge dx^1,\\ \omega_3 &= dx^0\wedge dx^3 - dx^1 \wedge dx^2.\end{aligned}$$

This shows that the connection is ASD. The energy desnsity is given by

$$|F_A(x)|^2 = \frac{48}{(1+|x|^2)^4}.$$

The ASD equation is invariant under translations and dilations of $\mathbb{R}^4$. In this way we generate a manifold $\mathcal{M}$ of solutions, with connection 1-form $A_{x_0,\lambda} = \sigma^*_{x_0,\lambda}A$ and curvature density

$$|F_{A_{x_0,\lambda}}(x)|^2 = \frac{48\lambda^4}{(\lambda^2+|x-x_0|^2)^4}.$$

The manifold $\mathcal{M}$ is parametrized by its center $x_0 \in \mathbb{R}^4$ and its scale $\lambda > 0$. For $\lambda \searrow 0$, the solutions concentrate at the center point x_0.

We now consider the problem of finding a good gauge for a given $U(r)$-connection. The gauge orbit of D_A is

$$\mathcal{G}(D_A) = \{g(D_A) = g\, D_A\, g^{-1} : g \in \mathcal{G}\}.$$

Putting $g(t) = e^{t\chi}$ where χ is a section of $\mathfrak{g}_E$, we compute

$$\frac{d}{dt}e^{t\chi}D_Ae^{-t\chi}|_{t=0} = -d_A\chi.$$

For two $U(r)$-connections D_A, D_B we introduce the L^2-distance

$$\operatorname{dist}(D_A, D_B)^2 = \int_M |D_A - D_B|^2\, d\mu_g.$$

The group $\mathcal{G}$ acts isometrically,

$$\operatorname{dist}(g(D_A), g(D_B)) = \operatorname{dist}(gD_Ag^{-1}, gD_Bg^{-1}) = g(D_A, D_B).$$

In is natural to look for a connection in the gauge orbit of D_B, which minimizes the L^2 distance to D_A, i.e. to solve the variational problem

$$\int_M |D_A - gD_Bg^{-1}|^2\, d\mu_g \longrightarrow \min \qquad \text{among all } g \in \mathcal{G}.$$

Assuming that D_B realizes this infimum, we compute

$$\begin{aligned} 0 &= \frac{d}{dt}\int_M |D_A - e^{t\chi} D_B e^{-t\chi}|^2 \, d\mu_g|_{t=0} \\ &= \frac{d}{dt}\int_M |e^{-t\chi} D_A e^{t\chi} - D_B|^2 \, d\mu_g|_{t=0} \\ &= 2\int_M \langle D_A - D_B, d_A\chi\rangle \, d\mu_g. \end{aligned}$$

Therefore D_B should satisfy the condition

$$d_A^*(D_B - D_A) = 0. \tag{17.64}$$

We say that D_B is in Coulomb gauge relative to D_A. Obviously this property is symmetric with respect to D_A, D_B. Due to possible concentrations, as for the BPST solutions, we will need the following local version.

Definition 17.2 (local Coulomb gauge) *Let $D_A = d + A$ be the local representation with respect to some orthonormal basis. Then D_A is in Coulomb gauge if it satisfies $d^*A = 0$.*

Here d^* is the standard co-differential of d. Under this gauge condition, the Yang-Mills equation can be written as follows (using formal notation):

$$\begin{aligned} 0 &= d_A^* F_A \\ &= (d^* + A)(dA + A * A) \\ &= \underbrace{(d^*d + dd^*)}_{=\Delta} A + A * DA + A * A * A \qquad \text{as } d^*A = 0. \end{aligned}$$

The proof of regularity for weak solutions $A \in W^{1,2}$ is similar to our discussion for the Yamabe equation, using Sobolev and Calderon-Zygmund estimates. We postpone the details.

