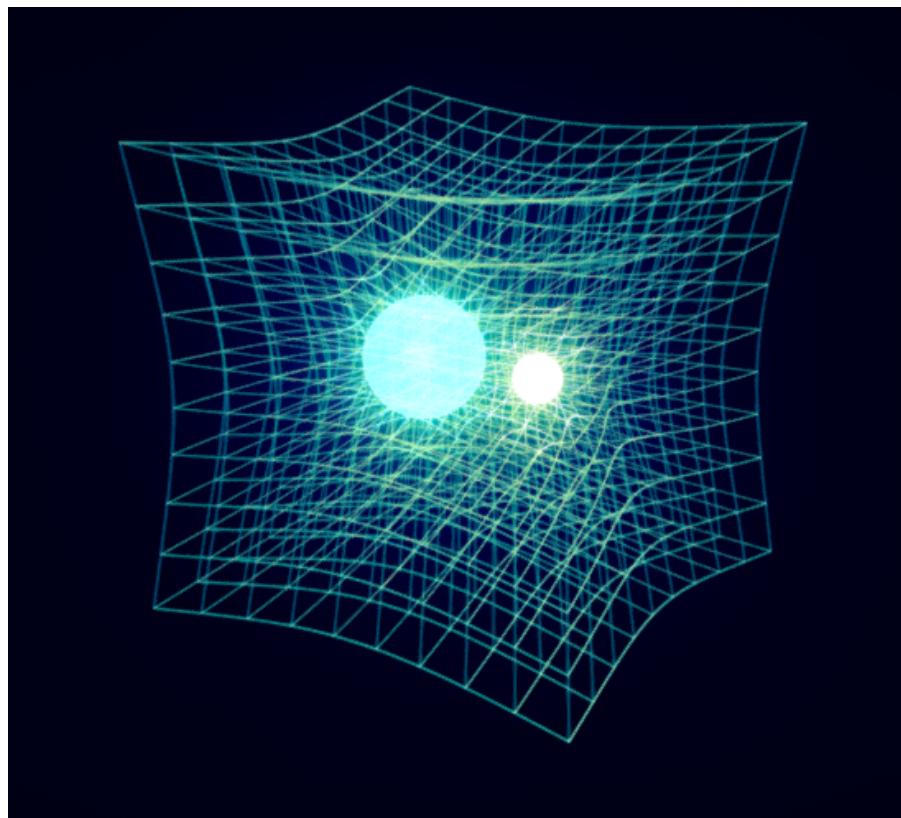


EINFÜHRUNG IN DIE ALLGEMEINE RELATIVITÄTSTHEORIE

VO 136.026, Priv. Doz. Dr. Herbert Balasin, WS 2025/26

Mitschrift von Ciel Frey



Inhaltsverzeichnis

| | |
|---|-----------|
| 1 Einführung | 1 |
| 1.1 Newtonsche Mechanik und Gravitationsphysik | 1 |
| 1.1.1 Inertialsystem, Galilei-Transformation und Galileisches Relativitätsprinzip (nicht Teil der VO) | 2 |
| 1.1.2 spezielle Relativitätstheorie im Überblick (nicht Teil der VO) | 3 |
| 1.2 Geraden in gekrümmten Räumen | 8 |
| 1.3 Newton und Einstein im Vergleich | 10 |
| 1.3.1 Übergang zu allgemeiner Relativitätstheorie (nicht Teil der VO) | 11 |
| 2 Minkowski-Geometrie und SRT | 14 |
| 2.1 Der Minkowski-Raum | 14 |
| 2.1.1 Weltlinien und Gleichzeitigkeit | 15 |
| 2.2 Lorentz-Transformation | 17 |
| 2.2.1 Gleichzeitigkeit, Zeitdilatation und Längenkontraktion | 17 |
| 2.2.2 Einheiten und einsteinsche Summenkonvention | 19 |
| 2.2.3 Herleitung generelle Lorentz-Transformation | 20 |
| 2.3 Minkowski-Metrik und Raumzeitabstand | 26 |
| 2.3.1 Mathematische Struktur der Minkowski-Metrik (nicht Teil der VO) | 29 |
| 3 Multilineare Algebra | 33 |
| 3.1 Vektorräume | 33 |
| 3.1.1 Linearkombination und lineare Unabhängigkeit | 35 |
| 3.1.2 Basis eines Vektorraums | 36 |
| 3.2 Dualraum | 41 |
| 3.2.1 Basis des Dualraums | 43 |
| 3.3 Bilineare Abbildungen | 44 |
| 3.4 Multilineare Abbildungen | 48 |
| 3.5 Doppeldualraum | 51 |

| | | |
|----------|--|------------|
| 3.6 | Gemischte multilineare Abbildungen | 53 |
| 3.7 | Basis- und Koordinatentransformationen | 55 |
| 3.8 | Abstrakte Indexschreibweise | 58 |
| 4 | Topologie | 64 |
| 4.1 | Der Begriff der Topologie | 64 |
| 4.2 | Konvergenz einer Folge | 66 |
| 4.3 | Stetigkeit | 68 |
| 4.4 | Topologisieren von Räumen | 70 |
| 4.5 | Spurtopologie (Teilmengentopologie) | 72 |
| 4.6 | Homöomorphismen | 73 |
| 4.7 | Trennungsaxiome | 75 |
| 4.8 | Die euklidische Topologie | 79 |
| 4.9 | Kompaktheit und Parakompaktheit | 83 |
| 5 | Mannigfaltigkeiten | 85 |
| 5.1 | Der Begriff der Mannigfaltigkeit | 85 |
| 5.2 | Kartenwechsel | 87 |
| 5.3 | Skalarfelder auf Mannigfaltigkeiten | 90 |
| 5.3.1 | Skalarfelder und Karten am Beispiel der stereographischen Projektion | 93 |
| 5.4 | Vektoren auf Mannigfaltigkeiten (Tangentialraum) | 97 |
| 5.4.1 | Dualraum des Tangentialraums (Kotangentialraum) | 103 |
| 6 | Differentialgeometrie und Einstingleichungen | 108 |
| 6.1 | Kovariante Ableitung | 108 |
| 6.1.1 | Koordinatenkovariante Ableitung auf Skalarfeldern | 108 |
| 6.1.2 | Koordinatenkovariante Ableitung auf Vektorfeldern | 109 |
| 6.1.3 | Von der koordinatenkovarianten Ableitung zur kovarianten Ableitung | 114 |
| 6.2 | Derivationen und Kommutator | 118 |
| 6.3 | Riemann tensor | 120 |
| 6.4 | Parallelverschiebung | 122 |
| 6.4.1 | Situation im \mathbb{R}^n | 122 |
| 6.4.2 | Verallgemeinerung auf Mannigfaltigkeiten | 123 |
| 6.4.3 | Konstruktion ∇ -paralleles Vektorfeld | 125 |
| 6.4.4 | Autoparallele Kurven und Geodätengleichung | 127 |
| 6.5 | Metrik und Levi-Civita-Ableitung | 127 |

| | |
|---|-----|
| 6.5.1 Eindeutigkeit der Levi-Civita Ableitung | 129 |
| 6.6 Newtonsche Deviationsgleichung | 130 |
| 6.7 Geodätische Deviationsgleichung | 132 |
| 6.8 Einsteinsche Feldgleichungen | 135 |

Vorwort

Diese Mitschrift ergänzt die bereits vorhandenen Mitschriften von Robert Berndl-Forstner (WS20/21) und Florian Lindenbauer (WS18/19), die hier zu finden sind:

<https://forum.fstph.at/t/einf-i-d-allgemeine-relativitaetstheorie-mitschrift-ws2018/3734>

<https://forum.fstph.at/t/latex-skriptum-allg-relativitaetstheorie/4021>

Zuerst gibt es eine kurze Einführung, wie es zur speziellen und allgemeinen Relativitätstheorie (kurz SRT und ART) kam. Anschließend wird SRT mit Raumzeit-Diagrammen und dem Minkowski-Raum beschrieben. In den darauffolgenden Kapiteln werden die mathematischen Tools vorgestellt, die für das Verständnis von ART benötigt werden. Zum Schluss folgt die Einsteingleichung. Einen tieferen Einblick in ART gewährt die im Sommersemester anschließende VO 136.002 Geometrie und Gravitation I.

An einigen Stellen habe ich zusätzliche Passagen eingefügt, die zwar nicht Teil der Vorlesung waren, mir aber beim Verständnis geholfen haben. Diese sind entsprechend mit „*nicht Teil der VO*“ oder durch einen grauen Strich am linken Rand markiert.

Das Bild des Titelblatts stellt die Raumzeitkrümmung in 3D dar.

Quelle: <https://sketchfab.com/3d-models/gravity-warping-space-time-749a8acac04049a89dc8d02054a02ee8>

Kapitel 1

Einführung

1.1 Newtonsche Mechanik und Gravitationsphysik

Die Gravitation zählt seit Jahrhunderten zu den zentralen Themen der Physik. Schon Isaac Newton formulierte nach der bekannten Anekdote mit dem fallenden Apfel das *Gravitationsgesetz*.

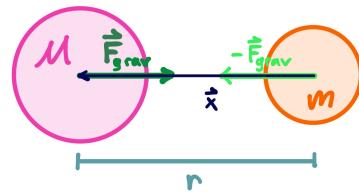
Der Betrag der Gravitationskraft/Schwerkraft lautet

$$F_{\text{grav}} = G \frac{Mm}{r^2}$$

wobei M und m die Massen zweier Körper, $r = |\vec{x}|$ ihr Abstand, \vec{x} der Ortsvektor von m relativ zu M und $G = 6,6743 \cdot 10^{-11} \frac{\text{m}^3}{\text{kg} \cdot \text{s}^2}$ die Gravitationskonstante.

In Vektorform lautet die Gravitationskraft

$$\vec{F}_{\text{grav}} = -G \frac{Mm}{r^3} \vec{x}$$



In Vektorkomponenten i geschrieben

$$F_{\text{grav}}^i = -G \frac{Mm}{r^3} x^i$$

Dieses Gesetz beschreibt bei Geschwindigkeiten, die wesentlich kleiner als die Lichtgeschwindigkeit sind und solange man sich nicht in der Nähe großer Massen aufhält, eine sehr gute Näherung der auftretenden Gravitationswechselwirkungen.

Das Gravitationsgesetz kann auch mithilfe des Gravitationspotentials (Skalarfeld) $\Phi(r) = -\frac{GM}{r}$ formuliert werden:

$$F_{\text{grav}}^i = -m \partial_i \Phi(x^k(t))$$

wobei $\partial_i := \frac{\partial}{\partial x^i}$ die partielle Ableitung nach der i -ten Raumkoordinate bezeichnet und somit $\partial_i \Phi = (\nabla \Phi)^i$ die Komponenten des Gradienten des Gravitationspotentials.

Zugleich verband Newton das Konzept der *Kraft* mit dem der *Beschleunigung* und formulierte damit die Grundlage der heute sogenannten *Newton'schen Mechanik*. Nach dem zweiten Newtonschen Axiom gilt

$$m \ddot{x}^i(t) = F^i(x^k(t))$$

Die an der jeweiligen Position des Körpers wirkende Kraft ist also Masse mal Beschleunigung.

Die Lösung dieser Gleichung beschreibt die Bahn $x^i(t, m)$ eines (punktformig gedachten) Körpers in Abhängigkeit von der Zeit und der Masse des Körpers. Wendet man das zweite Newtonsche Axiom auf die Gravitationskraft an, erhält man:

$$\mu \ddot{x}^i(t) = F_{\text{grav}}^i = -\mu \partial_i \Phi(x^k(t))$$

Die Masse m kürzt sich heraus, sodass sich die Bewegungsgleichung auch in der Form

$$\ddot{x}^i(t) = -\partial_i \Phi(x^k(t))$$

schreiben lässt. Die Bahn des Teilchens ist somit unabhängig von seiner Masse, also $x^i = x^i(t)$.

Diese Unabhängigkeit wird als das ***Klassische Äquivalenzprinzip*** oder auch ***schwache Äquivalenzprinzip*** bezeichnet, und geht auf Überlegungen von Galileo Galilei und seine ***Versuche mit Kugeln auf schießen Bahnen*** zurück

Alle Körper fallen im selben Gravitationsfeld im Vakuum gleich schnell, unabhängig von ihrer Masse. Die Beschleunigung hängt allein vom Gravitationsfeld ab.

Bei der Mondlandung 1971 wurde dies im Vakuum bestätigt: Hammer und Feder trafen gleichzeitig auf der Mondoberfläche auf: <https://www.youtube.com/watch?v=0o8TaPVsn9Y>

Wirken keine Kräfte ($F^i = 0$), so folgt aus dem zweiten Newtonschen Axiom mit zweimaliger Integration:

$$m \ddot{x}^i(t) = 0 \quad \Rightarrow \quad x^i(t) = x_0^i + \dot{x}_0^i t$$

Dies entspricht dem *ersten Newtonschen Axiom*: Ein kräftefreier Körper bleibt in Ruhe oder bewegt sich geradlinig mit konstanter Geschwindigkeit. Also ist die Bahn eine Gerade mit gleichförmiger Bewegung.

1.1.1 Inertialsystem, Galilei-Transformation und Galileisches Relativitätsprinzip (nicht Teil der VO)

Bewegungen können immer nur relativ zu etwas beschrieben werden: von etwas weg, zu etwas hin, ruhend relativ zu etwas. Ein **Bezugssystem** (oder Koordinatensystem) hat einen Bezugspunkt (Koordinatenursprung), bezüglich dessen die Bewegungen beobachtet werden. Es gibt ruhende und bewegte Bezugssysteme.

Definition 1.1. Ein **Inertialsystem** hat die folgenden äquivalenten Definitionen:

- ein Bezugssystem, in dem das erste Newtonsche Gesetz gilt: jeder kräftefreie Körper innerhalb dieses Bezugssystems befindet sich entweder in Ruhe oder bewegt sich mit konstanter Geschwindigkeit auf einer Geraden.
- ein nicht beschleunigtes Bezugssystem.

In einem Inertialsystem wirken also keine Scheinkräfte/Trägheitskräfte. Systeme, die sich relativ zu einem Inertialsystem beschleunigt oder rotierend bewegen, sind selbst keine Inertialsysteme.

Beispiel 1.2. Ein Zug bewegt sich mit konstanter Geschwindigkeit auf einer Strecke mit geraden Schienen. Ein Bezugssystem, das sich mit dem Zug bewegt, ist ein Inertialsystem. Sobald der Zug um eine Kurve fährt, müssen sich die äußeren Teile des Zugs beschleunigen, um dieselbe Strecke wie die inneren Teile zurückzulegen. Das mitgeführte Bezugssystem ist kein Inertialsystem mehr. Ebenso wenn der Zug auf der geraden Strecke abbremst oder beschleunigt. Durch die krummlinige Bewegung oder die Beschleunigung/das Abbremsen treten Trägheitskräfte auf.

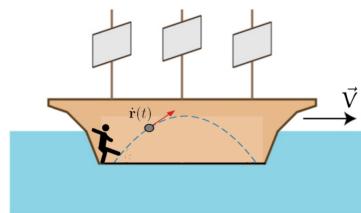
Satz 1.3 (Galileisches Relativitätsprinzip). Alle physikalischen Gesetze sind in allen Inertialsystemen gleich. Oder auch: Bewegt sich ein Inertialsystem I' mit konstanter Geschwindigkeit relativ zu einem anderen Inertialsystem I , so gelten in beiden dieselben physikalischen Gesetze.

Dieses Prinzip bedeutet, dass keine mechanischen Experimente durchgeführt werden können, die erkennen lassen, ob man sich in Ruhe oder in gleichförmiger Bewegung befindet.

Definition 1.4. Die **Galilei-Transformation** beschreibt den Zusammenhang zwischen den Raum- und Zeitkoordinaten zweier Inertialsysteme I und I' , die sich mit konstanter Geschwindigkeit \vec{v} zueinander bewegen, wobei \vec{r} die Position des Körpers und t die Zeit beschreibt. Sie lautet:

$$\begin{aligned}\vec{r}' &= \vec{r} - \vec{v}t && \text{Translation der Trajektorie des Körpers} \\ \vec{v}' &= \dot{\vec{r}}' = \dot{\vec{r}} - \vec{v} && \text{Geschwindigkeiten werden addiert} \\ t' &= t && \text{Zeit bleibt gleich / ist absolut}\end{aligned}$$

Beispiel 1.5. Ein Schiff bewegt sich geradlinig und gleichförmig mit Geschwindigkeit \vec{v} . Ein Fußballspieler unter Deck kann durch ein Experiment - etwa indem er einen Ball hochwirft - nicht feststellen, ob sich das Schiff bewegt oder ruht. Das Schiff bildet also ein Inertialsystem, und die Flugbahn des Balls kann mit der Galilei-Transformation in das ruhende System an Land überführt werden.



1.1.2 spezielle Relativitätstheorie im Überblick (nicht Teil der VO)

gutes Video: 4D Spacetime and Relativity explained simply and visually: <https://youtu.be/ZfR1Jc6Zglo>

Ausgangspunkt: Newtonsche Mechanik

In der newtonschen Mechanik wird von einem **absoluten Raum** und einer **absoluten Zeit** ausgegangen, die **unabhängig** voneinander existieren. Alle Beobachter messen dieselbe Zeit und dieselben Längen, und alle Bewegungen werden relativ zu einem Inertialsystem beschrieben.

Transformationen zwischen Inertialsystemen sind Galilei-Transformationen: lineare Transformationen, bei denen die Zeit unverändert bleibt und sich Geschwindigkeiten addieren.

Beispiel 1.6. Ein Zug fährt geradlinig mit konstanter Geschwindigkeit $v_1 = 100 \text{ km/h}$ (Inertialsystem). Eine

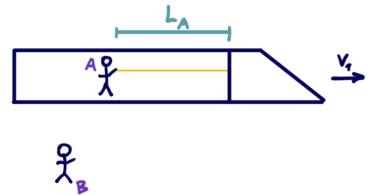
Person im Zug wirft einen Ball mit $v_2 = 10 \text{ km/h}$ nach vorne. Für einen stillstehenden Beobachter außerhalb des Zuges (Inertialsystem), bewegt sich der Ball dann mit der Geschwindigkeit $v = v_1 + v_2 = 110 \text{ km/h}$.

Konflikt: Lichtgeschwindigkeit

Aus den Maxwell-Gleichungen (Gesetze des Elektromagnetismus) folgt, dass die Lichtgeschwindigkeit $c = 299\,792 \text{ km/s}$ konstant ist. Experimente zeigen, dass diese Geschwindigkeit in allen Inertialsystemen denselben Wert hat (invariant ist). Damit widerspricht die Lichtgeschwindigkeit jedoch der Galilei-Transformation, da sie sich nicht addieren lässt.

Beispiel 1.7. Ein Zug bewegt sich geradlinig mit konstanter Geschwindigkeit $v_1 = \frac{c}{2}$. Eine Person A im Zug leuchtet mit einer Taschenlampe nach vorne auf eine Wand in der Entfernung L_A . Das Licht legt diese Strecke in der Zeit $t_A = L_A/c$ zurück.

Da $c = \text{konst.}$ bewegt sich das Licht auch für einen stillstehenden Beobachter B außerhalb des Zuges mit der Geschwindigkeit c und nicht mit $v_2 = c + v_1$. Das heißt Beobachter B muss die Zeit t_A und/oder die Länge L_A anders wahrnehmen. Tatsächlich tritt beides ein: Zeitdilatation und Längenkontraktion.



Für den außenstehenden Beobachter B trifft der Lichtstrahl später auf die Wand, weil sich die Wand weg bewegt. D.h. $t_B > t_A$ (Zeitdilatation). Außerdem nimmt Beobachter B den Zug kürzer wahr. D.h. $L_b < L_A$ (Längenkontraktion). Beides wird im Folgenden noch mehr erklärt.

Gedankenexperiment: Die Lichtuhr (Zeitdilatation)

gute Videos:

Time Dilation - Einstein's Special Relativity: <https://youtu.be/P1MG61R17Ks>

I never understood why you can't go faster than light - until now!: <https://youtu.be/Vitf8YaVXhc>

Special Relativity Part 2: Time Dilation and the Twin Paradox: <https://youtu.be/iIEeSiT3SI4>

Eine Lichtuhr besteht aus zwei Spiegeln, zwischen denen ein Photon hin und her springt. Jedes Mal, wenn das Photon einen Spiegel erreicht, tickt die Uhr um eine Einheit weiter. *Fun fact: Für eine Sekunde müssten dafür die Spiegel $L = c \cdot 1 \text{ s} = 299\,792 \text{ km}$ voneinander entfernt sein.*

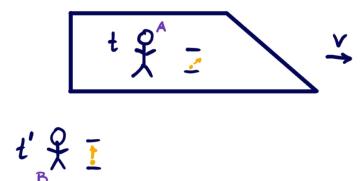


Person A (im mit $v = \frac{c}{2}$ fahrenden Zug; Inertialsystem I') und Person B (außerhalb des Zugs; Inertialsystem I) besitzen jeweils eine solche Lichtuhr. Im fahrenden Zug bewegt sich das Photon diagonal, da sich die Spiegel weiter bewegen. Also legt es, aus der Sicht von B, eine längere Strecke zurück, sodass die Uhr im Zug, aus Sicht von B, langsamer läuft.



Es gilt: $L = c \cdot t$ (Weg = Geschwindigkeit · Zeit)

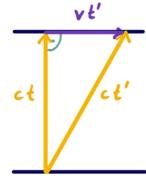
Sei t die Zeit, die aus der Sicht von B im Zug vergeht, und t' die Zeit, die B an seiner Lichtuhr abliest.



Dann ist mit dem Satz des Pythagoras

$$c^2 t^2 + v^2 t'^2 = c^2 t'^2$$

$$\Leftrightarrow t' = \sqrt{\frac{c^2 t^2}{c^2 - v^2}} = \sqrt{\frac{c^2 t^2}{c^2(1 - \frac{v^2}{c^2})}} = \frac{t}{\sqrt{1 - \frac{v^2}{c^2}}}$$



Dies ist die **Zeitdilatationsgleichung**.

Hierbei gilt:

- t ... Eigenzeit / proper time (Zeit von bewegtem System, aus der Sicht des ruhenden Beobachters; ruhender Beobachter sieht wie die Person *im Zug* ihre *eigene Zeit* hat)
- t' ... dilated time (Zeit, die vom ruhenden Beobachter für sich selbst gemessen wird)
- v ... Geschwindigkeit des bewegten Objekts relativ zum ruhenden Beobachter

Es gilt $\lim_{v \rightarrow c} t' = \infty$. Also braucht alles außer Licht unendlich viel Zeit, um Lichtgeschwindigkeit zu erreichen. Daher kann nichts, außer Licht, mit Lichtgeschwindigkeit unterwegs sein.

Der Faktor

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

heißt **Lorentz-Faktor**.

Damit gilt

$$t' = \gamma \cdot t$$

Da $\frac{v^2}{c^2} < 1$ für alles außer Licht, ist $\sqrt{1 - \frac{v^2}{c^2}} \in (0, 1)$ (Bruch) und somit $\gamma > 1$.

Aus Sicht der ruhenden Person *B* außerhalb des Zugs, die ihre Zeit mit t' misst, vergeht die Zeit der Person *A* im Zug mit $t = \frac{t'}{\gamma}$. Also ist die Zeit im Zug, aus Sicht von außerhalb, um den Faktor γ langsamer. (z.B. $t' = 2\text{ s}$ vergeht außen, dann vergeht, für $\gamma = 2$, die Zeit $t = 1\text{ s}$ innen)

| $\frac{v}{c}$ | γ |
|---------------|----------|
| 1% | 1,0005 |
| 10% | 1,005 |
| 50% | 1,2 |
| 87% | 2 |
| 99% | 7 |
| 99,9% | 22 |
| 99,99% | 70 |

Tabelle 1.1: Lorentz-Faktor zu Geschwindigkeit gegeben in Prozent der Lichtgeschwindigkeit.

$$1\% = 1\,079\,252\,848,8 \text{ km/h}$$

Für $v \ll c$ ist $\gamma \approx 1$, und die Zeitdilatation ist praktisch nicht bemerkbar - die Newtonsche Mechanik bleibt gültig.

Bemerkung 1.8. Zeitdilatation ist symmetrisch. Aus Sicht der Person im Zug bewegt sich der außenstehende Beobachter mit der Geschwindigkeit v in die entgegengesetzte Richtung. Damit gilt die Zeitdilatation auch hier, wobei nun t' die Zeit im Zug und $t = \frac{t'}{\gamma}$ die Zeit außerhalb ist. Also ist die Zeit außerhalb, aus der Perspektive von innen, um den Faktor γ langsamer.

Die Zeitdilatation beeinflusst tatsächlich den Alltag. Satelliten und GPS-Systeme müssen aufgrund ihrer hohen Geschwindigkeit und der allgemeinen Relativität (Gravitationseinfluss) regelmäßig korrigiert werden.

Längenkontraktion

gute Videos:

Time Dilation and Length Contraction | Special Relativity: <https://youtu.be/bArTzG3Mkml>

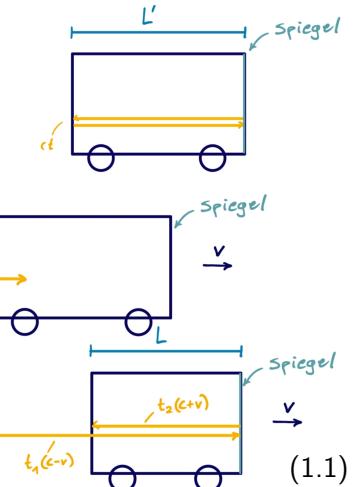
Special Relativity Part 3: Length Contraction: <https://youtu.be/FPzGAksFCbs>

In einem ruhenden Zug befindet sich an der Vorderseite ein Spiegel. Ein Photon wird von einer Quelle an der Rückseite auf den Spiegel geschossen und erreicht die Rückseite wieder nach Zeit t . Also ist die Länge des Zuges $L' = \frac{ct}{2}$ und die Zeit $t = 2\frac{L'}{c}$.

Bewegt sich der Zug nun mit Geschwindigkeit v , so braucht das Photon länger, um den vorderen Spiegel zu erreichen, da sich dieser weg bewegt, und weniger lang die Rückseite zu erreichen, da sich diese auf das Photon zubewegt. Für einen außenstehenden Beobachter braucht das Photon nun Zeit t für die Distanz L , wobei $t' = \gamma \cdot t$.

Setze t in $t = 2\frac{L'}{c}$ ein:

$$\frac{t'}{\gamma} = 2\frac{L'}{c}$$



Andererseits ist die Länge von der Rückseite zur Vorderseite $L = t_1(c - v)$ bzw. $t_1 = \frac{L}{c-v}$ und von der Vorderseite wieder zur Rückseite $L = t_2(c + v)$ bzw. $t_2 = \frac{L}{c+v}$.

Dann ist

$$t' = t_1 + t_2 = \frac{L}{c-v} + \frac{L}{c+v} = \frac{(c+v)L + (c-v)L}{(c+v)(c-v)} = L \frac{c+v+c-v}{c^2-v^2} = \frac{2Lc}{c^2-v^2} = \frac{2Lc}{c^2(1-\frac{v^2}{c^2})} = \frac{2L}{c}\gamma^2$$

Setze in (1.1) ein:

$$\begin{aligned} \frac{2L}{c}\gamma^2 \cdot \frac{1}{\gamma} &= 2\frac{L'}{c} \\ \Leftrightarrow L\gamma &= L' \end{aligned}$$

$$L = \frac{L'}{\gamma}$$

Dies ist die **Längenkontraktionsgleichung**.

Dabei ist:

- L' ... Eigenlänge / proper length (Länge des ruhenden Objekts; oft auch mit L_0 bezeichnet)
- L ... kontrahierte Länge (Länge des bewegten Objekts)

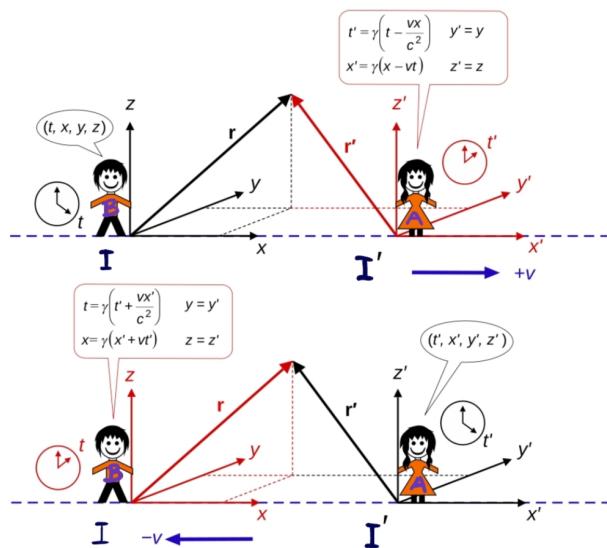
Das bedeutet: Ein bewegtes Objekt erscheint in Bewegungsrichtung kürzer, und zwar um den Faktor γ . Genauer: wenn der Zug eine Initiallänge L' im ruhenden Zustand hatte, hat er mit der Geschwindigkeit v die Länge $L = \frac{L'}{\gamma}$, die wegen $\gamma > 1$ kürzer ist als L' .

Lorentz-Transformation

Ist der Beobachter A in Inertialsystem I' mit konstanter Geschwindigkeit v in x -Richtung gegenüber einem anderen Beobachter B in Inertialsystem I bewegt, so hängen die Koordinaten (t', x', y', z') , die Beobachter A einem Ereignis zuschreibt, durch die *Lorentz-Transformation* mit den Koordinaten (t, x, y, z) des Beobachters B für dasselbe Ereignis zusammen, falls die beiden Bezugssysteme I und I' denselben Ursprung haben, das heißt, zum Zeitpunkt $t = t' = 0$ miteinander übereinstimmen.

$$\begin{aligned} t' &= \gamma \left(t - \frac{v}{c^2} x \right) \\ x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \end{aligned}$$

wobei $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ der Lorentz-Faktor ist.



Darstellung der Lorentz-Transformation: Zwei Beobachter A und B in unterschiedlichen Inertialsystemen I und I' . r ist der Ortsvektor des Ereignisses.

Quelle: https://en.wikipedia.org/wiki/Lorentz_transformation#/media/File:Lorentz_boost_x_direction_standard_configuration.svg

Gleichzeitigkeit

Auch Gleichzeitigkeit ist relativ. Eine Person im fahrenden Zug befindet sich in der Mitte und sendet zwei Lichtblitze aus - einen nach vorne, einen nach hinten. Für sie treffen beide gleichzeitig an den Enden des Wagens ein.



Ein außenstehender Beobachter sieht jedoch, dass der Lichtstrahl zuerst am hinteren Ende auftrifft und danach am vorderen, da sich die hintere Wand auf den Strahl zubewegt und die vordere sich weg bewegt.

Zusammenfassung

Die spezielle Relativitätstheorie beruht auf zwei Grundannahmen:

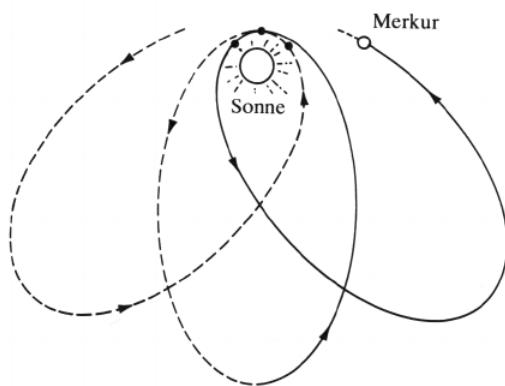
1. Das galileische Relativitätsprinzip: Alle physikalischen Gesetze sind in allen Inertialsystemen gleich.
2. Die Lichtgeschwindigkeit c ist konstant.

Daraus folgt, dass die Galilei-Transformation mit $t = t'$ nicht mehr gültig sein kann. Raum und Zeit sind nicht unabhängig, sondern bilden gemeinsam die *Raumzeit*. Jedes Objekt besitzt seine eigene Zeit (Eigenzeit) und seine eigene Länge.

Anstelle der Galilei-Transformation gilt nun die Lorentz-Transformation.

1.2 Geraden in gekrümmten Räumen

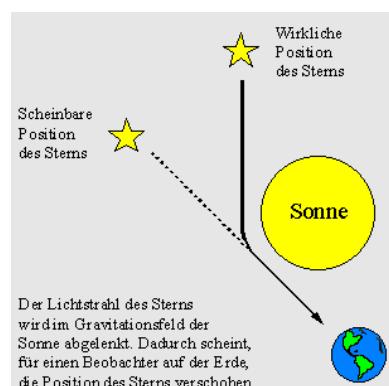
Das Gravitationsgesetz von Newton ist in den meisten Situationen eine sehr gute Näherung. Bei genauen Messungen oder in der Nähe großer Massen treten jedoch Abweichungen auf. Beispiele sind die Lichtableitung oder die Periheldrehung der Planeten. Solche Effekte lassen sich mit der Newtonschen Mechanik nicht erklären.



(a) Die Periheldrehung der Merkurbahn wird auch durch Raumzeitkrümmung verursacht.

Quellen: [Link](#)

I never understood why planets don't follow perfect orbits... until now! <https://youtu.be/S78h8zQwQe0?si=UaW3JNNzFQhHuBl7&t=1124>



(b) Lichtableitung
Quelle: <https://einstein-website.de/albert-einstein-biografie/>

Die allgemeine Relativitätstheorie (ART) ersetzt das Konzept der Gravitationskraft durch die Vorstellung einer gekrümmten Raumzeit. Massen verändern die **Geometrie** der Raumzeit. In dieser gekrümmten Raumzeit bewegen sich Körper auf den kürzesten Wegen zwischen zwei Punkten, also auf „Geraden“, die auf gekrümmten Flächen *Geodäten* genannt werden. Für einen Beobachter, der in der Raumzeit lebt, *scheint* es jedoch so, als würde eine Kraft auftreten (Gravitationskraft).

Beispiel 1.9 (Gedankenexperiment: Aufzug). Wir betrachten eine Person in einem geschlossenen Aufzug. Die Person hat keine Möglichkeit hinauszusehen.

Wenn der Aufzug abstürzt, schwebt die Person im Inneren schwerelos. Gegenstände bleiben an ihrem Ort oder bewegen sich geradlinig mit konstanter Geschwindigkeit weiter, solange keine Kraft auf sie wirkt. Der Zustand des freien Falls ist also lokal nicht von einem kräftefreien Raum zu unterscheiden.

Diese Beobachtung wird als **starkes Äquivalenzprinzip** oder **Einstineisches Äquivalenzprinzip** bezeichnet:

Ein frei fallendes Bezugssystem kann lokal (in einer hinreichend kleinen Umgebung) nicht von einem kräftefreien Bezugssystem unterschieden werden. Frei fallende Bezugssysteme sind lokal äquivalent zu Inertialsystemen. In jedem frei fallenden Bezugssystem gelten lokal dieselben physikalischen Gesetze wie in der Speziellen Relativitätstheorie.

Es bildet das Fundament der Allgemeinen Relativitätstheorie. Weitere [Quelle](#).

Global gilt diese Äquivalenz jedoch nicht. Zoomt man weiter heraus, bewegen sich Körper im freien Fall nicht mehr auf Geraden, sondern auf geradest-möglichen Strecken (Kurven), die durch die Gravitation verursacht werden. Beispiel: ein Stift bewegt sich geradlinig mit konstanter Geschwindigkeit in der Schwerelosigkeit der ISS. Weiter herausgezoomt, fällt die ISS um die Erde herum. Global betrachtet bewegt sich der Stift nicht auf einer Geraden.

Einstein erkannte, dass diese Beobachtung darauf hindeutet, dass die Bewegung von Körpern im Gravitationsfeld nicht durch eine Kraft, sondern durch die Krümmung der Raumzeit bestimmt wird. Körper bewegen sich also entlang von Geraden auf dieser gekrümmten Raumzeit.

Um eine solche Bewegung zu beschreiben, müssen wir klären, was eine *Gerade in einem gekrümmten Raum* bedeutet. Im euklidischen Raum ist sie die kürzeste Verbindung zwischen zwei Punkten. Allgemein gilt:

- Die kürzeste Verbindung zwischen zwei Punkten verläuft entlang einer Geraden.
- Der Tangentialvektor einer Geraden bleibt stets zu sich selbst parallel. Diese Eigenschaft nennt man *Autoparallelität*.

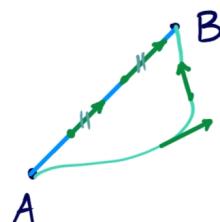


Abbildung 1.2: In blau: eine Gerade im flachen Raum, denn alle Tangentialvektoren sind parallel. In türkis: keine Gerade den mind. zwei Tangentialvektoren sind nicht parallel.

In gekrümmten Räumen wird diese Vorstellung schwieriger. Auf einer Kugeloberfläche liegt die kürzeste Verbindung zwischen zwei Punkten auf einem Großkreis (Kreis mit maximalem Umfang). Von außen betrachtet sind die Tangentialvektoren solcher Großkreise jedoch nicht autoparallel, siehe [Abbildung 1.3](#). Für eine Person, die auf der Kugeloberfläche lebt, erscheint die Umgebung lokal flach. Daher sind die Tangentialvektoren für diese Person autoparallel.

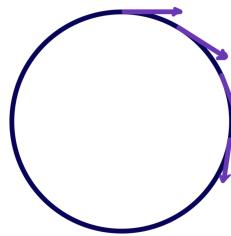


Abbildung 1.3: Tangentialvektoren auf Großkreisen sind von außen betrachtet nicht autoparallel

In einem flachen Raum kehrt ein Vektor, der entlang einer geschlossenen Kurve parallelverschoben wird, wieder zu seiner Ausgangsrichtung zurück. In einem gekrümmten Raum ist das nicht der Fall.

Starten wir mit einem Tangentialvektor auf dem Äquator und verschieben ihn entlang eines Großkreises bis zum Nordpol. Dann folgen wir einem anderen Großkreis zurück zum Äquator und dann entlang des Äquators zurück zum Ausgangspunkt. Obwohl der Vektor auf dem gesamten Weg nur parallel verschoben wurde, kehrt er nicht zur ursprünglichen Richtung zurück, sondern ist gegenüber seiner Anfangsrichtung verdreht. Das zeigt anschaulich, was Raumkrümmung bedeutet.

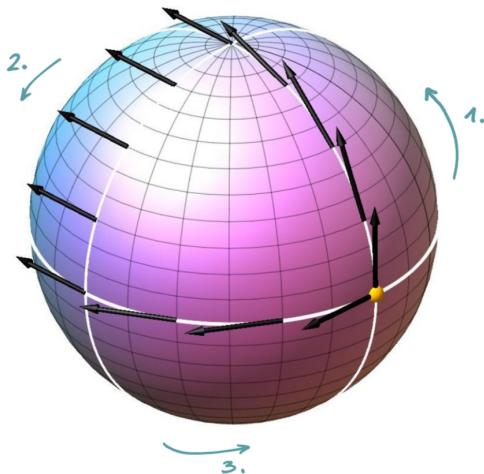


Abbildung 1.4: Parallelverschiebung

Quelle: ART Mitschrift Berndl-Forstner (WS20/21)

<https://forum.fstph.at/t/latex-skriptum-allg-relativitaetstheorie/4021>

1.3 Newton und Einstein im Vergleich

Newton und Einstein erklären Gravitation auf unterschiedliche Weise. Newton beschreibt sie als Kraft, die zwischen Massen wirkt. Einstein beschreibt sie als Folge der Raumzeitkrümmung.

Das lässt sich mit der Kugeloberfläche verdeutlichen. Zwei Körper bewegen sich auf Großkreisen, die durch den Nordpol verlaufen. Beide starten von der Äquatorebene in Richtung Pol. Ein Beobachter auf der Oberfläche weiß nicht, dass er sich in einem gekrümmten Raum befindet. Wenn sich die Körper annähern, deutet er das als gegenseitige Anziehung, also als Kraftwirkung.

Im Einstein'schen Bild bewegen sich beide Körper einfach entlang der kürzesten Wege auf der gekrümmten Oberfläche, also entlang der Großkreise. Die Näherung der beiden ist keine Folge einer Kraft, sondern der Geometrie.

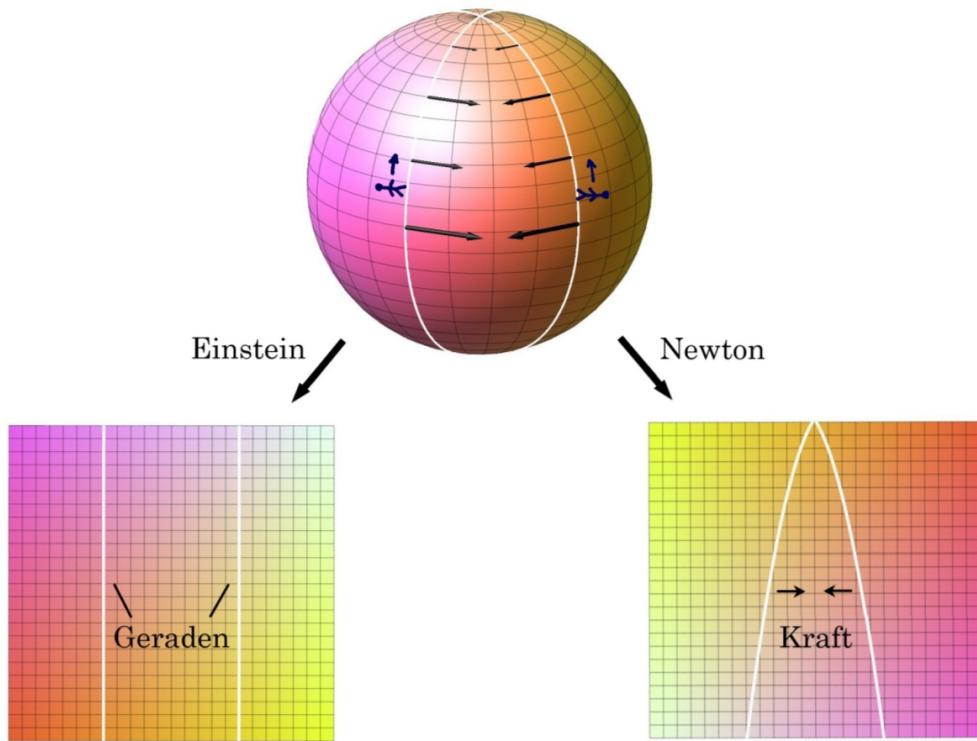


Abbildung 1.5: Vergleich zwischen Newton und Einstein: Während Newton Gravitation als Kraft interpretiert, versteht Einstein sie als Folge der Raumzeitkrümmung. Die Körper bewegen sich entlang der kürzesten Wege auf der gekrümmten Oberfläche.

Quelle: ART Mitschrift Berndl-Forstner (WS20/21)

<https://forum.fstph.at/t/latex-skriptum-allg-relativitaetstheorie/4021>

Die allgemeine Relativitätstheorie beschreibt genau diese Krümmung der Raumzeit. Sie erklärt Gravitation ohne das Konzept einer Kraft. In den folgenden Kapiteln wird der dafür notwendige mathematische Rahmen entwickelt. Am Ende steht die zentrale Gleichung der allgemeinen Relativitätstheorie: die *Einstein-Gleichung*.

1.3.1 Übergang zu allgemeiner Relativitätstheorie (nicht Teil der VO)

gute Videos:

General Relativity Explained simply & visually: <https://youtu.be/tzQC3uYL67U>

I finally rediscovered why matter curves spacetime! <https://youtu.be/qKZvAwcXceE>

General Relativity: The Curvature of Spacetime: <https://youtu.be/R7V3koyL7M>

Die spezielle Relativitätstheorie (SRT) gilt nur für Inertialsysteme, also für Bezugssysteme ohne Beschleunigung und insbesondere ohne Rotation oder Gravitation. Die allgemeine Relativitätstheorie (ART) erweitert diesen Rahmen auf beschleunigte Systeme und formuliert Gravitation im Sinne der Relativitätstheorie.

Konflikt: wieder die Lichtgeschwindigkeit

In Newtons Theorie wirkt Gravitation augenblicklich und über beliebig große Entferungen hinweg. Das widerspricht jedoch der Tatsache, dass nichts (also auch keine Kraftwirkung) schneller als die Lichtgeschwindigkeit sein kann.

Einstiens Gedankenexperiment

Eine Person in einem geschlossenen Raumschiff kann nicht unterscheiden, ob sich das Raumschiff

- auf der Erde befindet und die Person die Gewichtskraft des Gravitationsfeldes $F_G = mg$ erfährt (*schwere Masse*),
- oder ob sich das Raumschiff im Weltraum befindet und mit g nach oben beschleunigt wird (*träge Masse; Widerstand bei Beschleunigung*).

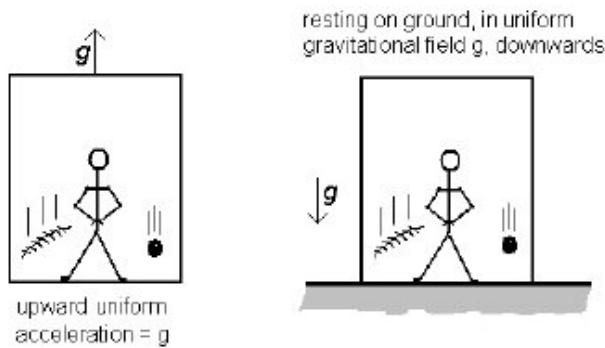


Abbildung 1.6: Gleichwertigkeit von Beschleunigung und Gravitation

Quelle: <https://asymptotia.com/wp-images/2008/02/equivalence1.jpg>

Schwere Masse und träge Masse sind also lokal äquivalent/gleichwertig. Dies ist eine weitere Formulierung des **starken Äquivalenzprinzips**.

Einstein führte diesen Gedanken weiter. Wenn Gravitation und Beschleunigung äquivalent sind, dann kann Gravitation nicht als Kraft verstanden werden. Massen verändern stattdessen die Geometrie der Raumzeit. Körper bewegen sich in dieser gekrümmten Raumzeit entlang des kürzesten Wegs zwischen zwei Punkten (Geodäte).

Gravitation ist somit Ausdruck der Raumzeitkrümmung: Sie beschreibt, wie sich die Geodäten in der Nähe von Massen verändern. Bewegung im Gravitationsfeld folgt also nicht einer Kraft, sondern der Struktur der Raumzeit selbst.

In der speziellen Relativitätstheorie ist die Raumzeit flach, in der allgemeinen Relativitätstheorie dagegen gekrümmmt.

Die allgemeine Relativitätstheorie sagte auch schwarze Löcher vorher und dass Uhren in der Nähe von großen Massen langsamer ticken.

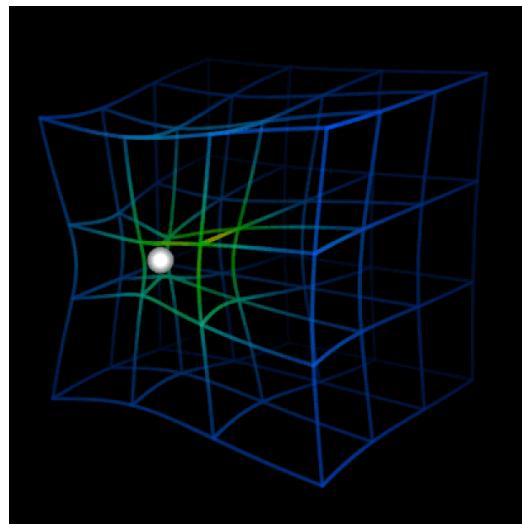


Abbildung 1.7: Krümmung der Raumzeit in 3D

Quelle (bewegtes Bild):

<https://lucasvb.tumblr.com/post/142549026838/figured-id-add-a-mass-in-there-see-also>

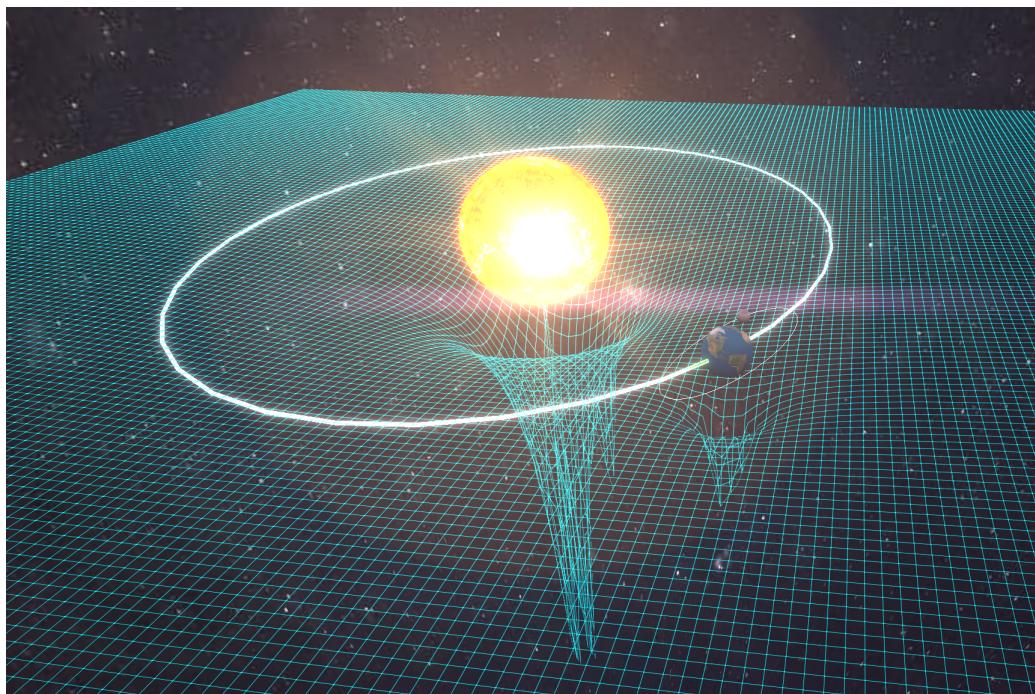


Abbildung 1.8: Krümmung der Raumzeit in 2D

Quelle: <https://hiteshsahu.com/lab/relativity/>

Kapitel 2

Minkowski-Geometrie und SRT

2.1 Der Minkowski-Raum

gute Videos:

4D Spacetime and Relativity explained simply and visually: <https://youtu.be/ZfR1Jc6Zgl>

Weltlinien und Minkowski-Diagramm - Spezielle Relativitätstheorie (5) | Peter Kroll: <https://youtu.be/ef7X9-1qhkA?si=v6eHwvrXmXKqFNKo&t=1339>

Der **Minkowski-Raum** \mathbb{M} ist die mathematische Grundlage der speziellen Relativitätstheorie. Er beschreibt die Raumzeit als einen vierdimensionalen Raum, in dem drei Raumrichtungen und eine Zeitrichtung gemeinsam betrachtet werden.

Die drei räumlichen Koordinaten x^i beschreiben den gewohnten euklidischen Raum. Die vierte Koordinate steht für die Zeit t . Sie wird mit der Lichtgeschwindigkeit c multipliziert, damit auf allen Achsen mit Längen gearbeitet werden kann: $ct = \frac{km}{s} \cdot s = km$.

Damit ist der Minkowski-Raum

$$\mathbb{M} := \{(\underbrace{ct}_{=x^0}, x^1, x^2, x^3) \in \mathbb{R}^4\}$$

Dennoch unterscheidet sich der Minkowski-Raum wesentlich von einem vierdimensionalen euklidischen Raum.

Zur Veranschaulichung verwendet man sogenannte *Minkowski-Diagramme* oder auch *Raumzeit-Diagramme*. Da es unmöglich ist, vier aufeinander normal stehende Achsen graphisch darzustellen, werden für solche Diagramme die Raumdimensionen von drei auf eine oder zwei reduziert, und eine Zeitachse eingeführt, die meistens nach oben zeigt.

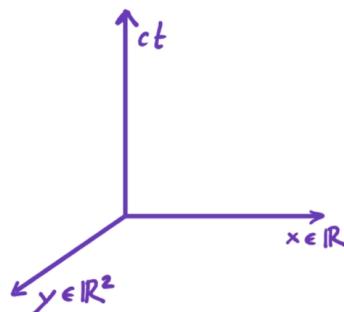


Abbildung 2.1: Minkowski-Raum

2.1.1 Weltlinien und Gleichzeitigkeit

Im Minkowski-Diagramm stellt jeder Punkt ein Teilchen oder auch ein Ereignis dar, also einen Ort zu einem bestimmten Zeitpunkt. Die Bahn eines Teilchens durch die Raumzeit wird durch eine sogenannte *Weltlinie* in diesem Diagramm beschrieben.

Ein Körper, der ruht, wird durch eine vertikale Linie dargestellt, da sich sein Ort in der Zeit nicht ändert. Ein Körper, der sich mit konstanter Geschwindigkeit bewegt, hat eine geneigte Weltlinie. Je größer die Geschwindigkeit, desto flacher verläuft die Linie. Parallelle Weltlinien beschreiben Körper, die sich in relativer Ruhe befinden. Auf der **Gleichzeitebene** erfahren alle Teilchen die gleiche Zeit/alle Ereignisse/Events finden zur gleichen Zeit statt.

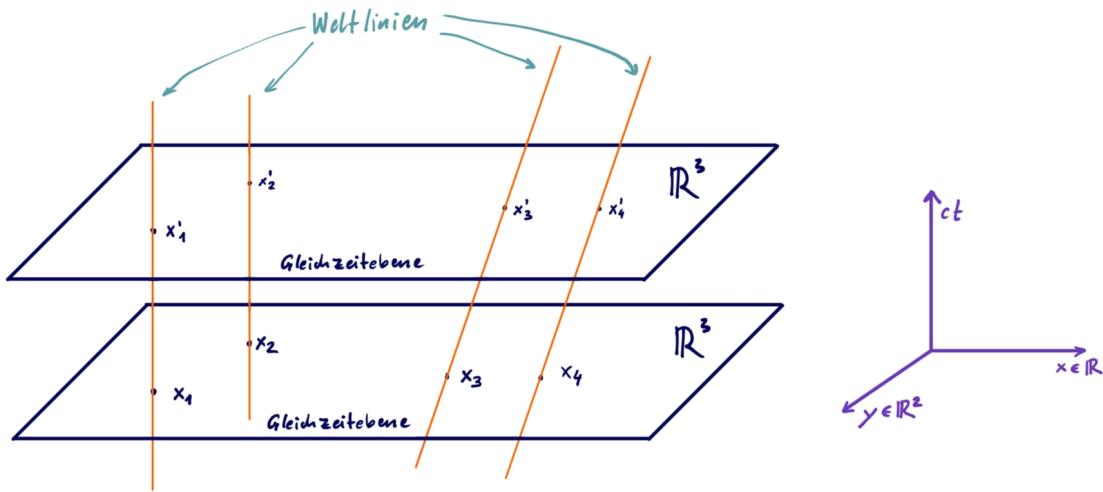


Abbildung 2.2: Weltlinien im Minkowski-Diagramm. x_1 & x_2 sind ruhende Körper, x_3 & x_4 bewegen sich gleichförmig mit derselben Geschwindigkeit.

Wenn sich x_3 nun mit der gleichen Geschwindigkeit v in alle Richtungen bewegt, entsteht ein Kegel, der sogenannte **Gleichgeschwindigkeitskegel**. Der Winkel θ des Gleichgeschwindigkeitskegels ist also durch die Geschwindigkeit bestimmt. Desto größer die Geschwindigkeit des Teilchens, desto größer der Winkel.

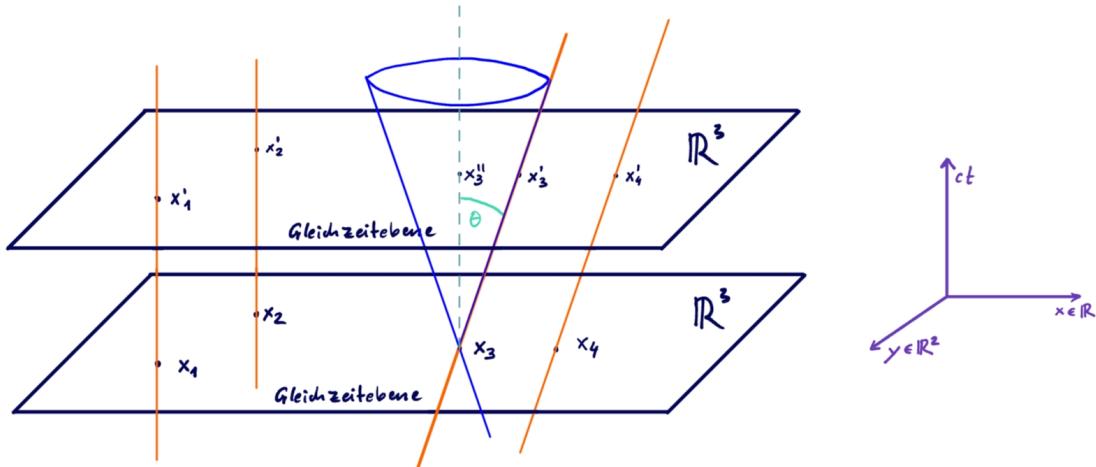


Abbildung 2.3: Gleichgeschwindigkeitskegel von x_3 . Punkt x_3'' liegt auf unbewegten Weltlinie von x_3 .

In der Newton-Mechanik werden Geschwindigkeiten addiert. Also kann eine unendlich große Geschwindigkeit auftreten (deutsche Autobahn). Das bedeutet, dass der Winkel eines Gleichgeschwindigkeitskegels $\theta = 90^\circ$ erreichen kann. Die wesentliche Erkenntnis der speziellen Relativitätstheorie ist, dass es eine maximale Geschwindigkeit gibt: die Lichtgeschwindigkeit c . Diese ist in allen Bezugssystemen gleich – also eine Invariantengeschwindigkeit. Also kann ein Gleichgeschwindigkeitskegel die 90° nur annähern, aber nicht erreichen.

Skaliere so, dass $v = c$ einem Winkel $\theta = 45^\circ$ entspricht. (Man könnte auch anders wählen; die Idee ist, dass der Winkel zwischen 0° und 90° liegen muss, wobei 0° und 90° nicht möglich sind – also nimmt man die Mitte.)

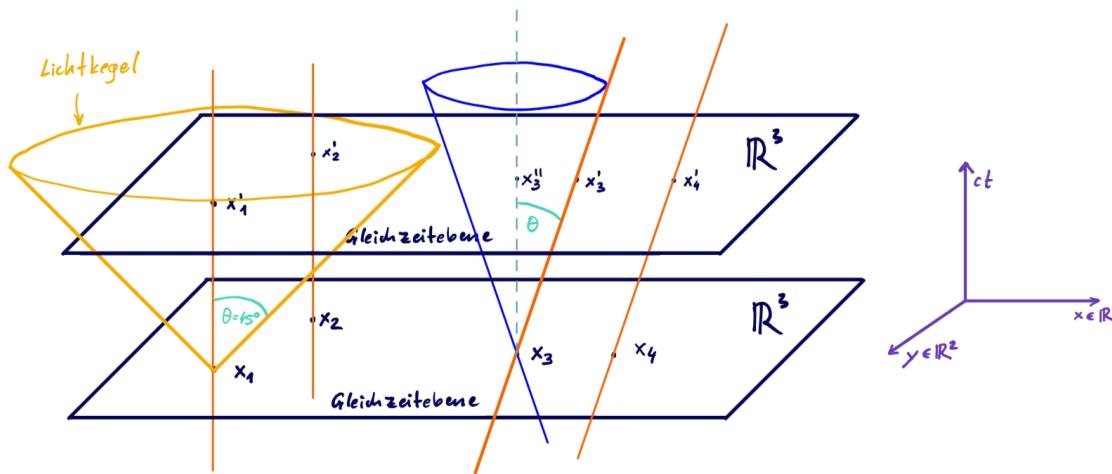


Abbildung 2.4: x_1 sendet Licht aus und erzeugt somit den Lichtkegel.