Theorem 17.2 (Uhlenbeck 1982) *There exist universal constants $\varkappa > 0$ and $C < \infty$ such that the following holds: if for $p \in [2,3]$ the connection D_A is in the class*

$$\mathcal{C}^{1,p}_{\varkappa} = \{D_A = d + A : A \in W^{1,p}(B^4, \Lambda^1\mathbb{R}^4 \otimes \mathfrak{u}(r)),\ \|F_A\|_{L^2(B^4)} \le \varkappa\}, \tag{17.65}$$

then there exists $g \in W^{2,p}(B^4, U(r))$ such that $D_B = gD_Ag^{-1}$ satisfies

$$d^*B = 0 \text{ in } B^4, \qquad B(\nu) = 0 \text{ on } \partial B^4, \tag{17.66}$$

$$\|B\|_{W^{1,2}} \le C\,\|F_A\|_{L^2}, \tag{17.67}$$

$$\|B\|_{W^{1,p}} \le C\,\|F_A\|_{L^p}. \tag{17.68}$$

Proof. We first note that the class $\mathcal{C}^{1,p}_{\kappa}$ is connected. Namely, any D_A connects to the trivial connection D via the path D_{A_λ}, $\lambda \in [0,1]$, where

$$A_\lambda(x) = (\tau_\lambda^* A)(x) \quad \text{for } \tau_\lambda(x) = \lambda x.$$

We have $A_\lambda(x) = \lambda A(\lambda x)$ and hence $F_{A_\lambda}(x) = \lambda^2 F_A(\lambda x)$, so that

$$\|F_{A_\lambda}\|_{L^2(B^4)} = \|F_A\|_{L^2(\lambda B^4)} \le \varkappa.$$

Our strategy is to show that the set of $D_A \in \mathcal{C}^{1,p}_\varkappa$ having a gauge $D_B = gD_Ag^{-1}$ with the desired properties (17.66), (17.67) and (17.68) is closed and open in $\mathcal{C}^{1,p}_\varkappa$. As the trivial connection belongs to this set, taking $g = \mathrm{E}_r$, the theorem follows.

Closedness of the set. This step may be expected as the easier part, since the equations (17.66), (17.67) and (17.68) are formulated as closed conditions. Let $D_{A_i} \to D_A$ in $\mathcal{C}^{1,p}_\varkappa$, and assume that for $g_i \in W^{2,p}(B^4, U(r))$ the connections $D_{B_i} = g_i D_{A_i} g_i^{-1}$ satisfy the three equations. Thus

$$\|B_i\|_{W^{1,2}} \leq C\|F_{A_i}\|_{L^2} \leq C\varkappa.$$

Using Hölder's inequality and Sobolev, we estimate independent of $i \in \mathbb{N}$

$$\begin{aligned}\|F_{A_i}\|_{L^p} &\leq \|dA_i\|_{L^p} + \|A_i \wedge A_i\|_{L^p}\\ &\leq \|dA_i\|_{L^p} + C\|A_i\|_{L^4}\|A_i\|_{L^{4p/(4-p)}}\\ &\leq \|dA_i\|_{L^p} + C\|A_i\|^2_{W^{1,p}} \leq \text{Const.}\end{aligned}$$

From equation (17.68) we obtain

$$\|B_i\|_{W^{1,p}} \leq C\|F_{A_i}\|_{L^p} \leq \text{Const.}$$

After passing to a subsequence, we have $B_i \to B$ weakly in $W^{1,p}(B^4)$, and B solves (17.66). Now $A_i \to A$ strongly in $W^{1,p}$ by assumption, which implies $F_{A_i} \to F_A$ in L^p. We conclude by lower semicontinuity of the norm under weak convergence

$$\|B\|_{W^{1,p}} \leq \liminf_{i\to\infty} \|B_i\|_{W^{1,p}} \leq C \lim_{i\to\infty} \|F_{A_i}|_{L^p} = C\|F_A\|_{L^p}.$$

The same argument applies for $\|B\|_{W^{1,2}}$, so that (17.67) and (17.68) are settled. It remains to show that B is obtained form A by a gauge transformation. We calculate

$$\begin{aligned}g_i D_{A_i} g_i^{-1} &= D_{A_i} - (d_{A_i} g_i)g_i^{-1}\\ &= d + A_i - (dg_i + [A_i, g_i])g_i^{-1}\\ &= d + g_i A_i g_i^{-1} - (dg_i)g_i^{-1}.\end{aligned}$$

Now $d + B_i = D_{B_i} = g_i D_{A_i} g_i^{-1}$, and we get

$$dg_i = g_i A_i - B_i g_i.$$

Since A_i, B_i are bounded in $W^{1,p}$, it is easy see that g_i is bounded in $W^{2,p}$, and the equation passes to limits showing that $D_B = gD_Ag^{-1}$.

Lemma 17.4 *There exist constants $\varkappa_1 > 0$, $C_1 < \infty$ such that if B solves (17.66) and*

$$\|B\|_{L^4} \leq \varkappa_1,$$

then $\|B\|_{W^{1,p}} \leq C_1\|F_B\|_{L^p}$ for any $p \in [2,3]$.

Proof. For B satisfying (17.66) one can compute the identity

$$\int_B |dB|^2 = \int_B |DB|^2 + \int_{\partial B} |B|^2.$$

This implies $\|B\|_{W^{1,2}} \le C\|dB\|_{L^2}$. Now one can prove, using Calderon-Zygmund estimates,

$$\|B\|_{W^{1,p}} \le C\|dB\|_{L^p}.$$

Combining Hölder and Sobolev as before, we can estimate

$$\begin{aligned}
\|B\|_{W^{1,p}} &\le C\|dB\|_{L^p} \\
&\le C\|F_B\|_{L^p} + C\|B \wedge B\|_{L^p} \\
&\le C\|F_B\|_{L^p} + C\|B\|_{L^4}\|B\|_{L^{4p/(4-p)}} \\
&\le C\|F_B\|_{L^p} + C\varkappa_1\|B\|_{W^{1,p}}.
\end{aligned}$$

The claim of the lemma follows by absorbing to the left hand side, choosing $C\varkappa_1 = \frac{1}{2}$. $\square$

Openness of the set. Let $D_B \in \mathcal{C}^{1,p}_{\varkappa}$ be a connection with (17.66), (17.67) and (17.68). For a given pertubation $\tilde{B} = B + b$, we want to find a gauge transformation $g = e^{\chi}$ such that $gD_{B+b}g^{-1} = d + g(B+b)g^{-1} - (dg)g^{-1}$ satisfies (17.66). To achieve the boundary condition, we assume that $\chi \in X_0$, $b \in Y_0$ where

$$\begin{aligned}
X_0 &= \{\chi \in W^{2,p}(B^4, \mathfrak{u}(r)) : d\chi(\nu) = 0 \text{ on } \partial B^4\}, \\
Y_0 &= \{b \in W^{1,p}(B^4, \mathfrak{u}(r)) : b(\nu) = 0 \text{ on } \partial B^4\}.
\end{aligned}$$

The Coulomb equation becomes

$$G(\chi, b) = d^*\big((de^{\chi})e^{-\chi} - e^{\chi}(B+b)e^{-\chi}\big) = 0.$$

To apply the implicit function theorem we need that $G(\chi, b)$ is of class C^1. Functions in $W^{2,2}$ may be discontinuous, therefore we assume from now on that $p > 2$. Linearizing $G(\chi, b)$ at the point $(0,0)$ yields the operator $L : X_0 \to Z = L^p(B^4, \mathfrak{u}(r))$ given by

$$L\chi = D_1G(0,0)\chi = d^*d\chi + \langle B, d\chi\rangle - \langle d\chi, B\rangle.$$

To have L invertible we need to put the conditions

$$\begin{aligned}
X_0^{\perp} &= \{\chi \in X_0 : \int_{B^4} \chi = 0\}, \\
Z^{\perp} &= \{\psi \in Z : \int_{B^4} \psi = 0\}.
\end{aligned}$$

Note that G maps into $Z^{\perp}$, in fact the divergence theorem gives

$$\int_{B^4} G(\chi, b) = -\int_{\partial B^4} (de^{\chi})(\nu)e^{-\chi} - e^{\chi}(B+b)(\nu)e^{-\chi}\big) = 0.$$