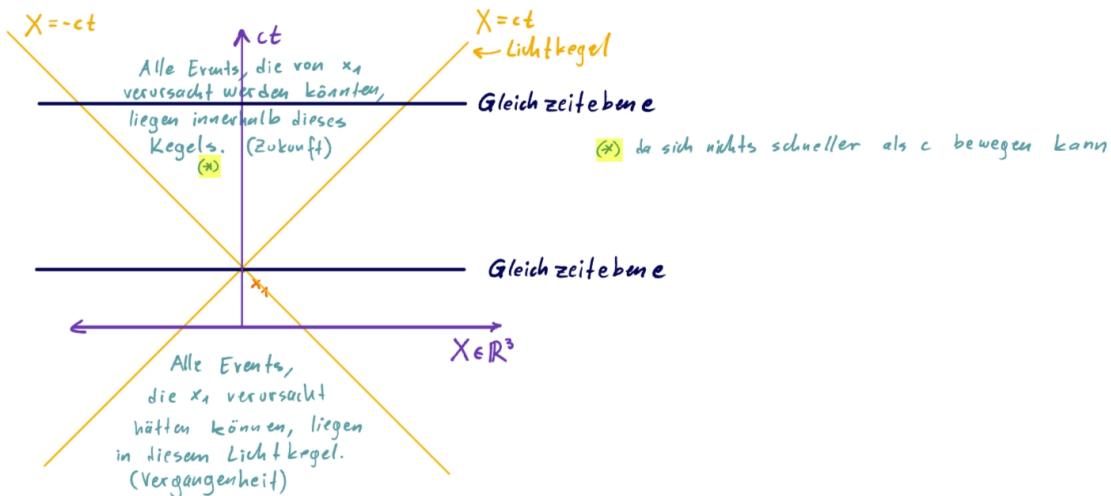


Abbildung 2.5: Zukunft und Vergangenheit.

2.2 Lorentz-Transformation

gute Videos:

Lorentz Transformations | Special Relativity Ch. 3: <https://www.youtube.com/watch?v=RhOpYtQG5wI>

Length Contraction and Time Dilation | Special Relativity Ch. 5: https://www.youtube.com/watch?v=-NN_m2yKAAk

Relativistische Längenkontraktion ▪ Garagenparadoxon ▪ Spezielle Relativitätstheorie(6)| Peter Kroll <https://youtu.be/fHGFZ2GVdkM?si=h1Ngn3ipegW7KE4h&t=770>

2.2.1 Gleichzeitigkeit, Zeitdilatation und Längenkontraktion

Bei Newton sind Zeit und Längen absolut und unabhängig. Wenn sich ein Teilchen bewegt, nimmt es bei Newton dieselbe Zeit und Länge war, wie ein ruhendes Teilchen. Die Gleichzeitebenen verändern sich nicht. Die spezielle Relativitätstheorie zeigt, dass Gleichzeitigkeit, Zeit und Länge vom Bezugssystem/Beobachter abhängen.

Gleichzeitigkeit hängt vom Bezugssystem/Beobachter ab. Sei ein Stab im Koordinatenursprung platziert. Der Stab bewegt sich gleichmäßig mit Geschwindigkeit v nach rechts. Betrachte zwei Bezugssysteme: ein fixes I und ein mitbewegtes I' . Im System I treten die Ereignisse x_1 und x_2 nicht gleichzeitig ein. Das Licht trifft x_1 viel früher. Im System I' treten sie gleichzeitig ein: die Linie x_1x_2 ist eine Gleichzeitigkeitsebene in I' (geht durch die Schnittpunkte der Weltlinien von A und E mit dem Lichtkegel). Alle Linien, die parallel dazu sind, sind auch Gleichzeitigkeitsebenen in I' .

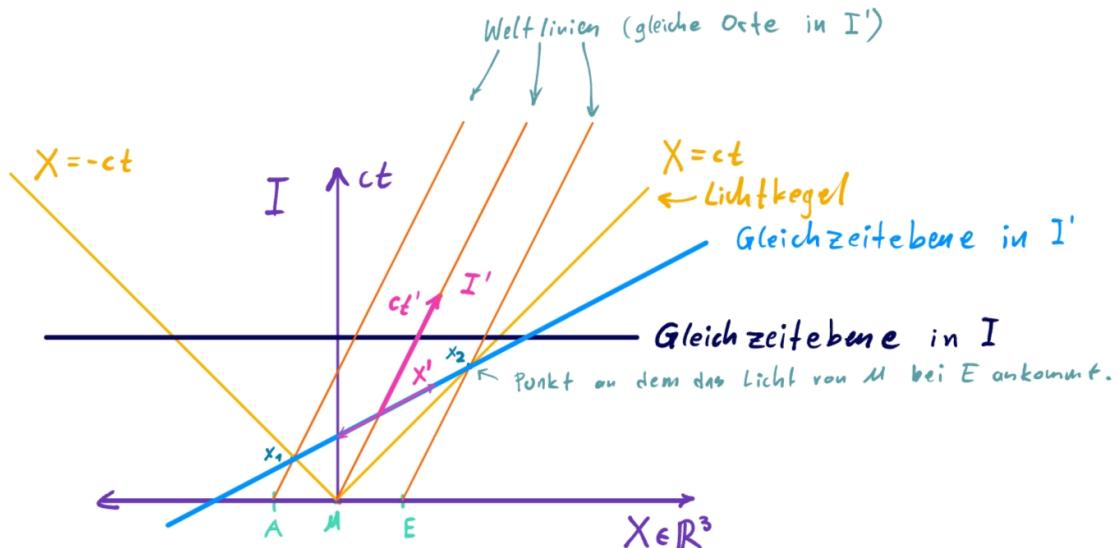
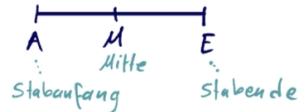


Abbildung 2.6: Die Beziehung zwischen I und I' wird durch eine Lorentz-Transformation beschrieben.

Zweite Möglichkeit zur Konstruktion einer Gleichzeitebene in I' :

Da wir den Lichtkegel mit 45° einzeichnen, ist der Abstand von t (zeitlicher Abstand) vom Ursprung zu einem Ereignis gleich dem Abstand von x (räumlicher Abstand) von t zum Lichtkegel (siehe Abbildung 2.7). Da die Lichtgeschwindigkeit im bewegten System I' ebenso schnell ist, muss auch dort der zeitliche Abstand t' und der räumliche Abstand x' zum Lichtkegel übereinstimmen.

Der Begriff "Abstand" wurde hier noch nicht korrekt eingeführt. Wie wir später sehen werden, können wir genaugenommen nur über ein Abstandsquadrat sprechen.

1. Zeichne bewegte Weltlinie ein.
2. Nehme einen beliebigen Punkt auf der bewegten Weltlinie her und trage den Vektor t' ein.
3. Trage anschließend, ausgehend von der Spitze vom Vektor t' , einen Vektor x' mit derselben Länge von t' so ein, dass dessen Spitze den Lichtkegel trifft.
4. Die Gleichzeitebenen im bewegten System I' sind nun diejenigen, die parallel zu x' sind.

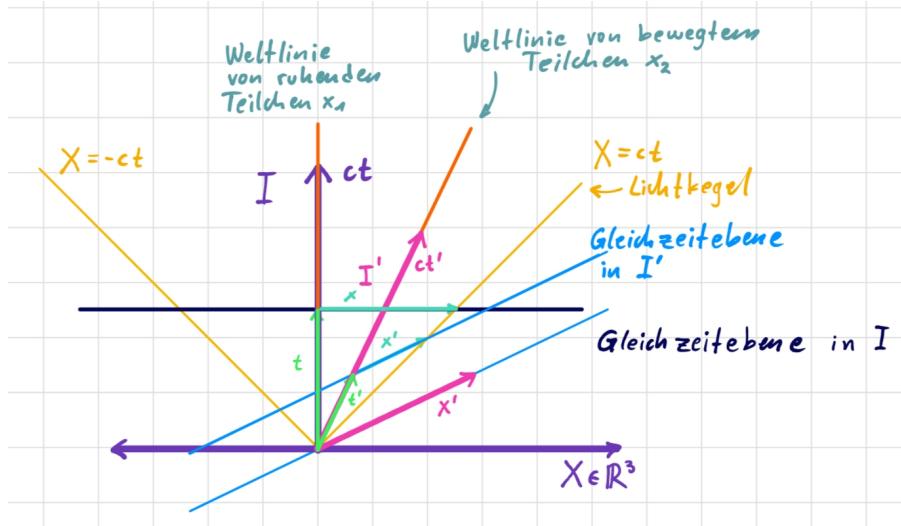


Abbildung 2.7: Konstruktion einer Gleichzeitebene in I' mit nur einer Weltlinie.

Zeit hängt vom Bezugssystem/Beobachter ab.

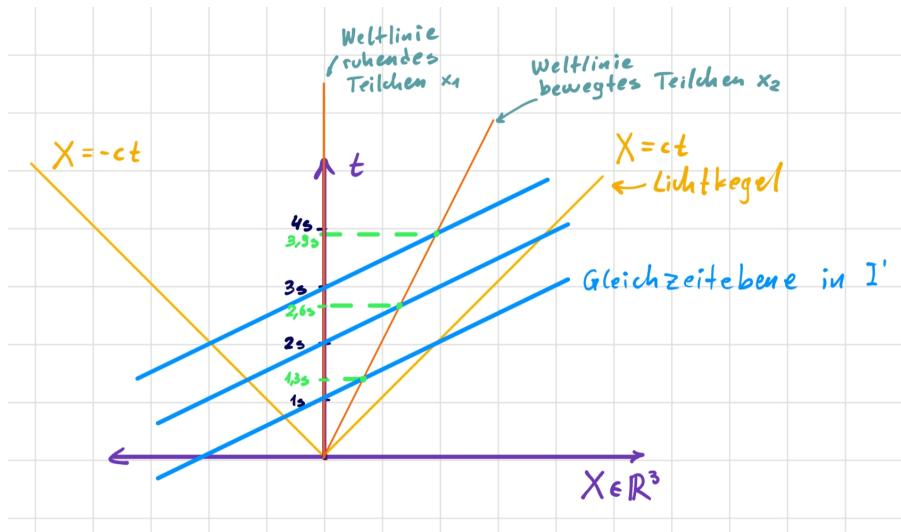


Abbildung 2.8: Zeitdilatation: Aus der Sicht von dem ruhenden x_1 , läuft die Zeit von sich selbst normal, aber die Zeit vom bewegten x_2 um 0,3 s langsamer.

Dies kann auch mit dem Satz des Pythagoras gezeigt werden. In Abbildung 2.9 gilt

$$t^2 < t^2 + x^2 = \bar{t}^2$$

Also läuft die Zeit \bar{t} im bewegten System langsamer als im ruhenden.

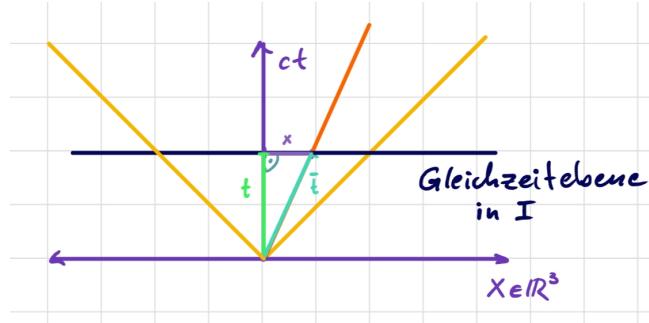


Abbildung 2.9: Zeitdilatation

Auch **Länge** ist beobachterabhängig. Die Länge eines Objekts ist der Abstand des Anfangs- und Endpunktes gleichzeitig gemessen. Die Länge bewegter Objekte wird aus der Sicht des ruhenden Beobachters kürzer.

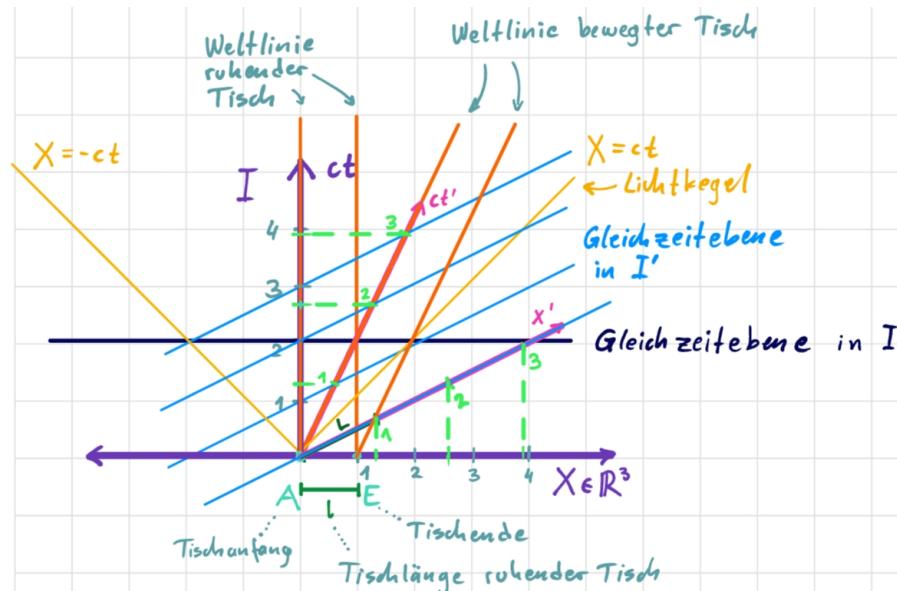


Abbildung 2.10: Längenkontraktion: Ein Tisch bewegt sich mit Geschwindigkeit v . Die ruhende Person nimmt die Länge l für den bewegten Tisch wahr, der im ruhenden Zustand die Länge L hatte. Da $l < L$, nimmt die ruhende Person den Tisch kürzer wahr. (Konstruktion der Gleichzeitebenen von I' gleich wie bei Abbildung 2.7)

2.2.2 Einheiten und einsteinsche Summenkonvention

Um die Notation zu vereinfachen, wird die Lichtgeschwindigkeit $c = 1$ gesetzt. Also messen wir nun die Zeit in Längen. Um die Formeln in SI-Einheiten zu erhalten, kann anschließend mit c multipliziert werden.

Definition 2.1 (Einstein'sche Summenkonvention). Über doppelt vorkommende Indizes wird summiert. Das

heißt, jedes Mal, wenn ein Index zweimal innerhalb eines Terms vorkommt, ist implizit die Summe über diesen Index gemeint.

$$\sum_{i=1}^3 a^i b^i \equiv a^i b^i \equiv \delta_{ij} a^i b^j$$

Achtung: Die Summenkonvention gilt nur, wenn Indizes genau *zweimal* vorkommen und nicht mehrfach!

$$a^i b^i c^i \neq \sum_{i=1}^3 a^i b^i c^i$$

2.2.3 Herleitung generelle Lorentz-Transformation

Die bereits im vorherigen Kapitel verwendete Lorentz-Transformation soll nun in der generellen Form (also mit Bewegung in x-, y- und z-Richtung) aus dem Galileischen Relativitätsprinzip [Satz 1.3](#) hergeleitet werden, **ohne** dass die Konstanz der Lichtgeschwindigkeit verwendet wird.

Die **Lorentz-Transformation** beschreibt eine Koordinatentransformation von einem Inertialsystem I in ein dagegen in x-, y- und z-Richtung mit der Geschwindigkeit v^i (also $\vec{v} = (v_x, v_y, v_z)$) bewegtes Inertialsystem \bar{I} . I hat die Koordinaten (t, x^i) , \bar{I} hat die Koordinaten (\bar{t}, \bar{x}^i)

$$\begin{aligned}\bar{t} &= \gamma \left(t - \frac{v^j x^j}{c^2} \right) && \text{wobei } \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} > 1 \\ \bar{x}^i &= x^i - \gamma v^i t + (\gamma - 1) \frac{v^i v^j x^j}{c^2}\end{aligned}$$

Da wir uns in Inertialsystemen befinden, gibt es keine beschleunigten Bewegungen. Teilchen bewegen sich auf geraden Weltlinien. (Eine beschleunigte Bewegung wäre eine gekrümmte Weltlinie). Ein Bezugssystemwechsel muss daher Geraden auf Geraden sowie Ebenen auf Ebenen abbilden. Mathematisch besagt dies, dass die Transformation **linear** ist.

Nach dem galileischen Relativitätsprinzip gibt es weder ausgezeichnete Orte noch ausgezeichnete Raumrichtungen (Homogenität (an allen Orten gleich) + Isotropie (in alle Richtungen gleich)). Daher darf die einzige ausgezeichnete räumliche Richtung, die in der Transformation auftreten kann, der Relativgeschwindigkeitsvektor v^i . D.h. die Transformation muss *drehinvariant* sein. In einem Bezugssystem, das gegenüber dem anderen verdreht ist, ändert sich die Physik nicht.

Gesucht ist also eine lineare Transformation der Form

$$\begin{pmatrix} \bar{t} \\ \bar{x}^i \end{pmatrix} = A(v^i) \cdot \begin{pmatrix} t \\ x^j \end{pmatrix}$$

Da \bar{t} ein Skalar ist und \bar{x}^i die Komponenten eines Vektors sind, hat die allgemeinste lineare Form die Gestalt

$$\begin{aligned}\bar{t} &= \alpha t + \beta^j x^j \\ \bar{x}^i &= \gamma^i t + M^{ij} x^j\end{aligned}$$

wobei α ein Skalar, β^j ein Vektor, γ^i ein Vektor und M^{ij} eine Matrix sind.

Wegen der Drehinvarianz darf β^j nur proportional zum einzigen ausgezeichneten Vektor v^j sein, also $\beta^j = b(v) v^j$, wobei die Koeffizienten nur vom Betrag $v = \sqrt{v^i v^i}$ abhängen können. Entsprechend setzen wir $a = a(v)$.

Analog folgt für den Vektorterm $\gamma^i = d(v) v^i$. Für den Matrixterm ergibt sich ebenso wegen der Drehinvarianz $M^{ij} = e(v) \delta^{ij} + f(v) v^i v^j$.

Wir erhalten

$$\begin{aligned}\bar{t} &= a t + b v^j x^j \\ \bar{x}^i &= d v^i t + e x^i + f v^i (v^j x^j)\end{aligned}$$

wobei $a(v)$, $b(v)$, $d(v)$, $e(v)$, $f(v)$ zunächst unbekannte Funktionen des Betrags der Relativgeschwindigkeit $v = \sqrt{v^i v^i}$ sind.

Diese Transformation ist dann in Matrixform:

$$\begin{pmatrix} \bar{t} \\ \bar{x}^i \end{pmatrix} = \underbrace{\begin{pmatrix} a & b v^j \\ d v^i & e \delta^{ij} + f v^i v^j \end{pmatrix}}_{= A(v^i)} \begin{pmatrix} t \\ x^j \end{pmatrix} \quad (2.1)$$

wobei $b v^j$ ein Zeilenvektor, $d v^i$ ein Spaltenvektor und $e \delta^{ij} + f v^i v^j$ eine 3×3 Matrix ist.

Ziel: Bestimmung Koeffizienten $a(v)$, $b(v)$, $d(v)$, $e(v)$, $f(v)$

Wir verwenden dazu folgende Bedingungen:

1. Bedingung: Die Transformation soll **stetig** sein. Weiters soll eine Rücktransformation mit $-v^i$ muss wieder dasselbe Bezugssystem liefern, also die **Zusammensetzung der beiden Lorentz-Transformationen** muss der **identischen Abbildung** entsprechen:

$$I \xrightarrow{v^i} \bar{I} \xrightarrow{-v^i} I = id$$

Das heißt, die Matrix $A(v^i)$ muss mit $A(-v^i)$ die Einheitsmatrix ergeben $A(-v^i)A(v^i) = E$:

$$\begin{pmatrix} a & -b v^k \\ -d v^i & e \delta^{ik} + f v^i v^k \end{pmatrix} \begin{pmatrix} a & b v^j \\ d v^k & e \delta^{kj} + f v^k v^j \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 & 0^k \\ 0^i & \delta^{ij} \end{pmatrix}$$

Ausmultiplizieren der Matrizen ergibt:

$$(I) \quad a^2 - bd v^2 = 1$$

$$(II) \quad ab v^j - bf v^2 v^j = b v^j (a - e - fv^2) = 0$$

$$(III) \quad ad v^i + df v^i v^2 = d v^i (a - e - fv^2) = 0$$

$$(IV) \quad -bd v^i v^j + e^2 \delta^{ij} + 2ef v^i v^j + f^2 v^i v^j v^2 = \delta^{ij}$$

Für $v = 0$ ist $I = \bar{I}$. Sei also $v \neq 0$.

Aus Gleichung (II) folgt, dass Fall 1: $b v^j = 0$ (also $b=0$) oder Fall 2: $a - e - fv^2 = 0$. Betrachte nur den Fall 2. Man kann zeigen, dass Fall 1 auf einen Spezialfall führt, der von den anderen Transformationen auch abgedeckt wird.

Aus Gleichung (III) folgt, dass Fall 1: $dv^i = 0$ (also $b=0$) oder Fall 2: $-a + e + fv^2 = 0$. Betrachte nur den Fall 2. Selbes Argument wie vorhin. Beachte, dass $-a + e + fv^2 = 0$ einfach $a - e - fv^2 = 0$ mal -1 ist.

Aus Gleichung (IV) folgt, dass $e^2 = 1$ und $-bd + 2ef + f^2v^2 = 0$.

Erhalte also

$$\begin{aligned} (\text{I}') \quad & a^2 - bd v^2 = 1 \\ (\text{II}') \quad & a - e - fv^2 = 0 \\ (\text{III}') \quad & e^2 = 1 \\ (\text{IV}') \quad & -bd + 2ef + f^2v^2 = 0 \end{aligned}$$

Aus Gleichung (III') folgt unmittelbar:

$$e^2 = 1 \quad \Rightarrow \quad e = \pm 1.$$

Für $v = 0$ soll die Transformation in die Identität übergehen:

$$\underbrace{\begin{pmatrix} a & b 0^j \\ d 0^i & e \delta^{ij} + f 0^i 0^j \end{pmatrix}}_{= A(0^i)} = \begin{pmatrix} 1 & 0^k \\ 0^i & \delta^{ij} \end{pmatrix}$$

Daher ist $e = 1$ bei $v = 0$. Da die Transformation stetig sein soll, soll e es auch sein und kann somit keine Sprungstelle besitzen, bei der zu -1 gewechselt wird. Daher ist $\forall v$:

$$e = 1$$

Aus (II') folgt nun mit $e = 1$:

$$a - 1 = fv^2 \quad \Rightarrow \quad \boxed{f = \frac{a - 1}{v^2}} \quad (2.2)$$

2. Bedingung: Ein Körper, der sich gleichförmig mit v^i bewegt, kann durch eine Transformation in Ruhe gebracht werden.

Wähle eine Transformation in das Bezugssystem, dass sich mit dem bewegten Körper mitbewegt. In diesem Fall ist dann die Geschwindigkeit des Körpers im neuen Bezugssystem gleich 0. Mit $x^j = v^j t$ (Körper bewegt sich mit Geschwindigkeit v) gilt (siehe (2.1) mit $e = 1$):

$$\begin{pmatrix} a & b v^j \\ d v^i & \delta^{ij} + f v^i v^j \end{pmatrix} \begin{pmatrix} t \\ v^j t \end{pmatrix} = \begin{pmatrix} \tau \\ 0^i \end{pmatrix} \quad (2.3)$$

Wobei τ die Zeit im mitbewegten Bezugssystem ist und 0^i die Position des mit v bewegten Objekts. Diese ist im Ursprung, da für $t = 0$ dort 0^i steht. Da sich die Position des Objekts im mitbewegten Bezugssystem nicht ändert, bleibt es bei der Position 0^i . τ ist übrigens die Eigenzeit die, vom ruhenden System im mitbewegten System beobachtet wird.

Ausmultiplizieren ergibt (wobei nur die Ortskomponente relevant ist):

$$\begin{aligned}
 & dv^i t + v^i t + f v^2 v^i t = 0^i \\
 \Leftrightarrow & v^i t (d + 1 + \underbrace{f v^2}_{= a - 1}) = 0^i \\
 \Leftrightarrow & v^i t (d + a) = 0^i \\
 \Leftrightarrow & d + a = 0 \quad \text{da i.A. } v \neq 0 \text{ und } t > 0 \\
 \Leftrightarrow & \boxed{d = -a}
 \end{aligned}$$

Setze $d = -a$ in (I') ein und forme auf b um:

$$\begin{aligned}
 & a^2 + abv^2 = 1 \\
 \Leftrightarrow & \boxed{b = \frac{1 - a^2}{av^2}}
 \end{aligned}$$

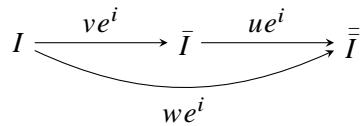
Wir haben nun alle Koeffizienten in Abhängigkeit von a und v bestimmt, womit sich die Lorentz-Transformation anschreiben lässt als:

$$A(v^i) = \begin{pmatrix} a & \frac{1-a^2}{av^2} v^j \\ -a v^i & \delta^{ij} + \frac{a-1}{v^2} v^i v^j \end{pmatrix}$$

Nun muss noch a bestimmt werden.

3. Bedingung: **Transitivität**: Zwei hintereinander ausgeführte Lorentz-Transformationen mit unterschiedlichen Geschwindigkeiten, aber in dieselbe Richtung, sollen wieder eine Lorentz-Transformation bilden.

Betrachte dazu einen Koordinatenwechsel zwischen mehreren Inertialsystemen. Das Inertialsystem \bar{I} bewege sich mit der Geschwindigkeit ve^i relativ zu I , und ein weiteres System $\bar{\bar{I}}$ bewege sich mit der Geschwindigkeit ue^i relativ zu \bar{I} . Somit bewegt sich $\bar{\bar{I}}$ relativ zu I mit der zusammengesetzten Geschwindigkeit we^i , wobei wir noch nicht wissen wie we^i aussieht.



Die Hintereinanderausführung der beiden Transformationen $A(v^i)$ und $A(u^i)$ muss der direkten Transformation mit der Geschwindigkeit w^i entsprechen. Das heißt:

$$A(ue^i) A(ve^i) = A(we^i)$$

Beziehungsweise

$$\begin{pmatrix} a_u & \frac{1-a_u^2}{a_u u^2} u^i \\ -a_u u e^k & \delta^{ki} + \frac{a_u - 1}{u^2} u^k u^i \end{pmatrix} \begin{pmatrix} a_v & \frac{1-a_v^2}{a_v v^2} v^j \\ -a_v v e^i & \delta^{ij} + \frac{a_v - 1}{v^2} v^i v^j \end{pmatrix} = \begin{pmatrix} a_w & \frac{1-a_w^2}{a_w w^2} w^j \\ -a_w w e^k & \delta^{kj} + \frac{a_w - 1}{w^2} w^k w^j \end{pmatrix}$$

wobei $a_v = a(v)$, $a_u = a(u)$ und $a_w = a(w)$

Aus der linken oberen Komponente erhalten wir die Gleichung:

$$a_w = a_v a_u - \frac{1 - a_u^2}{a_u u} a_u u \quad (2.4)$$

Der rechte untere Block ergibt:

$$\begin{aligned} -a_u u e^k \cdot \frac{1 - a_v^2}{a_v v} e^j + (\delta^{ki} + (a_u - 1)e^k e^i) \cdot (\delta^{ij} + (a_v - 1)e^i e^j) &= \delta^{kj} + (a_w - 1)e^k e^j \\ \left(-a_u u \frac{1 - a_v^2}{a_v v} + (a_v - 1) + (a_u - 1) + (a_v - 1)(a_u - 1) \right) e^k e^j + \underbrace{\delta^{ki} \delta^{ij}}_{= \delta^{kj}} &= \delta^{kj} + (a_w - 1)e^k e^j \end{aligned}$$

Vergleiche Koeffizienten von $e^k e^j$:

$$\begin{aligned} a_w - 1 &= -a_u u \frac{1 - a_v^2}{a_v v} + (a_v - 1) + (a_u - 1) + (a_v - 1)(a_u - 1) \\ &= -a_u u \frac{1 - a_v^2}{a_v v} + a_v a_u - 1 \end{aligned}$$

Damit folgt

$$a_w = -a_u u \frac{1 - a_v^2}{a_v v} + a_v a_u \quad (2.5)$$

Setze (2.4) und (2.5) gleich:

$$\begin{aligned} \cancel{a_v a_u} - \frac{1 - a_u^2}{a_u u} a_u u &= -a_u u \frac{1 - a_v^2}{a_v v} + \cancel{a_v a_u} \\ \Leftrightarrow \frac{a_u^2 - 1}{(a_u u)^2} &= \frac{a_v^2 - 1}{(a_v v)^2} =: K = \text{konstant} \end{aligned}$$

Da die linke Seite nur noch von u und die rechte nur noch von v abhängt und die Gleichheit für alle u und v gelten soll, kann diese nur eintreten, wenn sie konstant sind. ($\forall u, v \in \mathbb{R} : g(u) = g(v) \Rightarrow g \equiv \text{const}$)

$$\begin{aligned} a_v^2 - 1 &= K a_v^2 v^2 \\ \Leftrightarrow a_v^2 \left(1 - K v^2 \right) &= 1 \\ \Leftrightarrow a_v &= \pm \frac{1}{\sqrt{1 - K v^2}} \end{aligned}$$

Für $v \rightarrow 0$, also wenn nur die triviale Transformation vorliegt, muss der Matrixeintrag a_v gegen 1 gehen. Daher ist für uns nur der positive Zweig der Lösung relevant.

Also gilt:

$$a = \frac{1}{\sqrt{1 - K v^2}} \quad (2.6)$$

(2.2) $f = \frac{1 - a^2}{a v^2}$ lässt sich mit Hilfe von a und K ausdrücken:

$$\frac{1 - a^2}{a v^2} = -K a$$

Damit schreiben wir die Transformation in Matrixform als

$$\begin{pmatrix} \bar{t} \\ \bar{x}^i \end{pmatrix} = \begin{pmatrix} a & -Kav^j \\ -av^i & \delta^{ij} + \frac{a-1}{v^2} v^i v^j \end{pmatrix} \begin{pmatrix} t \\ x^j \end{pmatrix} \quad (2.7)$$

Damit haben wir die *Form* der Lorentz-Transformationen einzig aus dem Galileischen Relativitätsprinzip bereits bestimmt. Bisher ging die Konstanz der Lichtgeschwindigkeit NICHT ein. Wir erhalten eine Familien von Transformationen abhängig von K .

Für $K = 0$ ist $a = 1$ und wir erhalten die Galilei-Transformation mit einer absoluten Zeit ($t = \bar{t}$). Für $K > 0$ erhalten wir eine endliche Geschwindigkeit, die in allen Inertialsystemen gleich ist. Jetzt müssen wir nur noch K bestimmen.

Bestimmung von K über den Lichtkegel

Das [Michelson-Morley-Experiment](#) zeigte, dass die Lichtgeschwindigkeit konstant ist. Das kann nun verwendet werden, um den Wert von K zu bestimmen. Wir betrachten nun einen Punkt $s = (t, x^1, x^2, x^3) \in \mathbb{M}$ auf dem Lichtkegel. Da wir ohne Beschränkung der Allgemeinheit so skaliert haben, dass der Winkel des Lichtkegels bei 45° liegt, gilt im System I :

$$t^2 = x^i x^i$$

D.h. das Längenquadrat der zeitlichen Komponente stimmt mit dem der räumlichen Komponente überein.

Im System \bar{I} , das sich mit der Geschwindigkeit v relativ zu I bewegt, gilt ebenso

$$\bar{t}^2 = \bar{x}^i \bar{x}^i$$

Licht breitet sich in jedem Inertialsystem gleich schnell aus, daher gilt für einen Punkt auf dem Lichtkegel:

$$t^2 = x^i x^i \Leftrightarrow \bar{t}^2 = \bar{x}^i \bar{x}^i$$

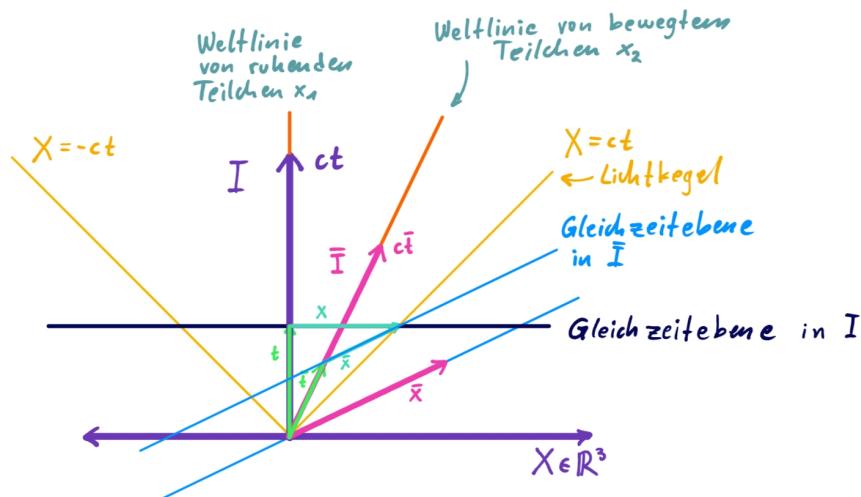


Abbildung 2.11: Die Längenquadrate von t und x stimmen in I überein genau dann wenn die Längenquadrate von \bar{t} und \bar{x} in \bar{I} übereinstimmen.

Mit der Transformation (2.7) gilt daher:

$$\begin{aligned}\bar{t} &= at - Kav^j x^j \\ \bar{x}^i &= -av^i t + x^i + \frac{a-1}{v^2} v^i v^j x^j\end{aligned}\tag{2.8}$$

Quadrieren wir diese beiden Gleichungen, so erhalten wir:

$$\begin{aligned}\bar{t}^2 &= a^2 t^2 - 2Ka^2 v^i x^i t + K^2 a^2 v^2 x^2 \\ \bar{x}^2 &= a^2 v^2 t^2 + x^2 + \frac{(a-1)^2}{v^2} (vx)^2 - 2avxt - 2a(a-1)vxt + 2(vx)^2 \frac{a-1}{v^2}\end{aligned}\tag{2.9}$$

Da auf dem Lichtkegel $t^2 = x^2$ und $\bar{t}^2 = \bar{x}^2$ gilt, obiges gleichsetzen und auch das im zweiten Term vorkommende x^2 auf t^2 umschreiben.

$$a^2 t^2 + K^2 a^2 (vx)^2 - 2a^2 Ktvx = \left(a^2 v^2 + 1\right) t^2 + \frac{(a-1)^2}{v^2} (vx)^2 - 2a^2 vxt + \frac{2(a-1)}{v^2} (vx)^2$$

Koeffizientenvergleich für t^2 Terme:

$$a^2 = a^2 v^2 + 1 \Rightarrow a^2(1 - v^2) = 1$$

Setze (2.6) ein:

$$\frac{1 - v^2}{1 - Kv^2} = 1 \Rightarrow \boxed{K = 1}$$

Es folgt die generelle Lorentz-Transformation:

Damit kann nun a bestimmt werden:

$$a = \frac{1}{\sqrt{1 - Kv^2}} = \frac{1}{\sqrt{1 - v^2}} \stackrel{c^2=1}{=} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} =: \gamma \dots \text{Lorentz-Faktor}$$

Setze $K = 1$ und $a = \gamma$ in (2.8) ein und erhalte schließlich:

$$\begin{aligned}\bar{t} &= \gamma t - \gamma v^j x^j = \gamma \left(t - v^j x^j\right) \stackrel{c^2=1}{=} \gamma \left(t - \frac{v^j x^j}{c^2}\right) \\ \bar{x} &= -\gamma t v^i + x^i + \frac{\gamma - 1}{v^2} v^i v^j x^j\end{aligned}$$

2.3 Minkowski-Metrik und Raumzeitabstand

Wir haben in der Herleitung der Lorentz-Transformation gesehen, dass für ein Punkt auf dem Lichtkegel in beiden Inertialsystemen I und \bar{I} gilt:

$$-t^2 + x^i x^i = 0 \Leftrightarrow -\bar{t}^2 + \bar{x}^i \bar{x}^i = 0\tag{2.10}$$

Tatsächlich gilt das sogar ganz allgemein, wie wir im Folgenden sehen werden. In (2.9) haben wir die Längenquadrate bestimmt:

$$\begin{aligned}\bar{t}^2 &= a^2 t^2 - 2K a^2 vxt + K^2 a^2 v^2 x^2 \stackrel{K=1}{=} a^2 t^2 - 2a^2 vxt + a^2 v^2 x^2 \\ \bar{x}^2 &= a^2 v^2 t^2 + x^2 + \frac{(a-1)^2}{v^2} (vx)^2 - 2avxt - 2a(a-1)vxt + 2\frac{a-1}{v^2} (vx)^2 \\ &= a^2 v^2 t^2 + x^2 + \frac{a^2 - 1}{v^2} (vx)^2 - 2a^2 vxt\end{aligned}$$

Damit können wir nun den Ausdruck $-\bar{t}^2 + \bar{x}^2$ bilden:

$$\begin{aligned}-\bar{t}^2 + \bar{x}^2 &= -a^2 t^2 + 2a^2 \cancel{vxt} - a^2 v^2 x^2 + a^2 v^2 t^2 + x^2 + \frac{a^2 - 1}{v^2} (vx)^2 - 2a^2 \cancel{vxt} \\ &= -t^2 a^2 (1 - v^2) + x^2 + (vx)^2 \frac{a^2 - 1}{v^2} \quad | \quad a^2 = \frac{1}{1 - v^2} \\ &= -t^2 + x^2 + (vx)^2 \left(a^2 + \frac{\frac{1}{1-v^2} - 1}{v^2} \right) \\ &= -t^2 + x^2 + (vx)^2 \left(a^2 + \frac{\frac{v^2}{1-v^2}}{v^2} \right) \\ &= -t^2 + x^2 + (vx)^2 \left(a^2 - a^2 \right) \\ &= -t^2 + x^2\end{aligned}$$

Damit haben wir gezeigt, dass der Ausdruck

$$-\bar{t}^2 + \bar{x}^2 = -t^2 + x^2$$

in allen Inertialsystemen denselben Wert besitzt. Dies gilt unabhängig davon, ob es sich um Lichtsignale handelt oder nicht.

Das bedeutet: Der 4-dimensionale Abstand $-t^2 + x^2$ ist eine vom Bezugssystem unabhängige Größe – eine *Lorentz-Invariante*. Er ändert sich nicht durch Transformation in ein anderes Inertialsystem.

Diese Invarianz motiviert die Definition des **Raumzeitabstandes** Δs^2 :

$$\Delta s^2 := -c^2 \Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$$

Mit der sogenannten *Minkowski-Metrik* $\eta_{\mu\nu}$, kann dies kompakt geschrieben werden:

$$\Delta s^2 = \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu,$$

wobei

$$(\eta_{\mu\nu}) := \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \text{diag}(-1, 1, 1, 1) \quad \mu, \nu = 0, 1, 2, 3$$

Sie hat die Signatur $(-, +, +, +)$. Der Minkowski-Raum \mathbb{M} ist ein Vektorraum, zusammen mit der Minkowski-Metrik $\eta_{\mu\nu}$.

Die Raumzeit-Koordinaten eines Ereignisses lassen sich also zu einem Vierervektor zusammenfassen, wobei $x^0 = ct$ die zeitliche Komponente und x^i die räumlichen Komponenten sind:

$$v = \begin{pmatrix} ct \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \in \mathbb{M}$$

Dann ist das **Längenquadrat** des Vektors v :

$$v^2 = \eta_{\mu\nu} v^\mu v^\nu = -c^2 t^2 + \vec{x}^2 = -c^2 t^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 \quad (2.11)$$

wobei jeweils über $\mu\nu$ summiert wird und die Summe weggelassen wurde (Einstiensche Summenkonvention).

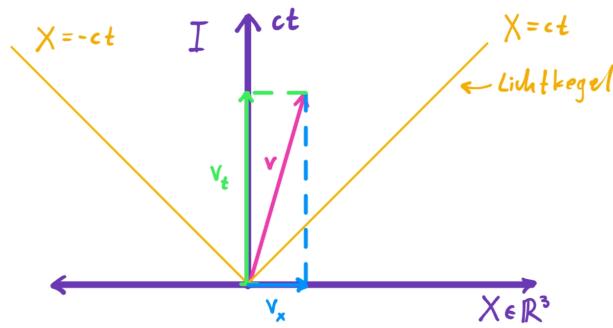


Abbildung 2.12: $v_t = ct$ zeitliche Komponente des Vektors v ; $v_x = (x^1, x^2, x^3) = \vec{x}$ räumliche Komponente vom Vektor v

Durch das negative Vorzeichen der Zeitkomponente kann der Raumzeitabstand eines Vektors im Minkowski-Raum positiv, negativ oder null werden:

- $v^2 < 0$: **zeitartig** – Ereignisse liegen innerhalb des Lichtkegels und werden auch *kausale* Vektoren genannt, da sie die mögliche Zukunft bedeuten.
- $v^2 = 0$: **lichtartig** – Ereignisse liegen auf dem Lichtkegel.
- $v^2 > 0$: **raumartig** – Ereignisse liegen außerhalb des Lichtkegels.

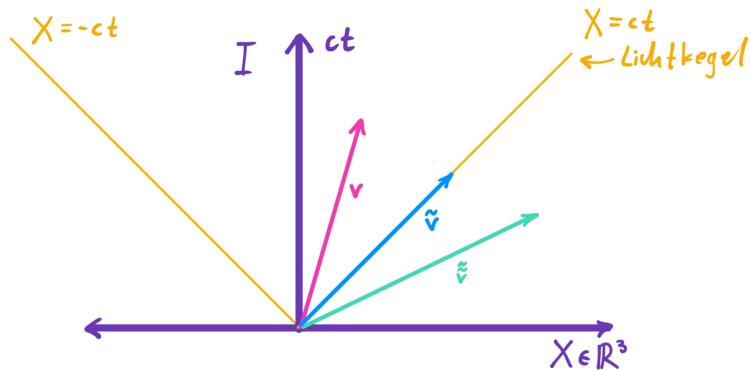


Abbildung 2.13: v ist zeitartig, \tilde{v} ist lichtartig, $\tilde{\tilde{v}}$ ist raumartig

2.3.1 Mathematische Struktur der Minkowski-Metrik (nicht Teil der VO)

Im Folgenden möchte ich die soeben kennengelernte Minkowski-Metrik mathematisch einordnen. Falls folgende Begriffe nicht bekannt sind, diesen Teil der Mitschrift überspringen, sie zu Ende lesen und am Schluss wieder diesen Abschnitt durchlesen. Alle Begriffe werden im Laufe der Mitschrift eingeführt.

TLDR: Die Minkowski-Metrik ist keine Metrik im Sinne der [metrischen Räume](#), sondern eine symmetrische, nicht entartete Bilinearform, die nicht positiv definit ist. Auf dem Minkowski-Vektorraum \mathbb{M} ist sie ein Pseudo-Skalarprodukt. Auf der flachen Raumzeit, aufgefasst als Mannigfaltigkeit, ist sie eine pseudo-riemannsche Metrik bzw. ein metrischer Tensor. Da sie negative Werte annehmen kann, induziert sie keine Norm, sondern ein Längenquadrat v^2 , welches ebenso negativ sein kann.

Definition 2.2 (reelles Skalarprodukt).

Quelle: [https://de.wikipedia.org/wiki/Skalarprodukt#Definition_\(Axiomatik\)](https://de.wikipedia.org/wiki/Skalarprodukt#Definition_(Axiomatik))

Sei V ein reeller Vektorraum. Eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$$

heißt **Skalarprodukt**, falls es eine positiv definite symmetrische Bilinearform ist, d.h. für alle $x, y, z \in V$ und $\lambda \in \mathbb{R}$ die folgenden Eigenschaften gelten:

1. **Bilinear:** linear in beiden Argumenten

$$\begin{aligned} \langle x + y, z \rangle &= \langle x, z \rangle + \langle y, z \rangle & \langle \lambda x, y \rangle &= \lambda \langle x, y \rangle \\ \langle x, y + z \rangle &= \langle x, y \rangle + \langle x, z \rangle & \langle x, \lambda y \rangle &= \lambda \langle x, y \rangle \end{aligned}$$

2. **Symmetrisch:**

$$\langle x, y \rangle = \langle y, x \rangle$$

3. **Positiv definit:**

$$\langle x, x \rangle \geq 0 \quad \text{und} \quad \langle x, x \rangle = 0 \Leftrightarrow x = 0$$

Ein Skalarprodukt liefert Winkelbegriffe/Orthogonalität und induziert durch Wurzelziehen eine Norm $\|x\| := \sqrt{\langle x, x \rangle}$. Die Norm induziert durch Differenzenbildung eine Metrik $d(x, y) := \|x - y\|$.

Bemerkung 2.3.

Quelle: https://de.wikipedia.org/wiki/Bilinearform#Nicht_ausgeartete_Bilinearform

Aus der positiven Definitheit folgt, dass jedes reelle Skalarprodukt **nicht entartet/nicht ausgeartet** ist, d. h.

$$\begin{aligned}\forall x \in V \setminus \{0\} \exists y \in V : \langle x, y \rangle \neq 0 \\ \forall y \in V \setminus \{0\} \exists x \in V : \langle x, y \rangle \neq 0\end{aligned}$$

Nicht entartet heißt: kein Vektor, außer der Nullvektor, ist zu allen anderen Vektoren orthogonal. Aus nicht entartet, folgt allerdings NICHT die positive Definitheit.

Definition 2.4.

Quelle: <https://matheplanet.com/default3.html?call=article.php?sid=1516> Sei V ein reeller Vektorraum. Eine Abbildung

$$(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$$

heißt **Pseudo-Skalarprodukt**, falls es eine nicht entartete symmetrische Bilinearform ist.

Quelle: https://www.thp.uni-koeln.de/~rk/tpI_18.html/vl22.pdf

Bei einem Pseudo-Skalarprodukt ist also $(x, x) < 0$ zugelassen. Somit erhalten wir aus einem Pseudo-Skalarprodukt nur ein "Normquadrat" $\|x\|^2$ (genauer: [quadratische Form](#) $q(x)$) und keine Norm

$$q(x) := (x, x)$$

Wobei der Name irreführend ist, da die quadratische Form dennoch negativ sein kann. Entsprechend erhalten wir auch keine Metrik, sondern

$$d(x, y)^2 := q(x - y)^2$$

Definition 2.5.

Quellen: https://de.wikipedia.org/wiki/Riemannsche_Mannigfaltigkeit#Definition

https://de.wikipedia.org/wiki/Pseudo-riemannsche_Mannigfaltigkeit

Sei M eine differenzierbare Mannigfaltigkeit. Ein **riemannsche Metrik** ist eine Abbildung

$$g : M \rightarrow \mathfrak{Bil}(T_p M, \mathbb{R}) : p \mapsto g_p$$

die jedem Punkt $p \in M$ eine positiv definite, symmetrische Bilinearform

$$g_p : T_p M \times T_p M \rightarrow \mathbb{R}$$

zuordnet. Wobei $T_p M$ der Tangentialraum (Vektorraum) an einem Punkt $p \in M$ ist.

Das heißt die riemannsche Metrik ist ein $(0,2)$ -Tensorfeld $g \in \mathcal{T}_2^0(M)$.

Quelle: [https://de.wikipedia.org/wiki/Riemannsche_Mannigfaltigkeit#Riemannsche_Mannigfaltigkeiten_als_metrische_Räume](https://de.wikipedia.org/wiki/Riemannsche_Mannigfaltigkeit#Riemannsche_Mannigfaltigkeiten_als_metrische_R%C3%A4ume)

Die riemannsche Metrik ist keine Metrik im Sinne der Theorie der [metrischen Räume](#), sondern punktweise ein Skalarprodukt auf den Tangentialräumen. Wie zuvor erwähnt, induziert es eine Metrik. Somit können riemannsche Mannigfaltigkeiten als metrische Räume verstanden werden. Die riemannsche Metrik legt Längen und Winkel fest.

Definition 2.6.

Quelle: https://de.wikipedia.org/wiki/Pseudo-riemannsche_Mannigfaltigkeit

Ist die Bilinearform g_p symmetrisch und nicht entartet, jedoch nicht positiv definit, so heißt g ein **pseudo-riemannsche Metrik**.

Die pseudo-riemannsche Metrik ist punktweise ein Pseudo-Skalarprodukt auf den Tangentialräumen.

In der Literatur ist der **metrische Tensor** nicht eindeutig definiert. Manchmal wird damit die riemannsche Metrik, manchmal die pseudo-riemannsche Metrik bezeichnet.

Definition 2.7.

Quellen: <http://www.vermessungsseiten.de/erdvermessung/node10.html>

<https://matheplanet.com/default3.html?call=article.php?sid=1516>

https://www.thp.uni-koeln.de/~rk/tpl_18.html/vl22.pdf

Der **Minkowski-Raum** $\mathbb{M} := \{(ct, x^1, x^2, x^3) \in \mathbb{R}^4\} \cong \mathbb{R}^4$ ist ein reeller Vektorraum, ausgestattet mit dem Pseudo-Skalarprodukt

$$\eta : \mathbb{M} \times \mathbb{M} \rightarrow \mathbb{R}$$

mit

$$\eta(x, y) := \underbrace{-c^2 t^2}_{=-x^0 y^0} + x^1 y^1 + x^2 y^2 + x^3 y^3 := \eta_{\mu\nu} x^\mu y^\nu$$

wobei

$$\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$$

Quelle: https://de.wikipedia.org/wiki/Raumzeit#Minkowski-Raum,_Vierervektoren

Die Matrix $\eta_{\mu\nu}$ ist also auf dem Minkowski-Vektorraum \mathbb{M} ein *Pseudo-Skalarprodukt*. Auf der flachen Raumzeit, aufgefasst als Mannigfaltigkeit ([Lorentzsche Mannigfaltigkeit](#)), ist sie eine *pseudo-riemannsche Metrik* bzw. ein *metrischer Tensor*. Somit ist sie keine Metrik im Sinne der Theorie der metrischen Räume. Da sie nicht positiv definit ist (indefinit) ist ebenso zugelassen

$$\eta(x, x) < 0$$

Somit erhalten ein "Normquadrat" $\|x\|^2$ (genauer: [quadratische Form](#) $q(x)$) und keine Norm

$$q(x) := \eta(x, x) = \eta_{\mu\nu} x^\mu x^\nu$$

Das Normquadrat kann somit negativ sein, womit der Name irreführend ist. Entsprechend erhalten wir auch *keine Metrik*, sondern ein *Längenquadrat/Abstandsquadrat* für $s_1 = (ct, x^1, x^2, x^3)$, $s_2 = (ct, y^1, y^2, y^3) \in \mathbb{M}$

$$ds^2 := \Delta s^2 := (s_1 - s_2)^2 = q(x - y)$$

Das Längenquadrat kann somit negativ sein. Wir erhalten also negative Längen. Dennoch wird η bzw. $\eta_{\mu\nu}$ als **Minkowski-Metrik** bezeichnet. Sie stellt die Beziehung von Teilchen/Ereignissen in der flachen Raumzeit dar.

Für das Längenquadrat eines Vektors $v \in \mathbb{M}$ schreiben wir

$$v^2 = \eta_{\mu\nu} v^\mu v^\nu$$

Das Längenquadrat $\Delta s^2 := \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu$ wird **Raumzeitabstand** genannt.

Quellen: <https://matheplanet.com/default3.html?call=article.php?sid=1516>

https://en.wikipedia.org/wiki/Orthonormal_basis

https://en.wikipedia.org/wiki/Orthonormal_basis#Examples

Mit einem Pseudo-Skalarprodukt kann eine Pseudo-Orthonormalbasis definiert werden.

Definition 2.8 (Pseudo-Orthonormalbasis). Eine Basis E^μ von V heißt **pseudo-orthonormal** bezüglich eines Pseudo-Skalarprodukts (\cdot, \cdot) , falls

$$(E^\mu, E^\nu) = g_{\mu\nu}$$

wobei $g_{\mu\nu}$ eine diagonale Matrix mit Einträgen ± 1 entsprechend der Signatur des Pseudo-Skalarprodukts ist. Siehe später in [Beispiel 3.64](#).

Kapitel 3

Multilineare Algebra

Um die allgemeine Relativitätstheorie mathematisch präzise formulieren zu können, benötigen wir Werkzeuge aus der sogenannten multilinear Algebra. Dieses Kapitel führt die wichtigsten Begriffe ein, auf denen später die Tensorrechnung aufbaut.

3.1 Vektorräume

Definition 3.1 (Nicht Teil der VO).

Ein **Körper** ist ein Tupel $(\Gamma, +, \cdot)$, bestehend aus einer Menge Γ und zwei inneren zweistelligen Verknüpfungen $+$ und \cdot (Addition und Multiplikation):

$$\begin{aligned}+ : \Gamma \times \Gamma &\rightarrow \Gamma, & (a, b) &\mapsto a + b \\ \cdot : \Gamma \times \Gamma &\rightarrow \Gamma, & (a, b) &\mapsto a \cdot b\end{aligned}$$

Für die Verknüpfungen gleiten die folgenden Axiome:

1. Addition

- (1.1) $\forall a, b, c \in \Gamma : a + (b + c) = (a + b) + c$ (Assoziativgesetz)
(1.2) $\forall a, b \in \Gamma : a + b = b + a$ (Kommutativgesetz)
(1.3) $\exists 0 \in \Gamma \forall a \in \Gamma : 0 + a = a$ (neutrales Element bzgl. $+$)
(1.4) $\forall a \in \Gamma \exists -a \in \Gamma : (-a) + a = 0$ (additives Inverses)

Damit ist $(\Gamma, +)$ eine *kommutative/abelsche Gruppe*.

2. Multiplikation

- (2.1) $\forall a, b, c \in \Gamma : a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativgesetz)
(2.2) $\forall a, b \in \Gamma : a \cdot b = b \cdot a$ (Kommutativgesetz)
(2.3) $\exists 1 \in \Gamma \forall a \in \Gamma : 1 \cdot a = a$ (neutrales Element bzgl. \cdot)
(2.4) $\forall a \in \Gamma \setminus \{0\} \exists a^{-1} \in \Gamma : a^{-1} \cdot a = 1$ (multiplikatives Inverses)

Damit ist $(\Gamma \setminus \{0\}, \cdot)$ ebenfalls eine kommutative/abelsche Gruppe.

3. Zusammenspiel von Addition und Multiplikation

$$(3.1) \quad \forall a, b, c \in \Gamma : a \cdot (b + c) = a \cdot b + a \cdot c \quad (\text{Linksdistributivgesetz})$$

$$(3.2) \quad \forall a, b, c \in \Gamma : (b + c) \cdot a = b \cdot a + c \cdot a \quad (\text{Rechtsdistributivgesetz})$$

Beispiele: $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$

Definition 3.2. Ein **Vektorraum** über einem Körper $(\Gamma, +, \cdot)$, kurz Γ , ist ein Tupel (V, \oplus, \odot) , bestehend aus einer Menge V , einer inneren zweistelligen Verknüpfung \oplus (Vektoraddition) und einer äußeren zweistelligen Verknüpfung \odot (Skalarmultiplikation).

$$\begin{aligned} \oplus : V \times V &\rightarrow V, & (u, w) &\mapsto u \oplus w \\ \odot : \Gamma \times V &\rightarrow V, & (\lambda, u) &\mapsto \lambda \odot u \end{aligned}$$

Für die Verknüpfungen gleiten die folgenden Axiome:

1. Vektoraddition

$$(1.1) \quad \forall u, v, w \in V : u \oplus (v \oplus w) = (u \oplus v) \oplus w \quad (\text{Assoziativgesetz})$$

$$(1.2) \quad \forall v, w \in V : v \oplus w = w \oplus v \quad (\text{Kommutativgesetz})$$

$$(1.3) \quad \exists 0 \in V \forall v \in V : 0 \oplus v = v \quad (\text{neutrales Element bzgl. } \oplus \text{ (Nullvektor)})$$

$$(1.4) \quad \forall v \in V \exists (-v) \in V : (-v) \oplus v = 0 \quad (\text{additives Inverses})$$

Damit bildet (V, \oplus) eine kommutative/abelsche Gruppe.

2. Skalarmultiplikation

$$(2.1) \quad \forall v \in V, \lambda, \mu \in \Gamma : (\lambda \cdot \mu) \odot v = \lambda \odot (\mu \odot v) \quad (\text{Assoziativgesetz})$$

$$(2.2) \quad \exists 1 \in \Gamma \forall v \in V : 1 \odot v = v \quad (\text{neutrales Element bzgl. } \odot \text{ (Skalar)})$$

3. Zusammenspiel von Vektoraddition und Skalarmultiplikation

$$(3.1) \quad \forall u, v \in V, \lambda \in \Gamma : \lambda \odot (u \oplus v) = (\lambda \odot u) \oplus (\lambda \odot v) \quad (\text{Linksdistributivgesetz})$$

$$(3.2) \quad \forall v \in V, \lambda, \mu \in \Gamma : (\lambda + \mu) \odot v = (\lambda \odot v) \oplus (\mu \odot v) \quad (\text{Rechtsdistributivgesetz})$$

Beispiele: $(\mathbb{R}^n, +, \cdot)$ über dem Körper \mathbb{R} , $(\mathbb{C}^n, +, \cdot)$ über dem Körper \mathbb{C} , $n \in \mathbb{N}$

Bemerkung 3.3. Wir verwenden ab jetzt für Körperaddition & Vektoraddition $+$ und Körpermultiplikation & Skalarmultiplikation \cdot , um die Notation zu vereinfachen.

Es folgt bereits und muss somit nicht extra gefordert werden:

Korollar 3.4.

$$\exists 0 \in \Gamma, \forall v \in V : 0 \cdot v = 0 \in V$$

Beweis.

$$\begin{aligned} 0 \cdot v &\stackrel{\text{neutr. El.}}{\equiv} (0 + 0) \cdot v \stackrel{\text{Rechts-}}{\equiv} \stackrel{\text{distr.ges.}}{=} 0 \cdot v + 0 \cdot v \quad | - 0 \cdot v \text{ (additives Inverses)} \\ &\Rightarrow 0 = 0 \cdot v \end{aligned}$$

□

3.1.1 Linearkombination und lineare Unabhängigkeit

Definition 3.5. Sei (V, Γ) ein Vektorraum und $(v_\alpha)_{\alpha \in I} \subseteq V$ ein Teilsystem von Vektoren (auch Familie oder Teilmenge von Vektoren genannt). Die Indexmenge ist beliebig und kann sogar überabzählbar unendlich sein. Ein Vektor $v \in V$ heißt **Linearkombination** dieses Teilsystems, wenn es Skalare $\lambda^\alpha \in \Gamma$ gibt mit

$$v = \sum_{\alpha \in I} \lambda^\alpha v_\alpha \quad \lambda^\alpha \neq 0 \text{ für endlich viele } \alpha$$

Auch hier verwendet der Physiker gerne die Einsteinsche Summenkonvention: $\sum_{\alpha \in I} \lambda^\alpha v_\alpha \equiv \lambda^\alpha v_\alpha$

Bemerkung 3.6. Bei der Forderung, dass nur endlich viele $\lambda^\alpha \neq 0$ sein dürfen, geht [Korollar 3.4](#) ein und sie ist notwendig, da rein algebraisch definierte Vektorräume ([Definition 3.2](#)) keinen Abstandsbegriff in Form einer Topologie oder Metrik besitzen. Somit existiert kein Konvergenzbegriff für unendliche Summen und Linearkombinationen müssen endlich sein. Man beachte, dass es für jeden Vektorraum mind. eine Topologie gibt, mit diesem er ein topologischer Vektorraum wird, wir hier jedoch ausschließlich die algebraische Definition betrachten. Siehe Kapitel [Topologie Bemerkung 4.44](#).

Definition 3.7. Sei (V, Γ) ein Vektorraum und $(v_\alpha)_{\alpha \in I} \subseteq V$ ein Teilsystem von Vektoren, wobei I beliebige Indexmenge. Seien $\lambda^\alpha \in \Gamma$.

- $(v_\alpha)_{\alpha \in I}$ heißt **linear unabhängig** (kurz l.u.) genau dann wenn

$$\lambda^\alpha v_\alpha = 0 \Leftrightarrow \forall \alpha \in I : \lambda^\alpha = 0 \quad (3.1)$$

Das heißt: der Nullvektor lässt sich nur mit einer Linearkombination von Vektoren aus $(v_\alpha)_{\alpha \in I}$ darstellen, in der *alle* Koeffizienten $\lambda^\alpha = 0$. Da alle Vektoren unterschiedlich sind, können sie nicht voneinander abgezogen werden sodass 0 rauskommt. Die Vektoren sind nicht parallel.

Bemerkung: die Rückrichtung von (3.1) gilt klarerweise immer.

Ist diese Bedingung nicht erfüllt, heißt das Teilsystem linear abhängig:

- $(v_\alpha)_{\alpha \in I}$ heißt **linear abhängig** (kurz l.a.) genau dann wenn

$$\exists \text{ endlich viele } \alpha \in I \text{ mit } \lambda^\alpha \neq 0 : \lambda^\alpha v_\alpha = 0$$

Das heißt: der Nullvektor lässt sich auch mit einer Linearkombination von Vektoren aus $(v_\alpha)_{\alpha \in I}$ darstellen, in der *nicht alle* Koeffizienten $\lambda^\alpha = 0$.

- Eine äquivalente Definition für linear abhängig ist:

$$\exists \beta \in I : v_\beta = \sum_{\alpha \in I \setminus \{\beta\}} \lambda^\alpha v_\alpha$$

Das heißt: mind. ein Vektor des Teilsystems kann durch eine Linearkombination der anderen Vektoren des Teilsystems dargestellt werden. Die Vektoren sind parallel.

Bemerkung 3.8. Vektorräume $(v_\alpha)_{\alpha \in I} \subseteq V$ sind per Definition linear abhängig: zu jedem Vektor $v_\beta \in V$ ist das additive Inverse $-v_\beta$ ebenso enthalten. Dann kann beispielsweise die Linearkombination gewählt werden $1 \cdot v_\beta + 1 \cdot (-v_\beta) + \sum_{\alpha \in I \setminus \{\beta\}} 0 \cdot v_\alpha = 0$.

Proposition 3.9. Sei $(v_\alpha)_{\alpha \in I} \subseteq V$ linear unabhängig. Wenn ein Vektor $v \in V$ durch diese Familie mit einer Linearkombination darstellbar ist ($v = \lambda^\alpha v_\alpha$), dann ist die Darstellung eindeutig (d.h. $(\lambda^\alpha)_{\alpha \in I}$ eindeutig).

Beweis. Annahme: es gibt zwei verschiedene Mengen von Koeffizienten $(\lambda^\alpha)_{\alpha \in I}$ und $(\mu^\alpha)_{\alpha \in I}$ durch die $v \in V$ ebenso darstellbar ist: $v = \lambda^\alpha v_\alpha = \mu^\alpha v_\alpha$.

Dann ist

$$\begin{aligned} 0 &= v - v = \lambda^\alpha v_\alpha - \mu^\alpha v_\alpha \stackrel{\text{Distr.ges.}}{=} (\lambda^\alpha - \mu^\alpha)v_\alpha \\ &\stackrel{v_\alpha \text{ l.u.}}{\Rightarrow} \lambda^\alpha - \mu^\alpha = 0 \Rightarrow \lambda^\alpha = \mu^\alpha, \forall \alpha \in I \end{aligned}$$

□

3.1.2 Basis eines Vektorraums

Wir können Vektoren also eindeutig durch eine Linearkombination aus linear unabhängigen Vektoren darstellen. Beispielsweise können mit zwei linear unabhängigen Vektoren alle anderen, die in derselben Ebene liegen, aufgespannt werden, aber keine, die aus der Ebene herauszeigen. Also stellt sich die Frage: Welche kleinste Menge von Vektoren reicht aus, um alle anderen im Raum zu erzeugen? Dies führt uns zu dem Begriff der Basis.

Definition 3.10 (nicht Teil der VO). Sei (V, Γ) ein Vektorraum und $U := (v_\alpha)_{\alpha \in I} \subseteq V$ ein Teilsystem von Vektoren, wobei I beliebige Indexmenge. Seien $\lambda^\alpha \in \Gamma$.

Dann heißt

$$\text{span}(U) := \{\lambda^\alpha v_\alpha \mid \lambda^\alpha \in \Gamma, v_\alpha \in U\}$$

die **lineare Hülle** von U . Andere Schreibweisen sind: $[U]$, $\langle U \rangle$

$\text{span}(U)$ ist die Menge aller Linearkombinationen der Vektoren v_α . Sie ist der kleinste lineare Teilraum, der U enthält (also $U \subseteq \text{span}(U)$). Wenn U ein linearer Teilraum ist, dann gilt: $U = \text{span}(U)$.

U heißt **Erzeugendensystem** von V , wenn gilt:

$$\text{span}(U) = V$$

Das bedeutet: U spannt den gesamten Raum V auf. Jeder Vektor $v \in V$ lässt sich als Linearkombination von Vektoren aus U darstellen.

Definition 3.11 (nicht Teil der VO). Sei (V, Γ) ein Vektorraum und $B \subseteq V$ ein Teilsystem von Vektoren.

B heißt **Basis** des Vektorraums V wenn B ein linear unabhängiges Erzeugendensystem von V ist.

Das bedeutet: Jeder Vektor in V lässt sich als Linearkombination von Vektoren aus B darstellen und die Vektoren von B sind linear unabhängig.

Bemerkung 3.12. Basen eines Vektorraums sind nicht eindeutig.

Beispiel 3.13 (nicht Teil der VO).

- Im Vektorraum \mathbb{R}^2 ist

$$\mathcal{B} = \{(1, 0), (0, 1)\}$$

eine Basis (es gibt auch andere). Diese Basis wird *Standardbasis* oder *kanonische Basis* genannt.

- \mathbb{C} als Vektorraum über \mathbb{R} :

$$\mathcal{B} = \{1, i\}$$

- \mathbb{C} als Vektorraum über \mathbb{C} :

$$\mathcal{B} = \{1\} \quad \text{oder auch} \quad \mathcal{B} = \{i\}$$

- \mathbb{C}^2 als Vektorraum über \mathbb{C} :

$$\mathcal{B} = \{(1, 0), (0, i)\} \quad \text{oder auch} \quad \mathcal{B} = \{(i, 0), (0, -i)\}$$

Nun stellt sich die Frage: Hat jeder Vektorraum eine Basis? Um das zu beantworten, benötigen wir erst mal weitere Definitionen.

Definition 3.14. Eine Relation $\leq \subseteq M^2$ auf einer Menge M heißt **Totale Ordnung** oder **lineare Ordnung** ((M, \leq)), wenn für alle $x, y, z \in M$ gilt:

1. $x \leq x$ (Reflexivität)
2. $x \leq y \wedge y \leq x \Rightarrow x = y$ (Antisymmetrie)
3. $x \leq y \wedge y \leq z \Rightarrow x \leq z$ (Transitivität)
4. $x \leq y \vee y \leq x$ (Totalität / Linearität)

Die Bedingung 4. bedeutet, dass alle Elemente der Menge M paarweise vergleichbar sind.

Definition 3.15. Eine Relation $\leq \subseteq M^2$ auf einer Menge M heißt **Halbordnung** oder **partielle Ordnung** ((M, \leq)), wenn für alle $x, y, z \in M$ obige Bedingungen 1., 2. und 3. erfüllt sind. Es fehlt also die Vergleichbarkeit aller Elemente aus M bzgl. \leq .

Beispiel 3.16 (Totalordnung – nicht Teil der VO).

- Die Mengen (\mathbb{N}, \leq) , (\mathbb{Z}, \leq) , (\mathbb{Q}, \leq) und (\mathbb{R}, \leq) sind totalgeordnet. Auch die leere Menge (\emptyset, \prec) ist mit der leeren Ordnung totalgeordnet (trivialerweise, da es keine Elemente gibt, die nicht vergleichbar wären).
- Ein weiteres Beispiel ist $(\mathbb{C}, \leq_{\text{lex}})$, wenn man \mathbb{C} als $\mathbb{R} \times \mathbb{R}$ auffasst und die *lexikographische Ordnung* verwendet:

$$(a, b) \leq_{\text{lex}} (c, d) \Leftrightarrow (a < c) \text{ oder } (a = c \wedge b \leq d).$$

Beispiel 3.17 (Halbordnung, aber keine Totalordnung).

- Betrachte die Potenzmenge

$$\mathcal{P}(\{1, 2, 3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

mit der Inklusionsrelation \subseteq . Dann ist $(\mathcal{P}(\{1, 2, 3\}), \subseteq)$ eine Halbordnung, aber keine Totalordnung, da z. B.

$$\{1\} \not\subseteq \{2\} \quad \text{und} \quad \{2\} \not\subseteq \{1\}.$$

Diese beiden Mengen sind also *nicht vergleichbar*.

- (nicht Teil der VO) Betrachte (\mathbb{C}, \leq) , wobei \mathbb{C} als Menge aufgefasst wird. Dann gilt:

$$1 + 0 \cdot i = 1 \leq 2 = 2 + i \cdot 0, \quad 0 + 1 \cdot i = i \leq 2 \cdot i = 0 + 2 \cdot i$$

aber

$$1 + 0 \cdot i = 1 \not\leq i = 0 + 1 \cdot i$$

Daher sind die komplexen Zahlen im Allgemeinen *nicht vergleichbar* und bilden somit keine Totalordnung.

Definition 3.18. Sei (M, \leq) eine halbgeordnete Menge und $T \subseteq M$ eine Teilmenge. Ein Element $s \in M$ heißt **obere Schranke** von T , wenn gilt:

$$\forall t \in T : t \leq s$$

Das bedeutet: s ist größer oder gleich jedem Element aus T (bezüglich der Ordnung \leq).

Definition 3.19. Ein Element $m \in M$ einer halbgeordneten Menge (M, \leq) heißt **maximal in M** , wenn gilt:

$$\forall x \in M : m \leq x \Rightarrow x = m$$

Das heißt: es gibt kein Element $x \in M$, das echt größer als m ist. Es kann mehrere Maximale Elemente in M geben. Diese sind nicht vergleichbar, da sonst eines größer wäre als das andere.

Definition 3.20. Sei (M, \leq) eine halbgeordnete Menge. Eine Teilmenge $T \subseteq M$ heißt **Kette**, wenn sie totalgeordnet ist.

Eine Kette ist also eine totalgeordnete Teilmenge einer halbgeordneten Menge. Die Bezeichnung Kette röhrt daher, dass die Elemente davon nacheinander aufgereiht werden können: $t_1 \leq t_2 \leq \dots \leq t_n$ - also eine Kette bilden.

Das Lemma von Zorn ist ein fundamentales Hilfsmittel in der linearen Algebra und der Mengenlehre. Es sichert die Existenz maximaler Elemente in bestimmten halbgeordneten Mengen.

Satz 3.21 (Lemma von Zorn).

Voraussetzungen:

- (A, \leq) halbgeordnete Menge
- Jede Kette von A besitzt eine obere Schranke, die in A liegt (muss nicht in der Kette liegen).

Behauptung: Dann hat A maximale Elemente bzgl. " \leq ".

Wir verwenden nun das Lemma von Zorn zu zeigen, dass jeder Vektorraum maximal linear unabhängige Teilsysteme besitzt und zeigen anschließend mit [Satz 3.24](#), dass ein maximal linear unabhängiges Teilsystem eine Basis ist. Insgesamt also, dass jeder Vektorraum eine Basis besitzt.

Satz 3.22. Jeder Vektorraum besitzt maximal linear unabhängige Teilsysteme.

Beweis. Sei (V, Γ) ein Vektorraum. Wir betrachten die Menge \mathcal{L} aller Teilmengen von V , die linear unabhängig sind:

$$\mathcal{L} = \{L \subseteq V \mid L \text{ ist linear unabhängig}\}.$$

Auf \mathcal{L} definieren wir eine Ordnungsrelation durch Inklusion:

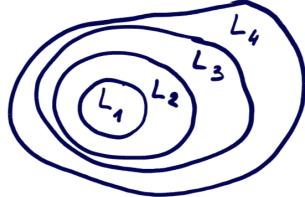
$$\forall L_1, L_2 \in \mathcal{L} : L_1 \leq L_2 \Leftrightarrow L_1 \subseteq L_2$$

(\mathcal{L}, \leq) ist eine halbgeordnete Menge, da die Inklusion reflexiv, antisymmetrisch und transitiv ist.

Um das Lemma von Zorn anwenden zu können, muss also noch gezeigt werden, dass jede Kette von \mathcal{L} eine obere Schranke besitzt, die in \mathcal{L} liegt.

Sei $\mathcal{K} = (L_\alpha)_{\alpha \in I} \subseteq \mathcal{L}$ eine Kette. (Also $L_1 \subseteq L_2 \subseteq L_3 \subseteq \dots$). \mathcal{K} ist eine Menge von Mengen. Dessen obere Schranke ist die Vereinigung:

$$L^* = \bigcup_{\alpha \in I} L_\alpha$$



Die Kette \mathcal{K}

Also muss noch gezeigt werden, dass $L^* \in \mathcal{L}$. Das heißt, es muss gezeigt werden, dass L^* linear unabhängig ist.

Sei $L^* = (v_\beta)_{\beta \in J}$. Für lineare Unabhängigkeit muss gezeigt werden:

$$\lambda^\beta v_\beta = 0 \Rightarrow \forall \beta \in J : \lambda^\beta = 0$$

Sei also $\lambda^\beta v_\beta = 0$, wobei nach der Definition der Linearkombination nur endlich viele $\lambda^\beta \neq 0$ sein können. Seien diese die $\lambda^{\beta_1}, \lambda^{\beta_2}, \dots, \lambda^{\beta_n}$. Betrachte die zugehörigen v_{β_i} .

Fall 1: Alle $v_{\beta_i}, i = 1, \dots, n$ liegen in einem L_α . Da $L_\alpha \in \mathcal{L}$ und \mathcal{L} aus linear unabhängigen Mengen besteht, ist also $v_{\beta_i}, i = 1, \dots, n$ linear unabhängig und somit $\lambda^{\beta_i} = 0$ für $i = 1, \dots, n$. Also $L^* = (v_\beta)_{\beta \in J}$ linear unabhängig.

Fall 2: Alle $v_{\beta_i}, i = 1, \dots, n$ liegen in unterschiedlichen L_{α_i} .

Da $(L_\alpha)_{\alpha \in I}$ eine Kette ist, sind die L_{α_i} paarweise durch Inklusion vergleichbar. Für $v_{\beta_1} \in L_{\alpha_1}$ und $v_{\beta_2} \in L_{\alpha_2}$ gilt also:

entweder $L_{\alpha_1} \subseteq L_{\alpha_2}$ und somit $v_{\beta_1}, v_{\beta_2} \in L_{\alpha_2}$

oder $L_{\alpha_2} \subseteq L_{\alpha_1}$ und somit $v_{\beta_1}, v_{\beta_2} \in L_{\alpha_1}$.

Also existiert ein Index, ab dem alle drin liegen:

$$\exists \alpha_N \in I \text{ mit } v_{\beta_1}, \dots, v_{\beta_n} \in L_{\alpha_N}$$

Wie in Fall 1 folgt nun, dass $L^* = (v_\beta)_{\beta \in J}$ linear unabhängig ist.

Nun kann das Lemma von Zorn angewandt werden: \mathcal{L} hat maximale Elemente bzgl. " \leq ".

Da \mathcal{L} aus linear unabhängigen Teilsystemen besteht, gibt es also maximal linear unabhängige Teilsysteme $M \subseteq V$. \square

Bemerkung 3.23. Neben der bereits genannten Definition für *Maximalität*, ist die folgende etwas intuitiver: eine l.u. Menge von Vektoren heißt *maximal*, wenn durch Hinzufügen eines weiteren Vektors die Menge linear abhängig wird.

Satz 3.24. Sei $M = (v_\alpha)_{\alpha \in I}$ eine maximale linear unabhängige Teilmenge des Vektorraums V . Dann lässt sich jeder Vektor $v \in V$ eindeutig als Linearkombination der Basisvektoren darstellen:

$$v = \lambda^\alpha v_\alpha$$

Das heißt: M ist ein linear unabhängiges Erzeugendensystem von V und somit eine Basis.

Beweis. Dazu unterscheiden wir zwei Fälle.

Fall 1: $v \in M$. Da M per Definition linear unabhängig ist, ist v nur über nur einen einzigen Vektor aus dem System $(v_\alpha)_{\alpha \in I}$ darstellbar: sich selbst.

$$v = v_{\alpha_0} = \lambda^{\alpha_0} v_{\alpha_0} + \lambda^{\alpha_i} v_{\alpha_i},$$

wobei $\lambda^{\alpha_0} = 1$ und alle übrigen $\lambda^{\alpha_i} = 0$ sind. Damit ist v also eindeutig als Linearkombination der Basisvektoren dargestellt.

Fall 2: $v \notin M$. Dann ist die Menge $\{v\} \cup M$ größer als M . Da sich M aber ein maximal linear unabhängiges Teilsystem von V ist, ist jede Vereinigung von M mit weiteren Elementen aus V nicht mehr linear unabhängig. Das bedeutet, dass eine nichttriviale Darstellung der Null existiert, die sowohl v als auch Elemente aus $v_\alpha \in M$ enthält. Also gilt:

$$\exists (\lambda^\alpha)_{\alpha \in I} \neq (0)_{\alpha \in I} : \lambda v + \lambda^\alpha v_\alpha = 0 \quad (3.2)$$

Nun unterscheiden wir wieder zwei Möglichkeiten:

- Wenn $\lambda = 0$, folgt mit (3.2), dass $\lambda^\alpha v_\alpha = 0$. Da $(v_\alpha)_{\alpha \in I}$ l.u. muss also $\lambda^\alpha = 0$ für alle $\alpha \in I$. Das steht jedoch im Widerspruch zur Annahme, dass $(\lambda^\alpha)_{\alpha \in I} \neq (0)_{\alpha \in I}$. Dieser Fall kann nicht eintreten.
- Wenn $\lambda \neq 0$, dann gilt:

$$\lambda v = -\lambda^\alpha v_\alpha \Rightarrow v = -\frac{\lambda^\alpha}{\lambda} v_\alpha =: \mu^\alpha v_\alpha$$

Damit ist v als Linearkombination der Vektoren $v_\alpha \in M$ darstellbar.

Mit Proposition 3.9 folgt die Eindeutigkeit. □

Bemerkung 3.25. Häufig wird die Basis mit $(E_\alpha)_{\alpha \in I}$ bezeichnet, und die Koeffizienten des Vektors v als dessen Komponenten:

$$v = v^\alpha E_\alpha$$

Diese Koeffizienten bezeichnet man als *Koordinaten* des Vektors bezüglich der Basis (E_α) .

Definition 3.26 (nicht Teil der VO). Die Mengen M und N heißen **gleichmächtig**, falls es eine Bijektion

$$f : M \rightarrow N$$

gibt.

Die **Mächtigkeit** oder **Kardinalität** einer Menge M ist die Äquivalenzklasse aller Mengen, die zu M gleichmächtig sind.

Wenn M endlich ist, dann ist existiert eine Bijektion zur Menge $\{1, 2, \dots, n\}$, wobei n die Anzahl der Elemente von M ist, und es ist $|M| = n$. Die Mächtigkeit entspricht also der Anzahl der Elemente der Menge.

Ist M nicht endlich, dann ist $|M| = \infty$.

Die Definition über die Bijektion ist notwendig, da für überabzählbar unendliche Mengen, die Elemente nicht gezählt werden können und man nicht von einer „Anzahl der Elemente der Menge“ sprechen kann.

Definition 3.27 (nicht Teil der VO). Sei (V, Γ) ein Vektorraum und \mathcal{B} eine Basis des Vektorraums.

Dann heißt

$$\dim_{\Gamma} V := |\mathcal{B}|$$

die **Dimension** des Vektorraums V über Γ .

Besitzt V keine endliche Basis, wird V **unendlich-dimensional** genannt.

Bemerkung 3.28 (nicht Teil der VO). Ein Vektorraum mit Dimension 0 besteht nur aus dem Nullpunkt, da jeder Vektorraum das Element 0 und sich selbst enthält. Also gilt insbesondere

$$\dim\{0\} = 0$$

- $\dim V = 1 \rightarrow$ Gerade durch den Ursprung
- $\dim V = 2 \rightarrow$ Ebene durch den Ursprung
- $\dim V = 3 \rightarrow$ unendlicher Würfel

In [Beispiel 3.13](#) lernten wir Basen von \mathbb{C} und \mathbb{R} kennen. Daher kennen wir nun ihre Dimension:

- $\dim_{\mathbb{R}} \mathbb{R}^2 = 2$
- $\dim_{\mathbb{R}} \mathbb{C} = 2$
- $\dim_{\mathbb{C}} \mathbb{C} = 1$
- $\dim_{\mathbb{C}} \mathbb{C}^2 = 2$

3.2 Dualraum

Definition 3.29. Seien V und W Vektorräume über dem Körper Γ . Eine Abbildung $\varphi : V \rightarrow W$ heißt **lineare Abbildung**, wenn für alle $v, w \in V$ und $\lambda \in \Gamma$ die folgenden Bedingungen erfüllt sind:

$$\varphi(v + w) = \varphi(v) + \varphi(w) \quad (\text{Additivität}) \quad (3.3)$$

$$\varphi(\lambda \cdot v) = \lambda \cdot \varphi(v) \quad (\text{Homogenität}) \quad (3.4)$$

Der **Raum der linearen Abbildungen** ist definiert als

$$L(V, W) = \{ \varphi : V \rightarrow W \mid \varphi \text{ ist linear} \}$$

Für zwei lineare Abbildungen $\varphi, \psi \in L(V, W)$ definieren wir außerdem Addition und Skalarmultiplikation durch:

$$+ : L(V, W) \times L(V, W) \rightarrow L(V, W) : (\varphi, \psi) \mapsto \varphi + \psi$$

$$\cdot : \Gamma \times L(V, W) \rightarrow L(V, W) : (\lambda, \varphi) \mapsto \lambda \cdot \varphi$$

Wir kennen eine Funktion, wenn wir ihre Wertetabelle kennen. Daher definieren wir, was die Abbildungen für $v \in V, \lambda \in \Gamma$ liefern:

$$(\varphi + \psi)(v) := \varphi(v) + \psi(v) \in W \quad (3.5)$$

$$(\lambda \cdot \varphi)(v) := \lambda \cdot \varphi(v) \in W \quad (3.6)$$

Definition 3.30. Der **Dualraum** V^\sim eines Vektorraums (V, Γ) ist definiert als die Menge aller linearen Abbildungen von V nach Γ :

$$V^\sim := L(V, \Gamma)$$

Jeder Vektorraum hat genau einen Dualraum.

Lemma 3.31. V^\sim ist ein Vektorraum über dem Körper Γ .

Beweis. Wir müssen nachweisen, dass die Rechengesetze eines Vektorraums erfüllt sind.

- Zunächst zeigen wir, dass $(\varphi + \psi)$ und $(\lambda \cdot \psi)$ linear sind, also $\varphi + \psi \in V^\sim$ und $\lambda \cdot \psi \in V^\sim$. Man sagt V^\sim ist *abgeschlossen bzgl. + und ·*.
 - Additivität für $(\varphi + \psi)$:
Für $v, w \in V$ gilt:

$$\begin{aligned} (\varphi + \psi)(v + w) &\stackrel{(3.5)}{=} \varphi(v + w) + \psi(v + w) \\ &\stackrel{(3.3)}{=} \varphi(v) + \varphi(w) + \psi(v) + \psi(w) \\ &= (\varphi(v) + \psi(v)) + (\varphi(w) + \psi(w)) \\ &\stackrel{(3.5)}{=} (\varphi + \psi)(v) + (\varphi + \psi)(w) \end{aligned}$$

- Homogenität für $(\varphi + \psi)$:
Ebenso gilt für $\lambda \in \Gamma$:

$$\begin{aligned} (\varphi + \psi)(\lambda v) &\stackrel{(3.6)}{=} \varphi(\lambda v) + \psi(\lambda v) \\ &\stackrel{(3.4)}{=} \lambda \varphi(v) + \lambda \psi(v) \\ &\stackrel{\substack{\text{Linksdistr.} \\ \text{ges. in } \Gamma}}{=} \lambda(\varphi(v) + \psi(v)) \\ &\stackrel{(3.6)}{=} \lambda(\varphi + \psi)(v) \end{aligned}$$

- Additivität und Homogenität für $(\lambda \cdot \psi)$ sind analog.

- Nun müssen die Vektorraumaxiome überprüft werden.
 - + bzgl. :

$$\begin{aligned} (\lambda \cdot (\varphi + \psi))(v) &\stackrel{(3.6)}{=} \lambda \cdot ((\varphi + \psi)(v)) \stackrel{(3.5)}{=} \lambda(\varphi(v) + \psi(v)) \\ &\stackrel{\substack{\text{Linksdistr.} \\ \text{ges. in } \Gamma}}{=} \lambda \varphi(v) + \lambda \psi(v) \stackrel{(3.6)}{=} (\lambda \cdot \varphi)(v) + (\lambda \cdot \psi)(v) \stackrel{(3.5)}{=} ((\lambda \cdot \varphi) + (\lambda \cdot \psi))(v) \end{aligned}$$

- Die restlichen Vektorraumaxiome für $(\varphi + \psi)$ und $(\lambda \cdot \psi)$ sind dem Leser überlassen :)

□

3.2.1 Basis des Dualraums

Sei (V, Γ) ein Vektorraum mit einer Basis $(E_\alpha)_{\alpha \in I}$. Für ein beliebiges $\varphi \in V^\sim$ definieren wir

$$\varphi_\alpha := \varphi(E_\alpha) \in \Gamma, \quad \alpha \in I$$

Satz 3.32. Sei (V, Γ) ein Vektorraum mit einer Basis $(E_\alpha)_{\alpha \in I}$. Für jede Familie $(\varphi_\alpha)_{\alpha \in I}$ von Skalaren aus Γ existiert genau eine lineare Abbildung

$$\varphi : V \rightarrow \Gamma$$

mit

$$\varphi(E_\alpha) = \varphi_\alpha \quad \forall \alpha \in I. \quad (3.7)$$

Das heißt $\varphi \in V^\sim$ wird eindeutig durch $(\varphi_\alpha)_{\alpha \in I}$ bestimmt.

Beweis.

- zz: φ wird durch φ_α bestimmt:

Für $v = v^\alpha E_\alpha$, mit $v^\alpha \in \Gamma$ gilt:

$$\varphi(v) = \varphi(v^\alpha E_\alpha) \stackrel{\varphi \text{ linear}}{=} v^\alpha \varphi(E_\alpha) \stackrel{(3.7)}{=} v^\alpha \varphi_\alpha.$$

- zz: φ ist eindeutig

Seien $\varphi, \psi \in V^\sim$ mit $\varphi(E_\alpha) = \psi(E_\alpha)$ für alle α , so folgt für jedes $v = v^\alpha E_\alpha \in V$

$$\varphi(v) = \varphi(v^\alpha E_\alpha) \stackrel{\varphi \text{ linear}}{=} v^\alpha \varphi(E_\alpha) = v^\alpha \psi(E_\alpha) \stackrel{\psi \text{ linear}}{=} \psi(v^\alpha E_\alpha) = \psi(v)$$

Also $\varphi = \psi$. D.h. Funktionen aus V^\sim , die auf einer Basis übereinstimmen, stimmen überall überein.

□

Wir können nun also zu vorgegebenen Zahlen $(\varphi_\alpha)_{\alpha \in I}$ auf eindeutige Weise eine zugehörige Funktion finden. Für die neutralen Elemente 0 und 1 des Vektorraums können wir damit eine Basis für den Dualraum konstruieren.

Satz-Definition 3.33. Sei (V, Γ) ein Vektorraum mit $\dim V < \infty$, einer Basis $(E_\alpha)_{\alpha \in I}$ und V^\sim sein Dualraum. Dann ist $(e^\alpha)_{\alpha \in I} \subseteq V^\sim$ mit

$$e^\alpha(E_\beta) = \delta_\beta^\alpha, \quad \alpha, \beta \in I \quad (3.8)$$

eine Basis von V^\sim , genannt **duale Basis**, wobei δ_β^α das Kronecker-Delta bezeichnet und eine Matrix ist:

$$\delta_\beta^\alpha = \begin{cases} 1, & \alpha = \beta \\ 0, & \alpha \neq \beta \end{cases}$$

Beweis.

- zz: Die Familie $(e^\alpha)_{\alpha \in I} \subseteq V^\sim$ ist linear unabhängig.

Also zz: $(\lambda_\alpha e^\alpha)(v) = 0 \Rightarrow \lambda_\alpha = 0 \ \forall \alpha \in I$

Nach [Satz 3.32](#) ist jede lineare Abbildung durch die Bilder der Vektoren einer Basis eindeutig bestimmt. Daher genügt es, die Auswertung von $e^\alpha \in V^\sim$ auf Basiselementen zu betrachten:

$$0 = (\lambda_\alpha e^\alpha)(E_\beta) \stackrel{(3.6)}{=} \lambda_\alpha e^\alpha(E_\beta) \stackrel{(3.8)}{=} \lambda_\alpha \delta_\beta^\alpha = \lambda_\beta \quad \forall \beta \in I$$

- zz: Jede Abbildung $\varphi \in V^\sim$ kann als Linearkombination von $(e^\alpha)_{\alpha \in I}$ dargestellt werden.

Also zz: $\forall v \in V : \varphi(v) = (\varphi_\alpha e^\alpha)(v)$

Wieder genügt es, die Auswertung von $\varphi \in V^\sim$ auf Basiselementen zu betrachten. Für $v = v^\alpha E_\alpha$ erhält man

$$\varphi(E_\alpha) = \varphi_\alpha = \varphi_\beta \delta_\alpha^\beta \stackrel{(3.8)}{=} \varphi_\beta e^\beta(E_\alpha) \stackrel{(3.6)}{=} (\varphi_\beta e^\beta)(E_\alpha)$$

Das letzte Gleichheitszeichen gilt nur, wenn es endlich viele φ_β gibt, da sonst die Definition der Linearkombination $\varphi_\beta e^\beta$ verletzt ist. Da $\varphi(E_\beta) = \varphi_\beta$ folgt somit, dass es dies nur der Fall ist, wenn es endlich viele E_α gibt, also $\dim V < \infty$ bzw. $|I| < \infty$.

□

Das heißt, sobald wir eine Basis in V wählen, erhalten wir direkt eine Basis in V^\sim dazu, wenn $\dim V < \infty$.

Bemerkung 3.34 (Nicht Teil der VO). Für jeden endlichdimensionalen Vektorraum V existieren zwar viele Vektorraum-Isomorphismen $V \cong V^\sim$ (linear+bijektive Abbildung), jedoch ist keiner davon *kanonisch*. Jeder davon ist basisabhängig oder benötigt eine zusätzliche Struktur (z.B. Skalarprodukt oder Bilinearform).

Beispiel eines basisabhängigen Isomorphismus:

Sei $(E_\alpha)_{\alpha \in I}$ eine Basis von V und $(e^\alpha)_{\alpha \in I}$ die dazu duale Basis, also $e^\alpha(E_\beta) = \delta_\beta^\alpha$, $\alpha, \beta \in I$.

Dann kann gezeigt werden, dass die Abbildung

$$\Phi : V \rightarrow V^\sim, \quad \Phi(v^\alpha E_\alpha) = v^\alpha e^\alpha$$

ein Isomorphismus ist. Wird eine andere Basis gewählt, so erhält man einen *anderen* Isomorphismus $V \rightarrow V^*$. Daher ist diese Abbildung nicht kanonisch.

Da $V \cong V^\sim$ wenn $\dim V < \infty$ ist somit $\dim V = \dim V^\sim$. Wenn $\dim V = \infty$ existieren keine bijektiv und linearen Abbildungen $V \rightarrow V^\sim$, nur injektive und lineare Abbildungen mit $\Phi(V) \subseteq V^\sim$. Man sagt V wird in V^\sim eingebettet. Dann ist $\dim V < \dim V^\sim$.

3.3 Bilineare Abbildungen

Wir betrachten wieder einen Vektorraum (V, Γ) . Der **Raum der bilinearen Abbildungen** von V nach Γ ist definiert als

$$\mathfrak{Bil}(V, \Gamma) := \{f : V \times V \rightarrow \Gamma \mid f \text{ ist linear in beiden Argumenten}\}.$$

Das heißt, eine Abbildung $f \in \mathfrak{Bil}(V, \Gamma)$, auch **Bilinearform** genannt, erfüllt für alle $v, v', w, w' \in V$ und $\lambda \in \Gamma$ die Bedingungen:

$$\begin{aligned} f(v + v', w) &= f(v, w) + f(v', w) \\ f(v, w + w') &= f(v, w) + f(v, w') \\ f(\lambda v, w) &= \lambda f(v, w) \\ f(v, \lambda w) &= \lambda f(v, w) \end{aligned}$$

Für zwei bilineare Abbildungen $f, g \in \mathfrak{Bil}(V, \Gamma)$ definieren wir ebenso Addition und Skalarmultiplikation durch:

$$\begin{aligned} + : \mathfrak{Bil}(V, \Gamma) \times \mathfrak{Bil}(V, \Gamma) &\rightarrow \mathfrak{Bil}(V, \Gamma) : (f, g) \mapsto f + g \\ \cdot : \Gamma \times \mathfrak{Bil}(V, \Gamma) &\rightarrow \mathfrak{Bil}(V, \Gamma) : (\lambda, f) \mapsto \lambda \cdot f \end{aligned}$$

Für $v, v' \in V, \lambda \in \Gamma$:

$$(f + g)(v, v') := f(v, v') + g(v, v') \in \Gamma \quad (3.9)$$

$$(\lambda \cdot f)(v, v') := \lambda \cdot f(v, v') \in \Gamma \quad (3.10)$$

Lemma 3.35. Der Raum $\mathfrak{Bil}(V, \Gamma)$ ist ein Vektorraum über Γ .

Beweis. Da wir in Lemma 3.31 bereits die Linearität der Addition und Skalarmultiplikation in einem Argument gezeigt haben, kann durch Festhalten je eines der beiden Argumente auf die Linearität in beiden Argumenten geschlossen werden.

Ebenso folgen anschließend wie zuvor die Vektorraumaxiome. \square

Definition 3.36. Seien V ein Vektorraum über Γ und V^\sim sein Dualraum. Für zwei Funktionen $\varphi, \psi \in V^\sim$ ist das **Tensorprodukt** definiert als

$$\begin{aligned} \otimes : V^\sim \times V^\sim &\rightarrow \mathfrak{Bil}(V, \Gamma) : (\varphi, \psi) \mapsto \varphi \otimes \psi \\ (\varphi \otimes \psi)(v, w) &:= \varphi(v) \cdot \psi(w) \in \Gamma, \quad v, w \in V \end{aligned} \quad (3.11)$$

Damit wird der **Tensorproduktraum zweiter Ordnung** in der VO definiert als

$$\bigotimes^2 V^\sim := V^\sim \otimes V^\sim := \left\{ f \in \mathfrak{Bil}(V, \Gamma) \mid f = \sum_{(\varphi, \psi) \in V^\sim \times V^\sim} \lambda^{(\varphi, \psi)} (\varphi \otimes \psi), \quad \lambda \in \Gamma, \text{ nur endlich viele } \lambda^{(\varphi, \psi)} \neq 0 \right\} \quad (3.12)$$

wobei $f \in \mathfrak{Bil}(V, \Gamma)$ wegen (3.11) und da die Verknüpfung von bilinearen Funktionen wieder bilinear ist.

$V^\sim \otimes V^\sim$ wird durch Addition und Skalarmultiplikation zum Vektorraum:

$$\left(\sum_{i,j} \lambda^{(i,j)} (\varphi_i \otimes \psi_j) \right) + \left(\sum_{l,k} \mu^{(k,l)} (\tilde{\varphi}_k \otimes \tilde{\psi}_l) \right) := \sum_{i,j} \lambda^{(i,j)} (\varphi_i \otimes \psi_j) + \sum_{l,k} \mu^{(k,l)} (\tilde{\varphi}_k \otimes \tilde{\psi}_l) \quad (\text{Addition}) \quad (3.13)$$

$$\mu \cdot \left(\sum_{i,j} \lambda^{(i,j)} (\varphi_i \otimes \psi_j) \right) := \sum_{i,j} (\mu \lambda^{(i,j)}) (\varphi_i \otimes \psi_j), \quad \mu \in \Gamma \quad (\text{Skalarmultiplikation}) \quad (3.14)$$

Mit der Definition der Vorlesung (3.12) folgt eine *mengentheoretische* Teilmengenbeziehung

$$V^\sim \otimes V^\sim \subseteq \mathfrak{Bil}(V, \Gamma) \quad (3.15)$$

Da $V^\sim \otimes V^\sim$ ein Vektorraum ist, kann auch gezeigt werden, dass er ein Untervektorraum von $\mathfrak{Bil}(V, \Gamma)$ ist.

In der Vorlesung wird das Tensorprodukt $\varphi \otimes \psi$ direkt als bilineare Funktion definiert. In der Literatur findet man auch die folgende Definition.

Quellen: [Multilinear Algebra by Alexander Duncan, 3. Kapitel Tensor Products, Definition 3.1.](#)

https://en.wikipedia.org/wiki/Tensor_product#Linearly_disjoint

Definition 3.37 (Nicht Teil der Vorlesung). Seien V ein Vektorraum über Γ und V^* sein Dualraum. Für zwei Funktionen $\varphi, \psi \in V^*$ ist das **Tensorprodukt** definiert als

$$\begin{aligned}\otimes : V^* \times V^* &\rightarrow V^* \otimes V^* : (\varphi, \psi) \mapsto \varphi \otimes \psi \\ (\varphi \otimes \psi)(v, w) &:= \varphi(v) \cdot \psi(w) \in \Gamma, \quad v, w \in V\end{aligned}$$

Damit wird der **Tensorproduktraum zweiter Ordnung** in der Literatur definiert als

$$\bigotimes^2 V^* := V^* \otimes V^* := \text{span} \left\{ \varphi \otimes \psi \mid \varphi, \psi \in V^* \right\} = \left\{ \sum_{i=1}^N \lambda^i (\varphi_i \otimes \psi_i) \mid N \in \mathbb{N}, \lambda^i \in \Gamma, \varphi_i, \psi_i \in V^* \right\}$$

Hier ist $\sum_{i=1}^N \lambda^i (\varphi_i \otimes \psi_i)$ eine *formale Summe*: die Ausdrücke $\varphi_i \otimes \psi_i$ werden zunächst *nicht* als Funktionen verstanden, sondern nur als Symbole. Man betrachtet sie also wie Bausteine eines neu konstruierten Vektorraums, auf denen Addition und Skalarmultiplikation wie gewohnt definiert sind. Erst durch die Abbildung

$$\Phi : V^* \otimes V^* \rightarrow \mathfrak{Bil}(V, \Gamma) : \varphi \otimes \psi \mapsto \varphi \cdot \psi \quad (3.16)$$

erhalten diese formalen Summen eine Bedeutung als *konkrete bilineare Abbildungen*.

Dann ist

$$\Phi(V^* \otimes V^*) = \left\{ f \in \mathfrak{Bil}(V, \Gamma) \mid f = \sum_{(\varphi, \psi) \in V^* \times V^*} \lambda^{(\varphi, \psi)} \Phi(\varphi \otimes \psi), \lambda \in \Gamma, \text{ nur endlich viele } \lambda^{(\varphi, \psi)} \neq 0 \right\}$$

Mit dieser Definition ist $\Phi(V^* \otimes V^*) \subseteq \mathfrak{Bil}(V, \Gamma)$.

Es kann gezeigt werden, dass Φ ein *kanonischer oder natürlicher Vektorraum-Isomorphismus* ist. Kanonisch heißt, der Isomorphismus ergibt sich allein aus der Struktur, die die beteiligten Räume schon von Natur aus haben. Er hängt nicht von der Wahl der Basis ab. Hier wird also die Definition (3.11) bzgl. dieses kanonischen Isomorphismus verstanden.

Ein *Vektorraum-Isomorphismus* ist definiert als lineare und bijektive Abbildung.

Das heißt:

- $(V^* \otimes V^*, +, \cdot)$ und $(\mathfrak{Bil}(V, \Gamma), +, \cdot)$ haben dieselbe **algebraische Struktur**: Addition und Skalarmultiplikation erfüllen dieselben Rechenregeln (Vektorraumaxiome) und werden, da Φ linear ist, strukturerhaltend abgebildet:

$\forall \varphi_1, \varphi_2, \psi_1, \psi_2 \in V^*$ und $\lambda \in \Gamma$ gilt:

$$\Phi((\varphi_1 \otimes \psi_1) + (\varphi_2 \otimes \psi_2)) = \Phi(\varphi_1 \otimes \psi_1) + \Phi(\varphi_2 \otimes \psi_2), \quad \Phi(\lambda \cdot (\varphi \otimes \psi)) = \lambda \cdot \Phi(\varphi \otimes \psi) \quad (3.17)$$

- $V^* \otimes V^*$ und $\mathfrak{Bil}(V, \Gamma)$ sind **gleichmächtig**, da Φ bijektiv ist: zu jedem $g \in \mathfrak{Bil}(V, \Gamma)$ kann genau ein $f \in V^* \otimes V^*$ gefunden werden, mit dem es durch $\Phi(f) = g$ identifiziert wird.
- Die beiden Räume sind also **nicht identisch**, verhalten sich jedoch gleich. Wenn $\dim V < \infty$ gilt $\Phi(V^* \otimes V^*) = \mathfrak{Bil}(V, \Gamma)$, siehe [Satz 3.39](#).

Satz 3.38 (mit [Definition 3.36](#)). Wenn der Vektorraum V endlich dimensional ist, ist der Tensorproduktraum bereits *mengentheoretisch* gleich dem Raum der bilinearen Abbildungen:

$$V^* \otimes V^* = \mathfrak{Bil}(V, \Gamma)$$

Beweis. \subseteq wurde bereits in (3.15) argumentiert. Daher zeigen wir nun \supseteq .

zz: $\forall f \in \mathfrak{Bil}(V, \Gamma) \Rightarrow f \in V^* \otimes V^*$

Sei $(E_\alpha)_{\alpha \in I}$ eine Basis des Vektorraums V und $f \in \mathfrak{Bil}(V, \Gamma)$. Wir definieren

$$f(E_\alpha, E_\beta) =: f_{\alpha\beta} \in \Gamma \quad (3.18)$$

Zeige: f wird durch $f_{\alpha\beta}$ eindeutig bestimmt.

Für beliebige Vektoren $v = v^\alpha E_\alpha$ und $w = w^\beta E_\beta$ gilt:

$$f(v, w) = f(v^\alpha E_\alpha, w^\beta E_\beta) \stackrel{f \text{ bilinear}}{=} v^\alpha w^\beta f(E_\alpha, E_\beta) \stackrel{(3.18)}{=} v^\alpha w^\beta f_{\alpha\beta}$$

Die Abbildung wird also in der Tat durch $f_{\alpha\beta}$ bestimmt. Die Beweisführung der Eindeutigkeit ist gleich wie die im Beweis von (3.32): zwei Funktionen, die auf einer Basis übereinstimmen, stimmen überall überein.

Kann (3.18) nun verwenden, um die Aussage zu zeigen:

Sei $(e^\alpha)_{\alpha \in I}$ die duale Basis. Dann gilt

$$\begin{aligned} f(E_\alpha, E_\beta) &\stackrel{(2.2)}{=} f_{\alpha\beta} = f_{\gamma\delta} \delta_\alpha^\gamma \delta_\beta^\delta \stackrel{\text{duale Basis}}{=} f_{\gamma\delta} e^\gamma(E_\alpha) e^\delta(E_\beta) \\ &\stackrel{(3.11)}{=} f_{\gamma\delta} (e^\gamma \otimes e^\delta)(E_\alpha, E_\beta) \stackrel{(3.14)}{=} (f_{\gamma\delta} e^\gamma \otimes e^\delta)(E_\alpha, E_\beta) \end{aligned}$$

Damit das letzte Gleichheitszeichen gilt, also die Definition der Linearkombination nicht verletzt wird, muss wieder $(E_\alpha)_{\alpha \in I}$ endlich sein, also $\dim V < \infty$.

Damit stimmen f und $f_{\gamma\delta} e^\gamma \otimes e^\delta$ auf einer Basis überein und stimmen somit überall überein. Also ist gezeigt, dass jede bilineare Abbildung f als endliche Summe

$$f = f_{\gamma\delta} e^\gamma \otimes e^\delta \in V^* \otimes V^*$$

geschrieben werden kann. Daraus folgt

$$\mathfrak{Bil}(V, \Gamma) \subseteq V^* \otimes V^*$$

□

Satz 3.39 (nicht Teil der Vorlesung mit Definition 3.37). Wenn der Vektorraum V endlich dimensional ist, ist der Tensorproduktraum isomorph zum Raum der bilinearen Abbildungen:

$$V^* \otimes V^* \cong \mathfrak{Bil}(V, \Gamma)$$

Das heißt die Abbildung (3.16) ist ein Vektorraum-Isomorphismus.

Wenn der Vektorraum V unendlich dimensional ist, ist Φ nicht surjektiv. Es gilt also nur $\Phi(V^* \otimes V^*) \subseteq \mathfrak{Bil}(V, \Gamma)$. Man sagt, $V^* \otimes V^*$ wird in $\mathfrak{Bil}(V, \Gamma)$ eingebettet.

Quelle: [Multilinear Algebra by Alexander Duncan, 3. Kapitel Tensor Products, Corollary 3.8](#).

Lemma 3.40. Sei $(E_\alpha)_{\alpha \in I}$ eine Basis von V und $(e^\alpha)_{\alpha \in I}$ die zugehörige duale Basis. Dann bildet die Menge $\{e^\alpha \otimes e^\beta\}$ eine Basis von $\mathfrak{Bil}(V, \Gamma)$.

Beweis.

- Wir haben im Beweis von [Satz 3.38](#) bereits gezeigt, dass jede bilineare Abbildung f als Linearkombination

$$f = f_{\alpha\beta} e^\alpha \otimes e^\beta$$

dargestellt werden kann. Daher ist $\{e^\alpha \otimes e^\beta\}$ ein Erzeugendensystem.

- zz: Die Familie $\{e^\alpha \otimes e^\beta\} \subseteq \mathfrak{Bil}(V, \Gamma)$ ist linear unabhängig.

Also zz: $\forall v, w \in V : (\lambda_{\alpha\beta} e^\alpha \otimes e^\beta)(v, w) = 0 \Rightarrow \lambda_{\alpha\beta} = 0 \ \forall \alpha, \beta \in I$

Nach [Satz 3.32](#) ist jede lineare Abbildung durch die Bilder der Vektoren einer Basis eindeutig bestimmt. Daher genügt es, die Auswertung von $e^\alpha \otimes e^\beta \in \mathfrak{Bil}(V, \Gamma)$ auf Basiselementen zu betrachten:

$$0 = (\lambda_{\alpha\beta} e^\alpha \otimes e^\beta)(E_\gamma, E_\delta) \stackrel{(3.36)}{=} \lambda_{\alpha\beta} e^\alpha(E_\gamma) \cdot e^\beta(E_\delta) \stackrel{(3.8)}{=} \lambda_{\alpha\beta} \delta_\gamma^\alpha \delta_\delta^\beta = \lambda_{\gamma\delta} \quad \forall \gamma, \delta \in I$$

□

3.4 Multilineare Abbildungen

Analog zum bilinearen Fall betrachten wir Abbildungen, die in mehreren Argumenten linear sind.

Definition 3.41. Der Raum der **p-multilinearen Abbildungen** ist gegeben durch

$$\mathfrak{Mult}_p(V, \Gamma) := \{f : \underbrace{V \times \cdots \times V}_{p \text{ mal}} \rightarrow \Gamma \mid f \text{ ist linear in jedem Argument}\}$$

Wir definieren wir Addition und Skalarmultiplikation durch

$$\begin{aligned} + &: \mathfrak{Mult}_p(V, \Gamma) \times \mathfrak{Mult}_p(V, \Gamma) \rightarrow \mathfrak{Mult}_p(V, \Gamma) : (f, g) \mapsto f + g \\ \cdot &: \Gamma \times \mathfrak{Mult}_p(V, \Gamma) \rightarrow \mathfrak{Mult}_p(V, \Gamma) : (\lambda, f) \mapsto \lambda \cdot f \end{aligned}$$

Für $v_i \in V, \lambda \in \Gamma$:

$$(f + g)(v_1, \dots, v_p) := f(v_1, \dots, v_p) + g(v_1, \dots, v_p) \in \Gamma \tag{3.19}$$

$$(\lambda \cdot f)(v_1, \dots, v_p) := \lambda \cdot f(v_1, \dots, v_p) \in \Gamma \tag{3.20}$$

womit durch Festhalten aller Argumente, bis auf eines und der Linearität in einem Argument gezeigt werden kann, dass $\mathfrak{Mult}_p(V, \Gamma)$ wieder ein Vektorraum über Γ ist.

Definition 3.42. Für $\varphi_1, \dots, \varphi_p \in V^\sim$ definieren wir das **p-fache Tensorprodukt**

$$\otimes : \underbrace{V^\sim \times \cdots \times V^\sim}_{p \text{ mal}} \rightarrow \mathfrak{Mult}_p(V, \Gamma) : (\varphi_1, \dots, \varphi_p) \mapsto \varphi_1 \otimes \cdots \otimes \varphi_p \tag{3.21}$$

Für $v_1, \dots, v_p \in V$ ist

$$(\varphi_1 \otimes \cdots \otimes \varphi_p)(v_1, \dots, v_p) := \varphi_1(v_1) \cdot \cdots \cdot \varphi_p(v_p) \in \Gamma \tag{3.22}$$

Damit wird der **Tensorproduktraum p-ter Ordnung** in der VO definiert als

$$\begin{aligned} \bigotimes^p V^\sim &:= \underbrace{V^\sim \otimes \cdots \otimes V^\sim}_{p \text{ mal}} \\ &:= \left\{ f \in \mathfrak{Mult}_p(V, \Gamma) \mid f = \sum_{(\varphi_1, \dots, \varphi_p) \in \underbrace{V^\sim \times \cdots \times V^\sim}_{p \text{ mal}}} \lambda^{(\varphi_1, \dots, \varphi_p)} (\varphi_1 \otimes \cdots \otimes \varphi_p), \lambda \in \Gamma, \text{ nur endl. viele } \lambda^{(\varphi_1, \dots, \varphi_p)} \neq 0 \right\} \end{aligned}$$

wobei $f \in \mathfrak{Mult}_p(V, \Gamma)$ da per [Definition 3.42](#) $\varphi_1 \otimes \cdots \otimes \varphi_p \in \mathfrak{Mult}_p(V, \Gamma)$ und da die Verknüpfung von multilinear Funktionen wieder multilinear ist.

Damit gilt wieder insbesondere (mengentheoretisch)

$$\bigotimes^p V^\sim \subseteq \mathfrak{Mult}_p(V, \Gamma) \quad (3.23)$$

$\mathfrak{Mult}_p(V, \Gamma)$ wird mit der Addition und Skalarmultiplikation zum Vektorraum. Es kann wieder gezeigt werden, dass $\bigotimes^p V^\sim$ dann ein Untervektorraum von $\mathfrak{Mult}_p(V, \Gamma)$ ist.

Bemerkung 3.43 (Nicht Teil der VO). Wieder kann obige Definition in der Literatur mit

$$\otimes : \underbrace{V^\sim \times \cdots \times V^\sim}_{p \text{ mal}} \rightarrow \underbrace{V^\sim \otimes \cdots \otimes V^\sim}_{p \text{ mal}} : (\varphi_1, \dots, \varphi_p) \mapsto \varphi_1 \otimes \cdots \otimes \varphi_p \quad (3.24)$$

und bzgl. des kanonischen Isomorphismus

$$\begin{aligned} \Phi : \bigotimes^p V^\sim &\rightarrow \mathfrak{Mult}_p(V, \Gamma) : \varphi_1 \otimes \cdots \otimes \varphi_p \mapsto \varphi_1 \cdot \dots \cdot \varphi_p \\ \text{mit } v_1, \dots, v_p \in V : \Phi(\varphi_1 \otimes \cdots \otimes \varphi_p)(v_1, \dots, v_p) &:= \varphi_1(v_1) \cdot \dots \cdot \varphi_p(v_p) \end{aligned} \quad (3.25)$$

gefunden werden.

Satz 3.44. Wenn der Vektorraum V endlich dimensional ist, ist der Tensorproduktraum p-ter Ordnung mengentheoretisch gleich dem Raum der p-multilinear Abbildungen:

$$\bigotimes^p V^\sim = \mathfrak{Mult}_p(V, \Gamma)$$

Beweis. \subseteq wurde bereits in (3.23) argumentiert. Daher zeigen wir nun \supseteq .

zz: $\forall f \in \mathfrak{Mult}_p(V, \Gamma) \Rightarrow f \in \bigotimes^p V^\sim$

Sei $(E_\alpha)_{\alpha \in I}$ eine Basis von V und $(e^\alpha)_{\alpha \in I}$ die duale Basis. Dann ist für $f \in \mathfrak{Mult}_p(V, \Gamma)$

$$f(E_{\alpha_1}, \dots, E_{\alpha_p}) =: f_{\alpha_1 \dots \alpha_p} \in \Gamma. \quad (3.26)$$

Der Beweis, dass f durch diese Darstellung eindeutig bestimmt ist, ist analog zu dem Abschnitt im Beweis von [Satz 3.39](#).

Sei nun $f \in \mathfrak{Mult}_p(V, \Gamma)$ und $(E_\alpha)_{\alpha \in I}$ eine Basis von V mit dualer Basis $(e^\alpha)_{\alpha \in I}$. Dann gilt analog zum

bilinearen Fall:

$$\begin{aligned}
 f(E_{\alpha_1}, \dots, E_{\alpha_p}) &\stackrel{(3.26)}{=} f_{\alpha_1 \dots \alpha_p} = f_{\beta_1 \dots \beta_p} \delta_{\alpha_1}^{\beta_1} \dots \delta_{\alpha_p}^{\beta_p} \\
 &\stackrel{\text{duale Basis}}{=} f_{\beta_1 \dots \beta_p} e^{\beta_1}(E_{\alpha_1}) \dots e^{\beta_p}(E_{\alpha_p}) \\
 &\stackrel{(3.22)}{=} f_{\beta_1 \dots \beta_p} (e^{\beta_1} \otimes \dots \otimes e^{\beta_p})(E_{\alpha_1}, \dots, E_{\alpha_p}) \\
 &\stackrel{\text{Skalar-mult.}}{=} (f_{\beta_1 \dots \beta_p} e^{\beta_1} \otimes \dots \otimes e^{\beta_p})(E_{\alpha_1}, \dots, E_{\alpha_p})
 \end{aligned}$$

Damit das letzte Gleichheitszeichen gilt, also die Definition der Linearkombination nicht verletzt wird, muss wieder $(E_\alpha)_{\alpha \in I}$ endlich sein, also $\dim V < \infty$.

Damit stimmen f und $f_{\beta_1 \dots \beta_p} e^{\beta_1} \otimes \dots \otimes e^{\beta_p}$ auf einer Basis überein und stimmen somit überall überein. Also ist gezeigt, dass jede multilinear Abbildung f als endliche Summe

$$f = f_{\beta_1 \dots \beta_p} e^{\beta_1} \otimes \dots \otimes e^{\beta_p} \in \bigotimes^p V^\sim$$

dargestellt werden kann. □

Satz 3.45 (Nicht Teil der VO mit [Bemerkung 3.43](#)). Wenn der Vektorraum V endlich dimensional ist, ist der Tensorproduktraum p -ter Ordnung isomorph zum Raum der p -multilinear Abbildungen:

$$\bigotimes^p V^\sim \cong \mathfrak{Mult}_p(V, \Gamma)$$

Das heißt die Abbildung Φ von (3.25) ist ein Vektorraum-Isomorphismus.

Wenn der Vektorraum V unendlich dimensional ist, ist Φ nicht surjektiv. Es gilt also nur $\Phi(\bigotimes^p V^\sim) \subseteq \mathfrak{Mult}_p(V, \Gamma)$. Man sagt, $\bigotimes^p V^\sim$ wird in $\mathfrak{Mult}_p(V, \Gamma)$ eingebettet.

Lemma 3.46. Sei $(E_\alpha)_{\alpha \in I}$ eine Basis von V und $(e^\alpha)_{\alpha \in I}$ die zugehörige duale Basis. Dann bildet die Menge $\{e^{\alpha_1} \otimes \dots \otimes e^{\alpha_p}\}$ eine Basis von $\mathfrak{Mult}_p(V, \Gamma)$.

Beweis.

- Wir haben im Beweis von [Satz 3.44](#) bereits gezeigt, dass jede multilinear Abbildung f als Linearkombination

$$f = f_{\alpha_1 \dots \alpha_p} e^{\alpha_1} \otimes \dots \otimes e^{\alpha_p}$$

dargestellt werden kann. Daher ist $\{e^{\alpha_1} \otimes \dots \otimes e^{\alpha_p}\}$ ein Erzeugendensystem.

- zz: Die Familie $\{e^{\alpha_1} \otimes \dots \otimes e^{\alpha_p}\} \subseteq \mathfrak{Mult}_p(V, \Gamma)$ ist linear unabhängig.

Also zz: $\forall v_1, \dots, v_p \in V : (\lambda_{\alpha_1 \dots \alpha_p} e^{\alpha_1} \otimes \dots \otimes e^{\alpha_p})(v_1, \dots, v_p) = 0 \Rightarrow \lambda_{\alpha_1 \dots \alpha_p} = 0 \ \forall \alpha_1, \dots, \alpha_p \in I$
Nach [Satz 3.32](#) ist jede lineare Abbildung durch die Bilder der Vektoren einer Basis eindeutig bestimmt.
Daher genügt es, die Auswertung von $e^{\alpha_1} \otimes \dots \otimes e^{\alpha_p} \in \mathfrak{Mult}_p(V, \Gamma)$ auf Basiselementen zu betrachten:

$$\begin{aligned}
 0 &= (\lambda_{\alpha_1 \dots \alpha_p} e^{\alpha_1} \otimes \dots \otimes e^{\alpha_p})(E_{\beta_1}, \dots, E_{\beta_p}) \stackrel{(3.36)}{=} \lambda_{\alpha_1 \dots \alpha_p} e^{\alpha_1}(E_{\beta_1}) \cdot \dots \cdot e^{\alpha_p}(E_{\beta_p}) \\
 &\stackrel{(3.8)}{=} \lambda_{\alpha_1 \dots \alpha_p} \delta_{\beta_1}^{\alpha_1} \cdot \dots \cdot \delta_{\beta_p}^{\alpha_p} = \lambda_{\beta_1 \dots \beta_p} \quad \forall \beta_1, \dots, \beta_p \in I
 \end{aligned}$$

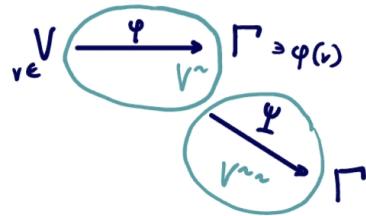
□

3.5 Doppeldualraum

Jeder Vektorraum hat einen Dualraum. Da der Dualraum V^\sim ein Vektorraum ist, hat auch dieser einen Dualraum.