Claim 1. *There exists $\varkappa_2 > 0$ such that whenever B solves (17.66) with $\|B\|_{L^4} < \varkappa_2$, then for any small $b \in Y_0$ the equation $G(\chi, b) = 0$ has a unique small solution $\chi \in X_0^{\perp}$.*

Proof (sketch). We show that $L : X_0^{\perp} \to Z^{\perp}$ is invertible, then the result follows by the implicit function theorem. By L^p-theory the operator $d^*d : X_0^{\perp} \to Z^{\perp}$ is an isomorphism. We cannot prove this here, but at least we indicate that d^*d has both kernel and co-kernel zero. If $d^*d\chi = 0$ then one obtains using $d\chi(\nu) = 0$

$$\int_{B^4} |d\chi|^2 = \int_{B^4} \langle d^*d\chi, \chi\rangle = 0.$$

Thus $\chi = 0$ as $\chi \in X_0^\perp$. Now assume there exists a $\psi \in Z^\perp$ such that

$$\int_{B^4} \langle d^* d\chi, \psi \rangle = 0 \text{ for all } \chi \in X_0^\perp.$$

If ψ was regular we would conclude $d^* d\psi = 0$ in B^4 and $d\psi(\nu) = 0$ on ∂B^4, hence $\psi = 0$. So much to the result for $d^* d$.

Using Hölder and Sobolev, the lower order term in L is now estimated as follows.

$$\|\langle B, d\chi \rangle - \langle d\chi, B \rangle\|_{L^p} \leq C\, \|B\|_{L^4} \|d\chi\|_{L^{4p/(4-p)}} \leq C\, \|B\|_{L^4}\, \|\chi\|_{W^{2,p}}.$$

For suitable $\varkappa_2 > 0$ the operator L is also invertible. □

Claim 2. *There exist constants $\varkappa > 0$, $C < \infty$ such that the set of connections $D_A \in \mathcal{C}^{1,p}_\varkappa$ admitting a $W^{2,p}$-gauge with (17.66), (17.67) and (17.68) is open. This proves the theorem in the case $p > 2$.*

Proof. We start with a constant $\varkappa > 0$, collecting the conditions along the proof. Let $A \in \mathcal{C}^{1,p}_\varkappa$, $g \in W^{2,p}(B^4, U(r))$ and $D_B = g D_A g^{-1}$ with (17.66), (17.67) and (17.68). Let $a \in W^{1,p}(B^4, \Lambda^1\mathbb{R}^4 \otimes \mathfrak{u}(r))$ be given, and put $b = gag^{-1}$. By (17.68) we have using Sobolev

$$\|B\|_{W^{1,p}} \leq C\, \|F_A\|_{L^p} \leq C\, \|A\|_{W^{1,p}}.$$

From $g \in U(r)$ and $dg = gA - Bg$ we then get

$$\|g\|_{W^{2,p}} \leq C(\|A\|_{W^{1,p}}). \tag{17.69}$$

Then further, again by Sobolev,

$$\|b\|_{W^{1,p}} \leq C\big(\|dg\|_{L^4} \|a\|_{L^{4p/(4-p)}} + \|a\|_{W^{1,p}}\big) \leq C(\|A\|_{W^{1,p}})\, \|a\|_{W^{1,p}}. \tag{17.70}$$

To apply Claim 1 with $b = gag^{-1}$ we need $b \in Y_0$, however b does not satisfy the boundary condition $b(\nu) = 0$. We handle this by first performing another gauge, based on the following auxiliary technical result.