Definition 3.47. Der **Doppeldualraum** (auch **Bidualraum**) eines Vektorraums (V, Γ) ist der Dualraum des Dualraums:

$$V^{\sim\sim} := (V^\sim)^\sim = \{\psi : V^\sim \rightarrow \Gamma : \psi \text{ ist linear}\}$$



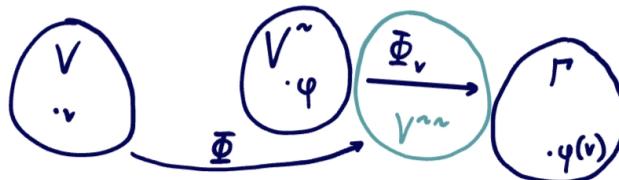
Wie hängen nun V und $V^{\sim\sim}$ zusammen?

Es gibt eine kanonische Einbettung

$$\Phi : V \rightarrow V^{\sim\sim} : v \mapsto \Phi_v$$

wobei das Punktauswertungsfunktional Φ_v folgendermaßen agiert

$$\Phi_v : V^\sim \rightarrow \Gamma : \varphi \mapsto \Phi_v(\varphi) := \varphi(v) \quad (3.27)$$



Lemma 3.48. Φ ist linear.

Beweis.

- zz: Φ ist wohldefiniert. D.h. $\Phi_v \in V^{\sim\sim}$.

Wir müssen zeigen, dass Φ_v linear ist.

Für $\varphi_1, \varphi_2 \in V^\sim$ und $\lambda \in \Gamma$ gilt:

$$\begin{aligned} \Phi_v(\varphi_1 + \varphi_2) &\stackrel{(3.27)}{=} (\varphi_1 + \varphi_2)(v) \stackrel{\varphi_1, \varphi_2 \text{ linear}}{=} \varphi_1(v) + \varphi_2(v) \stackrel{(3.27)}{=} \Phi_v(\varphi_1) + \Phi_v(\varphi_2) \\ \Phi_v(\lambda \cdot \varphi) &\stackrel{(3.27)}{=} (\lambda \cdot \varphi)(v) \stackrel{\varphi \text{ linear}}{=} \lambda \cdot \varphi(v) \stackrel{(3.27)}{=} \lambda \cdot \Phi_v(\varphi) \end{aligned}$$

- zz: Φ ist linear

Für $v, v_1, v_2 \in V$ und $\lambda \in \Gamma$ gilt:

$$\begin{aligned}\Phi_{v_1+v_2}(\varphi) &\stackrel{(3.27)}{=} \varphi(v_1 + v_2) \stackrel{\varphi \text{ linear}}{=} \varphi(v_1) + \varphi(v_2) \stackrel{(3.27)}{=} \Phi_{v_1}(\varphi) + \Phi_{v_2}(\varphi) \stackrel{\text{Addition in } V^{\sim\sim}}{=} (\Phi_{v_1} + \Phi_{v_2})(\varphi) \\ \Phi_{\lambda \cdot v}(\varphi) &\stackrel{(3.27)}{=} \varphi(\lambda \cdot v) \stackrel{\varphi \text{ linear}}{=} \lambda \cdot \varphi(v) \stackrel{(3.27)}{=} \lambda \cdot \Phi_v(\varphi) \stackrel{\text{Skalarmult. in } V^{\sim\sim}}{=} (\lambda \cdot \Phi_v)(\varphi)\end{aligned}$$

Da diese Gleichungen für alle $\varphi \in V^{\sim\sim}$ gelten, folgt:

$$\Phi_{v_1+v_2} = \Phi_{v_1} + \Phi_{v_2}, \quad \Phi_{\lambda \cdot v} = \lambda \cdot \Phi_v \quad (3.28)$$

□

Bemerkung 3.49. Da Φ linear (strukturerhaltend) ist, macht es also keinen Unterschied, ob ich zuerst in V addiere mit $v_1 + v_2$ und dann $\Phi_{v_1+v_2}$ auf φ anwende oder ob ich zuerst in $V^{\sim\sim}$ addiere mit $\Phi_{v_1} + \Phi_{v_2}$ und dann $(\Phi_{v_1} + \Phi_{v_2})$ auf φ anwende. Selbiges gilt für das Skalarprodukt. Daher müssen wir nicht mehr die Addition bzw. die Skalarmultiplikation in V von der Addition bzw. die Skalarmultiplikation in $V^{\sim\sim}$ unterscheiden. Somit agieren Elemente aus V gleich wie Elemente aus $V^{\sim\sim}$. Also können wir $v \in V$ mit $\Phi_v \in V^{\sim\sim}$ identifizieren. Wir definieren daher

$$v(\varphi) := \Phi_v(\varphi) = \varphi(v) \in \Gamma \quad (3.29)$$

Mit dieser Definition folgt dann die mengentheoretische Teilmengenbeziehung

$$V \subseteq V^{\sim\sim}$$

Da $V^{\sim\sim}$ ein Vektorraum ist, kann auch gezeigt werden, dass V mit dieser Definition ein Untervektorraum von $V^{\sim\sim}$ ist.

Bemerkung 3.50 (Nicht Teil der VO). Wenn obige Definition nicht verwendet wird, kann gezeigt werden, dass Φ ein kanonischer Isomorphismus ist und somit $V \cong V^{\sim\sim}$. Damit gibt es für jedes $\Phi_v \in V^{\sim\sim}$ genau ein $v \in V$, sodass $\Phi(v) = \Phi_v$ und für $\varphi \in V^{\sim\sim}$: $\Phi(v)(\varphi) = \Phi_v(\varphi)$. Da die Struktur durch die Linearität von Φ auch erhalten bleibt und für Φ keine zusätzliche Struktur erfordert bzw. von der Wahl der Basis unabhängig ist, wird in der Vorlesung direkt obige Definition gewählt.

Da $V \cong V^{\sim\sim}$ wenn $\dim V < \infty$ ist somit $\dim V = \dim V^{\sim\sim}$. Mit [Bemerkung 3.34](#) folgt dann $\dim V^{\sim} = \dim V^{\sim\sim}$. Wenn $\dim V = \infty$ ist $\dim V < \dim V^{\sim\sim}$.

Satz 3.51. Wenn der Vektorraum V endlich dimensional ist, ist er bereits *mengentheoretisch* gleich seinem Bidualraum:

$$V = V^{\sim\sim}$$

Beweis. \supseteq haben wir zuvor schon argumentiert. Wir zeigen nun \subseteq .

zz: $\Psi \in V^{\sim\sim} \Rightarrow \Psi \in V$

Sei also $\Psi \in V^{\sim\sim}$, $(E_\alpha)_{\alpha \in I}$ eine Basis von (V, Γ) und $(e^\alpha)_{\alpha \in I}$ die dazu duale Basis des Dualraums (V^{\sim}, Γ) .

Dann ist

$$\Psi(e^\alpha) =: \underbrace{\Psi^\alpha}_{\in \Gamma} = \Psi^\beta \delta_\beta^\alpha \stackrel{\text{kanon. Basis}}{=} \Psi^\beta e^\alpha(E_\beta) \stackrel{(3.29)}{=} \Psi^\beta E_\beta(e^\alpha) \stackrel{\text{Skalar. mult.}}{=} (\Psi^\beta E_\beta)(e^\alpha) \quad (3.30)$$

Wieder ist das letzte Gleichheitszeichen nur für endlich viele $(E_\alpha)_{\alpha \in I}$ zulässig. Zwei lineare Funktionale, die auf allen Basisvektoren übereinstimmen, stimmen auch auf allen Vektoren überein. Daher ist:

$$\Psi = \Psi^\beta E_\beta \in V \quad (3.31)$$

□

Definition 3.52. Der Raum der **p-multilinearen Abbildungen auf dem Dualraum** ist definiert als

$$\mathfrak{Mult}_p(V^\sim, \Gamma) := \{f : \underbrace{V^\sim \times \cdots \times V^\sim}_{p\text{-mal}} \rightarrow \Gamma \mid f \text{ ist linear in jedem Argument}\}$$

In vollkommener Analogie zu $\mathfrak{Mult}_p(V, \Gamma)$ gilt im endlichdimensionalen Fall:

$$\mathfrak{Mult}_p(V^\sim, \Gamma) = \bigotimes^p V^{\sim\sim} = \bigotimes^p V$$

und

$$(v_1 \otimes \cdots \otimes v_p)(\varphi_1, \dots, \varphi_p) = v_1(\varphi_1) \cdot \dots \cdot v_p(\varphi_p) = \varphi_1(v_1) \cdot \dots \cdot \varphi_p(v_p) \in \Gamma$$

3.6 Gemischte multilinear Abbildungen

In den bisherigen Abschnitten haben wir multivariate Abbildungen betrachtet, deren Argumente entweder nur aus dem Vektorraum V oder nur aus dem Dualraum V^\sim stammen. Für die Tensorrechnung in der Relativitätstheorie benötigen wir jedoch allgemeinere Abbildungen, die sowohl Vektoren aus V als auch Linearformen aus V^\sim als Argumente aufnehmen und in jedem Argument linear sind.

Definition 3.53. Sei (V, Γ) ein Vektorraum über dem Körper Γ . Der **Raum der gemischten p,q-multilinear Abbildungen** ist definiert als

$$\mathfrak{Mult}_{p,q}(V, V^\sim, \Gamma) := \{t : \underbrace{V \times \cdots \times V}_{p\text{-mal}} \times \underbrace{V^\sim \times \cdots \times V^\sim}_{q\text{-mal}} \rightarrow \Gamma \mid t \text{ linear in jedem Argument}\} \quad (3.32)$$

Ein Element $t \in \mathfrak{Mult}_{p,q}(V, \Gamma)$ bildet somit $v_1, \dots, v_p \in V$, $\varphi_1, \dots, \varphi_q \in V^\sim$ folgendermaßen ab:

$$(v_1, \dots, v_p, \varphi_1, \dots, \varphi_q) \mapsto t(v_1, \dots, v_p, \varphi_1, \dots, \varphi_q) \in \Gamma$$

Es kann wieder gezeigt werden, dass $\mathfrak{Mult}_{p,q}(V, V^\sim, \Gamma)$ mit komponentenweiser Addition und Skalarmultiplikation ein Vektorraum über Γ ist.

Sei (V, Γ) ein endlichdimensionaler Vektorraum mit Dualraum V^\sim . Wir haben zuvor gezeigt, dass dann

$$\mathfrak{Mult}_p(V, \Gamma) = \bigotimes^p V^\sim \quad \text{und} \quad \mathfrak{Mult}_q(V^\sim, \Gamma) = \bigotimes^q V^{\sim\sim} \stackrel{V=\Gamma}{=} \bigotimes^q V$$

Damit folgt

$$\mathfrak{Mult}_{p,q}(V, V^\sim, \Gamma) = \bigotimes^p V^\sim \otimes \bigotimes^q V =: \bigotimes_p^q V$$

Elemente $t \in \bigotimes_p^q V$ heißen **(q,p)-Tensoren** oder **Tensoren vom Typ (q,p)**. Also sind Tensoren nach dieser Definition (q,p) -multilineare Funktionen.

Für

$$\mathfrak{Mult}_{p,q}(V^\sim, V, \Gamma) := \{t : \underbrace{V^\sim \times \cdots \times V^\sim}_{p\text{-mal}} \times \underbrace{V \times \cdots \times V}_{q\text{-mal}} \rightarrow \Gamma \mid t \text{ linear in jedem Argument}\} \quad (3.33)$$

gilt auch

$$\mathfrak{Mult}_{p,q}(V^*, V, \Gamma) = \bigotimes_q^p V \otimes \bigotimes^q V^* = \bigotimes_q^p V \quad (3.34)$$

Entsprechend heißen Elemente $t \in \bigotimes_q^p V$ **(p,q)-Tensoren** oder **Tensoren vom Typ (p,q)**. Diese Definition werden wir ab jetzt verwenden.

Bemerkung 3.54 (Nicht Teil der VO). In der Literatur kann auch die folgende Definition für **(p,q)-Tensoren** gefunden werden.

Quelle: [https://en.wikipedia.org/wiki/Tensor_\(intrinsic_definition\)#Definition_via_tensor_products_of_vector_spaces](https://en.wikipedia.org/wiki/Tensor_(intrinsic_definition)#Definition_via_tensor_products_of_vector_spaces)

$$T_q^p(V) := \underbrace{V \otimes \cdots \otimes V}_{p\text{-mal}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q\text{-mal}} = \bigotimes^p V \otimes \bigotimes^q V^* = \mathfrak{Mult}_{p,q}(V^*, V, \Gamma) \quad (3.35)$$

Da für endlichdimensionale Vektorräume V, W gilt: $V \otimes W \cong_{\text{kanonisch}} W \otimes V$, stimmen dann die beiden Sichtweisen überein:

$$\mathfrak{Mult}_{p,q}(V^*, V, \Gamma) = \bigotimes^p V \otimes \bigotimes^q V^* \cong_{\text{kanonisch}} \bigotimes^q V^* \otimes \bigotimes^p V = \mathfrak{Mult}_{q,p}(V, V^*, \Gamma)$$

Bemerkung 3.55. Für $t \in \mathfrak{Mult}_{p,q}(V^*, V, \Gamma) = \bigotimes^p V \otimes \bigotimes^q V^* = \bigotimes_q^p V$ gilt:

$$\begin{aligned} t(\underbrace{\varphi_1, \dots, \varphi_p}_{\in V^*}, \underbrace{v_1, \dots, v_q}_{\in V}) &\in \Gamma \\ t = t_{b_1 \dots b_q}^{a_1 \dots a_p} \underbrace{E_{a_1} \otimes \cdots \otimes E_{a_p}}_{\in V} \otimes \underbrace{e^{b_1} \otimes \cdots \otimes e^{b_q}}_{\in V^*} \end{aligned}$$

- In der Darstellung $t(\varphi_1, \dots, \varphi_p, v_1, \dots, v_q)$ nimmt t also p **Argumente** aus dem Dualraum und q **Argumente** aus dem Vektorraum.
- In der Darstellung $t = t_{b_1 \dots b_q}^{a_1 \dots a_p} E_{a_1} \otimes \cdots \otimes E_{a_p} \otimes e^{b_1} \otimes \cdots \otimes e^{b_q}$ besteht t aus p **Faktoren** aus dem Vektorraum und q **Faktoren** aus dem Dualraum.
- p wird **kontravarianter Grad** genannt (Anzahl der Faktoren aus dem Vektorraum V). Vektoren $\in V$ tragen obere Indizes: $v = v^\alpha E_\alpha$.
- q wird **kovarianter Grad** genannt (Anzahl der Faktoren aus dem Dualraum V^*). Kovektoren $\in V^*$ tragen untere Indizes: $\varphi = \varphi_\alpha e^\alpha$.

Beispiel 3.56 (Nicht Teil der VO).

Quelle: <https://en.wikipedia.org/wiki/Tensor#Examples>

Je nach Definition sind untenstehende Gleichheitszeichen mengentheoretisch oder bzgl. eines kanonischen Isomorphismus zu verstehen.

- $\bigotimes_0^0 V = \Gamma$... Tensor 0. Stufe (Skalar)
- $\bigotimes_0^1 V = V$... Tensor 1. Stufe (Vektor)
- $\bigotimes_1^0 V = V^*$... Tensor 1. Stufe (Kovektor/lineares Funktional/Linearform)

- $\bigotimes_1^1 V = V \otimes V^\sim \cong V^\sim \otimes V = L(V, V)$... Tensor 2. Stufe (Matrix)
- $\bigotimes_0^2 V = V \otimes V = \mathfrak{Bil}(V^\sim, \Gamma) = \mathfrak{Mult}_{2,0}(V^\sim, V, \Gamma)$... Tensor 2. Stufe (Matrix)
- $\bigotimes_2^0 V = V^\sim \otimes V^\sim = \mathfrak{Bil}(V, \Gamma) = \mathfrak{Mult}_{0,2}(V^\sim, V, \Gamma)$... Tensor 2. Stufe (Matrix)
- $\bigotimes_1^2 V = V \otimes V \otimes V^\sim \cong V^\sim \otimes V \otimes V = \mathfrak{Mult}_{2,1}(V^\sim, V, \Gamma)$... Tensor 3. Stufe ("3D Matrix")
- $\bigotimes_2^1 V = V \otimes V^\sim \otimes V^\sim \cong V^\sim \otimes V^\sim \otimes V = \mathfrak{Mult}_{1,2}(V^\sim, V, \Gamma)$... Tensor 3. Stufe ("3D Matrix")
- $\bigotimes_2^2 V = V \otimes V \otimes V^\sim \otimes V^\sim \cong V^\sim \otimes V^\sim \otimes V \otimes V = \mathfrak{Mult}_{2,2}(V^\sim, V, \Gamma)$... Tensor 4. Stufe ("4D Matrix")

Sei V ein endlichdimensionaler Vektorraum mit Basis $(E_\beta)_{\beta \in I}$ und dualer Basis $(e^\alpha)_{\alpha \in I}$. Sei ein Tensor $t \in \mathfrak{Mult}_{p,q}(V^\sim, V, \Gamma) = \bigotimes^p V \otimes \bigotimes^q V^\sim = \bigotimes_q^p V$.

Jeder Vektor $v \in V$ und jeder Kovektor $\varphi \in V^\sim$ lässt sich bezüglich dieser Basen schreiben als

$$v = v^\beta E_\beta, \quad \varphi = \varphi_\alpha e^\alpha, \quad \text{wobei} \quad \delta_\beta^\alpha = e^\alpha(E_\beta)$$

Die Komponenten von t werden dann durch die Basisvektoren bestimmt

$$t(e^{\alpha_1}, \dots, e^{\alpha_p}, E_{\beta_1}, \dots, E_{\beta_q}) = t_{\beta_1 \dots \beta_q}^{\alpha_1 \dots \alpha_p} \in \Gamma \quad (3.36)$$

Für beliebige Vektoren $v_1, \dots, v_q \in V$ und Kovektoren $\varphi_1, \dots, \varphi_p \in V^\sim$ gilt dann

$$\begin{aligned} t(\varphi_1, \dots, \varphi_p, v_1, \dots, v_q) &= t(\varphi_{1,\alpha_1} e^{\alpha_1}, \dots, \varphi_{p,\alpha_p} e^{\alpha_p}, v_1^{\beta_1} E_{\beta_1}, \dots, v_q^{\beta_q} E_{\beta_q}) \\ &\stackrel{\text{Skalar}}{=} t(e^{\alpha_1}, \dots, e^{\alpha_p}, E_{\beta_1}, \dots, E_{\beta_q}) \varphi_{1,\alpha_1} \cdots \varphi_{p,\alpha_p} v_1^{\beta_1} \cdots v_q^{\beta_q} \\ &\stackrel{(3.36)}{=} \underbrace{t_{\beta_1 \dots \beta_q}^{\alpha_1 \dots \alpha_p}}_{\text{Koeffizienten}} \varphi_{1,\alpha_1} \cdots \varphi_{p,\alpha_p} v_1^{\beta_1} \cdots v_q^{\beta_q} \end{aligned}$$

Diese Gleichung beschreibt die Auswertung eines (p, q) -Tensors in Komponentenform.

3.7 Basis- und Koordinatentransformationen

Wir haben bisher mit einer festen Basis des Vektorraums bzw. Dualraums gearbeitet. Da Vektorräume mehrere Basen besitzen ist es oft notwendig, von einer Basis/Koordinatendarstellung in eine andere zu transformieren.

Seien dazu $(E_\alpha)_{\alpha \in I}$ und $(\tilde{E}_\alpha)_{\alpha \in I}$ zwei Basen desselben Vektorraums V . Um von einer Basis zur anderen überzugehen, verwenden wir eine Basistransformation. Es soll gelten:

$$\tilde{E}_\alpha = \Lambda_\alpha^\beta E_\beta \quad (3.37)$$

$$E_\beta = \Sigma_\beta^\gamma \tilde{E}_\gamma \quad (3.38)$$

wobei Λ_α^β und Σ_β^γ invertierbare Transformationsmatrizen sind und α angegeben werden muss, um zu wissen welches Basiselement zerlegt wurde.

Setzen wir (3.38) in (3.37) ein, so folgt

$$\tilde{E}_\alpha = \Lambda_\alpha^\beta \Sigma_\beta^\gamma \tilde{E}_\gamma \quad (3.39)$$

Andererseits kann \tilde{E}_α geschrieben werden als

$$\tilde{E}_\alpha = \delta_\alpha^\gamma \tilde{E}_\gamma \quad (3.40)$$

Da die Basisvektoren linear unabhängig sind folgt mit [Proposition 3.9](#), dass die Darstellung von \tilde{E}_α eindeutig ist und somit müssen die Matrizen aus [\(3.39\)](#) und [\(3.40\)](#) übereinstimmen:

$$\Lambda_\alpha^\beta \Sigma_\beta^\gamma = \delta_\alpha^\gamma \quad (3.41)$$

Analog erhält man durch Einsetzen in umgekehrter Reihenfolge:

$$\Sigma_\alpha^\beta \Lambda_\beta^\gamma = \delta_\alpha^\gamma$$

Die Matrizen Λ_α^β und Σ_α^β sind daher zueinander invers.

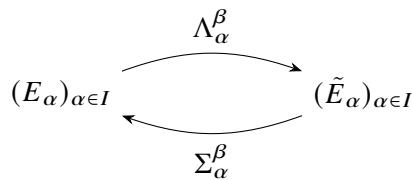


Abbildung 3.1: Basiswechsel im Vektorraum V

Betrachten wir nun die zugehörigen dualen Basen $(e^\alpha)_{\alpha \in I}$ und $(\tilde{e}^\alpha)_{\alpha \in I}$. Diese sind definiert durch

$$e^\alpha(E_\beta) = \delta_\beta^\alpha, \quad \tilde{e}^\alpha(\tilde{E}_\beta) = \delta_\beta^\alpha \quad (3.42)$$

Für die Transformation zwischen den beiden dualen Basen gelte:

$$\tilde{e}^\beta = \Xi_\gamma^\beta e^\gamma \quad (3.43)$$

Dann folgt:

$$\Lambda_\alpha^\gamma \Sigma_\gamma^\beta = \delta_\alpha^\beta \stackrel{\text{duale Basis}}{=} \tilde{e}^\beta(\tilde{E}_\alpha) \stackrel{(3.43)}{=} \Xi_\gamma^\beta e^\gamma(\tilde{E}_\alpha) \stackrel{(3.37)}{=} \Xi_\gamma^\beta e^\gamma \left(\underbrace{\Lambda_\alpha^\delta}_{\in \Gamma} E_\delta \right) \stackrel{\text{linear}}{=} \Xi_\gamma^\beta \Lambda_\alpha^\delta e^\gamma(E_\delta) \stackrel{\text{duale Basis}}{=} \Xi_\gamma^\beta \Lambda_\alpha^\delta \delta_\delta^\gamma = \Xi_\gamma^\beta \Lambda_\alpha^\gamma$$

und somit:

$$\begin{aligned} \Lambda_\alpha^\gamma \Sigma_\gamma^\beta &= \Xi_\gamma^\beta \Lambda_\alpha^\gamma \\ \Leftrightarrow \Sigma_\beta^\alpha &= \Xi_\beta^\alpha \quad \text{bzw. } \Xi_\beta^\alpha \cdot (\Sigma_\alpha^\gamma)^{-1} = \delta_\alpha^\gamma \end{aligned}$$

Die Erkenntnis ergibt wegen [\(3.42\)](#) durchaus Sinn.

Die duale Basis transformiert somit *invers* zur Vektorbasis. Die Transformation, die im Vektorraum von der alten Basis in die neue führt, bringt uns im Dualraum von der neuen Basis wieder zurück in die alte.

Kovariant (Kovektoren $\in V^\sim$): Die Komponenten transformieren gleich (ko = mit) der Basis (untere Indizes: $\varphi = \varphi_\alpha e^\alpha$).

Kontravariant (Vektoren $\in V$): Die Komponenten transformieren entgegengesetzt zur Basis (obere Indizes: $v = v^\alpha E_\alpha$).

$$\begin{array}{ccc}
 & \Xi_\alpha^\beta = \Sigma_\alpha^\beta & \\
 (e^\alpha)_{\alpha \in I} & \xrightarrow{\quad} & (\tilde{e}^\alpha)_{\alpha \in I} \\
 & \xleftarrow{\quad} & \\
 & (\Xi_\alpha^\beta)^{-1} = \Lambda_\alpha^\beta &
 \end{array}$$

Abbildung 3.2: Basiswechsel im Dualraum V^\sim

Zusammengefasst:

Für

$$\tilde{E}_\alpha = \Lambda_\alpha^\beta E_\beta \quad \text{und} \quad E_\alpha = \Sigma_\alpha^\beta \tilde{E}_\beta \quad \text{mit} \quad e^\alpha(E_\beta) = \delta_\beta^\alpha \quad \text{und} \quad \tilde{e}^\alpha(\tilde{E}_\beta) = \delta_\beta^\alpha$$

folgt

$$\Lambda_\gamma^\alpha \Sigma_\beta^\gamma = \delta_\beta^\alpha, \quad \tilde{e}^\alpha = \Sigma_\beta^\alpha e^\beta \quad \text{und} \quad e^\alpha = \Lambda_\beta^\alpha \tilde{e}^\beta$$

Transformation eines Tensors

Wir betrachten nun einen Tensor

$$t \in \bigotimes_q^p V = \bigotimes_q^p V \otimes \bigotimes^q V^\sim$$

Bezüglich der Basis $(E_\alpha)_{\alpha \in I}$ von V und der dualen Basis $(e^\alpha)_{\alpha \in I}$ lässt er sich schreiben als

$$t = \underbrace{t_{\beta_1 \dots \beta_q}^{\alpha_1 \dots \alpha_p}}_{\text{Koeffizienten}} E_{\alpha_1} \otimes \dots \otimes E_{\alpha_p} \otimes e^{\beta_1} \otimes \dots \otimes e^{\beta_q} \quad (3.44)$$

Da unsere Definition eines Tensors basisunabhängig ist, können wir denselben Tensor ebenso bezüglich einer beliebigen anderen Basis $(\tilde{E}_\alpha)_{\alpha \in I}$ von V und der zugehörige dualen Basis $(\tilde{e}^\alpha)_{\alpha \in I}$ darstellen:

$$t = \tilde{t}_{\nu_1 \dots \nu_q}^{\mu_1 \dots \mu_p} \tilde{E}_{\mu_1} \otimes \dots \otimes \tilde{E}_{\mu_p} \otimes \tilde{e}^{\nu_1} \otimes \dots \otimes \tilde{e}^{\nu_q} \quad (3.45)$$

Transformiere nun von der neuen Basis zurück in die alte mit den zuvor gefundenen Transformationsgleichungen

$$\tilde{E}_\mu = \Lambda_\mu^\alpha E_\alpha, \quad \tilde{e}^\nu = \Sigma_\beta^\nu e^\beta$$

Setze diese in (3.45) ein:

$$\begin{aligned}
 t &= \tilde{t}_{\nu_1 \dots \nu_q}^{\mu_1 \dots \mu_p} \Lambda_{\mu_1}^{\alpha_1} E_{\alpha_1} \otimes \dots \otimes \Lambda_{\mu_q}^{\alpha_q} E_{\alpha_q} \otimes \Sigma_{\beta_1}^{\nu_1} e^{\beta_1} \otimes \dots \otimes \Sigma_{\beta_p}^{\nu_p} e^{\beta_p} \\
 &\stackrel{\text{Tensorprod.}}{=} \tilde{t}_{\nu_1 \dots \nu_q}^{\mu_1 \dots \mu_p} \Lambda_{\mu_1}^{\alpha_1} \dots \Lambda_{\mu_p}^{\alpha_p} \Sigma_{\beta_1}^{\nu_1} \dots \Sigma_{\beta_q}^{\nu_q} \left(E_{\alpha_1} \otimes \dots \otimes E_{\alpha_p} \otimes e^{\beta_1} \otimes \dots \otimes e^{\beta_q} \right)
 \end{aligned} \quad (3.46)$$

Wir haben also mit (3.44) und (3.46) zwei Darstellungen vom Tensor t in derselben Basis gefunden. Nach da die Vektoren der Basis linear unabhängig sind ist die Darstellung nach [Proposition 3.9](#) eindeutig. Daher müssen die Koeffizienten übereinstimmen:

$$t_{\beta_1 \dots \beta_q}^{\alpha_1 \dots \alpha_p} = \tilde{t}_{\nu_1 \dots \nu_q}^{\mu_1 \dots \mu_p} \Lambda_{\mu_1}^{\alpha_1} \dots \Lambda_{\mu_p}^{\alpha_p} \Sigma_{\beta_1}^{\nu_1} \dots \Sigma_{\beta_q}^{\nu_q} \quad (3.47)$$

(3.47) wird **Tensor-Transformationsgesetz** genannt. Es beschreibt, wie sich die Komponenten eines (p, q) -Tensors beim Wechsel der Basis verändern.

Es ist der Ursprung der "Münchhausen Definition" eines Tensors: Ein Objekt heißt genau dann Tensor, wenn es wie ein Tensor transformiert.

3.8 Abstrakte Indexschreibweise

Wie wir gesehen haben, hängen die Komponenten eines Tensors von der Wahl der Basis ab, der Tensor selbst jedoch nicht. Die Komponenten tragen dabei obere und untere Indizes, entsprechend ihrem kontravarianten bzw. kovarianten Grad. Um den Typ eines Tensors unabhängig von einer konkreten Basis sichtbar zu machen, verwenden wir die sogenannte **abstrakte Indexschreibweise** (Penrose-Notation).

Dabei werden Tensoren mit *Symbolen* versehen, sogenannten abstrakten Indizes, die nicht für Zahlen stehen, sondern anzeigen, welchen ko- bzw. kontravarianten Rang der Tensor besitzt. Beispielsweise bedeutet

$$t \in \bigotimes_1^2 V = \bigotimes_1^2 V \otimes \bigotimes_1^1 V^\sim = \mathfrak{M}\text{ult}_{2,1}(V^\sim, V, \Gamma)$$

dass t zwei kontravariante und einen kovarianten Index besitzt, also ein Tensor vom Typ $(2, 1)$ ist. Dies lässt sich abstrakt schreiben als

$$t = t_c^{ab} \quad a, b, c \in I$$

wobei a , b und c *keine* Komponentenindizes sind, sondern lediglich den Typ/Stufe/Grad des Tensors angeben.

Auch die Objekte, auf die ein Tensor wirkt, tragen abstrakte Indizes. Um zu kennzeichnen, an welchem Index des Tensors ein Argument wirkt, verwendet man für das Argument denselben abstrakten Index.

Beispiel 3.57 (Vektor und Kovektor). Sei $v \in V = \bigotimes_0^1 V$ und $\varphi \in V^\sim = \bigotimes_1^0 V$. Dann ordnen wir zu

$$\begin{aligned} v &\rightarrow v^a \quad (\text{kontravarianter abstrakter Index } a) \\ \varphi &\rightarrow \varphi_a \quad (\text{kovarianter abstrakter Index } a) \end{aligned}$$

und es gilt

$$\varphi(v) = \varphi_a v^a = v^a \varphi_a = v(\varphi) \in \Gamma = \bigotimes_0^0 V$$

Abbildung und Argument tragen *denselben* abstrakten Index a , sodass klar ist, dass φ auf v wirkt. Dabei ist nur wichtig, dass sie denselben Index tragen. Da diese nur ein Symbol ist, kann dafür ein beliebiger gewählt werden: $\varphi_a v^a = \varphi_b v^b = \varphi_c v^c$.

Beispiel 3.58 (Tensorprodukt). Seien $\varphi, \psi \in V^\sim$. Dann ordnen wir $(\varphi \otimes \psi) \in \mathfrak{Bil}(V, \Gamma) = \bigotimes_2^0 V$ zu

$$(\varphi \otimes \psi) \rightarrow (\varphi \otimes \psi)_{ab} = \varphi_a \psi_b$$

Für die Auswertung von $(\varphi \otimes \psi)$ auf $v, w \in V$ gilt: $(\varphi \otimes \psi)(v, w) = \varphi(v) \cdot \psi(w) = \varphi_a v^a \psi_b w^b = \varphi_a \psi_b v^a w^b = \psi_b \varphi_a v^a w^b = \psi(w) \cdot \varphi(v)$. Also gilt $\varphi_a \psi_a = \psi_b \varphi_a$. Da aber $\varphi(v) \cdot \psi(w) \neq \psi(v) \cdot \varphi(w)$ ist somit $\varphi_a \psi_b \neq \psi_a \varphi_b$.

Beispiel 3.59 (Basis). Die Basis $(E_\alpha)_{\alpha \in I}$ des Vektorraums V kann in abstrakter Indexschreibweise als E_α^a notiert werden, wobei a der abstrakte Index und α der Zählindex ist, über den summiert wird. Die duale Basis wird entsprechend als $e^\alpha{}_a$ notiert. Es gilt somit:

$$e^\alpha(E_\beta) = e^\alpha_a E_\beta^a = E_\alpha^a e^\beta_a = \delta_\beta^\alpha$$

Die Zählindizes α, β , über die summiert wird, sind verschieden, aber der abstrakte Index a ist derselbe.

Beispiel 3.60 (Metrischer Tensor).

Quelle: [https://en.wikipedia.org/wiki/Metric_tensor_\(general_relativity\)#Definition](https://en.wikipedia.org/wiki/Metric_tensor_(general_relativity)#Definition)

Ein **metrischer Tensor** (z.B. die Minkowski-Metrik) auf einem Vektorraum V ist ein Tensor vom Typ $(0, 2)$ (kovarianter Tensor zweiter Stufe), der symmetrisch und nicht-entartet ist. In unserer Notation schreiben wir

$$g_{ab} \in \bigotimes_2^0 V = \mathfrak{Bil}(V, \Gamma)$$

Die Darstellung eines metrischen Tensors ist bezüglich einer Basis $(E_\alpha)_{\alpha \in I}$ wie üblich eindeutig

$$g_{ab} E_\alpha^a E_\beta^b = g_{\alpha\beta}$$

Für die Minkowski-Metrik gilt beispielsweise $\eta_{\alpha\beta} := \text{diag}(-1, 1, 1, 1)$.

Ein metrischer Tensor induziert die lineare Abbildung

$$g_{ab} : V \rightarrow V^\sim : v^a \mapsto v_b := g_{ab} v^a$$

und dessen Inverse

$$g^{ab} : V^\sim \rightarrow V : v_b \mapsto v^a := g^{ab} v_b$$

Der metrische Tensor identifiziert also ein Element aus dem Vektorraum V mit einem aus dessen Dualraum V^\sim . Allerdings sind v^a und v_a nicht dasselbe Objekt! Obige Operationen werden *Index-Transport* genannt.

Bemerkung 3.61 (Nicht Teil der VO).

Quelle: https://en.wikipedia.org/wiki/Musical_isomorphism#Discussion

Da $g \in \mathfrak{Bil}(V, \Gamma)$ ist durch das Festhalten eines Vektors $v \in V$ die Abbildung

$$g(v, \cdot) : V \rightarrow \mathbb{R}$$

linear im zweiten Argument und somit $g(v, \cdot) \in V^\sim$. Der metrische Tensor induziert daher die lineare Abbildung *flat*

$$\flat : V \rightarrow V^\sim : v^a \mapsto v^b := v_b := g(v, \cdot) = g_{ab} v^a$$

wobei für $v \in V^\sim, w \in V$ gilt $v^b(w) := v_b w^b := g(v, w) = g_{ab} v^a w^b$.

Wenn V endlich dimensional ist, ist \flat bijektiv und somit ein Isomorphismus. Dieser ist für V zusammen mit der Bilinearform g sogar ein *kanonischer* Isomorphismus. Seine Inverse ist der *sharp* Operator

$$\sharp : V^\sim \rightarrow V : v_b \mapsto v^\sharp := v^a := g^{-1}(v^b, \cdot) = g^{ab} v_b$$

wobei für $v \in V, \varphi \in V^\sim$ gilt $v^\sharp(\varphi) := v^a \varphi_a := g^{-1}(v^b, \varphi) = g^{ab} v_a \varphi_b$ und $g(v^\sharp, w) = v(w)$.

Da die beiden Operatoren \flat und \sharp aus der Musiknotation stammen, werden sie auch *musikalische Isomorphismen* genannt. Das Symbol \flat ("flat") bezeichnet das *Absenken* eines Tons, während \sharp ("sharp") ein *Anheben* eines Tons bedeutet. Analog dazu senkt der Operator \flat in der Tensorrechnung einen Index ab, und \sharp hebt einen Index an.

Die inverse Metrik g^{ab} ist entsprechend ein Tensor vom Typ $(2, 0)$,

$$g^{ab} \in \bigotimes_0^2 V = \mathfrak{Bil}(V^\sim, \Gamma)$$

und erfüllt

$$g^{ab} e_a^\alpha e_b^\beta = g^{\alpha\beta}$$

wobei e^β_a die duale Basis bezeichnet. Metrik und inverse Metrik stehen in der Beziehung

$$g^{ab} g_{bc} = g^{\alpha\beta} E_\alpha^a E_\beta^b g_{\gamma\delta} e_b^\gamma e_c^\delta = g^{\alpha\beta} g_{\gamma\delta} \underbrace{E_\alpha^a E_\beta^b}_{= \delta_\beta^\gamma} e_c^\gamma e_c^\delta = g^{\alpha\beta} g_{\beta\delta} E_\alpha^a e_c^\delta = \delta_\delta^\alpha \delta_c^\alpha = \delta_c^a.$$

Die inverse Minkowski-Metrik ist ebenso $\eta^{\alpha\beta} = \text{diag}(-1, 1, 1, 1)$.

Hier wurde die Metrik noch nicht als Längenbegriff eingeführt.

Definition 3.62 (Nicht Teil der VO). Sei $(V, \langle \cdot, \cdot \rangle)$ ein Skalarproduktraum, wobei das Skalarprodukt eine positiv definite symmetrische Bilinearform ist. Ein System von Vektoren $(E_\alpha)_{\alpha \in I} \subset V$ heißt **Orthonormalsystem** (ONS), wenn gilt

$$\langle E_\alpha, E_\beta \rangle = \begin{cases} 1, & \text{falls } \alpha = \beta \\ 0, & \text{falls } \alpha \neq \beta \end{cases} = \delta_\beta^\alpha$$

Das heißt: alle Vektoren in $(E_\alpha)_{\alpha \in I}$ sind normiert und paarweise orthogonal.

Ein Orthonormalsystem $M = (E_\alpha)_{\alpha \in I}$ heißt **Orthonormalbasis** (ONB) $\Leftrightarrow \overline{\text{span } M} = V$.

Definition 3.63 (Nicht Teil der VO). Sei $(V, (\cdot, \cdot))$ ein Pseudo-Skalarproduktraum, wobei das Pseudo-Skalarprodukt eine indefinite symmetrische nicht-entartete Bilinearform ist. Ein System von Vektoren $(E_\alpha)_{\alpha \in I} \subset V$ heißt **Pseudo-Orthonormalsystem** (ONS), wenn gilt

$$(E_\alpha, E_\beta) = \begin{cases} \pm 1, & \text{falls } \alpha = \beta \\ 0, & \text{falls } \alpha \neq \beta \end{cases}$$

Sei (V, g) ein endlichdimensionaler Vektorraum mit einem metrischen Tensor g und $(E_\alpha)_{\alpha \in I} \subset V$ eine Pseudo-ONB. Dann kann g damit dargestellt werden:

$$g(E_\alpha, E_\beta) = g_{ab} E_\alpha^a E_\beta^b = g_{\alpha\beta}$$

wobei $g_{\alpha\beta}$ die Komponenten der Metrik in dieser Pseudo-ONB bezeichnet. Obiges beschreibt die Transformation des metrischen Tensors durch eine Pseudo-ONB in eine Diagonalmatrix, da

$$g_{ab} E_\alpha^a E_\beta^b = \begin{cases} \pm 1, & \text{falls } \alpha = \beta \\ 0, & \text{falls } \alpha \neq \beta \end{cases}$$

Die Tatsache, dass jede nicht-ausgeartete symmetrische Bilinearform durch Wahl einer geeigneten Basis auf eine Diagonalform mit Einträgen ± 1 gebracht werden kann, folgt aus dem *Trägheitssatz von Sylvester*.

Beispiel 3.64 (Orthonormalbasis bzgl. der Minkowski-Metrik). Betrachte den Minkowski-Raum (\mathbb{M}, η) mit $\eta_{ab} = \text{diag}(-1, 1, 1, 1)$ und eine Pseudo-ONB $(E_\alpha)_{\alpha=0}^3$. Dann gilt

$$\eta(E_\alpha, E_\beta) = \eta_{ab} E_\alpha^a E_\beta^b = \eta_{\alpha\beta}$$

Insbesondere:

$$\eta_{ab} E_0^a E_0^b = \eta_{00} = -1$$

Für die räumlichen Basisvektoren ($i = 1, 2, 3$) gilt:

$$\eta_{ab} E_1^a E_1^b = \eta_{11} = +1, \quad \eta_{ab} E_2^a E_2^b = \eta_{22} = +1, \quad \eta_{ab} E_3^a E_3^b = \eta_{33} = +1$$

Dualbasis

Aus

$$\eta_{ab} E_\beta^a = E_{\beta b} \in V^\sim$$

folgt, dass es Koeffizienten λ_α gibt mit

$$\eta_{ab} E_\beta^a = E_{\beta b} = \lambda_\alpha e_b^\alpha, \quad \beta = 0, 1, 2, 3$$

Für den Zeitanteil ($\beta = 0$) gilt:

$$-1 = \eta_{ab} E_0^a E_0^b = \lambda_\alpha \underbrace{e_a^\alpha E_0^a}_{= \delta_a^\alpha} = \lambda_0$$

somit

$$e_a^0 = -E_{0a}$$

Ganz analog für die räumlichen Komponenten:

$$e_a^i = +E_{ia}, \quad i = 1, 2, 3$$

Indextransport

Ein Vektor $v \in V$ mit

$$v = v^\alpha E_\alpha$$

wird mittels der Metrik zu einem Kovektor

$$v_b = \eta_{ba} v^a = v_\beta e_b^\beta = \eta_{ab} v^\alpha E_\beta^\alpha = v^\alpha \eta_{ab} E_\alpha^a = v^\alpha E_{ab} = -v^0 e_b^0 + v^i e_b^i$$

Dies liefert den Zusammenhang zwischen Kovariant- und Kontravariantkomponenten:

$$\begin{aligned} \text{zeitartig: } \quad v_0 &= -v^0 \\ \text{räumlich: } \quad v_i &= v^i \quad (i = 1, 2, 3) \end{aligned}$$

Da Komponenten reine Zahlen sind, können sie gleichgesetzt werden, obwohl die Vektoren selbst in unterschiedlichen Räumen liegen.

Kompatibilität der Basis- und Dualbasis-Komponenten

Aus

$$E_{\alpha b} = \eta_{ba} E_\alpha^a = \lambda_{\alpha\beta} e_b^\beta$$

folgt durch Multiplikation mit E_γ^b :

$$E_\gamma^b E_\alpha^a \eta_{ba} = \eta_{\gamma\alpha} = \lambda_{\alpha\beta} \underbrace{e_b^\beta E_\gamma^b}_{= \delta_\gamma^\beta} = \lambda_{\alpha\gamma}$$

Damit folgt schließlich

$$E_{\alpha b} = \eta_{\alpha\beta} e_b^\beta$$

Die in den Dualraum transportierten Basisvektoren $E_{\alpha b}$ stimmen also mit den dualen Basisvektoren e_b^α überein, bis auf das charakteristische Minuszeichen im zeitartigen Anteil.

Kontraktion/Tensorverjüngung

Die Kontraktion eines Tensors reduziert dessen Stufe, indem jeweils ein kontravarianter und ein kovarianter Index "zusammengezogen" werden. Für einen Tensor

$$t \in \bigotimes_p^q V$$

erhält man durch Kontraktion einen Tensor

$$s \in \bigotimes_{p-1}^{q-1} V$$

Sei also

$$t(\varphi_1, \dots, \varphi_p, v_1, \dots, v_q) \in \Gamma,$$

wobei $\varphi_i \in V^\sim$ und $v_j \in V$. Die Kontraktion an der k -ten kovarianten und l -ten kontravarianten Stelle definiert man durch

$$s(\varphi_1, \dots, \varphi_{p-1}, v_1, \dots, v_{q-1}) := t(\varphi_1, \dots, \underbrace{e^\alpha}_{\text{k-te Stelle, } 1 \leq k \leq p}, \dots, \varphi_p, v_1, \dots, \underbrace{E_\alpha}_{\text{l-te Stelle, } 1 \leq l \leq q}, \dots, v_q)$$

Wir zeigen nun, dass dies unabhängig von der Wahl der Basis und somit wohldefiniert ist.

Betrachte einen Basiswechsel

$$\tilde{E}_\alpha = \Lambda_\alpha^\beta E_\beta, \quad \tilde{e}^\alpha = \Sigma_\beta^\alpha e^\beta,$$

mit inversen Transformationsmatrizen

$$\Sigma_\beta^\alpha \Lambda_\gamma^\beta = \delta_\gamma^\alpha.$$

Dann gilt:

$$\begin{aligned} t(\varphi_1, \dots, \tilde{e}^\alpha, \dots, \varphi_{p-1}, \tilde{E}_\alpha, \dots, v_{q-1}) &= t(\varphi_1, \dots, \Sigma_\beta^\alpha e^\beta, \dots, \varphi_{p-1}, v_1, \dots, \Lambda_\alpha^\beta E_\beta, \dots, v_{q-1}) \\ &= \Sigma_\beta^\alpha \Lambda_\gamma^\beta t(\varphi_1, \dots, e^\beta, \dots, \varphi_{p-1}, v_1, \dots, E_\gamma, \dots, v_{q-1}) \\ &= \delta_\beta^\gamma t(\varphi_1, \dots, e^\beta, \dots, \varphi_{p-1}, v_1, \dots, E_\gamma, \dots, v_{q-1}) \end{aligned}$$

Damit ist die Kontraktion basisunabhängig.

Kontraktion in abstrakter Indexschreibweise

In abstrakten Indizes lässt sich Kontraktion besonders einfach formulieren. Für einen Tensor

$$t_{b_1 \dots b_q}^{a_1 \dots a_p} \in \bigotimes_q^p V$$

ergibt sich folgende Rechenvorschrift für die Kontraktion des k -ten kontravarianten und des l -ten kovarianten Index:

$$t_{b_1 \dots c \dots b_q}^{a_1 \dots c \dots a_p} := t_{b_1 \dots b_l \dots b_q}^{a_1 \dots a_k \dots a_p} \underbrace{e_{a_k}^\alpha E_{\alpha}^{b_l}}_{=\delta_{a_k}^{b_l}} = S_{b_1 \dots b_{l-1}, b_{l+1} \dots b_q}^{a_1 \dots a_{k-1}, a_{k+1} \dots a_p} \in \bigotimes_{q-1}^{p-1} V$$

Die Basisvektoren bzw. Dualbasisvektoren schreiben wir als

$$E_\alpha^a \in \bigotimes_0^1 V, \quad e_a^\alpha \in \bigotimes_1^0 V,$$

mit

$$E_\alpha^a e_\alpha^\beta = \delta_\alpha^\beta \in \bigotimes_1^1 V$$

Für einen Vektor $v \in V$ gilt

$$v^a = v^\alpha E_\alpha^a.$$

Ebenso für die Wirkung eines Kovektors e^β auf einen Vektor v :

$$e^\beta(v) = e_\alpha^\beta v^\alpha = e_\alpha^\beta E_\alpha^a v^\alpha = \delta_\alpha^\beta v^\alpha = v^\beta$$

Schließlich folgt für die kombinierte Wirkung eines Basisvektors und eines Dualvektors:

$$(E_\alpha^b e_\beta^a) v^b = E_\alpha^b (e_\beta^a v_b) = v^\gamma \delta_\gamma^b \delta_\beta^a = v^a,$$

also insbesondere

$$E_\alpha^b e_\beta^a = \delta_\alpha^a.$$

Damit ist die abstrakte Indexnotation vollständig konsistent mit der ursprünglichen Definition der Kontraktion.

Bemerkung 3.65 (Nicht Teil der VO).

Ein Beispiel für Kontraktion ist die gewöhnliche Matrix–Vektor–Multiplikation. Eine Matrix A lässt sich als Tensor vom Typ $(1, 1)$, $(0, 2)$ oder $(2, 0)$ auffassen. Wähle für dieses Beispiel $(1, 1)$:

$$A_b^a \in \bigotimes_1^1 V$$

Ein Vektor v besitzt die Darstellung

$$v^b \in \bigotimes_0^1 V$$

Die Matrix–Vektor Multiplikation ist die Wirkung der Matrix auf den Vektor durch Kontraktion über den Index b :

$$A_b^a v^b = w^a \in \bigotimes_0^1 V.$$

Der kontrahierte Index verschwindet, sodass das Ergebnis wieder ein Vektor ist. Die Kontraktion reduziert den Tensorgrad von $(1, 1)$ und $(1, 0)$ zu $(1, 0)$.

Kapitel 4

Topologie

4.1 Der Begriff der Topologie

Die allgemeine Relativitätstheorie findet in gekrümmten Räumen, sogenannten Mannigfaltigkeiten, statt. Daher ist eine Verallgemeinerung des Begriffs der Geraden erforderlich. Dabei lösen wir uns von der Struktur des \mathbb{R}^n , insbesondere von seiner linearen (Vektorraum-) und seiner Metrikstruktur. Begriffe wie Konvergenz und Stetigkeit müssen deshalb unabhängig von einem Abstandsbegriff neu formuliert werden. Dies führt uns zum Begriff der Topologie.

Die Topologie bildet die Grundlage aller weiteren Strukturen: Differenzierbarkeit, Metrik, Krümmung und schließlich die Einsteinschen Feldgleichungen.

Bemerkung 4.1 (Mögliche Prüfungsfrage).

Wozu brauchen wir Topologie?

Antwort: siehe oben

Definition 4.2. Die **Potenzmenge** $\mathcal{P}(X)$ einer Menge X ist die Menge aller Teilmengen von X :

$$\mathcal{P}(X) := \{U \mid U \subseteq X\}.$$

Definition 4.3. Sei $X \neq \emptyset$ eine Menge (genannt Grundmenge). Eine Teilmenge $\mathfrak{X} \subseteq \mathcal{P}(X)$ heißt **Topologie** auf X , wenn folgende Eigenschaften gelten:

$$(O1) \quad \emptyset, X \in \mathfrak{X}$$

$$(O2) \quad O_1, O_2 \in \mathfrak{X} : O_1 \cap O_2 \in \mathfrak{X} \quad (\text{stabil unter endlichen Schnitten})$$

$$(O3) \quad (O_\alpha)_{\alpha \in I} \subseteq \mathfrak{X} : \bigcup_{\alpha \in I} O_\alpha \in \mathfrak{X} \quad (\text{stabil unter beliebigen Vereinigungen})$$

Die Elemente von \mathfrak{X} heißen **offene Mengen**. Daher wird \mathfrak{X} auch ein *System offener Mengen* genannt. Das Paar (X, \mathfrak{X}) heißt **topologischer Raum**. Das Zeichen \mathfrak{X} wird "Kurrent X" ausgesprochen.

Bemerkung 4.4 (Nicht Teil der VO). (O2) ist äquivalent zu:

$$(O2)' \quad O_1, \dots, O_n \in \mathfrak{X} : \bigcap_{i=1}^n O_i \in \mathfrak{X}$$

Bemerkung 4.5 (Prüfungsfrage).

Wie stelle ich fest, ob eine Menge offen ist?

Antwort: Ich muss mich zuerst fragen, bezüglich welcher Topologie ich die Offenheit prüfen muss. Anschließend prüfe ich, ob die gegebene Menge in der Topologie enthalten ist.

Beispiel 4.6. Sei X eine beliebige Grundmenge und $\mathfrak{Y} = \{\emptyset, X\}$ die sogenannte **indiskrete Topologie/Klumpentopologie**, wobei \mathfrak{Y} "Kurrent Y" ausgesprochen wird. Dann ist (X, \mathfrak{Y}) der indiskrete topologische Raum. Dies ist die kleinstmögliche Topologie auf X .

Wir überprüfen, ob es sich wirklich um eine Topologie handelt:

(O1) zz: $\emptyset, X \in \mathfrak{Y}$

Ist offensichtlich erfüllt, da \mathfrak{Y} genau aus diesen Mengen besteht.

(O2) zz: $O_1, O_2 \in \mathfrak{Y} : O_1 \cap O_2 \in \mathfrak{Y}$

Die offenen Mengen von \mathfrak{Y} sind \emptyset und X und es gilt

$$\emptyset, X \in \mathfrak{Y} : \emptyset \cap X = \emptyset \in \mathfrak{Y}$$

(O3) zz: $(O_\alpha)_{\alpha \in I} \subseteq \mathfrak{Y} : \bigcup_{\alpha \in I} O_\alpha \in \mathfrak{Y}$ Beliebige Vereinigungen aus $\{\emptyset, X\}$

Die offenen Mengen von \mathfrak{Y} sind \emptyset und X und es gilt

$$\emptyset, X \in \mathfrak{Y} : \emptyset \cup X = X \in \mathfrak{Y}$$

Damit ist \mathfrak{Y} eine Topologie.

Beispiel 4.7. Sei X eine beliebige Grundmenge und $\Theta = \mathcal{P}(X)$ die sogenannte **diskrete Topologie**. Dann ist (X, Θ) der diskrete topologische Raum. Dies ist die größtmögliche Topologie auf X .

Wir überprüfen wieder, ob es sich wirklich um eine Topologie handelt:

(O1) zz: $\emptyset, X \in \Theta$

Ist erfüllt, da $\emptyset, X \in \mathcal{P}(X)$.

(O2) zz: $O_1, O_2 \in \Theta : O_1 \cap O_2 \in \Theta$

Die offenen Mengen von $\Theta = \mathcal{P}(X)$ sind die Teilmengen von X . Der Schnitt zweier Teilmengen von X ist immer wieder Teilmenge von X und liegt somit in Θ .

(O3) zz: $(O_\alpha)_{\alpha \in I} \subseteq \Theta : \bigcup_{\alpha \in I} O_\alpha \in \Theta$

Auch beliebige Vereinigungen von Teilmengen von X sind wieder eine Teilmenge von X und liegen somit in Θ .

Damit ist Θ eine Topologie.

Definition 4.8. Seien τ und σ Topologien auf einer Menge X .

Gilt $\sigma \subseteq \tau$, so heißt τ **feiner** als σ (bzw. σ **größer** als τ).

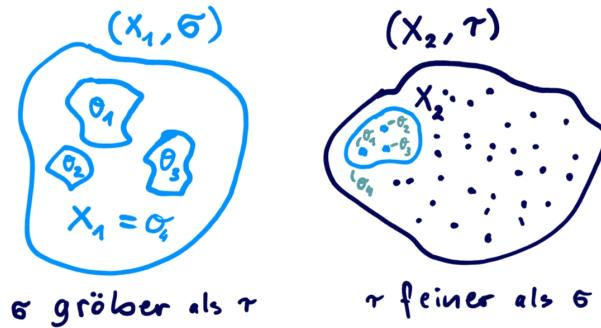


Abbildung 4.1: $\sigma \subseteq \tau$; feiner: enthält mehr offene Mengen; größer: enthält weniger offene Mengen

Damit ist die diskrete Topologie die feinstmögliche Topologie auf einer Menge, während die indiskrete Topologie die gröbstmögliche Topologie auf einer Menge ist. Für beliebige Topologien \mathfrak{X} gilt also $\mathfrak{Y} \subseteq \mathfrak{X} \subseteq \Theta$.

Bemerkung 4.9 (Mögliche Prüfungsfrage).

Gibt es Topologien überhaupt?

Antwort: Ja: Beispiele sind die diskrete und indiskrete Topologie.

Beispiel 4.10 (Nicht Teil der VO).

Betrachte $X = \{1, 2\}$ und $\mathfrak{X} = \{\emptyset, \{1\}, \{1, 2\}\} \subset \mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$ Dann gilt:

$$\begin{aligned} (O1) \quad & \emptyset, X \in \mathfrak{X} \\ (O2) \quad & \emptyset \cap \{1\} = \emptyset \in \mathfrak{X}, \\ & \emptyset \cap \{1, 2\} = \emptyset \in \mathfrak{X}, \\ & \{1\} \cap \{1, 2\} = \{1\} \in \mathfrak{X} \\ (O3) \quad & \emptyset \cup \{1\} = \{1\} \in \mathfrak{X}, \\ & \emptyset \cup \{1, 2\} = \{1, 2\} \in \mathfrak{X}, \\ & \{1\} \cup \{1, 2\} = \{1, 2\} \in \mathfrak{X} \end{aligned}$$

Damit ist \mathfrak{X} tatsächlich eine Topologie auf X . Hier wäre die indiskrete Topologie $\mathfrak{Y} = \{\emptyset, \{1, 2\}\}$ und die diskrete Topologie $\Theta = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$. Man sieht somit gut, dass $\mathfrak{Y} \subseteq \mathfrak{X} \subseteq \Theta$ gilt.

4.2 Konvergenz einer Folge

Erst durch die Ausstattung einer Menge mit einer Topologie wird es sinnvoll, über die Konvergenz von Folgen zu sprechen. Ohne eine solche zusätzliche Struktur gibt es keinen Begriff von "Nähe" und damit auch keinen Grenzwert.

Definition 4.11. Sei (X, \mathfrak{X}) ein topologischer Raum und $(x_n)_{n \in \mathbb{N}}$ eine Folge in X . Die Folge (x_n) heißt **konvergent gegen** $x \in X$ (bezüglich der Topologie \mathfrak{X}), in Zeichen $x_n \xrightarrow{\mathfrak{X}} x$, falls gilt:

$$\underbrace{\forall O \in \mathfrak{X} : x \in O}_{\text{für alle offenen Umgebungen von } x \text{ existiert ein Index } N_O \text{ ab dem alle Folgenglieder } x_n \text{ in } O \text{ liegen}} \quad \underbrace{\exists N_O \in \mathbb{N}}_{\text{für alle offenen Umgebungen von } x \text{ existiert ein Index } N_O \text{ ab dem alle Folgenglieder } x_n \text{ in } O \text{ liegen}} \quad \underbrace{\forall n \geq N_O \text{ sodass } x_n \in O}_{\text{für alle offenen Umgebungen von } x \text{ existiert ein Index } N_O \text{ ab dem alle Folgenglieder } x_n \text{ in } O \text{ liegen}}$$

für alle offenen Umgebungen von x existiert ein Index N_O ab dem alle Folgenglieder x_n in O liegen

Das bedeutet: Eine Folge konvergiert gegen x , wenn sich ihre Folgenglieder ab einem gewissen Index, beliebig

nahe bei x befinden. In der Topologie wird diese Nähe nicht über einen Abstand, sondern über offene Mengen beschrieben. Wir schreiben N_O , da der Index von der offenen Menge O abhängt.

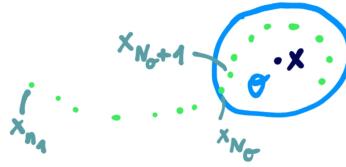


Abbildung 4.2: Konvergenz einer Folge $(x_n)_{n \in \mathbb{N}}$ gegen ein x .