Lemma 17.5 *There is a bounded operator $E : W^{1,p}(B^4) \to W^{2,p}(B^4)$ such that $\psi = E\varphi$ satsfies $\psi = 0$ and $d\psi(\nu) = \varphi$ on ∂B^4.*

Such an operator can be constructed using convolution. Writing $b = b_i(x)dx^i$ we apply the Lemma with $\varphi = b_i(x)x^i$. For $\psi = E\varphi$ we then have

$$\psi = 0 \text{ and } d\psi(\nu) = b_i(x)x^i = b(\nu) \quad \text{ on } \partial B^4.$$

Moreover we have the estimate

$$\|\psi\|_{W^{2,p}} \leq C\, \|\varphi\|_{W^{[}1,p} \leq \|b\|_{W^{1,p}} \leq C(\|A\|_{W^{1,p}})\, \|a\|_{W^{1,p}}. \tag{17.71}$$

Define $\tilde{b}$ by writing

$$e^\psi (D_B + b) e^{-\psi} = d + e^\psi (B + b) e^{-\psi} - (de^\psi) e^{-\psi} =: d + B + \tilde{b}.$$

On ∂B^4 we have $\tilde{b}(\nu) = b(\nu) - d\psi(\nu) = 0$. Moreover

$$\|\tilde{b}\|_{W^{1,p}} = \|e^{\psi} B e^{-\psi} - B + e^{\psi} b e^{-\psi} - (de^{\psi})e^{-\psi}\|_{W^{1,p}}.$$

Now $a \to 0$ in $W^{1,p}$ implies $b \to 0$ in $W^{1,p}$ and $\psi \to 0$ in $W^{2,p}$, so that also $\tilde{b} \to 0$ in $W^{1,p}$. By Sobolev and (17.67) we have

$$\|B\|_{L^4} \leq \|B\|_{W^{1,2}} \leq C\|F_A\|_{L^2} \leq C\varkappa.$$

For $C\varkappa \leq \varkappa_2$ and $\|a\|_{W^{1,p}}$ sufficiently small, Claim 1 yields a (small) function $\chi \in X_0^{\perp}$ such that the following connection solves (17.66):

$$D_{\tilde{B}} = e^{\chi} D_{B+\tilde{b}} e^{-\chi} = e^{\chi} e^{\psi} g(D_A + a) g^{-1} e^{-\psi} e^{-\chi}.$$

It remains to check that $\tilde{B}$ satisfies the estimates. We have

$$\|\tilde{B}\|_{L^4} \leq \|e^{\chi}(B + \tilde{b})e^{-\chi} - (de^{\chi})e^{-\chi}\|_{L^4} \leq C\left(\|B\|_{L^4} + \|\tilde{b}\|_{L^4} + \|d\chi\|_{L^4}\right).$$

By Sobolev, we know that $\tilde{b} \to 0$ and $d\chi \to 0$ in L^4 as $a \to 0$ in $W^{1,p}$. Moreover $\|B\|_{L^4} \leq C\varkappa$. Thus for $C\varkappa \leq \frac{1}{2}\varkappa_1$, and $\|a\|_{W^{1,p}}$ small, Lemma 17.4 applies to $\tilde{B}$. Therefore we have the estimates, for all $p \in [2,3]$,

$$\|\tilde{B}\|_{W^{1,p}} \leq C_1\|F_{\tilde{B}}\|_{L^p} = C_1\|F_{A+a}\|_{L^p}.$$

To sum up, for $p > 2$ the openness is obtained if $C\varkappa \leq \min(\varkappa_1, \varkappa_2)$.

The case $p = 2$. We choose $\tilde{A}_i \in C^\infty(\overline{B}^4)$ with $A_i \to A$ in $W^{1,2}$. Then $F_{A_i} \to F_A$ in L^2 by Sobolev. Applying the case $p = 3$, we obtain $g_i \in W^{2,3}(B^4, U(r))$ such that $D_{B_i} = g_i D_{A_i} g_i^{-1}$ satisfies

$$d^* B_i = 0 \text{ in } B^4, \quad B_i(\nu) = 0 \text{ on } \partial B^4.$$

We also have $\|B_i\|_{W^{1,2}} \leq \|F_{A_i}\|_{L^2}$. We now use the closedness assertion of the Theorem to obtain the desired gauge as limit. □

18 Appendix

Definition 18.1 (Teilung der Eins) *Sei M eine n-dimensionale differenzierbare Mannigfaltigkeit. Eine Teilung der Eins auf M ist eine Familie χ_i, $i \in I$, von glatten Funktionen mit folgenden Eigenschaften:*

(a) *das System* $\operatorname{spt}\chi_i$*, $i \in I$, ist lokal endlich: jedes $p \in M$ hat eine offene Umgebung U, so dass* $\operatorname{spt}\chi_i \cap U \neq \emptyset$ *nur für endlich viele $i \in I$.*

(b) $\sum_{i \in I} \chi_i \equiv 1$.

Die Teilung der Eins heißt der Überdeckung V_λ, $\lambda \in \Lambda$, untergeordnet, wenn es zu jedem $i \in I$ ein $\lambda \in \Lambda$ gibt mit $\operatorname{spt}\chi_i \subset V_\lambda$.

Beachte: der Träger der Nullfunktion ist nach Definition die leere Menge.

Theorem 18.1 (Teilung der Eins) *Sei M eine n-dimensionale differenzierbare Mannigfaltigkeit, und V_λ, $\lambda \in \Lambda$, eine offene Überdeckung von M. Dann gibt es eine untergeordnete Teilung der Eins χ_i, $i \in I$, mit $\operatorname{spt} \chi_i$ kompakt für alle $i \in I$.*

Bemerkung. Ist $K \subset M$ kompakt, so ist die Menge der $i \in I$ mit $\operatorname{spt} \chi_i \cap K \neq \emptyset$ endlich wegen Eigenschaft (a) in Definition 18.1. Für M kompakt ist also I endlich. Allgemein ist I abzählbar, da M durch eine Folge kompakter Mengen ausgeschöpft ist, siehe Satz 1.1.

Zum Beweis des Satzes fixieren wir eine Funktion $\chi \in C^\infty_c(\mathbb{R}^n)$ mit $\chi = 1$ auf $[-1,1]^n$ und $\operatorname{spt} \chi \subset (-2,2)^n$.

Beweis: Sei $\emptyset = G_0 \subset\subset G_1 \subset \ldots$ die Ausschöpfung aus Satz 1.1. Für $p \in M$ sei $i_p = \max\{i \in \mathbb{N}_0 : p \notin \overline{G}_i\}$. Es folgt $p \in G_{i_p+2} \backslash \overline{G}_{i_p}$. Wähle $\lambda_p \in \Lambda$ mit $p \in V_{\lambda_p}$. Es gibt dann eine Karte (U_p, φ_p) mit folgenden Eigenschaften:

(a) $U_p \subset V_{\lambda_p} \cap G_{i_p+2} \backslash \overline{G}_{i_p}$,

(b) $\varphi_p(p) = 0$ und $\varphi_p(U_p) \supset [-2,2]^n$.

Definiere $\eta_p \in C^\infty(M)$ durch

$$\eta_p = \begin{cases} \chi \circ \varphi_p & \text{auf } U_p, \\ 0 & \text{sonst.} \end{cases}$$

Nach Konstruktion ist $\operatorname{spt} \eta_p \subset\subset U_p \subset V_{\lambda_p}$, und es gilt

$$\operatorname{spt} \eta_p \cap G_k \neq \emptyset \quad \Rightarrow \quad G_{i_p+2} \backslash \overline{G}_{i_p} \cap G_k \neq \emptyset \quad \Rightarrow \quad k \geq i_p + 1.$$

Setze $W_p = \varphi_p^{-1}([-1,1]^n)$, also $\eta_p = 1$ auf W_p.

Wähle nun zu $j \in \mathbb{N}_0$ Punkte $p_{j,\nu} \in \overline{G}_{j+1} \backslash G_j$, $1 \leq \nu \leq N_j$, so dass $\overline{G}_{j+1} \backslash G_j$ durch die $W_{p_{j,\nu}}$ überdeckt wird. Wegen $p_{j,\nu} \notin G_j \supset \overline{G}_{j-1}$ ist $i_{p_{j,\nu}} \geq j - 1$. Damit gilt

$$\operatorname{spt} \eta_{p_{j,\nu}} \cap G_k \neq \emptyset \quad \Rightarrow \quad j \leq i_{p_{j,\nu}} + 1 \leq k.$$