Diese Definition zeigt bereits, dass Konvergenz stark von der vorliegenden Topologie abhängt. Insbesondere gilt:

- Je *größer* eine Topologie ist (weniger offene Mengen), desto weniger offene Mengen müssen die Konvergenzbedingung erfüllen, desto *schwächer* ist die Konvergenz.
- Je *feiner* eine Topologie ist (mehr offene Mengen), desto mehr offene Mengen müssen die Konvergenzbedingung erfüllen, desto *stärker* ist die Konvergenz.

Beispiel 4.12. Sei (X, \mathfrak{Y}) ein topologischer Raum mit der indiszierten Topologie $\mathfrak{Y} = \{\emptyset, X\}$. Sei $\tilde{x} \in X$ und $(x_n)_{n \in \mathbb{N}} \subset X$ eine bel. Folge. Wann konvergiert sie gegen \tilde{x} ? Die einzige in \mathfrak{Y} offene Menge, die \tilde{x} enthält, ist X . Also muss $(x_n)_{n \in \mathbb{N}}$ die folgende Bedingung erfüllen:

$$\text{für } X \in \mathfrak{Y} \exists N_X \in \mathbb{N} \forall n \geq N_X \text{ sodass } x_n \in X$$

Da aber die Folge bereits ganz in X liegt, tritt dies schon ab dem 1. Folgenglied ein. So ein Index existiert also und somit konvergiert die Folge gegen \tilde{x} . Da die Folge beliebig gewählt war, konvergieren also alle Folgen aus X gegen \tilde{x} . Dasselbe kann man nun für alle $x \in X$ durchspielen. Es folgt: alle Folgen in X konvergieren gegen alle $x \in X$. Der Grenzwert der Folgen ist nicht eindeutig. Die Konvergenz bezüglich der indiszierten Topologie besitzt somit nur sehr geringe Aussagekraft.

Jede Folge $(x_n)_{n \in \mathbb{N}} \subset X$ muss hier nur eine Bedingung erfüllen: die Folgenglieder müssen ab einem Index N_X in X liegen.

Beispiel 4.13. Sei (X, Θ) ein topologischer Raum mit der diskreten Topologie $\Theta = \mathcal{P}(X)$. Dann ist jede Teilmenge von X offen, insbesondere auch die einelementige Menge $\{x\}$.

Für Folge $(x_n)_{n \in \mathbb{N}} \subset X$, die gegen x konvergiert, muss daher gelten:

$$\exists N_{\{x\}} \in \mathbb{N} \forall n \geq N_{\{x\}} \text{ sodass } x_n \in \{x\}$$

also

$$x_n = x \quad \text{für alle } n \geq N_{\{x\}}$$

Eine Folge ist bezüglich der diskreten Topologie also genau dann konvergent, wenn sie ab einem gewissen Index konstant ist. Die Konvergenz ist hier eindeutig, aber sehr restriktiv: Nur sehr wenige Folgen konvergieren.

Es sei angemerkt, dass andere offene Mengen $O \in \Theta$ keine weiteren Einschränkungen sind, denn für $x \in O \in \Theta$ gilt bereits $\{x\} \subseteq O$ und $x_n \in \{x\} \subseteq O \forall n \geq N_{\{x\}}$.

Diese beiden Beispiele verdeutlichen den Zusammenhang zwischen Feinheit einer Topologie und Konvergenz: Je feiner die Topologie (je mehr offene Mengen), desto weniger Folgen konvergieren.

Bemerkung 4.14. Seien τ_1 und τ_2 Topologien auf X mit $\tau_1 \subseteq \tau_2$. Dann gilt:
Konvergiert eine Folge bezüglich τ_2 so konvergiert sie auch bezüglich τ_1 .

Grund: wenn eine Bedingung (offene Menge) weggelassen wird, wird die Folge natürlich weiterhin konvergieren.

4.3 Stetigkeit

Definition 4.15. Sei $f : X \rightarrow Y$. Das **Urbild** einer Teilmenge $B \subseteq Y$ unter f ist definiert durch

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\} \subseteq X$$

Die Menge $f^{-1}(B)$ ist immer definiert, f muss dafür nicht invertierbar sein.

Definition 4.16. Seien (X, \mathfrak{X}) und (Y, \mathfrak{Y}) topologische Räume.

Eine Abbildung $f : X \rightarrow Y$ heißt **stetig** : $\Leftrightarrow \forall O' \in \mathfrak{Y} : f^{-1}(O') \in \mathfrak{X}$.

Also hängt Stetigkeit ausschließlich von der zugrunde liegenden Topologie ab und benötigt keinen Abstandsgriff.

Anmerkung: $f^{-1}(O') \in \mathfrak{X}$ ist äquivalent zu $f^{-1}(O') \subseteq X \wedge f^{-1}(O')$ offen in \mathfrak{X} .

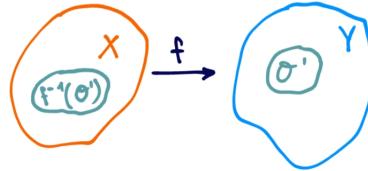


Abbildung 4.3: Stetigkeit

Bemerkung 4.17. Seien (X, \mathfrak{X}) und (Y, \mathfrak{Y}) topologische Räume und $f : X \rightarrow Y$ stetig.

- Werden weitere Mengen zur Topologie \mathfrak{X} hinzugefügt (die Topologie wird verfeinert), so bleibt f stetig.
- Werden Mengen aus der Topologie \mathfrak{Y} entfernt (die Topologie wird größer gemacht), so bleibt f stetig.

Beispiel 4.18. Die Funktion $f : (X, \mathfrak{X}) \rightarrow (Y, \underbrace{\{\emptyset, Y\}}_{\text{indirekte Top.}})$ ist stetig:

$f^{-1}(\emptyset) = \{x \in X \mid f(x) \in \emptyset\} = \emptyset$ und $f^{-1}(Y) = \{x \in X \mid f(x) \in Y\} = X$. Da immer $\emptyset, X \in \mathfrak{X}$ gilt, ist f stetig.

Die Funktion $f : (X, \underbrace{\mathcal{P}(X)}_{\text{diskrete Top.}}) \rightarrow (Y, \mathfrak{Y})$ ist stetig:

$\forall O' \in \mathfrak{Y} : f^{-1}(O') \subseteq X$ und somit ist $f^{-1}(O') \in \mathcal{P}(X)$.

Beispiel 4.19. Seien (X, \mathfrak{X}) und (Y, \mathfrak{Y}) beliebige topologische Räume und sei $y_0 \in Y$ fest. Die konstante Abbildung

$$c : X \rightarrow Y, \quad c(x) := y_0$$

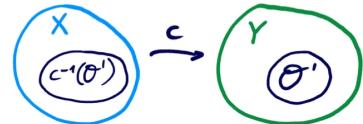
ist stetig.

Wir müssen also zeigen: $\forall O' \in \mathfrak{Y} : c^{-1}(O') \in \mathfrak{X}$

Sei $O' \in \mathfrak{Y}$. Dann gilt

$$c^{-1}(O') = \{x \in X \mid c(x) \in O'\} = \{x \in X \mid y_0 \in O'\}$$

Es treten zwei Fälle auf:



- Ist $y_0 \in O'$, so ist $c^{-1}(O') = X$
- Ist $y_0 \notin O'$, so ist $c^{-1}(O') = \emptyset$

Da immer $\emptyset, X \in \mathfrak{X}$ gilt, ist c stetig.

Proposition 4.20 (Zusammensetzung stetiger Funktionen ist stetig). Seien $(X, \mathfrak{X}), (Y, \mathfrak{Y}), (Z, \mathfrak{Z})$ topologische Räume und $f : X \rightarrow Y, g : Y \rightarrow Z$. Dann gilt

f und g stetig \Rightarrow die Hintereinanderausführung $g \circ f : X \rightarrow Z, (g \circ f)(x) := g(f(x))$ ist stetig.

Beweis. zz: $\forall O'' \in \mathfrak{Z} : (g \circ f)^{-1}(O'') \in \mathfrak{X}$

Sei $O'' \in \mathfrak{Z}$. Da g stetig ist, gilt

$$g^{-1}(O'') \in \mathfrak{Y}$$

Da f stetig ist, ist wiederum

$$\forall O' \in \mathfrak{Y} : f^{-1}(O') \in \mathfrak{X}$$

und somit

$$f^{-1}(g^{-1}(O'')) \in \mathfrak{X}$$

Das heißt, wir müssen zeigen:

$$f^{-1}(g^{-1}(O'')) = (g \circ f)^{-1}(O'')$$

Tatsächlich gilt diese Beziehung sogar ganz allgemein für beliebige Mengen $C \subseteq Z$. Wir zeigen, dass gilt $(g \circ f)^{-1}(C) \subseteq f^{-1}(g^{-1}(C))$ und $(g \circ f)^{-1}(C) \supseteq f^{-1}(g^{-1}(C))$.

$$x \in (g \circ f)^{-1}(C) \stackrel{\text{Def.}}{\Leftrightarrow} g(f(x)) \in C \stackrel{\substack{\text{Def.} \\ \text{Urbild}}}{\Leftrightarrow} f(x) \in g^{-1}(C) \stackrel{\substack{\text{Def.} \\ \text{Urbild}}}{\Leftrightarrow} x \in f^{-1}(g^{-1}(C))$$

Obige Richtung " \Rightarrow " zeigt " \subseteq " und " \Leftarrow " zeigt " \supseteq ".

□

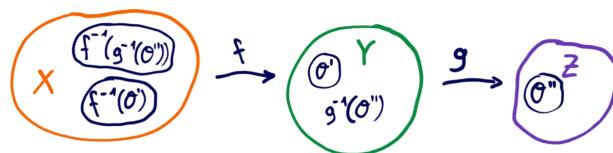


Abbildung 4.4: Situation von Proposition 4.20

Merksatz für $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$: Wenn man Kleidung anzieht, muss man diese in umgekehrter Reihenfolge wieder ausziehen.

Proposition 4.21 (Stetige Abbildungen erhalten Konvergenz). Seien (X, \mathfrak{X}) , (Y, \mathfrak{Y}) topologische Räume. Sei die Folge $(x_n)_{n \in \mathbb{N}} \subset X$ konvergent $x_n \xrightarrow{\mathfrak{X}} x \in X$ und $f : X \rightarrow Y$ stetig. Dann gilt

$$f(x_n) \xrightarrow{\mathfrak{Y}} f(x) \in Y$$

Beweis. zz: $\forall O \in \mathfrak{Y} : f(x) \in O \exists N_O \in \mathbb{N} \forall n \geq N_O$ sodass $f(x_n) \in O$

Sei $O \in \mathfrak{Y}$ beliebig mit $f(x) \in O$. Dann ist $x \in f^{-1}(O)$.

Da f stetig ist, gilt $f^{-1}(O) \in \mathfrak{X}$. Da $x_n \xrightarrow{\mathfrak{X}} x \in X$ gilt:

$$\exists N_{f^{-1}(O)} \in \mathbb{N} \forall n \geq N_{f^{-1}(O)} \text{ sodass } x_n \in f^{-1}(O)$$

Daraus folgt (mit Def. Urbild)

$$f(x_n) \in O \quad \forall n \geq N_{f^{-1}(O)}$$

also $f(x_n) \xrightarrow{\mathfrak{Y}} f(x)$. □

4.4 Topologisieren von Räumen

Sei (Y, \mathfrak{Y}) ein topologischer Raum und $f : X \rightarrow Y$. Wir suchen die gröbstmögliche Topologie, bezüglich dieser f stetig wird.

Da f stetig sein soll, muss für jede offene Menge $O' \in \mathfrak{Y}$ gelten, dass ihr Urbild unter f offen in X ist. Wir sammeln alle Urbilder der offenen Mengen in \mathfrak{Y} und stecken sie in \mathfrak{X}_f :

$$\mathfrak{X}_f := \{O \subseteq X \mid \exists O' \in \mathfrak{Y} \text{ mit } O = f^{-1}(O')\} = f^{-1}(\mathfrak{Y})$$

Wir werden anschließend zeigen, dass es sich hierbei um eine Topologie handelt. Also sind alle Mengen daraus offen und f somit stetig.

\mathfrak{X}_f holt in \mathfrak{Y} offene Mengen nach X .

Wir überprüfen nun, dass \mathfrak{X}_f tatsächlich die Axiome einer Topologie erfüllt.

$$\mathfrak{X}_f \subseteq \mathcal{P}(X) \checkmark$$

(O1) zz: $\emptyset, X \in \mathfrak{X}_f$

Da $\emptyset \in \mathfrak{Y}$ gilt, folgt

$$f^{-1}(\emptyset) = \{x \in X \mid f(x) \in \emptyset\} = \emptyset,$$

also $\emptyset \in \mathfrak{X}_f$.

Weiter gilt $Y \in \mathfrak{Y}$ und damit

$$f^{-1}(Y) = \{x \in X \mid f(x) \in Y\} = X,$$

also $X \in \mathfrak{X}_f$.

(O2) zz: $O_1, O_2 \in \mathfrak{X}_f : O_1 \cap O_2 \in \mathfrak{X}_f$

Also zz: $\exists O' \in \mathfrak{Y} : O_1 \cap O_2 = f^{-1}(O')$.

Seien $O_1, O_2 \in \mathfrak{X}_f$. Dann existieren offene Mengen $O'_1, O'_2 \in \mathfrak{Y}$ mit $O_1 = f^{-1}(O'_1), O_2 = f^{-1}(O'_2)$.

Somit ist

$$O_1 \cap O_2 = f^{-1}(O'_1) \cap f^{-1}(O'_2)$$

Da \mathfrak{Y} eine Topologie ist, gilt (O2): $O'_1 \cap O'_2 \in \mathfrak{Y}$. Also müssen wir zeigen

$$f^{-1}(O'_1) \cap f^{-1}(O'_2) = f^{-1}(O'_1 \cap O'_2)$$

Diese Beziehung kann sogar für beliebige Teilmengen $B_1, B_2 \subseteq Y$ gezeigt werden:

$$\begin{aligned} x \in f^{-1}(B_1) \cap f^{-1}(B_2) &\Leftrightarrow x \in f^{-1}(B_1) \wedge x \in f^{-1}(B_2) \\ &\stackrel{\text{Def. Urbild}}{\Leftrightarrow} f(x) \in B_1 \wedge f(x) \in B_2 \\ &\Leftrightarrow f(x) \in B_1 \cap B_2 \\ &\stackrel{\text{Def. Urbild}}{\Leftrightarrow} x \in f^{-1}(B_1 \cap B_2) \end{aligned}$$

Damit folgt

$$f^{-1}(B_1) \cap f^{-1}(B_2) = f^{-1}(B_1 \cap B_2)$$

Insbesondere gilt

$$O_1 \cap O_2 = f^{-1}(O'_1 \cap O'_2).$$

und es folgt

$$O_1 \cap O_2 \in \mathfrak{X}_f.$$

(O3) zz: $(O_\alpha)_{\alpha \in I} \subseteq \mathfrak{X}_f : \bigcup_{\alpha \in I} O_\alpha \in \mathfrak{X}_f$

Also zz: $\exists O' \in \mathfrak{Y} : \bigcup_{\alpha \in I} O_\alpha = f^{-1}(O')$

Sei $(O_\alpha)_{\alpha \in I} \subseteq \mathfrak{X}_f$. Dann existieren offene Mengen $(O'_\alpha)_{\alpha \in I} \subseteq \mathfrak{Y}$ mit

$$O_\alpha = f^{-1}(O'_\alpha) \quad \text{für alle } \alpha \in I.$$

Wir betrachten die Vereinigung:

$$\bigcup_{\alpha \in I} O_\alpha = \bigcup_{\alpha \in I} f^{-1}(O'_\alpha)$$

Da \mathfrak{Y} eine Topologie ist, gilt (O3): $\bigcup_{\alpha \in I} O'_\alpha \in \mathfrak{Y}$.

Also müssen wir zeigen

$$\bigcup_{\alpha \in I} f^{-1}(O_\alpha) = f^{-1}\left(\bigcup_{\alpha \in I} O_\alpha\right)$$

Diese Beziehung kann wieder sogar für beliebige Teilmengen $(B_\alpha)_{\alpha \in I} \subseteq \mathcal{P}(Y)$ gezeigt werden:

$$x \in \bigcup_{\alpha \in I} f^{-1}(B_\alpha) \Leftrightarrow \exists \alpha_0 \in I : x \in f^{-1}(B_{\alpha_0})$$

$$\stackrel{\text{Def. Urbild}}{\Leftrightarrow} \exists \alpha_0 \in I : f(x) \in B_{\alpha_0}$$

$$\Leftrightarrow f(x) \in \bigcup_{\alpha \in I} B_\alpha$$

$$\stackrel{\text{Def. Urbild}}{\Leftrightarrow} x \in f^{-1}\left(\bigcup_{\alpha \in I} B_\alpha\right)$$

Daraus folgt

$$\bigcup_{\alpha \in I} f^{-1}(B_\alpha) = f^{-1}\left(\bigcup_{\alpha \in I} B_\alpha\right)$$

Insbesondere gilt

$$\bigcup_{\alpha \in I} O_\alpha = f^{-1}\left(\bigcup_{\alpha \in I} O'_\alpha\right)$$

und damit

$$\bigcup_{\alpha \in I} O_\alpha \in \mathfrak{X}_f$$

Definition 4.22. Die so definierte Topologie \mathfrak{X}_f heißt die **initiale Topologie** bezüglich der Abbildung $f : X \rightarrow (Y, \mathfrak{Y})$.

Sie ist die *größte* Topologie auf X , für die f stetig ist. Jede andere Topologie auf X , bezüglich der f stetig ist, ist feiner als \mathfrak{X}_f .

Bemerkung 4.23. Das Konzept der initialen Topologie lässt sich unmittelbar auf Familien von Abbildungen $f_i : X \rightarrow Y_i$ verallgemeinern. In diesem Fall ist die initiale Topologie die größte Topologie auf X , bezüglich der alle f_i stetig sind.

4.5 Spurtopologie (Teilmengentopologie)

Sei (X, \mathfrak{X}) ein topologischer Raum und sei $U \subseteq X$ eine Teilmenge. Wir möchten U mit einer Topologie versehen, die in natürlicher Weise aus der gegebenen Topologie auf X hervorgeht.

Zu diesem Zweck betrachten wir die sogenannte **kanonische Einbettungsabbildung**

$$i_U : U \rightarrow X, \quad i_U(x) := x.$$



Diese Abbildung ordnet jedem Punkt aus U denselben Punkt in X zu und beschreibt damit die Einbettung der Teilmenge U in den umgebenden Raum X .

Sei $O \in \mathfrak{X}$. Dann gilt nach Definition des Urbildes:

$$i_U^{-1}(O) = \{x \in U \mid i_U(x) \in O\} = \{x \in U \mid x \in O\} = O \cap U \dots \text{repräsentiert die Spur von } O \text{ in } U$$

Definition 4.24. Die **Spurtopologie** oder **Teilmengentopologie** auf U ist definiert durch

$$\mathfrak{U} := i_U^{-1}(\mathfrak{X}) = \{O' \subseteq U \mid O' = O \cap U \wedge O \in \mathfrak{X}\} = \{O \cap U \mid O \in \mathfrak{X}\} = \mathfrak{X} \cap U$$

und macht (U, \mathfrak{U}) zu einem topologischen Raum.

Die offenen Mengen der Spurtopologie sind also genau diejenigen Mengen, die als Schnitt einer offenen Menge aus X mit U entstehen.

Die Spurtopologie beschreibt die "Spur", welche die Topologie \mathfrak{X} des umgebenden Raumes X auf der Teilmenge U hinterlässt. Sie ist die Einschränkung von \mathfrak{X} auf U .

4.6 Homöomorphismen

Wir betrachten zwei topologische Räume (X, \mathfrak{X}) und (Y, \mathfrak{Y}) . Wann sind zwei Räume aus topologischer Sicht **gleichartig** (d.h. besitzen dieselben topologischen Eigenschaften)?

Definition 4.25. Seien (X, \mathfrak{X}) und (Y, \mathfrak{Y}) topologische Räume. Eine Abbildung $f : X \rightarrow Y$ heißt **Homöomorphismus** : \Leftrightarrow

- f ist bijektiv. d.h. $\exists g : Y \rightarrow X$ (Inverse von f) sodass

$$g \circ f = \text{id}_X \text{ (also } f \text{ injektiv)} \quad \text{und} \quad f \circ g = \text{id}_Y \text{ (also } f \text{ surjektiv)}$$

- f und g sind stetig

In diesem Fall heißen die Räume (X, \mathfrak{X}) und (Y, \mathfrak{Y}) **homöomorph**. In Zeichen $\mathfrak{X} \simeq \mathfrak{Y}$.

Bemerkung 4.26. homo = gleich, morph = Gestalt \Rightarrow Homöomorphismus = gleichgestaltig. Homöomorphismen sind Isomorphismen auf topologischen Räumen. \mathfrak{X} und \mathfrak{Y} sind nicht gleich, aber aus topologischer Sicht ununterscheidbar. Sie haben dieselben topologischen Eigenschaften wie beispielsweise Kompaktheit, Zusammenhang und Trennungseigenschaften.

Ein Homöomorphismus dehnt, staucht, verbiegt, verzerrt einen Gegenstand so, dass er dennoch die gleichen topologischen Eigenschaften wie zuvor hat. Zerschneiden ist nur erlaubt, wenn die Teile später genau an der Schnittfläche wieder zusammenfügt werden.



Abbildung 4.5: Eine Tasse ist homöomorph zu einem Donut.

Quelle: https://commons.wikimedia.org/wiki/File:Topology_joke.jpg

Satz 4.27. (X, \mathfrak{X}) und (Y, \mathfrak{Y}) topologische Räume und $f : X \rightarrow Y$ ein Homöomorphismus mit der Inversen $g : Y \rightarrow X$. Wenn wir f nun auffassen als $f^{-1} : \mathcal{P}(Y) \rightarrow \mathcal{P}(X)$, ordnet f^{-1} jeder offenen Menge $O' \in \mathfrak{Y}$ genau ein $O \in \mathfrak{X}$ zu sodass $f^{-1}(O') = O$.

Also gilt $f^{-1}(\mathfrak{Y}) = \mathfrak{X}$ und analog dazu $g^{-1}(\mathfrak{X}) = \mathfrak{Y}$. Das heißt, die Topologien stehen in einer 1 zu 1 Beziehung, nicht nur die Punkte aus X und Y .

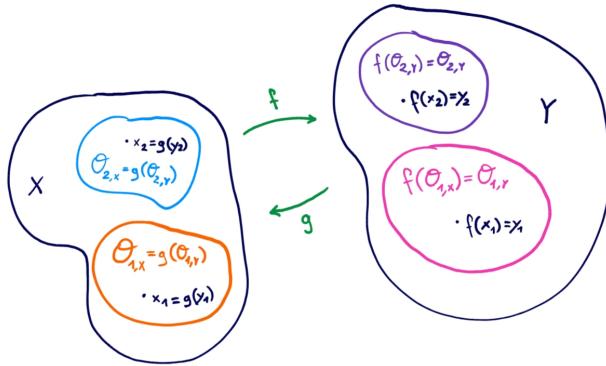


Abbildung 4.6: Homöomorphismus f zwischen (X, \mathfrak{X}) und (Y, \mathfrak{Y})

Beweis.

- zz: $g^{-1}(\mathfrak{X}) = \mathfrak{Y}$

Aufgrund der Stetigkeit von f gilt: $f^{-1}(\mathfrak{Y}) \subseteq \mathfrak{X}$

Wir möchten darauf g^{-1} anwenden. Daher zeigen wir die folgende allgemeine Beziehung.

Für beliebige Teilmengen $B_1, B_2 \subseteq Y$ gilt:

$$B_1 \subseteq B_2 \Rightarrow f^{-1}(B_1) \subseteq f^{-1}(B_2) \quad (4.1)$$

denn

$$x \in f^{-1}(B_1) \Leftrightarrow f(x) \in B_1 \subseteq B_2 \Rightarrow f(x) \in B_2 \Leftrightarrow x \in f^{-1}(B_2)$$

Wenden wir (4.1) nun an, so erhalten wir:

$$\begin{aligned} f^{-1}(\mathfrak{Y}) \subseteq \mathfrak{X} &\Rightarrow \underbrace{g^{-1}(f^{-1}(\mathfrak{Y}))}_{=(f \circ g)^{-1}(\mathfrak{Y})} \subseteq g^{-1}(\mathfrak{X}) \\ &\Leftrightarrow \underbrace{(f \circ g)^{-1}(\mathfrak{Y})}_{=id_Y} \subseteq g^{-1}(\mathfrak{X}) \\ &\Leftrightarrow \mathfrak{Y} \subseteq g^{-1}(\mathfrak{X}) \end{aligned}$$

Da g stetig ist, gilt: $g^{-1}(\mathfrak{X}) \subseteq \mathfrak{Y}$

Zusammen folgt daher:

$$\mathfrak{Y} = g^{-1}(\mathfrak{X}) \quad (4.2)$$

- zz: $f^{-1}(\mathfrak{Y}) = \mathfrak{X}$

$$f^{-1}(\mathfrak{Y}) \stackrel{(4.2)}{=} \underbrace{f^{-1}(g^{-1}(\mathfrak{X}))}_{(g \circ f)^{-1}} = \underbrace{(g \circ f)^{-1}(\mathfrak{X})}_{=id_X} = \mathfrak{X}$$

□

4.7 Trennungsaxiome

Quellen: https://www.old.mathematik.tu-dortmund.de/~lschwach/WS13/Proseminar/Dirk_Lauschner.pdf
<https://www.tuwien.at/index.php?eID=download&t=f&f=204097&token=564fbc343ac3892d97b2713d52d540fd084e8ca5>
https://www.math.tugraz.at/~ganster/lv_materialien/trennungseigenschaften.pdf

Trennungseigenschaften geben Auskunft darüber, ob "genügend offene Mengen" vorhanden sind, um Punkte bzw. Teilmengen voneinander zu "trennen". Aus dem Vorhandensein gewisser Trennungseigenschaften ergeben sich meist sehr wichtige Konsequenzen.

Definition 4.28. Ein topologischer Raum (X, \mathfrak{X}) heißt **T_0 -Raum** : \Leftrightarrow

$$\forall x, y \in X \text{ mit } x \neq y \exists O \in \mathfrak{X} : (x \in O \wedge y \notin O) \vee (y \in O \wedge x \notin O)$$

Das heißt: die Topologie kann verschiedene Punkte unterscheiden.



Abbildung 4.7: T_0

Definition 4.29. Ein topologischer Raum (X, \mathfrak{X}) heißt **T_1 -Raum** : \Leftrightarrow

$$\forall x, y \in X \text{ mit } x \neq y \exists O_x, O_y \in \mathfrak{X} : (x \in O_x \wedge y \in O_y) \wedge (x \notin O_y \wedge y \notin O_x)$$

Das heißt: je zwei verschiedene Punkte können durch offene Mengen getrennt werden.

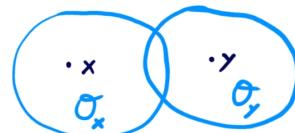


Abbildung 4.8: T_1

Im Gegensatz zu T_0 verlangt T_1 , dass es offene Mengen um *beide* Punkte gibt, die jeweils den anderen Punkt ausschließen. Also anstelle des logischen "oder" von T_0 gilt hier das logische "und".

Es gilt:

$$T_1 \Rightarrow T_0 \quad \text{aber} \quad T_0 \not\Rightarrow T_1$$



Abbildung 4.9: T_0 aber nicht T_1

Definition 4.30. Eine Menge $A \subseteq X$ ist **abgeschlossen** : $\Leftrightarrow A^c := X \setminus A \in \mathfrak{X}$ (das Komplement ist offen)

Bemerkung 4.31. In der Vorlesung wird statt $X \setminus A$ geschrieben $X - A$.

Bemerkung 4.32. Die leere Menge \emptyset und die Gesamtmenge X sind immer offen und abgeschlossen zugleich: Per Definition ist immer $\emptyset, X \in \mathfrak{X}$ (also die Mengen sind offen) und es gilt: $\emptyset^c = X \setminus \emptyset = X, X^c = X \setminus X = \emptyset$ (also sind die Mengen abgeschlossen).

Fun fact: Wenn \emptyset und X die einzigen Mengen sind, die offen und abgeschlossen sind, ist der Raum **zusammenhängend** (kann nicht in disjunkte nichtleere offene Mengen zerlegt werden).

Proposition 4.33. In einem T_1 -Raum sind alle einpunktigen Mengen abgeschlossen.

Beweis. zz: $\forall x \in X : X \setminus \{x\} \in \mathfrak{X}$

Sei $x \in X$ beliebig. Sei dazu $x' \in X \setminus \{x\}$, also $x' \neq x$. Da der Raum T_1 ist, existiert $O_{x'} \in \mathfrak{X}$ mit

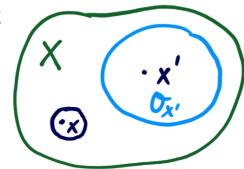
$$x' \in O_{x'} \quad \text{und} \quad x \notin O_{x'}.$$

Damit gilt

$$O_{x'} \subseteq X \setminus \{x\}$$

Somit auch

$$\{x'\} \subseteq O_{x'} \subseteq X \setminus \{x\}$$



Diese Kettenbeziehung bleibt durch Vereinigung über $x' \in X \setminus \{x\}$ erhalten

$$\underbrace{\bigcup_{x' \in X \setminus \{x\}} \{x'\}}_{=X \setminus \{x\}} \subseteq \underbrace{\bigcup_{x' \in X \setminus \{x\}} O_{x'}}_{=X \setminus \{x\}} \subseteq \underbrace{\bigcup_{x' \in X \setminus \{x\}} X \setminus \{x\}}_{=X \setminus \{x\}}$$

Wir erhalten

$$X \setminus \{x\} \subseteq \bigcup_{x' \in X \setminus \{x\}} O_{x'} \subseteq X \setminus \{x\}$$

Also

$$X \setminus \{x\} = \bigcup_{x' \in X \setminus \{x\}} O_{x'} \stackrel{(O3)}{\in} \mathfrak{X}$$

Damit ist $\{x\}$ abgeschlossen. □

Definition 4.34. Ein topologischer Raum (X, \mathfrak{X}) heißt **T_2 -Raum** oder **Hausdorff-Raum** : \Leftrightarrow

$$\forall x, y \in X \text{ mit } x \neq y \exists O_x, O_y \in \mathfrak{X} : (x \in O_x \wedge y \in O_y) \wedge (O_x \cap O_y = \emptyset)$$

Das heißt: verschiedene einpunktige Mengen lassen sich durch disjunkte offene Mengen trennen.

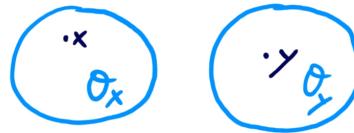


Abbildung 4.10: T_2

Hier werden die beiden Punkte nicht nur getrennt, sondern sogar durch disjunkte offene Mengen. Also eine Verschärfung von T_1 . Insbesondere gilt:

$$T_2 \Rightarrow T_1 \quad \text{aber} \quad T_1 \not\Rightarrow T_2$$

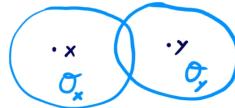


Abbildung 4.11: Ist T_1 aber nicht T_2 .

Proposition 4.35. In einem Hausdorff-Raum (X, \mathfrak{X}) ist der Grenzwert einer konvergenten Folge eindeutig.

Beweis. Wir zeigen dies mit Kontraposition:

Grenzwert einer konvergenten Folge ist nicht eindeutig \Rightarrow nicht T_2

Angenommen, eine Folge $(x_n)_{n \in \mathbb{N}}$ konvergiert sowohl gegen x als auch gegen y mit $x \neq y$.

Aus der Konvergenz folgt, dann es Indizes N_x, N_y gibt mit

$$\begin{aligned} \forall O_x \in \mathfrak{X} \text{ mit } x \in O_x \exists N_x \in \mathbb{N} \forall n \geq N_x \text{ sodass } x_n \in O_x \\ \forall O_y \in \mathfrak{X} \text{ mit } y \in O_y \exists N_y \in \mathbb{N} \forall n \geq N_y \text{ sodass } x_n \in O_y \end{aligned}$$

Daher müsste es einen Index $N^* = \max\{N_x, N_y\}$ geben, für den gilt

$$\forall n \geq N^* : x_n \in O_x \wedge x_n \in O_y$$

und somit ist

$$O_x \cap O_y \neq \emptyset$$

Also ist der topologische Raum nicht Hausdorff. □

Definition 4.36. Ein topologischer Raum (X, \mathfrak{X}) heißt **T_3 -Raum** : \Leftrightarrow

$$\forall x \in X, A \subseteq X \text{ abgeschlossen}, x \notin A \exists O_x, O_A \in \mathfrak{X} : (x \in O_x \wedge A \subseteq O_A) \wedge (O_x \cap O_A = \emptyset)$$

Das heißt: Einpunktige und abgeschlossene Mengen lassen sich durch disjunkte offene Mengen trennen.



Abbildung 4.12: T_3

Definition 4.37. Ein topologischer Raum (X, \mathfrak{X}) heißt **T_4 -Raum** : \Leftrightarrow

$$\forall A, B \subseteq X \text{ abgeschlossen mit } A \cap B = \emptyset \exists O_A, O_B \in \mathfrak{X} : (A \subseteq O_A \wedge B \subseteq O_B) \wedge (O_A \cap O_B = \emptyset)$$

Das heißt: disjunkte abgeschlossene Mengen lassen sich durch disjunkte offene Mengen trennen.

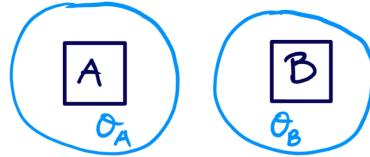


Abbildung 4.13: T_4

Definition 4.38 (Nicht Teil der VO).

Ein topologischer Raum (X, \mathfrak{X}) heißt

- **regulär**, wenn er $T_1 \wedge T_3$ ist
- **normal**, wenn er $T_1 \wedge T_4$ ist

In der Literatur ist "regulär", "normal" und T_4 nicht einheitlich definiert, siehe [hier](#).

Bemerkung 4.39. Zwischen den Trennungsaxiomen bestehen folgende wichtige Zusammenhänge:

$$\begin{aligned} T_1 + T_3 = T_2 : & \quad \text{einpunktige Mengen sind abgeschlossen } (T_1) \\ & + \text{einpunktige Mengen lassen sich von abgeschlossene Mengen trennen } (T_3) \\ & \Rightarrow \text{einpunktige Mengen lassen sich voneinander trennen } (T_2) \end{aligned}$$

$$\begin{aligned} T_1 + T_4 = T_3 : & \quad \text{einpunktige Mengen sind abgeschlossen } (T_1) \\ & + \text{abgeschlossene Mengen lassen sich voneinander trennen } (T_4) \\ & \Rightarrow \text{einpunktige \& abgeschlossene Mengen lassen sich voneinander trennen } (T_3) \end{aligned}$$

Beispiel 4.40. Es gibt topologische Räume, die T_3 sind, aber weder T_1 noch T_2 erfüllen (also $T_3 \not\Rightarrow T_2$ und $T_3 \not\Rightarrow T_1$).

Sei $X := \{x, y\}, \mathfrak{X} := \{\emptyset, X\}$ die indiskrete Topologie auf X .

- Offene Mengen: \emptyset und $\{x, y\} = X$.
- Abgeschlossene Mengen: $\emptyset^c = \{x, y\} = X$ und $X^c = \emptyset$.

Der Raum (X, \mathfrak{X}) ist T_3 :

- trenne x und \emptyset : $x \in X$ und $\emptyset \subseteq \emptyset$
- trenne y und \emptyset : $y \in X$ und $\emptyset \subseteq \emptyset$

Der Raum ist *nicht* T_1 : Versuche x und y zu trennen. Es gibt nur eine offene Menge, in der bereits sowohl x als auch y liegen: X . Somit können x und y nicht durch verschiedene offene Mengen getrennt werden.

Der Raum auch *nicht* T_2 : Versuche x und y zu trennen. Da es keine verschiedenen offenen Mengen gibt, die x bzw. y enthalten, gibt es insbesondere keine verschiedenen disjunkten offenen Mengen, die dies tun.

Bemerkung 4.41. Metrische Räume sind T_2 und T_4 . Da $T_2 \Rightarrow T_1 \Rightarrow T_0$ und $T_1 + T_4 = T_3$ erfüllen sie also alle hier genannten Trennungsaxiome. Der \mathbb{R}^n ist ein metrischer Raum.

4.8 Die euklidische Topologie

Wir wollen nun untersuchen, mit welcher Topologie der Raum \mathbb{R}^n *standardmäßig* ausgestattet wird, sodass die aus der reellen Analysis bekannten Begriffe von Konvergenz und ε - δ -Stetigkeit mit der topologischen Struktur übereinstimmen. Diese Topologie heißt die **euklidische Topologie** oder auch **Standardtopologie des \mathbb{R}^n** .

Definition 4.42. Eine Abbildung

$$d : X \times X \rightarrow \mathbb{R}$$

heißt **Metrik**, falls für alle $x, y, z \in X$ gilt:

$$\begin{aligned} d(x, y) &\geq 0 \quad \text{und} \quad d(x, y) = 0 \Leftrightarrow x = y && (\text{positive Definitheit}) \\ d(x, y) &= d(y, x) && (\text{Symmetrie}) \\ d(x, y) &\leq d(x, z) + d(z, y) && (\text{Dreiecksungleichung}) \end{aligned}$$

Ein Raum (M, d) , der mit einer Metrik d versehen ist, heißt **metrischer Raum**.

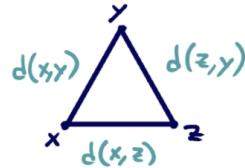


Abbildung 4.14: Dreiecksungleichung

Der Abstand zwischen zwei Punkten $x, y \in \mathbb{R}^n$ wird durch die **euklidische Metrik** beschrieben

$$d(x, y) := \|x - y\|_2 := \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

wobei $\|\cdot\|_2$ die **euklidische Norm** ist.

Mithilfe dieses Abstands definieren wir sogenannte **offene Kugeln**. Für einen Mittelpunkt $x \in \mathbb{R}^n$ und einen Radius $\varepsilon > 0$ sei

$$K(x, \varepsilon) := \{y \in \mathbb{R}^n \mid \|y - x\|_2 < \varepsilon\}$$

Wir definieren die **euklidische Topologie** \mathfrak{R} als die Menge aller Teilmengen $O \subseteq \mathbb{R}^n$, die sich als Vereinigung offener Kugeln $K_\alpha := K(x_\alpha, \varepsilon_\alpha)$ schreiben lassen:

$$\mathfrak{R} := \{O \subseteq \mathbb{R}^n \mid O = \bigcup_{\alpha \in I} K_\alpha\}$$

Dann heißt $(\mathbb{R}^n, \mathfrak{R})$ **euklidischer Raum**.

Diese Definition liefert zunächst lediglich eine Kandidatenmenge. Wir müssen nun zeigen, dass \mathfrak{R} tatsächlich eine Topologie ist. Dazu überprüfen wir die drei Topologieaxiome.

(O1) zz: $\emptyset, \mathbb{R}^n \in \mathfrak{R}$

Wir definieren

$$\emptyset := \bigcup_{\alpha \in \emptyset} K_\alpha$$

Damit ist $\emptyset \in \mathfrak{R}$.

Für jedes $x \in \mathbb{R}^n$ existiert ein $n \in \mathbb{N}$ mit $\|x\|_2 < n$, also $x \in K(0, n)$ und

$$\begin{aligned} & \{x\} \subseteq K(0, n) \\ \Leftrightarrow & \underbrace{\bigcup_{x \in \mathbb{R}^n} \{x\}}_{=\mathbb{R}^n} \subseteq \underbrace{\bigcup_{x \in \mathbb{R}^n} K(0, n)}_{=K(0, n)} \\ \Leftrightarrow & \mathbb{R}^n \subseteq K(0, n) \end{aligned}$$

Außerdem gilt

$$K(0, n) \subseteq \bigcup_{n \in \mathbb{N}} K(0, n) \subseteq \mathbb{R}^n$$

Somit folgt

$$\mathbb{R}^n = \bigcup_{n \in \mathbb{N}} K(0, n) \in \mathfrak{R}$$

(O2) zz: $\forall O_1, O_2 \in \mathfrak{R} : O_1 \cap O_2 \in \mathfrak{R}$

Braucht mehr Überlegungen, da der Durchschnitt von Kugeln nicht notwendigerweise eine Kugel sein muss. Seien nun $O_1, O_2 \in \mathfrak{R}$. Dann existieren Familien offener Kugeln $(K_\alpha)_{\alpha \in I}$ und $(K_\beta)_{\beta \in J}$ mit

$$O_1 = \bigcup_{\alpha \in I} K_\alpha, \quad O_2 = \bigcup_{\beta \in J} K_\beta.$$

Der Schnitt lässt sich schreiben als

$$O_1 \cap O_2 = \bigcup_{\substack{\alpha \in I \\ \beta \in J}} K_\alpha \cap K_\beta$$

Fall 1: für $\alpha \in I, \beta \in J : K_\alpha \cap K_\beta = \emptyset \in \mathfrak{R} \checkmark$

Fall 2: für $\alpha \in I, \beta \in J : K_\alpha \cap K_\beta \neq \emptyset$. Dieser Fall benötigt mehr Arbeit.

Dazu werden wir zeigen, dass für $y \in K_\alpha \cap K_\beta$ und ein spezielles $\delta^* > 0$ gilt:

$$K_\alpha \cap K_\beta = \bigcup_{y \in K_\alpha \cap K_\beta} K(y, \delta^*) \in \mathfrak{R}$$

Sei also $y \in K_\alpha \cap K_\beta$ beliebig.

Dann gelten die Abstände

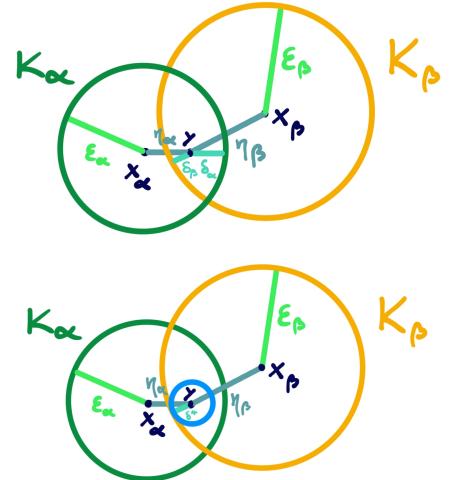
$$\eta_\alpha := \|y - x_\alpha\|_2 < \varepsilon_\alpha \quad \eta_\beta := \|y - x_\beta\|_2 < \varepsilon_\beta$$

Wir setzen

$$\delta_\alpha := \varepsilon_\alpha - \eta_\alpha > 0, \quad \delta_\beta := \varepsilon_\beta - \eta_\beta > 0,$$

und definieren

$$\delta^* := \min\{\delta_\alpha, \delta_\beta\}$$



Wir zeigen nun:

$$K(y, \delta^*) \subseteq K_\alpha \cap K_\beta$$

Dafür müssen wir zeigen, dass $\forall z \in K(y, \delta^*) : z \in K_\alpha \wedge z \in K_\beta$

Also müssen wir zeigen: $\|z - x_\alpha\|_2 < \varepsilon_\alpha \wedge \|z - x_\beta\|_2 < \varepsilon_\beta$

Sei dazu $z \in K(y, \delta^*)$ beliebig. Dann gilt

$$\begin{aligned} \|z - x_\alpha\|_2 &= \|z - y + y - x_\alpha\|_2 \stackrel{\Delta \text{ Unigl.}}{\leq} \|z - y\|_2 + \|y - x_\alpha\|_2 \\ &< \delta^* + \eta_\alpha \\ &\leq \delta_\alpha + \eta_\alpha \\ &= \varepsilon_\alpha - \eta_\alpha + \eta_\alpha = \varepsilon_\alpha \end{aligned}$$

Also ist $z \in K_\alpha$. Die gleiche Rechnung liefert

$$\|z - x_\beta\|_2 < \varepsilon_\beta,$$

also $z \in K_\beta$.

Damit gilt

$$K(y, \delta^*) \subseteq K_\alpha \cap K_\beta.$$

und auch

$$\begin{aligned} \{y\} &\subseteq K(y, \delta^*) \subseteq K_\alpha \cap K_\beta \\ \Leftrightarrow \underbrace{\bigcup_{y \in K_\alpha \cap K_\beta} \{y\}}_{=K_\alpha \cap K_\beta} &\subseteq \underbrace{\bigcup_{y \in K_\alpha \cap K_\beta} K(y, \delta^*)}_{=K_\alpha \cap K_\beta} \subseteq \underbrace{\bigcup_{y \in K_\alpha \cap K_\beta} K_\alpha \cap K_\beta}_{=K_\alpha \cap K_\beta} \end{aligned}$$

Somit folgt:

$$K_\alpha \cap K_\beta = \bigcup_{y \in K_\alpha \cap K_\beta} K(y, \delta^*) \in \mathfrak{R}.$$

Damit haben wir erhalten

$$O_1 \cap O_2 = \bigcup_{\substack{\alpha \in I \\ \beta \in J}} K_\alpha \cap K_\beta = \bigcup_{\alpha \in I} \bigcup_{\beta \in J} K(y, \delta^*) \in \mathfrak{R}.$$

(O3) zz: $\forall (O_\alpha)_{\alpha \in I} \subseteq \mathfrak{R} : \bigcup_{\alpha \in I} O_\alpha \in \mathfrak{R}$

Sei nun $(O_\alpha)_{\alpha \in I} \subseteq \mathfrak{R}$. Dann existieren für jedes α offene Kugeln $(K_\beta)_{\beta \in I_\alpha}$ mit

$$O_\alpha = \bigcup_{\beta \in I_\alpha} K_\beta.$$

Daraus folgt

$$\bigcup_{\alpha \in I} O_\alpha = \bigcup_{\alpha \in I} \bigcup_{\beta \in I_\alpha} K_\beta \in \mathfrak{R}.$$

Damit sind alle drei Topologieaxiome erfüllt, und \mathfrak{R} ist eine Topologie auf \mathbb{R}^n .

Definition 4.43. Ein Vektorraum V über dem Körper $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, versehen mit einer Topologie \mathfrak{X} , heißt **topologischer Vektorraum**, wenn die algebraischen Operationen Vektoraddition und Skalarmultiplikation stetig sind.

Bemerkung 4.44 (Nicht Teil der VO). Jeder Vektorraum kann, versehen mit der indiskreten Topologie, zu einem topologischen Vektorraum gemacht werden, da die algebraischen Operationen Vektoraddition und Skalarmultiplikation bzgl. der indiskreten Topologie stetig sind.

Bemerkung 4.45 (Nicht Teil der VO).

Sei $(V, \|\cdot\|)$ ein Vektorraum mit einer Norm.

1. Die Norm induziert durch Differenzenbildung eine Metrik

$$d(x, y) := \|x - y\|, \quad x, y \in V.$$

Damit wird der Vektorraum (V, d) ein metrischer Raum.

2. Mit dieser Metrik definieren wir für $x \in V$ und $\varepsilon > 0$ die offene Kugel

$$U_\varepsilon(x) := \{y \in V \mid d(x, y) < \varepsilon\}$$

Damit heißt eine Teilmenge $M \subseteq V$ offen, falls

$$\forall x \in M \exists \varepsilon > 0 : U_\varepsilon(x) \subseteq M.$$

3. Über diese offenen Mengen induziert die Metrik nun auf V eine Topologie:

$$\mathfrak{X} := \{M \subseteq V \mid M \text{ ist offen}\}$$

4. Mit dieser Topologie wird der Vektorraum V zu einem *topologischen Vektorraum* (V, \mathfrak{X}) .

Die so entstehende Topologie heißt die von der Norm induzierte **Normtopologie**. Das heißt, jeder Normierte Raum kann zu einem metrischen Raum gemacht werden und jeder metrische Raum kann zu einem topologischen Raum gemacht werden. Damit gelten alle Aussagen über topologische Räume auch für metrische und normierte Räume.

Bemerkung 4.46. In endlichdimensionalen Vektorräumen sind alle Normen äquivalent. Daher induzieren

die zugehörigen Metriken dieselbe Topologie: im \mathbb{R}^n und \mathbb{C}^n wird diese die **euklidische Topologie** oder auch die **Standardtopologie** des \mathbb{R}^n und \mathbb{C}^n genannt.

4.9 Kompaktheit und Parakompaktheit

Kompaktheit und Parakompaktheit sind zentrale Begriffe der Topologie und spielen später insbesondere bei der Theorie der Mannigfaltigkeiten eine wichtige Rolle.

Definition 4.47. Sei (X, \mathfrak{X}) ein topologischer Raum. Eine Familie offener Mengen $(O_\alpha)_{\alpha \in I} \subseteq \mathfrak{X}$ heißt **offene Überdeckung** von X , falls

$$X = \bigcup_{\alpha \in I} O_\alpha$$

Eine offene Überdeckung heißt **endlich**, wenn es endlich viele Indizes $\alpha_1, \dots, \alpha_N \in I$ gibt mit

$$X = O_{\alpha_1} \cup \dots \cup O_{\alpha_N}$$

Definition 4.48. Ein topologischer Raum (X, \mathfrak{X}) heißt **kompakt**, wenn gilt: zu jeder offenen Überdeckung von X existiert eine endliche Subüberdeckung/Teilüberdeckung. Das heißt

$$\forall (O_\alpha)_{\alpha \in I} \subseteq \mathfrak{X} \text{ mit } X = \bigcup_{\alpha \in I} O_\alpha \exists O_{\alpha_1}, \dots, O_{\alpha_N} \in \mathfrak{X} : X = O_{\alpha_1} \cup \dots \cup O_{\alpha_N}$$

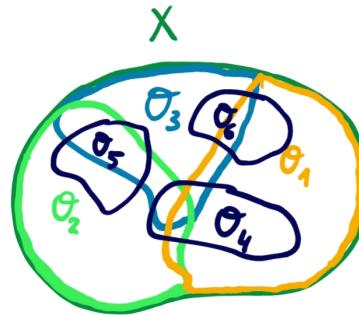


Abbildung 4.15: $\bigcup_{i=1}^3 O_i$ ist eine endliche Teilüberdeckung von X

Bemerkung 4.49. In der Literatur findet man die Definition der Kompaktheit teils auch mit der Hausdorff-Bedingung.

Bemerkung 4.50. Im \mathbb{R}^n gilt: Eine Teilmenge $M \subseteq \mathbb{R}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist (Satz von Heine-Borel). Insbesondere ist \mathbb{R}^n selbst nicht kompakt, während z.B. die Oberfläche einer Kugel kompakt ist.

Definition 4.51. Seien $(O_\alpha)_{\alpha \in I}$ und $(O_\beta)_{\beta \in J}$ offene Überdeckungen von X . Die Überdeckung $(O_\beta)_{\beta \in J}$ heißt **Verfeinerung** von $(O_\alpha)_{\alpha \in I}$, falls

$$\forall \beta \in J \exists \alpha \in I : O_\beta \subseteq O_\alpha$$

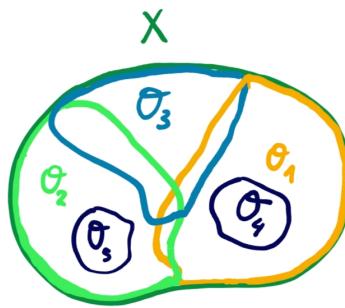


Abbildung 4.16: Offene Überdeckung $X = \bigcup_{i=1}^3 O_i$; Verfeinerung $X = \bigcup_{i=1}^5 O_i$, denn: $O_1 \subseteq O_1$, $O_2 \subseteq O_2$, $O_3 \subseteq O_3$, $O_4 \subseteq O_1$, $O_5 \subseteq O_2$

Definition 4.52. Eine offene Überdeckung $(O_\alpha)_{\alpha \in I}$ von X heißt **lokal endlich**, wenn zu jedem Punkt $x \in X$ eine Umgebung $U(x)$ existiert, die nur endlich viele Mengen der Überdeckung schneidet, d.h.

$$\forall x \in X \exists U(x) \in \mathfrak{X} \text{ mit } x \in U(x) : |\{\alpha \in I \mid O_\alpha \cap U(x) \neq \emptyset\}| < \infty$$

Definition 4.53. Ein topologischer Raum (X, \mathfrak{X}) heißt **parakompakt**, wenn zu jeder offenen Überdeckung von X eine lokal endliche Verfeinerung existiert.

Bemerkung 4.54. Parakompaktheit ist eine schwächere Eigenschaft als Kompaktheit. Insbesondere gilt:

$$\text{kompakt} \Rightarrow \text{parakompakt}$$

Metrische Räume sind parakompakt.

Kapitel 5

Mannigfaltigkeiten

5.1 Der Begriff der Mannigfaltigkeit

In der Allgemeinen Relativitätstheorie wird Gravitation durch gekrümmte Räume beschrieben, die mathematisch als Mannigfaltigkeiten modelliert werden.

Definition 5.1. Eine (**topologische**) **Mannigfaltigkeit** ist ein topologischer Raum (M, \mathfrak{M}) mit den folgenden Eigenschaften:

- M ist ein Hausdorff-Raum (Folgen haben eindeutigen Grenzwert)
- M ist parakompakt (um lokale Begriffe zu globalisieren)
- M ist lokal homöomorph zum $(\mathbb{R}^n, \mathfrak{R})$ ("dem \mathbb{R}^n sehr ähnlich/ sieht im Herangezoomten aus wie der \mathbb{R}^n ")

Die Zahl n heißt dann die **Dimension** der Mannigfaltigkeit M .

Die **lokale Homöomorphie** zum \mathbb{R}^n bedeutet, dass für jeden Punkt $p \in M$ eine offene Umgebung existiert, die (mit der von M induzierten Spurtopologie) homöomorph zum euklidischen Raum ist:

$$\forall p \in M \exists O_p \in \mathfrak{M} \text{ mit } p \in O_p : (O_p, \mathfrak{M} \cap O_p) \simeq (\mathbb{R}^n, \mathfrak{R})$$

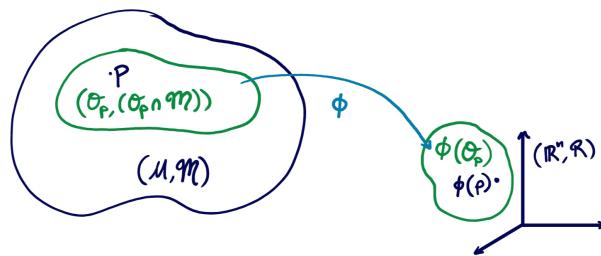


Abbildung 5.1: Homöomorphismus ϕ

Wir können also die ganze Mannigfaltigkeit mit offenen Mengen überdecken, die homöomorph zum \mathbb{R}^n sind.

Bemerkung 5.2. In der Allgemeinen Relativitätstheorie betrachtet man in der Regel 4-dimensionale Mannigfaltigkeiten, wobei drei Dimensionen räumlich und eine zeitlich interpretiert werden.

Definition 5.3. Das Paar (O, ϕ) bestehend aus einer offenen Menge $O \in \mathfrak{M}$ und einem Homöomorphismus

$$\phi : O \rightarrow \phi(O) \subseteq \mathbb{R}^n, \quad \phi(p) = (x^1(p), \dots, x^n(p))$$

heißt **Karte** der Mannigfaltigkeit M .

Da ϕ ein Homöomorphismus ist, besitzt die Abbildung eine stetige Inverse

$$\phi^{-1} : \phi(O) \subseteq \mathbb{R}^n \rightarrow O$$

Definition 5.4. Eine Familie von Karten

$$(O_\alpha, \phi_\alpha)_{\alpha \in I}$$

heißt ein **Atlas** der Mannigfaltigkeit M , falls gilt:

$$M = \bigcup_{\alpha \in I} O_\alpha$$

Eine Mannigfaltigkeit kann sich somit also eine Ansammlung von \mathbb{R}^n 's vorstellen.

Beispiel 5.5. Die Erdoberfläche ist eine 2-dimensionale Mannigfaltigkeit. Jede Region O_i der gekrümmten Erdoberfläche kann mit einer Karte (O_i, ϕ_i) auf eine Ebene $\phi_i(O_i) \subseteq \mathbb{R}^2$ abgebildet werden. Ein Atlas ist die Sammlung all dieser Karten, bei der auch Überlappungen zugelassen werden.

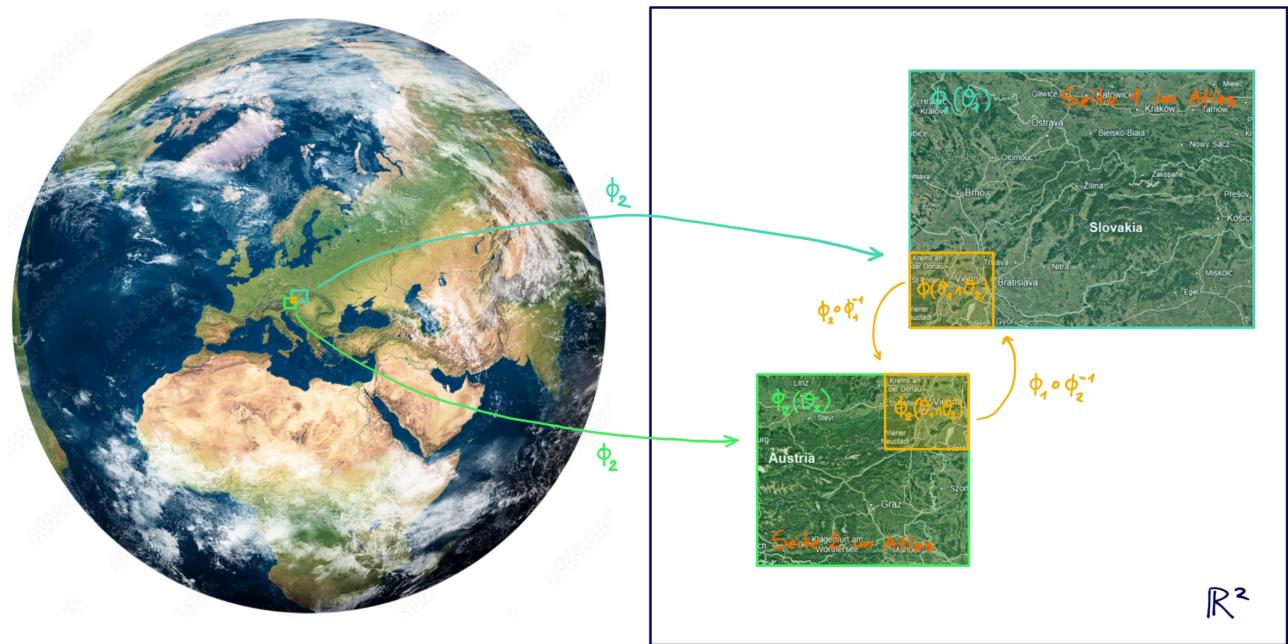
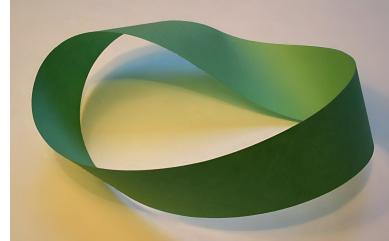


Abbildung 5.2: Die Erdoberfläche als Mannigfaltigkeit. Der gelbe Bereich auf der Kugel (links) ist die Überlappung $O_1 \cap O_2$, der gelbe Bereich rechts im \mathbb{R}^2 ist die Abbildung davon unter ϕ_1 und ϕ_2 . Um zwischen $\phi_1(O_1 \cap O_2)$ und $\phi_2(O_1 \cap O_2)$ zu wechseln, also auf die andere Karte zu wechseln, benötigt es die *Kartenwechselabbildung*. Diese beschreibt auf welcher anderen Seite im Atlas die Karte weiter geht.

Also sind Mannigfaltigkeiten eine Ansammlung von \mathbb{R}^n 's, die auf geeignete Weise miteinander verklebt sind.

Beispiel 5.6. Weitere Beispiele für Mannigfaltigkeiten sind der \mathbb{R}^n , \mathbb{C}^n , die Kleinflasche und das Möbiusband.



(a) Die Kleinflasche besitzt nur eine Seite und keine Innen- *und* Außenseite. Sie besitzt auch keinen Rand.

Quelle:

[https://de.laborxing.com/
produkten/kleinsche-flasche](https://de.laborxing.com/produkten/kleinsche-flasche)

(b) Das Möbiusband besitzt ebenso nur eine Seite und keine Innen- *und* Außenseite, aber hat einen Rand.

Quelle:

[https://de.wikipedia.org/wiki/
M%C3%B6biusband#/media/Datei:
M%C3%B6bius_Strip.jpg](https://de.wikipedia.org/wiki/M%C3%B6biusband#/media/Datei:M%C3%B6bius_Strip.jpg)

5.2 Kartenwechsel

Überlappen sich zwei offene Umgebungen $O_\alpha, O_\beta \in \mathfrak{M}$, die homöomorph zum \mathbb{R}^n sind, so wird dessen Durchschnitt durch verschiedene Karten abgebildet. Der Übergang zwischen diesen Darstellungen erfolgt über die **Kartenwechselabbildung**.

Definition 5.7. Für zwei Karten (O_α, ϕ_α) und (O_β, ϕ_β) mit $O_\alpha \cap O_\beta \neq \emptyset$ ist der **Kartenwechsel**

$$\psi_{\beta\alpha} := \phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(O_\alpha \cap O_\beta) \subseteq \mathbb{R}^n \rightarrow \phi_\beta(O_\alpha \cap O_\beta) \subseteq \mathbb{R}^n$$

Der inverse Kartenwechsel ist gegeben durch

$$\psi_{\alpha\beta} := \phi_\alpha \circ \phi_\beta^{-1}$$

und es gilt

$$\psi_{\alpha\beta} \circ \psi_{\beta\alpha} = \text{id}_{\phi_\alpha(O_\alpha \cap O_\beta)} \quad \psi_{\beta\alpha} \circ \psi_{\alpha\beta} = \text{id}_{\phi_\beta(O_\alpha \cap O_\beta)}$$

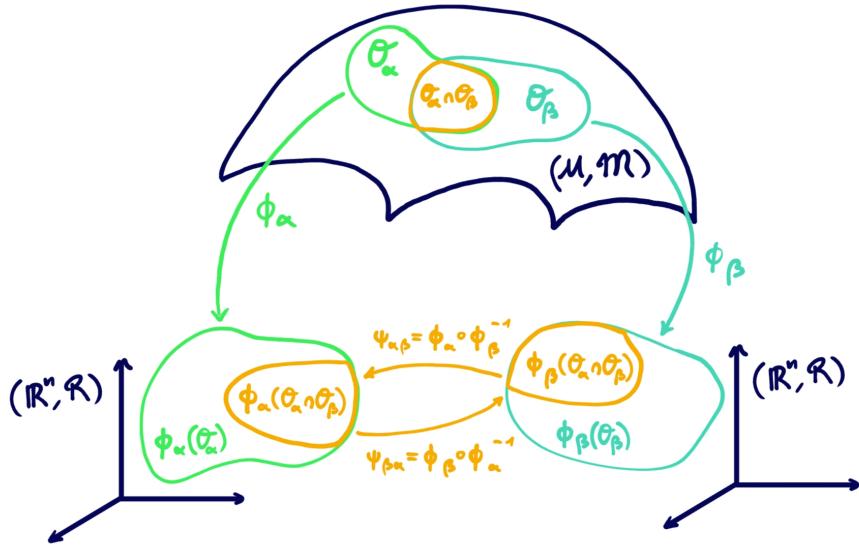


Abbildung 5.4: Kartenwechsel zwischen zwei Karten

Seien drei offene Mengen $O_\alpha, O_\beta, O_\gamma \in \mathfrak{M}$, die homöomorph zum \mathbb{R}^n sind, mit

$$O_\alpha \cap O_\beta \cap O_\gamma \neq \emptyset,$$

so können drei Kartenwechsel definiert werden:

$$\begin{aligned}\psi_{\beta\alpha} &= \phi_\beta \circ \phi_\alpha^{-1}, \\ \psi_{\gamma\beta} &= \phi_\gamma \circ \phi_\beta^{-1}, \\ \psi_{\gamma\alpha} &= \phi_\gamma \circ \phi_\alpha^{-1}.\end{aligned}$$

Proposition 5.8 (Dreifachdurchschnittsbedingung). Das Diagramm der Kartenwechsel kommutiert, d.h.

$$\psi_{\gamma\alpha} = \psi_{\gamma\beta} \circ \psi_{\beta\alpha} \quad (5.1)$$

Beweis. Es gilt

$$\psi_{\gamma\beta} \circ \psi_{\beta\alpha} = (\phi_\gamma \circ \phi_\beta^{-1}) \circ (\phi_\beta \circ \phi_\alpha^{-1}) \stackrel{\text{assoziativ}}{=} \underbrace{\phi_\gamma \circ (\phi_\beta^{-1} \circ \phi_\beta)}_{=\text{id}} \circ \phi_\alpha^{-1} = \phi_\gamma \circ \phi_\alpha^{-1} = \psi_{\gamma\alpha}$$

□

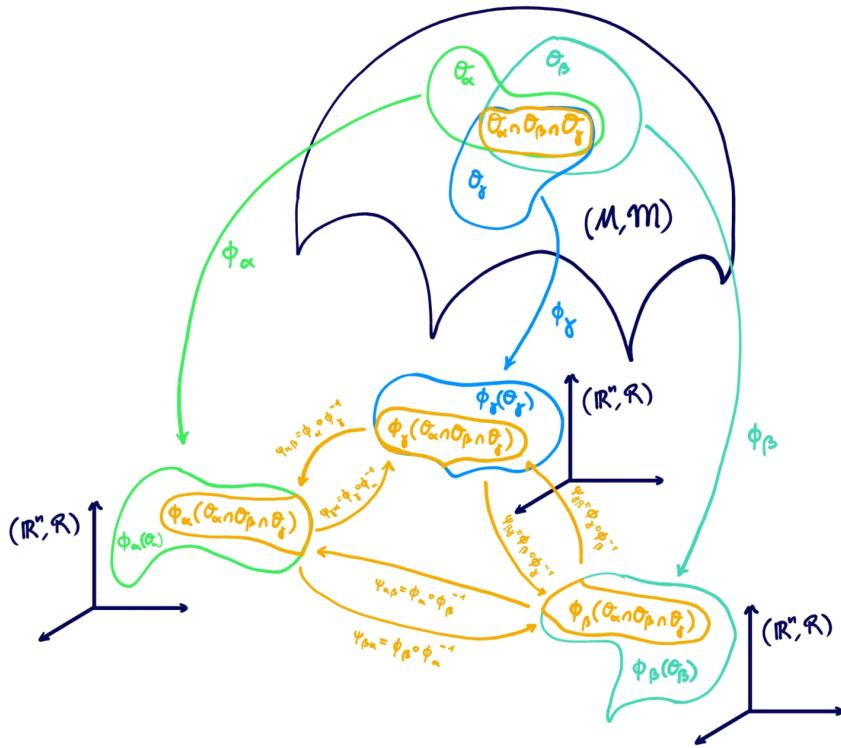


Abbildung 5.5: Kartenwechsel zwischen drei Karten

Diese Dreifachdurchschnittsbedingung lässt sich auf beliebige Mehrfachüberlappungen verallgemeinern.

Die Vierfachdurchschnittsbedingung ist $\psi_{\alpha\delta} = \psi_{\alpha\beta} \circ (\psi_{\beta\gamma} \circ \psi_{\gamma\delta})$. Diese folgt direkt aus der Dreifachdurchschnittsbedingung:

$$\psi_{\alpha\beta} \circ (\psi_{\beta\gamma} \circ \psi_{\gamma\delta}) \stackrel{(5.1)}{=} \psi_{\alpha\beta} \circ \psi_{\beta\delta} \stackrel{(5.1)}{=} \psi_{\alpha\delta}$$

Das kann induktiv weitergeführt werden zur Mehrfachdurchschnittsbedingung

$$\psi_{\alpha_0 \alpha_k} = \psi_{\alpha_0 \alpha_1} \circ \psi_{\alpha_1 \alpha_2} \circ \cdots \circ \psi_{\alpha_{k-1} \alpha_k}$$

Bemerkung 5.9. Eine Karte beschreibt, wie O im \mathbb{R}^n dargestellt wird. Beispielsweise durch kartesische Koordinaten $\phi_\alpha(p) = (x, y, z)$ oder durch Zylinderkoordinaten $\phi_\beta(p) = (r, \phi, z)$. Der Kartenwechsel $\psi_{\alpha\beta}$ ist dann die Transformation von den Zylinderkoordinaten in kartesische Koordinaten: $\psi_{\alpha\beta}(r, \phi, z) = (r \cdot \cos\phi, r \cdot \sin\phi, z) = (x, y, z)$.

Kartenwechsel in Koordinaten

Sei $p \in O_\alpha \cap O_\beta$. Dann besitzt p bezüglich der Karten (O_α, ϕ_α) und (O_β, ϕ_β) die Koordinatendarstellungen

$$x_\alpha^i := \phi_\alpha^i(p) \in \mathbb{R}^n \quad \text{und} \quad x_\beta^i := \phi_\beta^i(p) \in \mathbb{R}^n$$

Obwohl x_α und x_β im Allgemeinen verschieden sind, repräsentieren beide denselben Punkt der Mannigfaltigkeit

$$\phi_\alpha^{-1}(x_\alpha^i) = \phi_\beta^{-1}(x_\beta^i) = p \in M$$

Zwischen diesen Koordinaten besteht eine bijektive Beziehung, gegeben durch die Kartenwechselabbildungen

$$x_\alpha^i = \psi_{\alpha\beta}^i(x_\beta^k) = (\phi_\alpha \circ \phi_\beta^{-1})^i(x_\beta^k) \quad \text{und} \quad x_\beta^i = \psi_{\beta\alpha}^i(x_\alpha^k) = (\phi_\beta \circ \phi_\alpha^{-1})^i(x_\alpha^k)$$

Statt der Kartenwechselabbildung (Koordinatenwechselfunktion) können die Koordinaten x_α^i auch als Funktion der Koordinaten einer anderen Karte x_β^k aufgefasst werden

$$\begin{aligned} x_\beta^k(x_\alpha^i) &:= \psi_{\beta\alpha}^k(x_\alpha^i) = x_\beta^k \\ x_\alpha^k(x_\beta^i) &:= \psi_{\alpha\beta}^k(x_\beta^i) = x_\alpha^k \end{aligned}$$

Definition 5.10. Für zwei Karten (O, ϕ) und $(\tilde{O}, \tilde{\phi})$ mit $O \cap \tilde{O} \neq \emptyset$ ist der **Kartenwechsel in Koordinatendarstellung** gegeben durch

$$\psi : \phi(O \cap \tilde{O}) \subseteq \mathbb{R}^n \rightarrow \tilde{\phi}(O \cap \tilde{O}) \subseteq \mathbb{R}^n : x^k \mapsto \tilde{x}^i(x^k) := \psi^i(x^k) = \tilde{x}^i$$

5.3 Skalarfelder auf Mannigfaltigkeiten

Da wir nun Mannigfaltigkeiten kennen, können wir Skalarfelder definieren.

Definition 5.11. Sei (M, \mathfrak{M}) eine Mannigfaltigkeit. Eine Abbildung

$$F : M \rightarrow \mathbb{R}$$

heißt **reellwertiges Skalarfeld** auf M .

Skalarfelder bilden also jeden Punkt p einer Mannigfaltigkeit auf einen *Skalar* $F(p)$ ab. Dabei kann es sich beispielsweise um ein Temperaturfeld handeln, das jedem Punkt im Raum eine Temperatur zuweist.

Ein Skalarfeld F ist zunächst nur als abstrakte Abbildung auf der Mannigfaltigkeit definiert. Um ein Skalarfeld rechnerisch zu untersuchen, betrachten wir daher seine Darstellung in lokalen Koordinaten, die durch die Karten der Mannigfaltigkeit gegeben sind.

Definition 5.12. Sei (M, \mathfrak{M}) eine Mannigfaltigkeit, $F : M \rightarrow \mathbb{R}$ ein Skalarfeld und (O_α, ϕ_α) eine Karte der Mannigfaltigkeit. Dann heißt die numerische Funktion

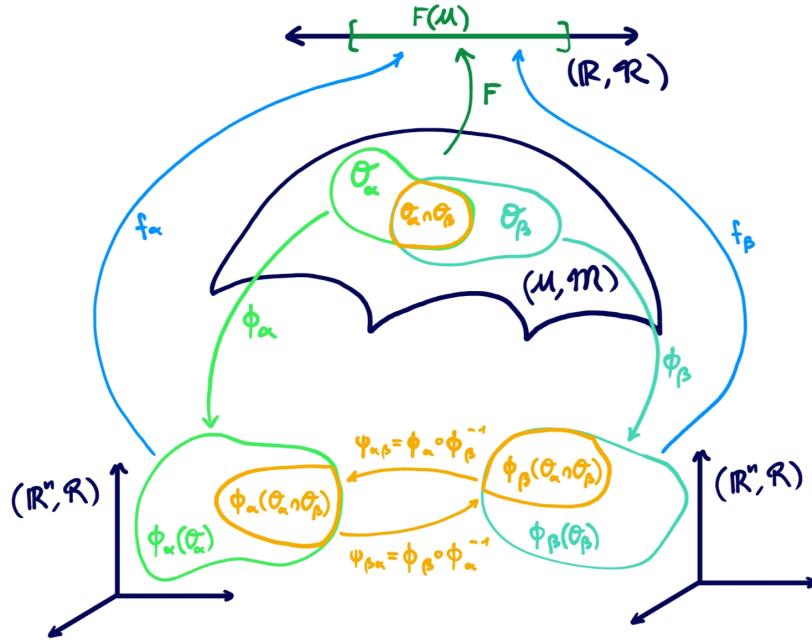
$$f_\alpha := F \circ \phi_\alpha^{-1} : \phi_\alpha(O_\alpha) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$$

lokaler Repräsentant oder auch **lokale Darstellung** von F in der α -Karte.

f_α ist keine Koordinatentransformation, sondern die Darstellung eines Skalarfeldes in den Koordinaten der Karte (O_α, ϕ_α) . Ob diese Koordinaten kartesisch, zylindrisch oder sphärisch sind, hängt allein von der gewählten Karte ab.

Entsprechend erhält man zu einem Atlas $(O_\alpha, \phi_\alpha)_{\alpha \in I}$ eine Familie numerischer Funktionen

$$(f_\alpha)_{\alpha \in I} \text{ mit } f_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$$



Sollen lokale Darstellungen $(f_\alpha)_{\alpha \in I}$ ein gemeinsames (globales) Skalarfeld F beschreiben, dürfen sie nicht unabhängig voneinander gewählt werden. Überlappen sich zwei Karten O_α und O_β , so müssen die entsprechenden Funktionen auf dem Überlappungsbereich zueinander passen.

Proposition 5.13 (Kompatibilitätsbedingung). Seien (O_α, ϕ_α) und (O_β, ϕ_β) zwei Karten mit $O_\alpha \cap O_\beta \neq \emptyset$. Seien f_α und f_β lokale Darstellungen desselben Skalarfeldes F . Dann gilt

$$\forall x \in \phi_\beta(O_\alpha \cap O_\beta) : f_\beta(x) = (f_\alpha \circ \psi_{\alpha\beta})(x) \quad (5.2)$$

wobei $\psi_{\alpha\beta} := \phi_\alpha \circ \phi_\beta^{-1}$

Beweis. Dies geht aus der Definition der lokalen Darstellung hervor.

Sei $x \in \phi_\beta(O_\alpha \cap O_\beta)$. Dann ist

$$f_\beta(x) = (F \circ \phi_\beta^{-1})(x) = (F \circ \underbrace{(\phi_\alpha^{-1} \circ \phi_\alpha)}_{=\text{id}} \circ \phi_\beta^{-1})(x) = ((F \circ \phi_\alpha^{-1}) \circ (\phi_\alpha \circ \phi_\beta^{-1}))(x) = (f_\alpha \circ \psi_{\alpha\beta})(x)$$

□

Die beiden lokalen Darstellungen f_α und f_β beschreiben denselben physikalischen/geometrischen Wert, nur in verschiedenen Koordinaten, da sie vom selben globalen Skalarfeld F stammen.

Skalarfelder sind kartensunabhängig (koordinatenunabhängig)

Die lokale Darstellung f_α eines Skalarfeldes F erlaubt es, den abstrakten Begriff einer Funktion auf einer Mannigfaltigkeit auf bekannte numerische Funktionen zurückzuführen. Dabei ist zu beachten, dass ein Punkt $p \in M$ je nach Karte durch unterschiedliche Koordinaten beschrieben wird.

Sei $p \in O_\alpha \cap O_\beta$. Dann besitzt p bezüglich der Karten (O_α, ϕ_α) und (O_β, ϕ_β) die Koordinatendarstellungen, die mit der Kartenwechselabbildung (Koordinatenwechselfunktion) [Definition 5.10](#) ineinander umgerechnet

werden können

$$\begin{aligned}x_\beta^k(x_\alpha^i) &= x_\beta^k \\x_\alpha^k(x_\beta^i) &= x_\alpha^k\end{aligned}$$

Die Kompatibilitätsbedingung der lokalen Darstellungen stellt sicher, dass die numerischen Funktionen f_α und f_β denselben Funktionswert liefern, wenn sie auf Koordinaten angewandt werden, die demselben Punkt p entsprechen:

$$f_\alpha(x_\alpha^i) = f_\alpha(\psi_{\alpha\beta}^i(x_\beta^k)) = (f_\alpha \circ \psi_{\alpha\beta})(x_\beta^k) \stackrel{(5.2)}{=} f_\beta(x_\beta^k) \quad (5.3)$$

Dies entspricht der physikalischen Forderung, dass eine skalare Größe – etwa eine Temperatur – unabhängig von der gewählten Karte denselben Wert besitzt. Die numerischen Funktionen hängen von der gewählten Karte (Koordinatendarstellung) ab, der physikalische Wert jedoch nicht.

Rekonstruktion des Skalarfeldes aus der lokalen Darstellung

Umgekehrt kann zu gegebenen lokalen Darstellungen $(f_\alpha)_{\alpha \in I}$, $f_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ das zugehörige Skalarfeld F gewonnen werden.

Für einen Punkt $p \in M$ existiert eine Karte (O_α, ϕ_α) mit $p \in O_\alpha$, sodass wir definieren können

$$F(p) := f_\alpha(\phi_\alpha^i(p)) = f_\alpha(x_\alpha^i)$$

Liegt der Punkt p zusätzlich in einer weiteren Karte (O_β, ϕ_β) , so können wir ebenso definieren

$$F(p) := f_\beta(\phi_\beta^i(p)) = f_\beta(x_\beta^i) \stackrel{(5.3)}{=} (f_\alpha \circ \psi_{\alpha\beta})(x_\beta^i) = f_\alpha(x_\alpha^i)$$

Da per Definition der Mannigfaltigkeit für jeden Punkt $p \in M$ eine Karte existiert, gilt obiges für alle $p \in M$. Der Wert des Skalarfeldes ist somit unabhängig von der gewählten Karte und eindeutig durch die lokalen Darstellungen bestimmt.

Diese Konstruktion ist analog zur Darstellung eines Vektors in verschiedenen Basen: Die einzelnen Komponenten hängen von der gewählten Basis ab, der Vektor selbst, als abstraktes Objekt, jedoch nicht. Ebenso hängen die lokalen Darstellungen f_α von der Wahl der Karte ab, während das Skalarfeld F kartensunabhängig ist.

Mit der Kartenwechselabbildung (Koordinatenwechselfunktion) (5.10) kann (5.3) nun folgendermaßen dargestellt werden

$$f_\alpha(x_\alpha^i) = f_\beta(x_\beta^j) = f_\beta(x_\beta^j(x_\alpha^i))$$

Definition 5.14. Eine Mannigfaltigkeit heißt **C^k -differenzierbare Mannigfaltigkeit**, wenn alle Kartenwechselabbildungen ψ k -mal stetig differenzierbar sind, d.h. $\psi \in C^k$.

Eine Mannigfaltigkeit heißt **glatte Mannigfaltigkeit** oder **C^∞ -Mannigfaltigkeit**, wenn alle Kartenwechselabbildungen unendlich oft stetig differenzierbar sind, d.h. $\psi \in C^\infty$.

Wir fordern hier glatte Mannigfaltigkeiten, da andernfalls auch die lokalen Darstellungen von Skalarfeldern nur C^k wären, da sie über

$$f_\alpha = F \circ \phi_\alpha^{-1}$$

definiert sind. D.h. das Skalarfeld wäre lokal nur k -mal stetig differenzierbar.

Als anschauliches Beispiel für ein Skalarfeld betrachten wir eine Wetterkarte. Bei einer solchen wird jedem Punkt eines zweidimensionalen Raumes (der Landkarte) ein Skalar zugeordnet, beispielsweise die Temperatur

oder der Luftdruck. Formal handelt es sich dabei um eine Abbildung, die jedem Punkt einen reellen Wert zuweist.

Im Alltag kann eine Wetterkarte meist sinnvoll interpretiert werden, ohne die gekrümmte Gestalt der Erdoberfläche, also die zweidimensionale Mannigfaltigkeit, zu berücksichtigen.

5.3.1 Skalarfelder und Karten am Beispiel der stereographischen Projektion

Als Beispiel betrachten wir die Kugeloberfläche S^2 (zweidimensionale Mannigfaltigkeit). Zur Beschreibung von S^2 genügen bereits zwei Karten, die gemeinsam einen Atlas bilden: die Nordpolprojektion und die Südpolprojektion.

Jede dieser Projektionen bildet alle Punkte der Kugeloberfläche, mit Ausnahme des jeweiligen Projektionspols, auf eine Ebene ab, die am gegenüberliegenden Pol tangential anliegt.

Die Abbildung erfolgt dabei folgendermaßen: Die direkte Verbindungsline zwischen dem Projektionspol und einem Punkt $p \in S^2$ wird so weit verlängert, bis sie die entsprechende Tangentialebene schneidet. Der Schnittpunkt ist das Bild von p in der Ebene.

Damit gilt:

- Die Nordpolprojektion ϕ_N (vom Norpol ausgehend) bildet $S^2 \setminus \{N\}$ auf eine Ebene in $\mathbb{R}^2 \setminus \{N\}$ ab.
- Die Südpolprojektion ϕ_S (vom Südpol ausgehend) bildet $S^2 \setminus \{S\}$ auf eine Ebene in $\mathbb{R}^2 \setminus \{S\}$ ab.

Zusammen beschreiben diese beiden Karten die gesamte Mannigfaltigkeit S^2 und bilden somit einen Atlas.

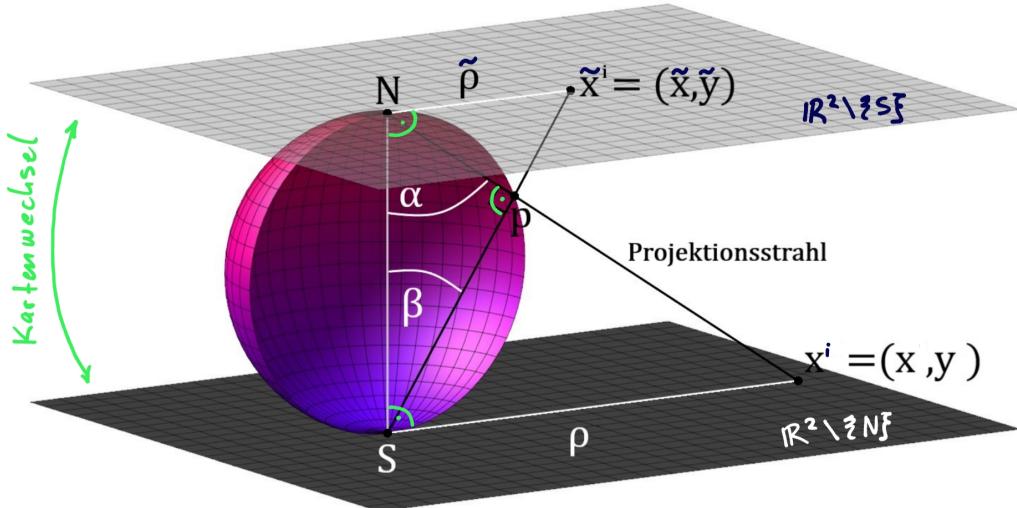


Abbildung 5.6: stereographische Projektion in 3D Ansicht

Quelle: ART Mitschrift Berndl-Forstner (WS20/21)

<https://forum.fstph.at/t/latex-skriptum-allg-relativitaetstheorie/4021>

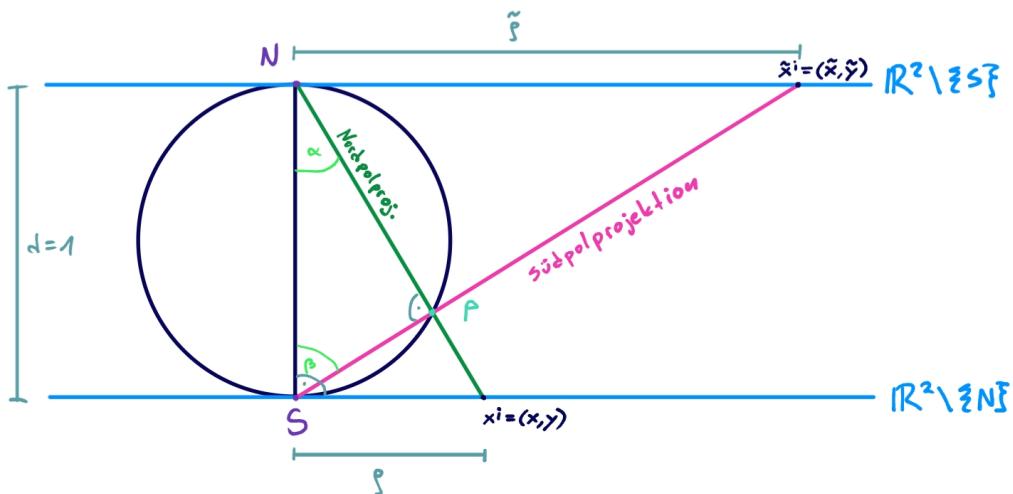
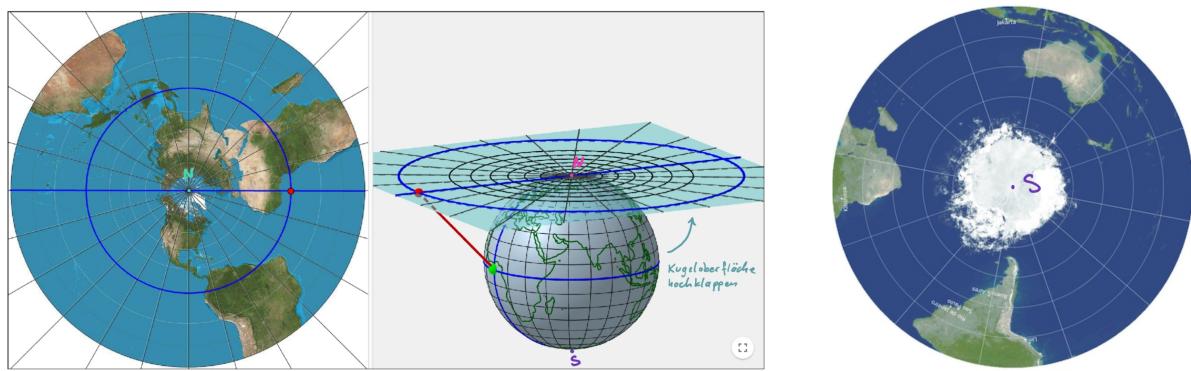


Abbildung 5.7: stereographische Projektion in 2D Ansicht



(a) Südpolprojektion

(b) Nordpolprojektion

Quelle: <https://www.geogebra.org/m/d6jucfnc>

Abbildung 5.8: Projektion der Erde



Abbildung 5.9: Nordpolprojektion mithilfe von Licht

Herleitung der Gestalt der Kartenwechselabbildung

Sei $p \in S^2$. Dann ist $\phi_N(p) = x^i \in \mathbb{R}^2 \setminus \{N\}$ und $\phi_S(p) = \tilde{x}^i \in \mathbb{R}^2 \setminus \{S\}$. Für die Koordinaten von x^i und \tilde{x}^i verwenden wir Polarkoordinaten. Dabei ist der Winkel (Längengrad) φ für beide Projektionen identisch. Der Unterschied zwischen den beiden Darstellungen liegt ausschließlich im radialen Abstand ρ bzw. $\tilde{\rho}$, zum Süd-

bzw. Nordpol, welchen wir als Koordinatenursprung für die jeweilige Ebene festlegen. Um die Rechnungen zu vereinfachen, setzen wir den Durchmesser der Kugel auf $d = 1$.

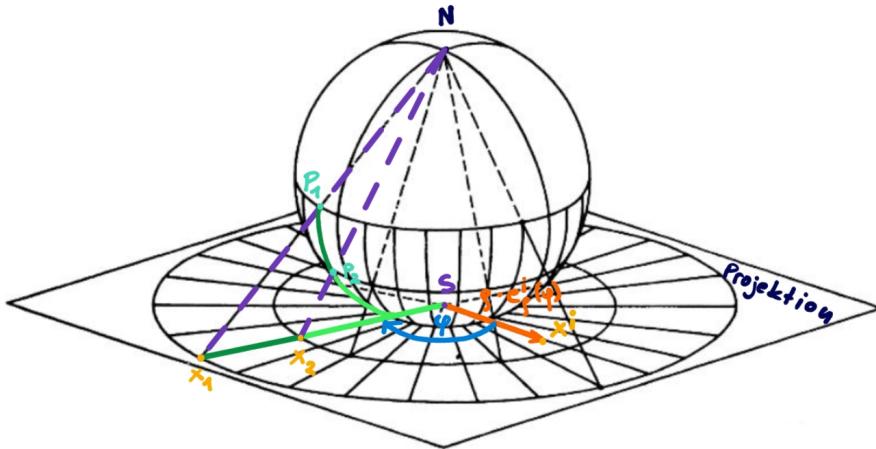


Abbildung 5.10: Die Position eines Punktes x^i in der projizierten Ebene wird durch den Winkel φ und Radius ρ beschrieben: $\rho \cdot e_\rho^i(\varphi) = x^i$

Aus der Geometrie ergibt sich

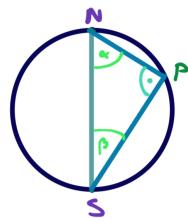
$$\tan \alpha = \frac{\rho}{d} = \frac{\rho}{1} = \rho, \quad \tan \beta = \tilde{\rho}$$

Da p auf der Kugeloberfläche liegt, liegt dieser Punkt auf einem Thaleskreis (verbindet man einen bel. Punkt auf einem Halbkreis mit dessen Endpunkten schließen die beiden Strecken einen rechten Winkel ein). Daher stehen die Verbindungsstrecken \overline{Np} und \overline{Sp} im rechten Winkel aufeinander. Da die Summe der Winkel in einem Dreieck $180^\circ = \pi$ beträgt, ist somit $\beta = \pi - \frac{\pi}{2} - \alpha = \frac{\pi}{2} - \alpha$. Somit folgt

$$\tilde{\rho} = \tan \beta = \tan\left(\frac{\pi}{2} - \alpha\right) = \frac{1}{\tan \alpha} = \frac{1}{\rho} \quad (5.4)$$



Thaleskreis



Sei $e_\rho^i(\varphi) = (\cos \varphi, \sin \varphi)$ der Einheitsvektor in radialer Richtung im \mathbb{R}^2 . Dann sind die Koordinaten des Bildpunktes bezüglich der Nordpolprojektion

$$\phi_N(p) = x^i = \rho e_\rho^i(\varphi)$$

Es folgt die Kartenwechselabbildung zwischen Nord- und Südpolprojektion

$$x^i = \rho e_\rho^i(\varphi) \stackrel{(5.4)}{=} \frac{1}{\tilde{\rho}} e_\rho^i(\varphi) \stackrel{\frac{1}{\rho} \cdot \tilde{\rho}^i = e_\rho^i(\varphi)}{=} \frac{1}{\tilde{\rho}^2} \tilde{\rho}^i \stackrel{\tilde{\rho}^2 = |\tilde{x}^m|^2}{=} \frac{1}{|\tilde{x}^m|^2} \tilde{x}^i \quad \tilde{x}^i \in \mathbb{R}^2 \setminus \{0\}.$$

und

$$\tilde{x}^i = \frac{1}{x^m x^m} x^i, \quad x^i \in \mathbb{R}^2 \setminus \{0\}.$$

Zusammengefasst: Auf dem Überlappungsbereich der beiden Karten sind die Koordinaten durch die Kartenwechselabbildung miteinander verknüpft. Da dort $\tilde{\rho} = \frac{1}{\rho}$ gilt, können die beiden Darstellungen ineinander umgerechnet werden.

Darstellung eines Skalarfeldes

Wir betrachten nun ein Skalarfeld $F : S^2 \rightarrow \mathbb{R}$ auf der Mannigfaltigkeit S^2 . Bezüglich der Nordpolprojektion ϕ_N sei das Skalarfeld lokal dargestellt durch

$$f := F \circ \phi_N^{-1} : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x^i) := \frac{1}{1 + \rho^2} = \frac{1}{1 + x^i x^i}$$

Da es sich um ein Skalarfeld handelt, muss es auch bezüglich der anderen Karte darstellbar sein. Bezeichnet \tilde{f} die lokale Darstellung bezüglich der Südpolprojektion ϕ_S

$$\tilde{f} := F \circ \phi_S^{-1} : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \tilde{f}(x^i) := \frac{1}{1 + \tilde{\rho}^2} = \frac{1}{1 + \tilde{x}^i \tilde{x}^i}$$

Dann muss für $p \in S^2$ gelten

$$F(p) = \tilde{f}(\tilde{x}^i) = f(x^i)$$

und es kann die Darstellung des Skalarfeldes bezüglich der anderen Karten angeschrieben werden

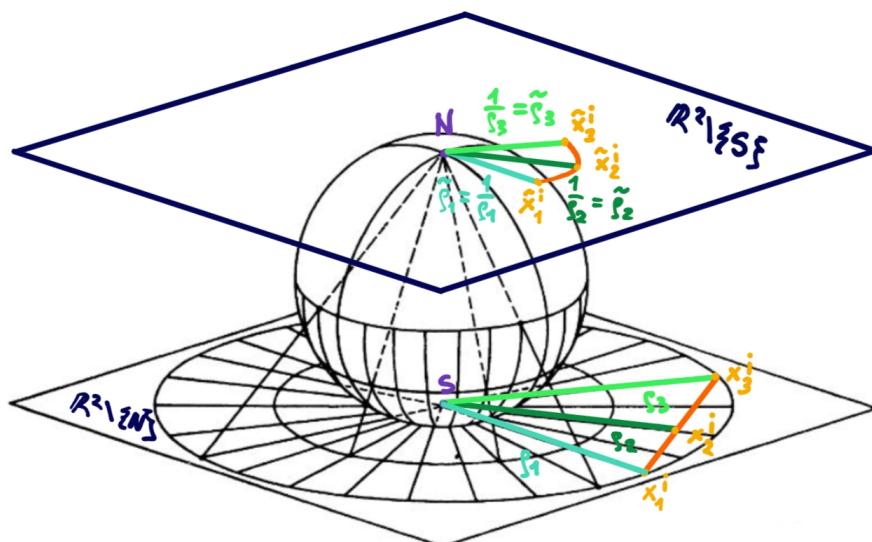
$$\tilde{f}(\tilde{x}^i) = f(x^i) = \frac{1}{1 + \rho^2} \stackrel{(5.4)}{=} \frac{1}{1 + \frac{1}{\tilde{\rho}^2}} = \frac{\tilde{\rho}^2}{1 + \tilde{\rho}^2} = \frac{\tilde{x}^i \tilde{x}^i}{1 + \tilde{x}^i \tilde{x}^i}$$

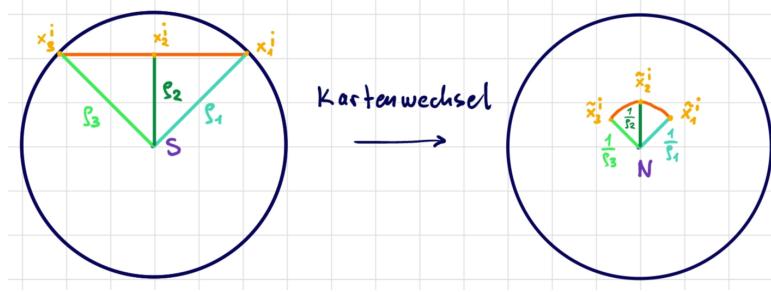
Für $\rho \rightarrow \infty$ geht die Funktion $f(x^i) = \frac{1}{1+\rho^2}$ gegen 0. Dieser Grenzfall entspricht geometrisch dem Annähern an den Nordpol. (Der Nordpol ist hier wie der Punkt im unendlichen bei der Riemannschen Zahlenkugel.)

Da die Nord- und Südpolprojektion zusammen einen Atlas der Kugeloberfläche bilden, existiert ein globales Skalarfeld F auf S^2 , dessen lokale Darstellungen durch f und \tilde{f} gegeben sind.

Die Mannigfaltigkeit muss nicht als eingebettetes Objekt im \mathbb{R}^3 aufgefasst werden. Die Beschreibung durch Karten und Kartenwechsel enthält bereits die vollständige geometrische Information über das Skalarfeld auf S^2 .

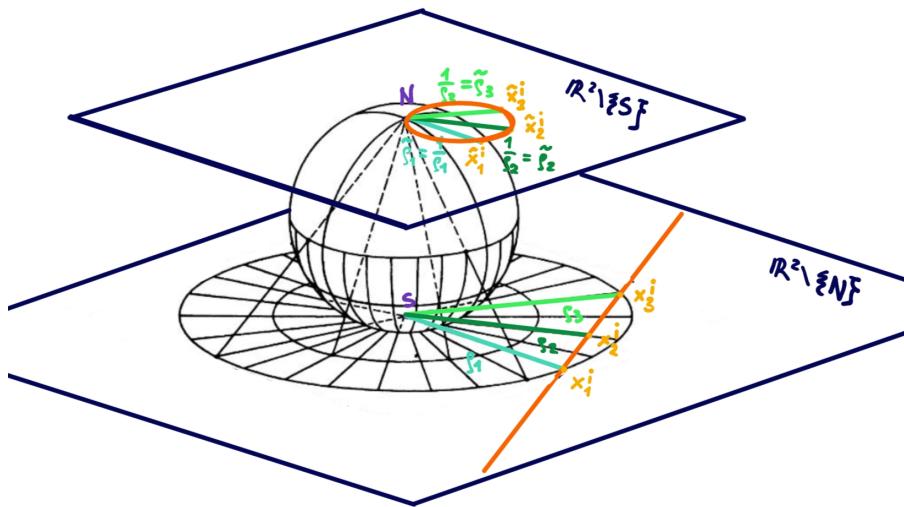
Anhand der stereographischen Projektion erkennt man sehr deutlich, dass verschiedene Karten eine und dieselbe Mannigfaltigkeit in den \mathbb{R}^n auf sehr unterschiedliche Weise abbilden können. Obwohl stets dieselbe Punktmenge der Mannigfaltigkeit (M, \mathfrak{M}) beschrieben wird, können die Bilder im Zielraum völlig verschieden aussehen. Vgl. Abbildung 5.8 bzw. auch die Tatsache, dass die Kartenwechselabbildung nicht linear ist: Wenn wir eine Gerade in $\mathbb{R}^2 \setminus N$ durch die Kartenwechselabbildung $\psi : x^i = \rho e_\rho^i(\varphi) \mapsto \tilde{x}^i = \frac{1}{\rho} e_\rho^i(\varphi)$ in $\mathbb{R}^2 \setminus S$ abbilden, wird diese Gerade gekrümmmt.





Wenn wir diese Gerade nun ins Unendliche verlängern, wird der Abstand ρ der Punkte auf der Geraden zu S immer größer und somit $\tilde{\rho} = \frac{1}{\rho}$ immer kleiner. Es bildet sich in $\mathbb{R}^2 \setminus S$ ein Kreis. Beim Grenzübergang $\rho \rightarrow \infty$ gilt $\lim_{\rho \rightarrow \infty} \tilde{\rho} = 0$, wir treffen auf den Nordpol.

Das heißt, jede Gerade wird durch die Kartenwechselabbildung auf einen Kreis abgebildet, der durch den Pol verläuft. Außerdem werden Kurven auf Kurven abgebildet.



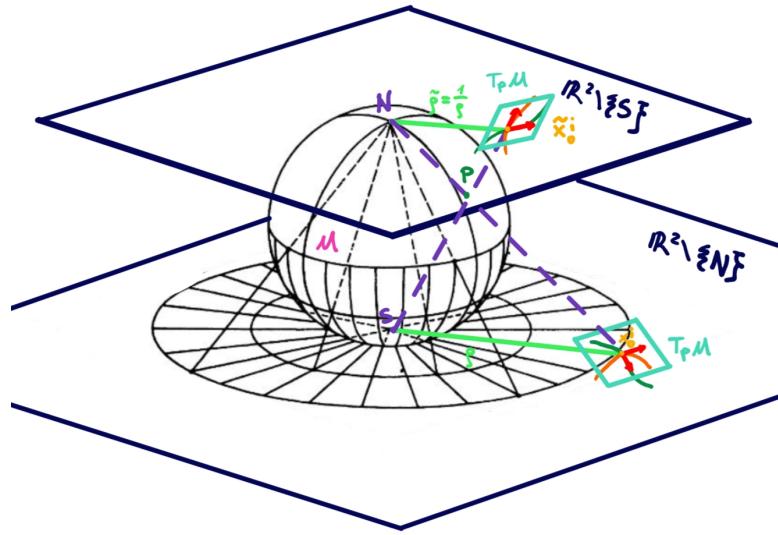
5.4 Vektoren auf Mannigfaltigkeiten (Tangentialraum)

Ziel dieser Sektion ist es, Vektorfelder bzw. allgemein Tensorfelder auf Mannigfaltigkeiten zu definieren. Vektorfelder bzw. Tensorfelder auf dem \mathbb{R}^n bilden auf einen Vektorraum V bzw. Tensorproduktraum T_s^r ab

$$v : \mathbb{R}^n \rightarrow V \quad t : \mathbb{R}^n \rightarrow T_s^r$$

Wenn wir dies auf allgemeine Mannigfaltigkeiten übertragen möchten, benötigen wir einen Vektorraum bzw. Tensorraum auf der Mannigfaltigkeit, in den die Vektorfelder bzw. Tensorfelder abgebildet werden.

Wie wir anhand der stereographischen Projektion gesehen haben, ist die Kartenwechselabbildung im allgemeinen nicht linear. Somit wird die Vektorraumstruktur beim Kartenwechsel nicht erhalten. Es stellt sich nun die Frage, wie wir den Kartenwechsel linearisieren können. Der Kartenwechsel bildet Kurven auf Kurven ab. Wir betrachten nun mehrere Kurven, die durch einen fixierten Punkt $p \in M$ bzw. in Koordinatendarstellungen x_0^i in einer Karte verlaufen und dessen Tangentialvektoren $v^i := \dot{x}_0^i$ in diesem Punkt. Wie sich herausstellen wird, spannen diese Tangentialvektoren einen Vektorraum auf, den *Tangentialraum*. Wir erhalten also für jeden Punkt $p \in M$ einen Vektorraum. Das Vektorraumkonzept wird lokalisiert.

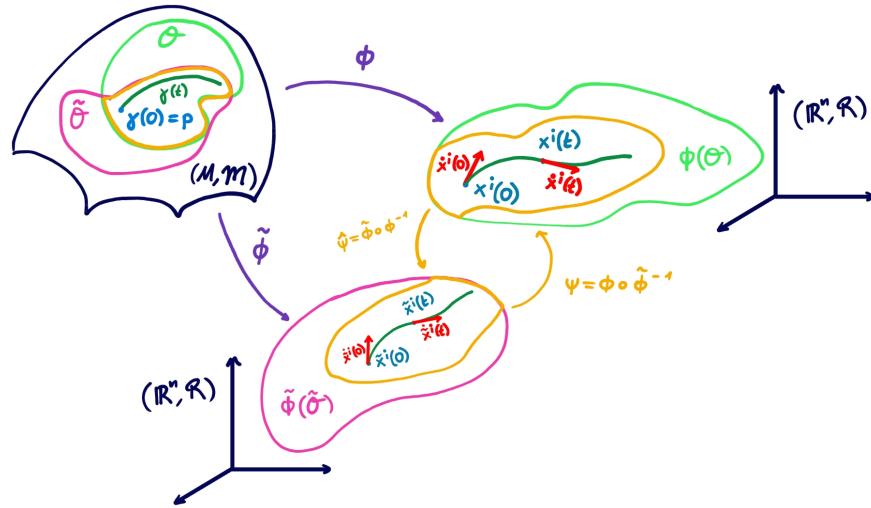


Sei nun (M, \mathfrak{M}) eine beliebige Mannigfaltigkeit. Wir betrachten eine Kurve $\gamma : I \rightarrow M$ mit $I \subset \mathbb{R}$ beliebiges Intervall, die bei $p = \gamma(0)$ differenzierbar ist. Wir wählen zwei Karten (O, ϕ) , $(\tilde{O}, \tilde{\phi})$ mit $\gamma(I) \subset O \cap \tilde{O}$, dann erhält die Kurve in Koordinaten die Darstellung

$$x^i(t) = \phi^i(\gamma(t)) \quad \tilde{x}^i(t) = \tilde{\phi}^i(\gamma(t))$$

Dann ist der Tangentialvektor der Kurve in der jeweiligen Karte an $t = 0$ gegeben durch:

$$v^i := \frac{d}{dt} \Big|_{t=0} x^i(t) = \dot{x}^i(0) \quad \tilde{v}^i := \frac{d}{dt} \Big|_{t=0} \tilde{x}^i(t) = \dot{\tilde{x}}^i(0)$$



Wir suchen nun nach einer linearen Beziehung zwischen v^i und \tilde{v}^i .

Auf dem Überlappungsbereich hängen die Koordinaten über den Kartenwechsel zusammen. Wie in [Definition 5.10](#) beschrieben können wir die „neuen“ Koordinaten als Funktionen der „alten“ auffassen:

$$\tilde{x}^i(x^k) := \psi^i(x^k) = \tilde{x}^i \in \phi^k(O \cap \tilde{O})$$

Wenden wir nun die Kettenregel auf $\tilde{x}^i(t) = \tilde{x}^i(x^k(t))$ an, so erhalten wir:

$$\tilde{v}^i = \frac{d}{dt} \Big|_{t=0} \tilde{x}^i(t) = \frac{d}{dt} \Big|_{t=0} \tilde{x}^i(x(t)) = \underbrace{\frac{\partial \tilde{x}^i}{\partial x^l}(x^k(0))}_{=:x_0^k} v^l = \frac{\partial \tilde{x}^i}{\partial x^l}(x_0^k) v^l \quad (5.5)$$

Die Matrix $\frac{\partial \tilde{x}^i}{\partial x^l}$ ist die **Jacobi-Matrix** des Kartenwechsels. Sie gibt an, wie sich v^i unter einem Kartenwechsel verändert und ist **linear**. Somit ist diese Transformation linear in v . Sie hängt nur vom Punkt p (bzw. x_0) ab, nicht von der konkreten Kurve γ . Wichtig: wir beschränken uns hier auf den einen Punkt x_0^i ! Global bleibt der Kartenwechsel im Allgemeinen nichtlinear.

Es kann viele verschiedene Kurven geben, die durch denselben Punkt p bzw. in Koordinaten x_0^i laufen und dort dieselbe „Anfangsrichtung“ v^i besitzen. Wir fassen diese zusammen mit der folgenden **Äquivalenzrelation**:

$$x_1^i(t) \sim x_2^i(t) : \Leftrightarrow \begin{cases} x_1^i(0) = x_2^i(0) =: x_0^i \\ \dot{x}_1^i(0) = \dot{x}_2^i(0) = v^i \end{cases}$$

Die Äquivalenzklasse ist dann $[x^i(t)]_\sim = \{y^i(t) \mid y^i(0) = x_0^i, \dot{y}^i(0) = v^i\}$.

Diese Äquivalenz ist **kartenunabhängig** und hängt nur von $p \in M$ ab. D.h. $x_1^i(t) \sim x_2^i(t) \Leftrightarrow \tilde{x}_1^i(t) \sim \tilde{x}_2^i(t)$. Das folgt aus dem Transformationsgesetz (5.5): Wenn zwei Kurven in einer Karte denselben Tangentialvektor besitzen, dann haben sie wegen der linearen Jacobi-Transformation auch in jeder anderen Karte denselben Tangentialvektor:

Transformieren wir in x_0^k die Tangentialvektoren $v_1^i = v_2^i$ zweier Kurven mit (5.5)

$$\tilde{v}_1^i = \frac{\partial \tilde{x}^i}{\partial x^l}(x_0^k) v_1^l \quad \tilde{v}_2^i = \frac{\partial \tilde{x}^i}{\partial x^l}(x_0^k) v_2^l$$

Da die Jacobi-Matrix nur von x_0^k bzw. $p \in M$ abhängt, wird dieselbe Matrix auf $v_1^i = v_2^i$ angewandt und wir erhalten $\tilde{v}_1^i = \tilde{v}_2^i$.

Koordinatenunabhängig geschrieben:

$$\gamma_1(t) \sim \gamma_2(t) : \Leftrightarrow \begin{cases} \gamma_1(0) = \gamma_2(0) \\ \dot{\gamma}_1(0) = \dot{\gamma}_2(0) \end{cases}$$

Definition 5.15. Der **Tangentialraum** an einem Punkt $p \in M$ ist die Menge aller obiger Äquivalenzklassen von Kurven γ durch p , die in p differenzierbar sind:

$$T_p M := \{[\gamma]_\sim \mid \gamma : I \rightarrow M \text{ differenzierbar in } p \in M, \gamma(0) = p\}$$

Vektorraumstruktur des Tangentialraums

Damit der Tangentialraum zu einem Vektorraum wird, muss er bzgl. der Addition und der Skalarmultiplikation abgeschlossen sein. Zudem muss die Addition und die Skalarmultiplikation unabhängig von der Wahl der Karte sein.

In einer fixierten Karte (O, ϕ) betrachten dazu zwei Kurven γ_x und γ_y mit Koordinaten $x^i(t)$ und $y^i(t)$ bzgl. dieser Karte, die beide durch den Punkt $\gamma_x(0) = \gamma_y(0) = p \in O \subset M$ bzw. in Koordinaten x_0^i verlaufen und in diesem Punkt die Tangentialvektoren

$$v = [\gamma_x]_\sim \in T_p M \text{ mit Komponenten } v^i := \dot{x}_0^i \in \mathbb{R}$$

$$w = [\gamma_y]_\sim \in T_p M \text{ mit Komponenten } w^i := \dot{y}_0^i \in \mathbb{R}$$

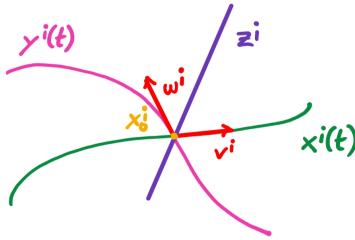
besitzen.

Um die Summe der beiden Tangentialvektoren zu definieren, konstruieren wir eine dritte Kurve γ_z mit Koordinaten $z^i(t)$, die ebenso durch den Punkt x_0^i verläuft

$$z^i(t) := x_0^i + t(v^i + w^i) \quad (\text{also eine Gerade})$$

Diese Kurve besitzt den Tangentialvektor

$$\dot{z}^i(0) = v^i + w^i$$



Wir definieren die Summe $v + w := [\gamma_z] \in T_p M$ als die Äquivalenzklasse der Kurve γ_z .

Es bleibt zu zeigen, dass diese Definition unabhängig von der Wahl der Karte ist. Dazu betrachten wir eine zweite Karte $(\tilde{O}, \tilde{\phi})$ mit $p \in \tilde{O}$ und die entsprechenden Koordinatendarstellungen nach dem Kartenwechsel ψ (siehe (5.4))

$$\begin{aligned}\tilde{x}^i(t) &= \psi^i(x^k(t)) = \tilde{x}^i(x^k(t)) \\ \tilde{y}^i(t) &= \psi^i(y^k(t)) = \tilde{y}^i(y^k(t)) \\ \tilde{z}^i(t) &= \psi^i(z^k(t)) = \tilde{z}^i(z^k(t))\end{aligned}$$

Nach dem Transformationsgesetz (5.5) für Tangentialvektoren gilt

$$\tilde{v}^i = \frac{\partial \tilde{x}^i}{\partial x^k}(x_0^m) v^k, \quad \tilde{w}^i = \frac{\partial \tilde{x}^i}{\partial x^k}(x_0^m) w^k,$$

Für die Kurve γ_z ergibt sich entsprechend

$$\tilde{z}^i(0) = \frac{\partial \tilde{x}^i}{\partial x^k}(x_0^m) (v^k + w^k) = \tilde{v}^i + \tilde{w}^i$$

Damit besitzt die Kurve $\tilde{z}^i(t)$ in der neuen Karte genau den Tangentialvektor, der der Summe der beiden transformierten Tangentialvektoren entspricht. Die Addition ist somit unabhängig von der Wahl der Karte.

Folglich ist die Addition von Tangentialvektoren eine koordinatenunabhängige Operation. Der Beweis für die Skalarmultiplikation wird mit $z^i(t) := x_0^i + (\lambda v^i)t$ geführt und geht analog.

Folglich ist der Tangentialraum $T_p M$ ein **Vektorraum**. Allerdings haben wir bisher eine **koordinaten- bzw. basisabhängige Darstellung** der Elemente von $T_p M$:

$v \in T_p M$ (Tangentialvektor) $\Leftrightarrow \exists$ eine Kurve in Koordinaten $x^i(t)$ mit $\phi^i(p) = x^i(0) =: x_0^i$ für ein $p \in O \subset M$ und $v^i = \dot{x}^i(0)$ und der Kartenwechseltransformation $\tilde{v}^i = \frac{\partial \tilde{x}^i}{\partial x^k}(x_0^m) v^k$.

Koordinaten- bzw. basisunabhängige Darstellung vom Tangentialraum

Schreiben wir den Tangentialvektor mit Hilfe eines abstrakten Index a , so ist der Vektor selbst unabhängig

von der Wahl der Karte. Bezuglich einer Karte (O, ϕ) lässt sich v^a als Linearkombination von Komponenten v^i und Basiselementen E_i^a darstellen:

$$v^a = v^i E_i^a \quad (5.6)$$

In einer anderen Karte $(\tilde{O}, \tilde{\phi})$ gilt entsprechend

$$v^a = \tilde{v}^i \tilde{E}_i^a$$

Die Komponenten eines Tangentialvektors transformieren beim Kartenwechsel gemäß (5.5)

$$\tilde{v}^i = \frac{\partial \tilde{x}^i}{\partial x^k}(x_0^m) v^k$$

Somit ist

$$v^a = \tilde{v}^i \tilde{E}_i^a = v^k \frac{\partial \tilde{x}^i}{\partial x^k}(x_0^m) \tilde{E}_i^a \stackrel{(5.6)}{=} v^i E_i^a$$

Also

$$\frac{\partial \tilde{x}^i}{\partial x^k}(x_0^m) \tilde{E}_i^a = E_i^a$$

bzw.

$$\tilde{E}_i^a = \frac{\partial x^k}{\partial \tilde{x}^i}(x_0^m) E_k^a \quad (5.7)$$

Also ist das Transformationsverhalten der Basisvektoren invers zum Transformationsverhalten der Tangentialvektoren.

Wir suchen nun ein natürliches Objekt mit genau diesem Transformationsverhalten. Betrachte dazu den Kartenwechsel

$$\tilde{x}^k(t) = \tilde{x}^k(x^i(t))$$

und dessen Ableitung

$$\partial_i := \frac{\partial}{\partial x^i} \stackrel{\text{Ketten-}}{\underset{\text{regel-}}{=}} \underbrace{\frac{\partial}{\partial \tilde{x}^k} \frac{\partial \tilde{x}^k}{\partial x^i}}_{=: \tilde{\partial}_k} = \tilde{\partial}_k \frac{\partial \tilde{x}^k}{\partial x^i}$$

Also

$$\partial_i|_{x_0^m} = \tilde{\partial}_k|_{x_0^m} \underbrace{\frac{\partial \tilde{x}^k}{\partial x^i}(x_0^m)}_{= \partial_i \tilde{x}^k|_{x_0^m}} \quad (5.8)$$

bzw.

$$\tilde{\partial}_i|_{x_0^m} = \partial_k|_{x_0^m} \underbrace{\frac{\partial x^k}{\partial \tilde{x}^i}(x_0^m)}_{= \tilde{\partial}_i x^k|_{x_0^m}} \quad (5.9)$$

Die partiellen Ableitungen sind immer bzgl. x_0^m zu verstehen. Um die Schreibweise zu vereinfachen wird $|_{x_0^m}$ ab jetzt weggelassen.

Somit wird (5.8) und (5.9) zu:

$$\boxed{\partial_i = \partial_i \tilde{x}^k \tilde{\partial}_k} \quad (5.10)$$

$$\tilde{\partial}_i = \tilde{\partial}_i x^k \partial_k \quad (5.11)$$

Vergleiche (5.7) mit (5.11): die Operatoren ∂_i besitzen genau das Transformationsverhalten der Basiselemente E_i^a und sind von der Struktur des Vektorraums her ununterscheidbar.

Wir identifizieren daher

$$E_i^a \leftrightarrow \partial_i^a$$

Damit kann jeder Tangentialvektor geschrieben werden als

$$v^a = v^i \partial_i^a = \tilde{v}^i \tilde{\partial}_i^a \quad (5.12)$$

Der Tangentialraum am Punkt $p \in M$ lässt sich somit **basis- und koordinatenunabhängig** als Vektorraum darstellen als

$$T_p M = \{v^a \mid v^a = v^i \partial_i^a \text{ in jeder Karte } x^i : p \mapsto x^i(p) \text{ die } p \text{ enthält}\}$$

Bemerkung 5.16 (Nicht Teil der VO).

Besser verständlich würde ich es so anschreiben:

$$T_p M = \left\{ v^a \mid \forall \text{ Karten } (O, \phi) \text{ mit } p \in O, \phi^i(p) = x_0^m : v^a = v^i \partial_i^a \Big|_{x_0^m} \right\}$$

Also Vektoren v^a , die in **jeder** Karte (O, ϕ) mit $p \in O$, $\phi^m(p) = x_0^m$ darstellbar sind als $v^a = v^i \partial_i^a \Big|_{x_0^m}$.