Also sind die Träger $\operatorname{spt} \eta_{p_{j,\nu}}$ mit $j \in \mathbb{N}_0$, $1 \leq \nu \leq N_j$, lokal endlich. Definiere schließlich

$$\eta = \sum_{j=0}^{\infty} \sum_{1 \leq \nu \leq N_j} \eta_{p_{j,\nu}},$$

und erhalte die Teilung der Eins $\chi_{j,\nu} = \eta_{p_{j,\nu}}/\eta$. □

Theorem 18.2 (Teilung der Eins mit gleicher Indexmenge) *Sei M eine n-dimensionale differenzierbare Mannigfaltigkeit, und V_λ, $\lambda \in \Lambda$, eine offene Überdeckung von M. Dann gibt es eine untergeordnete Teilung der Eins φ_λ, $\lambda \in \Lambda$.*

Im Gegensatz zu Satz 18.1 kann hier nicht verlangt werden, dass $\operatorname{spt} \varphi_\lambda$ kompakt ist. Das sieht man schon am Fall $M = \mathbb{R}$ mit Überdeckung $U = \mathbb{R}$.

Beweis: Seien χ_i, $i \in I$, wie in Satz 18.1. Wähle für alle $i \in I$ ein $\lambda(i) \in \Lambda$ mit $\operatorname{spt}\chi_i \subset V_{\lambda(i)}$, und setze

$$I(\lambda) = \{i \in I : \lambda(i) = \lambda\} \quad \text{und} \quad \varphi_\lambda = \sum_{i \in I(\lambda)} \chi_i.$$

φ_λ ist in $C^\infty(M)$ wohldefiniert, da das System $\operatorname{spt}\chi_i$ lokal endlich ist. Ist $\operatorname{spt}\varphi_\lambda \cap U \neq \emptyset$, so folgt $\operatorname{spt}\chi_i \cap U \neq \emptyset$ für ein $i \in I(\lambda)$. Da $I(\lambda) \cap I(\mu) = \emptyset$ für $\lambda \neq \mu$, ist das System $\operatorname{spt}\varphi_\lambda$ lokal endlich. Weiter gilt für $p \in M$

$$\sum_{\lambda \in \Lambda} \varphi_\lambda(p) = \sum_{\lambda \in \Lambda} \sum_{i \in I(\lambda)} \chi_i(p) = \sum_{i \in I} \chi_i(p) = 1.$$

Schließlich gibt es zu $p \in \operatorname{spt}\varphi_\lambda$ ein $i \in I(\lambda)$ mit $p \in \operatorname{spt}\chi_i$, somit gilt $p \in V_{\lambda(i)} = V_\lambda$, das heißt die Teilung der Eins ist der Überdeckung V_λ, $\lambda \in \Lambda$, untergeordnet. $\square$

Diese Variante wird in Satz 6.3 und in Satz 8.3 gebraucht.

Literatur

[1] Gallot, S., Hulot, D., Lafontaine, J.: *Riemannian Geometry*, 2. edition, Springer Universitext, 1990.

[2] Gromoll, D., Klingenberg, W., Meyer, W.: *Riemannsche Geometrie im Großen*, Springer Lecture Notes in Mathematics, 1975.

[3] Jost, J.: *Riemannian Geometry and Geometric Analysis*, 3. edition, Springer Universitext, 2002.

[4] Spivak, M.: *A Comprehensive Introduction to Differential Geometry I-V*, 3. edition, Publish of Perish, 1999.

[5] do Carmo, M.: *Riemannian Geometry*, Birkhäuser 1992.

[6] Warner, F.: *Foundations of Differentiable Manifolds and Lie Groups*, Springer Graduate Text, 1996.

[7] Hirzebruch, F., Scharlau, W., *Einführung in die Funktionalanalysis*, B.I. Hochschultaschenbuch, 1971.

[8] Kuwert, E. *Elementare Differentialgeometrie*, Vorlesungsskript, Freiburg 2021.