Bemerkung 5.17. Aus (5.12) erhalten wir wieder (5.5):

$$v^a = v^i \partial_i^a \stackrel{(5.10)}{=} v^i \partial_i^a \tilde{x}^k \tilde{\partial}_k^a = \left(\partial_i^a \tilde{x}^k v^i \right) \tilde{\partial}_k^a \stackrel{(5.12)}{=} \tilde{v}^k \tilde{\partial}_k^a$$

Somit folgt das **Transformationsverhalten der Komponenten des Tangentialvektors beim Kartenwechsel** (5.5)

$$\tilde{v}^k = \partial_i^a \tilde{x}^k v^i \quad v^k = \tilde{\partial}_i^a x^k \tilde{v}^i \quad (5.13)$$

Also muss man sich nur (5.12) und die Kettenregel (5.10) merken und kann sich daher das Transformationsverhalten der Tangentialvektoren herleiten.

Wir können nun den erarbeiteten Vektorraum verwenden, um Vektorfelder auf Mannigfaltigkeiten zu definieren.

Definition 5.18. Sei (M, \mathfrak{M}) eine C^k -Mannigfaltigkeit. Ein C^k -**Vektorfeld** ist dann

$$v^a : M \rightarrow T_p M : p \mapsto v^a(p) \quad \forall p \in M$$

Das Vektorfeld bildet also jeden Punkt $p \in M$ auf einen Tangentialvektor $v^a(p) \in T_p M$ ab.

Für jede Karte (O, ϕ) mit Koordinaten x^i besitzt v^a lokal eine Darstellung

$$v^a = v^i(x^k) \partial_i^a \Big|_{x_0^m}$$

wobei die **numerischen Funktionen/Vektorfelder (Komponenten des Vektorfelds)**

$$v^i : v^a \circ (\phi^i)^{-1} : \phi^i(O) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$$

k -mal stetig differenzierbar sind.

Definition 5.19 (Nicht Teil der VO).

Die disjunkte Vereinigung aller Tangentialräume von M heißt **Tangentialbündel** und wird definiert als

$$TM := \bigcup_{p \in M} T_p M$$

D.h.

$$TM = \bigcup_{p \in M} T_p M \quad \text{und} \quad T_p M \cap T_{\tilde{p}} M = \emptyset \quad (\text{paarweise disjunkt})$$

Mit dem Tangentialbündel lautet die Definition:

Definition 5.20 (Nicht Teil der VO).

Sei (M, \mathfrak{M}) eine C^k -Mannigfaltigkeit und die kanonische Projektion $\pi : TM \rightarrow M$. Ein C^k -**Vektorfeld** ist dann

$$v^a : M \rightarrow TM : p \mapsto v^a(p) \in T_p M$$

mit

$$\pi \circ v^a = \text{id}_M$$

wobei die numerischen Funktionen/Vektorfelder (Komponenten des Vektorfelds) k -mal stetig differenzierbar sind.

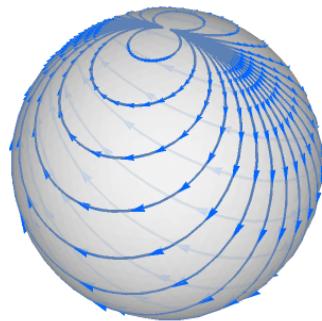


Abbildung 5.11: Vektorfeld auf der Mannigfaltigkeit S^2

5.4.1 Dualraum des Tangentialraums (Kotangentialraum)

Wir erinnern uns an die Definition von (r,s) -Tensoren (3.34)

$$\mathfrak{Mult}_{r,s}(V^\sim, V, \Gamma) = \bigotimes^r V \otimes \bigotimes^s V^\sim = \bigotimes_s^r V =: T_s^r$$

Also brauchen wir für Tensorfelder auf Mannigfaltigkeiten noch das Konzept des Dualraums vom Tangentialraum. Da der Tangentialraum $T_p M$ ein Vektorraum ist, besitzt er einen Dualraum. Diesen bezeichnen wir als den **Kotangentialraum** $(T_p M)^\sim$.

Ein Kovektor $\omega_a \in (T_p M)^\sim$ ist somit eine lineare Abbildung

$$\omega_a : T_p M \rightarrow \mathbb{R} : v^a \mapsto \omega_a v^a$$

Bezüglich einer Basis E_i^a des Tangentialraums existiert eine duale Basis e_a^i des Kotangentialraums, definiert durch die Dualitätsrelation

$$e_a^i(E_j^a) = \delta_j^i \quad (5.14)$$

Damit lässt sich ein Kovektor schreiben als

$$\omega_a = \omega_i e_a^i$$

Da wir das Transformationsverhalten der Basis E_i^a des Tangentialraums kennen (siehe (5.7)) und in section 3.7 gezeigt haben, dass die duale Basis e_a^i stets invers zur Vektorbasis transformiert, folgt für einen Kartenwechsel $\tilde{x}^i = \tilde{x}^i(x^k)$

$$\begin{aligned} \tilde{e}_a^i &= \underbrace{\frac{\partial \tilde{x}^i}{\partial x^k}(x_0^m)}_{=\partial_k \tilde{x}^i|_{x_0^m}} e_a^k \\ &= \partial_k \tilde{x}^i|_{x_0^m} \end{aligned}$$

Wir suchen wieder nach einem Objekt, das obiges Transformationsverhalten aufweist. Dazu betrachten wir das totale Differential

$$\begin{aligned} d\tilde{x}^i|_{x_0^m} &= \underbrace{\frac{\partial \tilde{x}^i}{\partial x^k}(x_0^m) dx^k|_{x_0^m}}_{=\partial_k \tilde{x}^i|_{x_0^m}} \\ &= \partial_k \tilde{x}^i|_{x_0^m} \end{aligned} \quad (5.15)$$

Die Ableitungen sind wieder immer bzgl. dem Punkt x_0^m zu verstehen. Wir werden nun wieder $|_{x_0^m}$ weglassen, um die Schreibweise zu vereinfachen. Damit wird (5.15) zu

$$d\tilde{x}^i = \partial_k \tilde{x}^i dx^k \quad (5.16)$$

bzw.

$$dx^i = \tilde{\partial}_k x^i d\tilde{x}^k \quad (5.17)$$

Wir sehen, dass dx^i exakt dasselbe Transformationsverhalten wie die dualen Basiselemente e_a^i besitzt und somit von der Struktur des Vektorraums her ununterscheidbar sind. Wir identifizieren daher

$$e_a^i \leftrightarrow dx_a^i$$

Damit kann jeder Kovektor ω_a bzgl. verschiedenen Basen zerlegt werden

$$\omega_a = \omega_i dx_a^i = \tilde{\omega}_i d\tilde{x}_a^i \quad (5.18)$$

Somit kann der Kotangentialraum am Punkt $p \in M$ **basis- und koordinatenunabhängig** dargestellt werden als

$$(T_p M)^\sim = \{\omega_a \mid \omega_a = \omega_i dx_a^i \text{ in jeder Karte } x^i : p \mapsto x^i(p) \text{ die } p \text{ enthält}\}$$

Bemerkung 5.21 (Nicht Teil der VO).

Besser verständlich würde ich es so anschreiben:

$$(T_p M)^\sim = \left\{ \omega_a \mid \forall \text{Karten } (O, \phi) \text{ mit } p \in O, \phi^m(p) = x_0^m : \omega_a = \omega_i dx_a^i|_{x_0^m} \right\}$$

Also Kovektoren ω^a , die in **jeder** Karte (O, ϕ) mit $p \in O$, $\phi^m(p) = x_0^m$ darstellbar sind als $\omega_a = \omega_i dx_a^i|_{x_0^m}$.

Bemerkung 5.22 (Dualitätsrelation).

Mit der Identifikation $E_i^a \leftrightarrow \partial_i^a$ und $e_a^i \leftrightarrow dx_a^i$ kann die Dualitätsrelation (5.14) nun geschrieben werden als

$$dx_a^i(\partial_j^a) = \delta_j^i$$

Dies ist die kanonische Paarung zwischen Basis und dualer Basis und *nicht* als Ableitungsgleichung zu interpretieren.

Bemerkung 5.23 (Transformation Kovektor bzgl. Kartenwechsel).

Ein Kovektor ω_a kann bezüglich zweier Karten (O, ϕ) und $(\tilde{O}, \tilde{\phi})$ geschrieben werden als

$$\omega_a = \omega_i dx_a^i = \tilde{\omega}_i d\tilde{x}_a^i \quad (5.19)$$

Setze das Transformationsverhalten der Basiselemente (5.16) bzw. (5.17) in (5.19) ein

$$\omega_i dx_a^i = \omega_i \tilde{\partial}_k x^i d\tilde{x}_a^k, \quad \tilde{\omega}_i d\tilde{x}_a^i = \tilde{\omega}_i \partial_k \tilde{x}^i dx_a^k$$

und vergleiche die Koeffizienten vor dx_a^k bzw. $d\tilde{x}_a^k$ mit (5.19), so folgt das **Transformationsverhalten der Komponenten beim Kartenwechsel**

$$\boxed{\omega_k = \partial_k \tilde{x}^i \tilde{\omega}_i} \quad \boxed{\tilde{\omega}_k = \tilde{\partial}_k x^i \omega_i} \quad (5.20)$$

Das Transformationsverhalten der Komponenten ergibt sich somit *automatisch* aus dem Transformationsverhalten der Basisobjekte dx_a^i und muss nicht auswendig gelernt werden.

Bemerkung 5.24 (Interpretation von Differentialen und Koordinatenunabhängigkeit).

Die Einführung der Differentiale dx_a^i bzw. der Ableitungsoperatoren ∂_i^a ist als *Merkregel für das Transformationsverhalten* der Komponenten von Vektoren v^i und Kovektoren w_i zu verstehen.

Wir haben nach natürlichen Objekten gesucht, deren Transformationsverhalten beim Kartenwechsel genau dem der Komponenten entspricht:

- die partiellen Ableitungsoperatoren ∂_i^a besitzen genau das Transformationsverhalten der Basiselemente E_i^a des Vektorraums (Kettenregel)
- die totalen Differentiale dx_a^i besitzen dasselbe Transformationsverhalten wie die Basiselemente e_a^i des Dualraums

Die Komponenten sind *keine intrinsischen Eigenschaften*, sondern hängen von der gewählten Basis ab. Wechselt man die Karte, so ändern sich sowohl die Komponenten v^i bzw. w_i als auch die zugehörigen Basisobjekte ∂_i^a bzw. dx_a^i und zwar so, dass sich die Änderungen gegenseitig kompensieren und sich somit das Objekt selbst, der Vektor v^a bzw. der Kovektor ω_a , nicht ändert. (Vergleiche (5.12) und (5.18))

Bemerkung 5.25 (Koordinatenunabhängigkeit der dualen Paarung).

Wir zeigen, dass gilt

$$\omega_a v^a = \omega_i v^i = \tilde{\omega}_j \tilde{v}^j$$

$$\begin{aligned} \omega_a v^a &= \omega_i v^i \stackrel{(5.13)}{=} \partial_i \tilde{x}^j \tilde{\omega}_j \tilde{\partial}_l x^i \tilde{v}^l = \underbrace{\partial_i \tilde{x}^j \tilde{\partial}_l x^i}_{=\frac{\partial \tilde{x}^j}{\partial x^i} \frac{\partial x^i}{\partial \tilde{x}^l} = \delta_l^j} \tilde{\omega}_j \tilde{v}^l = \delta_l^j \tilde{\omega}_j \tilde{v}^l = \tilde{\omega}_j \tilde{v}^j \end{aligned}$$

Ausgehend vom Tangential- und Kotangentialraum ist nun der Tensorproduktraum vom Typ (r, s) am Punkt p

$$\bigotimes_s^r T_p M = \underbrace{T_p M \otimes \cdots \otimes T_p M}_{r\text{-mal}} \otimes \underbrace{(T_p M)^\sim \otimes \cdots \otimes (T_p M)^\sim}_{s\text{-mal}}$$

Ein Tensor $t_{b_1 \dots b_s}^{a_1 \dots a_r} \in \bigotimes_s^r T_p M$ besitzt bezüglich einer Karte, die p enthält, die Darstellung

$$\begin{aligned} t_{b_1 \dots b_s}^{a_1 \dots a_r} &= t_{j_1 \dots j_s}^{i_1 \dots i_r} \partial_{i_1}^{a_1} \otimes \cdots \otimes \partial_{i_r}^{a_r} \otimes dx_{b_1}^{j_1} \otimes \cdots \otimes dx_{b_s}^{j_s} \\ &\stackrel{(3.22)}{=} t_{j_1 \dots j_s}^{i_1 \dots i_r} \partial_{i_1}^{a_1} \cdot \dots \cdot \partial_{i_r}^{a_r} \cdot dx_{b_1}^{j_1} \cdot \dots \cdot dx_{b_s}^{j_s} \\ &= \tilde{t}_{j_1 \dots j_s}^{i_1 \dots i_r} \tilde{\partial}_{i_1}^{a_1} \cdot \dots \cdot \tilde{\partial}_{i_r}^{a_r} \cdot d\tilde{x}_{b_1}^{j_1} \cdot \dots \cdot d\tilde{x}_{b_s}^{j_s} \end{aligned}$$

Nun können wir Tensorfelder auf Mannigfaltigkeiten global und kartunabhängig definieren.

Definition 5.26. Sei (M, \mathfrak{M}) eine C^k -Mannigfaltigkeit. Ein **(r,s) - C^k -Tensorfeld** ist dann

$$t_{b_1 \dots b_s}^{a_1 \dots a_r} : M \rightarrow \bigotimes_s^r T_p M : p \mapsto t_{b_1 \dots b_s}^{a_1 \dots a_r}(p) \quad \forall p \in M$$

Das Tensorfeld bildet also jeden Punkt $p \in M$ auf einen Tensor $t_{b_1 \dots b_s}^{a_1 \dots a_r}(p) \in \bigotimes_s^r T_p M$ ab.

Für jede Karte (O, ϕ) mit Koordinaten x^i besitzt $t_{b_1 \dots b_s}^{a_1 \dots a_r}$ lokal eine Darstellung

$$t_{b_1 \dots b_s}^{a_1 \dots a_r} = t_{j_1 \dots j_s}^{i_1 \dots i_r} \partial_{i_1}^{a_1} \cdot \dots \cdot \partial_{i_r}^{a_r} \cdot dx_{b_1}^{j_1} \cdot \dots \cdot dx_{b_s}^{j_s}$$

wobei die **numerischen Funktionen/Tensorfelder (Komponenten des Tensorfelds)**

$$t_{j_1 \dots j_s}^{i_1 \dots i_r} : t_{b_1 \dots b_s}^{a_1 \dots a_r} \circ \phi^{-1} : \phi(O) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$$

k -mal stetig differenzierbar sind.

Sei $T_s^r(T_p M) := \bigotimes_s^r T_p M$ Tensorproduktraum des Vektorraums $T_p M$.

Definition 5.27 (Nicht Teil der VO).

Die disjunkte Vereinigung aller Tensorprodukträume von M heißt **(r,s) -Tensorbündel** und wird definiert als

$$T_s^r(TM) := \bigcup_{p \in M} T_s^r(T_p M)$$

Mit dem (r,s) -Tensorbündel lautet die Definition:

Definition 5.28 (Nicht Teil der VO).

Sei (M, \mathfrak{M}) eine C^k -Mannigfaltigkeit und die kanonische Projektion $\pi_s^r : T_s^r(TM) \rightarrow M$.

Ein **(r,s) - C^k -Tensorfeld** ist dann

$$t_{b_1 \dots b_s}^{a_1 \dots a_r} : M \rightarrow T_s^r(TM) : p \mapsto t_{b_1 \dots b_s}^{a_1 \dots a_r}(p) \in T_p M$$

mit

$$\pi_s^r \circ t_{b_1 \dots b_s}^{a_1 \dots a_r} = \text{id}_M$$

wobei die numerischen Funktionen/Tensorfelder (Komponenten des Tensorfelds) k -mal stetig differenzierbar sind.

Ein Vektorfeld ist dabei ein Spezialfall eines Tensorfeldes mit kontravariantem Grad 1 und kovariantem Grad 0.

Definition 5.29. Der Raum der **(r,s) - C^∞ -Tensorfelder** auf der Mannigfaltigkeit M ist definiert als

$$\mathcal{T}_s^r(M) := \left\{ t_{b_1 \dots b_s}^{a_1 \dots a_r} \text{ ist ein } C^\infty\text{-Tensorfeld} \right\}$$

Insbesondere gilt:

$\mathcal{T}_0^0(M) = C^\infty(M)$ ist der Raum der C^∞ -Skalarfelder

$\mathcal{T}_0^1(M)$ ist der Raum der C^∞ -Vektorfelder

$\mathcal{T}_1^0(M)$ ist der Raum der C^∞ -Kovektorfelder (1-Formen)

$\mathcal{T}_1^1(M)$ ist der Raum der $(1,1)$ - C^∞ -Tensorfelder

Kapitel 6

Differentialgeometrie und Einstingleichungen

6.1 Kovariante Ableitung

Im flachen Raum \mathbb{R}^n (euklidische Geometrie) wird ein Vektor $v^a = v^i \partial_i^a$ entlang einer Kurve $\gamma(t)$ mit Koordinaten $\gamma^i(t) = x^i(t)$ als *parallel* betrachtet, wenn er sich entlang der Kurve nicht ändert. Da die Basisvektoren ∂_i^a im \mathbb{R}^n ortsunabhängig sind, reduziert sich diese Bedingung auf die Konstanz der Komponenten:

$$\frac{d}{dt} v^i(x^m(t)) = 0$$

Die gewöhnliche Ableitung misst in diesem Fall genau die Abweichung von der Parallelität.

Auf einer gekrümmten Mannigfaltigkeit ist diese Situation jedoch nicht mehr so einfach. Ein Tangentialvektor entlang einer Kurve $\gamma(t)$ besitzt die Darstellung

$$v^a(x^m(t)) = v^i(x^m(t)) \partial_i^a|_{x^m(t)}$$

Bei einer Ableitung nach dem Kurvenparameter t ändern sich nun nicht nur die Komponenten $v^i(x^m(t))$, sondern auch die Basisvektoren $\partial_i^a|_{x^m(t)}$ selbst, da diese vom Punkt $p = \gamma(t) \in M$ via $\phi^i(\gamma(t)) = x^i(t)$ abhängen.

Die gewöhnliche Ableitung misst daher nicht ausschließlich eine tatsächliche geometrische Änderung des Vektors, sondern enthält zusätzliche Beiträge, die durch den Wechsel der lokalen Basis entstehen. Eine verschwindende gewöhnliche Ableitung ist somit im Allgemeinen nicht gleichbedeutend mit Parallelität.

Ziel der kovarianten Ableitung ist es, eine Ableitung zu definieren, die genau die Änderung eines Vektors misst, die einer Abweichung von der Parallelität entspricht. Sie erlaubt es, den Begriff der Parallelverschiebung auf gekrümmten Mannigfaltigkeiten sinnvoll zu definieren.

6.1.1 Koordinatenkovariante Ableitung auf Skalarfeldern

Wir betrachten zunächst ein Skalarfeld $f \in \mathcal{T}_0^0(M)$, also eine Abbildung

$$f : M \rightarrow \mathbb{R}$$

die bezüglich einer Karte (O, ϕ) mit Koordinaten x^i lokal durch $f(x^i)$ dargestellt wird. Auf diese Darstellung können wir die partielle Ableitung anwenden:

$$\partial_i f := \frac{\partial f}{\partial x^i}$$

Diese Größe ist jedoch *nicht* koordinatenunabhängig. Wir suchen daher eine Ableitung, deren Wirkung auf Skalarfelder unabhängig von der Wahl der Karte ist.

Sei

Definition 6.1. Die **koordinatenkovariante Ableitung eines Skalarfeldes** ist definiert als

$$\partial_a : \mathcal{T}_0^0(M) \rightarrow \mathcal{T}_1^0(M) : f \mapsto \partial_a f := \partial_i f dx_a^i \quad (6.1)$$

$\partial_a f$ ist also ein Kovektorfeld. Hierbei wirken Ableitungssymbole nur auf das direkt folgende Objekt, sofern nicht durch Klammern anders gekennzeichnet.

Die partielle Ableitung wirkt auf der lokalen Darstellung des Skalarfelds, die koordinatenkovariante Ableitung wirkt auf dem Skalarfeld selbst.

Wir zeigen nun, dass diese Definition tatsächlich koordinatenunabhängig ist, also $\tilde{\partial}_a f = \partial_a f$.

Seien dazu (O, ϕ) und $(\tilde{O}, \tilde{\phi})$ zwei Karten mit Kartenwechsel $\tilde{x}^k = \tilde{x}^k(x^i)$ bzw. $x^k = x^k(\tilde{x}^i)$. Gilt für die lokale Darstellung des Skalarfeldes

$$f(x^k) = f(x^k(\tilde{x}^i))$$

Mit der Kettenregel folgt

$$\tilde{\partial}_i f = \frac{\partial f}{\partial \tilde{x}^i}(x^k(\tilde{x}^i)) = \frac{\partial x^k}{\partial \tilde{x}^i} \frac{\partial f}{\partial x^k} = \tilde{\partial}_i x^k \partial_k f \quad (6.2)$$

Zugleich gilt für das totale Differential

$$dx_a^i = dx_a^i(\tilde{x}^k) = \frac{\partial x^i}{\partial \tilde{x}^k} d_a \tilde{x}^k = \tilde{\partial}_k x^i d_a \tilde{x}^k \quad (6.3)$$

Setzt man beides zusammen, so ergibt sich

$$\begin{aligned} \tilde{\partial}_a f &= \tilde{\partial}_i f dx_a^i \stackrel{(6.2)}{=} \tilde{\partial}_i x^k \partial_k f dx_a^i = \underbrace{\tilde{\partial}_i x^k d_a \tilde{x}^i}_{\stackrel{(6.3)}{=} dx_a^k} \partial_k f = dx_a^k \partial_k f = \partial_a f \end{aligned} \quad (6.4)$$

Damit ist die koordinatenkovariante Ableitung auf Skalarfeldern unabhängig von der Wahl der Karte.

Bemerkung 6.2. Da aus (6.2) folgt, dass $\partial_i f$ beim Kartenwechsel kontravariant transformiert und wir bereits wissen, dass dx^i kovariant transformiert, hätte man sofort schließen können, dass die Koordinatenkovariante Ableitung auf Skalarfeldern koordinatenunabhängig ist, da sich die jeweilige Koordinatenabhängigkeit in der Kombination $\partial_i f dx_a^i$ aufhebt.

6.1.2 Koordinatenkovariante Ableitung auf Vektorfeldern

Nachdem wir gesehen haben, dass alle koordinatenkovarianten Ableitungen auf Skalarfeldern in allen Koordinaten übereinstimmen, untersuchen wir nun, ob dies auch für Vektorfelder gilt.

Sei $v^a = v^i(x^m) \partial_i^a \in \mathcal{T}_0^1(M)$ ein Vektorfeld.

Definition 6.3. Die **koordinatenkovariante Ableitung eines Vektorfeldes** bezüglich der Karte (O, ϕ) ist definiert durch

$$\partial_b : \mathcal{T}_0^1(M) \rightarrow \mathcal{T}_1^1(M) : v^a \mapsto \partial_b v^a := \partial_j v^i dx_b^j \partial_i^a$$

$\partial_b v^a$ ist also ein Tensorfeld.

Mit der Leibnizregel können wir schreiben:

$$\partial_b v^a = \partial_b(v^i \partial_i^a) = \partial_b v^i \partial_i^a + v^i \partial_b \partial_i^a \quad (6.5)$$

Vergleicht man diesen Ausdruck mit [Definition 6.3](#), so folgt

$$\partial_b v^i = \partial_j v^i dx_b^j \quad \text{und} \quad \partial_b \partial_i^a = 0$$

$\partial_b \partial_i^a = 0$ bedeutet, dass diese Basisvektoren ∂_i^a bezüglich der koordinatenkovarianten Ableitung ∂_b konstant sind. Da jede Karte/Koordinatensystem ihre eigene Basis besitzt, liefert jedes Koordinatensystem seinen eigenen koordinatenkovarianten Ableitungsoperator. Also gilt im Allgemeinen für verschiedene Karten (O, ϕ) und $(\tilde{O}, \tilde{\phi})$: $\partial_b \tilde{\partial}_i^a \neq 0$, $\tilde{\partial}_b \partial_i^a \neq 0$ aber $\tilde{\partial}_b \tilde{\partial}_i^a = 0$.

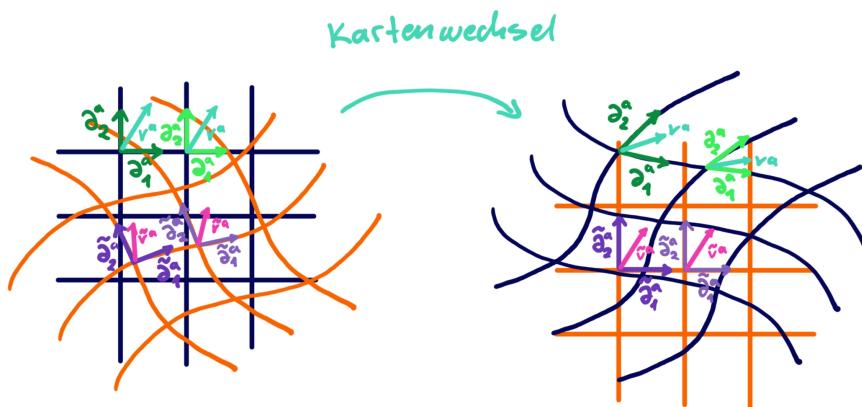


Abbildung 6.1: Wenn wir einen Vektor v im ungetildeten Koordinatensystem parallel verschieben, dann ändern sich dessen Basisvektoren ∂_i^a nicht bzgl. des eigenen Ableitungsoperators ∂_b , also $\partial_b \partial_i^a = 0$, aber im allgemeinen schon bzgl. des Ableitungsoperators $\tilde{\partial}_b$ des getildeten Koordinatensystems, also $\tilde{\partial}_b \partial_i^a \neq 0$.

Dasselbe gilt auch für das getildete Koordinatensystem. Parallelverschiebung hängt also vom Koordinatensystem ab.

Somit ist die koordinatenkovarianten Ableitung auf Vektorfeldern *abhängig* von der Wahl der Karte. Also gilt im Allgemeinen $\partial_b v^a \neq \tilde{\partial}_b v^a$. Wir haben somit keinen globalen Ableitungsbegriff.

Transformationsverhalten der koordinatenkovarianten Ableitung auf einem Vektorfeld

Wir betrachten dazu zwei Karten (O, ϕ) und $(\tilde{O}, \tilde{\phi})$ mit Koordinaten x^i bzw. \tilde{x}^i . Das Vektorfeld kann bezüglich beider Basen dargestellt werden:

$$v^a = v^i \partial_i^a = \tilde{v}^i \tilde{\partial}_i^a$$

Wir berechnen nun $\partial_b v^a$ in der Tilde Darstellung:

$$\begin{aligned}
 \partial_b v^a &= \partial_b (\tilde{v}^i \tilde{\partial}_i^a) \\
 &\stackrel{\text{Leibniz}}{=} \underbrace{\partial_b \tilde{v}^i}_{\substack{\text{ist num. VF} \\ \text{skalare Fkt.} \\ \stackrel{(6.4)}{=} \tilde{v}^i}} + \underbrace{\tilde{v}^i \partial_b}_{\tilde{\partial}_i^a} \underbrace{\tilde{\partial}_i^a}_{\substack{\text{Basis-} \\ \text{Trafo} \tilde{\partial}_i x^j \partial_j^a}} \\
 &= \tilde{\partial}_b \tilde{v}^i \tilde{\partial}_i^a + \tilde{v}^i \partial_b (\tilde{\partial}_i x^j \partial_j^a) \tag{6.6}
 \end{aligned}$$

Da

$$\tilde{\partial}_b (\tilde{v}^i \tilde{\partial}_i^a) = \tilde{\partial}_b \tilde{v}^i \tilde{\partial}_i^a + \tilde{v}^i \underbrace{\tilde{\partial}_b \tilde{\partial}_i^a}_{=0} \Rightarrow \tilde{\partial}_b (\tilde{v}^i \tilde{\partial}_i^a) = \tilde{\partial}_b \tilde{v}^i \tilde{\partial}_i^a$$

und

$$\partial_b (\tilde{\partial}_i x^j \partial_j^a) = \partial_b \tilde{\partial}_i x^j \partial_j^a + \tilde{\partial}_i x^j \underbrace{\partial_b \partial_j^a}_{=0} \Rightarrow \partial_b (\tilde{\partial}_i x^j \partial_j^a) = \partial_b \tilde{\partial}_i x^j \partial_j^a$$

folgt

$$\begin{aligned}
 (6.6) &= \tilde{\partial}_b (\tilde{v}^i \tilde{\partial}_i^a) + \underbrace{\tilde{v}^i}_{=v^a} \underbrace{=\tilde{v}^j \delta_j^i = \tilde{v}^j \tilde{\partial}_j^c d\tilde{x}_c^i = v^c d\tilde{x}_c^i} \partial_b \tilde{\partial}_i x^j \partial_j^a \\
 &= \tilde{\partial}_b v^a + d\tilde{x}_c^i v^c \partial_b \tilde{\partial}_i x^j \underbrace{\partial_j^a}_{\substack{\text{Basis-} \\ \text{Trafo} \partial_j \tilde{x}^k \tilde{\partial}_k^a}} \\
 &= \tilde{\partial}_b v^a + d\tilde{x}_c^i v^c \underbrace{\partial_b \tilde{\partial}_i x^j \partial_j \tilde{x}^k \tilde{\partial}_k^a}_{\substack{\text{skalare Fkt.} \\ \tilde{\partial}_b \tilde{\partial}_i x^j \partial_j \tilde{x}^k \tilde{\partial}_k^a}} \\
 &= \tilde{\partial}_b v^a + d\tilde{x}_c^i v^c \underbrace{\tilde{\partial}_b}_{\stackrel{(6.1)}{=}\tilde{\partial}_l} \underbrace{\tilde{\partial}_i x^j \partial_j \tilde{x}^k \tilde{\partial}_k^a}_{\tilde{\partial}_l \tilde{\partial}_i x^j \tilde{\partial}_j \tilde{x}^k \tilde{\partial}_k^a} \\
 &= \tilde{\partial}_b v^a + d\tilde{x}_c^i d\tilde{x}_b^l \tilde{\partial}_l \tilde{\partial}_i x^j \partial_j \tilde{x}^k \tilde{\partial}_k^a v^c
 \end{aligned}$$

Also insgesamt

$$\partial_b v^a = \tilde{\partial}_b v^a + \tilde{\partial}_l \tilde{\partial}_i x^j \partial_j \tilde{x}^k \tilde{\partial}_k^a d\tilde{x}_c^i d\tilde{x}_b^l v^c$$

Wir definieren das **Christoffel-Symbol**

$$\Gamma_{cb}^a := \underbrace{\tilde{\partial}_l \tilde{\partial}_i x^j}_{=\frac{\partial^2 x^j}{\partial \tilde{x}^l \partial \tilde{x}^i}} \partial_j \tilde{x}^k \tilde{\partial}_k^a d\tilde{x}_c^i d\tilde{x}_b^l \tag{6.7}$$

und erhalten damit

$$\partial_b v^a = \tilde{\partial}_b v^a + \underbrace{\Gamma_{cb}^a v^c}_{\substack{\text{linearer Zusatzterm} \\ (\text{Korrekturterm})}}$$

(6.8)

bzw.

$$(\partial_b - \tilde{\partial}_b)v^a = \Gamma_{cb}^a v^c \in \mathcal{T}_1^1(M) \quad (6.9)$$

wobei a, b, c abstrakte Indizes und i, j, k, l Zählindizes sind.

Zusammengefasst erhalten wir:

- Auf Skalarfeldern stimmen alle koordinatenkovarianten Ableitungen überein.
- Auf Vektorfeldern unterscheiden sich die koordinatenkovarianten Ableitungen verschiedener Koordinatensysteme um einen linearen Zusatzterm. Dieser Zusatzterm beschreibt die Änderung der Basisvektoren beim Kartenwechsel und kann als Maß für die Abweichung von Parallelität interpretiert werden.

Bemerkung 6.4. Das Christoffel-Symbol Γ_{cb}^a ist ein lokal im Kartenüberlapp definierter Ausdruck, der vom Kartenwechsel abhängt. Er beschreibt die Differenz zweier Ableitungsoperatoren (6.9), von denen mindestens einer eine koordinatenkvariante Ableitung ist. Das Christoffel-Symbol ist nur *lokal* ein Tensor. Global gesehen, ist es kein Tensor da es bzgl. dem Kartenwechsel nicht wie ein Tensor transformiert. Erst die Differenz (6.9) ist (global) ein Tensor, der sogenannte **Differenzentensor**, da dabei die auftretenden nicht-tensoriellen Anteile wegfallen. [Quelle](#)

Bemerkung 6.5 (Nicht Teil der VO).

Quelle: https://en.wikipedia.org/wiki/Christoffel_symbols#Transformation_law_under_change_of_variable

Angenommen, Γ_{cb}^a aus (6.7) wäre ein Tensor vom Typ (1,2). Dann müssten seine Komponenten beim Kartenwechsel $\hat{x}^i = \hat{x}^i(\tilde{x}^j)$ ohne zusätzliche Terme transformieren:

$$\hat{\Gamma}_{fe}^d = \tilde{\partial}_a \hat{x}^d \hat{\partial}_f \tilde{x}^c \hat{\partial}_e \tilde{x}^b \Gamma_{cb}^a$$

Es gilt aber

$$\hat{\Gamma}_{fe}^d = \tilde{\partial}_a \hat{x}^d \hat{\partial}_f \tilde{x}^c \hat{\partial}_e \tilde{x}^b \Gamma_{cb}^a + \underbrace{\hat{\partial}_f \hat{\partial}_e \tilde{x}^a \tilde{\partial}_a \hat{x}^d}_{\text{inhomogener Zusatzterm}}$$

Bemerkung 6.6 (Nicht Teil der VO).

Quellen: https://en.wikipedia.org/wiki/Christoffel_symbols

https://einsteinerativelyeasy.com/index.php?option=com_content&view=article&id=25&catid=10&Itemid=112

<https://profoundphysics.com/christoffel-symbols-a-complete-guide-with-examples/>

Es gilt $\Gamma_{cb}^a(p) \in \mathbb{R}$ für Punkte aus dem Kartenüberlapp $p \in O \cap \tilde{O} \subset M$. Als Abbildung gilt $\Gamma_{cb}^a \in C^\infty(O \cap \tilde{O})$.

Für ein besseres Verständnis betrachte ich das Christoffel-Symbol Γ_{cb}^a in **Koordinatendarstellung**.

Betrachte wieder (6.5)

$$\partial_b v^a = \partial_b(v^i \partial_i^a) = (\partial_b v^i) \partial_i^a + v^i \underbrace{\partial_b \partial_i^a}_{=\Gamma_{ib}^k \partial_k^a} = \underbrace{(\partial_b v^i)}_{\text{Komponenten}} \underbrace{\partial_i^a}_{\text{Ableitung Basisvektor}} + \underbrace{v^i}_{\text{Komponenten}} \underbrace{\Gamma_{ib}^k \partial_k^a}_{\text{Änderung der Basisvektoren}}$$

Also in koordinatenabhängiger/basisabhängiger Darstellung

$$\partial_b \partial_i^a := \Gamma_{ib}^k \partial_k^a \in \mathcal{T}_0^1(M)$$

wobei

- der Zählindex i gibt den Basisvektor an, dessen Änderung betrachtet wird (∂_i^a)
- der abstrakte Index b gibt die Richtung an, entlang der abgeleitet wird
- der Zählindex k gibt den Basisvektor ∂_k^a an, in dessen Richtung die Ableitung vom Basisvektor ∂_i^a zeigt (Vektorrichtung des Zusatzterms zeigt)

Beispiel in Polarkoordinaten (siehe anschließend genauer):

$$\partial_\varphi e_r = 0 \cdot e_r + \frac{1}{r} e_\varphi = \Gamma_{r\varphi}^r \cdot e_r + \Gamma_{r\varphi}^\varphi \cdot e_\varphi$$

Das heißt, bei einer Ableitung eines Vektors $\partial_b v^a = \partial_b(v^i \partial_i^a)$ in Richtung b beschreibt das Christoffel-Symbol Γ_{ib}^k die Änderung des Basisvektors ∂_i^a , indem es den Koeffizienten der Komponente des Basisvektors ∂_k^a angibt, in dessen Richtung die Ableitung vom Basisvektor ∂_i^a zeigt. Aus diesem Grund können Christoffel-Symbole selbst in flachen Räumen auftreten, sofern die Koordinatenbasis nicht konstant ist (z. B. in Polarkoordinaten).

Das Christoffel-Symbol misst *nicht* die Änderung der Komponenten v^i eines Vektorfeldes.

Beispiel 6.7 (Nicht Teil der VO).

Quelle: https://en.wikipedia.org/wiki/Polar_coordinate_system#Differential_geometry

Wir betrachten die Ebene \mathbb{R}^2 als differenzierbare Mannigfaltigkeit. In kartesischen Koordinaten (x, y) sind die Koordinatenbasisvektoren

$$e_x = \partial_x = (1, 0), \quad e_y = \partial_y = (0, 1)$$

ortsunabhängig, sodass ihre Ableitungen verschwinden:

$$\partial_x \partial_x = \partial_x \partial_y = \partial_y \partial_x = \partial_y \partial_y = 0$$

Bei einem Vektorfeld in kartesischen Koordinaten $v(x, y) = v_1(x, y)\partial_x + v_2(x, y)\partial_y$ ändern sich bei der Ableitung somit nur die Komponenten $v_1(x, y)$ und $v_2(x, y)$ und nicht die Basisvektoren ∂_x und ∂_y .

Die Christoffel-Symbole sind somit $\Gamma_{xx}^x = \Gamma_{xx}^y = \Gamma_{xy}^x = \Gamma_{xy}^y = \Gamma_{yx}^x = \Gamma_{yx}^y = \Gamma_{yy}^x = \Gamma_{yy}^y = 0$.

Transformieren wir in Polarkoordinaten (r, φ) (führen einen Kartenwechsel durch) mit

$$x = r \cos \varphi, \quad y = r \sin \varphi$$

lauten die Basisvektoren

$$e_r = \partial_r = \cos \varphi \partial_x + \sin \varphi \partial_y, \quad r e_\varphi = \partial_\varphi = -r \sin \varphi \partial_x + r \cos \varphi \partial_y$$

Diese Basisvektoren sind nun *ortsabhängig*. Betrachten wir nun ihre Ableitungen

$$\partial_r \partial_r = \partial_r (\cos \varphi \partial_x + \sin \varphi \partial_y) = 0 \cdot \partial_r + 0 \cdot \partial_\varphi \rightarrow \Gamma_{rr}^r = 0, \quad \Gamma_{rr}^\varphi = 0$$

$$\partial_\varphi \partial_r = \partial_\varphi (\cos \varphi \partial_x + \sin \varphi \partial_y) = -\sin \varphi \partial_x + \cos \varphi \partial_y = 0 \cdot \partial_r + \frac{1}{r} \partial_\varphi \rightarrow \Gamma_{\varphi r}^r = 0, \quad \Gamma_{\varphi r}^\varphi = \frac{1}{r}$$

$$\partial_r \partial_\varphi = \partial_r (-r \sin \varphi \partial_x + r \cos \varphi \partial_y) = -\sin \varphi \partial_x + \cos \varphi \partial_y = 0 \cdot \partial_r + \frac{1}{r} \partial_\varphi \rightarrow \Gamma_{r\varphi}^r = 0, \quad \Gamma_{r\varphi}^\varphi = \frac{1}{r}$$

$$\partial_\varphi \partial_\varphi = \partial_\varphi (-r \sin \varphi \partial_x + r \cos \varphi \partial_y) = -r \cos \varphi \partial_x - r \sin \varphi \partial_y = -r \cdot \partial_r + 0 \cdot \partial_\varphi \rightarrow \Gamma_{\varphi\varphi}^r = -r, \quad \Gamma_{\varphi\varphi}^\varphi = 0$$

Daraus folgt, dass sich bei einem Vektorfeld in Polarkoordinaten $v(r, \varphi) = v_1(r, \varphi)\partial_r + v_2(r, \varphi)\partial_\varphi$ bei der Ableitung sowohl die Komponenten $v_1(r, \varphi)$ und $v_2(r, \varphi)$ als auch die Basisvektoren ∂_r und ∂_φ ändern.

Obwohl der Raum selbst flach ist, treten diese Terme aufgrund der nichtlinearen Koordinatenwahl auf (ist ein Kreis!).

Im \mathbb{R}^n gibt es nicht den Gradienten, sondern jedes Koordinatensystem definiert seinen eigenen Gradienten in dessen Koordinaten.

6.1.3 Von der koordinatenkovarianten Ableitung zur kovarianten Ableitung

Ausgehend von der koordinatenkovarianten Ableitung definieren wir nun den Begriff der kovarianten Ableitung. Sie soll in koordinatenunabhängiger Weise beschreiben, wie Tensorfelder auf einer Mannigfaltigkeit parallel verschoben bzw. differenziert werden. Wie bereits bei der koordinatenkovarianten Ableitung erhöht auch die kovariante Ableitung den kovarianten Grad eines Tensorfeldes um eins, worin sich die Namensgebung begründet.

Definition 6.8. Sei (M, \mathfrak{M}) eine glatte Mannigfaltigkeit. Eine **kovariante Ableitung** ist eine Abbildung

$$\nabla_a : \mathcal{T}_s^r(M) \longrightarrow \mathcal{T}_{s+1}^r(M) : t_{c_1 \dots c_s}^{b_1 \dots b_r} \mapsto \nabla_a t_{c_1 \dots c_s}^{b_1 \dots b_r}$$

mit folgenden Eigenschaften:

(K1) ∇_a ist linear:

$$\nabla_a (t_{c_1 \dots c_s}^{b_1 \dots b_r} + \tilde{t}_{c_1 \dots c_s}^{b_1 \dots b_r}) = \nabla_a t_{c_1 \dots c_s}^{b_1 \dots b_r} + \nabla_a \tilde{t}_{c_1 \dots c_s}^{b_1 \dots b_r}$$

sowie

$$\nabla_a (\lambda \cdot t_{c_1 \dots c_s}^{b_1 \dots b_r}) = \lambda \cdot \nabla_a t_{c_1 \dots c_s}^{b_1 \dots b_r}, \quad \lambda \in \mathbb{R}$$

(K2) ∇_a erfüllt die Leibniz-Regel (Produktregel):

$$\nabla_a (t_{c_1 \dots c_q}^{b_1 \dots b_p} \otimes \tilde{t}_{e_1 \dots e_s}^{d_1 \dots d_r}) = (\nabla_a t_{c_1 \dots c_q}^{b_1 \dots b_p}) \otimes \tilde{t}_{e_1 \dots e_s}^{d_1 \dots d_r} + t_{c_1 \dots c_q}^{b_1 \dots b_p} \otimes (\nabla_a \tilde{t}_{e_1 \dots e_s}^{d_1 \dots d_r})$$

(K3) ∇_a stimmt auf Skalarfeldern lokal (d.h. nach Wahl einer Karte) mit der koordinatenkovarianten Ableitung überein:

$$\nabla_a f = \underbrace{\partial_a f}_{\text{lokal definiert}}, \quad \forall f \in \underbrace{C^\infty(M)}_{= \mathcal{T}_0^0(M) \text{ (Skalarfelder)}}$$

∇_a ist auf der ganzen Mannigfaltigkeit definiert (global). ∂_a ist nur lokal definiert (erst nach Wahl einer Karte). Sobald eine Karte gewählt wird, dürfen sich beide Ableitungen auf Skalarfeldern nicht unterscheiden.

Es ist egal, ob man zuerst mit ∇_a ableitet und dann eine Karte wählt oder ob zuerst eine Karte gewählt wird und dann mit ∂_a abgeleitet wird.

(K4) Satz von Schwarz (Torsionsfreiheit/Symmetrie):

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f, \quad \forall f \in C^\infty(M)$$

Jede koordinatenkovariante Ableitung ist eine Kovariante Ableitung, die Umkehrung gilt im Allgemeinen nicht. Die kovariante Ableitung liefert uns einen globalen Ableitungsbegriff im Gegensatz zur koordinatenkovarianten Ableitung.

Die kovariante Ableitung ist nicht kanonisch (geht nicht aus der Struktur der Mannigfaltigkeit hervor). Sie ist basisabhängig/kartenabhängig. Daher gibt es viele verschiedene kovariante Ableitungen auf derselben Mannigfaltigkeit.

Wie sieht die Differenz von zwei kovarianten Ableitungen aus? Wir betrachten dafür zuerst, wie sie auf Skalarfelder wirkt.

Lemma 6.9. Seien ∇_a und $\tilde{\nabla}_a$ zwei kovariante Ableitungen auf einer Mannigfaltigkeit M und ein Vektorfeld $v^b \in \mathcal{T}_0^1(M)$.

Dann können Skalarfelder $f \in C^\infty(M)$ aus der Wirkung der Differenz der kovarianten Ableitungen herausgezogen werden

$$(\tilde{\nabla}_a - \nabla_a)(fv^b) = f \cdot (\tilde{\nabla}_a - \nabla_a)v^b \quad (6.10)$$

D.h. die Differenz der kovarianten Ableitungen ist $C^\infty(M)$ -linear.

Beweis.

$$(\tilde{\nabla}_a - \nabla_a)f \stackrel{K3}{=} (\tilde{\partial}_a - \partial_a)f = \tilde{\partial}_a f - \partial_a f \stackrel{\substack{(6.4) \\ \text{stimmen auf} \\ \text{Überlapp} \\ \text{überein}}}{\equiv} 0 \quad (6.11)$$

Thus

$$(\tilde{\nabla}_a - \nabla_a)(fv^b) \stackrel{K2}{=} \underbrace{(\tilde{\nabla}_a - \nabla_a)f \cdot v^b}_{=0} + f \cdot (\tilde{\nabla}_a - \nabla_a)v^b = f \cdot (\tilde{\nabla}_a - \nabla_a)v^b$$

□

Bemerkung 6.10. Beachte allerdings, dass Skalarfelder aus der kovarianten Ableitungen nicht herausgezogen werden können

$$\nabla_a(fv^b) \stackrel{K2}{=} \nabla_a f \cdot v^b + f \cdot \nabla_a v^b$$

Betrachte nun die Differenz in Koordinaten/in einer Karte

$$\begin{aligned} (\tilde{\nabla}_a - \nabla_a)v^b &= (\tilde{\nabla}_a - \nabla_a)(v^i \partial_i^b) \\ &\stackrel{(6.10)}{=} \underbrace{v^i}_{=v^i \delta_j^i = v^i \partial_j^c dx_c^i} \underbrace{(\tilde{\nabla}_a - \nabla_a)\partial_i^b}_{=v^c dx_c^i} \\ &= dx_c^i (\tilde{\nabla}_a - \nabla_a)\partial_i^b v^c =: C_{ca}^b v^c \end{aligned}$$

Das heißt, in Koordinaten (lokal) ist die Differenz ein Tensor C_{ca}^b . Damit folgt die koordinatenabhängige Definition des **Differenzentensors** in der Karte x^i

$$C_{ca}^b = \underbrace{dx_c^i}_{\text{lokal definiert}} (\tilde{\nabla}_a - \nabla_a) \underbrace{\partial_i^b}_{\text{lokal definiert}} \quad (6.12)$$

Proposition 6.11. C_{ca}^b ist unabhängig von der Wahl der Karte und somit global definiert. Insbesondere ist C_{ca}^b tatsächlich ein $(1,2)$ -Tensorfeld.

Beweis. Wähle eine Karte $\hat{x}^i = \hat{x}^i(x^k)$.

$$\text{zz: } \hat{C}_{ca}^b = C_{ca}^b$$

Es gilt

$$\begin{aligned}
 \hat{C}_{ca}^b &= d\hat{x}_c^i (\tilde{\nabla}_a - \nabla_a) \underbrace{\hat{\partial}_i^b}_{\substack{\text{Basis} \\ \text{trafo} \hat{\partial}_i x^k \partial_k^b}} \\
 &= d\hat{x}_c^i (\tilde{\nabla}_a - \nabla_a) \left(\underbrace{\hat{\partial}_i x^k}_{\in C^\infty(M)} \hat{\partial}_k^b \right) \\
 &\stackrel{(6.10)}{=} \underbrace{\hat{\partial}_i x^k d\hat{x}_c^i}_{\substack{\text{Basis} \\ \text{trafo} dx_c^k}} (\tilde{\nabla}_a - \nabla_a) \partial_k^b \\
 &= dx_c^k (\tilde{\nabla}_a - \nabla_a) \partial_k^b \\
 &= C_{ca}^b
 \end{aligned}$$

□

Damit können wir nun definieren

Definition 6.12. Seien ∇_a und $\tilde{\nabla}_a$ zwei kovariante Ableitungen auf einer Mannigfaltigkeit M und ein Vektorfeld $v^b \in \mathcal{T}_0^1(M)$. Der Unterschied zwischen den beiden definiert den **Differenzentensor**

$$C_{ca}^b v^c := (\tilde{\nabla}_a - \nabla_a)v^b = \tilde{\nabla}_a v^b - \nabla_a v^b$$

Also reicht es bereits aus, eine kovariante Ableitung zu kennen, um für ein Vektorfeld v^b alle anderen von der Mannigfaltigkeit zu erhalten mit

$$\boxed{\tilde{\nabla}_a v^b = \nabla_a v^b + C_{ca}^b v^c} \quad (6.13)$$

Wir werden später nachweisen, dass es sich bei $C_{ca}^b v^c$ tatsächlich um einen Tensor handelt.

Lemma 6.13 (Wirkung kovariante Ableitung auf Kovektorfeld). Seien ∇_a und $\tilde{\nabla}_a$ zwei kovariante Ableitungen und $\omega_b \in \mathcal{T}_1^0(M)$ ein Kovektorfeld. Dann gilt

$$(\tilde{\nabla}_a - \nabla_a) \omega_b = -C_{ba}^c \omega_c$$

Beweis. Seien $\omega_b \in \mathcal{T}_1^0(M)$ und $v^b \in \mathcal{T}_0^1(M)$ beliebig. Dann ist

$$(\tilde{\nabla}_a - \nabla_a) \left(\underbrace{\omega_b v^b}_{\in T_0^0(M)} \right) \stackrel{(6.11)}{=} 0$$

Mit der Leibniz-Regel ergibt sich andererseits auch

$$\begin{aligned}
 0 &= (\tilde{\nabla}_a - \nabla_a)(\omega_b v^b) = (\tilde{\nabla}_a - \nabla_a)\omega_b \cdot v^b + \omega_b \cdot \underbrace{(\tilde{\nabla}_a - \nabla_a)v^b}_{= C_{ba}^c v^b} \\
 &\Leftrightarrow (\tilde{\nabla}_a - \nabla_a)\omega_b \cdot v^b = -\omega_b C_{ba}^c v^b
 \end{aligned}$$

Da diese Gleichung für alle Vektorfelder v^b gilt, folgt

$$(\tilde{\nabla}_a - \nabla_a) \omega_b = -C_{ba}^c \omega_c$$

□

Wir erhalten also nun auch alle anderen kovarianten Ableitungen, die auf einen Kovektor wirken, mit der Transformation

$$\tilde{\nabla}_a \omega_b = \nabla_a \omega_b - C_{ba}^c \omega_c$$

(vgl. (6.13))

Proposition 6.14. Der Differenztensor ist in den unteren Indizes symmetrisch:

$$C_{ab}^c = C_{ba}^c$$

Beweis. Sei $f \in C^\infty(M)$. Aus der Torsionsfreiheit (K4) folgt

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f \quad \text{und} \quad \tilde{\nabla}_a \tilde{\nabla}_b f = \tilde{\nabla}_b \tilde{\nabla}_a f$$

Also ist

$$\nabla_a \nabla_b f - \nabla_b \nabla_a f = \tilde{\nabla}_a \tilde{\nabla}_b f - \tilde{\nabla}_b \tilde{\nabla}_a f$$

Umordnen ergibt

$$\begin{aligned} \underbrace{(\tilde{\nabla}_a - \nabla_a) \nabla_b f}_{\stackrel{(6.13)}{=} -C_{ba}^c \nabla_c f} &= \underbrace{(\tilde{\nabla}_b - \nabla_b) \nabla_a f}_{\stackrel{(6.13)}{=} -C_{ab}^c \nabla_c f} \\ \Leftrightarrow -C_{ba}^c \nabla_c f &= -C_{ab}^c \nabla_c f \end{aligned}$$

Da diese Gleichung für alle Kovektorfelder $\nabla_c f$ gilt, folgt

$$C_{ab}^c = C_{ba}^c$$

□

Abschließender Überblick

Die kovarianten Ableitungen ∇_a sind auf der gesamten Mannigfaltigkeit (d.h. global) definiert und somit unabhängig von der Wahl einer Karte. Sie sind jedoch **nicht** eindeutig. Die koordinatenkovarianten Ableitungen ∂_a sind bezüglich einer gewählten Karte (d.h. lokal) definiert. ∇_a angewandt auf ein Skalarfeld stimmt lokal (nach Wahl einer Karte) mit ∂_a überein.

Die Differenz zweier kovarianter Ableitungen

$$C_{ca}^b v^c = (\tilde{\nabla}_a - \nabla_a) v^b$$

ist global definiert, während die Koordinatendarstellung

$$C_{ca}^b = dx_c^i (\tilde{\nabla}_a - \nabla_a) \partial_i^b$$

nur lokal bezüglich der gewählten x^i -Karte existiert.

Damit sind kovariante Ableitungen leicht zu konstruieren: Kennt man eine kovariante Ableitung ∇_a , so erhält man alle anderen mit dem Differenzentensor C_{ca}^b :

$$\tilde{\nabla}_a v^b = \nabla_a v^b + C_{ca}^b v^c, \quad \tilde{\nabla}_a \omega_b = \nabla_a \omega_b - C_{ba}^c \omega_c$$

Bemerkung 6.15. Die Menge der kovarianten Ableitungen bildet einen affinen Raum. Wählt man eine kovariante Ableitung als Ursprung, so entsprechen die Tensorfelder C_{ca}^b den Verbindungsvektoren, mit deren Hilfe alle anderen kovarianten Ableitungen erreicht werden.

6.2 Derivationen und Kommutator

Definition 6.16. Eine **Algebra** (A, \cdot) ist ein Vektorraum A zusammen mit einer Abbildung,

$$\cdot : A \times A \rightarrow A, \quad (a, b) \mapsto a \cdot b$$

die bilinear ist, d.h. für alle $v, v_1, v_2, w, w_1, w_2 \in A$ und alle $\lambda \in \mathbb{R}$ gilt:

$$\begin{aligned} (v_1 + v_2) \cdot w &= v_1 \cdot w + v_2 \cdot w \\ v \cdot (w_1 + w_2) &= v \cdot w_1 + v \cdot w_2 \\ (\lambda v) \cdot w &= \lambda(v \cdot w) = v \cdot (\lambda w) \end{aligned}$$

Beispiel 6.17. Beispiele sind die Tensoralgebra mit dem Tensorprodukt, Polynomalgebren mit dem Cauchyprodukt oder der \mathbb{R}^3 mit dem Kreuzprodukt.

Definition 6.18. Eine **Derivation** auf einer Algebra A ist eine lineare Abbildung

$$D : A \rightarrow A,$$

die die Leibniz-Regel erfüllt:

$$\forall a, b \in A : D(a \cdot b) = Da \cdot b + a \cdot Db$$

Die Menge aller Derivationen wird mit $\text{Der}(A)$ bezeichnet.

Beispiel 6.19. Die koordinatenkovariante Ableitung und kovariante Ableitung sind Derivationen.

Proposition 6.20. Die Menge aller Derivationen $\text{Der}(A)$ bildet einen Vektorraum.

Die Komposition von Derivationen $D_1 \circ D_2$ ist im Allgemeinen keine Derivation, da zusätzliche Terme auftreten. Beim anschließend definierten Kommutator fallen diese zusätzlichen Terme weg.

Definition 6.21. Die bilineare Abbildung

$$* : \text{Der}(A) \times \text{Der}(A) \rightarrow \text{Der}(A), \quad D_1 * D_2 := [D_1, D_2] := D_1 \circ D_2 - D_2 \circ D_1$$

heißt der **Kommutator**. Damit wird $(\text{Der}(A), *)$ zu einer Algebra.

Wir kürzen ab jetzt ab mit $D_1 \circ D_2 = D_1 D_2$.

Beweis. Wir müssen zeigen, dass diese Definition wohldefiniert ist. Die Bilinearität folgt aus der Bilinearität der Derivation.

Also bleibt zz: Der Kommutator zweier Derivationen ist wieder eine Derivation.

Seien $D_1, D_2 \in \text{Der}(A)$ und $a, b \in A$.

zz: der Kommutator erfüllt die Leibniz-Regel:

$$[D_1, D_2](a \cdot b) = [D_1, D_2]a \cdot b + a \cdot [D_1, D_2]b$$

- Betrachte die linke Seite

$$[D_1, D_2](a \cdot b) = (D_1 D_2 - D_2 D_1)(a \cdot b) = (D_1 D_2)(a \cdot b) - (D_2 D_1)(a \cdot b) = D_1(D_2(a \cdot b)) - D_2(D_1(a \cdot b))$$

Dann ist

$$D_1(D_2(a \cdot b)) = D_1(D_2a \cdot b + a \cdot D_2b) = D_1D_2a \cdot b + D_2a \cdot D_1b + D_1a \cdot D_2b + a \cdot D_1D_2b$$

Analog erhält man

$$D_2(D_1(a \cdot b)) = D_2D_1a \cdot b + D_1a \cdot D_2b + D_2a \cdot D_1b + a \cdot D_2D_1b$$

Die Komposition D_1D_2 ist daher im Allgemeinen *keine* Derivation, da im Vergleich zur Leibniz-Regel zwei zusätzliche Terme auftreten.

$$[D_1, D_2](a \cdot b) = D_1(D_2(a \cdot b)) - D_2(D_1(a \cdot b)) = D_1D_2a \cdot b - D_2D_1a \cdot b + a \cdot D_1D_2b - a \cdot D_2D_1b$$

- Betrachte die rechte Seite

$$\begin{aligned} [D_1, D_2]a \cdot b + a \cdot [D_1, D_2]b &= (D_1 D_2 - D_2 D_1)a \cdot b + a \cdot (D_1 D_2 - D_2 D_1)b \\ &= D_1 D_2 a \cdot b - D_2 D_1 a \cdot b + a \cdot D_1 D_2 b - a \cdot D_2 D_1 b \end{aligned}$$

Also ist die linke Seite gleich der rechten Seite und die Behauptung folgt. \square

Definition 6.22. Fixiere $D \in \text{Der}(A)$. Wir definieren eine neue Abbildung

$$\mathcal{D} : \text{Der}(A) \rightarrow \text{Der}(A), \quad \mathcal{D}(D_1) := [D, D_1]$$

Diese Abbildung ist selbst eine Derivation auf der Algebra $\text{Der}(A)$ (*Derivation auf Derivationen*).

Insbesondere gilt für den Kommutator die Leibniz-Regel

$$\mathcal{D}(D_1 * D_2) = \mathcal{D}(D_1) * D_2 + D_1 * \mathcal{D}(D_2)$$

beziehungsweise ausgeschrieben

$$[D, [D_1, D_2]] = [[D, D_1], D_2] + [D_1, [D, D_2]]$$

Aus dieser Leibniz-Regel folgt unmittelbar mit $[A, B] = -[B, A]$ die **Jacobi-Identität**:

$$[D_1, [D_2, D_3]] + [D_2, [D_3, D_1]] + [D_3, [D_1, D_2]] = 0$$

Also ist die Jacobi-Identität ein Spezialfall der Leibniz-Regel.

Bemerkung 6.23. Die geschlungene Derivation \mathcal{D} wird hier nur zur Motivation eingeführt. Sie erklärt, warum die Jacobi-Identität aus der Leibniz-Regel folgt, und dient als Vorbereitung für die Lie-Algebra der Derivationen, die in der anschließenden Vorlesung [Geometrie und Gravitation I](#) eine wichtige Rolle spielt.

6.3 Riemannntensor

Da die Allgemeine Relativitätstheorie gekrümmte Räume betrachtet, benötigen wir ein Maß für diese Krümmung. Dieses Maß ist der sogenannte **Riemannntensor** oder **Riemannsche Krümmungstensor**.

Definition 6.24. Der **Riemannntensor** ist definiert durch die Wirkung des Kommutators zweier kovarianter Ableitungen auf ein Vektorfeld:

$$R_{dab}^c v^d := [\nabla_a, \nabla_b] v^c$$

Dabei bezeichnen die eckigen Klammern den **Kommutator** der kovarianten Ableitungen

$$[\nabla_a, \nabla_b] := \nabla_a \nabla_b - \nabla_b \nabla_a$$

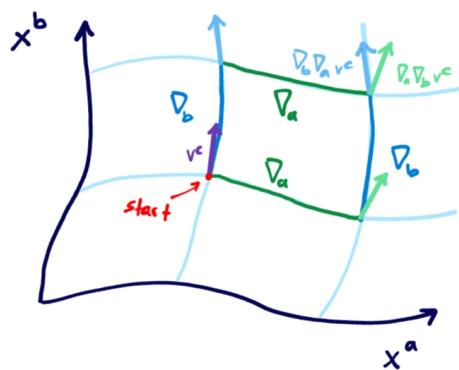


Abbildung 6.2: Der Riemannntensor beschreibt den Unterschied zwischen der Ableitung ∇_b von v^c in Richtung b und der anschließenden Ableitung ∇_a dieses Ergebnisses in Richtung a vs. der Ableitung ∇_a von v^c in Richtung a und der anschließenden Ableitung ∇_b dieses Ergebnisses in Richtung b .

Anders formuliert: er beschreibt den Unterschied von der infinitesimalen Bewegung von v^c zuerst in Richtung b und dann in Richtung a vs zuerst in Richtung a und dann in Richtung b . Wenn es keinen Unterschied gibt, ist der Raum flach. In diesem Sinn ist der Riemannntensor ein Maß für die Krümmung.

[Quelle](#)

Der Riemannntensor misst somit die Veränderung der kovarianten Ableitung in Bezug auf sich selbst und charakterisiert sie so.

Proposition 6.25. Kommutator der kovarianten Ableitungen ist $C^\infty(M)$ -linear, d.h. Skalarfelder können herausgezogen werden

$$[\nabla_a, \nabla_b](fv^c) = f \cdot [\nabla_a, \nabla_b]v^c \quad (6.14)$$

Beweis. Nach der vierten Eigenschaft der kovarianten Ableitung (Torsionsfreiheit / Satz von Schwarz) gilt für alle Skalarfelder $f \in C^\infty(M)$:

$$\begin{aligned} \nabla_a \nabla_b f &= \nabla_b \nabla_a f \\ \Leftrightarrow (\nabla_a \nabla_b - \nabla_b \nabla_a)f &= 0 \\ \Leftrightarrow [\nabla_a, \nabla_b]f &= 0 \end{aligned}$$

Damit erhalten wir für ein Vektorfeld v^c und ein Skalarfeld f :

$$[\nabla_a, \nabla_b](fv^c) = \underbrace{[\nabla_a, \nabla_b]f \cdot v^c + f \cdot [\nabla_a, \nabla_b]v^c}_{=0} = f \cdot [\nabla_a, \nabla_b]v^c$$

□

Wir leiten nun die **Koordinatendarstellung des Riemannntensors** her.

$$\begin{aligned} R_{dab}^c v^d &= [\nabla_a, \nabla_b]v^c = [\nabla_a, \nabla_b](v^i \partial_i^c) \\ &\stackrel{(6.14)}{=} \underbrace{v^i}_{=v^i \delta_j^i = v^i \partial_j^d dx_d^i} [\nabla_a, \nabla_b] \partial_i^c \\ &= v^i \delta_j^i = v^i \partial_j^d dx_d^i \\ &= dx_d^i [\nabla_a, \nabla_b] \partial_i^c v^d \end{aligned}$$

Damit folgt die koordinatenabhängige Definition des Riemannntensors in der Karte x^i

$$R_{dab}^c = \underbrace{dx_d^i}_{\text{lokal definiert}} [\nabla_a, \nabla_b] \underbrace{\partial_i^c}_{\text{lokal definiert}} \quad (6.15)$$

Proposition 6.26. R_{dab}^c ist unabhängig von der Wahl der Karte und somit global definiert. Insbesondere ist R_{dab}^c tatsächlich ein $(1,3)$ -Tensorfeld.

Beweis. Wähle eine Karte $\hat{x}^i = \hat{x}^i(x^k)$.

$$\text{zz: } \hat{R}_{dab}^c = R_{dab}^c$$

Es gilt

$$\begin{aligned} \hat{R}_{dab}^c &= d\hat{x}_d^i [\nabla_a, \nabla_b] \underbrace{\hat{\partial}_i^c}_{\substack{\text{Basis} \\ \text{trafo} \hat{\partial}_i x^k \partial_k^c}} \\ &= d\hat{x}_d^i [\nabla_a, \nabla_b] \left(\underbrace{\hat{\partial}_i x^k}_{\in C^\infty(M)} \partial_k^c \right) \\ &\stackrel{(6.14)}{=} \underbrace{\hat{\partial}_i x^k}_{\substack{\text{Basis} \\ \text{trafo} dx_d^k}} d\hat{x}_d^i [\nabla_a, \nabla_b] \partial_k^c \\ &= dx_d^k [\nabla_a, \nabla_b] \partial_k^c \\ &= R_{dab}^c \end{aligned}$$

□

Da die lokalen Darstellungen auf Überlappungsgebieten übereinstimmen, definiert R_{dab}^c ein globales Tensorfeld.

Definition 6.27. Der **Ricci-Tensor** ist die Spur des Riemannntensors

$$R_{ab} := R_{acb}^c$$

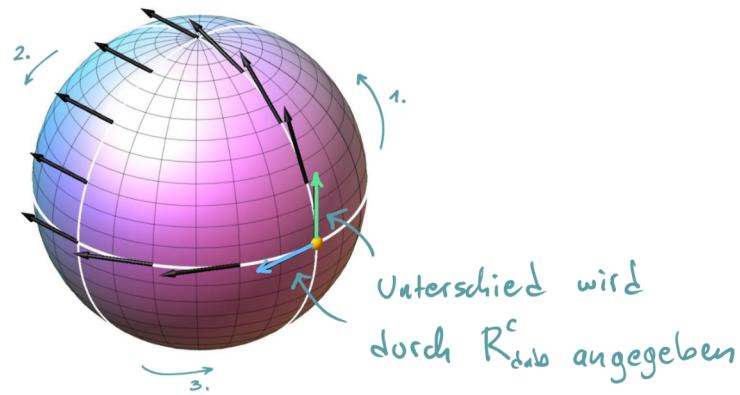


Abbildung 6.3: Der Riemann-tensor scheint also unser Dilemma aus Kapitel 1 mit der Parallelverschiebung im gekrümmten Raum zu lösen. Mehr dazu im nächsten Kapitel.

6.4 Parallelverschiebung

Die folgenden Kapitel wurden im WS25/26 in der letzten Vorlesung im Schnelldurchlauf durchgemacht und sind daher hier nicht detailliert ausgeführt. Genauer sind sie in den Mitschriften von Robert Berndl-Forstner (WS20/21) und Florian Lindenbauer (WS18/19), zu finden:

<https://forum.fstph.at/t/einf-i-d-allgemeine-relativitaetstheorie-mitschrift-ws2018>

<https://forum.fstph.at/t/latex-skriptum-allg-relativitaetstheorie>

Da wir nun die kovariante Ableitung eingeführt haben, können wir sie verwenden, um zu definieren, wie Parallelität im gekrümmten Raum zu verstehen ist.

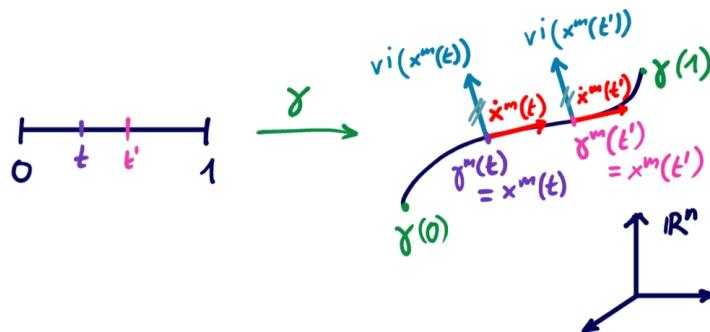
6.4.1 Situation im \mathbb{R}^n

Wir führen das zu Beginn des Kapitels motivierte Konzept der Parallelverschiebung nun genauer ein.

Sei eine parametrisierte Kurve

$$\gamma : [0, 1] \rightarrow \mathbb{R}^n$$

mit Koordinaten $\gamma(t) = x^m(t)$ und $v^b = v^i \partial_i^b$ ein Vektorfeld, welches wir mit $v^b(x^m(t)) = v^i(x^m(t)) \partial_i^b|_{x^m(t)}$ entlang der Kurve betrachten.



Das Vektorfeld heißt **parallel entlang der Kurve**, wenn es ausgewertet entlang der Kurve überall in dieselbe Richtung zeigt und denselben Betrag hat. Das heißt, wenn es sich entlang der Kurve nicht ändert. Also

$$\frac{d}{dt} (v^b(x^m(t))) = 0$$

Da die Basisvektoren ∂_i^b im \mathbb{R}^n ortsunabhängig sind, reduziert sich diese Bedingung auf die Konstanz der Komponenten:

$$\frac{d}{dt} v^i(x^m(t)) = 0$$

Die gewöhnliche Ableitung misst in diesem Fall genau die Abweichung von der Parallelität.

Mit der Kettenregel erhalten wir andererseits

$$0 = \frac{d}{dt} (v^i(x^m(t))) = \dot{x}^j(t) \partial_j v^i(x^m(t)) \quad (6.16)$$

wobei $\dot{x}^j(t)$ die Komponenten des Tangentialvektors $\dot{\gamma}(t)$ in der gewählten Karte sind.

Parallelverschiebung entlang einer Kurve im \mathbb{R}^n bedeutet also, dass ein Vektor, der verschoben wurde, weiterhin in dieselbe Richtung zeigt und denselben Betrag hat.

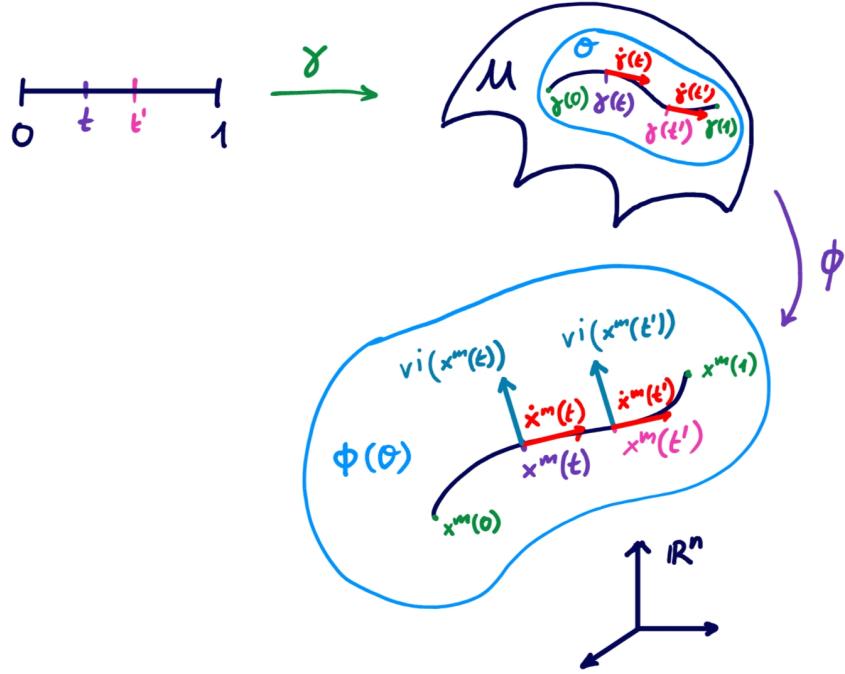
6.4.2 Verallgemeinerung auf Mannigfaltigkeiten

Sei $\gamma : [0, 1] \rightarrow M$ eine Kurve auf einer Mannigfaltigkeit.

Wir erinnern uns: ein Vektorfeld auf einer Mannigfaltigkeit ist definiert als eine Abbildung, die jedem Punkt $p \in M$ einen Tangentialvektor $v^a(p) \in T_p M$ zuordnet. Wenn wir dieses nun entlang einer Kurve γ und abgebildet mit einer Karte betrachten, hat es die folgende Darstellung

$$v^a(x^m(t)) = v^i(x^m(t)) \partial_i^a|_{x^m(t)}$$

Bei einer Ableitung nach dem Kurvenparameter t ändern sich nun nicht nur die Komponenten $v^i(x^m(t))$, sondern auch die Basisvektoren $\partial_i^a|_{x^m(t)}$ selbst, da diese vom Punkt $p = \gamma(t) \in M$ via $\phi^i(\gamma(t)) = x^i(t)$ abhängen.



Die gewöhnliche Ableitung misst daher nicht ausschließlich eine tatsächliche geometrische Änderung des Vektors, sondern enthält zusätzliche Beiträge, die durch den Wechsel der lokalen Basis entstehen. Eine verschwindende gewöhnliche Ableitung ist somit im Allgemeinen nicht gleichbedeutend mit Parallelität.

Plan: die partielle Ableitung durch die kovariante Ableitung ersetzen.

Im \mathbb{R}^n ist der Tangentialvektor der Kurve γ

$$u^a(t) := \dot{\gamma}^i(t) \partial_i^a = \dot{x}^i(t) \partial_i^a \quad (6.17)$$

Damit kann (6.16) angeschrieben werden

$$\begin{aligned} 0 &= \frac{d}{dt} v^i(x^m(t)) \partial_i^b \\ &= \dot{x}^j \partial_j v^i \partial_i^b \\ &= \dot{x}^i \underbrace{\delta_i^j}_{=\partial_i^a dx_a^j} \partial_j v^i \partial_i^b \\ &= \dot{x}^i \partial_i^a dx_a^j \partial_j v^i \partial_i^b \\ &= \underbrace{\dot{x}^i \partial_i^a}_{=u^a} \underbrace{\partial_j v^i dx_a^j \partial_i^b}_{=\partial_a v^b} \\ &= u^a \partial_a v^b \end{aligned} \quad (6.18)$$

Damit erhalten wir für Parallelität im flachen Raum eine Äquivalenz bzgl. der koordinatenkovarianten Ableitung des Vektorfelds

$$\frac{d}{dt} (v^i(x^m(t))) = 0 \Leftrightarrow u^a \partial_a v^b = 0 \quad (6.19)$$

Die koordinatenkovariante Ableitung ist lokal definiert. Nun globalisieren wir das ganze mit der kovarianten Ableitung.

Definition 6.28. Sei $\gamma : [0, 1] \rightarrow M$ eine Kurve auf einer Mannigfaltigkeit mit Tangentialvektor $u^a(t) := \dot{\gamma}^a(t) \in T_{\gamma(t)}M$ bzw. in Koordinaten $u^a(t) = \dot{x}^i(t) \partial_i^a$. Ein Vektorfeld v^b heißt **∇ -parallel** entlang γ , falls

$$(u^a \nabla_a) v^b = 0 \quad (6.20)$$

Zur Vereinfachung verwenden wir die Kurzschreibweise

$$(u \nabla) := (u^a \nabla_a)$$

Bemerkung 6.29 (Nicht Teil der VO).

Quelle: https://de.wikipedia.org/wiki/Substantielle_Ableitung#Herleitung

Sei $\vec{x} = (x, y, z)$ ein Ortsvektor, $\vec{u} = (u_x, u_y, u_z) = (\dot{x}, \dot{y}, \dot{z})$ ein Geschwindigkeitsvektor und $\vec{\Phi}(\vec{x}, t) = (\Phi_x, \Phi_y, \Phi_z)$ ein Vektorfeld. Die **konvektive Ableitung** ist der Operator

$$(\vec{u} \cdot \nabla) = u_x \frac{\partial}{\partial x} + u_y \frac{\partial}{\partial y} + u_z \frac{\partial}{\partial z} = \dot{x} \frac{\partial}{\partial x} + \dot{y} \frac{\partial}{\partial y} + \dot{z} \frac{\partial}{\partial z}$$

wirkt auf das Vektorfeld $\vec{\Phi}$ komponentenweise:

$$(\vec{u} \cdot \nabla) \vec{\Phi} = \begin{pmatrix} u_x \frac{\partial \Phi_x}{\partial x} + u_y \frac{\partial \Phi_x}{\partial y} + u_z \frac{\partial \Phi_x}{\partial z} \\ u_x \frac{\partial \Phi_y}{\partial x} + u_y \frac{\partial \Phi_y}{\partial y} + u_z \frac{\partial \Phi_y}{\partial z} \\ u_x \frac{\partial \Phi_z}{\partial x} + u_y \frac{\partial \Phi_z}{\partial y} + u_z \frac{\partial \Phi_z}{\partial z} \end{pmatrix}$$

Dies ist eine räumliche Ableitung ohne Zeitanteil. Sie beschreibt die Änderung des Vektorfeldes $\vec{\Phi}$ entlang der Richtung \vec{u} .

Existieren ∇ -parallele Vektorfelder überhaupt?

Ja! Wir konstruieren nun eines.

6.4.3 Konstruktion ∇ -paralleles Vektorfeld

Wir fixieren nun eine Karte (O, ϕ) , in der das betrachtete Kurvenstück γ liegt.

Definition 6.30. Sei (O, ϕ) eine Karte. Die Differenz der global definierten kovarianten Ableitung ∇_a , ausgewertet bzgl. dieser Karte und der lokal definierten koordinatenkovariante Ableitung ∂_a bzgl. dieser Karte ist der lokal definierte **Differenzentensor** Γ_{ca}^b

$$\Gamma_{ca}^b v^c := (\nabla_a - \partial_a) v^b \quad \forall v^b \in \mathcal{T}_0^1(O) \quad (6.21)$$

Ist ∇ die Levi-Civita-Ableitung, so ist Γ_{ca}^b das Christoffel-Symbol. In Koordinaten

$$\Gamma_{ca}^b = \Gamma_{jk}^i \partial_i^b dx_c^j dx_a^k$$

Damit kann die globale kovariante Ableitung ausgewertet bzgl. der fixierten Karte durch die zugehörige

lokale koordinatenkvariante Ableitung plus Zusatzterm geschrieben werden

$$\begin{aligned} 0 = (u\nabla)v^b &= (u^a\nabla_a)v^b = \underbrace{(u^a\partial_a)v^b}_{\stackrel{(6.18)}{=} \frac{d}{dt}v^k(t)\partial_k^b} + \underbrace{u^a\Gamma_{ca}^b v^c}_{=\dot{x}^l\Gamma_{jl}^k v^j \partial_k^b} \\ &= \left(\frac{d}{dt}v^k(t) + \Gamma_{jl}^k \dot{x}^l v^j \right) \partial_k^b \end{aligned}$$

Dieser Ausdruck ist aufgrund der linearen Unabhängigkeit der Basisvektoren nur dann null, wenn alle Komponenten gleich null sind. D.h.

$$\frac{d}{dt}v^k(t) + \underbrace{\Gamma_{jl}^k(x^m(t))\dot{x}^l(t)}_{=H_j^k(t)} v^j(t) = 0 \quad (6.22)$$

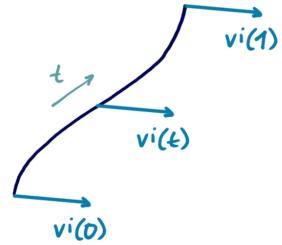
$$\dot{v}^k(t) + H_j^k(t) v^j(t) = 0 \quad (6.23)$$

Diese Gleichung erinnert an die zeitabhängige Schrödingergleichung mit dem zeitabhängigen Hamiltonoperator $H_j^k(t)$. Wenn wir diese lösen, erhalten wir ein Vektorfeld v^a , das entlang der Kurve ∇ -parallel ist (ein parallel verschobenes Vektorfeld). Man sagt auch, das Vektorfeld ist zu sich selbst parallel.

Die Gleichung kann mit einem **Zeitentwicklungsoperator/Parallelverschieber** U gelöst werden. Wir schreiben die Lösung von (6.23) in der Form

$$v^i(t) := U_j^i(t) v^j(0) \quad (6.24)$$

wobei $v^j(0)$ der Anfangsvektor am Startpunkt der Kurve ist und $U_j^i(t)$ der Parallelverschieber entlang γ .



Setzen wir (6.24) in (6.23) ein, so erhalten wir

$$\begin{aligned} 0 &= \frac{d}{dt}(U_j^i(t)v^j(0)) + H_j^k(t)U_i^j(t)v^i(0) \\ &= \dot{U}_j^i(t)v^j(0) + U_j^k(t)\underbrace{\frac{d}{dt}v^j(0)}_{=0} + H_j^k(t)U_i^j(t)v^i(0) \\ &= (\dot{U}_j^i(t) + H_j^k(t)U_i^j(t))v^i(0) \end{aligned}$$

Da der Anfangsvektor $v^i(0)$ beliebig ist, folgt die Differentialgleichung für den Parallelverschieber:

$$\dot{U}_j^i(t) + H_j^k(t)U_i^j(t) = 0 \quad (6.25)$$

Als Anfangsbedingung gilt

$$v^i(0) = U_j^i(0)v^j(0) \Rightarrow U_i^j(0) = \delta_i^j \quad (6.26)$$

da bei $t = 0$ keine Verschiebung stattfindet, muss der Vektor derselbe bleiben.

Wenn wir Gleichung (6.25) lösen, so erhalten wir den Parallelverschieber $U(t)$. Anschließend können wir mit (6.24) ein Vektorfeld konstruieren, das entlang γ ∇ -parallel ist (∇ -konstant ist).

Bemerkung 6.31. Der Parallelverschieber $U(t)$ definiert eine lineare Abbildung zwischen Tangentialräumen:

$$U(t) : T_{\gamma(0)}M \rightarrow T_{\gamma(t)}M \quad (6.27)$$

Er erlaubt es, Vektoren in verschiedenen Tangentialräumen entlang einer gegebenen Kurve miteinander zu identifizieren und als „parallel“ zu vergleichen.

6.4.4 Autoparallele Kurven und Geodätengleichung

Definition 6.32. Eine Kurve $\gamma : I \rightarrow M$ heißt **autoparallel** (bezüglich ∇), wenn das Vektorfeld, bestehend aus ihren Tangentenvektoren $u^a = \dot{\gamma}^a(t)$ (in Koordinaten $u^a = \dot{x}^i(t) \partial_i^a$), ∇ -parallel entlang der Kurve ist, d.h.

$$(u^a \nabla_a) u^b = 0 \quad (6.28)$$

gilt.

In Komponentenform mit $u^i = \dot{x}^i(t)$ ergibt sich aus (6.28) bzw. (6.22)

$$\boxed{\ddot{x}^i(t) + \Gamma_{jk}^i(x(t)) \dot{x}^j(t) \dot{x}^k(t) = 0} \quad (6.29)$$

Diese Gleichung heißt **Geodätengleichung**. Sie ist die Bewegungsgleichung eines Teilchens im Gravitationsfeld.

Im flachen Raum gilt $\Gamma_{jk}^i = 0$, sodass sich die Gleichung (6.29) vereinfacht zu

$$\frac{d}{dt}(u^i(t)) = \dot{x}^i(t) = 0$$

also zu einer Geraden $x(t) = x_0 + \dot{x}_0 t$.

Autoparallele Kurven sind daher die Verallgemeinerung von Geraden auf gekrümmte Räume.

6.5 Metrik und Levi-Civita-Ableitung

gutes Video: I finally understood Einstein's Metric Tensor! <https://www.youtube.com/watch?v=bV7KeKAPBQc>

Um Geodäten als kürzeste Verbindung zwischen zwei Punkten im gekrümmten Raum zu beschreiben, brauchen wir einen Abstandsbegriff.

Definition 6.33. Eine **Metrik** (genauer: **metrischer Tensor/pseudo-riemannsche Metrik**) auf einer Mannigfaltigkeit M ist ein $(0,2)$ -Tensorfeld

$$g_{ab} \in \mathcal{T}_2^0(M)$$

d.h.

$$\forall p \in M : \quad g_{ab} : \begin{cases} M \rightarrow \bigotimes_2^0 T_p M \cong \mathfrak{Bil}(T_p M, \mathbb{R}) \\ p \mapsto g_{ab}(p) \end{cases}$$

wobei $g_{ab}(p)$ eine nicht-entartete, symmetrische, indefinite Bilinearform ist, mit

$$g_{ab}(p) : T_p M \times T_p M \rightarrow \mathbb{R}$$

Mit symmetrisch meinen wir

$$g_{ab}(p) = g_{ba}(p)$$

Damit ist $g_{ab}(p)$ ein **Pseudo-Skalarprodukt** auf dem Tangentialraum $T_p M$.

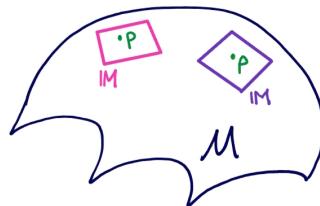
In jedem Punkt $p \in M$ gibt es eine geeignete Basis E_α^a von $T_p M$, sodass $g_{ab}(p)$ in dieser Basis Diagonalgestalt annimmt:

$$g_{ab}(p)(E_\alpha^a, E_\beta^b) = g_{ab} E_\alpha^a E_\beta^b = \begin{cases} \pm 1 & \alpha = \beta, \\ 0 & \alpha \neq \beta, \end{cases} = g_{\alpha\beta} = \text{diag}(\pm 1, \dots, \pm 1)$$

Mit einer geeigneten Basis kann also in jedem $T_p M$ der metrische Tensor g_{ab} zur Minkowski-Metrik transformiert werden

$$g_{ab} E_a^a E_b^b = \eta_{\alpha\beta} = \text{diag}(-1, 1, 1, 1)$$

Damit haben wir auf der Mannigfaltigkeit M ein Tensorfeld g_{ab} definiert, das in jedem Tangentialraum $T_p M$ eine Minkowski-Metrik induziert. Jeder Tangentialraum wird somit zu einem Minkowski-Raum \mathbb{M} . Die Raumzeit ist lokal flach (Minkowski-Raum), aber global gekrümmmt.



Die einzelnen Minkowski-Räume sind jedoch zunächst unabhängig voneinander. Wir möchten sie mit dem Paralleltransport in Verbindung bringen.

Wir betrachten eine Kurve γ mit Tangentenvektor in Koordinaten bzgl. einer Karte (O, ϕ)

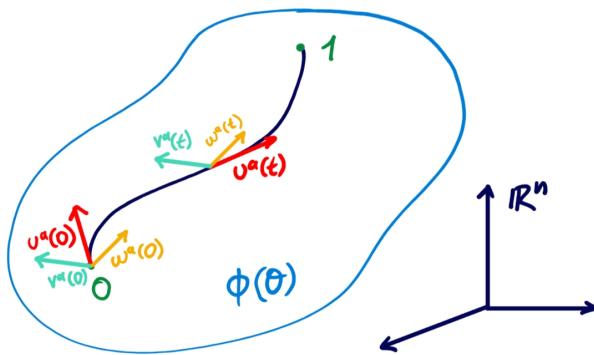
$$u^a = \dot{x}^i(t) \partial_i^a \quad (6.30)$$

sowie zwei Vektorfelder v^a und w^a , die entlang der Kurve ∇ -parallel sind:

$$(u\nabla)v^a = 0, \quad (u\nabla)w^a = 0$$

wobei

$$(u\nabla) := u^c \nabla_c$$



Wir fordern nun außerdem, dass sich auch das innere Produkt (genauer: dessen Pseudo-Skalarprodukt) $g_{ab}(p)(v^a, w^a) = g_{ab}v^a w^a \in \mathbb{R}$ entlang der Kurve nicht ändert. Dies ist skalar und repräsentiert den Winkel zwischen den beiden Vektoren. D.h. wir fordern, dass sich der Winkel zweier Vektoren bei Parallelverschiebung entlang einer Kurve nicht ändert.

Das heißt wir fordern

$$(u\nabla)(g_{ab}v^a w^b) = 0$$

Wendet man die Leibniz-Regel an und benutzt die Parallelität von v^a und w^a ,

$$\begin{aligned} 0 &= (u\nabla)(g_{ab}v^aw^b) = u^c\nabla_c(g_{ab}v^aw^b) \\ &= u^c\nabla_cg_{ab}v^aw^b + g_{ab}u^c\nabla_cv^aw^b + g_{ab}v^au^c\nabla_cw^b \end{aligned}$$

Die letzten beiden Terme verschwinden wegen (6.30). Übrig bleibt

$$u^c\nabla_cg_{ab}v^aw^b = 0 \quad (6.31)$$

Für $u^a(0)$, $v^a(0)$ und $w^a(0)$ folgt $\nabla_ag_{bc}(0) = 0$. Da Anfangswerte beliebig gewählt werden können und diese Beziehung für alle gelten soll, muss

$$\nabla_cg_{ab} = 0 \quad (6.32)$$

Das heißt, der metrische Tensor ist bezüglich dieser kovarianten Ableitung konstant. Eine kovariante Ableitung ∇_a die (6.32) erfüllt wird **metrisch kompatibel** genannt.

Definition 6.34. Eine kovariante Ableitung ∇_a heißt **Levi-Civita Ableitung** relativ zu $g_{ab} : \Leftrightarrow \nabla_cg_{ab} = 0$

Diese Bedingung wählt also eine ganz bestimmte kovariante Ableitung aus. Das ist die Ableitung, die in der Allgemeinen Relativitätstheorie verwendet wird.

6.5.1 Eindeutigkeit der Levi-Civita Ableitung

Seien ∇ und $\tilde{\nabla}$ zwei Levi-Civita Ableitungen. Ihre Differenz ist ein Tensor

$$C_{ab}^c := (\tilde{\nabla}_a - \nabla_a)$$

Für Kovektoren gilt dann

$$\nabla_a\omega_b = \tilde{\nabla}_a\omega_b - C_{ab}^c\omega_c \quad (6.33)$$

Dies kann auf beliebig viele kovariante Indizes fortgeführt werden. Somit gilt beim metrischen Tensor:

$$0 = \nabla_ag_{bc} = \tilde{\nabla}_ag_{bc} - C_{ba}^m g_{mc} - C_{ca}^m g_{bm}$$

Mit Index-Transport erhalten wir

$$\tilde{\nabla}_ag_{bc} = C_{cba} + C_{bca}$$

Durch zyklische Permutation der Indizes und Ausnutzung der Symmetrie (folgt aus Torsionsfreiheit der kovarianten Ableitung)

$$C_{ab}^c = C_{ba}^c$$

folgt schließlich (Rechenschritte hier aus Zeitgründen ausgelassen)

$$C_{ab}^c = \frac{1}{2}g^{cd}(\tilde{\nabla}_ag_{bd} + \tilde{\nabla}_bg_{ad} - \tilde{\nabla}_dg_{ab})$$

Definition 6.35. Sei (M, g_{ab}) eine pseudo-riemannsche Mannigfaltigkeit und ∇_a die zu g_{ab} gehörige Levi-Civita-Ableitung. Eine Kurve γ heißt **Geodäte**, wenn sie autoparallel bezüglich ∇_a ist, d.h.

$$(u^b\nabla_b)u^a = 0$$

Geodäten sind damit die Verallgemeinerung von Geraden auf gekrümmte Räume. Sie sind der kürzeste Weg zwischen zwei Punkten auf einer gekrümmten Mannigfaltigkeit.

Bemerkung 6.36. Mit der Levi-Civita-Ableitung haben wir nun ein geometrisches Werkzeug, das die zuvor mit $T_p M$ unabhängig voneinander definierten lokalen Minkowski-Räume \mathbb{M} miteinander in Beziehung setzt. Sie definiert einen Paralleltransport entlang von Kurven, der das Pseudo-Skalarprodukt und somit Winkel erhält. Auf diese Weise können Vektoren in verschiedenen Tangentialräumen sinnvoll miteinander verglichen werden. Die Levi-Civita-Ableitung wird daher auch als **Levi-Civita-Zusammenhang** bezeichnet.

6.6 Newtonsche Deviationsgleichung

Wir betrachten eine Familie von frei fallenden Objekten im Gravitationsfeld, die (nebeneinander) in Richtung Erdoberfläche fallen. Sie fallen alle in Richtung Erdmittelpunkt, daher verlaufen die Bahnlinien nicht nur „parallel nach unten“, sondern die Körper bewegen sich während des Falls auch aufeinander zu.

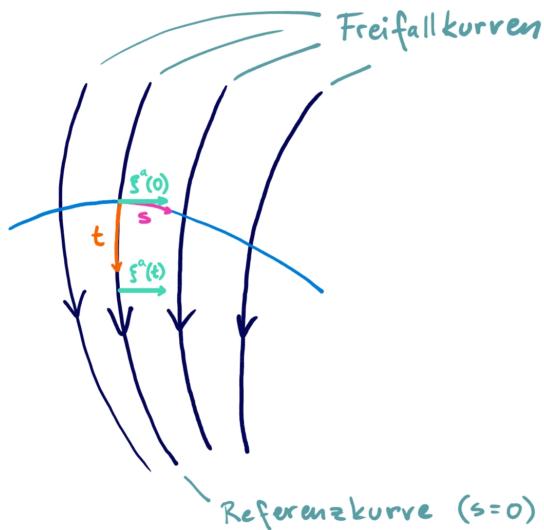
Die Teilchenbahnen hängen von zwei Parametern ab:

- t : Zeitparameter entlang einer Bahn
- s : Selektorparameter, der angibt, welche Freifallkurve betrachtet wird

Die Bahnen werden beschrieben durch

$$x^i = x^i(t, s)$$

Die Kurve $s = 0$ dient als Referenzkurve.



Nach dem zweiten Newtonschen Axiom gilt für ein Teilchen der Masse $m(s)$

$$m(s) \ddot{x}^i(t, s) = F^i(x(t, s))$$

Somit gilt für ein Teilchen newtonschen Gravitationsfeld

$$m(s) \ddot{x}^i(t, s) = -m(s) \partial_i \Phi(x^m(t, s))$$

Die Masse kürzt sich heraus, sodass die Bahnkurve unabhängig von der Masse des Teilchens ist:

$$\ddot{x}^i(t, s) = -\partial_i \Phi(x^m(t, s)) \quad (6.34)$$

Um zu beschreiben, wie sich benachbarte Freifallkurven relativ zueinander bewegen, definieren wir das **Zeigerfeld**

$$\xi^i(t) := \left. \frac{\partial x^i(t, s)}{\partial s} \right|_{s=0}$$

Dieses Vektorfeld ist bzgl. der Referenzkurve ($s = 0$) definiert und zeigt zur nächsten Freifallkurve. Es beschreibt, wie stark und in welche Richtung das Teilchen $x^i(t, 0)$ aufgrund der Gravitation abgelenkt wird.

Wir differenzieren die Bewegungsgleichung (6.34) nach s und werten sie bei $s = 0$ aus:

$$\begin{aligned} & \left. \frac{\partial}{\partial s} \right|_{s=0} (\ddot{x}^i(t, s) + \partial_i \Phi(x^m(t, s))) = 0 \\ \Leftrightarrow & \left. \frac{\partial^3 x^i(t, s)}{\partial s \partial t^2} \right|_{s=0} + \left. \frac{\partial}{\partial s} \right|_{s=0} (\partial_i \Phi(x^m(t, s))) = 0 \end{aligned}$$

Da die Ableitungen nach t und s vertauschen, erhalten wir

$$\left. \frac{\partial^3 x^i(t, s)}{\partial t^2 \partial s} \right|_{s=0} = \ddot{\xi}^i(t)$$

Mit der Kettenregel folgt

$$\begin{aligned} \left. \frac{\partial}{\partial s} \right|_{s=0} (\partial_i \Phi(x^m(t, s))) &= \left. \frac{\partial \partial_i \Phi(x^m(t))}{\partial x^j} \frac{\partial x^j(t, s)}{\partial s} \right|_{s=0} \\ &= \partial_j \partial_i \Phi(x^m(t)) \xi^j(t) \end{aligned}$$

Damit ergibt sich die **Newtonsche Deviationsgleichung**:

$$\boxed{\ddot{\xi}^i(t) + \partial_j \partial_i \Phi(x^m(t)) \xi^j(t) = 0} \quad (6.35)$$

Die Gleichung beschreibt die relative Beschleunigung aufgrund der Gravitation zweier benachbarter, frei fallender Teilchen. Also die Ablenkung, welche ein Teilchen unter Einwirkung des Gravitationsfeldes erfährt.

- Φ : Gravitationspotential (Skalarfeld)
- $\partial_i \Phi$: Gradient des Gravitationspotentials: Gravitationsfeld/newtonsche Gravitationsfeldstärke (Vektorfeld)
- $\partial_i \partial_j \Phi$: räumliche Änderung des Gravitationsfelds

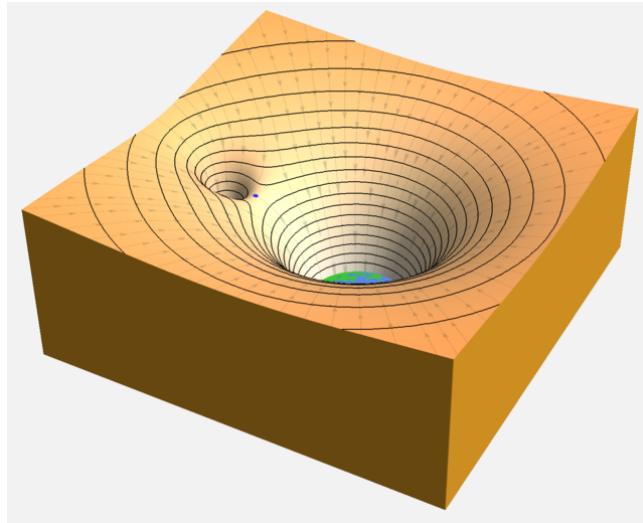


Abbildung 6.4: Gravitationspotential und Gravitationsfeld von Erde und Mond dargestellt. Der skalare Wert gibt die Höhe an, die Pfeile geben die Richtung an.

Quelle: <https://physikbuch.schule/work-in-the-gravitational-field.html>

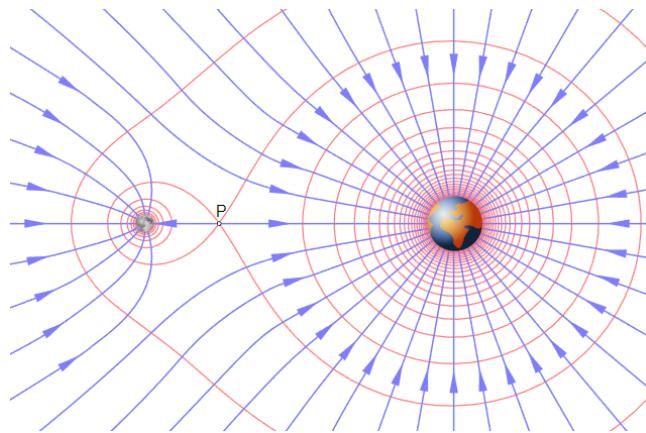


Abbildung 6.5: Gravitationsfeld von Erde und Mond

Quelle: <https://physikbuch.schule/work-in-the-gravitational-field.html>

Ist das Gravitationsfeld $\partial_i \Phi$ lokal konstant/homogen, so verschwindet dort $\partial_i \partial_j \Phi$ und benachbarte Freifallkurven bleiben parallel (wenn wir beispielsweise zwei Gegenstände im Vakuum dicht nebeneinander fallen lassen). In inhomogenen Gravitationsfeldern hingegen beschreibt die Deviationsgleichung die **Gezeitenkräfte**.

Diese Gleichung stellt das Newtonsche Analogon zur *Geodäten-Deviationsgleichung* der Allgemeinen Relativitätstheorie dar.

6.7 Geodätische Deviationsgleichung

Wir betrachten nun die Freifallkurven im Kontext der allgemeinen Relativitätstheorie. Hier bewegen sich frei fallende Körper auf kürzesten Wegen durch die gekrümmte Raumzeit, also auf *Geodäten* bzw. *autoparallelen Kurven*.

Seien die autoparallelen Kurven (Freifallkurven) wieder parametrisiert mit

- t : Zeitparameter entlang einer Bahn (Eigenzeit der Teilchen)
- s : Selektorparameter, der angibt, welche Freifallkurve betrachtet wird

Sei wieder das Zeigervektorfeld zwischen benachbarten Geodäten definiert als

$$\xi^i(t) := \left. \frac{\partial x^i(t, s)}{\partial s} \right|_{s=0}$$

und das Tangentialvektorfeld der jeweiligen Freifallkurve

$$u^i(t, s) = \frac{dx^i(t, s)}{dt}$$

wobei u^a zeitartig ist mit Längenquadrat

$$u^2 := g_{ab} u^a u^b < 0$$

Wir normieren das Tangentialvektorfeld sodass

$$u^2 = -1$$

Da die Freifallkurven autoparallel sind, gilt (Tangentialvektorfeld ist ∇ -konstant)

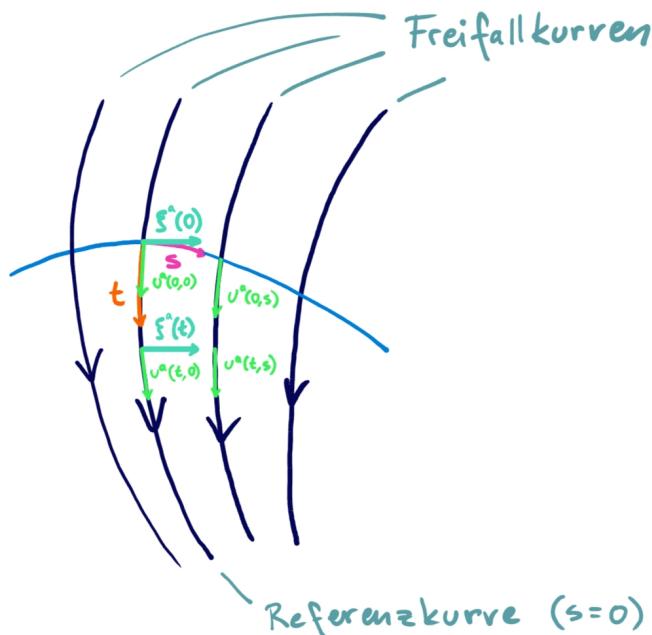
$$(u\nabla)u^a = 0 \quad (6.36)$$

wobei ∇ die Levi-Civita Ableitung ist.

Wir erinnern uns, dass der metrische Tensor g_{ab} invariant bzgl. der Levi-Civita Ableitung ist ($\nabla_c g_{ab} = 0$). Daher und wegen der Autoparallelität der Kurve verändert sich das Längenquadrat von u^a entlang der Kurve nicht:

$$(u\nabla)(u^2) = (u\nabla)(g_{ab}u^a u^b) = \underbrace{(u\nabla)g_{ab}}_{=0} u^a u^b + g_{ab} \underbrace{(u\nabla)u^a u^b}_{=0} + g_{ab} u^a \underbrace{(u\nabla)u^b}_{=0} = 0$$

Damit bleibt ein zu Beginn zeitartiger Tangentenvektor entlang der gesamten Freifallkurve zeitartig. Paralleltransport ändert also keine Längen- oder Kausalstruktur.



u^a ist ein Strömungsfeld, entlang dessen die Teilchen auf den Freifallkurven transportiert werden. Die Änderung des Strömungsfeldes durch das Zeigervektorfeld ξ^a wird durch die sogenannte **konvektive Ableitung** bzw. **Lie-Ableitung** beschrieben

$$L_u \xi^a = \underbrace{(u \nabla) \xi^a}_{\substack{\text{Änderung des Vektorfelds } \xi^a \\ \text{entlang von } u^a}} - \underbrace{(\xi \nabla) u^a}_{\substack{\text{Änderung des Vektorfelds } u^a \\ \text{entlang von } \xi^a}}$$

In WS25/26 hatten wir keine Zeit die Lie-Ableitung genauer einzuführen.

Wir fordern, dass das Zeigervektorfeld ξ^a mit der Strömung u^a mitgeführt wird, d.h. sich relativ zur Strömung nicht ändert:

$$L_u \xi^a = 0$$

beziehungsweise

$$(u \nabla) \xi^a = (\xi \nabla) u^a \quad (6.37)$$

Da wir das relativistische Analogon zur Newton-Deviationsgleichung (6.35) herleiten wollen und in dieser die zweite Ableitung von ξ auftritt, bestimmen wir nun die zweite Ableitung des Zeigervektorfelds entlang der Weltlinie, d.h. $(u \nabla)^2 \xi^a$.

Mit (6.37) schreiben wir

$$\begin{aligned} (u \nabla)^2 \xi^a &= (u \nabla)((\xi \nabla) u^a) \\ &= (u \nabla) \xi^c \nabla_c u^a + \xi^c \underbrace{(u \nabla)}_{=u^d \nabla_d} \nabla_c u^a \end{aligned} \quad (6.38)$$

Nach der Definition des Riemann-Tensors gilt

$$[\nabla_a, \nabla_b] v^c = (\nabla_a \nabla_b - \nabla_b \nabla_a) v^c = R_{dab}^c v^d$$

Dies können wir auf den zweiten Term von (6.38) anwenden und die Ableitungen $\nabla_d \nabla_c$ vertauschen

$$\begin{aligned} u^d \nabla_d \nabla_c u^a - u^d \nabla_c \nabla_d u^a &= u^d R_{bdc}^a u^b \\ \Leftrightarrow u^d \nabla_d \nabla_c u^a &= u^d \nabla_c \nabla_d u^a + u^d R_{bdc}^a u^b \end{aligned}$$

Somit ist

$$\begin{aligned} (6.38) &= (u \nabla) \xi^c \nabla_c u^a + \xi^c \underbrace{u^d \nabla_d \nabla_c u^a}_{\substack{\text{Leibniz} \\ \nabla_c (\underbrace{u^d \nabla_d u^a}_{(6.36)} - \nabla_c u^d \nabla_d u^a)}} + \xi^c u^d R_{bdc}^a u^b \\ &\Leftrightarrow (u \nabla) \xi^c \nabla_c u^a - \underbrace{\xi^c \nabla_c u^d \nabla_d u^a}_{=(\xi \nabla)} + \xi^c u^d R_{bdc}^a u^b \\ &\Leftrightarrow \underbrace{\left((u \nabla) \xi^d - (\xi \nabla) u^d \right)}_{(6.37) 0} \nabla_d u^a + R_{bdc}^a u^b u^d \xi^c \end{aligned}$$

Es folgt somit

$$(u \nabla)^2 \xi^a = \underbrace{R_{bdc}^a u^b u^d \xi^c}_{=-R_{bcd}^a}$$

So erhalten wir die **geodätische Deviationsgleichung**

$$(u\nabla)^2\xi^a + R_{bcd}^a u^b u^d \xi^c = 0$$

Sie beschreibt, wie sich benachbarte frei fallende Teilchen in einer gekrümmten Raumzeit relativ zueinander bewegen. Die Rolle der zweiten Ableitung des Gravitationspotentials in der Newtonschen Deviationsgleichung wird nun vom Riemann-Krümmungstensor übernommen. Sie macht sichtbar, dass Gravitation in der Allgemeinen Relativitätstheorie durch Raumzeitkrümmung beschrieben wird.

6.8 Einsteinsche Feldgleichungen

Wir vergleichen die Newtonsche und die geodätische Deviationsgleichung.

Einstens Idee: in schwachen Gravitationsfeldern, also bei wenig Raumzeitkrümmung, gilt (Newton-Limes)

$$R_{bcd}^a u^b u^d = \partial^a \partial_c \Phi \quad (6.39)$$

wobei wir die Indizes a, c anstelle von i, j verwenden, da wir die Newtonsche Deviationsgleichung in eine raumzeitliche Form überführen.

Aus der Newton-Mechanik folgt im Gravitationspotential Φ (für Einheitsmasse) die Poisson-Gleichung

$$\partial^2 \Phi = \Delta \Phi = 4\pi \rho_m \quad (6.40)$$

wobei ρ_m die **Materie-/Energiedichte** bezeichnet und Δ der Laplace Operator.

Um in (6.39) auf die zweite Ableitung zu kommen, müssen wir aus dem Krümmungstensor ein Objekt zweiter Stufe bauen. Dazu kontrahieren wir die Indizes a und c (Spurbildung):

$$R_{bd} := R_{bad}^a \quad (6.41)$$

R_{bd} wird **Ricci-Tensor** genannt.

Damit gilt

$$R_{bab}^a u^b u^d = R_{bd} u^b u^d = \partial^2 \Phi = \Delta \Phi = 4\pi \rho_m$$

Die Materie-/Energiedichte ρ_m kann mit dem **Energie-Impuls-Tensors** T_{ab} geschrieben werden als

$$\rho_m = T_{ab} u^a u^b$$

Der **Energie-Impuls-Tensor** beschreibt die Dichte und den Fluss von Energie und Masse in der Raumzeit. Salopp gesagt die Materie- und Energieverteilung in der Raumzeit. Er ist eine Verallgemeinerung des Stress-Tensors.

Damit ergibt sich als erster Ansatz (Vorstufe Einsteinsche Feldgleichung 1914)

$$R_{ab} = 4\pi T_{ab}$$

Warum stimmt diese Gleichung nicht?

Physikalisch gilt lokale Energie-Impuls-Erhaltung (Energie-Impuls-Tensor ist divergenzfrei):

$$\nabla^a T_{ab} = 0$$

Nähmen wir $R_{ab} = 4\pi T_{ab}$, müsste also auch $\nabla^a R_{ab} = 0$ gelten. Aus der (kontrahierten) Bianchi-Identität folgt jedoch

$$\nabla^a R_{ab} = \frac{1}{2} \nabla_b R \quad (6.42)$$

wobei $R := g^{ab} R_{ab}$ der **Krümmungsskalar**, der im Allgemeinen nicht konstant ist, und daher $\nabla_b R \neq 0$, also $\nabla^a R_{ab} \neq 0$.

Einstein betrachtete daher den Ausdruck

$$\begin{aligned} \nabla^a \left(R_{ab} - \frac{1}{2} g_{ab} R \right) &= \nabla^a R_{ab} - \frac{1}{2} \underbrace{\nabla^a (g_{ab} R)}_{R+g_{ab}\nabla^a R} = \nabla^a R_{ab} - \frac{1}{2} g_{ab} \underbrace{\nabla^a R}_{=g^{ab}\nabla_c} \\ &= \underbrace{\nabla^a g_{ab}}_{\substack{\text{Levi-Civita} \\ \text{Ableitung}}} R + g_{ab} \nabla^a R \\ &\stackrel{0}{=} \nabla^a R_{ab} - \frac{1}{2} \underbrace{g_{ab} g^{ac}}_{=\delta_b^c} \nabla_c R = \nabla^a R_{ab} - \frac{1}{2} \nabla_b R \stackrel{(6.42)}{=} 0 \end{aligned}$$

Somit ist der sogenannte **Einstein-Tensor**

$$G_{ab} := R_{ab} - \frac{1}{2} g_{ab} R$$

divergenzfrei.

Einstiens finaler Ansatz lautet daher

$$G_{ab} = 4\pi\alpha T_{ab}$$

Die Konstante α wird durch den Newton-Limes festgelegt; man erhält $\alpha = 2$ und somit die endgültige Form der **Einsteinschen Feldgleichungen** (1915):

$$G_{ab} = 8\pi T_{ab}$$

Die Einsteinsche Feldgleichung beschreibt die Wechselwirkung zwischen Materie (Masse) bzw. Energie (mit Energie-Impuls-Tensor T_{ab}) und der Geometrie der Raumzeit (mit Einstein-Tensor G_{ab}), wobei Materie die Raumzeit krümmt und diese Krümmung die Bewegung von Objekten (wie Planeten oder Licht) bestimmt. Es ist ein dynamisches Rückkopplungssystem.

Die Lösung dieser Gleichung liefert die Raumzeitmetrik g_{ab} . Erst durch g_{ab} erhalten Begriffe wie zeitartig, raumartig oder lichtartig überhaupt eine Bedeutung, da sie von der Metrik abhängen.

Der Physiker John Archibald Wheeler brachte es knapp auf den Punkt:

“Materie sagt dem Raum, wie er sich krümmen soll. Der Raum sagt der Materie, wie sie sich bewegen soll.“

Lösen der Gleichung

gute Links: https://en.wikipedia.org/wiki/Einstein_field_equations#Mathematical_form

<https://profoundphysics.com/einstein-field-equations-fully-written-out-what-do-they-look-like-expanded/>

G_{ab} und T_{ab} sind Tensoren zweiter Stufe mit Indizes $a, b = 0, 1, 2, 3$. Das heißt, es sind 4×4 Matrizen. Für jedes Indexpaar (a, b) gibt es eine Gleichung, also insgesamt 16 Gleichungen. Aufgrund der Symmetrie $G_{ab} = G_{ba}$ bzw. $T_{ab} = T_{ba}$ bleiben nur 10 Gleichungen übrig. Mit der (kontrahierten) Bianchi-Identität kann die Zahl der unabhängigen Gleichungen weiter auf 6 reduziert werden.

Die Einsteinschen Feldgleichungen sind gekoppelte nichtlineare partielle Differentialgleichungen zweiter Ordnung. Die partiellen Ableitungen stecken in G_{ab} .

Idee: Masse/Materieverteilung vorgeben

Um die Gleichung zu lösen, könnte man meinen, zunächst eine Materieverteilung T_{ab} vorzugeben und anschließend die Feldgleichungen nach g_{ab} lösen. Allerdings muss für die gewählte Materieverteilung $\nabla^a T_{ab} = 0$ gelten. Die Levi-Civita-Ableitung ∇^a hängt von der gesuchten Metrik g_{ab} ab, daher kann erst nachdem die Metrik bestimmt wurde, geprüft werden, ob die zuvor gewählte Materieverteilung eine physikalisch zulässige war. Diese Methode ist sehr ineffizient.

Idee: Metrik vorgeben

Eine andere Idee wäre es, eine Metrik g_{ab} vorgeben, daraus G_{ab} zu berechnen und anschließend den Energie-Impuls-Tensor zu bestimmen mit

$$T_{ab} = \frac{1}{8\pi} G_{ab}$$

Allerdings sind auch hier wieder nicht alle Ergebnisse für T_{ab} physikalisch zulässig, da die Materie-/Energiedichte für alle Beobachter positiv sein muss ([schwache Energiebedingung](#)) d.h.

$$\rho_m = T_{ab} u^a u^b \geq 0 \quad \forall \text{ zeitartigen } u^a$$

Diese Methode ist ebenso sehr ineffizient.

Vakuumlösung

Ein Spezialfall ist das **Vakuum**, in dem es keine Materie gibt, also ist der Energie-Impuls Tensor (Materieverteilung)

$$T_{ab} = 0$$

Dann folgt $G_{ab} = 0$ bzw. $R_{ab} = 0$. Der Riemann-Krümmungstensor kann weiterhin von Null verschieden sein (Raumzeit nicht flach). Karl Schwarzschild fand 1916 eine **Vakuumlösung** der Feldgleichungen. Ihre spätere Interpretation zeigte, dass sie ein schwarzes Loch beschreibt, mit Ereignishorizont und einer Region unbeschränkt wachsender Raumzeitkrümmung.

Kosmologische Lösungen und kosmologische Konstante

guter Link: https://en.wikipedia.org/wiki/Einstein_field_equations#Cosmological_constant

Mit der Annahme des **kopernikanischen Prinzips** existieren Lösungen der Einsteinschen Feldgleichungen, die das Universum auf großen Skalen (Galaxiehaufen, Supercluster) beschreiben. Es lautet:

- **Homogenität:** Auf hinreichend großen Skalen sind physikalische Größen (z.B. Ladungsdichte, Materiedichte) auf Gleichzeitflächen ortsunabhängig.
- **Isotropie:** Auf hinreichend großen Skalen gibt es keine ausgezeichneten Raumrichtungen. Das Universum schaut auf hinreichend großen Skalen in alle Richtungen gleich aus.

Unter diesen Annahmen vereinfachen sich die Einsteinschen Feldgleichungen stark und bilden ein dynamisches System von einem Freiheitsgrad. Dieser ist der **Skalenfaktor** $a(t)$. Das Universum kann sich daher insgesamt nur ausdehnen oder zusammenziehen.

Einstein stellte fest, dass die Einsteinschen Feldgleichungen unter diesen Annahmen *kein statisches Universum* zulassen, sondern notwendigerweise ein dynamisches. Um dennoch ein statisches Universum zu ermöglichen, fügte er einen zusätzlichen Term hinzu:

$$G_{ab} + \Lambda g_{ab} = 8\pi T_{ab}$$

wobei Λ die **kosmologische Konstante** ist.

Dieser Zusatzterm ist mathematisch konsistent, da sowohl der Einstein-Tensor G_{ab} als auch der metrische Tensor g_{ab} kovariant divergenzfrei sind:

$$\nabla^a G_{ab} = 0, \quad \nabla^a g_{ab} = 0.$$

Die Forderung der lokalen Energie-Impuls-Erhaltung $\nabla^a T_{ab} = 0$ bleibt somit erhalten.

Spätere Beobachtungen (insbesondere die [Rotverschiebung](#) ferner Galaxien durch Edwin Hubble, 1929) zeigten jedoch, dass sich das Universum tatsächlich ausdehnt. Einstein verwarf daraufhin die kosmologische Konstante und bezeichnete ihre Einführung später als seine „größte Eselei“.

Heute spielt die kosmologische Konstante erneut eine Rolle: kosmologische Beobachtungen Mitte der 1990ern deuten auf eine beschleunigte Expansion des Universums hin, die durch eine positive kosmologische Konstante erklärt werden kann. In diesem Kontext kann Λ als Energie des Vakuums interpretiert werden (vgl. dunkle Energie).

Auf der Skala von Galaxien oder kleineren Systemen ist der Einfluss der kosmologischen Konstante vernachlässigbar.

Bemerkung 6.37 (nicht Teil der VO).

Quelle: https://en.wikipedia.org/wiki/Einstein_field_equations#Mathematical_form

How Mass WARPS SpaceTime: Einstein's Field Equations in Gen. Relativity | Physics for Beginners: <https://youtu.be/FJnTItLVIqQ>

Die Einsteinschen Feldgleichungen lauten in SI-Einheiten

$$G_{ab} + \Lambda g_{ab} = \kappa T_{ab}$$

wobei $\kappa = \frac{8\pi G}{c^4}$ die **einsteinsche Gravitationskonstante** oder **Einsteinkonstante** mit der Newtonschen Gravitationskonstante G und der Lichtgeschwindigkeit c ist. In der Vorlesung wird $c = G = 1$ verwendet.

Fazit

Die **zentrale Aussage der Allgemeinen Relativitätstheorie** lautet: Gravitation ist keine Kraft, sondern eine Konsequenz der Raumzeitkrümmung. Frei fallende Objekte bewegen sich unter Einfluss der Gravitation auf geradest möglichen Bahnen (Geodäten). Die Raumzeitmetrik g_{ab} entsteht aus dem Wechselspiel zwischen Geometrie und Materie.

Weiterführende Literatur

- R. M. Wald, *General Relativity*.
- B. Schutz, *A First Course in General Relativity*.
- S. Carroll, *Spacetime and Geometry*.